Reliability test information for production assurance and reliability management based on ISO 20815:2018 - Uncertainty linked to test requirements

Jon T. Selvik
University of Stavanger and NORCE – Norwegian Research Centre, Stavanger, Norway;
Corresponding author: E-mail: jon.t.selvik@uis.no

Endre Willmann
Willmann Engineering, Oslo, Norway; E-mail: endre@willmann.no

Alex Green
Chevron Energy Technology Company, Aberdeen, UK; E-mail: Alex.green@chevron.com

As part of production assurance and reliability management, particularly in the engineering phase, reliability testing can be carried out to obtain relevant information related to the system performance. The testing represents a type of physical examination of the item, where the required level of reliability is to be demonstrated for the relevant environment and lifetime. For some items, an accelerated testing is required to obtain this information, as the failure characteristics might not be revealed from regular reliability life testing. However, as noted in the international standard ISO 20815:2018 on production assurance and reliability management, the obtained information could sometimes be misleading. One should be careful when using information from the reliability tests, as the test requirements are often misunderstood, and it could be difficult to apply these consistently in planning and execution; for example, from qualification testing. Hence, to achieve useful information, it is important to handle uncertainties in an appropriate way.

In this paper, we discuss the claim above, particularly why it is important to capture the uncertainties for the different types of reliability tests, within the context of ISO 20815:2018. Focus is on the reporting and usefulness of the test results. A common finding to several types of reliability tests is that there are different ways of interpreting and using the requirements, which strongly influence both assumptions made and the conclusions, and are thus influencing the usefulness of the information. In the paper, we give some advice on how to handle different uncertainty aspects in the test reporting and in use in a production assurance and reliability management context.

Keywords: Reliability testing, standardization, reliability, uncertainty, production assurance.

1. Introduction

1.1 Reliability testing context

For production facilities in the petroleum, petrochemical and natural gas industries, being the systems and industries focused upon in this paper, reliability testing of equipment could be required in the development and design phases, as input to the prediction of future reliability performance of systems. This is especially relevant in situations where equipment is of new design, when it is coming from new vendors or when it cannot show successful demonstration for the specific environment or for the relevant lifetime.

The information obtained from the reliability testing is part of overall reliability management. It captures a specter of coordinated activities undertaken to direct and control the reliability performance. These activities cover identification, analysis, treatment, and monitoring of system or component reliability throughout the relevant life cycles considered. When limited to development and design phases, the concept 'reliability engineering' is sometimes used instead, representing a subset of the full reliability management specter; including different types of

Proceedings of the 30th European Safety and Reliability Conference and the 15th Probabilistic Safety Assessment and Management Conference
Edited by Piero Baraldi, Francesco Di Maio and Enrico Zio
Published by Research Publishing, Singapore.
ISBN: 978-981-14-8593-0; doi:10.3850/978-981-14-8593-0

reliability testing, such as development-reliability testing, qualification testing and various accelerated tests (see Section 3 for a description of reliability testing).

The reliability performance relates to the capacity of the system or components to meet production performance objectives, i.e. the demands for delivery or performance for a specified lifetime. Having equipment with high availability, functioning on demand is obviously important in order to meet these objectives; giving high attention to information about reliability performance for design and execution of production assurance.

For the industries mentioned above, ISO 20815:2018 gives normative guidance on how to implement a 'production assurance programme' (PAP) to achieve the objectives. The standard defines 'production assurance' as: "*activities implemented to achieve and maintain a performance that is at its optimum in terms of the overall economy and at the same time consistent with applicable framework conditions*".

It is noted that this concept stretches beyond oil and gas production, and cover a broad set of activities influencing costs and requiring high reliability performance, such as for example drilling, installation and maintenance operations. Meaning that also associated equipment, such as e.g. drilling equipment and ROVs, are part of the production assurance and reliability management scope. In addition, software and human aspects are also included.

1.2 Reliability test requirement claim in ISO 20815:2018

Reliability testing of such equipment produce information regarding the performance and capacity of the production system, and clearly is relevant beyond the development and design phases. According to Kleyner (2008), reliability testing is the cornerstone in any reliability-engineering program. However, usefulness of the information obtained is about how well it reflects the actual lifetime degradation and failure characteristics, which in practice, might be difficult to reproduce. The usefulness is challenged by the claim in the ISO 20815:2018, that: "*careful consideration is needed when developing reliability testing programs as the requirements are often misunderstood and can be difficult to plan and implement*".

In this paper, we discuss this claim, where our objective is to present some insights from industrial experience, substantiating how interpretation of testing requirements can be a source of uncertainty and in practice lead to inappropriate input to reliability management and production assurance. For the PAP, it is important to take into account the uncertainty especially for important input parameters (Barabady et al. 2010). Focus is thus on the usefulness of the test results, and this relates to the reporting and the way to express uncertainties for the results. ISO 20815:2018, as a main reference document, gives some guidance on how to handle uncertainties in a production assurance and reliability management context, and is a natural starting point for the discussion.

The paper is structured as follows. Section 2 gives a brief overview of the guidance provided by ISO 20815:2018 on uncertainty handling. Section 3 gives an overview of different types of tests and some relevant requirements and standards applied for reliability testing. Then, Section 4 gives a relevant subsea example with discussions, followed by concluding remarks in Section 5.

2. Handling of uncertainty based on ISO 20815:2018

While the establishment and execution of the PAP is a mandatory activity outlined in ISO 20815:2018, the guidance on uncertainty handling is informative. Yet, the handling is implicitly required through the link to quality management, as uncertainty is an integral part of risk according to ISO 9000:2015 and ISO 9001:2015. This indicating the handling is important for achievement of the production performance objectives.

In a way, reliability testing is in itself a measure to reduce uncertainty by ensuring demonstration of the required functionality and robustness of critical equipment. However, this requires that the testing is performed and reported in an appropriate way, making the information useful and applicable. Following that, as pointed out in ISO 20815:2018, assumptions and reliability information obtained should be traceable and relevant for the environment and time period considered.

In Item I.9 of ISO 20815:2018, it is recommended that, in addition to variability (i.e. aleatory or stochastic uncertainty) aspects, where confidence or prediction intervals may be produced, "*factors contributing to uncertainty as a result of the way the system performance is modelled should be covered*". Confidence or prediction intervals fail to capture the model uncertainty. Especially when dealing with accelerated testing, when the output is amplified by extrapolation, it is important to consider the model uncertainty. For the modelling, the "assumptions made" are essential. Brissaud et al. (2012), in the context of production performance of oil and gas facilities, describes model development performed using the information, data and assumptions given in the study basis, including framing of the scope, system description and collection of reliability information. This framing of the system and the assumptions made can build on weak foundation and be sensitive for the reliability predictions. The effects of a weak model are not revealed from standard sensitivity testing or importance testing, and should be reflected upon for the PAP. Despite this, no explicit recommendation is given by ISO 20815: 2018 on how to express this uncertainty.

Regarding assumptions, the standard refers specifically to Selvik and Aven (2011); suggesting a way to address the uncertainties related to assumptions made in the assessment, capturing also model uncertainties. The suggested way is a qualitative procedure, which ranks the assumptions (referred to as 'uncertainty factors' in the reference) based on assessment of the level of uncertainty and sensitivity for each of them. And where high uncertainty characterises situations where one or more of the following conditions are met:

- The assumptions made represent strong simplifications.
- Data are not available, or are unreliable.
- There is lack of agreement/consensus among experts.
- The phenomena involved are not well understood; models are non-existent or known/believed to give poor predictions.

We refer to Selvik and Aven (2011) for further details on this. See also discussion on how to assess assumptions in

3. Reliability testing and requirements

3.1 Reliability testing

ISO 20815:2018 establish a link between reliability testing and requirements by describing reliability testing as the activities covering the physical demonstration of reliability requirements in the intended environment over a simulated design life. In practice, this can include a range of potential testing programmes from (a) early tests to establish functional capabilities supporting basic design performance requirements, (b) qualification tests that are more designed to demonstrate technical specifications and performance requirements at specific operating and environmental conditions can be met, and (c) tests more specifically designed to demonstrate that the aforementioned performance requirements can meet reliability and life expectations (Kececioglu 2002). In reality, these distinctions are often not explicitly stated.

In ISO 20815:2018, reliability testing is described distinctly from safety testing (e.g. periodic testing and testing of equipment for safety purposes), quality testing (e.g. FAT – factory acceptance test, SAT – Site acceptance test and SIT – system integration test), various small-scale tests, laboratory tests and software tests. Different types of reliability tests may be needed to support the planning and execution of the PAP.

Issues related to problems such as long operating life and adequately simulating reliability life can be addressed through use of accelerated life testing techniques. Accelerated tests may shorten the time needed to complete the testing, by overstressing (accelerating) one or several of the test parameters, typically temperature and pressure. For overview of accelerated testing, see e.g. Vassiliou et al. (2008) and Hobbs (2008). See also relevant standards; IEC 61123:2019, IEC 61163-1:1998 and IEC 61163-2:1998; and to technical report ISO/TR 16194:2017. The latter defines accelerated life testing as a: "*process in which a component is forced to fail more quickly than it would have under normal use conditions and which provides information about component's life characteristics*".

3.2 Test requirements

Two key references for subsea equipment that formulates requirements related to reliability testing in the oil and gas industries are the two API recommended practice (RP) documents 17N:2018 (addendum 1) and 17Q:2018. These provide guidance on reliability management and relevant techniques, for qualification and management of all types of equipment applied in subsea operations, and for all life cycle phases, for example related to design life. These are particularly relevant for development of new fields, extensions, upgrades and modifications. The RPs provide additional guidance to augment company specific requirements and practices. And offers strong recommendations on how to achieve good engineering practice in a subsea operations context, and is widely adopted by the industry.

Related to testing, particularly API RP 17Q:2018, and Annex A "Statistical analysis of test data" particularly, contains example statistical methods for estimating reliability data for equipment based on performance verification test data. The examples below (see Section 4) address such verification test data for a particular type of equipment.

These requirements and verification methods reflect project and company objectives. In most oil and gas capital projects, the end user or asset owner (operator), delivers technical and functional specifications to technology developers or vendors. Complex subsea machinery such as pumps, compressors and power systems are examples where reliability requirements and performance targets may be associated with a design life, and are often stated in terms of survivability or reliability over a given time periods that are estimated based on anticipated overhauls or replacement times. Additional performance targets may include system availability and related requirements associated with maintainability and equipment replacement times. It is good practice is to ensure that requirements are not vague, and that reliability and failure definitions are associated with the technology (equipment) functions, the operating environment and the type of events classified as failure (O'Connor and Kleyner 2012).

Reliability testing is, however, usually aimed at addressing failure modes at component level, which have been identified via various failure mode identification studies, engineering calculation and knowledge supported by empirical data history for similar fielded items. There are a number of challenges, some of them are described in 3.3 below.

It was stated earlier that reliability requirements and performance targets are usually developed by the end user, purchaser or asset owner. reliability targets for lower level components then need to be calculated via reliability allocation or apportion methods. This is usually carried out by the technology developer who should, in theory, demonstrate how the component reliability targets support the higher unit/package or system level requirements. The system requirements in turn support the overall development targets often expressed as a form of production availability or system utilization metric in capital projects.

3.3 Requirement related challenges

For complex technologies such as subsea power systems and speed drives, selection of the best reliability test for components is not a straightforward task. Recommended tests may yield accelerated test results that are given in terms of 'mean time to failure' (MTTF), which do not lend themselves well to the usual reliability allocation techniques, especially where there are complicated component and functional redundancy arrangements. So, in addition to new and specific tests being designed for the actual components, there may be a need for innovation around how the relationship between, say electrical/electronic component reliability and the system reliability and the specified performance targets.

A typical issue to be addressed is then not the difficulty of obtaining a result for the component test, but actually understanding and quantifying what the component performance target should actually be. For systems with 100's of similar components in complex arrangements this can be a significant challenge, for example understanding whether the target is 99.9% or 99.999% for a given operating period, and at what level of confidence levels does the item need to be demonstrated at.

Another difficult issue is related to test times. For new or modified technologies, including proven equipment intended to be used under different operating, performance or environmental conditions, there are challenges in terms of how long it takes to complete and analyse tests. In capital projects, qualification activities are usually carried out prior to detailed engineering phases and the point where final investment decisions are made. The API RP 17N/Q:2018 Technology Readiness Level (TRL) recommends that items that have achieved TRL 4 or greater are considered proven technology (not necessarily field proven however). This is the level where the technology has met product validation and environmental condition testing requirements. There can, however, be friction between the time needed to design and complete the testing needed to provide evidence of TRL 4 or above, and the project schedule. For example, where a test may take several months, which threatens to overrun the project schedule milestones.

An example of this challenge is given in the following section.

4. Subsea valve examples and discussions

In this example, we will first consider the reliability of subsea large bore valves subject to reliability testing. In the recent years, the expected field life of major subsea gas field developments have increased, and design life up to 50 years have been specified for subsea production systems, including the subsea large bore valves. Typical valve diameters for these are in the range 8 to 14 inches, and even larger valve sizes exist. These valves can be installed in water depths greater than 1,000 meters, and eventual failure and repair of these valves is associated with severe costs.

Historically, typical design life and actual time in service have been in the order of 10 to 15 years for such items. Furthermore, it should be noted that the average operational time for subsea valves registered in the industry recognized subsea reliability database OREDA (see oreda.com) is much lower, i.e. around 5 years.

We may then question whether it a reasonable assumption that historic reliability data is applicable for prediction of long-term performance of the new valves models, which are designed for 50 years continuous operation on the seabed?

A similar observation is made in Kleyner (2008), where it is pointed to the discrepancy between reliability predictions reflecting the inherent reliability of the product, on one hand, and the reliability required from the testing, on the other, often being higher; which could obviously be confusing and lead to misunderstandings.

As mentioned in the previous section, API RP 17N (2018) covers example methods for how to use verification test data in estimation of equipment reliability. Although it is explicitly stated that the methods are inappropriate for "reverse use" to determine the level of testing needed to achieve a certain performance requirement, anyhow this is commonly being done by engineers facing the problem of defining the number of test cycles required to demonstrate compliance with reliability requirements.

The methods are typically based on chi-squared statistics, involving several terms and parameters, which may be difficult to interpret practically by product engineers, including "degrees of freedom", confidence levels, "lower limit" expressions for reliability, required reliability at end of lifetime etc.

For the mentioned 50 years design life, several subsea valves are rarely operated, representative expectations can be in order of 4 times per year in average, representing in total 200 cycles over the field life.

Simplistically, assuming that reliability performance is dictated by number of cycles only, and that representative test and test conditions is feasible, a typical question is how many "failure free" test cycles are required "to qualify the valve for 50 years operation".

Below, a range of statistical interpretations observed in projects is given for the number of cycles required:

- 1,315 cycles: The reliability of the valve is minimum 90% after 50 years at 50% confidence level
- 4,371 cycles: The reliability of the valve is minimum 90% after 50 years at 90% confidence level
- 11,678 cycles: The reliability of the valve is minimum 95% after 50 years at 95% confidence level
- 91,639 cycles: The reliability of the valve is minimum 99% after 50 years at 99% confidence level

Even before taking into account the actual failures mechanisms and failure modes of the valves, a wide range and variation of test requirements originates from differences in the interpretation of risk acceptance and statistical terms and parameters. It is reasonable to believe that the same range/variation will be evident when interpreting the reliability test results, for instance when utilizing test data to derive failure rates or other reliability parameters as input to higher level reliability or availability models.

Strictly speaking, conducting a "50-years cycle test" for a conventional valve, i.e. with reliability demonstrated for 10 to 15 years in service, is likely to be successful, because the number of operating cycles are normally not the limiting factor for such "high reliability" valves. If the assumption of test applicability is not valid, obviously, the confidence in the test results will decrease when seen from the user's point of view.

Normally, a selection of techniques and methods described by ISO 20815:2018 will be applied by the engineers in order to zoom into the dominating failure modes, failure mechanisms and technical risks relevant for the new valve designs, and reliability test programs are defined accordingly. A specific threat in the valve example considered here is the failure "sticking valve when cycled after a sustained period in fixed position". How to construct a practicable reliability test to verify that the valve can be cycled, with sufficient degree of confidence, after several years in fixed position?

Related to the issue of significant time needed to complete the testing, we refer to a case of a directional control valve in a subsea pump, where the expected test would have impacted the project schedule as it was estimated to take over a year to complete. Leveraging the methods provided in Annex F of API RP 17N:2018, and Annex A of API RP 17Q:2018, a new test was satisfactorily designed which maintained the reliability targets, but which allowed adjustment of the confidence level, number of acceptable failures that could occur during the test and crucially, a reduction in the number of test cycles. This significantly reduced the time needed to demonstrate reliability performance targets from over a year to around three months.

One common strategy for mitigating the uncertainty associated with a test programme being "off target" is to utilize reliability methods such as FMECA combined with highly accelerated life testing (HALT) to identify and remove "weak spots" and dominating failure modes/mechanisms from the design, followed by regular accelerated life testing (ALT) or valve cycle testing to develop (and build further confidence in) reliability test data.

The many aspects of uncertainty demonstrated by the subsea valve examples are real, and represents a threat to the usefulness of test results. In some industry projects where the total picture of uncertainty is properly addressed, and found overwhelming to the project team; quantitative measures have been disregarded completely, and instead the project has performed "qualitative assessments", by considering the strength and value of reliability testing activities specifically and production assurance or design & development management activities in general. The outcome can be claims (or assumptions) such as "the reliability of the new design is at least better than the previous design", substantiated by the full range of reliability performance related activities performed.

Nevertheless, the aspects of uncertainty are still relevant, and several key questions remain:

- Is the test actually applicable to cover the relevant failure mechanisms related to extended design life?
- How confident are we that the dominating failure modes and mechanisms have been identified and mitigated, and what is subsequently the impact on reliability performance predictions?
- Which reliability data to use as input to overall system models, and how can the total picture of uncertainty be communicated/transferred?
- What is the actual confidence level in the results, taking all aspects of uncertainty into account?

These are key aspects when discussing the usefulness of the results. Hypothetically, a significant step forward would be to reverse the questions to derive a checklist (wish list) of supporting information to accompany the test results, such as:

- The results should state applicability of the results, with limitations and assumptions.
- Analysis approach, methodology and scope (boundary) should be summarized, and known gaps and simplifications should be listed etc.
- Recommended data and guidance on how to use data should be provided.

- Explicitly define the confidence level associated with predictions. Additionally, emphasize other aspects of uncertainty, and list the main uncertainty factors.

In many situations, the tests are conducted with intention of demonstrating very high reliability. This intention, combined with perhaps a strong manufacturer incentive, could easily drive favourable assumptions when interpreting testing requirements. For example, when selecting the appropriate reliability model, parameters could be selected based on assumptions that the limited sample size are able to reflect the reliability on an accurate level as long as the desired reliability level is achieved with confidence. This could be challenged by for example a questionable practise to re-test items failing based on claims that the particular item failing was not representative. It is not avoidable that reliability testing has a considerable uncertainty aspect, and that the interpretation of requirement is part of that.

This claim is supported by Kleyner (2008), pointing to different assumptions and complexities of reliability demonstration linked to uncertainty, several which can be linked to the requirements: e.g. the demonstration of reliability can require use of sophisticated theoretical understanding of reliability testing models, leading to simplifications and assumptions; and again, the transformation of design life in field to testing design may be confusing – going from perhaps a couple of weeks of testing to 20 years of operation; and then there are issues related to how well the testing reflects actual environment, which may include failure factors not assumed present or assumed as irrelevant. In brief, the uncertainty of assumptions related to testing quality and the reliability demonstrated should be expressed.

A way to do this is as suggested in ISO 20815:2018 and in Section 3, i.e. a simple qualitative assessment of whether the assumptions related to the testing requirements build on strong background knowledge and how sensitive these are to the result of the testing.

5. Concluding remarks

Despite the considerable uncertainty, in the context of production assurance and reliability management, reliability testing can produce useful information related to the system and reliability performance. For some items, an accelerated testing is required to obtain this information, adding yet a layer of potential uncertainty.

In this paper, we have focused on uncertainty related to the requirements of testing based on the claim given in the international standard ISO 20815:2018 on production assurance and reliability management, that the obtained information could sometimes be "misleading". The discussion provided in this paper supports this, and presents reasons why one in management of the PAP should be careful when using information from the reliability tests, as the test requirements is likely to be misunderstood, making it difficult to apply these consistently in planning and execution; for example, when performing accelerated testing. Some relevant issues discussed above are:

- Unclear reliability relationships between items at component and system levels
- Unclear reliability targets
- A need to shorten time (duration) for testing
- Discrepancy between reliability targets and actual performance
- How to express the results from testing

A key is to handle uncertainties in an appropriate way, including the ones coming from assumptions made in the selection of reliability model, test design, etc.

There are obviously variations, although we only to a limited extent have covered the range of company specific requirements. Nevertheless, there are clearly different ways of interpreting and using the requirements, which strongly influence both assumptions made and the results, and which cause uncertainties being relevant to express in order to achieve useful information.

Acknowledgments

This paper may be relevant as contribution to standardization work. The authors are currently members of the ISO TC67 working group "Reliability Engineering and Technology". However, the paper may not reflect the opinion of this group and is not produced as part of any standardization initiative.

References

API RP 17N (Addendum 1). 2018. Subsea Production System Reliability, Technical Risk, and Integrity Management. American Petroleum Institute (API).

API RP 17Q. 2018. Recommended Practice on Subsea Equipment Qualification (2nd edition, May 2018). American Petroleum Institute (API).

Barabady, J, Markeset, T and Kumar, U. 2010. A framework for improvement of production plant performance using production assurance programs. *International Journal of Systems Assurance Engineering and Management*, Springer, March, vol. 1(1), pp 59–65.

Brissaud , F, Varela, H, Declerck, B and Bouvier, N. 2012. Production availability analysis for oil and gas facilities: Concepts and procedure. Proceedings of the 11th International Probabilistic Safety Assessment and Management Conference and the Annual European Safety and Reliability Conference, Helsinki, Finland, June 25-29.

Hobbs, G.K. 2008. HALT and HASS overview: The new quality and reliability paradigm. In: Mistra, K.B. (ed.), *Handbook of Performability Engineering*. Springer.

IEC 61123.2019. Reliability testing - Compliance test plans for success ratio. International Electrotechnical Commission (IEC).

IEC 61163-1. 1998. Reliability stress screening - Part 1: Repairable assemblies manufactured in lots. International Electrotechnical Commission (IEC).

IEC 61163-2. 1998. Reliability stress screening - Part 2: Electronic components. International Electrotechnical Commission (IEC).

ISO 20815. 2018. Petroleum, petrochemical and natural gas industries — Production assurance and reliability management. International Organization for Standardization (ISO).

ISO 9000. 2015. Quality management systems — Fundamentals and vocabulary. International Organization for Standardization (ISO).

ISO 9001. 2015 Quality management systems — Requirements. International Organization for Standardization (ISO).

ISO/TR 16194. 2017. Pneumatic fluid power — Assessment of component reliability by accelerated life testing — General guidelines and procedures. International Organization for Standardization (ISO).

Kececioglu, DB. 2002. *Reliability & Life Testing Handbook*, DEStech Publications.

Kleyner, A. 2008. Reliability demonstration in product validation testing. In: Mistra, K.B. (ed.), *Handbook of Performability Engineering*. Springer.

O'Connor, PDT and Kleyner, A. 2012. Practical Reliability Engineering. Wiley.

Selvik, JT and Aven, TA. 2011. A framework for reliability and risk centered maintenance. *Reliability Engineering and System Safety*, Elsevier, 96(2), pp 324-331.

Vassiliou, P, Mettas A and El-Azzouri T. 2008. Quantitative accelerated life testing and data analysis. In: Mistra, K.B. (ed.), *Handbook of Performability Engineering*. Springer.