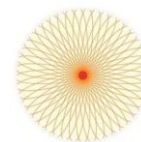




SINTEF



PDS



Report

Guideline for follow-up of Safety Instrumented Systems (SIS) in the operating phase

An APOS project report

Author(s):

Solfريد Håbrekke (SINTEF), Stein Hauge (SINTEF) and Mary Ann Lundteigen (NTNU)

Report No:

2023:00107

Client(s):

Multiclient

Report

Guideline for follow-up of Safety Instrumented Systems (SIS) in the operating phase

KEYWORDS

SIS follow-up
Maintenance
Functional safety

VERSION

3

DATE

2023-03-22

AUTHOR(S)

Solfrid Håbrekke (SINTEF), Stein Hauge (SINTEF) and Mary Ann Lundteigen (NTNU)

CLIENT(S)

Multiclient

CLIENT'S REFERENCE

Erik Korssjøen

PROJECT NO.

102020273

NO. OF PAGES

59

SUMMARY


This guideline describes a best practise for follow-up of safety instrumented systems (SIS) during operation of a process facility. It covers management of functional safety, operation, maintenance, monitoring, and management of change. Methods for updating failure rates and optimising test intervals are presented.

The document is an update of SINTEFs SIS follow-up guideline published in 2008 and 2021.

PREPARED BY

Solfrid Håbrekke, Stein Hauge, Mary Ann Lundteigen

SIGNATURE


Solfrid Håbrekke (Mar 21, 2023 15:55 GMT+1)

CHECKED BY

Maria V. Ottermo

SIGNATURE



APPROVED BY

Anita Øren

SIGNATURE



REPORT NO.

2023:00107

ISBN

978-82-14-07939-5

CLASSIFICATION

Unrestricted

CLASSIFICATION THIS PAGE

Unrestricted

Document history

VERSION	DATE	VERSION DESCRIPTION
1	2008-12-01	Edition 1 – SINTEF report A8788
2.1	2020-01-10	First draft of Edition 2
2.2	2020-07-07	Second draft of Edition 2 based on comments on first draft
2	2021-02-15	Final version of Edition 2 based on comments on second draft and workshop
3	2023-03-22	Edition 3 - based on comments to Edition 2

Table of contents

Preface	5
1 Introduction	6
1.1 Objective	6
1.2 Updates of this guideline compared to Ed. 1 (2008) and Ed. 2 (2021)	6
1.3 Motivation and target groups	6
1.4 Abbreviations	7
1.5 Notation	9
1.6 Regulatory requirements	10
2 Requirements from the project phase as input to SIS follow-up	11
3 SIS follow-up activities in operation	14
3.1 Main activities	14
3.2 Facility specific SIS follow-up procedure	16
3.3 Responsibilities and competency requirements	16
3.4 Outputs from and frequency of SIS follow-up activities	18
3.5 Handling of SIS failures	22
3.6 Handling of bypasses (inhibits and overrides)	22
3.7 Follow-up of demand rates, diagnostic coverage, and proof test coverage	23
3.8 Maintenance activities incl. proof testing	23
4 Preparations for SIS follow-up	25
4.1 Equipment groups and corresponding failure rates	25
4.2 Required input for optimising test intervals	25
4.3 Follow-up of individual components	26
4.4 Performance requirements and indicators	26
4.4.1 Dangerous Undetected (DU) failure rate from design	27
4.4.2 Expected (or acceptable) number of DU failures	27
4.4.3 PFD requirement for a complete SIF	27
4.4.4 PFD requirement / PFD budget for groups of components	28
4.4.5 Failure fraction (FF)	28
5 Method for updating failure rates and optimising proof test intervals	29
5.1 Assumptions	30
5.2 Select observation period	31
5.3 Review and classify SIS failure notifications	33

5.4	Identify outliers (Bad actors)	35
5.4.1	Component outliers.....	35
5.4.2	Equipment group outliers.....	35
5.5	Assess mitigating measures to remove systematic failure cause(s).....	36
5.6	Update failure rates	36
5.7	Compare with performance requirement(s)	40
5.8	Update test intervals.....	41
5.8.1	M1 – Method for optimising test interval based on failure rate	42
5.8.2	M2 – Method for optimising test interval based on PFD budget.....	43
5.8.3	M3 – Method for optimising test interval based on PFD requirement of SIF	44
5.8.4	Qualitative aspects to consider when changing the test interval	45
6	Proof testing alternatives.....	48
6.1	Crediting demands or activations as proof tests	48
6.2	Impact of partial stroke testing (PST)	49
7	References	51
A	Estimating DU failure rate using operational experience only	53
B	Multi-sample estimators – Failure rates based on data from two or more facilities	55
C	Crediting mitigating measures by adjusting the number of systematic DU failures	57

Preface

The work described in the report has been carried out as part of the research project “Automated process for follow-up of safety instrumented systems” (APOS). We would like to thank everyone for comments and valuable input to this work. The APOS project has received funding from the PETROMAKS 2 programme, The Research Council of Norway and PDS-forum.¹



This report is an update of the 2008 and 2021 editions of the guideline for follow-up of safety instrumented systems (SIS) in the operating phase. The present edition (3rd) also includes input from working groups in ISO 14224 and ISO/TR 12489. The main purpose of the APOS project has been to simplify and standardize reporting and classification of SIS failures, including the classification of safety equipment, and to provide a basis for increased automation and standardisation of SIS follow-up. The APOS project comprises seven related activities:

1. H1: Guidelines for standardised equipment classification and failure reporting [1]
2. H2: Potential for automated follow-up of safety equipment [2]
3. **H3: Guideline for follow-up of Safety Instrumented Systems (SIS) in the operating phase (this report)**
4. H4: Standardised/electronic SRS format [3]
5. H5: Information model for functional safety [4]
6. H6: Project summary and presentation
7. H7: PDS Data handbook, 2021 Edition [5]

Trondheim March 2023

¹ For more information about PDS-forum and the APOS project, reference is made to: <https://pds-forum.com/>

1 Introduction

1.1 Objective

The main objective of this guideline is to provide a best practise for follow-up of safety instrumented systems (SIS) in the operational phase of a process facility. This includes conformance to safety integrity level (SIL) requirements to ensure that the SIS performance is maintained throughout the operational lifetime of the facility. Since SIS are used to implement many of the safety barriers on a process facility, the guideline also applies for barrier management in the operational phase.

Focus is on SIS follow-up during operation, covering management of functional safety, SIS operation, SIS maintenance, SIS monitoring and SIS management of change. In particular, the guideline describes methods for updating failure rates and optimising test intervals based on operational experience. The main application of the guideline is for onshore and offshore (including subsea) SIS equipment. However, most of the recommendations are also relevant for non-instrumented safety equipment (e.g., process safety valves and fire dampers).

It should be noted that in Norway, the Offshore Norge guideline 070 [6] on application of IEC 61508 [7] and IEC 61511 [8] is widely applied. This guideline is in the following denoted GL-070 and the current guideline can be considered as a supplement to GL-070.

1.2 Updates of this guideline compared to Ed. 1 (2008) and Ed. 2 (2021)

This guideline is an update of the guideline published in 2008 [9] and 2021 [10]. The first edition (Ed. 1) of the guideline has been widely adapted by the Norwegian petroleum industry and used as a basis for developing SIS follow-up systems by several operators. Important updates of Edition 2 as compared to the first edition were:

- More detailed description of management of functional safety, including activities, responsibilities, and competence requirements.
- Possibility of having more than one observation period when updating failure rates and test intervals.
- Focus on specific follow-up of outliers (bad actors) and repeating failures.
- Possibility of crediting the implementation of mitigating measures to prevent systematic failures.
- One recommended approach for updating failure rates; using Bayesian estimation.
- Three approaches for optimising test interval based on updated failure rates.
- Proof test alternatives, such as crediting demands or activations as tests.
- Appendices including alternative method for updating failure rates based solely on operational experience, description of multi-sample estimator, and correction factors for mitigating measures.

In the preparation of this third edition of the guideline, we have received comments from persons participating in the maintenance of ISO 14224 and ISO TR 12489. Most of these comments have been incorporated.

1.3 Motivation and target groups

This guideline aims to improve the efficiency and quality of SIS follow-up by:

- Providing a standardised approach for updating failure rates and optimising test intervals considering relevant qualitative aspects.
- Facilitating improved utilisation and sharing of operational data between operators and between operators and engineering companies, consultants, and SIS vendors.

The guideline is relevant for several categories of personnel working with SIS. Target groups include:

- Personnel responsible for SIS follow-up and operational barrier management.
- Personnel responsible for developing and configuring SIS follow-up systems and applications.
- Personnel performing data analysis and assessment regarding SIS follow-up.
- Personnel classifying and/or performing quality assurance of failure data.
- Personnel performing maintenance and writing notifications and work orders related to SIS (to inform about how the data is used and the importance of high-quality failure registration and classification).

1.4 Abbreviations

ALARP	–	As Low As Reasonably Practicable
APOS	–	Automated Process for Follow-up of Safety Instrumented Systems
ASR	–	Automatic Shutdown Report
BPCS	–	Basic Process Control System
CAP	–	Critical Action Panel
CCF	–	Common Cause Failure
CCR	–	Central Control Room
CE	–	Conservative Estimate
C&E	–	Cause and Effect
CMMS	–	Computerised Maintenance Management System
DC	–	Diagnostic Coverage
DD	–	Dangerous Detected
DU	–	Dangerous Undetected
ESD	–	Emergency Shutdown
FF	–	Failure Fraction (Reporting measure specified by PSA)
FMECA	–	Failure Mode, Effect and Criticality Analysis
FSA	–	Functional Safety Assessment
FSM	–	Functional Safety Management
FSMP	–	Functional Safety Management Plan
GL	-	Guideline
HAZID	–	Hazard Identification Study
HAZOP	–	Hazard and Operability study
HMI	–	Human Machine Interface
HSE	–	Health, Safety and Environment
IMS	–	Information Management System
IEC	–	International Electrotechnical Commission
ISO	–	International Organisation for Standardisation
LOPA	–	Layer Of Protection Analysis
MoC	–	Management of Change
MTTF	–	Mean Time To Failure
MTTR	–	Mean Time To Restore
NOROG	–	Norwegian Oil&Gas Association



O&M	–	Operation and Maintenance
PFD	–	Probability of Failure on Demand
PFD _{avg}	-	Average PFD.
PFH	–	Probability of Failure per Hour
PHA	–	Process Hazard Analysis
P&ID	–	Process and Instrument Diagram
PSA	–	The Petroleum Safety Authority Norway
PST	–	Partial Stroke Test
PTC	–	Proof Test Coverage
QRA	–	Quantitative Risk Analysis
RNNP	–	Trends in risk level in the Norwegian petroleum activity
SFF	–	Safe Failure Fraction
SAS	–	Safety and Automation System
SAT	–	Safety Analysis Table
SCD	–	System Control Diagram
SIL	–	Safety Integrity Level
SIF	–	Safety Instrumented Function
SIS	–	Safety Instrumented System
SRS	–	Safety Requirement Specification
VDU	–	Visual Display Unit



1.5 Notation

Parameter	Denomination	Description
$E(x)$	-	Number of expected DU failures
α_i	-	Uncertainty parameter of prior knowledge, based on period $i - 1$
β_i	hours	Uncertainty parameter of prior knowledge, based on period $i - 1$
$\lambda_{DU,0}$	per hour	Design DU failure rate
λ_{DU} (or $\lambda_{DU,i}$)	per hour	DU failure rate (for period i)
$\lambda_{DU-CE,i}$	per hour	Conservative estimate of DU failure rate for period i
$\lambda_{DU,i}^{90U}$ or $\lambda_{DU,i}^{70U}$	per hour	Upper 90% or 70% bound for $\lambda_{DU,i}$ based on one-sided confidence interval (using maximum likelihood estimate) or credibility interval (using Bayesian estimate)
λ_{DU-op}	per hour	Updated DU failure rate based solely on facility specific operational experience
n (or n_i)	-	No. of tags (during period i)
x (or x_i)	-	No. of DU failures (during period i)
t (or t_i)	hours	Calendar time (for period i)
T (or T_i)	hours	Aggregated operating time; $T = n \cdot t$ (or during period i ; $T_i = n_i \cdot t_i$)
τ (or τ_i)	hours	Test interval (optimised based on $\lambda_{DU,i}$)
PFD	-	The average probability of failure on demand, i.e. used with the same meaning as PFD_{avg}
$PFD-SIF_{operational}$	-	Calculated PFD for SIF based on updated failure rates
$PFD-Group_{operational}$	-	Calculated PFD for equipment group component based on updated failure rate
$PFD_{SIF-Req.}$	-	PFD requirement for the complete SIF
PFD_t	-	PFD budget/target allocated to an equipment group
$z_{\epsilon,v}$	-	Upper ϵ percentile of the χ^2 -distribution with v degrees of freedom, i.e., $P(\chi^2 > z_{\phi,v}) = \epsilon$

1.6 Regulatory requirements

This guideline may help operators comply with requirements specified by the Petroleum Safety Authority (PSA) Norway. Some relevant PSA regulations related to SIS-follow-up are listed below [11], [12]. Note that in the PSA guidelines to the Activities regulations §26 on Safety systems [13], it is stated that IEC 61508-1 Ch. 7.6, IEC 61508-2 Ch. 7.6 [7] and Offshore Norge GL-070 Ch. 10–11 [6] should be used. It should also be noted that the guidelines to the Technical and operational regulations, applicable to onshore facilities only, refer to the IEC 61508 [7] and 61511 [8] standards for safety functions and safety systems.

The Management regulations §5 Barriers

Barriers shall be established that at all times can

- a) identify conditions that can lead to failures, hazard and accident situations,
- b) reduce the possibility of failures, hazard and accident situations occurring and developing,
- c) limit possible harm and inconveniences.

Where more than one barrier is necessary, there shall be sufficient independence between barriers.

The operator or the party responsible for operation of an offshore or onshore facility, shall stipulate the strategies and principles that form the basis for design, use and maintenance of barriers, so that the barriers' function is safeguarded throughout the offshore or onshore facility's life.

Personnel shall be aware of what barriers have been established and which function they are intended to fulfil, as well as what performance requirements have been defined in respect of the concrete technical, operational or organisational barrier elements necessary for the individual barrier to be effective.

Personnel shall be aware of which barriers and barrier elements are not functioning or have been impaired.

The Management regulations §19 Collection, processing, and use of data

The responsible party shall ensure that data of significance to health, safety and the environment are collected, processed, and used for

- a) monitoring and checking technical, operational, and organisational factors,
- b) preparing measurement parameters, indicators, and statistics,
- c) carrying out and following up analyses during various phases of the activities,
- d) building generic databases,
- e) implementing remedial and preventive measures, including improvement of systems and equipment.

Requirements shall be set as regards the quality and validity of the data, based on the relevant need.

The Management regulations §21 Follow-up

The responsible party shall follow-up to ensure that all elements in its own and other participants' management systems have been established and function as intended, and that a prudent level exists for health, safety, and the environment.

This follow-up shall contribute to identify technical, operational, or organisational weaknesses, failures, and deficiencies.

The Activities regulations §26 Safety systems

The measures and restrictions that are necessary for maintaining the safety systems' barrier functions in the event of overbridging, disconnection, or other impairment, shall be set in advance. The compensatory measures shall be implemented as rapidly as possible when such impairment occurs.

The status of all safety systems shall be known by and available for relevant personnel at all times.

2 Requirements from the project phase as input to SIS follow-up

An important part of SIS follow-up is to ensure that all relevant SIS requirements including SIS performance objectives, assumptions, prerequisites, use premises and constraints from the design and pre-ops. phases are fulfilled during operation. Such information is e.g., contained in the safety requirement specification (SRS), operation and maintenance (O&M) procedures, proof test procedures, system control diagrams (SCDs), piping and instrument diagrams (P&IDs), cause and effects (C&Es) tables, operational instructions for alarm management, etc.

Table 1 provides important documents and analyses with requirements and assumptions related to SIS follow-up (based on GL-070 Table 8.5) [6].

Table 1: Important documents and analyses relevant for SIS follow-up

Documents/Analyses	Description	SIS follow-up relevance
Quantitative risk analysis (QRA)	Systematic approach to understand and estimate the risk from the planned operation to people, environment, and assets. The risk level is compared to overall risk acceptance criteria.	Contains assumptions related to SIS operation and performance.
Process hazard analyses (PHAs) Hazard identification study (HAZID) Hazard operability study (HAZOP) Safety analyses tables ² (SATs)	Systematic approaches for identifying and evaluating process hazards – without and with the safety related functions present.	Provides recommendations and requirements for operation, some of which may be related to SIS operation, e.g., bypasses during start-up, operator response upon a process deviation, etc.
SIL allocation report or SIL identification report	The identification of SIFs and the SIL requirements are documented in the SIL allocation report. The allocation may be based on e.g., GL-070 minimum SIL requirements, layer of protection analysis (LOPA) or risk graph.	Contains assumptions related to SIS operation. When using LOPA and/or risk graph for establishing the SIL requirements, these analyses will typically include several assumptions relevant for operational follow-up, e.g., manual intervention by operators or automatic response from other protection systems. When credit is taken for alternative protection systems, it should be ensured that the systems are sufficiently independent.
Safety requirement specification (SRS)	Initially derived from the determination of SIFs and allocation of SIL requirements at a facility. Remains a living document which is kept updated along with changes in facility operation, modifications, and premises.	Contains key information, such as main requirements and assumptions for designing and operating the identified SIFs and associated SIS components. The SRS should be updated during operation in case of major changes, e.g., to failure rates, test intervals, SIL requirements, response times, etc.

² Referencing specific approach (SAT) in ISO 10418 - Petroleum and natural gas industries — Offshore production installations — Process safety systems [14]



Documents/Analyses	Description	SIS follow-up relevance
Supplier SIL documentation (SIL certificate, Safety Manual)	The supplier SIL documentation addresses the Safety Manual delivery as intended in IEC 61511 and IEC 61508. It provides the necessary information about how to safely apply a device or assembly and integrate it into a complete SIF.	Addresses operational and maintenance prerequisites, failure rates and other reliability data, fault detection and constraints regarding configuration and operating environment.
SIL compliance report or SIL verification report	Demonstrates conformance with SRS and SIL requirements for all SIFs.	Requires update during operation, i.e., when: <ul style="list-style-type: none"> • SIFs are modified, or new SIFs are introduced. • An operational failure rate is significantly different from assumed failure rate in design (e.g. factor 2 or more or factor 0.5 or less difference in failure rate). • A test interval is extended beyond what is specified in the SRS / SIL compliance report.

Table 2 identifies SRS requirements and assumptions that are of *special relevance* with respect to SIL follow-up in operation, including safety integrity requirements, functional requirements as well as additional descriptions and information related to the plant, systems, SIFs, and equipment. Other SRS requirements and assumptions will however also apply, and the full content of an SRS is described in Offshore Norge GL-070 App. E [6]. Also, a specification for an electronic standardised SRS has been developed as part of the APOS project [3].

Functional safety: Part of the overall safety relating to the process and the process control system which depends on the correct functioning of the SIS and other protection layers (IEC 61511).

Safety integrity: The ability of the SIS to perform the required SIF as and when required. Ability includes both the functional response (e.g., closing a specified valve within a specified time) and the likelihood that the SIS will act as required (IEC 61511).

Table 2: SRS assumptions and requirements especially relevant for SIL follow-up

IEC 61511 req. #	Requirement description
a	A <u>description</u> of all the SIFs (/SIS) necessary to achieve the required functional safety (e.g., a cause-and-effect diagram, logic narrative)
b	A list of the <u>plant input and output devices</u> related to each SIF which is clearly identified by the plant means of equipment identification (e.g., field tag list, also see section 4.1)
f	The assumed <u>sources of demand</u> and <u>demand rate</u> on each SIF
g	Requirements relating to <u>proof test intervals</u>
l	<u>Response time requirements</u> for each SIF to bring the process to a safe state within the process safety time.
j	The required <u>SIL and mode of operation</u> (demand/continuous) for each SIF, including allocated PFD (or PFH) requirement if relevant
k	A description of SIS process measurements, range, accuracy and their <u>trip points</u>
l	A description of <u>SIF process output actions</u> and the criteria for successful operation, e.g., leakage rate for valves
o	Requirements relating to <u>energize or de-energize to trip</u> for each SIF
q	Maximum allowable <u>spurious trip rate</u> for each SIF.
r	<u>Failure modes</u> for each SIF and <u>desired [automatic] responses</u> of the SIS (e.g., alarms, automatic shut-down)
u	A description of the <u>modes of operation</u> of the plant and requirements relating to SIF operation within each mode
y	The <u>mean repair time</u> which is feasible for the SIS, taking into account the travel time, location, spares holding, service contracts, environmental constraints

3 SIS follow-up activities in operation

This chapter describes how to prepare for and execute SIS follow-up activities during operation. Key aspects of SIS follow-up are to monitor and maintain adequate SIS performance throughout the operational lifetime. The required SIS performance is given by the functional safety and safety integrity requirements as described in the SRS. The preparation starts during design, installation, and commissioning, while the execution covers the phases of operation, maintenance, and modifications.

3.1 Main activities

The main functional safety activities associated with SIS in the operational phase are (IEC 61511 [8], GL-070 [6]):

- Management of functional safety
- SIS operation
- SIS maintenance
- SIS performance monitoring, verification, and analysis
- SIS management of change (MoC)

The activities listed above form what we in this report refer to as *SIL-follow-up*. The relationships between the activities are shown in Figure 1.

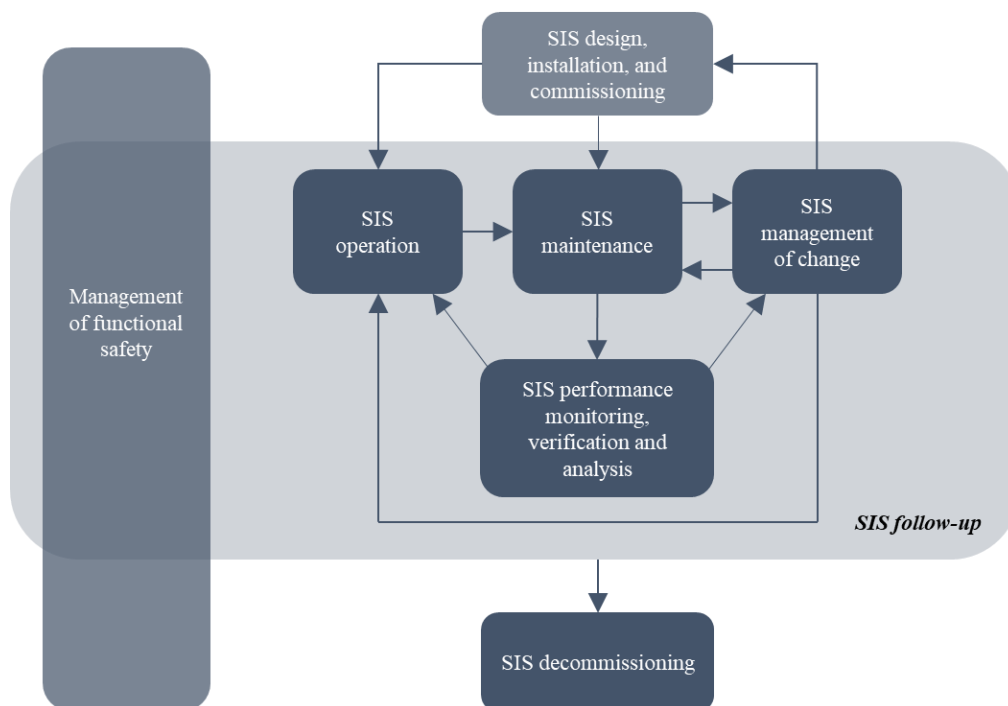


Figure 1: Main activities of SIS follow-up.

The tasks of the functional safety lifecycle activities are described in further details in IEC 61511-1, section 16: “Operation and maintenance” and section 17: “SIS modifications” [8]. The main tasks are briefly described below:

Functional safety management (FSM) focuses on the development and maintenance of procedures and practises relating to SIS operation, maintenance, follow-up, and competence/training development. Specific activities include:

- Establish and maintain the facility specific SIS follow-up procedure. See section 3.2.
- Define responsibilities and competence requirements. See section 3.3.
- Establish and maintain the (facility specific) functional safety management plan (FSMP) related to SIS follow-up. FSMP is not explained in more detail in this report and reference is made to GL-070 App. E.5 [6].
- Plan and execute functional safety assessment (FSA) stage 4 and 5. For a discussion of FSA, reference is made to GL-070, chapter 6 [13].
- Assessing overall adherence to design premises and assumptions as described in the SRS and other relevant documents (see chapter 2).

SIS performance monitoring, verification and analysis includes analyses and verification of the SIS performance requirements, such as PFD and SIL requirements, failure rates, demand rates and spurious trip rates. This activity also includes verification of the output requirements of all activities in all life cycle phases of the facility (e.g. FSA). See section 4.4 and chapter 6.

SIS operation includes normal interaction with the SIS during operation, i.e., monitoring the real-time status of SIS, start-up and shutdown, casual observation and of SIS equipment, reporting of identified failures, degraded states and non-conformities, initiation of maintenance requests or SIS modifications, implementation of mitigating measures if the SIS is degraded or unavailable and manage setting, resetting and status tracking of bypasses. See section 3.5 – 3.7 and GL-070, chapter 10 [6].

SIS maintenance includes scheduled proof testing, inspections, failure recording, repair, and overhaul, as well as replacements. Simple consequence- and possibly failure cause analysis (e.g. root cause analysis or 5-why analyses) to identify mitigating measures is also considered as part of the maintenance activities. SIS shall be regularly proof tested and maintained according to the SRS and as specified in the computerised maintenance management system (CMMS). In addition, proof testing is required after replacement/installation of new components and after firmware upgrade of logic solvers. SIS maintenance may, on an overall level be further split into:

- *SIS maintenance management and planning* focusing on ensuring the adequacy and quality of procedures for regular testing and maintenance of SIS devices, monitoring the quality of failure reporting, and planning of future SIS maintenance and testing activities.
- *SIS maintenance execution* focusing on the execution of regular testing and maintenance of SIS devices, including failure reporting through notifications and work orders.

This report does not give further details on SIS testing and maintenance execution practises and reference is made to GL-070 App. F.4 for more information.

SIS Management of Change (MoC) focuses on sufficient planning, review, approval, and documentation of SIS modifications, to ensure that the safety integrity is maintained with given modifications and possible new requirements and assumptions. SIS MoC helps ensure that changes in SIS hardware, software, procedures, and work practises are carefully evaluated and controlled prior to implementation. Introducing a new



sensing/measuring principle or changing the set point of transmitters are two examples of possible changes. A dedicated MoC procedure should be applied when deciding the need for changes, evaluating the scope and cost-benefit of changes, and establishing a plan for implementation according to the SIS lifecycle phases. IEC 61511 emphasizes the need to return to the relevant lifecycle phase to capture all relevant implications of the modification. In some cases, this implies going back to the hazards and risk analysis, and possible update of functional or safety integrity requirements. Software modifications should always be treated as a SIS modification, and the implementation should follow the software development requirements in IEC 61508-3 [7] or IEC 61511-1, Section 12 [8]. This report does not give further details on the MoC process, and for SIS modifications in general, reference is made to chapter 11 in GL-070 [6] and ISO 20815 appendix C.5 [15].

The IEC standard does not describe in detail how the different life-cycle activities shall be organised within each company but specifies that safety planning shall take place to define all relevant functional safety activities that are required to be carried and further specify the roles or organisational units to carry out these activities.

3.2 Facility specific SIS follow-up procedure

It is recommended to specify how the SIS follow-up activities are carried out, for example in terms of a facility specific SIS follow-up procedure. Following the requirements to the SIS follow-up activities, the SIS follow-up procedure should cover:

- Identification of project documentation (e.g., SRS and SIL verification report) that should be kept updated during operation. See chapter 2.
- Definition of roles (persons/departments) and responsibilities in SIS activities. See section 3.3.
- Definition of competence requirements for each role involved in SIS activities. Plan and organise courses and training (e.g., e-learning, classroom, webinar, feedback) to maintain and validate the competency requirements. See section 3.4.
- Description of SIS follow-up activities and corresponding procedures and references. See section 3.5.
- Description of performance requirements for SIS monitoring. See section 4.4.
- Description of methods and approaches related to SIS follow-up. See chapters 4–5.

3.3 Responsibilities and competency requirements

The SIS follow-up procedure must be complemented by organizational measures relating to assigning responsibilities and ensuring that involved personnel have sufficient competence.

- According to chapter 5 in IEC 61511-1 concerning management of functional safety: "Persons, departments, organisations, or other units responsible for carrying out and reviewing each of the SIS safety life-cycle phases shall be identified and be informed of the responsibilities assigned to them" (see IEC61511-1, section 5.2.2.1) [8].
- Furthermore, IEC 61511-1 (section 5.2.2.2) emphasizes that everyone involved in lifecycle activities must have sufficient knowledge of functional safety to be accountable. Assigning someone the role as SIL responsible to follow-up the requirements to functional safety management in the operational phase may be an efficient way to enforce this IEC 61511 requirement [8].

Relevant competence requirements for functional safety related tasks may be decomposed as shown in Table 3. The table is based on GL-070, Section 5.3.2 [6] but with some adjustments. It has been attempted to indicate some difference in level of competence by introducing the three categories:

- *A: Advanced knowledge*, implying that the personnel should have a detailed understanding of the objective and importance of the task as well as the necessary skills to execute of the task.
- *B: Basic knowledge*, implying that personnel involved should have a basic understanding of the objective and importance of the task as well as some basic skills on how to execute of the task.
- *I: Informed*, implying that personnel involved should be informed about the objective and importance of the task, but do not need any specific skills on how to execute the task.

Table 3: Competence requirements for defined activities

Competence requirement	FSM	SIS perf. m.v.a.	SIS operation	SIS maint. mgmt.	SIS maint. exec.	SIS MoC
Be able to explain the role and purpose of SIFs in prevention and mitigation of hazardous events-	A	A	B	B	I	A
Be able to retrieve and check adherence to SIL requirements and SIL allocation assumptions.	B	A		B		B
Understand the functionality of the SIS equipment and be familiar with SIS operation procedures.	B	B	A	B	B	B
Be familiar with safety requirements, such as the SRS, and know how to find SIS requirements and information about operational and environmental constraints.	A	A	A	A	I	A
Be familiar with relevant regulations and standards, e.g., PSA regulations, IEC 61508/61511 and GL-070, and now where to find them.	A	A	B	B	I	A
Be aware of SIS related documentation, systems and application, and information that shall be kept updated during operation.	A	A	A	B	I	A
Understand why regular testing and reporting of failures are needed for SIS performance monitoring.	B	A	B	A	A	I
Understand maintenance and test procedures, including the component's safety function and fail/pass criteria.	B	B	B	A	A	B
Know how to use procedures, systems, and applications for failure recording and classification.	A	A	B	A	A	I
Understand and use the taxonomy for detection methods, possible failure modes and relevant failure causes for the different equipment groups.	I	A	B	A	B	I
Understand the importance of detailed and precise failure reporting in the CMMS	B	A	B	A	A	I
Be familiar with possible maintenance actions necessary to restore the equipment back to "as good as new" condition.	I	B	I	A	B	I
Understand basic concepts used in reliability assessments such as unavailability, failures rates, and PFD.	B	A	-	I	-	I
Perform PFD calculations to verify performance requirements.	B	A	-	-	-	-

Competence requirement	FSM	SIS perf. m.v.a.	SIS operation	SIS maint. mgmt.	SIS maint. exec.	SIS MoC
Understand the difference between dangerous vs. safe failure and detected vs. undetected failures.	B	A	B	B	B	B
Understand the importance of identification and avoidance of systematic failures, repeating failures, bad actors and common cause failures.	A	A	B	B	B	I
Be able to perform failure cause analysis, e.g., root cause analysis or 5-why, to investigate the failure cause and to identify mitigating measures.	B	A	I	I	I	-
Assess results from failure reporting and performance verification (PFD) to determine new test intervals and possible mitigating measures	B	A	I	B	-	-

The competence can be achieved in various ways, such as through external SIS/SIL courses, internal courses (classroom courses, e-learning, etc.), formalized training, and on the job training.

3.4 Outputs from and frequency of SIS follow-up activities

Table 4 provides further guidance on each of the SIS follow-up activities (based on GL-070 Table 10.1 [6]), including examples of outputs and frequency of execution. Following the requirement in IEC 61511 to ensure clear responsibilities, it is recommended to assign one responsible for each activity or group of related tasks. As already mentioned, IEC 61511-1 leaves it up to each company how to assign and organise this [8]. In cases where 3rd party personnel are involved in SIS follow-up, it is especially important to ensure proper allocation and implementation of responsibilities and work processes.

Table 4: Guidance on SIS follow-up activities, outputs, and frequencies (based on GL-070 Table 10.1 [6]).

Main activity	Tasks	Outputs (examples)	Frequency
General / FSM	<p>Make sure that personnel involved in SIS-related work have necessary competence.</p> <p>Make sure that tasks related to SIS follow-up have been allocated to relevant disciplines and disciplinary functions in the organization</p>	<ul style="list-style-type: none"> • Overview of personnel involved in SIS related activities and their fields of competency • SIS follow-up activities responsibility allocation matrix 	Continuously
SIS operation	<p>SIS operation during normal conditions, during bypasses (e.g., due to maintenance) and in degraded mode (i.e., when equipment has failed):</p> <ul style="list-style-type: none"> • Ensure that SIS operation is performed according to procedures, that operational constraints and assumptions are fulfilled, and that operating personnel has access to updated documentation. • Ensure that all failures and degraded states revealed upon operation, i.e., from diagnostic (in CCR), condition monitoring systems, real demands, or random observations (e.g. log rounds/walkarounds or corrective maintenance on other nearby equipment) are reported (in CMMS). See section 3.5. • Identify and suggest necessary actions and mitigating measures upon degraded barriers/SIS or abnormal operating situations (e.g. in specific types of notifications in the CMMS). • Log and control inhibits and overrides, particularly with respect to loss of barriers and safety critical elements. See section 3.6. 	<ul style="list-style-type: none"> • Audits/reviews of SIS operational practices • Logs that identify status of inhibits/overrides/bypasses of SIFs • (High quality) notifications in CMMS based on failures detected through online monitoring (diagnostics and CM systems) and incidental observations • Records of demand rates and spurious trip rates 	Continuously
	<p>Ensure follow-up of and continuously improvement of SIS operation:</p> <ul style="list-style-type: none"> • Ensure that the competency of personnel who work with SIS is adequate. Ensure that relevant and competent personnel are involved in day-to-day SIS follow-up activities – both at the facility and onshore. • Ensure that operations and maintenance procedures and documentation are updated, available and used as intended. • Ensure that systems, and work practises contribute to avoidance of and control with systematic failures, e.g. to identify and follow-up outliers/bad actors (see section 5.4). 	<ul style="list-style-type: none"> • SIS training matrix and records • Regular audits/reviews of SIS operational procedures and documentation • Change requests for SIS modifications 	Continuously



Main activity	Tasks	Outputs (examples)	Frequency
	<ul style="list-style-type: none"> Assist and interact with maintenance on SIS related questions, e.g. handling of abnormal and degraded SIS, discussing mitigating measures, or updating procedures and documentation. Identify and evaluate the need for SIS modifications or changes in procedures based on process changes, non-conformities, performance deviations, integrity degradation, and reported failures and suggested actions and mitigating measures from operating personnel. 		
SIS O&M	Assure access control to the SIS, including secure use of keys and passwords, such that only authorised personnel can access the SIS hardware and software.	Instructions for access control	Continuously
SIS maintenance	<p>Perform maintenance and proof testing, see section 3.7:</p> <ul style="list-style-type: none"> Ensure that maintenance and proof testing is performed according to CMMS and the maintenance and test procedures. Ensure that the information in CMMS and maintenance procedures is correct (e.g. quality assured) and updated. Ensure that failures revealed upon maintenance and testing are correctly reported in the CMMS (e.g. by regularly review and quality assure failure notifications). Report all failures and non-conformities from maintenance activities and test results. Repair or replace failed and degraded equipment. Suggest and initiate mitigating measures if relevant. Initiate failure cause analysis if relevant. Improve maintenance supportability and testability if relevant (and possible). 	<ul style="list-style-type: none"> Regular audits/reviews of SIS testing and maintenance practices (High quality) notifications in CMMS based on failures detected during testing and maintenance Continuous quality review of reported notifications Labelling of notifications with respect to failure type (DD, DU, S) and cause (random, systematic, CCF) Completed notifications and work orders Completed tests and repairs 	Continuously
SIS monitoring, verification, and analysis	Ensure that operational data are properly registered to monitor the SIL requirements. Quality assure SIS failure notifications and work orders. If required, perform failure (re)classification and supplement the notifications.	<ul style="list-style-type: none"> Provide input and feedback to technical personnel who performs failure reporting Regular review and possible correction of reported notifications and concerning their quality/completeness Completed failure classification 	Weekly or bi-weekly
	After shutdown / SIS activation, go through relevant system logs and reports to identify failures and degraded states and if required initiate failure notifications, work orders, or other follow-up activities (e.g. analyses, discussion, mitigating measures, etc.).	<ul style="list-style-type: none"> Review meeting to go through last periods failure notifications 	On demand

Main activity	Tasks	Outputs (examples)	Frequency
		<ul style="list-style-type: none"> Initiation of failure cause analysis for repeating failure occurrences 	
	Verify the SRS requirements: <ul style="list-style-type: none"> Ensure that performance requirements and corresponding performance indicators are verified regularly. See sections 4.4 and 5.7. Ensure that methods for updating failure rates and test intervals are in place. See chapter 5. Perform analyses to consider the possibility for updating proof test intervals. See chapter 5. Verify SRS requirements such as demand rates, spurious trip rates, diagnostic coverage, and proof test coverage. See section 3.8. Update relevant data in documents and/or applications. Verify that the SIS is operated in line with other assumptions and prerequisites from the project phase (max repair times, response times, operating and environmental conditions, useful life, maintenance, etc.). 	<ul style="list-style-type: none"> Calculation of updated failure rates based on facility specific operational history Recommendations concerning new test intervals SRS / operational review verification report 	From monthly to annually
	Take corrective actions if the actual performance deviates from the specified performance. See MoC activity below.	Change requests for SIS modifications	Continuously
	Consider performing FSA stages 4 and 5 as operational experience is gained and/or as a result of major modifications. See GL-070 Ch. 6.	FSA report	When required
SIS modification (MoC)	Based on SIS operation, maintenance and verification activities, audits, and FSA, identify any deviations and areas of improvements, assess the need for and implement necessary mitigating measures and changes (e.g., update proof test intervals, update maintenance and inspection programs, improve operational and maintenance procedures, implement stricter access criteria to SIS, perform design modifications, etc.).	<ul style="list-style-type: none"> Decision of new test intervals Modification properly justified through MoC documentation New and improved test procedures 	Continuously
	Ensure that only approved changes are made to the SIS – both with respect to technical updates, procedure updates, document updates, and competence updates.	Approved change report.	Prior to modification

3.5 Handling of SIS failures

SIS failures revealed during operation, testing and maintenance must be properly registered, classified and documented in the notifications and work orders in the CMMS. For high quality SIS follow-up, as detailed as possible description of the failure should be given in the free text fields of the notification / work order. Correct registration and automation of detection method and failure mode, together with well-written information about failure causes and maintenance actions, improves quality and efficiency of SIS follow-up. The failure information in the CMMS is used to update failure rates, to identify possible failure causes and mitigating measures, and to optimise test intervals. It is therefore essential that maintenance personnel performing failure reporting are properly trained and motivated to ensure high quality of notifications and work orders.

All failures or impairments should be recorded in the CMMS and initiate a repair action. (PSA activity regulations require dangerous detected (DD) failures to be corrected immediately). It must be ensured that *all* dangerous undetected (DU) failures, also those revealed from other systems such as IMS and condition monitoring systems, are registered in the CMMS.

Procedures should be established describing necessary actions in the event of a SIS failure or impairment. Some operators have established (tag specific) documentation with pre-defined compensating measures to be implemented in case of SIS failures or impairments. When repair of a critical failure for some reason is delayed, necessary compensating measures should be identified and implemented.

DU failures: Failures that prevent the component to perform its safety function (dangerous failure) and not revealed immediately/automatically. The failures are typically revealed during proof tests, on demand or casually. E.g., a blowdown valve that fails to open within the response time requirement upon test (PDS/APOS).

DD failures: Dangerous failures revealed immediately/automatically, e.g., by diagnostic self-tests (PDS/APOS).

3.6 Handling of bypasses (inhibits and overrides)

Bypasses, inhibits, and overrides are sometimes necessary during maintenance activities or start-up. However, the use of such means should be strictly controlled. One reason is that failures may be introduced due to improper setting or resetting. In addition, a hazardous situation may occur upon a process demand while the SIS is temporarily unavailable. Thus, procedures should be in place for use of bypasses, including:

- Provisions for use of bypasses.
- Instructions needed for setting, suspension, and verification of suspension.
- Precautions that should be taken to avoid introducing failures.
- Routines for logging and visualisation of status on bypasses in barrier panel, SIS panel, safety, and automation systems (SAS), and CMMS.
- Routines for communicating status on bypasses from one shift to the next.

Bypass: An action taken to override, defeat, disable, or inhibit a SIF and may be necessary to avoid process disturbances e.g., during testing.

Override: Bypass of an output, i.e., giving different action than intended. **Inhibit:** Stops an action, typically from an input element.



3.7 Follow-up of demand rates, diagnostic coverage, and proof test coverage

SRS requirements and operational assumptions such as demand rates, spurious trip rates, diagnostic coverage, etc. should also be monitored and verified.

Demand rate:

It is particularly important to regularly estimate demand rates during operation as the assumed demand rates in the SRS are prerequisites for the determination of the PFD/SIL requirement. Demand rates should only include real hazardous events and not false or scheduled activations.

Diagnostic coverage (DC) – the fraction of dangerous failures detected by automatic diagnostic tests:

Automatic diagnostic testing may reveal dangerous failures, particularly for sensor and logic elements. Assumptions regarding DC should be followed-up to ensure that the assumed 'instrument' DC is also the "actual" DC: Are there procedures to act "immediately" upon dangerous failure alarms and are the alarms immediately visible to the operator? Are the dangerous failures revealed by diagnostic testing "immediately" repaired? If dangerous failures are not acted upon "immediately", they should be considered as undetected in the reliability calculations [10].

Proof test coverage (PTC) – the fraction of dangerous undetected failures revealed by proof test:

When the reliability of the SIS is estimated, it is often assumed that proof testing is perfect and has 100% coverage, i.e., all dangerous failures not detected during normal operation are assumed to be revealed during a proof test. A proof test will however often not fully represent a real demand. For example, the testing of gas detectors will not be performed during a real gas leakage exposure, and a pressure transmitter may not be tested by increasing the pressure of a tank to the set point. See also ISO TR 12489, section 14.2.4 for PFD calculation upon imperfect proof testing [16].

Proof test: Periodic test performed to detect [all] dangerous hidden (DU) failures in a SIS so that, if necessary, a repair can restore the system to an 'as new' condition or as close as practical to this condition (IEC 61511).

Note: If the proof test can detect all dangerous hidden failures, the proof test coverage is 100 %. If the proof test is not able to detect all dangerous hidden failures, the proof test coverage is less than 100 %.

3.8 Maintenance activities incl. proof testing

Proof testing is performed to verify the function of a SIF. Proof tests of parts of the SIF (e.g., a pressure transmitter test) using inhibits and/or overrides are often considered sufficient provided that all parts of the function are verified through individual tests (or separate testing programmes). The SIF may include redundant elements, e.g., voting of sensors or multiple solenoids acting on same valve, in which case it shall be ensured that all parts of the SIF are tested and verified. In such cases it is important that only the component(s) that are covered by the test are being credited as tested.

After a proof test there is always a possibility that the equipment is not restored to "as good as new" condition e.g., due to ageing, incorrect calibration, or incorrect restoration of tested equipment. Procedures should be in place to prevent failures initiated during proof test. Other preventive maintenance activities than proof test (e.g., inspection, periodic replacement, lubrications, periodic verification of diagnostic error messages not alarmed in CCR, condition monitoring, etc.) may be necessary to ensure that the equipment is restored to an "as good as new" condition. It is then relevant to consult the SRS, the Supplier SIL documentation or the vendors, e.g., to check for assumptions regarding maintenance.

Other maintenance activities also reveal failures and contribute to improved reliability. Therefore, when considering change in the proof test intervals, consideration should also be made to other maintenance activities that may be affected by such a change.

4 Preparations for SIS follow-up

This chapter describes important prerequisites that must be in place as part of SIS follow-up, including the definition of equipment groups and suitable performance indicators.

Table 2 in chapter 2 included SRS requirements to be maintained throughout operation. These requirements and other relevant information should preferably be linked to their respective SIFs (and corresponding tags) in an information model or a database. Each tag must also be allocated to an equipment group and its failure rate, as well as a test interval.

4.1 Equipment groups and corresponding failure rates

Since the PFD of a SIF is mainly a function of the component failure rates and the proof test intervals, these two parameters should be focused during SIF follow-up. Equipment groups of comparable components must be established to ensure adequate confidence in the failure rate estimates. Hence, every component/tag must be allocated to an equipment group for which a corresponding average failure rate is estimated.

Equipment group: A group of comparable components on a facility, assumed to have similar functionality and comparable failure rate and the same failure rate distribution. Examples of equipment groups are IR gas detectors and blowdown valves (PDS/APOS).

SIS equipment groups are presented in [1] where three levels of equipment grouping are suggested. Commonly, equipment groups are defined with basis in so called safety critical elements [1], for example by defining groups of all level transmitters, all pressure transmitters, all gas detectors, etc. However, for safety critical elements where reliability performance will vary based on e.g., design, application, process and/or environmental conditions, follow-up should be done on a more detailed level (level 3). Example of such level 3 equipment groups are IR gas detectors, nuclear level transmitters and ball type shutdown valves. Hence, each tag will belong to a group of safety critical elements (level 2) but may also be further filtered by specifying a set of reliability influencing properties (level 3). See Hauge et al. [1] and ISO 14224 [17] for a more detailed discussion of equipment grouping and reliability influencing properties.

Appropriate equipment group level will in practice often be a trade-off between sufficient aggregated operating time (number of components) and capturing the important properties influencing the failure rate.

Operating time: The time interval during which a component is in an operating state. Operating time includes actual operation of the equipment or the equipment being available for performing its required function (ISO 14224).

It may also be of interest to compare the failure rates with other facilities. This is particularly relevant for operators with several comparable facilities (see section 5.4).

4.2 Required input for optimising test intervals

In section 5.8, three methods for optimising test interval, based on operational experience, are presented. The methods require both a common and some separate input – which are summarized in Table 5. It will differ between operators and facilities which method is preferred (e.g., whether SIF calculations readily are available).

Table 5: Required input for optimising test intervals for the three methods.

Method		Required input	Common preparation / input
M1	Method based on equipment group failure rate	<ul style="list-style-type: none"> The design failure rate for the equipment group and corresponding test interval(s) have been shown to fulfil the PFD/SIL requirements of the respective SIFs. 	<ul style="list-style-type: none"> Design failure rate for the equipment group – to perform failure rate update based on operational experience.
M2	Method based on equipment group PFD budget/target	<ul style="list-style-type: none"> Common PFD budget allocated for all components within the same equipment group, either: <ul style="list-style-type: none"> a fixed value (target) based on company specific PFD requirements, the strictest PFD budget allocated to the components in the equipment group from the PFD requirements of the respective SIFs. 	
M3	Method based on the SIF PFD requirement	<ul style="list-style-type: none"> PFD requirement of the SIF. SIF calculation tool (e.g., in Excel). 	

4.3 Follow-up of individual components

In addition to quantitative follow-up on equipment group level and/or SIF level, individual components should also be followed-up qualitatively, especially with respect to identifying repeating failures and outliers, performing failure cause analysis, and specifying mitigating measures.

Typical failure rate for a SIS component corresponds to a mean time to failure (MTTF) of 50–100 years. The likelihood of one component failing several times during a limited period (e.g., two years) is therefore low. When the same component fails more than once within such a short period, the failures are typically systematic in which more testing will not always improve the PFD and focus should rather be on failure cause identification and removal. It is therefore important to identify individual outliers, examine the underlying failure cause(s) and follow-up these components on an individual basis. See also sections 5.4–5.5.

4.4 Performance requirements and indicators

SIS performance shall be followed-up to verify that the experienced (or measured) safety integrity is acceptable as compared to the SRS requirements. Performance indicators should therefore be defined. Below, some SIL/PFD related requirements and corresponding performance indicators are discussed.

Note that a performance requirement in practice will be a threshold value of an associated performance indicator. E.g., the threshold value is the *maximum acceptable* number of DU failures per year for a specified valve population. The performance indicator is then the *measured* number of experienced DU failures per year.

Performance indicator: A quantitative measure of the *experienced* safety integrity of an equipment group or a SIF. Examples of performance indicators are number of DU failures, the failure rate, and PFD (APOS).

Performance requirement (or threshold value): The *maximum acceptable* value of the corresponding performance indicator (APOS).

4.4.1 Dangerous Undetected (DU) failure rate from design

The rate of DU failures from the design phase ($\lambda_{DU,0}$) is the (a priori) failure rate used to verify the PFD requirement from the SRS and can therefore be considered as a performance requirement. The associated performance indicator is then the updated DU failure rate from operation. The Bayesian failure rate estimation that *combines* the (a priori) design failure rate with additional (a posteriori) operational experience is recommended (see section 5.6), but the failure rate based solely on operational data may also be used (see Appendix A).

4.4.2 Expected (or acceptable) number of DU failures

The expected (or acceptable) number of DU failures is a performance requirement that can be derived from the DU failure rate assumed from design. By assuming a homogenous equipment group with n components, constant failure rate, and observation period t , the performance requirement becomes:

$$E(x) \approx n \cdot t \cdot \lambda_{DU,0} = T \cdot \lambda_{DU,0}$$

Here, $E(x)$ is the expected number of DU failures, $\lambda_{DU,0}$ is the DU failure rate *from design* (assumed to be constant and initially the same for all n component), and T is the aggregated time in operation for the population. Note that we here assume that each of the n components are tested or activated at least once during the observation period t . The performance indicator is then the number of experienced DU failures for the n components during the period t .

EXAMPLE

On a facility there are 500 smoke detectors with a DU failure rate from design, $\lambda_{DU,0} = 0.2 \cdot 10^{-6}$ per hour. Then, for the smoke detector population:

$$E(x) \approx n \cdot t \cdot \lambda_{DU,0} = 500 \cdot 8760 \text{ hours} \cdot 0.2 \cdot 10^{-6} \text{ hours}^{-1} \approx 1$$

Hence, the *expected* number of failures during one year of operation will be approximately one (1), which can then be used as a requirement for maximum acceptable number of DU failures per year for the smoke detector population.

Note: If the smoke detectors are tested more frequently than once per year, the same performance requirement can be applied. However, if the detectors are tested only every second year, then a requirement of 2 failures *per two years* should apply.

4.4.3 PFD requirement for a complete SIF

Based on the required SIL and associated PFD requirement as specified in the SRS, a PFD performance requirement for each complete SIF will be applicable for facilities designed according to IEC 61508/61511 [7], [8]. The corresponding performance indicator will then be the calculated PFD from the updated DU failure rates and the current test intervals for the components in the SIF.

4.4.4 PFD requirement / PFD budget for groups of components

The PFD can also be used as a performance requirement for groups of components with a defined PFD budget or target, PFD_t , see Table 5. The corresponding performance indicator is then the PFD for *the associated equipment group* calculated from the updated DU failure rate and the current test interval for the equipment group or component under consideration.

EXAMPLE

On a facility there are 50 process shutdown (PSD) valves that are part of SIL 1 rated loops, but with an additional specified PFD requirement *for the SIF* ranging from 0.015-0.025. The valves have an assumed DU failure rate from design of $\lambda_{DU,0} = 2.3 \cdot 10^{-6}$ per hour. The valves are *initially* tested each year, and the PFD budget for the valves has been defined as:

$$PFD \approx \lambda_{DU,0} \cdot \frac{\tau}{2} = 2.3 \cdot 10^{-6} \text{ hours}^{-1} \cdot \frac{8760}{2} \text{ hours} = 0.01.$$

Hence, the PFD requirement or PFD budget for the PSD valve population is 0.01. So, when estimating the updated average failure rate for the valve population *during operation*, this failure rate shall, together with the current test interval be verified against the PFD criterion of 0.01.

4.4.5 Failure fraction (FF)

Trends in risk level in the Norwegian petroleum activity (RNNP) [18] as reported by the PSA uses the failure fraction (FF) as a performance indicator to monitor the experienced integrity of SIS equipment. Some operators also use FF as a performance requirement. FF is defined as the *ratio between* the number of DU failures revealed during proof testing and the corresponding number of proof tests performed and must not be confused with measures such as SFF and DC. In RNNP the failure fraction cannot be directly related to the PFD since the test interval is not explicitly reported. An example of an FF performance requirement for an equipment group is 1% (i.e. a maximum fraction of one failure per 100 functional tests are allowed).

For equipment where most DU failures are detected upon testing, FF is a realistic performance indicator, e.g. for gas detectors and deluge valves. However, for equipment such as valves, fire doors, fire and gas dampers and pumps where DU failures are often revealed in-between tests, the FF is not considered a suitable performance indicator.

5 Method for updating failure rates and optimising proof test intervals

This chapter describes methods for updating DU failure rates and optimising proof test intervals for an equipment group (see section 3.5).

In the rest of this chapter 'failure rate' refers to the 'DU failure rate' and 'test interval' refers to 'proof test interval'.

The recommended method for updating failure rates is a Bayesian approach, where new operational experience is used in combination with prior knowledge about the failure rate. The approach may be iterative with several possible observation periods. Then, data from the previous observation period is used to obtain prior knowledge (input parameters) to the current observation period. For the first observation period, the prior knowledge is based on the design failure rate (see Figure 5).

Observation period: The interval of time (calendar time) between the start date and end date for failure data collection (ISO 20815). (Corresponds to *surveillance period* in ISO 14224).

Figure 2 illustrates the overall approach and the associated activities together with recommended frequency of each activity. Section 5.1 lists the assumptions. Section 5.2 gives some recommendations regarding selection of observation period. Section 5.3 gives a brief description of the review and quality assurance of SIS failure notifications. Section 5.4 discusses treatment of outliers (bad actors). Section 5.5 presents a simplified method for crediting mitigating measures. Section 5.6 presents the method for updating failure rates. Sections 5.7 discusses comparison with performance requirements. Section 5.8 presents methods for optimising test intervals and a qualitative checklist for updating intervals.

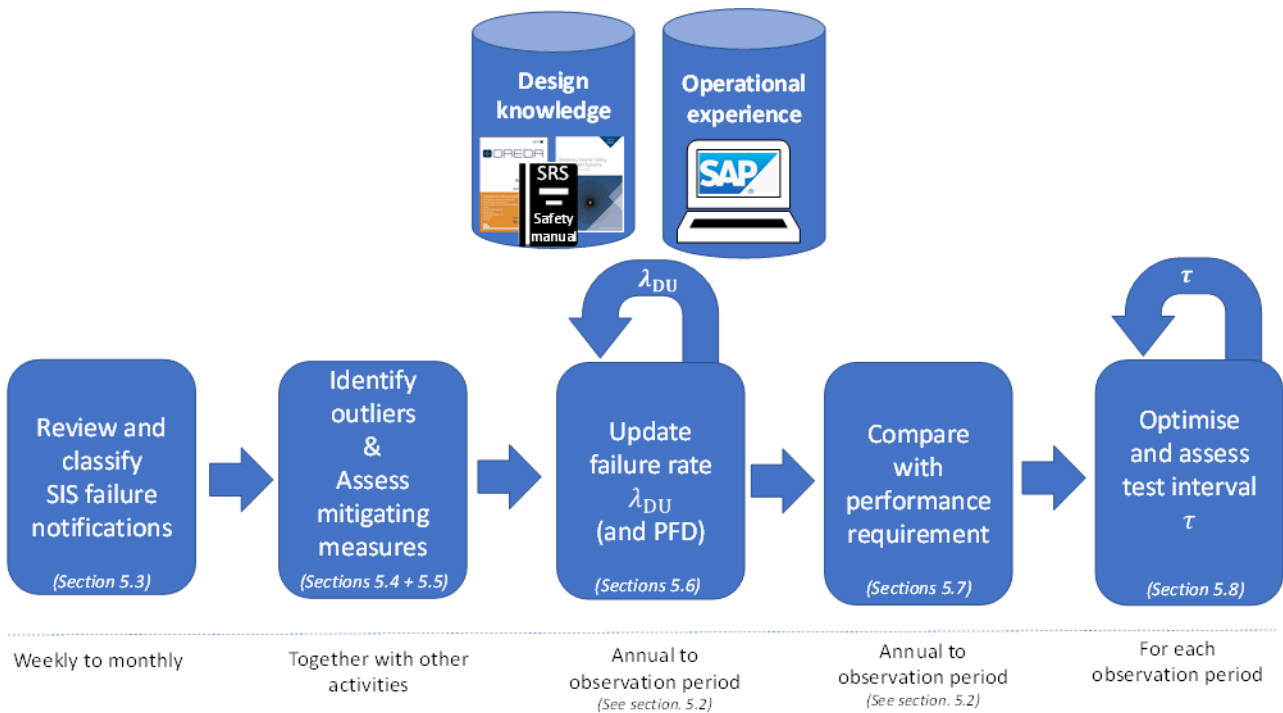


Figure 2: Overall method for updating failure rates and optimising test intervals.

5.1 Assumptions

The methods for updating failure rates and optimising test intervals are based on some assumptions summarised in Table 6. The validity of these assumptions should be verified prior to changing proof test interval or using the updated failure rates for other purposes. Deviations from the assumptions may require re-assessment of the failure rate estimation and the quantitative test interval optimisation. The table gives recommendations on necessary actions in case of such a re-assessment.

Table 6: Summary of assumptions regarding updating of failure rates and optimising test intervals.

Assumption for the equipment group	If assumption not fulfilled...
All components have been proof tested or activated at least once during the observation period.	...extend observation period or exclude components not activated.
All failures have been registered, classified and quality assured.	...perform additional review of possible failure events (notifications, automatic shutdown reports, etc.)
The components should as far as possible be considered as homogenous. (An equipment group with components from different facilities is considered as inhomogeneous.) All components within the same equipment group have a common functionality and comparable failure rate.	...consider more narrow groups. For instance, level transmitters may be split into groups depending on their measuring principle. ...identify possible outliers (bad actors), see section 5.4 ...see Appendix B for inhomogeneous samples. Inhomogeneous samples may reduce the precision and relevance of the updated failure rate.
The lifetime of a component ³ has an exponential distribution, and the failure rate is assumed constant (not time-dependent) during the observation period, i.e., the component is assumed to be within its useful life period.	... disregard systematic DU failures initiated during installation or commissioning ("burn-in failures" or failures due to specific ageing problems (beyond useful life period). Also, consider excluding bad actors and repeating failures, see section 5.5
The design failure rate is based on operational experience from comparable facilities. Ref also IEC 61511-1 (subclause 11.9.3) stating that the applied reliability data shall be credible, traceable, documented and justified and shall be based on field feedback from similar devices used in similar operating environment.	...if available use failure rates based on operational experience (such as PDS data) as design failure rates. ...gather more operational experience.
Time to perform proof test, repair or replace any failed component is negligible compared to the time between proof tests (or activations).	...adjust aggregated operating time, see section 5.2.

³ This assumption is generally valid for electrical, electronic, and programmable-electronic components, but in this guideline we also made the same assumption for some mechanical components such as valves. The rationales are: The valves are assumed to be in their useful life period (where the failure rate can be assumed constant). It is further assumed that a modification request is initiated in case the valves exceeds this phase, e.g. in case that inspections reveal an increased degradation over time or in cases where the calculated failure rates are increasing steadily over several observation periods.

5.2 Select observation period

The length of each observation period for an equipment group should provide *sufficient* aggregated operating time to increase the confidence of the updated failure rate, and in particular, to provide a sound statistical basis for recommending updated test intervals.

When to start a new observation period?

A new observation period for the equipment group could be initiated upon major changes in operation and maintenance, e.g.:

- When the proof test interval is changed.
- When a modification that affects the performance of the components is performed, e.g., several components have been replaced with new components with assumed improved reliability.
- When previous operational experiences are considered less relevant, e.g., due to removal of failure causes of previous DU failures.

There may be several observation periods (with various length) for an equipment group throughout the operational life of the facility. As illustrated in Figure 3, some components may not be operative from the beginning and new components may be added due to a modification at the facility.

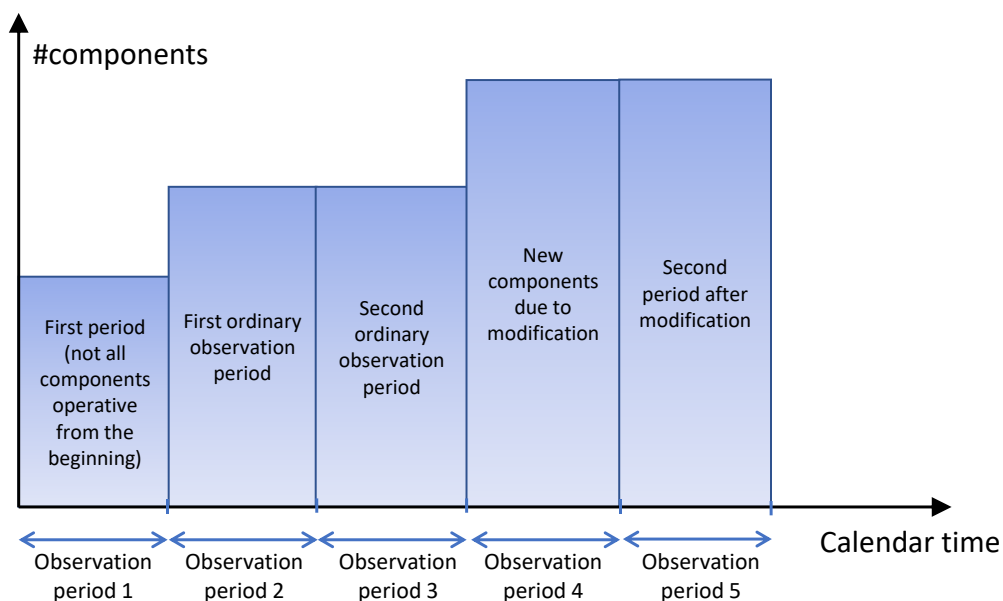


Figure 3: Variation in number of components between observation periods.

Note that it is also possible to continuously update the failure rate or apply shorter periods than the suggested observation periods for updating failure rates, e.g., due to annual failure/operational reviews, RNNP reporting (and subsequent failure rate updates and verification of performance requirements). It is then recommended to merge some of these shorter periods into observation periods of sufficient length when performing test interval optimisation (see below).

Aggregated operating time for an equipment group

The aggregated operating time for an equipment group within period i is denoted T_i , and is determined by the length of the observation period (t_i) and the number of components (n_i) within the group.

The following should be considered when calculating the aggregated operating time for an observation period:

- adjust for removed or added components during the period or since last period.
- exclude components not tested or activated in the period.
- exclude the operating time contribution from components that have been out of service or in passive standby during the period.

Sufficient operational experience to perform test interval optimisation

To obtain the necessary confidence in the test interval optimisation, *the operational experience from a new observation period i should preferably have a higher weight than the prior failure rate, $\lambda_{DU,i-1}$* . This implies that⁴:

$$\lambda_{DU,i-1} \cdot T_i > 1$$

which e.g., implies that if the input failure rate $\lambda_{DU,i-1}$ is $1 \cdot 10^{-6}$ per hour, then the aggregated operating time for period i should preferably exceed 10^6 hours.

Based on the above, the *minimum length of an observation period* should be:

$$t_i = \frac{1}{n_i \cdot \lambda_{DU,i-1}}$$

The corresponding suggested *minimum number of components within an equipment group* is:

$$n_i = \frac{1}{t_i \cdot \lambda_{DU,i-1}}$$

Table 7 gives suggested minimum number of components – for various input failure rates and observation periods. For instance, if the failure rate from the previous period (or from design) is $1.5 \cdot 10^{-6}$ for an equipment group with 60 components, the first observation period should as a minimum roughly be two years (60 is more than 38 but less than 76⁵). Alternatively, if the equipment group had 76 components, one year is minimum length for the observation period.

⁴Assuming (for simplicity) that the conservative estimate of the input failure rate is twice the input failure rate (assumption for period 1), see section 5.6. Then the Bayesian estimate for the updated failure rate becomes: $\lambda_{DU,i} = (1 + x_i) / (\frac{1}{\lambda_{DU,i-1}} + T_i)$ where x_i is the number of DU failures reported in the current observation period. The parameters of the Gamma distribution are recognized as $\alpha = 1$ and $\beta = 1/\lambda_{i-1}$. As we see, the weight is dependent of the operational experience within the current period (together with the input failure rate). We now consider the numerator (representing the no. of DU failures) and the denominator (representing the operating time), separately:

- If $x_i = 1$, the number of DU failures in period i is weighted 50% compared to the prior data. In general, the weight of the current period with respect to the DU failures, can be expressed as $1 - \frac{1}{1+x_i}$.
- If $\lambda_{DU,i-1} \cdot T_i = 1$, the operating time from period i is weighted 50% compared to the prior data (from previous periods). In general, the weight of period i regarding the operating time can be expressed as $1 - \frac{1}{1+\lambda_{DU,i-1} \cdot T_i}$.

⁵More precisely the observation period should as a minimum be $t_i = \frac{1}{n_i \cdot \lambda_{DU,i-1}} = \frac{1}{60 \cdot 1.5 \cdot 10^{-6}} = 1\,111$ hours \approx

15 months.



Table 7: Minimum number of components (n_i) suggested for an equipment group for period i for examples of failure rates ($\lambda_{DU,i-1}$) based on the previous period.

$\lambda_{DU,i-1}$	Years			
	1	2	3	4
	n_i			
$0.5 \cdot 10^{-6}$	228	114	76	57
$1.0 \cdot 10^{-6}$	114	57	38	29
$1.5 \cdot 10^{-6}$	76	38	25	19
$2.0 \cdot 10^{-6}$	57	29	19	14
$2.5 \cdot 10^{-6}$	46	23	15	11
$3.0 \cdot 10^{-6}$	38	19	13	10

5.3 Review and classify SIS failure notifications

SIS failure review should be performed regularly to verify SIL conformance (e.g., every year or every second year). However, quality assurance of failure reporting and classification in the CMMS should be carried out more frequently (preferably every second week). This will ease communication with personnel that are involved in failure reporting and thereby improve the quality of the reporting.

All failures within each equipment group should be classified into DU failures, DD failures, spurious/safe failures, or non-critical failures (such as degraded failures). Identification of failure causes, corresponding mitigating measures and their possible effects are also important, particularly for failures classified as DU. Repeating failures and common cause failures (CCFs) should be assessed in detail for all types of failures.

Degraded failure: Failures where the component functionality is reduced but still intact, and which over time *may develop into* a dangerous (or a safe) failure. E.g., reduced flow through a deluge valve that may develop into a DU failure (PDS/APOS).

As shown in Figure 4, failure causes can be split into random hardware failures and systematic failures. For random hardware failures, increased test interval is the main measure for reduced PFD. Systematic failures, can unlike random failures, be prevented “once and for all” if the specific causes that lead to the failure are removed (e.g., changing the calibration procedure or avoiding sandblasting in the specific area). Focus for systematic failures should therefore rather be on failure cause identification and removal rather than change of test interval.

Random hardware failures: Failures that occur due to normal ageing and degradation.

Systematic failures: failure related to a pre-existing fault, which consistently occurs under particular conditions, and which can only be eliminated by removing the fault by a modification of the design, manufacturing process, operating procedures, documentation or other relevant factors (IEC 61511-1).

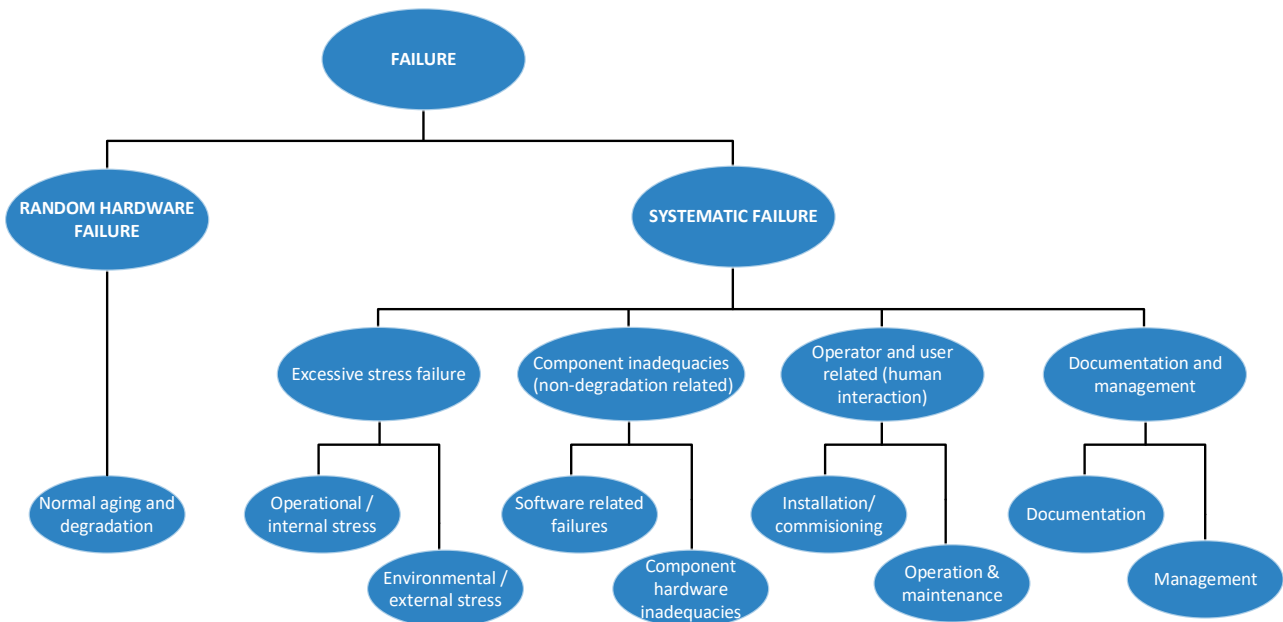


Figure 4: Random hardware failures vs systematic failures. Classification by cause [1]

During the first period after operation start-up or after a major modification, an increased number of DU failures is often experienced, e.g., due to inherent quality problems or installation and commissioning errors ("burn-in" failures). Such failures and their corresponding failure causes are especially important to identify and remove to prevent possible re-appearance.

Reporting and classifying failures consistently over time is a challenge. Therefore, it may be helpful to prepare a list of typical failures for each equipment group, and their suggested classification (DU, DD, degraded, safe/spurious or non-critical), to ensure equal classification of the same type of failures across persons, teams, and facilities (and ideally also between operators).

The following should be considered when classifying and counting DU failures:

- Whether a DU failure is a random hardware failure or a systematic failure. For systematic DU failures, see section 5.5.
- Whether there are any degraded failures that may develop into a DU failure. One should especially consider such degraded failures when suggesting an extended test interval.
- Whether some DU failures are repeating failures. Repeating failures of a component within a limited period of e.g. one year may be counted as one DU failure. Special follow-up of the relevant component(s) is required, see section 5.4.1.
- Whether any CCFs of DU failures are revealed. CCFs should be counted as one DU failure⁶.

Repeating failure: Two or more similar DU failures (normally systematic failures) for the same component due to the same cause within a further limited time period, e.g. within a test interval of one year (PDS/APOS).

Common cause failure: Systematic failures of two or more similar ("coupled") components (e.g., same equipment, location, same vendor, etc.) due to the same cause and within the same test interval (PDS/APOS).

⁶ The beta value may be re-assessed based on operational data together with a checklist for establishing beta values, see e.g., SINTEF (2015) [19] or IEC 61508-6 App. D [7].

5.4 Identify outliers (Bad actors)

5.4.1 Component outliers

A component that fails much more frequently compared to the rest of the equipment group, is called an *outlier* or a *bad actor*. If a component fails multiple times within the same test interval this is probably due to a specific problem such as design weaknesses, operational/maintenance issues, or environmental impact. Therefore, a component with repeating failures (*due to the same cause*) is also an outlier.

Component outlier (bad actor): A component that has experienced two or more DU failures (of any cause, but normally systematic ones) within the same test interval (APOS).

How to identify component outliers?

A component should be considered an outlier when *two or more DU failures of any cause are revealed for the component within the same test interval*. A typical test interval will be one year but can also be more (up to five years for control logic units with a very low failure rate)⁷.

How to handle outliers?

If an outlier has been identified, the following actions should be considered:

1. Assess the criticality of the component, e.g., single components in all SIFs and redundant components in SIFs with higher SIL requirements are the more critical.
2. If considered critical: perform root cause analysis of the problem(s).
3. Consider reducing the test interval for the component until the root cause is identified and removed.
4. Assess if the same problem is relevant for other components. Consider also to reduce the test interval for these components until the root cause is found.

5.4.2 Equipment group outliers

An equipment group for a given facility may be considered an *equipment group outlier* if its failure rate deviates significantly from similar equipment groups on other facilities. The failure rate of an equipment group outlier is either very high (e.g., due to specific design or operational issues), or very low (e.g., due to good design or possible underreporting).

Equipment group outlier: An equipment group population on a specific facility with a significantly higher experienced failure rate than the comparable equipment group on other facilities (APOS).

How to identify equipment group outliers?

To identify if an equipment group is an outlier, sufficient operational experience should be available:

- *Number of DU failures multiplied with the aggregated operating time should preferably exceed 3 million hours for the equipment group, i.e., $T \cdot x \geq 3 \cdot 10^6$ hours* (see Appendix A footnote 14).

The equipment group can then be considered as an outlier if:

- The estimated average failure rate is *outside the 90% OREDA multi-sample uncertainty interval* (see Appendix B) – when data from other facilities are available.

⁷ Two DU failures within five years corresponds to a failure frequency about 50 times higher than expected for SIS equipment with a failure rate of $1 \cdot 10^{-6}$ failures per hour.

- The estimated average failure rate is *more than 4 times the generic failure rate*⁸ (e.g., the PDS failure rate) – when data from other facilities are NOT available.

How to handle equipment group outliers?

If an equipment group outlier with a very high failure rate has been identified, the following actions should be considered:

1. Review the failure classification and reporting practice.
2. Consider more frequent testing and actions for the group until root cause is (significantly) removed.
3. Perform root cause analysis of the problem(s).

Review of failure classification and reporting practice is also highly relevant if a very low failure rate is experienced.

5.5 Assess mitigating measures to remove systematic failure cause(s)

When mitigating measures have been implemented to prevent re-occurrence of systematic DU failures, and it is sufficiently documented that the failures will not re-occur, the contribution from these DU failures can be disregarded in the quantitative analysis. An adjusted no. of DU failures may then form the input to the failure rate update and the subsequent test interval optimisation.

When disregarding DU failures from the failure data sample, the measures should be documented thoroughly, including:

- The failure cause.
- The effects of the measure (on removal of the DU failure / failure cause).
- How the measure is implemented.
- How the measure is followed-up to maintain its effect.

Appendix C describes a method to assess and document credit for mitigating measures when there is uncertainty regarding the documented effect of the measures (and/or where the measure only partially removes the failure cause).

5.6 Update failure rates

The Bayesian failure rate estimation that *combines* the (a priori) design failure rate with additional (a posteriori) *facility specific* operational experience is recommended. This approach is also less dependent on the amount of operational experience and gives limited fluctuations from observation period to observation period. An alternative approach for updating failure rates based solely on facility specific operational experience is also given in Appendix A.

⁸ The factor 4 is derived from an analysis of the data in the PDS data handbook 2021 [5]. The uncertainty interval of the multi-sample estimator has been estimated for several equipment groups and compared with their recommended (generic) values from the OREDA handbook [20]. It was found that the upper 90% uncertainty value was about 2–4 times the recommended value. A corresponding factor to identify group outliers with significantly lower average failure rates has not been possible to derive.

The Bayesian approach combines new operational experience with prior knowledge of the failure rate to obtain an updated failure rate. The prior knowledge is represented by an *input failure rate* and its *conservative estimate (CE)*. The conservative estimate expresses the uncertainty of the input failure rate.

The required input data for updating failure rate of an equipment group is:

Parameter	Denomination	Description
$\lambda_{DU,i-1}$	per hour	Input failure rate – from design ($i = 1$) or previous period ($i > 1$)
$\lambda_{DU-CE,i-1}$	per hour	Conservative estimate of input failure rate (as a means of expressing failure rate uncertainty)
n_i		No. of tags within equipment group in operation during period i
x_i	-	No. of DU failures within the equipment group during period i
t_i	hours	Length of period i
T_i	hours	Aggregated operating time for period i , $T_i = n_i \cdot t_i$

Figure 5 illustrates the prior estimates for the Bayesian update for some consecutive observation periods. The updated failure rate from the previous observation period and its conservative estimate, are the inputs for the next failure rate update. For the first observation period, the input is the failure rate from design with its conservative estimate.

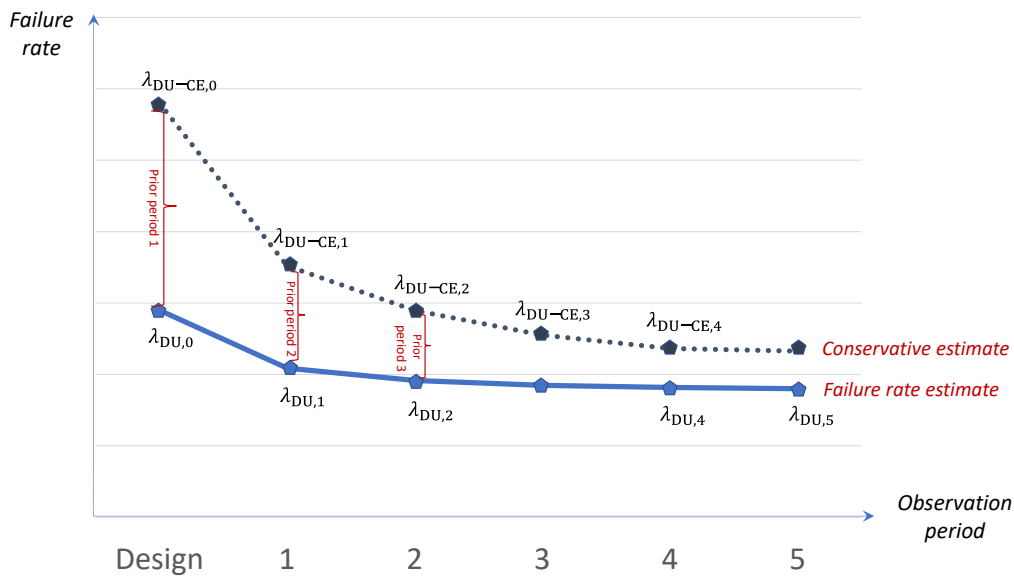


Figure 5: Observation periods from operation start of facility.

The estimate for the updated failure rate for period i is [21]:

$$\lambda_{DU,i} = \frac{\alpha_i + x_i}{\beta_i + T_i} \quad (1)$$

where the uncertainty parameters β_i and α_i are [22]:

$$\beta_i = \frac{\lambda_{DU,i-1}}{(\lambda_{DU-CE,i-1} - \lambda_{DU,i-1})^2} \quad (2)$$

and

$$\alpha_i = \beta_i \cdot \lambda_{DU,i-1} = \frac{\lambda_{DU,i-1}^2}{(\lambda_{DU-CE,i-1} - \lambda_{DU,i-1})^2} \quad (3)$$

Suggested choices for $\lambda_{DU-CE,i-1}$, for period $i = 1$ and periods $i > 1$, respectively, are:

For period $i = 1$:

$$\lambda_{DU-CE,0} = 2 \cdot \lambda_{DU,0} \quad (4)$$

This reduces equations (2) and (3) to $\beta_i = \frac{1}{\lambda_{DU,0}}$ and $\alpha_i = 1$, respectively.

For period $i > 1$:

$$\lambda_{DU-CE,i-1} = \lambda_{DU,i-1}^{90U} = \frac{Z_{0.10, 2(\alpha_{i-1} + x_{i-1})}}{2(\beta_{i-1} + T_{i-1})} \quad (5)$$

When suggesting a conservative estimate for the input failure rate for period $i > 1$, two main considerations must be taken:

- 1) $\lambda_{DU-CE,i-1}$ must be *low enough* such that any random fluctuations in failure rate between the observation periods are reduced.
- 2) $\lambda_{DU-CE,i-1}$ must be *high enough* such that the prior information (previous failure rate) does not totally outweigh the observations from the present observation period.

These two considerations are somewhat contradictory, since a large $\lambda_{DU-CE,i-1}$ (like $\lambda_{DU-CE,0}$) prevents that prior information is given a too high weight, but may cause large failure rate fluctuations between periods, whereas a small $\lambda_{DU-CE,i-1}$ ($\lambda_{DU-CE,i-1}$ close to $\lambda_{DU,i-1}$) prevents random fluctuations but places a high weight on previous data. The choice of $\lambda_{DU-CE,i-1}$ should therefore ideally be made for every period – based on the assessed relevance of the prior knowledge for the current period. The more relevant the prior information is, the closer to $\lambda_{DU,i-1}$ the $\lambda_{DU-CE,i-1}$ should be chosen. Such a choice is however not straightforward, and a pragmatic approach has therefore been applied: Different cases with varying number and length of periods, design failure rates, number of DU failures and aggregated operating time in each period have been investigated. These cases indicate that choosing a conservative estimate as the upper bound of the 90% one-sided credibility interval⁹ for the failure rate $\lambda_{DU,i-1}$ gives reasonable results in most cases. Here $z_{0.10, \nu}$ is the upper 10% percentile of the χ^2 -distribution with ν degrees of freedom, i.e., $P(\chi^2 > z_{0.10, \nu}) = 0.10$. In situations where ν is not an integer, an interpolation in the χ^2 -distribution may be performed.

Note that based on the above discussion, other quantiles than 90% can be applied in (5) depending on the confidence in the previous failure rate(s).

⁹ *Credibility* interval is the term used with Bayesian estimation of failure rates, while *confidence* interval applies to failure rates elsewhere (hence in practice credibility interval and confidence interval expresses the same).

Also note that the above discussion indicates that unless major changes in operation or other conditions that are expected to affect the failure rate significantly, take place, the most straightforward approach with respect to updating the failure rate, will be to consider all data as *one* observation period, i.e., period 1 with the conservative estimate from equation (4).

EXAMPLE

There are 35 blowdown valves on a facility that has been in operation for three years. During this period one DU failure has been revealed for the blowdown valves (solenoids excluded). We also know the assumed failure rate from design, $\lambda_{DU,0} = 2.9 \cdot 10^{-6}$.

First, we calculate the aggregated operating time for the period:

$$T_1 = 3 \cdot 8760 \cdot 35 = 9.2 \cdot 10^5 \text{ hours.}$$

The conservative estimate of the input failure rate is $\lambda_{DU-CE,0} = 2 \cdot \lambda_{DU,0} = 5.8 \cdot 10^{-6}$ per hour. Further, $\beta_1 = \frac{1}{2.9 \cdot 10^{-6}} = 3.5 \cdot 10^5$ and $\alpha_1 = 1$. An estimate for the updated failure rate for the period then becomes:

$$\lambda_{DU,1} = \frac{\alpha_1 + x_1}{\beta_1 + T_1} = \frac{1+1}{3.5 \cdot 10^5 + 9.2 \cdot 10^5} = 1.6 \cdot 10^{-6} \text{ per hour.}$$

It is now considered to initiate a new observation period, due to some operational changes that may affect the failure rate. We then check if the operational experience from period 1 is sufficient, i.e., if $\lambda_{DU,0} \cdot T_1 > 1$:

$$\lambda_{DU,0} \cdot T_1 = 2.9 \cdot 10^{-6} \cdot 9.2 \cdot 10^5 = 2.8.$$

Since this number is greater than 1, we can conclude that the operational experience in the first period of three years is sufficient to consider this period as a separate observation period (period 1).

Assume that for the new observation period (i.e., period 2), the 35 blowdown valves have been in operation for *two additional years* where 2 new DU failures have been revealed for the valves (solenoids excluded). Then the aggregated operating time for period 2 becomes: $T_2 = 2 \cdot 8760 \cdot 35 = 6.1 \cdot 10^5$ hours. The conservative estimate of the input failure rate $\lambda_{DU,1}$ for this new period based on prior knowledge from period 1 is:

$$\lambda_{DU-CE,1} = \frac{z_{0.10, 2(\alpha_1+x_1)}}{2(\beta_1+T_1)} = \frac{z_{0.10, 2(1+1)}}{2(3.5 \cdot 10^5 + 9.2 \cdot 10^5)} = \frac{z_{0.10, 4}}{2.5 \cdot 10^6} = \frac{7.8}{2.5 \cdot 10^6} = 3.1 \cdot 10^{-6} \text{ per hour.}$$

Note that the 10% percentile of the χ^2 -distribution with 4 degrees of freedom, $z_{0.10,4}$, can be found e.g., by Excel and its function CHISQ.INV.RT(0.10, 4).

To calculate an estimate for the updated failure rate, we first calculate β_2 and α_2 :

$$\beta_2 = \frac{\lambda_{DU,1}}{(\lambda_{DU-CE,1} - \lambda_{DU,1})^2} = \frac{1.6 \cdot 10^{-6}}{(3.1 \cdot 10^{-6} - 1.6 \cdot 10^{-6})^2} = 7.1 \cdot 10^5 \text{ and}$$

$$\alpha_2 = \beta_2 \cdot \lambda_{DU,1} = 7.1 \cdot 10^5 \cdot 1.6 \cdot 10^{-6} = 1.1.$$

Then, the failure rate estimate from the second observation period becomes:

$$\lambda_{DU,2} = \frac{\alpha_2 + x_2}{\beta_2 + T_2} = \frac{1.1 + 2}{7.1 \cdot 10^5 + 6.1 \cdot 10^5} = 2.3 \cdot 10^{-6} \text{ per hour.}$$

Hence, the new estimated failure rate has increased as compared to the failure rate from period 1 – but is still below the design failure rate.

Note that more operational data should preferably be gathered for observation period 2 prior to initiating a third observation period as $\lambda_{DU,1} \cdot T_2 = 1.6 \cdot 10^{-6} \cdot 6.1 \cdot 10^5 < 1$. If one additional year of operation is gathered, i.e., $T_2 = T_1 = 9.2 \cdot 10^5$, then the criteria is fulfilled.

5.7 Compare with performance requirement(s)

Operational data can be used in relevant performance indicators to compare with performance requirements such as (see section 4.4):

1. **Design failure rate:** The updated failure rate is compared with the design failure rate $\lambda_{DU,0}$:

$$\text{Is } \lambda_{DU,i} \leq \lambda_{DU,0} ?$$

2. **The expected (acceptable) number of DU failures:** The number of revealed/experienced DU failures is compared with the expected number of DU failures derived from the design failure rate:

$$\text{Is } x_i \leq \lambda_{DU,0} \cdot T_i ?$$

3. **SIL/PFD requirement for the SIF:** The PFD calculated from the updated failure rates and the current test intervals is compared with the PFD requirement for the SIF:

$$\text{Is } \text{PFD-SIF}_{\text{operational}} \leq \text{PFD requirement of SIF} ?$$

4. **PFD budget for an equipment group (component):** The PFD calculated from the updated failure rate and the current test interval is compared with the predefined PFD budget or target:

$$\text{Is } \text{PFD-Group}_{\text{operational}} \leq \text{PFD}_t ?$$

5. **Failure fraction:** The calculated failure fraction is compared with the failure fraction criterion:

$$\text{Is } \frac{x_{i-\text{test}}}{\# \text{ proof tests}} \leq \text{FF criterion} ?$$

When comparing with performance requirements, the following general guidelines apply:

- If the performance indicator is *within the requirement*, it is considered acceptable and less frequent testing may be considered. If relevant, the possibility of eliminating failure causes (the ALARP principle) should anyhow be considered.
- If the performance indicator is *outside the requirement*, mitigating measures should be considered, including the need for more frequent testing.

As Low As Reasonable Practicable (ALARP) principle: The residual risk shall be reduced as far as reasonably practicable. If the risk is said to be ALARP, it must be possible to demonstrate that the cost involved in reducing the risk further would be grossly disproportionate to the benefit gained.

5.8 Update test intervals

In this section, three methods for updating the test interval are presented. Reference is also made to section 4.2 for required input to each method and section 5.1 for underlying assumptions and prerequisites.

- M1. Method based on the assumed design failure rate and design test interval for the equipment group.
- M2. Method based on the allocated PFD budget for the equipment group.
- M3. Method based on the PFD requirement of the SIF.

Note that method M3 is not restricted to one equipment group and considers failure rates and test intervals of all components in a SIF.

Generally, the test interval may be updated if operational experience proves that the equipment reliability differs significantly from what was assumed from design or observed from previous observation period(s). It should be emphasised that changing the test interval is a decision which needs to be substantiated by extensive quantitative as well as qualitative arguments, see section 4.7.3.

The three methods are conservative in the sense that the 70% upper one-sided credibility interval value of the failure rate is applied, and that sufficient operational experience is required before the test interval can be changed¹⁰. The upper 70% value for period i is denoted $\lambda_{DU,i}^{70U}$.

The new proposed test interval should be rounded towards the closest *allowable* (predefined) *test interval* (see example below). *Allowable* test intervals here include (M=Month): 1M, 2M, 3M, 4M, 6M, 9M, 12M, 18M, 24M, 36M, 48M, etc.

Note that even if the algorithms allow for more than doubling of the test interval, it is generally not recommended to more than double the test interval from one observation period to another. Maximum recommended length of test interval also depends on the type of equipment: Mechanical components, such as shutdown valves, dampers, and pumps, must be operated regularly to prevent moving parts from becoming stuck. Components without mechanical parts, such as sensors, transmitters, and logic elements, require less frequent activation and can have a higher test interval.

¹⁰ This 70% upper value expresses the conservativeness or confidence required to allow for a change of test interval, whereas the upper 90% values as suggested in section 5.6, provides a conservative estimate of the *input failure from the previous observation period*. Note that for change of test interval, a 90% upper value *could* also have been used, however making the method more conservative.

5.8.1 M1 – Method for optimising test interval based on failure rate

Input data:

The required input data for period i for an equipment group is:

Parameter	Denomination	Description
$\lambda_{DU,i-1}$	per hour	Input failure rate (from previous period)
$\lambda_{DU-CE,i-1}$	per hour	Conservative estimate of the input failure rate
$\lambda_{DU,0}$	per hour	Design failure rate
τ_0	hours	Design test interval
x_i	-	DU failures revealed within period i
T_i	hours	Aggregated operating time within period i

Algorithm:

The updated 70% failure rate is compared with the design failure rate (and test interval). Find the highest allowable test interval corresponding to the design failure rate and design test interval:

$$\tau_i \leq \frac{\lambda_{DU,0} \cdot \tau_0}{\lambda_{DU,i}^{70U}} = \frac{2 \cdot \lambda_{DU,0} \cdot \tau_0 \cdot (\beta_i + T_i)}{z_{0.30,2(\alpha_i + x_i)}} \text{ hours} \quad (6)$$

where β_i and α_i are given by equations (2) and (3), respectively. The highest possible test interval should then be rounded *towards* the nearest *allowable test interval* (see above).

EXAMPLE

Consider the 35 blowdown valves (here assumed to have a constant failure rate during the observation period) that has been operated for three years with one DU failure experienced. The design failure rate is $\lambda_{DU,0} = 2.9 \cdot 10^{-6}$ per hour and the design test interval is $\tau_0 = 6M$.

Following the above algorithm, the highest possible (theoretical) test interval (in hours) is:

$$\frac{2 \cdot \lambda_{DU,0} \cdot \tau_0 \cdot (\beta_1 + T_1)}{Z_{0.30,2}(\alpha_1 + x_1)} = \frac{2 \cdot 2.9 \cdot 10^{-6} \cdot 4380 \cdot (3.5 \cdot 10^5 + 9.2 \cdot 10^5)}{Z_{0.30,2}(1+1)} = \frac{3.2 \cdot 10^4}{4.9} = 6.5 \cdot 10^3 \text{ hours.}$$

$6.5 \cdot 10^3$ hours correspond to 8.9 months. Taking a conservative approach, the nearest allowable (i.e. predefined) test interval is $\tau_1 = 6M$. Note that a less conservative approach (combined with a qualitative consideration) could allow rounding towards the nearest allowable test interval, in this case $\tau_1 = 9M$

Assuming a new observation period (period 2) with three years of operation and two DU failures. The highest possible test interval based on additional data from period 2 now becomes:

$$\frac{2 \cdot \lambda_{DU,0} \cdot \tau_0 \cdot (\beta_2 + T_2)}{Z_{0.30,2}(\alpha_2 + x_2)} = \frac{2 \cdot 2.9 \cdot 10^{-6} \cdot 4380 \cdot (7.1 \cdot 10^5 + 9.2 \cdot 10^5)}{Z_{0.30,2}(1.1+2)} = \frac{4.1 \cdot 10^4}{7.2} = 5.7 \cdot 10^3 \text{ hours.}$$

$5.7 \cdot 10^3$ hours correspond to 7.8 months. Again, taking a conservative approach, the nearest test interval is $\tau_2 = 6M$, whereas a less conservative approach may allow a choice of 9M since this is the nearest allowable interval.

Since the valves and the corresponding solenoids are tested together, the test interval of the solenoids should also be assessed prior to increasing the test interval of the blowdown valves. If zero DU failures have been revealed for the blowdown valve solenoids, the test interval can be increased as estimated for the blowdown valves. If DU failures have been revealed for the solenoids, there are two alternative approaches:

1. Solenoids are assessed together with the blowdown valves (merging the failure rates and DU failures). Then the blowdown valve and the corresponding solenoid(s) count as one component giving a single test interval suggestion.
2. Solenoids are assessed as a separate group (together with other solenoids). The minimum estimated test interval when comparing the solenoids and the blowdown valves, becomes the combined suggested test interval.

5.8.2 M2 – Method for optimising test interval based on PFD budget

Input data:

The required input data for period i for an equipment group is:

Parameter	Denomination	Description
PFD_t	-	PFD budget or target for the equipment group
$\lambda_{DU,i-1}$	per hour	Input failure rate (from previous period)



$\lambda_{DU-CE,i-1}$	per hour	Conservative estimate of the input failure rate
x_i	-	DU failures revealed within period i
T_i	hours	Aggregated operating time within period i

Algorithm:

The updated PFD calculated from the updated 70% failure rate is compared to the PFD budget for the equipment group. Find the highest allowable test interval that fulfils the PFD budget (PFD_t):

$$\tau_i \leq 2 \cdot \frac{PFD_t}{\lambda_{DU,i}^{70U}} = \frac{4 \cdot PFD_t \cdot (\beta_i + T_i)}{z_{0.30,2(\alpha_i + x_i)}} \text{ hours} \quad (7)$$

where β_i and α_i are given by equations (2) and (3), respectively. The highest possible test interval should then be rounded *towards* the nearest *allowable test interval* (see above).

EXAMPLE

Consider again the 35 blowdown valves that has been operated for three years with one DU failure experienced. The design test interval is $\tau_0 = 6M$ and the PFD budget for the blowdown valves is here assumed to be 0.01. The design failure rate is $\lambda_{DU,0} = 2.9 \cdot 10^{-6}$ per hour.

Following the above algorithm, the highest possible (theoretical) test interval is:

$$\frac{4 \cdot PFD_t \cdot (\beta_1 + T_1)}{z_{0.30,2(\alpha_1 + x_1)}} = \frac{4 \cdot 0.01 \cdot (3.5 \cdot 10^5 + 9.2 \cdot 10^5)}{z_{0.30,2(1+1)}} = \frac{5.1 \cdot 10^4}{z_{0.30,4}} = \frac{5.1 \cdot 10^4}{4.9} = 1.0 \cdot 10^4 \text{ hours.}$$

$1.0 \cdot 10^4$ hours correspond to 13.7 months. The nearest allowable test interval is then 12M. Hence, the new proposed test interval becomes $\tau_1 = 12M$.

Observe from the above examples, that methods M1 and M2 may give different result. Since only one DU failure has been observed during the observation period, the PFD budget of 0.01 allows for a new test interval of 12 months for the M2-method example. If the PFD budget had been based on the design failure rate and the design test interval of 6 months, i.e., $PFD_t = 2.9 \cdot 10^{-6} \cdot \frac{4380}{2} = 0.0064$, the two methods would have had the same starting point (or initial conditions) and would have given the same new test interval (6 or 9 months depending on degree of conservativeness). Consequently, it is important to understand that initial conditions and especially the choice of PFD budget will affect the suggested test interval. An increased PFD budget will inevitably allow for less frequent testing (and vice versa).

5.8.3 M3 – Method for optimising test interval based on PFD requirement of SIF**Input data:**

The required input data for period i is:



Parameter	Denomination	Description
$\text{PFD}_{\text{SIF-Req.}}$	-	PFD requirement for the complete SIF
$\lambda_{\text{DU},i-1}$	per hour	Input failure rate for each equipment group represented in the SIF
$\lambda_{\text{DU-CE},i-1}$	per hour	Conservative estimate of the input failure rate for each equipment group
x_i	-	DU failures revealed within period i for each equipment group
T_i	hours	Aggregated operating time within period i for each equipment group

Algorithm:

The updated PFD calculated from the updated 70% failure rates is compared to the PFD requirement for the SIF. Find the optimised combination of allowable (and reasonable) test intervals that fulfils the PFD requirement ($\text{PFD}_{\text{SIF-Req.}}$) for the complete SIF:

A simple expression for the optimised test interval can only be given if all components in the SIF have a *common test interval*. The algorithm is analogue to method M2 in section 5.8.2: E.g., for a SIF comprising three single components, one sensor, one logic solver and one final element with the 70% failure rates $\lambda_{\text{DU-Sensor}}^{70\text{U}}$, $\lambda_{\text{DU-LogicSolver}}^{70\text{U}}$ and $\lambda_{\text{DU-FinalElement}}^{70\text{U}}$, respectively, the highest allowable test interval for the SIF is¹¹:

$$\tau_i \leq 2 \cdot \frac{\text{PFD}_{\text{SIF-Req.}}}{\lambda_{\text{DU-Sensor},i}^{70\text{U}} + \lambda_{\text{DU-LogicSolver},i}^{70\text{U}} + \lambda_{\text{DU-FinalElement},i}^{70\text{U}}} \quad (8)$$

If some components in the SIF have *different test intervals*, there may be several combinations of updated test intervals for the components in the SIF. This requires a SIF model (e.g., in Excel or other tools/applications) that calculates the PFD based on the updated failure rates and test intervals specified for each component:

- For the failure rates, the updated 70% failure rates should be entered.
- For the test intervals, e.g., τ_{S} , τ_{LS} , τ_{FE} , only allowable test intervals should be entered.
- The selected test intervals are then the optimised combination (with respect to cost, maintenance, etc.) of test intervals such that the updated PFD with the 70% failure rates is within the SIF PFD requirement:

$$\text{PFD}(\lambda_{\text{DU-Sensor},i}^{70\text{U}}, \lambda_{\text{DU-LogicSolver},i}^{70\text{U}}, \lambda_{\text{DU-FinalElement},i}^{70\text{U}}, \tau_{\text{Sensor}}, \tau_{\text{LogicSolver}}, \tau_{\text{FinalElement}}) \leq \text{PFD}_{\text{SIF-Req.}}$$

5.8.4 Qualitative aspects to consider when changing the test interval

Changing the test interval should not rely on quantitative considerations alone. Necessary qualitative aspects are given in the checklist below and are summarised in the flow chart in Figure 6. Based on an evaluation of these qualitative aspects, a change of the test interval may be recommended – or postponed until more insight is gained or more data has been collected.

Before implementing *new test interval (reduced or extended)* – consider:

- Whether the *data is valid*, i.e., if the data applies for equipment that are in use today and forward.

¹¹ Note that this expression only applies for simple SIFs where failure rates for input, logic and output can be added up. Also, ideally it should be demonstrated that the sum of several 70% bounds gives a 70% bound (for the SIF).

- Whether assumptions in Table 6 are fulfilled, particularly if the assumption of a *constant failure rate* is still valid when the test interval is extended, or if it is likely that the component reach the wear out period before next test.
- Whether the test interval is in line with the *vendor recommendation* and the rationale behind the vendor recommendation. If necessary, the vendor may be consulted.

Before implementing *more frequent testing (reduced test interval)* – consider:

- Whether *systematic DU failures* seem predominant. In such case, rather than reducing the test interval, failure cause analysis should be performed so that corrective means can be identified and implemented for the components in question.
- Whether *mitigating measures* have already been implemented and it can be documented that similar DU failures will not re-occur, or the probability of re-occurrence is low.
- Whether the *equipment group* should be re-defined, e.g., extracting outliers.
- Whether more testing can *increase the failure rate*, due to additional wear and/or human errors.
- Whether there are *available resources* to execute more frequent testing.

Before implementing *less frequent testing (extended test interval)* – consider:

- Whether dangerous *degraded failures* have been revealed. In such cases, the test interval should not be extended, or these failures should be analysed, and mitigating measures implemented to ensure that they not develop into DU failures before next planned proof test.
- Whether a change of test interval may affect the *frequency of other associated maintenance activities* (inspection, lubrication, etc.) that can influence the component reliability.
- Whether there are any *secondary effects* from extending the test interval, e.g., build-up of materials/contamination in valve seat or cavity for normally opened valves. It is often recommended that shutdown valves are activated at least annually to avoid sticking, while detectors and transmitters with high diagnostic coverage and limited risk for drifting do not necessarily need to be tested that often.

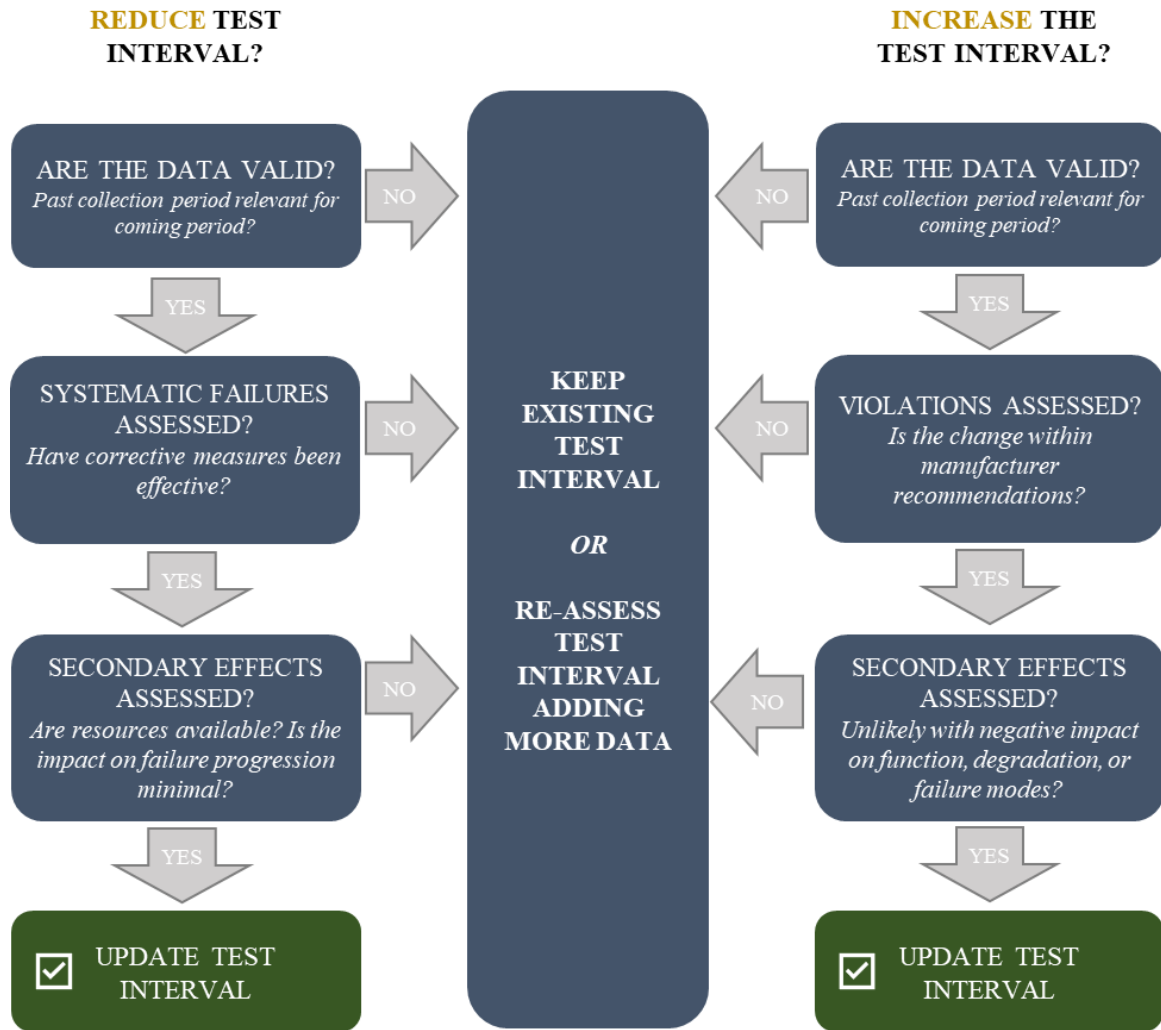


Figure 6: Flow chart for qualitative assessment prior to test interval update.

6 Proof testing alternatives

This chapter briefly discusses some special issues related to proof testing:

- Premises for crediting demands or activations as a proof test.
- The possibility of testing only a sample of components within a population.
- The impact of partial stroke testing.

6.1 Crediting demands or activations as proof tests

In addition to proof testing, there will be different operational events (shutdowns, trips, etc.) that result in an activation of the SIS. These activations can be credited as tests, given that certain premises are fulfilled.

Offshore Norge GL-070, section F.6, suggests an approach where actual demands or other activations that trigger the safety action can be credited as a proof test and lists the following prerequisites:

- The demand and/or other activation occurs within the second half¹² of the current proof test interval.
- The demand and/or other activation provides equivalent information as a proof test.
- The complete safety function of the equipment is verified in the demand/activation. If not, the residual functionality must be verified upon the next scheduled proof test. E.g., an unplanned shutdown of a valve with leakage requirement will only verify closure of the valve while the leakage rate requirements must be verified during the next scheduled leakage test.

Figure 7 illustrates the PFD (both average and time-dependent) for two test intervals (assuming 100% proof test coverage). A demand/activation is here experienced in the second half of the first test interval such that the planned test at time τ can be postponed until time 2τ . The blue dotted line shows the time-dependent PFD when the demand/activation is not credited, whereas the blue dash-dotted line shows the time-dependent PFD when the demand/activation is credited. The red lines illustrate the average PFDs for the two cases. Note that the average PFD for the postponed proof testing (red solid line) will vary depending on when the demand occurs.

¹² See reasoning behind second half of test interval in Offshore Norge 070 guideline:

<https://offshorenorge.no/retningslinjer/arkiv/helse-arbeidsmiljo-og-sikkerhet/teknisk-sikkerhet/070-guidelines-for-the-application-of-iec-61508-and-iec-61511-in-the-petroleum-activities-on-the-continental-shelf/>

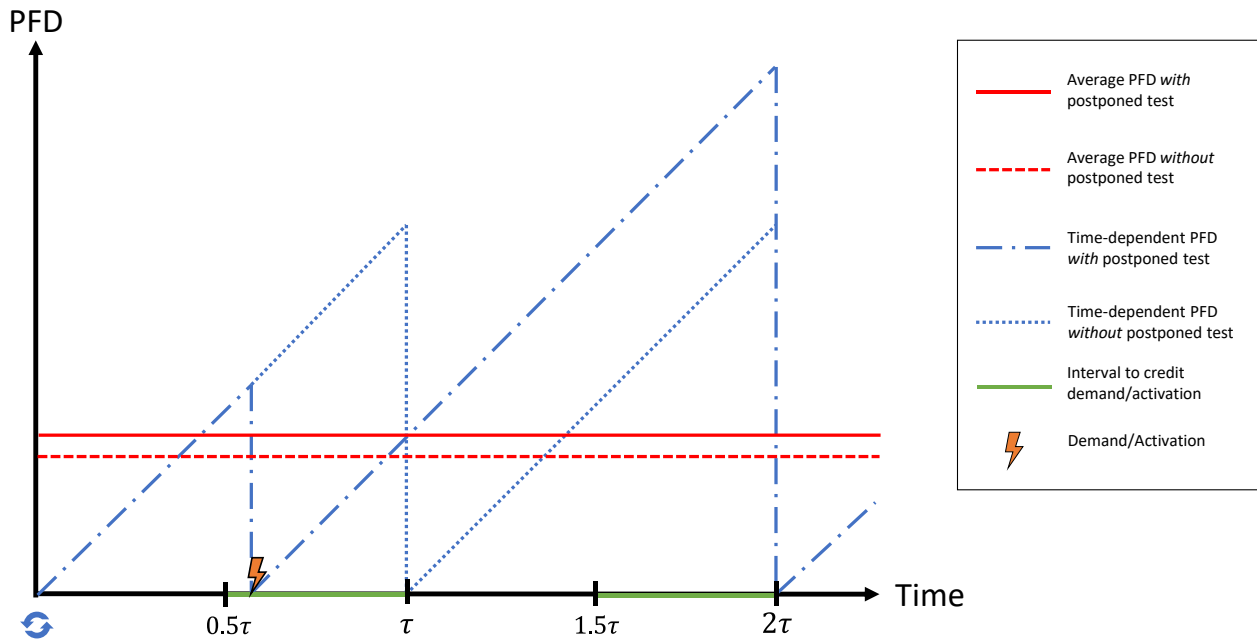


Figure 7: Time-dependent (and average) PFD with and without postponed proof testing.

For the case where the demand/activation is credited and the testing is postponed, the average PFD increases 25% as a worst case or 8% on average, compared to the situations where the proof test was not postponed. Note that although there is a limited increase in the average PFD when the test is postponed, the time dependent PFD will be temporary higher for a time before the postponed proof test (or the next activation) occurs. Whether this is acceptable or not (and the SIL requirement is still fulfilled) should be assessed for each specific case. Thus, utilising shutdowns or activations as means of proof testing should be done with some care and it must be ensured that the above listed prerequisites are always fulfilled.

6.2 Impact of partial stroke testing (PST)

Partial testing can be utilized to detect DU failures, for instance partial stroke testing (PST) of valves. PST is not a complete proof test, but a partial test that supplements full proof testing. The average PFD is then improved because some of the DU failure modes are tested and revealed more frequently than during the complete proof test (e.g., the solenoid valve functionality is tested every third month as compared to the annual complete valve closure test).

A simplified PFD formula including the effect of PST, is shown below. Let $PST_{coverage}$ be the assessed coverage of the PST, τ_{PST} the PST interval and τ the complete proof test interval. Then the average PFD for a single component becomes:

$$PFD = PST_{coverage} \cdot \left(\lambda_{DU} \cdot \frac{\tau_{PST}}{2} \right) + (1 - PST_{coverage}) \cdot \left(\lambda_{DU} \cdot \frac{\tau}{2} \right)$$

PST will reduce the need for (complete) proof testing but are not a substitute. How much the proof test can be extended if additional PST is introduced, is dependent of the PST coverage, the current proof test interval, and the PST interval. Based on a quantitative criterion saying that the overall PFD shall not be reduced when introducing PST, the maximum (new) proof test interval, τ_{new} , for a single component becomes:

$$\tau_{\text{new}} \leq \frac{\tau - \tau_{\text{PST}} \cdot \text{PST}_{\text{coverage}}}{1 - \text{PST}_{\text{coverage}}}$$

where τ is the current (complete) proof test interval.

When utilizing partial testing, it must be ensured that the reliability is still according to the requirements, i.e., that the proof test interval is within the requirement and that the DU failure rate is adjusted for possible introduction of new DU failures due to partial testing.

EXAMPLE

A single component has DU failure rate $\lambda_{\text{DU}} = 1.0 \cdot 10^{-6}$ per hour, and 12M proof test interval. The average PFD of this component without additional partial stroke test is then $\approx 4.4 \cdot 10^{-3}$.

Assume that additional partial stroke test is implemented with 60% coverage and 4M interval. Then the proof test interval can, based on quantitative optimisation, only be extended to 24M as:

$$\frac{12 - 4 \cdot 0.60}{0.4} = 24.$$

This results in a new average PFD (with $\tau_{\text{PST}} = 2920$ hours and $\tau = 17520$ hours) of:

$$\begin{aligned} \text{PFD} &= \text{PST}_{\text{coverage}} \cdot \left(\lambda_{\text{DU}} \cdot \frac{\tau_{\text{PST}}}{2} \right) + (1 - \text{PST}_{\text{coverage}}) \cdot \left(\lambda_{\text{DU}} \cdot \frac{\tau}{2} \right) \\ &= 0.6 \cdot \left(1.0 \cdot 10^{-6} \cdot \frac{2920}{2} \right) + 0.4 \cdot \left(1.0 \cdot 10^{-6} \cdot \frac{17520}{2} \right) \approx 4.4 \cdot 10^{-4} \end{aligned}$$

As seen, the new PFD when implementing partial stroke testing is not exceeding the original PFD with 12M proof test and no partial testing.

7 References

1. Hauge S, Håbrekke S, Lundteigen MA, Lee S, Ottermo MV. Guidelines for standardised failure reporting and classification of safety equipment failures in the petroleum industry, Ed. 1 (open version) (APOS H1). 2023. Report No.: SINTEF Report 2023:00108.
2. Lee S, Ottermo MV, Hauge S, Håbrekke S, Lundteigen MA. Potential for automated follow-up of safety equipment. 2023. Report No.: SINTEF Report 2023:00110.
3. Hauge S, Kvam, E, APOS working group. Specification for standardised electronic SRS. 2023.
4. Hauge S, Lundteigen MA, Ottermo MV, Lee S, Petersen S. Information model for functional safety (APOS H5). 2023. Report No.: SINTEF Report 2023:00109.
5. Ottermo MV, Hauge S, Håbrekke S. Reliability data for safety equipment. PDS data handbook. 2021.
6. NOROG070. Application of IEC 61508 and IEC 61511 in the Norwegian Petroleum Industry (Recommended SIL requirements) [Internet]. 2020. Available from: <https://www.norskoljeoggass.no/contentassets/adc7e1512f90400cb7fe9f314600bed6/norwegian-oil-and-gas-guidelines-070-rev-3-june-2018.pdf>
7. IEC 61508. IEC 61508 - Functional Safety of Electrical/Electronic/Programmable Electronic Safety-related Systems. 2010.
8. IEC. IEC 61511 Functional safety - Safety instrumented systems for the process industry sector. Geneva, Switzerland: International Electrotechnical Commission; 2016.
9. SINTEF. Guideline for follow-up of Safety Instrumented Systems (SIS) in the operating phase. 2008. Report No.: SINTEF report A8788.
10. Håbrekke S, Hauge S, Lundteigen MA. Guideline for follow-up of Safety Instrumented Systems (SIS) in the operating phase. 2021.
11. PSA. Petroleum Safety Authority Norway, The Activity Regulations [Internet]. Norway; 2021. Available from: <https://www.ptil.no/en/regulations/all-acts/?forskrift=613>
12. PSA. Petroleum Safety Authority Norway, The Management Regulations [Internet]. Norway; 2021. Available from: <https://www.ptil.no/en/regulations/all-acts/?forskrift=611>
13. PSA. PSA Norway - Guidelines regarding the activities regulations. 2020.
14. ISO. ISO-10418 - Petroleum and natural gas industries - Offshore production installations - Process safety systems. 2019.
15. ISO. ISO-20815:2018, Petroleum, petrochemical and natural gas industries — Production assurance and reliability management [Internet]. Available from: <https://www.iso.org/standard/69983.html>
16. ISO/TR 12489:2013, Petroleum, petrochemical and natural gas industries — Reliability modelling and calculation of safety systems [Internet]. Available from: <https://www.iso.org/standard/51456.html>

17. NS-EN ISO 14224, Petroleumsindustri, petrokjemisk industri og naturgassindustri - Innsamling og utveksling av pålitelighets- og vedlikeholdsdata for utstyr [Internet]. 2016. Available from: <https://www.standard.no/>
18. PSA. Petroleum Safety Authority Norway, Trends in risk level (RNNP) [Internet]. [cited 2021 Jan 1]. Available from: <https://www.ptil.no/en/technical-competence/rnnp/>
19. SINTEF. Common Cause Failures in Safety Instrumented Systems,. 2015. Report No.: Report no. A26922.
20. OREDA. Offshore reliability data handbook. NO 1322 Høvik, Norway: OREDA Participants, Available from: Det Norske Veritas; 2015.
21. Rausand M, Hoyland A. System Reliability Theory: Models, Statistical Methods, and Applications [Internet]. Wiley; 2003. Available from: <https://books.google.no/books?id=gkUWz9AA-QEC>
22. Vatn J. Procedures for updating test intervals based on experience data. In 2006.

A Estimating DU failure rate using operational experience only

This appendix presents the approach for estimating DU failure rates using operational experience only.

To provide necessary confidence in the DU failure estimate, this approach should *only* be applied when there is *sufficient operational data*¹³. If T is the aggregated operating time and x the number of DU failures, the following should be fulfilled:

$$T \cdot x \geq 3 \cdot 10^6 \text{ hours.}$$

In case of *insufficient operational data*, i.e., $T \cdot x < 3 \cdot 10^6$, it is necessary either:

- to combine the operational data with prior knowledge of the DU failure rate (Bayesian approach), see section 5.6, or
- to collect more operational data.

The required input data for updating the DU failure rate for an equipment group based on operational experience only, λ_{DU-op} , is:

Parameter	Denomination	Description
n	-	No. of tags within equipment group that have been in operation
x	-	No. of DU failures within the equipment group
t	hours	Operating time

The DU failure rate estimate based solely on operational experience, is given by (e.g. ISO 14224, sect. C.3.2) [17]:

$$\lambda_{DU-op} = \frac{\text{Number of DU failures}}{\text{Aggregated operating time}} = \frac{x}{n \cdot t} = \frac{x}{T} \quad (9)$$

¹³ A possible “cut off point” for using only operational data, may be when the confidence in the λ_{DU-op} based solely on operational experience equals the confidence in the design DU failure rate $\lambda_{DU,0}$, i.e., when the statistical confidence in λ_{DU-op} is comparable to the confidence in the input DU failure rate $\lambda_{DU,0}$ then it is justifiable to apply only operational experience. So, when will this occur? Representative OREDA failure rates for SIS equipment (detectors, sensors, and valves) shows that typically, the upper 95% percentile in the uncertainty interval for the critical failure modes are 2–3 times the mean value of the DU failure rate. We therefore state that for cases where the upper 95% percentile of λ_{DU-op} based on operational experience is approximate 3 times the mean value λ_{DU-op} , we may use the λ_{DU-op} value solely. This condition is normally fulfilled if $T \cdot x > 3 \cdot 10^6$ hours. Then it will be possible to derive at an updated DU failure rate λ_{DU-op} with sufficiently confidence. Note that in case of only one experienced DU failure, the above condition will strictly speaking not be fulfilled (the ratio will be closer to 5). However, as the operating time must exceed as much as $3 \cdot 10^6$ hours, this is considered sufficient to rely solely on operational experience.

EXAMPLE

Assume there are 35 blowdown valves on a facility. The facility has been in operation for three years. During this period one DU failure is revealed.

The aggregated operating time multiplied with the number of DU failures for this period is: $3 \cdot 8760 \cdot 35 \cdot 1 = 9.2 \cdot 10^5$ hours, which is less than $3 \cdot 10^6$. Hence, it is recommended either to apply the Bayesian failure rate estimation or to gather more operational experience.

After two more years of operation, one new DU failure is revealed. The aggregated operating time multiplied with the number of DU failures for the now five years long period becomes: $5 \cdot 8760 \cdot 35 \cdot 2 \approx 3 \cdot 10^6$ hours. Hence, there is sufficiently operational data to estimate the DU failure rate based solely on operational experience:

The estimate for the updated DU failure rate becomes:

$$\lambda_{\text{DU-op}} = \frac{2}{5 \cdot 8760 \cdot 35} = \frac{2}{1.5 \cdot 10^6} = 1.3 \cdot 10^{-6} \text{ per hour.}$$

B Multi-sample estimators – Failure rates based on data from two or more facilities

For selected equipment groups, the operator may want to combine data across several facilities, to establish an "average" failure rate estimate and associated uncertainty bounds.

The following assumptions and notation apply for multi-sample estimation:

- Data from $k \geq 2$ different facilities shall be merged for an equipment group.
- At least one failure is revealed.
- From facility no. j , there have been revealed n_j failures during a period with aggregated operating time T_j . $j = 1, \dots, k$.

The multi-sample is either assumed to be *homogeneous*, i.e., the failure rate is the same across all facilities, or *inhomogeneous*, i.e., each facility has its own failure rate due to different operational and environmental conditions, maintenance, types of equipment, etc.

B.1 Homogeneous samples

Assuming k homogeneous samples from the k facilities have a common failure rate $\lambda = \lambda_{DU-op}$, the failure rate can be estimated by:

$$\lambda_{DU-op} = \frac{\sum_{j=1}^k n_j}{\sum_{j=1}^k T_j}$$

The corresponding 90% confidence interval is:

$$\left(\frac{1}{2 \sum_{j=1}^k T_j} z_{0.95,2 \sum_{j=1}^k n_j}, \quad \frac{1}{2 \sum_{j=1}^k T_j} z_{0.05,2(\sum_{j=1}^k n_j+1)} \right)$$

Here $z_{0.95,v}$ and $z_{0.05,v}$ denote the upper 95% and 5% percentiles, respectively, of the χ^2 -distribution with v degrees of freedom, i.e., $P(\chi^2 > z_{0.95,v}) = 0.95$ and $P(\chi^2 > z_{0.05,v}) = 0.05$.

The samples should be carefully checked if they are homogeneous before the samples are merged and the above formulas are applied. E.g., the equipment on every facility should be exposed to approximately the same environmental and operational impact and maintenance activities.

When merging several homogenous populations, the aggregated operational time may become long, and the associated confidence interval will, depending also on number of failures, become correspondingly short, see Figure 8. If homogeneity is not the case, this short confidence interval will underestimate the uncertainty in the failure rate estimate. For such cases, the multi-sample estimator for inhomogeneous or heterogeneous samples should be considered.

B.2 Inhomogeneous or heterogeneous samples

A heterogeneous sample is the combination of several more or less homogeneous samples. The multi-sample estimator applicable for inhomogeneous samples, assumes that facility no. j has its own constant failure rate $\lambda_{DU,j}$. The failure rate is assumed to be a random variable, that can take different values for the different samples.

The OREDA multi-sample estimator, θ , is calculated as follows (OREDA, 2015 [20]):

1. $\hat{\theta} = \frac{\text{Total number of failures}}{\text{Aggregated op. time}} = \frac{\sum_{j=1}^k n_j}{\sum_{j=1}^k T_j}$
2. $S_1 = \sum_{j=1}^k T_j$ and $S_2 = \sum_{j=1}^k T_j^2$
3. $V = \frac{(n_j - \hat{\theta})^2}{\sum_{j=1}^k T_j} \approx \sum_{j=1}^k \frac{n_j^2}{T_j} - \hat{\theta}^2 S_1$
4. **If** $\frac{V - (k-1)\hat{\theta}}{S_1^2 - S_2} \cdot S_1 > 0$ **Then** $\sigma^2 = \frac{V - (k-1)\hat{\theta}}{S_1^2 - S_2} \cdot S_1$. **Else** $\sigma^2 = \sum_{j=1}^k \frac{\left(\frac{n_j}{T_j} - \hat{\theta}\right)^2}{k-1}$
5. $\theta = \frac{1}{\sum_{j=1}^k \frac{1}{\hat{\theta} + \sigma^2}} \cdot \sum_{j=1}^k \left(\frac{1}{\frac{\hat{\theta}}{T_j} + \sigma^2} \cdot \frac{n_j}{T_j} \right)$

Note: The above multi-sample estimator is also applicable when merging equipment from *the same* facility that may have different failure rates or equipment attributes, e.g., various types of detectors.

The 90% *uncertainty interval* is as follows:

$$\left(\frac{\sigma^2}{2\theta} Z_{0.95,v}, \frac{\sigma^2}{2\theta} Z_{0.05,v} \right)$$

where $v = \frac{2\theta^2}{\sigma^2}$.

Note: The above uncertainty interval should not be misinterpreted as a confidence interval. A confidence interval decreases with more data, which is not the case for uncertainty intervals like this.

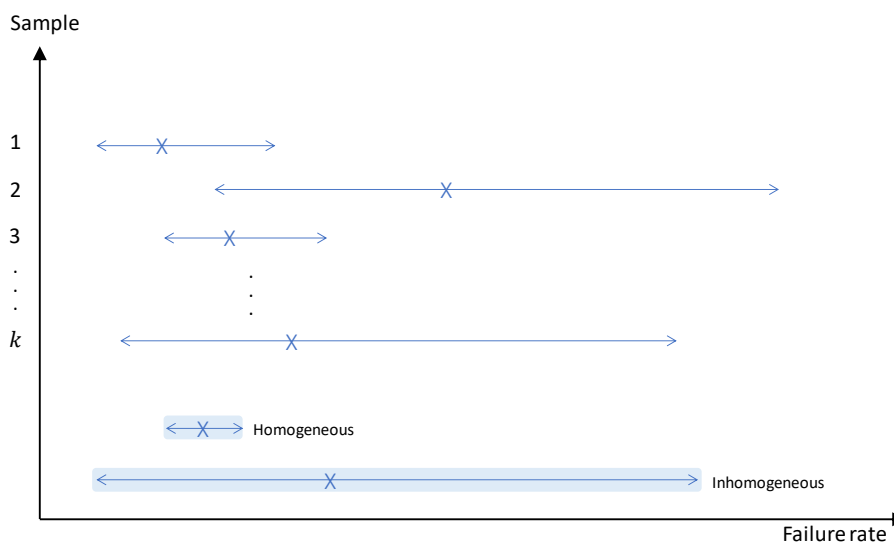


Figure 8: Uncertainty in failure rate estimates for homogeneous vs. inhomogeneous multi-samples.

C Crediting mitigating measures by adjusting the number of systematic DU failures

Systematic DU failures may re-occur if no measures are implemented to reduce or eliminate the failure cause. When suitable measures are implemented, systematic DU failures may be (partly) removed from the failure data for DU failure rate estimation and test interval optimisation.

Assume that X DU failures, $x^{(1)}, x^{(2)}, \dots, x^{(X)}$, have been revealed for an equipment group during period i , i.e., $x_i = X$. If mitigating measures have been implemented, we adjust a "future equivalent" to the number of DU failures by (based on Vatn (2006)):

$$x = \gamma^{(1)}x^{(1)} + \gamma^{(2)}x^{(2)} + \dots + \gamma^{(N)}x^{(N)}$$

where $\gamma^{(i)}$; $i = 1, \dots, X$ are correction factors due to the anticipated effect of implemented measures on each DU failure. The correction factor may be the same for some DU failures, e.g., for DU failures with a common cause, or distinct for all or some DU failures.

An analysis should be performed to identify systematic DU failures and their causes. This analysis will provide necessary input to identify and implement the correct mitigating measures.

Table 8 suggests values for the correction factor for a specific type of DU failure and its corresponding failure cause. An example is also given for fail to close of shutdown valves due to freezing of actuators.

Table 8: Suggested correction factor values and corresponding explanations and examples (based on Vatn (2006) [22]).

γ	Effect	Explanation	<i>Example: Freezing of actuators</i>
1.0	No/Low effect	No measures implemented <u>or</u> the implemented measures will have limited effect on the failure cause of this DU failure.	<i>Measures have been implemented to reduce the failure cause of other types of failures, but none of these has any anticipated effect on this specific DU failure.</i>
0.5	Medium effect	Measures implemented are expected to significantly reduce the failure cause of this DU failure. The effect of the measure (considering experience from other facilities, expert judgements, etc.) should be documented.	<i>Weekly inspection of the actuators exposed to freezing during the winter months is implemented in the inspection program.</i>
0.0	High effect	Measures implemented are expected to eliminate the failure cause. The specific measures (new design, maintenance program/procedures, etc.) and the effect (no further such DU failures will be experienced) should be documented, e.g., by FMECA review.	<i>Heat tracing has been installed and it can be documented that winter period will cause no or limited freezing problems.</i>

Table 9 lists some examples of how to treat typical DU failures for gas detectors, transmitters, and shutdown valves. From the table we see that to exclude a DU failure from the failure data / test interval optimisation



(high effect), the failure cause needs to be known. Also note that the measures should be properly documented (see Table 8) to credit for medium or high effect.

Table 9: Examples of DU failures and corresponding classification into failure type, how to include the DU failure in the test interval optimisation and follow-up of the DU failure.

Component	Failure and failure cause description	Mitigating measure	Effect	γ	Comment
Gas detectors	<i>Detector does not respond on test gas due to unknown detector failure. The gas detector was replaced.</i>	Investigate if other detectors have experienced the same problem.	No/Low	1.0	Assumed as a random hardware failure.
	<i>Detector is found to be covered/wrapped (randomly or on preventive maintenance) and has been like this since commissioning. Plastic was removed.</i>	Investigate if other detectors (not tested yet) are also covered and remove plastic on covered detectors.	High	0.0	This is a systematic failure from commissioning that will not occur in the future.
	<i>Detector does not provide alarm in CCR when exceeding the HH level due to a software error / SAS error introduced upon installation of new detector.</i>	No actions besides correcting this software error are performed.	No/Low	1.0	-
		Investigate if other detectors installed at the same time have the same software error.	Medium	0.5	
		Ensure failures do not re-occur, e.g., updating procedures for installation of detectors.	High	0.0	
	Level transmitters	<i>Transmitter with LL alarm displays level measurements significantly higher than the actual level in vessel. Another measuring principle would probably have provided more reliable measurements.</i>	No actions.	No/Low	1.0
Change measuring principle.			Medium	0.5	
<i>Transmitter with LL alarm displays level measurement significantly higher than the actual level in vessel due to incorrect calibration after replacement of transmitter.</i>		No actions besides correcting the failure are performed.	No/Low	1.0	-
		Ensure failures do not re-occur, e.g., updating calibration procedures.	High	0.0	



Component	Failure and failure cause description	Mitigating measure	Effect	γ	Comment
Pressure transmitters	<i>Isolation valve on transmitter tubing is in such a position that the transmitter is not functioning (revealed either randomly or upon preventive maintenance). The failure was introduced on the last proof test.</i>	No actions besides correcting the failure are performed.	No/Low	1.0	-
		Ensure failures do not re-occur, e.g., updating test/maintenance procedures.	High	0.0	
Shutdown valves	<i>Valve does not close completely upon shutdown signal due to dirt or corrosion (from process or environment).</i>	Reduce the impact from process/environment.	Medium	0.5	The failure is a systematic failure unless it can be documented that the process/environmental impact is inside the valve's design envelope.
		Reduce the cause of dirt/corrosion.	High	0.0	
	<i>Valve does not close within the response time requirement due to unknow cause.</i>	No actions due to unknown cause.	No/Low	1.0	Consider performing root cause analysis, particularly, if the problem is repeating.
	<i>Valve does not close within the response time requirement due to weak actuator / design issues.</i>	No actions besides correcting the failure.	No/Low	1.0	-
		Replacing actuators with new actuator of same type.	Medium	0.5	
		Ensure failures do not re-occur, e.g., changing design of actuators.	High	0.0	
	<i>Valve closes one second above the response time requirement. The closing time was adjusted on the previous test to be exactly on the response time requirement.</i>	No actions besides adjusting response time.	No/Low	1.0	-
		Update procedures for adjusting response times.	Medium	0.5	