An Examination of the Effectiveness of Proof Testing Regimes for Emergency Shutdown Valves

Steven Ashwell Kriescher

Submitted to Swansea University in fulfilment of the requirements for the degree of

Doctor of Philosophy

2025

Abstract

The emergency shutdown valve is a critical element of a typical safety instrumented function. Consequently, both effective proof testing and an appropriate test interval are imperative to ensure the desired functionality, and hence a reliably safe operation of a high-hazard process facility.

This research investigates the application of a Fuzzy Inference System (FIS) to model data uncertainty in Failure Modes, Effects, and Diagnostic Analysis (FMEDA) for safety-critical systems. The research aims to address the limitations of traditional FMEDA methods in managing subjective expert data. A novel FIS-based framework was developed that employs trapezoidal membership functions to encode expert knowledge and analyse proof test coverage. Validation was performed using both real and synthetic datasets, achieving strong regression and low error metrics, with an overall R^2 value of 0.96 and RMSE of 7, thus confirming the accuracy of the model.

The variation in proof test coverage predictions between the FMEDA and FIS methods was observed to be up to 3% for a full stroke test and up to 22% for a partial stroke test. A comparison of the FMEDA and FIS proof test coverage results suggests that the FIS approach supports further optimisation of emergency shutdown valve maintenance by integrating full and partial proof tests, extending proof test intervals, and enhancing plant uptime without compromising safety. The FIS results indicate a 100% improvement in both test intervals, which could extend the typical partial stroke test interval from six months to one year and the usual full stroke test interval from one year to two years.

This work demonstrates that fuzzy inference systems can be employed to overcome issues relating to FMEDA data subjectivity allowing emergency shutdown valve proof test coverage estimates to be determined for both full and partial stroke tests. This research contributes to the integration of expert judgement in safety-related assessments, laying the foundation for future studies to validate and refine the model further.

Declarations

This work has not previously been accepted in substance for any degree and is not being concurrently submitted in candidature for any degree.

Signed...[redacted].................. Date..........................24 April 2025

This thesis is the result of my own investigations, except where otherwise stated. Other sources are acknowledged by footnotes giving explicit references. A bibliography is appended.

Signed...[redacted].................. Date..........................24 April 2025

I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loan, and for the title and summary to be made available to outside organisations.

Signed....[redacted].................... Date..........................24 April 2025

The University's ethical procedures have been followed and, where appropriate, that ethical approval has been granted.

Signed..[redacted]...................... Date..........................24 April 2025

Acknowledgements

I would like to thank my academic supervisors, Dr. Roderick Thomas and Dr. Christopher Phillips, for their invaluable guidance and constructive feedback throughout my research journey. Their expertise and mentorship have been instrumental in shaping the direction and quality of this work.

I wish to extend my heartfelt gratitude to my industrial supervisor, Dr. David J. Smith, for his constant support and encouragement, and for providing industry-specific insights and essential guidance during the development and application of this research.

I would like to thank Dr. Neil Mac Parthaláin for his invaluable advice and guidance on fuzzy inference systems.

I also wish to thank Dr. William Goble for generously providing valuable resources and training in support of this research.

I would like to express my gratitude to several industry experts and organisations that provided invaluable support for this research. Although they prefer to remain unnamed, their contributions and provision of resources have been invaluable to the successful completion of this thesis.

Finally, I would like to extend my heartfelt gratitude to my family for their unwavering support and encouragement throughout this journey. To my dear mother and children, thank you for your patience, understanding, and constant belief in me. A special thanks goes to my partner, Ela, who has witnessed the challenges of this work first-hand and provided devoted support during the most difficult moments.

Contents

List of Figures

List of Tables

Abbreviations

AI	Artificial Intelligence
ALARP	As Low As Reasonably Practicable
BPCS	Basic Process Control System
CBM	Condition Based Maintenance
CCPS	Centre for Chemical Process Safety
CDF	Cumulative Density Function
CMMS	Computerised Maintenance Management System
COG	Centre Of Gravity Method
COIN	Corporate Operational Information
COMAH	Control of Major Accident Hazards
CRD	Component Reliability Database
DOP	Delayed Operation
DTT	De-energise To Trip
ESD	Emergency Shut Down
ESDV	Emergency Shutdown Valve
FIS	Fuzzy Inference System
FIT	Failure in Time
FMEA	Failure Modes and Effects Analysis
FMEDA	Failure Modes, Effects and Diagnostics Analysis
FST	Full Stroke Test
FTA	Fault Tree Analysis
FTC	Fail To Close
HAZOP	Hazard and Operability
HEP	Human Error Probability
HEART	Human Error Assessment and Reduction Technique
HFT	Hardware Fault Tolerance
HRA	Human Reliability Assessment
HSE	Health and Safety Executive
IAM	Industrial Asset Management
LCP	Leak in the Closed Position
LNG	Liquefied Natural Gas
LPG	Liquefied Petroleum Gas
LOPA	Layer of Protection Analysis
MAST	Maximum Allowable Stem Torque
MSE	Mean Square Error
OREDA	Offshore Reliability Equipment Data
OSHA	Occupational Safety and Health Administration
PDF	Probability Density Function
$\mathrm{PFD}_{\mathrm{avg}}$	Probability of Failure on Demand Average
PLC	Programmable Logic Controller
PPM	Planned Preventive Maintenance

PST	Partial Stroke Test
PTC	Proof Test Coverage
PTI	Proof Test Interval
RMSE	Root Mean Square Error
RESDV	Riser Emergency Shutdown Valve
RRF	Risk Reduction Factor
R^2	R-squared (coefficient of determination)
SIF	Safety Integrity Function
SIL	Safety Integrity Level
SIS	Safety Instrumented System
SOV	Solenoid Operated Valve
SPAR-H	Standardised Plant Analysis Risk-Human Reliability Analysis
SRS	Safety Requirements Specification
TSK	Takagi-Sugeno-Kang
TSO	Tight Shut Off

Chapter 1

Introduction

Safety instrumented systems are essential for preventing adverse incidents, ensuring the safe operation of industrial processes, and safeguarding people, equipment, and the environment. This thesis examines the maintenance procedures associated with these systems, with a specific focus on the concept of proof testing, which informs maintenance protocols. A significant area of discussion concerns the selection of parameter values, such as proof test coverage, which influence reliability calculations and, in turn, inform the determination of appropriate proof test intervals. The use of inappropriate or inaccurate proof test coverage values can result in either over-testing or under-testing. Published estimates of proof test coverage are frequently based on expert judgement and can be somewhat subjective, thereby introducing uncertainty into the analysis.

1.1 Hazards and Incidents in the Process Industry

Process sector industries such as oil and gas, and chemical handle large inventories of hazardous substances, in some cases at elevated temperatures and pressures, and therefore have an inherent potential for major accidents involving hazards such as fire, explosion, and toxic release [1]. Process safety is aimed at preventing major accidents involving hazardous materials and loss of containment, while occupational safety protects workers from daily work hazards such as slips, trips, and falls. In this context, a commonly accepted definition of process safety is from the Centre for Chemical Process Safety (CCPS):

"A disciplined framework for managing the integrity of hazardous operating systems and processes by applying good design principles engineering and operating practices. It deals with the prevention and control of incidents that have the potential to release hazardous materials or energy. Such incidents can cause toxic effects, fire or explosion and could ultimately result in serious injuries, property damage, lost production and environmental impact" [2, p. 1].

The number of major industrial accidents in the process industry sector (hydrocarbon and chemical) over the last four decades is significant. The main contributing factor in a number of the cases has been maintenance (or, rather, the lack of or inadequacy of) [3]. These accidents include the Bhopal Gas Tragedy (1984), the Phillips 66 Disaster (1989), the Texaco Refinery Explosion (1994), the Texas City Refinery Explosion (2005), the Deepwater Horizon Explosion (2010) and more recently the Philadelphia Refinery Explosion (2019).

Table 1.1 provides examples of major industrial accidents and associated losses. Some accidents can also have a significant environmental impact, such as the Deepwater Horizon accident. This was the largest oil spill in the history of marine oil drilling operations, with a discharge of approximately 4.9 million barrels of oil into the sea over an 87-day period, before it was finally capped [4].

Table 1.1: Examples of Major Accident Losses [4–8]

Accident	**Safety impact**	**Cost impact**
Texaco Refinery Explosion, Pembrokeshire (1994)	26 injuries	£120 million[1]
BP Texas City Refinery, Explosion, Texas (2005)	15 fatalities, 180 injuries	US$3 billion
Deepwater Horizon Explosion, Gulf of Mexico (2010)	11 fatalities, 17 injuries	US$65 billion

[1] adjusted for inflation.

History has shown that major industrial accidents can have severe consequences for human life, the environment, and the financial stability of the operating company. This highlights the importance of ensuring that potentially hazardous events are prevented and mitigated appropriately through well-engineered processes and safety systems, and effective asset management.

1.2 Achieving Process Safety: The importance of Functional Safety

Process safety is a globally recognised discipline, with regulatory frameworks established across many industrialised nations. In the UK, the Control of Major Accident Hazards (COMAH) Regulations [9] and the Offshore Safety Case Regulations (SCR) [10], derived from the European Seveso III Directive [11], require operators to demonstrate that risks are reduced to As Low As Reasonably Practicable (ALARP). In the United States, a comparable framework exists under the Occupational Safety and Health Administration's (OSHA) Process Safety Management (PSM) standards [12, 13], which mandate the identification, evaluation, and control of hazards associated with highly hazardous chemicals. ALARP is achieved when the time, practicality, and cost of further reduction measures become unreasonably disproportionate to the additional risk reduction obtained [14, 15].

Figure 1.1 shows the tolerability of risk model. Application of the ALARP principle involves specifying two sets of risk tolerance criteria. The first set of criteria (the dividing line between unacceptable and tolerable regions, usually called the *maximum tolerable risk* is a minimum requirement that must be met. This refers to a level of risk that is clearly unacceptable to most people, necessitating action to reduce it. The second set of criteria, the dividing line between the tolerable and broadly acceptable regions, usually called *broadly acceptable risk* is a goal that may not be reached, but toward which progress must be made until risk reduction measures involve grossly disproportionate sacrifices. This refers to a level of risk that is minimal and not concerning, effectively a virtually safe level widely considered acceptable, requiring no further action for reduction. A range of risks is tolerable between the two sets of criteria. The residual risk, or the risk remaining after implementing controls, should ideally fall within the broadly acceptable region or at the lower end of the tolerable region [15].

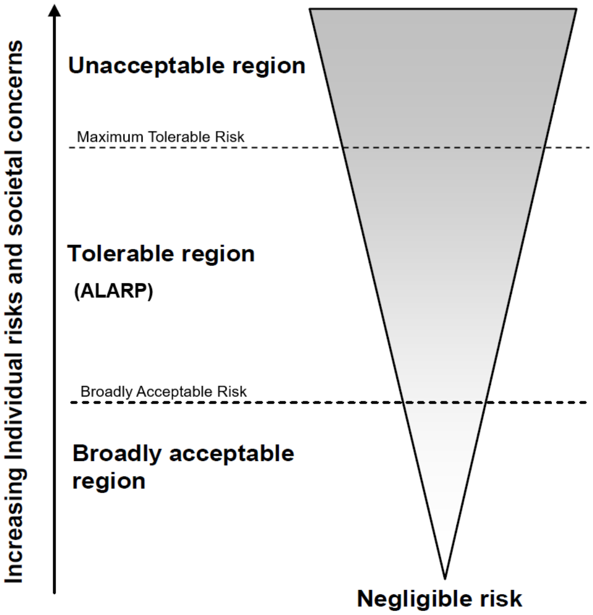


Figure 1.1: Tolerability of Risk - ALARP Principle [14–17]

Figure 1.2 illustrates a layer of protection methodology as a means to achieving ALARP high-hazard facilities. Each layer of protection is designed to be independent from each other, so that failure of one layer does not impact the operation or performance of another layer. Layers can be preventative, such as safety instrumented systems, or mitigating, such as fire and gas systems. Preventative layers are intended to reduce the likelihood of hazardous events occurring, while mitigating layers aim to minimise the consequences should such events occur. Safety instrumented systems provide protection from process excursions caused by initiating events such as failures in the Basic Process Control System (BPCS) or human error, hence preventing escalation of potentially hazardous events.

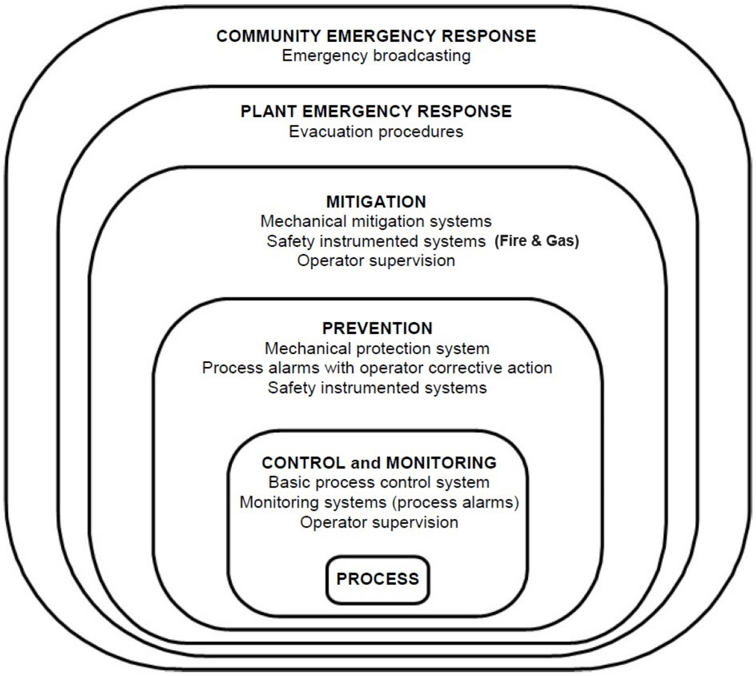

COMMUNITY EMERGENCY RESPONSE
Emergency broadcasting
PLANT EMERGENCY RESPONSE
Evacuation procedures
MITIGATION
Mechanical mitigation systems
Safety instrumented systems (Fire & Gas)
Operator supervision
PREVENTION
Mechanical protection system
Process alarms with operator corrective action
Safety instrumented systems
CONTROL and MONITORING
Basic process control system
Monitoring systems (process alarms)
Operator supervision
PROCESS

Figure 1.2: Layers of Protection [17–19]

Process safety is achieved, in part, by the implementation of functional safety utilising safety instrumented systems. Figure 1.3 shows the architecture of a safety instrumented system which consists of sensors, logic solvers and final elements.

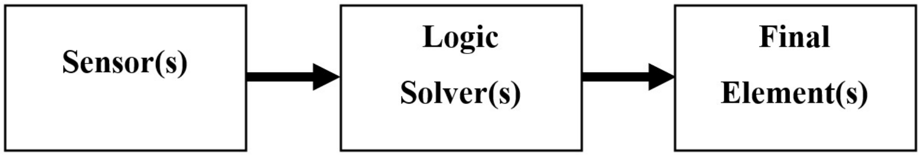


Figure 1.3: Safety Instrumented System Architecture [17]

A Safety Instrumented System (SIS) has a number of individual Safety Instrumented Functions (SIFs) that protect a process plant/facility from potentially hazardous events. SIFs are designed to protect against specific hazardous events for example, overfilling a storage tank or over pressurising a pipeline, thereby preventing a loss of containment. These systems typically culminate in the shutdown of the process that prevents the flow of a hazardous material.

Measurements of process variables such as pressure, temperature, flow, and level are achieved with sensors, typically in the form of process transmitters. The transmitters send hardwired signals to the logic solver providing real-time process measurements. A logic solver typically consists of a Programmable Logic Controller (PLC) and/or a hardwired system that make decisions based on inputs from process variable measurements. The logic solver then generates one or more outputs to the final elements as a mechanism to prevent a specific hazardous event from occurring. Final elements typically include Emergency Shutdown Valves (ESDVs) and/or contactors or relays used to stop pump motors. State-of-the-art sensors and PLCs are often equipped with onboard diagnostics, enabling some dangerous failures to be detected during operation. In contrast, ESDVs typically lack such diagnostic capabilities. The limited of diagnostic coverage emphasises the importance of proof testing ESDVs compared to other components of a SIF.

1.2.1 Industrial Asset Management and the Concept of Proof Testing

Maintenance falls under the umbrella of Industrial Asset Management (IAM), which is the comprehensive approach to the management of physical assets of a company. IAM includes developing, maintaining, modifying, and decommissioning assets in the most cost-effective manner, with guidance provided in standards such as the BS ISO 55000 [20] series. As previously discussed, many major industrial accidents have occurred due to insufficient, inappropriate, or, in some cases, a complete lack of maintenance. This section discusses the types of maintenance that are available and identifies the importance of carrying out maintenance at the appropriate frequency.

Figure 1.4 illustrates the two general types of maintenance undertaken on an industrial facility, these are corrective maintenance and preventive maintenance.

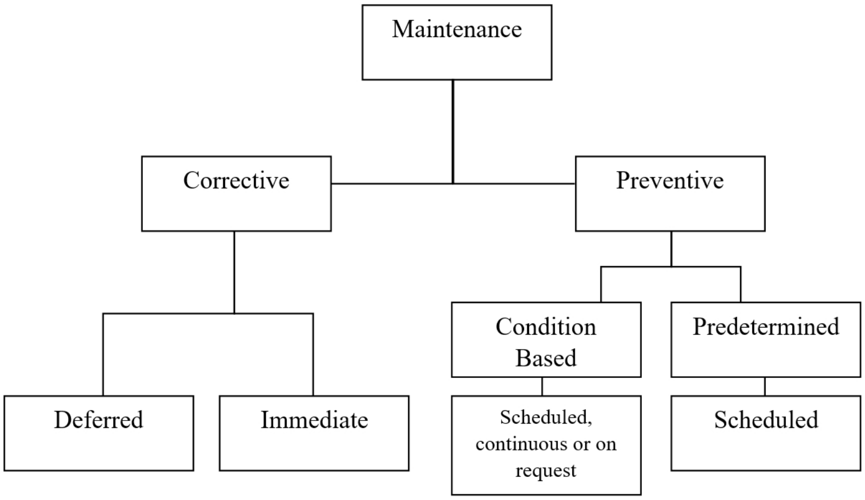


Figure 1.4: Maintenance Types [21]

Corrective maintenance, also known as repair or breakdown maintenance, is a type of task typically performed after equipment failure. Corrective maintenance is considered as all the activities that restore failed or broken down assets to their normal working condition, i.e., it is reactive in nature. Whenever an asset fails, it is replaced, repaired, or restored to its original design intent.

There are two main methodologies of preventive maintenance: predetermined maintenance and Condition-Based Maintenance (CBM). Predetermined maintenance, also known as Planned Preventive Maintenance (PPM), is time-based and consists mainly of standard maintenance activities carried out at fixed intervals. Preventive maintenance can be defined as scheduled tasks aimed at preserving an asset's functionality and lifespan, such as cleaning, adjustments, lubrication, and component replacement [22].

Condition-Based Maintenance (CBM) is a predictive maintenance strategy that monitors the real-time condition of an asset to determine what maintenance tasks are required. Unlike PPM, which uses calendar-based maintenance or other means to determine when to schedule and perform maintenance, CBM dictates that maintenance should only be performed when real-time indicators show signs of decreasing performance [23]. Figure 1.5 illustrates the typical objective and subjective techniques and tools used for condition monitoring, including more advanced 'simulated' techniques.

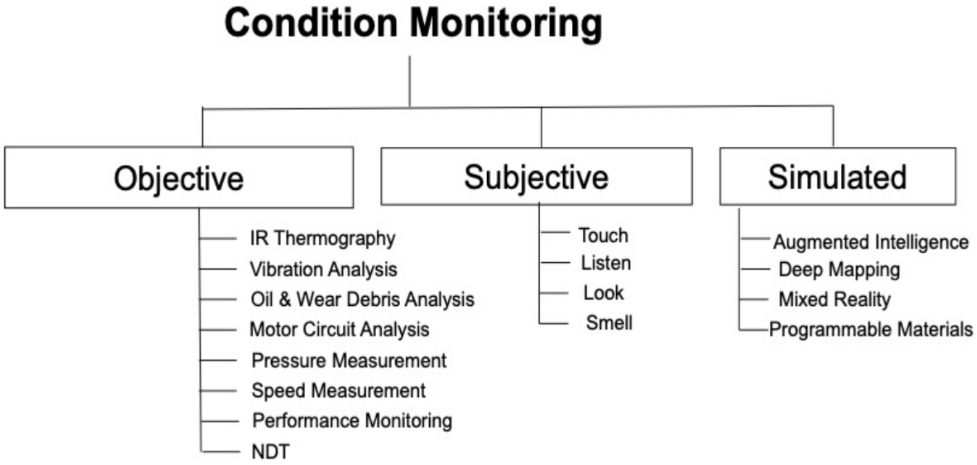


Figure 1.5: Condition monitoring techniques [22]

Proof testing in the context of process engineering is a form of PPM through which periodic functional tests are carried out on safety-related equipment such as safety instrumented systems. ESDV proof testing utilises a range of measurements and observations as an indicator for the likely condition and functionality of the component, these include measuring the closing and/or opening speed, smoothness of operation is observed and in some cases pressure measurements are used to determine seat leakage rates. Any faults found from testing are then rectified, restoring the system to its original condition. This is a form of *Opportunistic Maintenance*, in which an opportunity has arisen to repair or replace the equipment having suitable maintenance resource already on location. There is also potential for *Design-Out Maintenance (DOM)* usually a strategy after commissioning a new installation or equipment, and typically a one-off activity that allows modifications to be made allowing improved performance of the equipment. Hence, proof testing is essential in high-hazard process industries because it assures the reliability of safety instrumented systems.

Proof testing procedures are carried out at fixed intervals, which are quantitatively determined through reliability calculations. These calculations incorporate several parameters, including equipment failure rates and proof test coverage, which is influenced by the type of proof test regime applied. It is essential that these calculations use appropriate parameter values to ensure that proof test intervals are suitably defined, balancing the trade-off between over-testing and under-testing. Using inappropriate parameter values can lead to reduced test intervals, resulting in more frequent tests, increased plant downtime, and additional hours required for planning and execution. On the contrary, if the proof test intervals are calculated to be too infrequent, there is potential for safety consequences, i.e., the safety system fails to operate on demand hence leading to a hazardous event. Considering that a plant may have tens or even hundreds of safety-critical devices, the potential implications of incorrect proof test intervals can be significant, both in terms of maintenance costs and safety.

Figure 1.6 illustrates the cost implications associated with maintenance intervals, showing that the optimal point occurs where the direct costs of preventive maintenance intersect with the indirect costs of corrective maintenance.

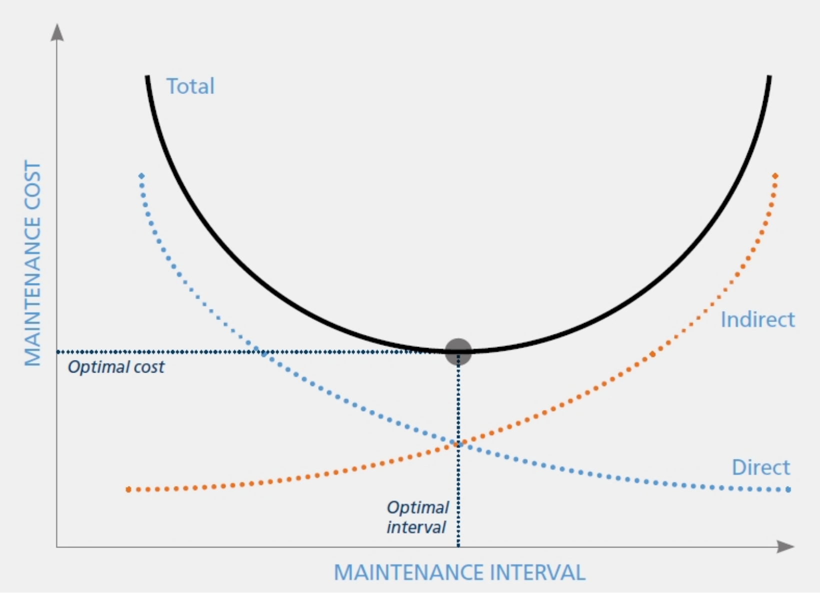


Figure 1.6: Maintenance cost v Maintenance interval [24]

1.2.2 Proof Testing ESDVs

In contrast to equipment used in process control, safety instrumented functions generally operate in low demand mode, i.e., less than one demand per year, which means that ESDVs can remain inactive in the field for extended periods. The purpose of proof testing is to reveal potentially dangerous failures, which are not revealed by equipment on-board diagnostics. For ESDVs, diagnostics are often limited or unavailable, making proof testing the most reliable method for detecting faults. The effectiveness of proof testing for detecting a failure is defined as Proof Test Coverage (PTC), which is the percentage of dangerous failures that are revealed during a proof test.

There are various proof test methods, but they are generally classified as either full stroke or partial stroke test types. The types of proof test methods seen industrially are typically imposed by the limitations of the plant design and its operation, for example, it may be disruptive or impractical to test ESDVs when the process is operating, so an offline test may be carried out or a partial test when the plant is online. The demands placed on the ESDV during testing may not fully reflect the conditions it will face when it must operate in a real scenario. For example, partial stroke testing involves limited valve movement and process conditions, such as flow, temperature, and pressure, can differ from those experienced during actual operation when the test is conducted offline. This results in an imperfect and subjective test that does not accurately reflect the true operational demands of the valve. However, it still serves as a reliable tool for guiding maintenance decisions.

Accurately estimating credible values of Proof Test Coverage (PTC) is essential for determining appropriate proof test intervals. Poor estimation of PTC can result in test frequencies that are either too low, potentially compromising safety, or too high, leading to unnecessary maintenance costs. The cost of proof testing includes factors such as man hours (for planning and execution), test equipment, and plant downtime. Plant downtime being the dominant factor, with figures in the region of £17,000 per hour for a process plant [25]. Assuming that a single proof test of an emergency shutdown valve (ESDV) takes approximately four hours, there can be a significant difference in annual maintenance costs between conducting tests twice per year and conducting them quarterly. For example, the estimated annual cost is approximately £34,000 for biannual testing, compared to £272,000 for quarterly testing. As previously discussed, a plant or facility may contain numerous ESDVs. When the total number of ESDVs is considered, the cumulative cost of proof testing can become substantial, highlighting the importance of using appropriate estimates of proof test coverage. A typical LNG terminal may operate up to 100 ESDVs, while oil refineries can have several hundred in use.

1.2.3 Failure Modes, Effects and Diagnostics Analysis (FMEDA)

Estimating proof test coverage can be carried out in a number of ways as previously described, semi-quantitatively or quantitatively. A commonly used method of estimating proof test coverage is using a Failure Modes, Effects, and Diagnostics Analysis (FMEDA). This is a systematic approach at the component level that requires an understanding of how each component can fail, its impact on the device itself, and whether the failure mode can be detected by diagnostics or revealed during testing. Part of the analysis requires the assessor to assign failure rates to the failure modes. Mechanical equipment failure rate databases exist, but these have not been standardised, which has allowed a small number of industry-based organisations such as OREDA [26], SINTEF [27], exida [28] and Technis [29] to collect failure rate data to create their own databases which are made available to the industry.

1.3 Aims and Objectives

The root cause of many major industrial accidents is the inadequate maintenance of plant and equipment [1]. Maintenance and testing of safety-related devices, such as Emergency Shutdown Valves (ESDVs) is paramount in achieving process safety [30]. The scope of this research is limited to ESDVs and does not include other elements of safety instrumented function, such as sensors and logic solvers, as these have diagnostics and are not as dependent on proof testing as ESDVs. Proof test intervals for ESDVs are determined based on a number of parameters such as Proof Test Coverage (PTC), yet published estimates of PTC vary substantially and there is no industry-wide consensus. A review of the literature has identified Failure Modes, Effects, and Diagnostics Analysis (FMEDA) as the preferred technique for estimating PTC, and is a recognised industry standard method, based on IEC 60812 [31]. The review also identified that fuzzy inference systems are effective in solving problems where there is subjectivity in the data.

The main aim of this research is to develop an inference system that addresses the subjective nature of the data used in an emergency shutdown valve FMEDA, allowing reasonable proof test coverage estimates to be determined. The research will focus on the use of the FMEDA technique using data obtained from a number of industry experts, combined with the use of an inference system as a means of handling the implicit uncertainty in the data. Fuzzy inference systems have been used previously with FMEAs but not with FMEDA; this is the originality of this research.

1.4 Thesis Layout

This chapter has offered a brief overview of the significance of correctly estimating proof test coverage for safety instrumented systems, particularly Emergency Shutdown Valves (ESDVs), while also emphasising the subjectivity inherent in component failure rate data. The remainder of this thesis will be presented in the following format.

- Chapter 2 - A literature review that encompasses functional safety and the requirements relating to proof testing and proof test coverage, as well as detailed information on the construction and operation of ESDVs. An examination of methods for estimating proof test coverage and methods for managing data subjectivity is also included.

- Chapter 3 - This chapter describes the methods used to estimate Proof Test Coverage (PTC) and presents the development of the design criteria for the Fuzzy Inference System (FIS).

- Chapter 4 - This chapter discusses a method of estimating proof test coverage utilising maintenance and test data for ESDVs collected from the industry, yielding valuable results for use and comparison with the results in later chapters.

- Chapter 5 - This chapter employs Failure Modes, Effects, and Diagnostic Analysis (FMEDA) to estimate proof test coverage, drawing on various industry data sources, including component failure databases, and expert judgement.

- Chapter 6 - Building on the previous chapter, this chapter investigates Fuzzy Inference Systems as a method for managing subjectivity in the data utilised during the FMEDA process, enabling the determination of inferred proof test coverage estimates.

- Chapter 7 - This chapter explores the influence of human error on proof testing and, consequently, how it impacts proof test coverage.

- Chapter 8 - Conclusions are drawn from this work, and recommendations for future study are presented.

Chapter 2

Literature Review

2.1 Introduction

This chapter provides an in-depth review of the literature on Emergency Shutdown Valves (ESDV) Proof Test Coverage (PTC) and approaches to managing data uncertainty, with a particular focus on Fuzzy Inference Systems (FIS). It begins with a general overview of functional safety and includes additional details on the structure and application of FIS. Figure 2.1 provides an overview of the structure of this chapter.

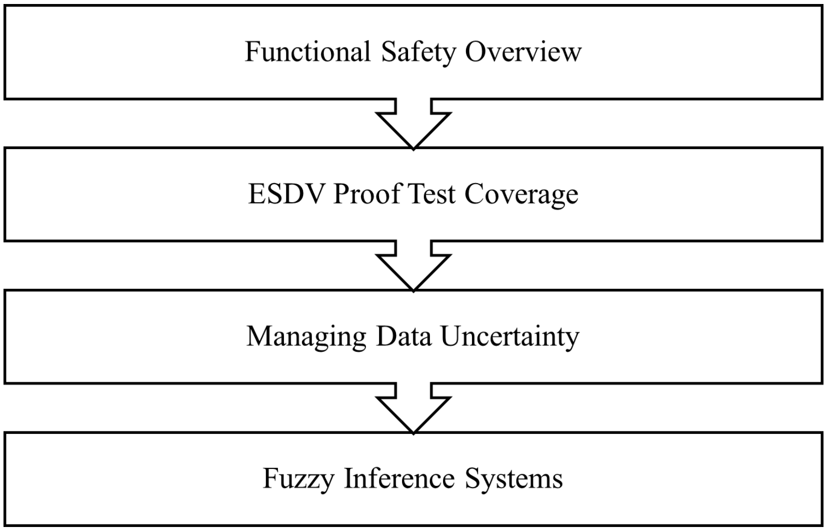


Figure 2.1: Chapter Two Structure

2.2 An Overview of Functional Safety

Functional safety is defined as being part of the overall safety that relates to the process and the Basic Process Control System (BPCS), which depends on the correct functioning of the Safety Instrumented System (SIS) and other protective layers [17].

Functional safety guidance, such as IEC 61508 [32] and IEC 61511 [17], are benchmarked within the COMAH directive [9]. These standards are seen as best practice, offering guidance on the application of appropriate safety engineering activities in the hydrocarbon and chemical industries.

As prescribed by functional safety standards, implementation is achieved using a life cycle approach from the early stage of undertaking a hazard and risk analysis to decommissioning. Figure 2.2 illustrates a typical SIS life cycle, beginning with the Hazard and Risk Analysis phase and progressing to the Operation and Maintenance phase as the end goal.

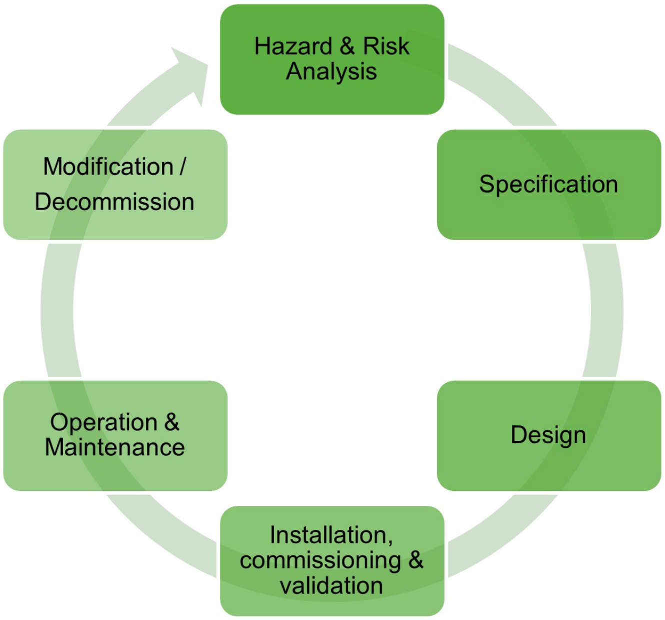

Hazard & Risk Analysis
Specification
Design
Installation, commissioning & validation
Operation & Maintenance
Modification / Decommission

Figure 2.2: Safety Instrumented System Life Cycle [17]

The following provides a brief overview of each phase of the life cycle [33]:

- Hazard and risk analysis - perform hazard and risk analysis, typically Hazard and Operability Study (HAZOP) [34], Layer of Protection Analysis (LOPA) [18, 35, 36], or Fault Tree Analysis (FTA) [17, 37], to define the required functionality and integrity for each protective function and the allocation of safety functions required to achieve the necessary risk reduction to specific protection layers. Safety Integrity Level (SIL) targets are determined for specific safety instrumented functions.

- Specification - develop a Safety Requirements Specification (SRS) [17] to achieve the required functionality and integrity while meeting plant targets for reliability, maintainability, and operability.

- Design - the hardware and software of the SIS is designed to meet the requirements of the safety requirements specification. Consideration of how Safety Instrumented Functions (SIFs) will be tested and maintained.

- Installation, commissioning, and validation - installation of the SIS and other protective layers. Followed by commissioning and validation to ensure that the SIS meets the intended design and specification.

- Operation and maintenance - specify the requirements necessary to ensure safe operation across all modes, including start-up, normal operations, abnormal conditions, and shutdown. Maintain the installed equipment, inspect and proof test to verify its intended functionality and confirm it meets the target integrity and reliability standards.

- Modification - to include analysis of the data collected from hazard and risk assessments, plant operations, and maintenance activities to assess if modifications are required to maintain or enhance protective functions.

Throughout the life cycle, emphasis is placed on key phase verification and functional safety management activities, including competency management, planning, auditing, functional safety assessments, and configuration management [17].

2.2.1 Safety Instrumented Systems

A Safety Instrumented System (SIS) comprises multiple Safety Instrumented Functions (SIFs), each designed to protect a process plant or facility from potentially hazardous events. SIFs are designed to protect against specific hazardous events for example, overfilling a storage tank or over pressurising a pipeline, hence preventing a loss of containment. These systems typically culminate in the shutdown of the process that prevents the flow of a hazardous material. Figure 2.3 illustrates a typical SIS architecture consisting of sensor(s), logic solver(s) and final element(s) [17]:

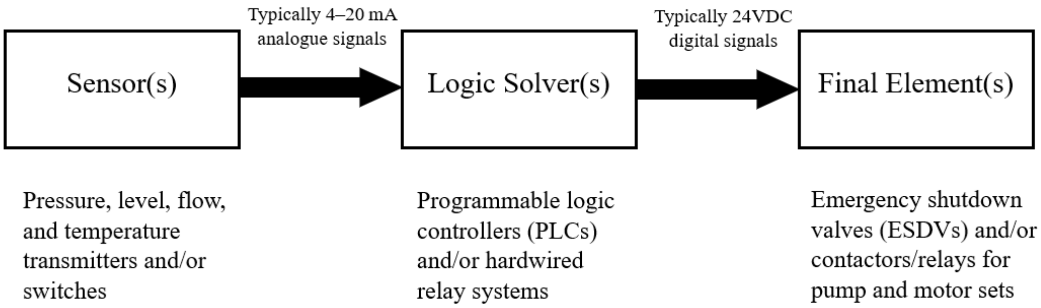

Typically 4–20 mA analogue signals
Typically 24VDC digital signals
Sensor(s)
Logic Solver(s)
Final Element(s)
Pressure, level, flow, and temperature transmitters and/or switches
Programmable logic controllers (PLCs) and/or hardwired relay systems
Emergency shutdown valves (ESDVs) and/or contactors/relays for pump and motor sets

Figure 2.3: SIS Architecture

Sensors typically include process transmitters (pressure, temperature, flow, level). Logic solvers typically include Programmable Logic Controllers (PLCs) and/or hard-wired systems. Final elements typically include Emergency Shutdown Valves (ESDVs), and/or contactors/relays (for pump motor stops). State-of-the-art sensors and PLCs have onboard diagnostics allowing a quantity of dangerous failures to be diagnosed when in operation, unlike ESDVs which do not have onboard diagnostics. The lack of diagnostics emphasises the importance of proof testing ESDVs compared to other components of a Safety Instrumented Function (SIF).

Figure 2.4 illustrates an example of a storage tank overfill protection SIF. In this scenario, a storage tank is being filled with a highly flammable liquid, Liquefied Natural Gas (LNG), where overfilling could result in a loss of containment and potentially hazardous consequences such as fire and/or explosion.

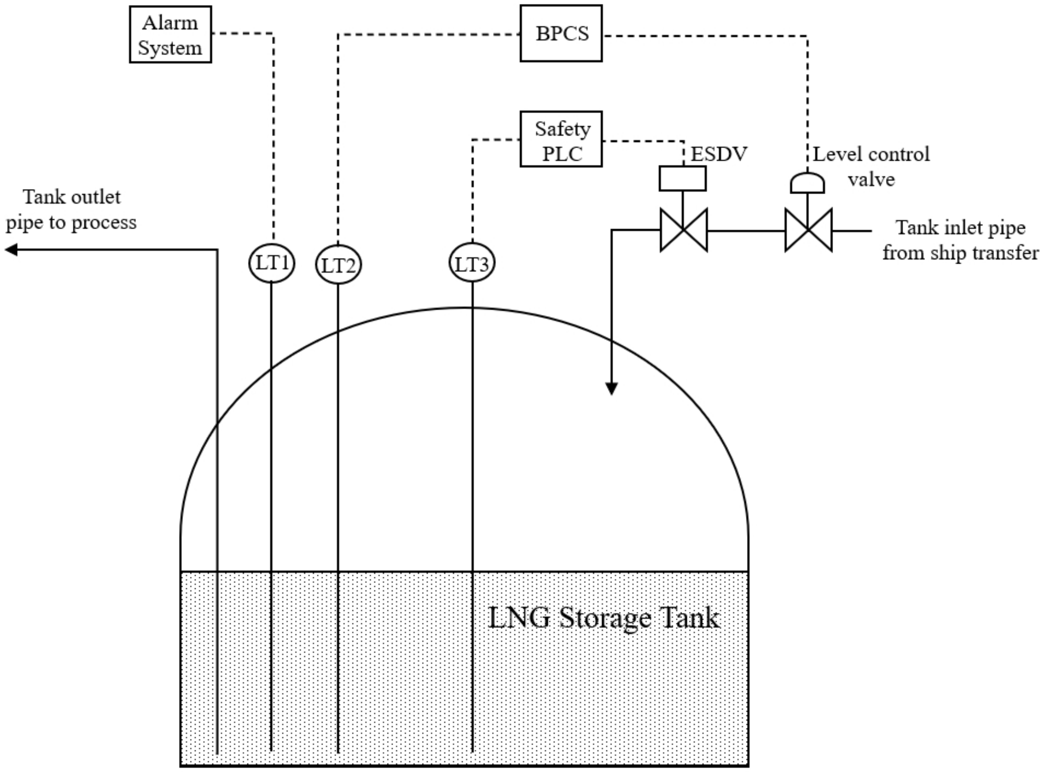


Figure 2.4: LNG Storage Tank with a Level Control System and Overfill Protection

The Basic Process Control System (BPCS) controls the level within the LNG tank by measuring the actual level using level transmitter LT2 and comparing the measurement with a defined control set-point. The level control valve is modulated accordingly to maintain the desired level in the tank. Sensors are typically transmitters, also known as Self-Monitoring, Analysis, and Reporting Technology (SMART) devices and are capable of continuously monitoring process measurements and have on-board diagnostics. The high-level alarm generated by the LT1 level transmitter alerts the process operator in the control room if the BPCS fails to maintain the LNG level within the desired range. In the event that the level control system and the high-level alarm system on the tank fails, the safety instrumented function, which is an independent layer of protection, will prevent overfilling. The level transmitter (LT3) measures the liquid level and sends the signal to the safety PLC which continuously monitors the level. When the tank liquid level is above a certain level, known as a trip set point, the safety PLC action is to initiate the closing of the ESDV to prevent further filling of the tank.

Figure 2.5 shows a High Integrity Pressure Protection System (HIPPS). The HIPPS is a Safety Instrumented Function (SIF) that protects a process line from over-pressurising. Overpressure can damage the process line, potentially causing a rupture and loss of containment of hazardous materials, which may escalate into a dangerous event such as a fire and/or explosion. Overpressure may be caused by, for example, failure of equipment downstream or human error, i.e., closing an isolation valve in error. Upon detection of high pressure in the process line, the logic solver activates the closure of the ESDVs. The HIPPS configuration has multiple inputs, i.e., pressure transmitters (PT1, PT2, PT3) and multiple outputs, i.e., emergency shutdown valves (ESDV1, ESDV2). Multiple inputs and outputs provide redundancy, enabling the system to tolerate a single device failure, a concept known as Hardware Fault Tolerance (HFT). This redundancy enhances the integrity of the SIF, thereby improving overall system reliability. The pressure transmitters are configured in a two-out-of-three (2oo3) voting arrangement requiring two of three transmitters to register a high pressure before activation, or ”tripping”, of the system, and thereby causing the ESDVs to close. The 2oo3 configuration also lends itself to reducing spurious trips, i.e., unintended activation of the safety instrumented function caused by a transmitter fault. ESDVs are configured in a one-out-of-two (1oo2) voting arrangement requiring either valve to close to ensure a safe state is achieved.

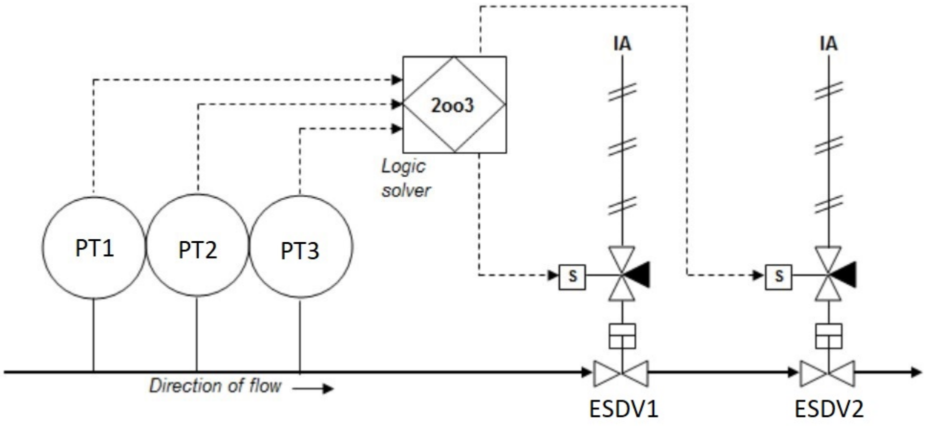


Figure 2.5: HIPPS Architecture [38]

As part of ongoing maintenance, SIFs must be proof tested periodically to reveal dormant dangerous failures, i.e., failures which are neither self-evident nor detected by automatic diagnostics [17]. The frequency at which a device is proof tested is determined by calculating the unavailability of the SIF, expressed as the Probability of Failure on Demand Average ($\mathrm{PFD_{avg}}$).

Table 2.1 shows the Safety Integrity Levels (SILs), for demand mode of operation, and corresponding $\mathrm{PFD_{avg}}$ and Risk Reduction Factor (RRF) values. As discussed previously, the SIL target for a safety instrumented function is determined prior to the specification and design phase as part of the hazard and risk assessment process.

Table 2.1: Safety Integrity Level, $\mathrm{PFD_{avg}}$ and Risk Reduction Factor (RRF) [17]

Safety Integrity Level	$\mathrm{PFD_{avg}}$	Risk Reduction Factor
1	$\geq 10^{-2}$ to $< 10^{-1}$	> 10 to ≤ 100
2	$\geq 10^{-3}$ to $< 10^{-2}$	> 100 to ≤ 1000
3	$\geq 10^{-4}$ to $< 10^{-3}$	> 1000 to ≤ 10000
4	$\geq 10^{-5}$ to $< 10^{-4}$	> 10000 to ≤ 100000

Parameters considered in the $\mathrm{PFD_{avg}}$ calculations include equipment failure dangerous undetected failure rates (λ_{DU}), Proof Test Interval (PTI), Proof Test Coverage (PTC) and Mission Time (MT). The equation for $\mathrm{PFD_{avg}}$, in its basic form, is given by [25, 39]:

$$PFD_{avg} = PTC \times \lambda_{\mathrm{DU}} \times \frac{PTI}{2} + (1 - PTC) \times \lambda_{\mathrm{DU}} \times \frac{MT}{2} \tag{2.1}$$

MT is the period of time between when the SIF (or device) is put into service and when it is replaced or completely refurbished to ”as new” condition [40].

Smith [41] provides extended $\mathrm{PFD_{avg}}$ equations specifically designed for testing regimes that incorporate both partial and full stroke testing.

Proof Test Coverage (PTC) is determined by the ratio of the sum of the revealed dangerous undetected failure rates (*Revealed* λ_{DU}) to the sum of the total dangerous undetected failure rates (*Total* λ_{DU}), and expressed as a percentage [25]:

$$PTC = \frac{\mathit{Revealed}\ \lambda_{DU}}{\mathit{Total}\ \lambda_{DU}} \tag{2.2}$$

2.2.2 ESDV Assembly Components

An Emergency Shutdown Valve (ESDV) is a complex electromechanical device comprising numerous components, all of which are potentially susceptible to failure in service. This section outlines the main components of an ESDV: the Solenoid-Operated Valve (SOV), actuator, and valve. Figure 2.6 shows a schematic of the main components of an ESDV assembly, and Figure 2.7 shows a photo of an ESDV assembly in the field.

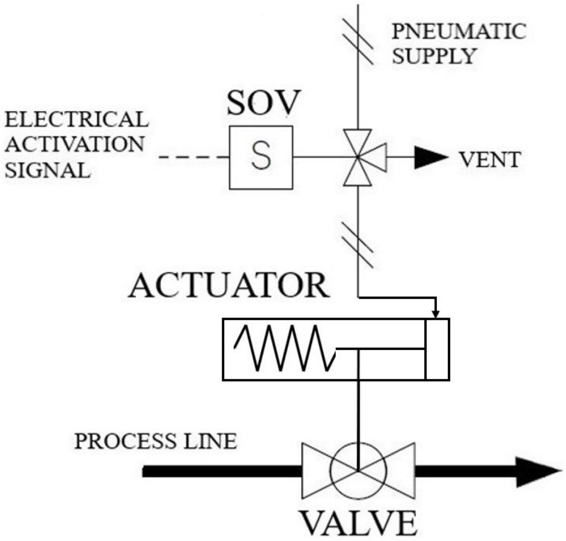


Figure 2.6: ESDV Assembly Schematic

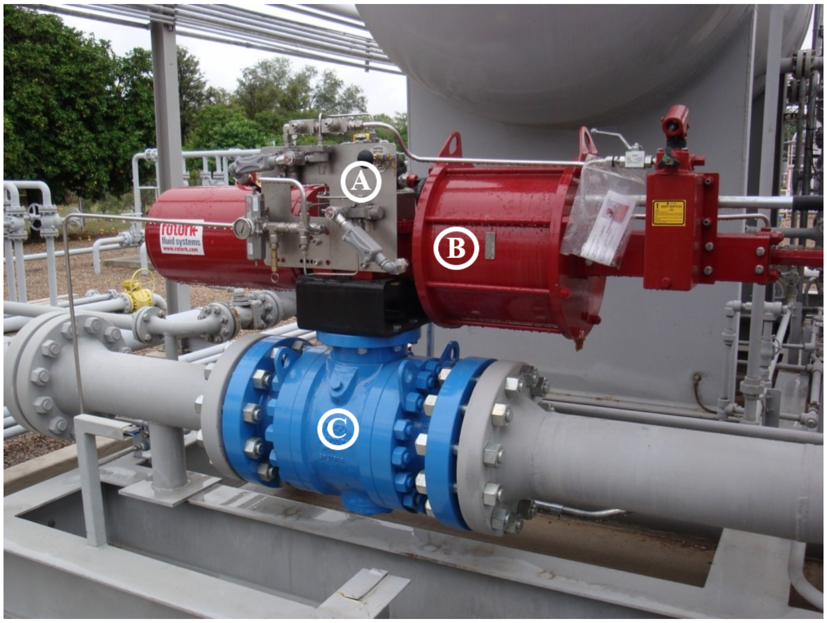


Figure 2.7: ESDV assembly in the field (A - SOV, B - Actuator, C - Valve) [42]

An SOV is a device that, upon activation by an electrical signal, delivers pressurised air to an actuator. The actuator cylinder becomes pressurised and its piston moves in a linear motion. The linear motion is converted to angular motion, via a rack-and-pinion or scotch-yoke mechanism, allowing the valve to turn through 90 degrees to its driven closed or open position depending on its fail safe mode. Removal of the electrical signal allows the actuator cylinder to vent to atmosphere, allowing the actuator to return to its original position via spring force, hence opening or closing the valve.

2.2.2.1 Solenoid-Operated Valve (SOV)

An SOV is an electromechanical device that controls the supply of air to the actuator, upon receiving an electrical signal from the logic solver, and determines whether or not a valve is activated (opened or closed) by providing pressured air to the valve actuator. As described earlier, an SOV is used to switch pressurised air to the actuator to allow the actuator to open or close the valve. Figure 2.8 shows a commonly used SOV in ESD applications, a pneumatic 3/2 direct-acting poppet type. A poppet is a type of valve that is used in many mechanical systems, such as pneumatic or hydraulic devices. It consists of a movable component, typically a disk or tapered plug, that seals against a valve seat to control the flow of gases or fluids.

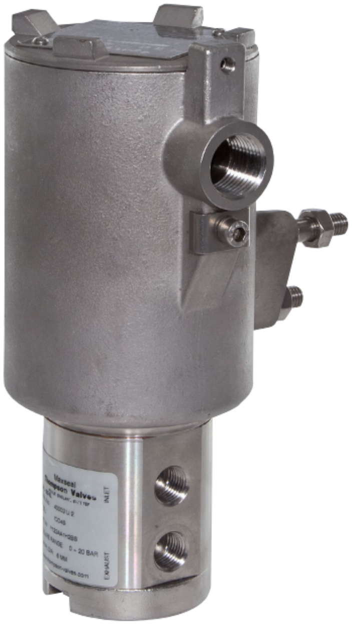


Figure 2.8: Solenoid-Operated Valve (SOV) [43]

3/2 relates to: three ports (pressure, control and exhaust) and two positions (energised and de-energised). An SOV used in safety applications is normally designed so that it is De-energised To Trip (DTT), meaning that in normal operation the device is energised allowing air to flow to the actuator, maintaining the valve in its open position. In the event of a demand or a failure in signal integrity, the SOV will de-energise, causing the valve to close. Operation of an SOV is via a solenoid coil which is energised with an electrical signal from the safety PLC (typically 24VDC). The magnetic field generated by the coil moves the armature which is connected to the poppet via a stem, thus allowing instrument air to be supplied to the actuator.

Figure 2.9 shows the energised solenoid position. Pneumatic air supplied from the instrument air supply is connected to the pressure port (P). The control port (C) is connected to the supply port of the actuator. In this position the actuator cylinder will be pressurised by the instrument air flow from pressure port (P) to control port (C).

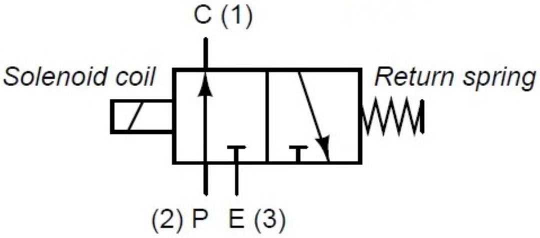


Figure 2.9: Energised position - with electrical signal applied

Figure 2.10 shows the de-energised solenoid position. In this position, the actuator cylinder will vent through the exhaust port of the SOV, allowing flow from the control port (C) to the exhaust port (E). The exhaust port (E) vents to atmosphere.

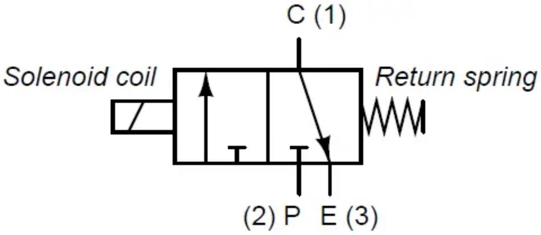


Figure 2.10: De-energised position - without electrical signal applied

Figure 2.11 illustrates a more detailed view of a SOV.

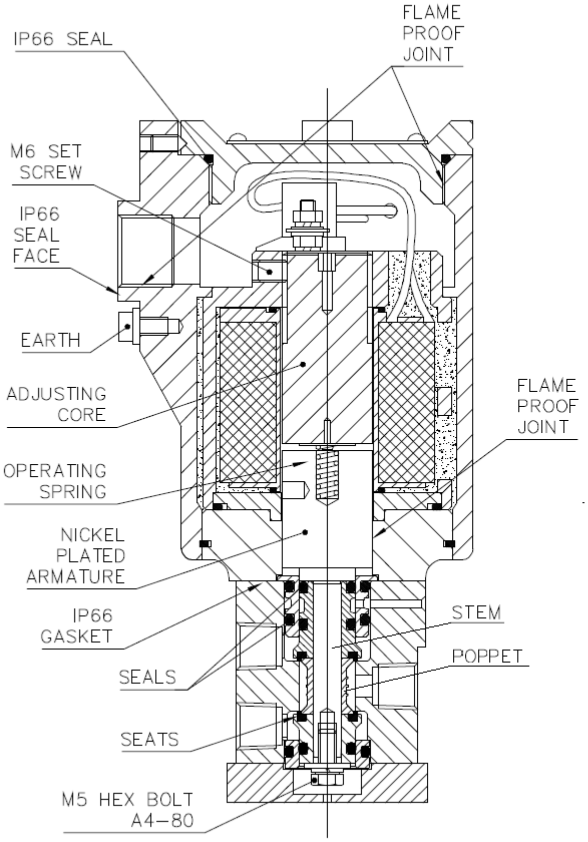


Figure 2.11: Solenoid-Operated Valve (SOV) - detailed view [43]

2.2.2.2 Actuator

An actuator is typically a pneumatic powered device that supplies force and motion to open or close a valve via the air supplied from the SOV. This movement is normally countered by a spring that holds the actuator in a normally closed or normally open position unless the actuator receives a pneumatic signal. An actuator is a mechanical device that converts linear movement of the actuator piston into rotational torque, to allow movement of a quarter-turn valve. Actuators are typically either single-acting or double-acting. Single-acting actuators use air pressure to move the piston in one direction and rely on a spring for the return stroke. In contrast, double-acting actuators require air pressure to move the piston in both directions and do not use a spring. Single acting actuators are typically used for SIS ESDVs as they are fail-safe by design, loss of air supply will allow the valve to move to its safe state via spring force. A commonly used actuator, the single-acting scotch yoke type, is made of three main components:

- Housing containing the yoke mechanism.
- Pressure cylinder containing the piston.

- Spring enclosure.

When air pressure is supplied into the cylinder chamber, a linear force is applied to the surface of the piston. Due to the movement of the sliding block in the yoke slot, an anti-clockwise rotation is generated on the yoke and the valve is driven to the open position. At the same time, the actuator spring is compressed. Figure 2.12 illustrates when air pressure is removed, the spring is automatically released, the yoke moves clockwise, driving the piston back to the starting position and thus the valve to the fail-safe closed position.

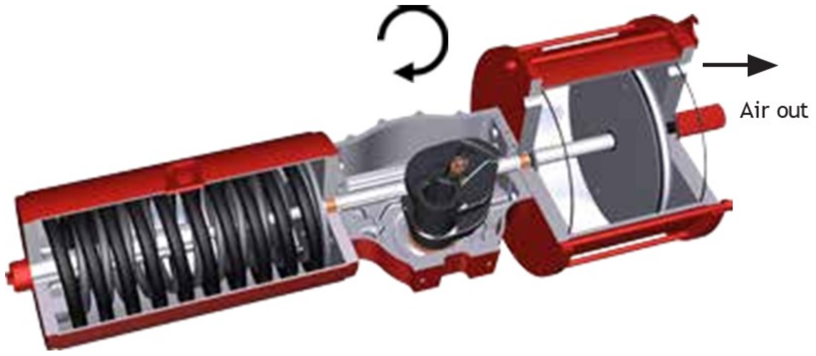


Figure 2.12: Scotch Yoke Actuator [44]

Figure 2.13 illustrates a more detailed view of a scotch yoke actuator.

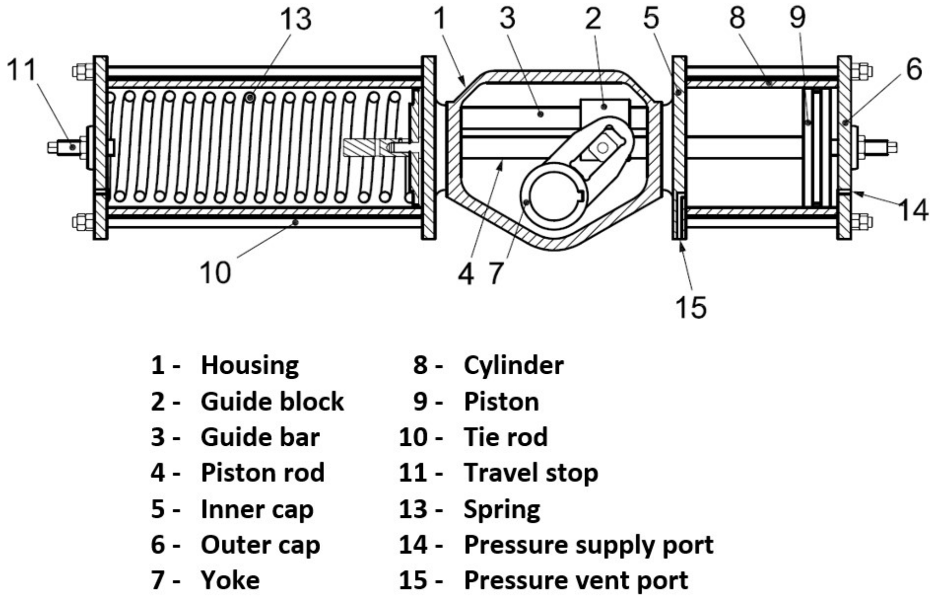


Figure 2.13: Actuator - detailed view [45]

2.2.2.3 Valve

A valve is a device for controlling the flow of material in a pipeline and contains the components that are wetted by the process fluid. Figure 2.14 illustrates a ball valve, a commonly used valve in SIS applications. A ball valve is characterised by a long service life and can provide a reliable sealing (two sealing faces) over its life span, even when not in use for long periods.

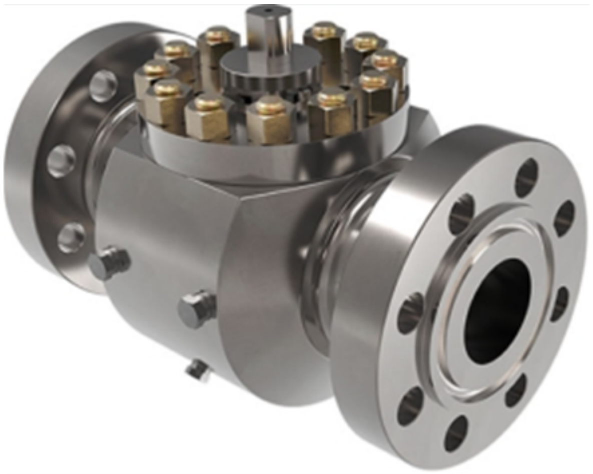


Figure 2.14: Ball Valve [46]

A valve is made up of a number of components, typically:

- Body – the pressure containing part of the valve, which contains the components that contact the fluid.
- Bonnet – closure for the valve body through which the stem passes.
- Obturator – a ball, disc, gate, or plug that is positioned in the flow stream to permit or prevent flow with either linear or rotary motion.
- Stem – the connector from the actuator to the inside of the valve – transmits force to move the obturator.
- Seat ring – the surface that the valve plug contacts when the valve is closed, thus forming the seal.

Within these components there are subcomponents such as springs, bearings, and seals which can fail individually thus causing the failure of that component.

Figure 2.15 shows the operation of a ball valve, with the ball moving from the open to the closed position (from left to right), noting that this is a manually operated ball valve as opposed to an actuated ball valve.

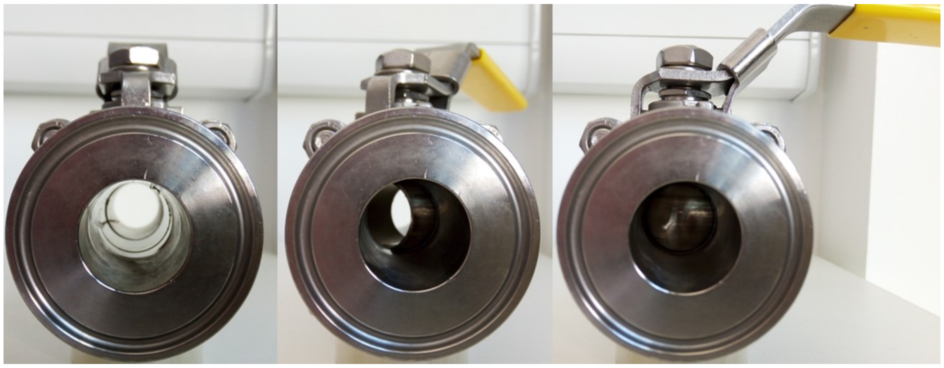


Figure 2.15: Ball valve operation – sequence from left to right shows the valve transitioning from fully open to fully closed.

Figure 2.16 illustrates a detailed view of a quarter-turn top entry full-bore, trunnion-mounted ball valve with seat rings and springs.

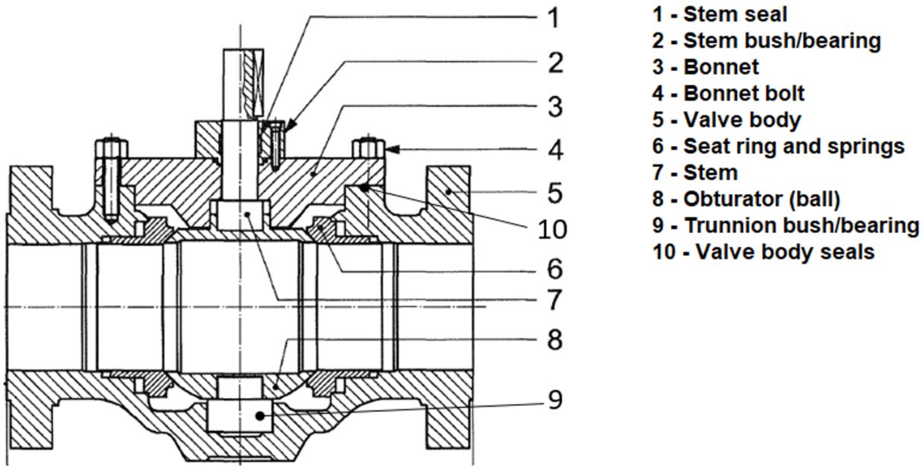


Figure 2.16: Ball Valve - detailed view [45]

Full bore indicating that the internal diameter of the valve is the same as the pipe to which it is connected, used where the system pressure drop is a factor, in contrast to a reduced bore which has a smaller diameter than the pipe to which it is connected. A trunnion is used for larger valve sizes and is a supporting mechanism that allows for more stability of the ball, hence reducing misalignment issues. A valve without a trunnion is commonly known as a "floating" type.

An ESDV is a complex electromechanical device designed to act as a critical safety barrier in process facilities. Its complexity arises from the integration of various components, including springs, seals, mechanical linkages, bearings, and bushes each of which can contribute to its overall functionality. This intricate design inherently introduces numerous potential failure modes, such as mechanical wear, seal degradation, component binding, and electrical faults. These potential failures underscore the importance of rigorous testing, maintenance, and reliability analysis to ensure that the ESDV performs its intended safety function effectively under all operational conditions.

2.2.3 Proof Test Methods

Technis [76], Stewart [47], NAMUR [48] and Dearden [49] provide PTC values for various types of proof tests, ranging from partial stroke testing to full stroke testing conducted under operating conditions, including seat leakage testing. Partial stroke testing is an option for a continuously running plant, as a full stroke test is not possible without shutting the plant down [50]. Furthermore, Brissaud et al. [51] suggests a combination of partial and full testing as a means to optimise the performance of safety-related systems. Stewart [47] identifies five different ESDV assembly proof tests, which include:

1. Partial stroke test - the valve is typically moved from 5% to 20% [52–54], online (under process operating conditions).
2. Full stroke test – the valve moved to its fully closed (or open) position, offline (not under process operating conditions).
3. Full stroke test at process operating conditions – as per test (2) with the valve online.
4. Full stroke test and leak test – with the valve offline.
5. Full stroke test at process operating conditions and leak test – as per test (4) with the valve online.

Full stroke testing under operating conditions, when the plant is live and the valve is online, is generally considered undesirable from an operator's perspective, as it disrupts process conditions and may potentially lead to a plant shutdown. In practice, offline Full Stroke Testing (FST) and online Partial Stroke Testing (PST) are the most commonly used test regimes in industry. As discussed, partial stroke testing is usually implemented in facilities where the process is continuously operating for extended periods between planned shutdowns, such as refineries and chemical plants. The trade-off is that a FST offers greater proof test coverage but can disrupt plant operations, whereas a PST provides lower proof test coverage but is less disruptive as it can be conducted without affecting plant uptime. Depending on the plant's design and operation, proof testing can be optimised by combining full and partial tests, ensuring maximum proof test coverage while minimising plant downtime.

2.2.3.1 Process Safety Time

Process safety time is a parameter defined as the time period between a failure that occurs in the process or the basic process control system (with the potential to cause a hazardous event) and the occurrence of the hazardous event if the Safety Instrumented Function (SIF) is not performed [17].

Calculating process safety time allows the response time of the SIF to be determined which is a key design parameter and a SIF safety requirement [17]. The ESDV is specified with a closing time to align with the SIF response time. The specified closing time then serves as a proof test metric for assessing the ESDV's performance and verifying the Delayed Operation (DOP) failure mode. Closing time is also a good measure of valve performance, as historical data will provide insight into potential degradation.

2.2.3.2 Tight Shut-Off (TSO)

Technis [76], Stewart [47], NAMUR [48] and Dearden [49] have also included test regimes for Tight Shut-Off (TSO) applications. TSO refers to the ESDV's ability to completely stop the flow of a fluid when it is in the closed position, ensuring no measurable leakage past the valve. TSO is typically an operational requirement for the isolation of plant sections rather than a means to suppress a potentially hazardous event; typically, simply throttling the flow is sufficient for a Safety Instrumented Function (SIF) [49]. This is a common misconception, as company standards often mandate TSO without adequately considering the specific requirements of the SIF. Specifying TSO leads to onerous testing and maintenance; therefore, to achieve an effective proof test, the test must include a seat leakage test that meets the specified TSO seat leakage class. In reality, this type of testing is difficult to implement as additional equipment needs to be considered for test and verification purposes, which typical process piping designs do not normally incorporate. If online seat leak testing is not feasible, the test can be performed with the valve out of service in a workshop environment. In this setup, the valve is pressurised on the inlet side, while the leak rate is monitored on the outlet side. Removing and reinstalling equipment can introduce additional risks, such as human error during reinstallation and the need for recommissioning and testing before the equipment is returned to service. Therefore, it is essential to ensure that ESDVs are properly specified to meet the requirements of the SIF.

2.2.4 Partial Stroke Testing (PST) in More Detail

Partial Stroke Testing (PST) complements Full Stroke Testing (FST) rather than replacing it. The reliability improvement obtained by introducing PST can be used in two ways [55]:

1. To improve safety: PST is added to the initially scheduled FSTs. This leads to a reduction in the calculated PFD_{avg}. With a lower PFD_{avg}, the safety improves.

2. To reduce costs: The potential reliability improvement is used to extend the interval between FSTs so that the calculated PFD_{avg} is kept unchanged. As a

result, operating and maintenance costs may be reduced as less man-hours and less plant downtime.

An important aspect of functional safety is the ability to detect device failure, e.g., incipient, degraded, and functional, so that identified failures can be corrected. For actuated valves, such as ESDVs, complete online testing is limited in many applications. The extension of fixed equipment maintenance intervals, i.e., plant outages, has resulted in reduced off-line test opportunities, and the actuated valve must meet the required reliability in the operating environment.
PST is a method of proof testing that requires the valve to be stroked typically 5% to 20% of its travel, allowing testing to be carried out online without impacting the process. PST can be performed on-line with a wide variety of equipment and can be executed either manually or automatically. The focus of this research relates to manual periodic PST. PST may provide early detection of valve failures and provide indications of developing valve movement failures.
The main advantages of performing PST is the ability to test a valve assembly when the process is online without having to bypass the process. The limitations of a PST is that it is not possible to verify certain failure modes in comparison to an FST. For example, it is not possible to verify the following dangerous failure modes:

- Fail to Close (FTC) - although it is possible to verify stuck in the open or closed position, it is not possible to determine whether the valve would fully close. Therefore, it is possible to partially verify FTC.

- Delayed Operation (DOP) - it is not possible to verify if the valve assembly would close in the specified time, but it is possible to verify the closure time for 20% of the valve travel.

- Leak in the Closed Position (LCP) - it is not possible to verify valve seat leakage using PST.

2.2.4.1 Partial Stroke Testing (PST) Methods

The following describes a number of different methods of carrying out a partial stroke test, both manual and automated type tests:

- Mechanical limiting - Mechanical limiting methods involve the installation of a mechanical device to limit the degree of ESDV travel [56]. When mechanical limiting methods are used, the ESDV is not available for process shutdown during the partial stroke test. Mechanical devices used for partial stroke testing include a valve collar, valve jack, mechanical jammer, or an actuator-integrated mechanical stop.

- Travel limiting devices - These devices include any pneumatically or hydraulically engaged mechanical travel limiting devices that are internal to the ESDV actuator. These limiting methods may involve the addition of tandem pistons within the actuator, where adjustable tandem pistons act as pneumatically engaged travel stops for the primary pistons. While the primary pistons enable full actuator stroking, the tandem pistons provide adjustable stroke limitation [58].

- Valve positioner - These devices use an electro-pneumatic or electro-hydraulic positioner to move the valve to a pre-determined point. An option for the automatic PST function requires the use of the intelligent valve positioner instead of a Solenoid-Operated Valve (SOV). PST can be conducted manually or automatically as a manufacturer standard function of the positioner. When the positioner receives the start trigger, the ESDV moves to a preset position and then returns to the original position, with continuous monitoring of the ESDV position and travel time [57]. Positioners transmit diagnostic information to maintenance personnel through both local and online systems.

- Solenoid-Operated Valve (SOV) - Several different approaches have been used to perform a PST by removing or reapplying power as required to allow SOVs to change state. The test can be conducted by operating a field-mounted switch, which de-energises the solenoid coil for as long as the switch is held. The field operator monitors the ESDV position and releases the switch once the ESDV movement is confirmed. When the ESDV moves, it can be inferred that the SOV has successfully vented. One of the risks is that the operator may hold the switch too long, allowing the ESDV to disrupt the process, resulting in a possible plant shutdown [56]. The use of pneumatic travel limit switches to restrict the travel of the ESDV can be used to prevent such scenarios. An example of this type of configuration with limited actuator travel is shown in Appendix A.1.

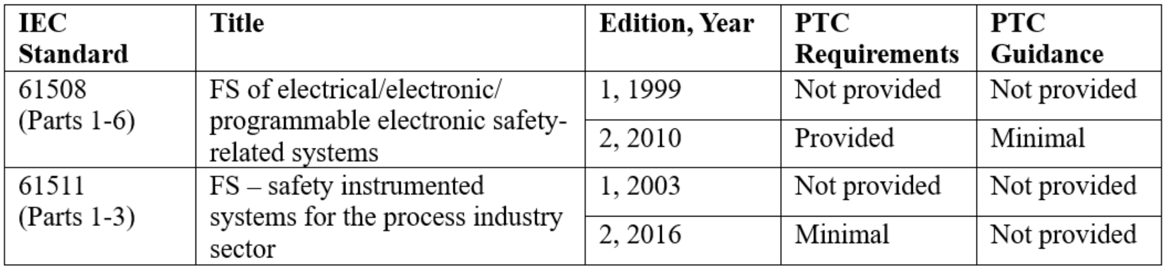
  During PST with a simplex SOV, the solenoid is de-energised and then re-energised. Failure of the SOV to reset will cause the ESDV to fully close and shut down the plant. Using redundant solenoid valves can effectively prevent this problem [58].

2.3 ESDV Proof Test Coverage

2.3.1 Importance of Proof Test Coverage and Guidance

As highlighted by Gubert et al. [59] the majority of the literature on the subject of proof testing tends to ignore Proof Test Coverage (PTC), hence the assumption that proof tests are perfect and capable of detecting all failure modes. This overestimation can lead to very optimistic results. This approach is also reflected in industry, many operators do not consider PTC in SIF Probability of Failure on Demand Average (PFD_{avg}) calculations. This issue is very much historical, as during the infancy of functional safety, even associated international standards, IEC 61508 [32] and IEC 61511 [17], overlooked PTC. These standards have since evolved with revisions that now include PTC requirements, albeit with limited guidance. Table 2.2 shows the evolution of functional safety standards in relation to PTC:

Table 2.2: IEC Functional Safety standards historical PTC requirements and guidance



IEC Standard	Title	Edition, Year	PTC Requirements	PTC Guidance
61508 (Parts 1-6)	FS of electrical/electronic/ programmable electronic safety-related systems	1, 1999	Not provided	Not provided
		2, 2010	Provided	Minimal
61511 (Parts 1-3)	FS – safety instrumented systems for the process industry sector	1, 2003	Not provided	Not provided
		2, 2016	Minimal	Not provided

The effect of imperfect testing has a negative impact on the SIF PFD_{avg}, and after each proof test the PFD_{avg} increases; eventually, after several intervals of proof testing, the PFD_{avg} is significantly higher due to the poor effectiveness of the proof test [60]. It is important that the SIF design phase incorporates a design that produces the most effective proof test and that the effectiveness of the test, i.e., PTC is incorporated within the SIF PFD_{avg} calculations [61]. Likewise, Hokstad et al. [62] places emphasis on incorporating PTC in the SIF PFD_{avg} calculations if the proof tests are imperfect. Similarly, Green and Bell [25], Abonyi et al. [63] and Rielly [64] demonstrate by calculation, using various values of PTC, the effects on PFD_{avg}. In some cases, there can be a significant difference, of an order of magnitude, in the results, which emphasises the importance of including accurate estimates of PTC in the calculations. Jin et al. [65] also emphasises the effects of imperfect proof tests on system PFD_{avg} using a modelling tool presented by Hauge et al. [66] which differs from the IEC 61508 [32] approach – but the results are comparable. There is consensus that guidance on determining PTC is limited; Lundteigen and Rausand [67] suggest that many authors on the subject of PFD_{avg} calculations have discussed how to include PTC in SIF PFD_{avg} calculations, but not how to determine test coverage. Similarly, Jin et al. [68] states that the guidance is sparse and that the authors have discussed approaches to determine the PTC that are based on expert judgement, and that there is little guidance available on how to estimate PTC, and the values used in many SIF PFD_{avg} calculations are assumptions. Furthermore, Smith [69] adds that there are several sources that offer guidance on the subject, but due to the lack of data related to the results of the proof test, it is more likely that they are based on subjective judgement rather than empirical data.

2.3.2 Methods for Estimating PTC

Several methods are used to determine the PTC, including semi-quantitative and quantitative methods. Several authors have identified the use of a quantitative Failure

Modes and Effects Analysis (FMEA) as a means of determining proof test coverage.

FMEA is a qualitative method that identifies potential failure modes, their causes, and effects on a system. In contrast, Failure Modes, Effects and Diagnostics Analysis (FMEDA) extends FMEA by adding quantitative data such as failure rates and diagnostic coverage, making it suitable for safety-critical applications. FMEDA is a technique originally developed to determine the diagnostic coverage of electronic equipment, by identifying the failure modes of electronic components and their effects on the system and whether the failures are detected or not by diagnostics [70]. The technique requires component failure rates to be assigned to failure modes which allows the distribution of failures and the overall diagnostic coverage to be determined. The FMEDA technique as described by Stewart [47] can also be used to determine proof test coverage of mechanical systems such as ESDVs. Similarly, Easton [71] and Green and Bell [25] agree that FMEDA is an appropriate tool for this purpose. Furthermore, Lundteigen and Rausand [67] suggest the use of an FMEA-type study to determine proof test coverage for partial stroke testing valves. Following the work of Bukowski and van Beurden [72], estimates of the effectiveness of the proof test can be determined using FMEDA and evaluating dangerous undetected failures that can be revealed from specific proof test procedures.

Easton [71] and Green and Bell [25] suggest that the preferred approach is to use IEC 61508 compliant equipment and refer to the manufacturer's safety manual for PTC values. Easton [71] also suggested that, in cases where the valve safety manual provides limited information, the operator (end user) should conduct an FMEDA. If this is not feasible, alternative methods for estimating PTC are recommended, such as the approach presented by Technis [76]. Valve manufacturer safety manuals are a valuable resource and a good starting point; however, the proof test information they provide often does not align with the specific requirements of the end user and the test procedures are frequently vague in their descriptions. For example, it is common for manufacturers to provide a PTC value for a full stroke proof test that requires verification of valve seat leakage. In reality, it is much the case that the valve proof test does not include this verification, hence the quandary of knowing what value is appropriate for PTC. With regard to the operator conducting an FMEDA and comparing the output against the proposed proof test procedure to determine the PTC – this requires expertise that is not readily available to the end user.

2.3.3 Failure Rate Data

Failure data used in the FMEDA, identified by O'Brien et al. [73], are based on a number of sources, including component failure rate databases, equipment failure rate databases, equipment manufacturer field return data, and cyclic test data. Since cyclic testing assumes constant dynamic operation, it is suitable only for safety functions with high demand rates (at least once a week) [74]. NAMUR [48] uses failure rate data from user experience gained with fault data management and industrial fault databases. Abdelrhafour et al. [75] provides a number of methods, one of which is

FMEDA, but where there is a lack of plant equipment failure data, generic failure rate data, such as data from Offshore and Onshore Reliability Data (OREDA) [26], are used instead. Similarly, Smith [69] suggests using Failure Rate Data In Perspective (FARADIP) [29] data for failure modes and failure rates. Valve service companies offer another potential avenue for valve failure rate data. These types of companies typically record the findings of the evaluation of a failed valve and the remedial work required to return to "as new" condition (this can be both in the field with the valve installed on the plant or in the workshop). Due to the large uncertainties of equipment failure rates and varying failure modes, precise values of proof test coverage are not possible [49]. Similarly, Isenburg [77] argues that the absence of a standardised database for mechanical component failure rates introduces subjectivity, which in turn affects the reliability of the assessment outcomes. Data sources can include generic failure rates from component and equipment databases, such as OREDA [26], SINTEF [27] and FARADIP.FOUR [29], and field data (from operators – maintenance and test data), manufacturers data (FMEDA and warranty returns) and service company data. This multi-source approach has been adopted by NAMUR [48], whereby generic PTC values are derived from a combination of sources, including manufacturer FMEDA results, user experience with fault data management, industrial fault databases, and theoretical modelling.

A comparison of equipment failure rates from the OREDA [26], SINTEF [27], FARADIP [29], and SILSafe [78] databases reveals significant variations, indicating a degree of uncertainty. The dangerous failure rates for an ESDV assembly range from approximately 200 FIT to 2500 FIT. Experts also have different opinions on the failure rates of mechanical components. The impact of this is that the proof test coverage estimates are likely to vary, which is highlighted in the next section.

2.3.4 Variation in Proof Test Coverage Estimates

A commonly used ESDV assembly proof test is the full stroke test, which is a test that is offline, i.e., not under process conditions, with a timed closure and does not include a seat leakage test. Figure 2.17 illustrates the wide range of published Proof Test Coverage (PTC) estimates for a full stroke test.

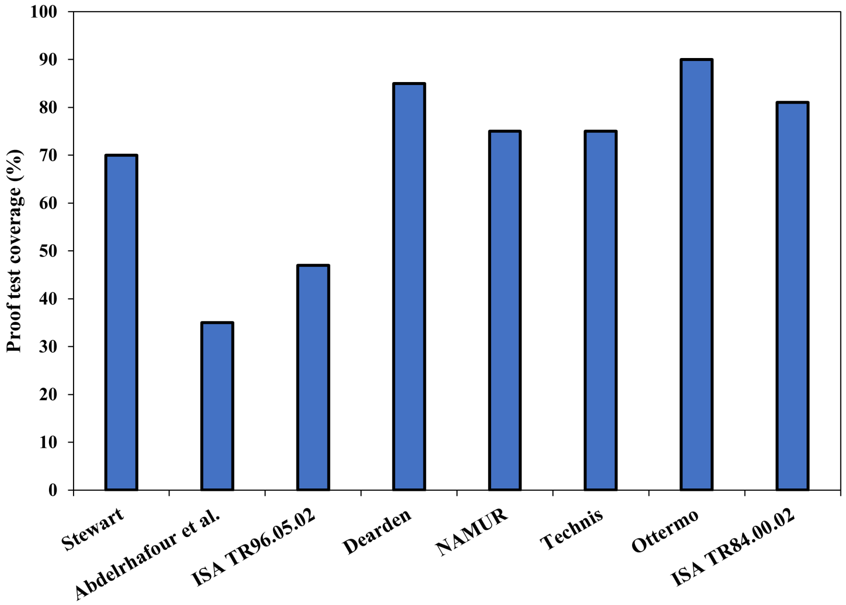


Figure 2.17: Published Full stroke test PTC estimates. Stewart [47], Abdelrhafour et al. [75], ISA TR96 [79], Dearden [49], NAMUR [48], Technis [76], Ottermo [27], ISA TR84 [80]

The PTC estimates were determined by the following methods: Abdelrhafour et al. [75], Technis [76] and Dearden [49] use qualitative credit scoring methods, Stewart [47] quantitative (FMEDA) and ISA TR84 [80], ISA TR96 [79], NAMUR [48] and Ottermo [27] use a combination of quantitative and semi-quantitative methods. It should be noted that the method presented by Abdelrhafour et al. [75] and assumes seat leakage rate testing to ANSI FCI 70-2 [81] is part of the proof test, hence the particularly low estimation of PTC. Likewise, ISA TR96 [79] assumes that the valve has a defined seat leakage rate requirement, such as tight shut-off.

In addition, the estimate presented by Ottermo [27] has been derived from the NAMUR [48] estimate but includes additional factors. The literature review confirms that there is a wide range of estimated values of PTC for the same type of proof test, i.e., from 35% to 90%. The variance if applied to $\mathrm{PFD}_{\mathrm{avg}}$ calculations produces a significant difference in the result, almost an order of magnitude, and can impact the proof test interval, which can be a concern from both a safety and an operating cost perspective. Although there is a wide variance in PTC values, there is some agreement that proof testing should not be assumed 100% [73]. Furthermore, PTC is a range from 0% to 100%, but is generally between 40% and 90% [47], the PTC value depends on the type of test being performed. An interesting observation is that some authors include in their guidelines a specified percentage of the PTC to be allocated to physical inspection of the equipment, with values ranging from 1% to 20%. This approach seems misguided in terms of in situ testing as inspection would not normally be invasive to the extent of revealing any undetected failures.

A full stroke test does not typically test the integrity of the valve seats by leakage testing; hence the majority of dangerous failures of an ESDV assembly that cannot be revealed are associated with the valve itself. The SOV and actuator operation can be tested with a high degree of confidence as quoted by O'Brien et al. [73], suggesting following PTC estimates; SOV 95%, actuator 95%, valve 55%. O'Brien et al. [73] also include PTC values for different valve types such as gate, ball, and butterfly valve.

Figure 2.18 illustrates the wide range of published PTC estimates for partial stroke testing, similar to the variability observed for full stroke testing.

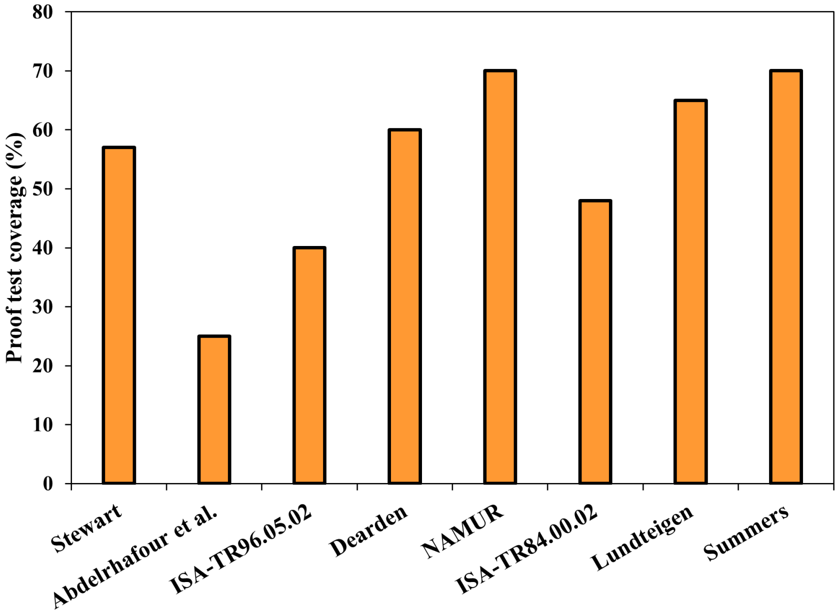


Figure 2.18: Published Partial stroke test PTC estimates. Stewart [47], Abdelrhafour et al. [75], ISA TR96 [79], Dearden [49], NAMUR [48], ISA TR84 [80], Lundteigen [67], Summers [56]

2.4 Managing Uncertainty in Data

In reliability and safety assessments, managing uncertainty is a critical concern, particularly when dealing with incomplete, imprecise, or subjective data. A range of methods has been developed to address this challenge, broadly categorised into established mathematical frameworks and more recent AI-based techniques. The following subsection first focuses on conventional methods, including probabilistic approaches, fuzzy set theory, possibility theory, and evidence theory. This is followed by a discussion of AI-based approaches that have emerged as powerful tools for handling uncertainty in complex and data-limited environments.

2.4.1 Conventional Approaches to Managing Uncertainty in Data

As highlighted in section (2.3.3) there is varying opinion on failure rate data associated with ESDVs, hence the data are subjective indicating a degree of uncertainty. More work is required to manage the uncertainty of the data, as discussed in this section.

According to Pate-́Cornell [82] there are two sources of uncertainty, *Aleatory uncertainty* or *Epistemic uncertainty.* Aleatory uncertainty arises from the natural variability of the physical world and represents the inherent randomness found in nature. It exists naturally, independent of human knowledge. Conversely, epistemic uncertainty is a systematic uncertainty that is due to a lack of knowledge and the subsequent ability to model and measure the studied system. When data are available, epistemic uncertainty can be expressed using probabilities and reduced by collecting additional information about the system being studied [83]. In contrast to aleatory uncertainty, which is inherent and cannot be reduced, epistemic uncertainty can be diminished through better understanding, improved data, and more accurate modelling techniques.

According to Yiping et al.[84] the main theories for managing data uncertainty include probability theory, fuzzy theory, info-gap theory, and derived uncertainty theory. Salahdine et al. [83] suggest that there are several methods to handle uncertainty in data and avoid imprecise decisions. These are classified into four categories: probabilistic, fuzzy set, evidence-based, and possibility-based theories. Figure 2.19 provides an overview of these theories.

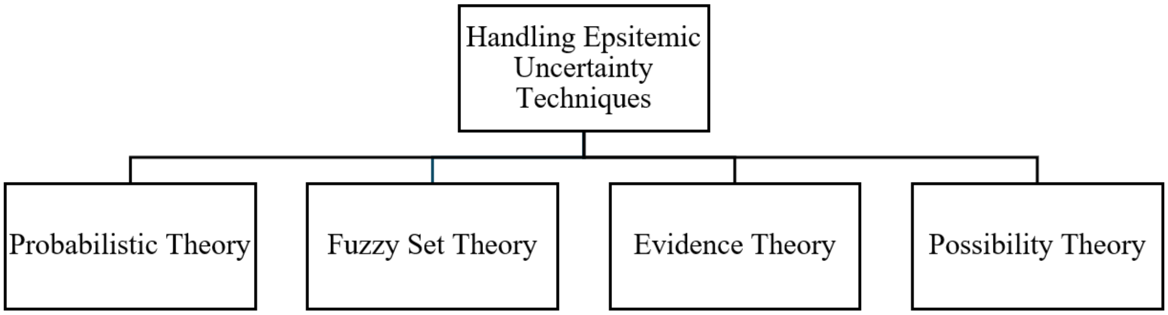


Figure 2.19: Theories for handling epistemic uncertainty [83]

Yiping et al.[84] state that probability theory is the most widely used method in almost every field. Monte Carlo and Bayesian methods are two common probability theory approaches for addressing epistemic uncertainty, which arises from limited knowledge or data in a system.

Monte Carlo (MC) methods refer to a broad category of stochastic simulation techniques widely applied to solve optimisation and inference problems in science and engineering. MC methods operate by generating a large set of potential parameter values, allowing integrals to be approximated by sampling averages. In practice, these parameter values are obtained by physically replicating the experiment or by probabilistically modelling it and generating random samples [85].

The Bayesian method is a data analysis approach grounded in the Bayes theorem. In which existing knowledge about parameters is updated within a statistical model using observed data. This prior knowledge is represented as a prior distribution and, when combined with observational data through a likelihood function, produces the posterior distribution. The posterior distribution can then be used to make predictions about future events [86].

Fuzzy set theory, introduced by Lotfi Zadeh in 1965, extends classical set theory to handle uncertainty and vagueness in data by allowing partial membership in sets [87]. Unlike traditional sets where an element belongs or does not belong, i.e., its binary (0 or 1), fuzzy sets assign a degree of membership to each element, ranging from 0 (no membership) to 1 (full membership) [89]. This flexibility makes fuzzy set theory especially useful in situations where boundaries are not sharply defined, such as linguistic or subjective assessments. Widely applied in fields such as control systems, decision making, and data analysis, fuzzy set theory provides a structured way to incorporate imprecise information into mathematical models.

Evidence theory, also called theory of belief functions or Dempster-Shafer Theory (DST), is a mathematical theory of evidence. Developed by Dempster [90] and further expanded by Shafer [91], the Dempster-Shafer theory can be interpreted as a generalisation of probability theory where probabilities are assigned to sets as opposed to mutually exclusive singletons. In traditional probability theory, evidence is associated with only one possible event. In DST, evidence can be associated with multiple possible events, for example, sets of events. As a result, evidence in DST can be meaningful at a higher level of abstraction without having to resort to assumptions about events within the evidential set [92].

Possibility theory is a method to mitigate uncertainty and incomplete or imprecise data in multi-source information [93]. Possibilistic methods assess the degree of feasibility (or possibility) of events rather than their probability, and they are often applied when dealing with imprecise or incomplete information. Possibilistic logic incorporates certainty and possibility degrees that are not compositional for all logical connectives and are applied to classical formulas, which means that they involve only clear, precise and well-defined statements, without ambiguity [94]. Possibilistic logic traditionally applies to classical logic statements that are either true or false. Then, the possibilities and the certainty degrees are assigned to these definitive statements to reflect confidence levels, but the statements themselves remain clear and unambiguous in their content. This logic framework is used to manage uncertainty by attaching these degrees to statements, providing a way to handle partial belief or feasibility in a structured manner.

Fuzzy set theory is good at handling human ambiguity by modelling epistemic uncertainty through fuzzy sets with membership functions. Salahdine et al. [83] suggest that fuzzy theory is easy to implement and interpret.

Fuzzy set theory when used in a Fuzzy Inference System (FIS) offers a distinct approach to handling uncertainty compared to other methods such as Monte Carlo, Bayesian, and Dempster-Shafer methods. FIS is particularly effective when dealing with subjective and vague expert input [95, 96], as it translates linguistic terms into numerical values through fuzzy sets. Unlike probabilistic approaches that require precise probabilities or distributions [97], FIS can manage qualitative assessments, which are often easier to obtain in practice and more reflective of real-world expert knowledge.

An FIS tends to be more intuitive and interpretable than probabilistic methods [98]. The IF-THEN rule structure in FIS aligns closely with human reasoning, making it easier to explain the decision-making process to stakeholders. In contrast, Monte Carlo and Bayesian methods may involve complex statistical models and computations that can make results more difficult to interpret.

Although Bayesian and Dempster-Shafer approaches rely on probability distributions or belief functions [99], FIS uses degrees of membership, which do not require precise probabilities. This flexibility is advantageous when data are limited or uncertain, allowing FIS to provide reliable outputs even with minimal or ambiguous input data. Kulkarni and Johnson [100] propose a Dempster-Shafer theory approach to FMEA to aggregate responses from multiple sources, team members, and experts. Combining responses overcomes the subjective issue of assigning weights to different team members. However, the implementation of DST to FMEA highlighted some issues using the Dempster rule of combination.

An FIS is computationally less intensive than Monte Carlo simulations, which require numerous iterations, or Bayesian analysis, which often requires advanced algorithms to estimate posterior distributions [101]. For applications requiring real-time decision making or when computational resources are limited, an FIS can provide faster results with fewer computational demands.

An FIS does not strictly require data distributions and can work effectively with imprecise or limited data. Probabilistic methods such as Bayesian and Monte Carlo often assume well-defined data distributions [101], which may not always be available or reliable. In comparison, FIS adapts to the variability within the input data using fuzzy sets, providing robustness regardless of data limitations.

Although an FIS offers a powerful means of capturing expert knowledge and managing uncertainty, it is not without limitations. One key limitation lies in the subjectivity involved in defining membership functions and rule sets, which may vary between experts and introduce inconsistency [102]. Furthermore, as the number of input variables increases, the IF-THEN rule base can grow exponentially, making the system increasingly complex and difficult to manage. Traditional FIS models also lack adaptive learning capabilities and do not automatically update based on new data unless integrated with machine learning techniques [98]. Despite these limitations, the FIS remains a highly valuable tool in scenarios where data are sparse or uncertain, and expert judgement plays a critical role.

In summary, a fuzzy inference system is especially valuable when dealing with ambiguous uncertainty, meaning uncertainty that does not stem from random variability but rather from vagueness, imprecision, or incomplete knowledge. This makes a fuzzy inference system ideal for applications where practical decision making relies on expertise or intuition rather than purely statistical or probabilistic data.

2.4.2 AI Techniques for Managing Uncertainty in Data

Artificial Intelligence (AI) has become increasingly important in reliability engineering, predictive modelling, and decision support, particularly in situations involving incomplete, uncertain, or noisy data. Given its growing role across these domains, it is now almost essential to review the range of AI techniques available, particularly those capable of handling uncertainty in engineering decision making. Several AI methods have emerged as particularly relevant for such applications, each with unique strengths and limitations.

Artificial Neural Networks (ANNs) are widely used for capturing non-linear relationships in complex datasets. They are especially effective in pattern recognition and predictive modelling. Despite being powerful learners, ANNs are often criticised for their "black-box" nature and lack of transparency. Zhang et al. [103] provided a comprehensive review of their forecasting capacity, noting that while ANNs can outperform traditional models, they require careful tuning and large datasets to be effective.

To improve the stability and accuracy of predictions, ensemble learning methods such as Bagging and Boosting have been developed. Bagging reduces variance by training multiple models on bootstrapped data subsets and averaging the results. Boosting reduces bias by sequentially focussing on errors made by previous models. Dietterich [104] highlighted the robustness of ensemble methods, particularly in scenarios with noisy data.

Reinforcement Learning (RL) is a dynamic learning approach in which an agent interacts with an environment to learn optimal actions through trial and error, guided by a reward signal. Kaelbling et al. [105] provided a foundational survey of RL, highlighting its potential for sequential decision making in uncertain and adaptive environments.

In contrast to data-intense models, Gray System Theory (GST) is designed for systems with small samples and poor or incomplete information. Introduced by Deng [106], GST uses differential equations to forecast trends with minimal data. Lin and Liu [107] suggest its utility in forecasting and decision analysis, particularly where statistical assumptions of traditional models cannot be satisfied.

Finally, Digital Twin (DT) modelling integrates real-time data from physical systems with virtual models to simulate, predict, and optimise system behaviour. This approach has gained traction in smart manufacturing and asset management. Tao et al. [108] reviewed the state-of-the-art in industrial digital twins, identifying their value in predictive maintenance, operational efficiency, and decision support under uncertainty.

Together, these AI methods form a toolkit for addressing various dimensions of uncertainty in engineering contexts, ranging from real-time system simulation to data-driven prediction and decision making under incomplete information.

The next section provides a more detailed discussion of fuzzy inference systems.

2.5 A Comprehensive Overview of Fuzzy Inference Systems

Fuzzy inference systems can manage and process imprecise, vague, or ambiguous data effectively. This aspect is fundamental in real-world applications where data might be imprecise or subjective.

Fuzzy inference is the process of formulating the mapping from a given input to an output using a set of rules. The mapping then serves as a foundation for making decisions or identifying patterns. The two main types of fuzzy inference systems are: Mamdani [109] and Sugeno [110], also known as the Takagi-Sugeno-Kang (TSK) fuzzy inference system. These types of inference systems vary in the way their outputs are determined.

With Mamdani-type inference, the output membership functions are fuzzy sets. Following the aggregation process, there is a fuzzy set for each output variable, which requires defuzzification. The Sugeno method of fuzzy inference is similar to the Mamdani method; for example, the first two parts of the fuzzy inference process, fuzzifying the inputs and applying the fuzzy operator, are exactly the same. The main difference between Mamdani and Sugeno fuzzy inferences is that the output membership functions are only linear or constant for Sugeno-type fuzzy inference [111].

The Mamdani rule-based fuzzy inference system is the most commonly used system and has two types, Type-1 and Type-2. A traditional Type-1 membership function has a single membership value or any value in the universe of discourse. Consequently, although a Type-1 membership function represents the degree of membership within a specific linguistic set, it does not account for the uncertainty in that degree of membership. To model this uncertainty, Type-2 [112] interval membership functions can be used. In these Type-2 membership functions, the degree of membership is represented by a range of values. The most commonly used out of the two types is the Mamdani Type-1 inference system due to its simplicity of application.

A fuzzy set consists of a universe of discourse and a membership function that maps every element in the universe of discourse to a membership value between 0 and 1. Formally defined, suppose X is the domain or universe of discourse, and $x \in X$ is a specific element of the domain X, then, the fuzzy set A is characterised by a membership mapping function:

$$\mu_{\mathrm{A}}(x) : X \rightarrow [0.1] \tag{2.3}$$

Therefore, for all x$\in X$, $\mu A(x)$ indicates the certainty with which element x belongs to fuzzy set A.

The membership function graphically represents the degree of participation for each input. The membership function is used to associate the degree of membership of each of the elements in the domain of the corresponding fuzzy set. Figure 2.20 illustrates the various components of a fuzzy set.

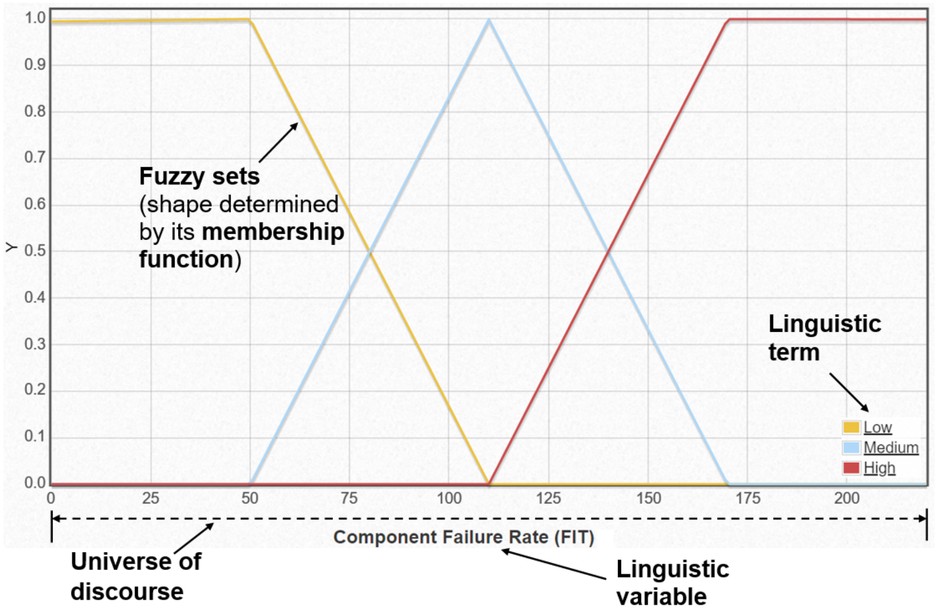


Figure 2.20: Fuzzy set components

Membership functions for fuzzy sets can be of any shape; the most commonly used are triangular, trapezoidal, and Gaussian [113]. The choice of membership function can be guided by the characteristics of the data and the desired behaviour the Fuzzy Inference System (FIS) is expected to display. The triangular membership function is considered the most simplistic and computationally efficient and is used in systems where data is limited. The trapezoidal membership function provides more flexibility than the triangular one when there is a need to capture a wider range of inputs as having the same membership grade, i.e., where a range of values can be considered fully belonging to a set. The Gaussian membership function is applied when the data follows a normal distribution or when smooth, gradual transitions between fuzzy sets are anticipated. In practice, a combination of these membership functions can be used to achieve a balance between effectiveness and smoothness, depending on the specific requirements of the FIS.
The graphical representation of the most common shapes, as described by Sabri et al. [114], is presented as follows.

Triangular function, defined as:

$$\mu_A(x) = \begin{cases} 0 & x < \alpha_{min} \\ \dfrac{x - \alpha_{min}}{\beta - \alpha_{min}} & x \in (\alpha_{min}, \beta) \\ \dfrac{\alpha_{max} - x}{\alpha_{max} - \beta} & x \in (\beta, \alpha_{max}) \\ 0 & x > \alpha_{max} \end{cases} \tag{2.4}$$

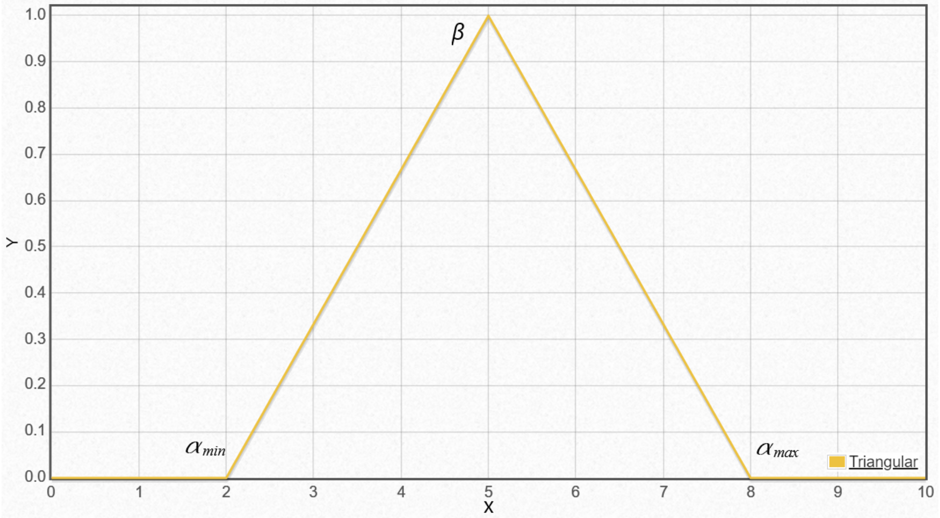


Figure 2.21: Triangular Membership Function

Trapezoidal function, defined as:

$$\mu_A(x) = \begin{cases} 0 & x \leq \alpha_{min} \\ \dfrac{x - \alpha_{min}}{\beta_1 - \alpha_{min}} & x \in (\alpha_{min}, \beta_1) \\ \dfrac{\alpha_{max} - x}{\alpha_{max} - \beta_2} & x \in (\beta_2, \alpha_{max}) \\ 0 & x \geq \alpha_{max} \end{cases} \tag{2.5}$$

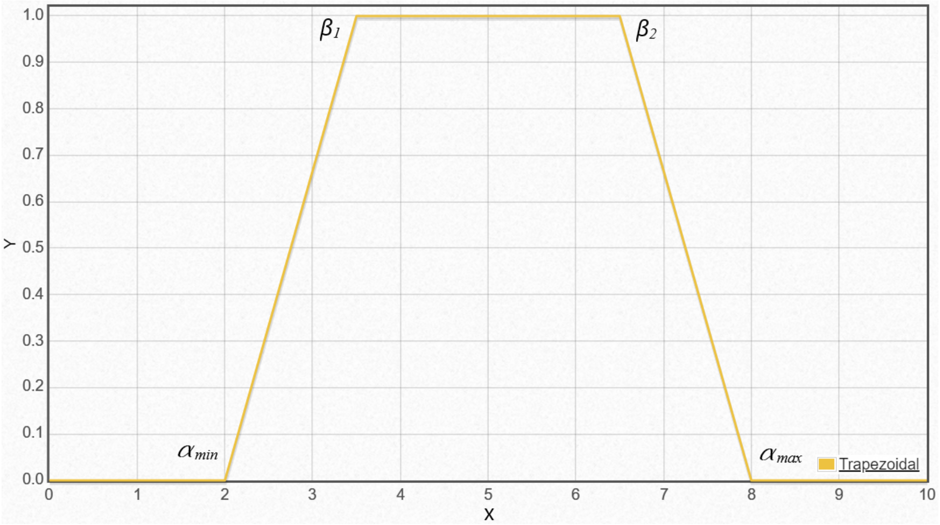


Figure 2.22: Trapezoidal Membership Function

Gaussian function, defined as:

$$\mu_A(x) = e^{-\alpha(x-\beta)^2} \tag{2.6}$$

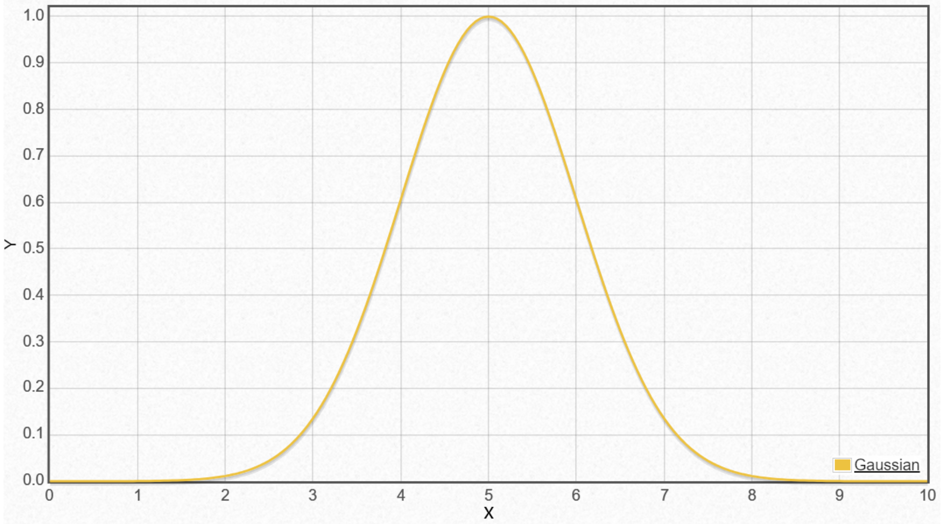


Figure 2.23: Gaussian Membership Function

2.5.1 The Mamdani Fuzzy Inference System

Figure 2.24) shows the main elements of a Mamdani (Type-1) Fuzzy Inference System (FIS) as used in this research. In this case, the system inputs are the failure rates of the ESDV component and the probability of revealing a failure during a proof test. The output is the weighted revealed failure rate which is a consequence of the two inputs.

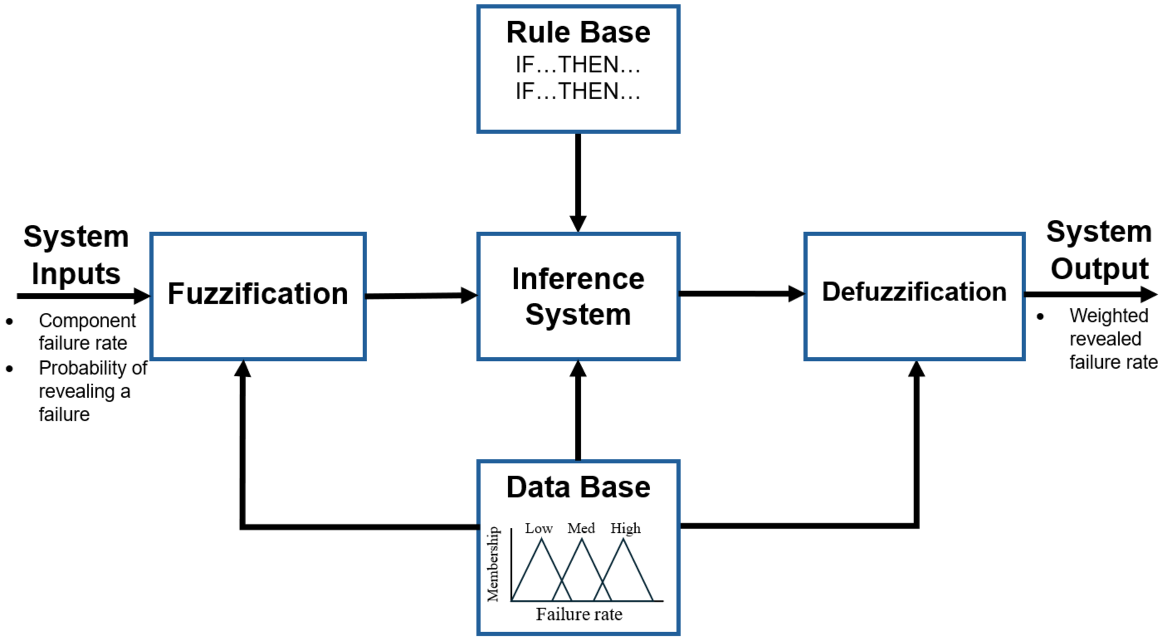


Figure 2.24: Mamadani FIS

The crisp inputs are converted to linguistic terms using membership functions, a process called fuzzification. Figure 2.25 shows an example of a fuzzy set (triangular type). This fuzzy set represents the input of the component failure rate and is divided into five membership functions representing the universe of discourse (0 to 220 FIT) using linguistic terms from Very Low (failure rate) to Very High (failure rate). A Failure in Time (FIT) is defined as one failure in one billion operating hours. The degree of membership in a fuzzy set is indicated by the y axis from 0 to 1, where 0 suggests that there is no membership in a fuzzy set and 1 is membership 100%. For example, if the component failure rate input is 50 FIT then it is both a member of the *Very Low* and *Low* fuzzy sets, approximately 0.6 and 0.35, respectively.

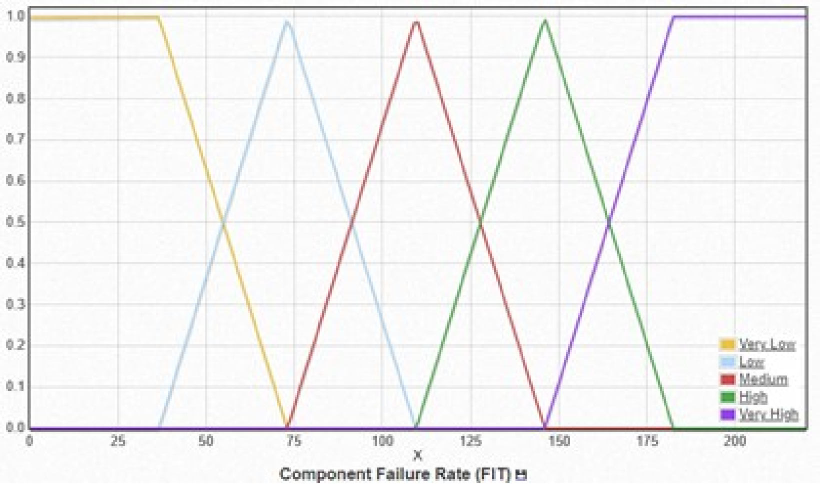


Figure 2.25: Triangular membership function

As discussed, membership functions can differ in shape depending on the application, with the most common types being triangular, trapezoidal, and Gaussian. The number of membership functions can also vary, which is also application-specific. Higher accuracy is obtained by using a greater number of membership functions.

Following the fuzzification process, the linguistic terms representing the input are subjected to an IF-THEN rule base. The rules are based on application-specific input/output criteria and are designed to activate outputs, which are also linguistic variables. Taking the universe of discourse for the probability of revealing the failure is in the range 0.5 to 1. Multiple rules can be activated with a single set of inputs. An example rule is given:

IF the **Component Failure Rate** is *Low* AND the **Probability of Revealing the Failure** is *High* THEN the **Weighted Revealed Failure Rate** is *Low*

Two processes that take place within the inference system: Implication and Aggregation [20]. Implication is the process of combining two or more inputs, also known as Antecedents. Implication is dependent on the operator used in the rules, for example if an AND operator is used as in the example above then the MINIMUM value (degree of membership) of the two inputs is taken forward to create a Consequent, the output. Subsequently, if an OR operator is used, then the MAXIMUM value is taken forward. The example in the following, Figure 2.26, shows the AND operator based on component failure rate = 50 FIT and probability of revealing the failure = 1. Two rules are activated. The minimum degree of membership of the two inputs for each rule, shown with a green dotted line from the component failure rate input to the output (the weighted revealed failure rate).

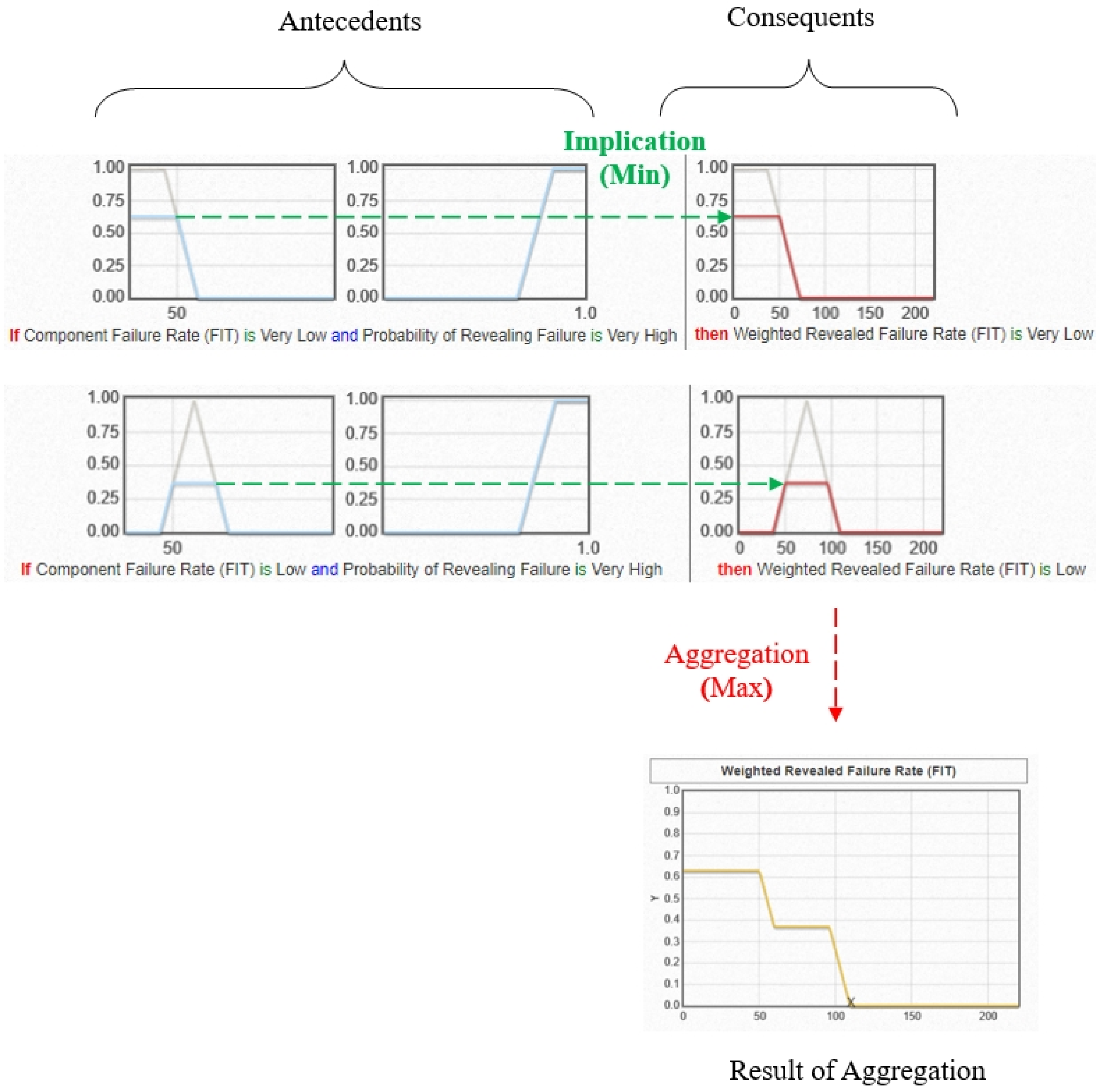

Antecedents
Consequents
Implication
(Min)
1.00
0.75
0.50
0.25
0.00
50
1.0
0
50
100
150
200
If Component Failure Rate (FIT) is Very Low and Probability of Revealing Failure is Very High
then Weighted Revealed Failure Rate (FIT) is Very Low
If Component Failure Rate (FIT) is Low and Probability of Revealing Failure is Very High
then Weighted Revealed Failure Rate (FIT) is Low
Aggregation
(Max)
Weighted Revealed Failure Rate (FIT)
Y
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
X
Result of Aggregation

Figure 2.26: Demonstration of two activated rules, showing how inputs are selected, implications are applied, and outputs are aggregated in the fuzzy inference process.

Aggregation is the process of merging the outputs of the two or more rules to form a common fuzzy output, using an OR operator, which in this case is the resultant Weighted revealed failure rate. This is the most common type of Mamdani inference system, also known as the Min-Max.

Defuzzification is the final step in the process through which the linguistic output of the rules is converted back into a crisp value. The most common method is by calculating the centroid (centre of area or centre of gravity) of the resultant geometric shape from aggregation. Other methods of defuzzification are available such as max-membership, weighted average, and mean-max membership. An example of centroid defuzzification is shown in Figure 2.27. The centroid is shown by the red dot and the arrow indicating the crisp output value, in this case 44 FIT.

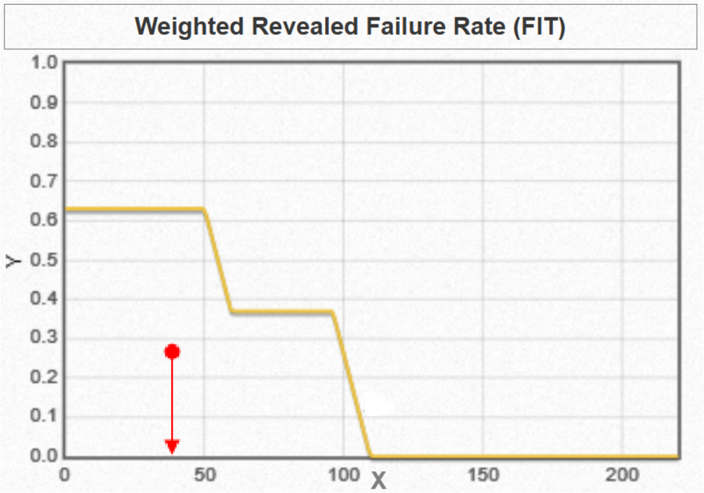


Figure 2.27: Centroid defuzzification - the red arrow showing the crisp output value

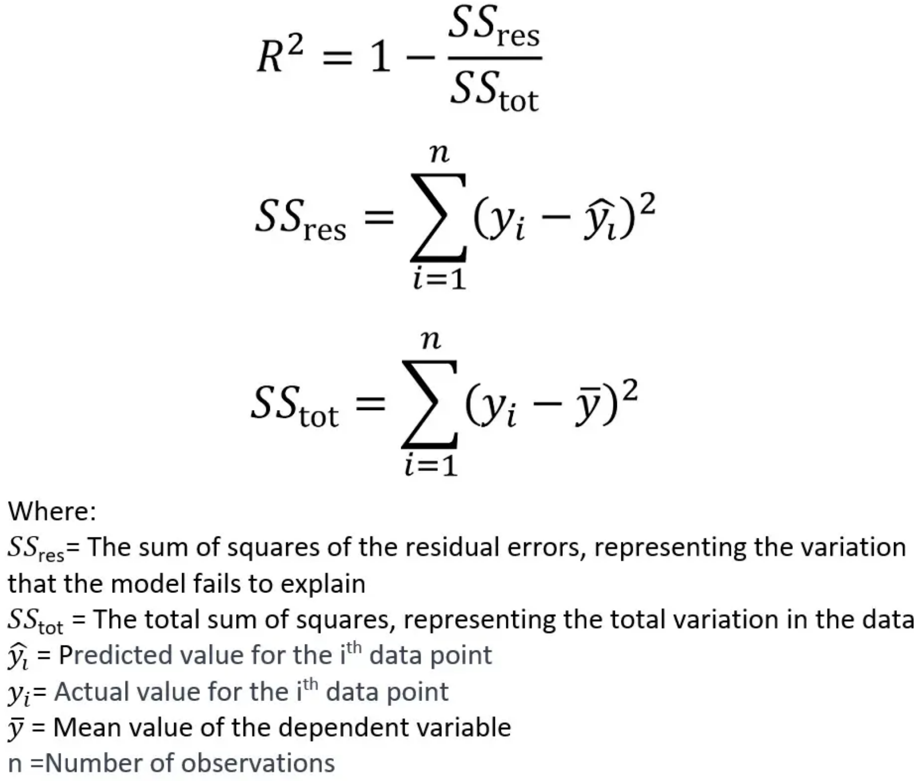
The next section discusses the performance measures and metrics used to evaluate the FIS design.

2.6 Performance Measures and Metrics used for Model Evaluation

Safavi et al. suggest that the performance of Fuzzy Inference Systems (FIS) designs should be evaluated against the coefficient of determination (R^2), Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) [115]. However, Mean Squared Error (MSE) is often preferred over Mean Absolute Error (MAE) in modelling FIS because MSE penalises larger errors more heavily due to the squaring of differences. This sensitivity can be advantageous in FIS modelling, where the goal is often to minimise significant deviations and improve model precision. The emphasis of MSE on larger errors aids in optimising fuzzy rule parameters, particularly when outliers or significant deviations are critical in the application context. According to Guillaume [116], FIS designs are typically evaluated and compared based on their Mean Square Error (MSE) performance. As proposed by Ye and Wei [117] when modelling an FIS, MSE serves as a measure of the average squared difference between the observed and predicted values, while R^2 assesses the proportion of variance in the dependent variable that was predictable from the independent variables.

2.6.1 The Coefficient of Determination (R-squared or R^2)

R^2 is a statistical measure that represents the proportion of the variance in the dependent variable that is predictable from the independent variables in a regression model. It provides an indication of the goodness-of-fit of the model, quantifying how well the independent variables explain the variability of the dependent variable [118]. The R^2 equation is given by [119]:



$$R^2 = 1 - \frac{SS_{\text{res}}}{SS_{\text{tot}}}$$

$$SS_{\text{res}} = \sum_{i=1}^{n} (y_i - \hat{y_i})^2$$

$$SS_{\text{tot}} = \sum_{i=1}^{n} (y_i - \bar{y})^2$$

Where:
SS_{res}= The sum of squares of the residual errors, representing the variation that the model fails to explain
SS_{tot} = The total sum of squares, representing the total variation in the data
$\hat{y_i}$ = Predicted value for the i$^{\text{th}}$ data point
y_i= Actual value for the i$^{\text{th}}$ data point
$\bar{y}$ = Mean value of the dependent variable
n =Number of observations

(2.7)

Interpretation of Values:

- Low R-squared (close to 0): Suggests that the model does not explain much of the variability of the response data around its mean. This might indicate that the chosen predictors do not have a significant linear relationship with the response variable or that important predictors are missing.
- High R-squared (close to 1): The model suggests that it explains a large part of the variability in the response data around its mean. This is typically a sign that the predictors chosen have a strong linear relationship with the response variable.

2.6.2 Mean Square Error (MSE)

Mean Squared Error (MSE) measures the squared differences between actual and predicted values, indicating how closely the line of best fit aligns with a data set. MSE is always positive because the squaring of the differences removes negative signs, providing a clear measure of prediction accuracy [118]. The MSE equation is given by [119]:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(\hat{y}_{i} - y_{i})^{2} \tag{2.8}$$

Where:
$\hat{y}_{i}$ = Predicted value for the i^{th} data point
y_i= Actual value for the i^{th} data point
n = number of observations

Interpretation of values:

- Low MSE: Indicates that the model's predictions are close to the actual values, reflecting a high accuracy.
- High MSE: Indicates that the model's predictions deviate significantly from the actual values, suggesting lower accuracy and potential room for improvement.

2.6.3 Root Mean Squared Error (RMSE)

The Root Mean Squared Error (RMSE) is a commonly used metric to evaluate the accuracy of a model in predicting quantitative data. Measures the average magnitude of errors between predicted and actual values, giving insight into how well a model performs. RMSE provides a measure of the average magnitude of errors in the same units as the target variable. The RMSE equation is given by [119]:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_{i} - y_{i})^{2}} \tag{2.9}$$

Where:
$\hat{y}_{i}$ = Predicted value for the i^{th} data point
y_i= Actual value for the i^{th} data point
n = number of observations

Interpretation of values:

- Low RMSE: Indicates that the model's predictions are close to the actual values, reflecting a high accuracy.

- High RMSE: Indicates that the model's predictions deviate significantly from the actual values, suggesting lower accuracy and potential room for improvement.

Lower RMSE values indicate better model performance, but the acceptability of an RMSE value, as with MSE, must be judged in the context of the data and the specific application.

Although R^2 can be a useful initial measure to evaluate the performance of an FIS, it should be accompanied by other metrics to gain a complete understanding of the performance of the system. A multifaceted approach facilitates a better understanding of both the relationship between the predicted outputs (FIS) and the actual outputs (FMEDA), as well as the accuracy of the predictions.

In addition to accuracy and performance metrics, recent literature highlights the importance of evaluating fuzzy systems based on their uncertainty, interpretability, and robustness. Uncertainty analysis focusses on how well the fuzzy model handles imprecise, vague, or incomplete input data, an inherent feature of many expert-based systems. Techniques such as Type-2 fuzzy systems [120] and Monte Carlo simulations [121] have been used to model and assess uncertainty in fuzzy logic applications. Possibility theory has also been applied to account for imprecision in knowledge-based systems [122].

Interpretability remains a critical attribute, particularly when fuzzy systems are intended to support expert judgement. Gacto et al. [123] provide a comprehensive overview of interpretability measures for linguistic fuzzy rule-based systems, categorising them based on complexity, e.g., number of rules, rule length, and semantic aspects such as how logically aligned and distinct the fuzzy sets are. Their framework is widely used as a benchmark for evaluating and improving the interpretability of fuzzy models.

Evaluating the robustness and sensitivity of fuzzy systems is essential to ensure reliable performance, particularly in environments with high levels of uncertainty or variability in input data. Robustness refers to a system's ability to maintain acceptable performance under disturbances or parameter changes, while sensitivity analysis examines how input variations influence system output behaviour.

Robustness in fuzzy reasoning refers to the stability of inference outcomes when subject to small perturbations in input data or membership functions. Li et al. [124] proposed a systematic approach to quantify this robustness by measuring how variations in fuzzy sets influence reasoning results. Their method provides a robustness index based on the logical equivalence between original and perturbed reasoning outcomes, offering a structured way to assess the resilience of fuzzy systems in uncertain environments.

In addition to robustness analysis, sensitivity analysis provides insight into the influence of individual inputs on the output of a fuzzy model. Javidan et al. [125] conducted a fuzzy sensitivity analysis within a structural performance context, applying a local sensitivity framework based on fuzzy random theory. Their study showed how changes in uncertain input parameters could be systematically linked to variations in performance metrics, using visual tools and numerical metrics to evaluate the impact. This approach helps identify critical parameters and improve model transparency and reliability.

2.7 Gap Analysis and Aims of the Work

The differing opinions regarding ESDV proof test coverage estimates create a risk of incorrectly calculating ESDV proof test intervals. This could result in ESDV proof testing being conducted too frequently, leading to higher costs for the operating company, or too infrequently, creating a potential safety risk if the ESDV fails to operate when required. Current approaches to estimating proof test coverage include quantitative methods, such as Failure Modes, Effects, and Diagnostics Analysis (FMEDA), as well as semi-quantitative and qualitative techniques, including credit scoring systems. Although FMEDA is considered the most effective method, it still involves some subjectivity, particularly in relation to component failure rate data. Given the importance of this aspect of functional safety, the objective is to develop a model that can estimate ESDV proof test coverage using the FMEDA technique while accounting for the subjectivity of the data.

There is a consensus that effective proof testing and quantification of accurate PTC estimates are critical to achieving functional safety [25, 47, 71, 72]. It has been demonstrated that the effect of PTC on PFD_{avg} calculations is significant, and thus utilising an inappropriate PTC value can have a considerable effect on the end result. Proof tests must be designed to be as effective as possible, ensuring that the maximum number of dangerous failures is revealed. The effectiveness of a proof test appears to be governed by the design, installation, and operational budget; elaborate test instrumentation and extensive testing required to achieve high levels of effectiveness come at a cost. The literature review also indicates that guidance on determining PTC is limited, with only a small number of sources offering estimates for PTC.

From the review it is apparent that the current published values of PTC vary, and the degree of variance when incorporated into SIF PFD_{avg} calculations is significant, i.e., almost an order of magnitude different between one end of the spectrum to the other. It is also evident that manufacturers' equipment safety manuals can be an unreliable source of PTC. Furthermore, the expectation that the operator should conduct an analysis, such as an FMEDA, to determine PTC appears inappropriate, given the level of expertise required to carry out such an assessment. With these points in mind, it is clear that more research is required to develop a model that can be used to determine accurate estimates of PTC.

A fuzzy inference system may be valuable when dealing with ambiguous uncertainty, which means uncertainty that does not stem from random variability, but rather from vagueness, imprecision, or incomplete knowledge. This makes a fuzzy inference system ideal for applications where practical decision making relies on expertise or intuition rather than purely statistical or probabilistic data.

As emphasised in this chapter, ESDVs are essential components of safety-critical systems, which require periodic proof testing to reveal hidden failures and support condition-based maintenance. The intervals for conducting tests are influenced by various factors, including Proof Test Coverage (PTC), which has been shown to vary in the published literature. FMEDA was identified as the preferred method for estimating PTC; however, the component failure rate data used in the analysis appear to be subjective. Methods for managing data uncertainty include various methods including Monte Carlo, Bayesian, Dempster-Shafer, and fuzzy inference systems. Fuzzy inference systems were shown to be a suitable method for this application and have been applied to similar techniques such as FMEA and FMECA previously.

FMEDA often involves significant subjectivity, particularly in the estimation of component failure rates and the probability of revealing a failure due to gaps in empirical data. Although modern AI techniques such as neural networks and ensemble methods are effective for data-driven modelling, they typically require large datasets and offer limited transparency, both of which are problematic in safety-critical applications. Fuzzy Inference Systems (FIS), in contrast, are specifically designed to handle imprecise, qualitative input using linguistic rules, making them particularly well suited to the FMEDA context. FIS allows expert judgement to be formalised in a mathematically structured but interpretable way, maintaining traceability of assumptions while explicitly modelling the uncertainty and vagueness inherent in expert-driven assessments. Moreover, its rule-based structure naturally aligns with the logical framework already employed in FMEDA, offering an effective bridge between qualitative reasoning and quantitative estimation.

With these knowledge gaps identified, this thesis aims to estimate ESDV proof test coverage through multiple methods, focussing on FMEDA and using a fuzzy inference system to manage data uncertainty.

The following chapter describes the methodology for this research, including the development of fuzzy inference system design criteria.

Chapter 3

Methodology

This chapter presents the methodology used to estimate Proof Test Coverage (PTC), detailing the underlying principles, data sources, and analytical techniques applied throughout the study. Three distinct methods were used to estimate PTC for emergency shutdown valves. Each method differed in its approach to data, level of subjectivity, and underlying assumptions. Together, they demonstrate the progression from traditional approaches toward a more structured method for dealing with uncertainty and expert judgement. Figure 3.1 provides an overview of the structure of this chapter.

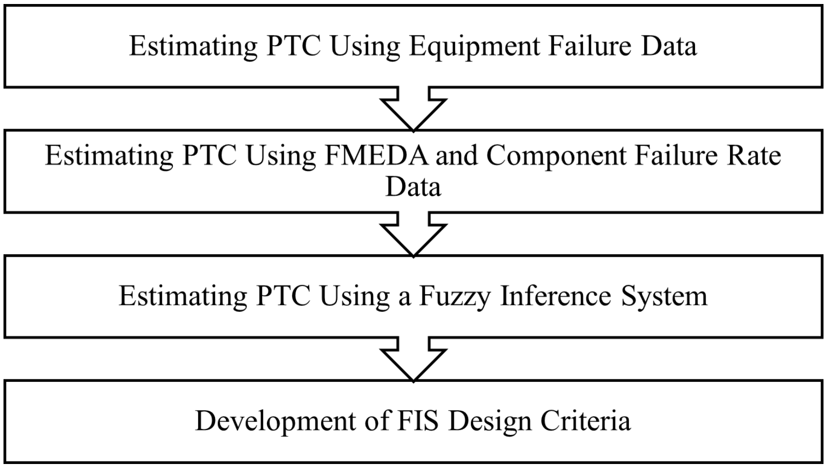

Estimating PTC Using Equipment Failure Data
Estimating PTC Using FMEDA and Component Failure Rate Data
Estimating PTC Using a Fuzzy Inference System
Development of FIS Design Criteria

Figure 3.1: Chapter Three Structure

3.1 Estimating PTC Using Equipment Failure Data

The first method used high-level equipment failure rate data collected from industry sources. This approach reflects conventional practice, whereby aggregated failure data for the complete final element (e.g., valve assembly) are used to estimate PTC. The data were obtained from operational records from real industrial installations. PTC was calculated by comparing the number of failures detected during maintenance and proof testing with the total dangerous failures recorded over the operational period. While this approach is based on real-world data, it lacks detail and may hide how specific components or testing limitations affect the results.

3.2 Estimating PTC Using FMEDA and Component Failure Rate Data

The second method applied a component-level Failure Modes, Effects, and Diagnostic Analysis (FMEDA) using detailed failure mode and failure rate data. Data was sourced from exida's Component Reliability Data (CRD), which provides failure data for a wide range of components used in ESDVs. This data was compared with expert-derived FMEDA figures, highlighting the variability and subjectivity inherent in FMEDA-based estimates. The comparison underscored the potential inconsistencies that can arise from relying solely on expert judgement, even when structured within an FMEDA framework. The Proof Test Coverage (PTC) was calculated by summing the revealed dangerous undetected failure rates and dividing by the total dangerous undetected failure rate, allowing for a granular and quantitative estimate. This method, while more precise, depends heavily on the availability of high-quality component-level data and consistent interpretation by analysts.

3.3 Estimating PTC Using a Fuzzy Inference System

The third method introduced a fuzzy logic approach to address the uncertainty and subjectivity encountered in the FMEDA process. A Mamdani-type Fuzzy Inference System (FIS) was developed using JuzzyOnline and MATLAB using two input variables: component failure rate and the probability of revealing the failure during the proof test. An initial test set was created using randomised input combinations to examine the behaviour of the system across the full input space. The FIS output was then validated against expert-derived failure rates to assess its accuracy. Regression analysis was used to compare the predictions of the fuzzy model with expert estimates, focussing on metrics such as R^2, MSE and RMSE. This approach allowed subjective assessments to be modelled in a structured, transparent, and repeatable manner, thereby offering a potential enhancement to traditional FMEDA methods.

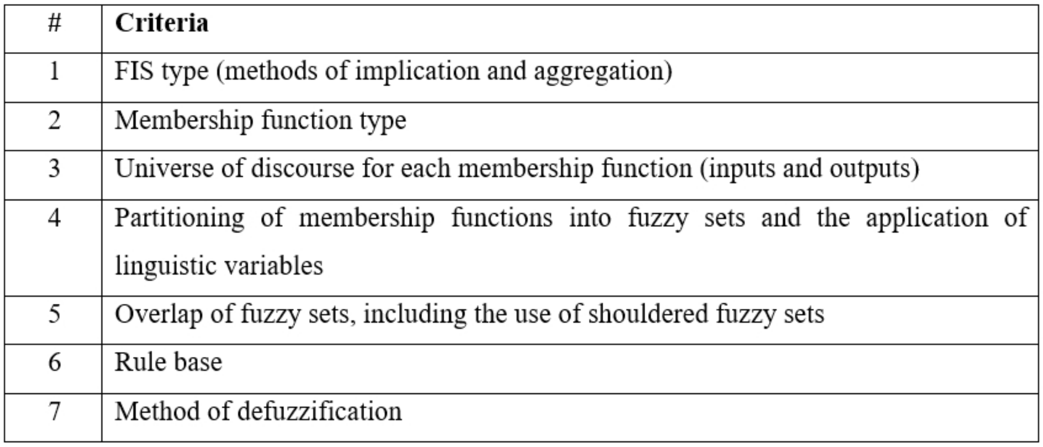
The following section describes the development of the fuzzy inference system design criteria.

3.4 Development of Fuzzy Inference System Design Criteria

Given the novelty of the Fuzzy-FMEDA approach, no published research was identified that applies a Fuzzy Inference System (FIS) to FMEDA. Therefore, the literature review concentrated on related techniques, including FMEA and FMECA (terms often used interchangeably in the literature), where fuzzy logic has been more widely explored. A FMECA traditionally has three inputs (Severity, Occurrence, and Detection), whereas FMEDA has two (Component failure rate and Probability of revealing the failure). There are similarities between the inputs of FMECA and FMEDA, for example, Occurrence is the equivalent of Component failure rate and Detection the equivalent of Probability of revealing the failure. Research papers relating to the application of fuzzy inference systems to Fault Tree Analysis (FTA) were also reviewed.

Research papers were reviewed to identify the most commonly used design criteria for an FIS applied to an FMEA type assessment. The identified design criteria were used to form the basis for the design of the FMEDA fuzzy inference system. Table 3.1 outlines the key design criteria required for an FIS, which are elaborated on in the discussion that follows.

Table 3.1: FIS design criteria



#	**Criteria**
1	FIS type (methods of implication and aggregation)
2	Membership function type
3	Universe of discourse for each membership function (inputs and outputs)
4	Partitioning of membership functions into fuzzy sets and the application of linguistic variables
5	Overlap of fuzzy sets, including the use of shouldered fuzzy sets
6	Rule base
7	Method of defuzzification

3.4.1 FIS Type

The most commonly employed fuzzy inference system for FMEA applications is the Mamdani (Type-1) rule-based type, which is evident in all reviewed research papers.

The Mamdani FIS is a tool for logical deductive inference used to analyse the outcomes of a model structure in terms of a set of IF-THEN rules [126]. The IF-THEN rules use AND and/or OR operators which can be used to combine the input levels expressed in terms of their linguistic terms, known as implication. The AND operator selects the minimum degree of membership of the inputs, and the OR operator the maximum. Where multiple rules are activated, aggregation of rules are also realised using the AND or OR operators. In this fuzzy inference design the AND operator is utilised for combining the linguistic terms of rules and the OR operator for aggregation of multiple rules. The type of FIS used in this study is known as a Mamdani Min-Max inference system.

3.4.2 Membership Function Type

Membership functions are a means of converting crisp inputs into fuzzy (linguistic) terms. Common types of membership functions used in fuzzy inference systems include triangular, trapezoidal, and Gaussian, as illustrated in Figure 3.2, Figure 3.3 and Figure 3.4, respectively.

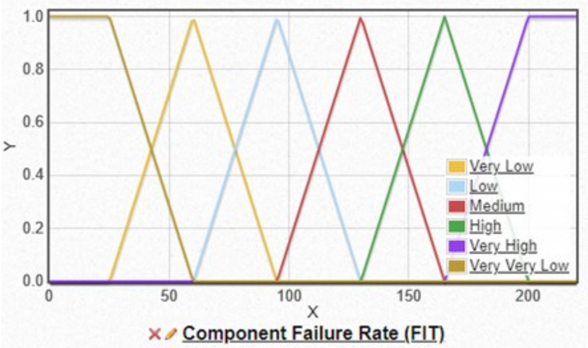


Figure 3.2: Fuzzy set with six triangular membership functions

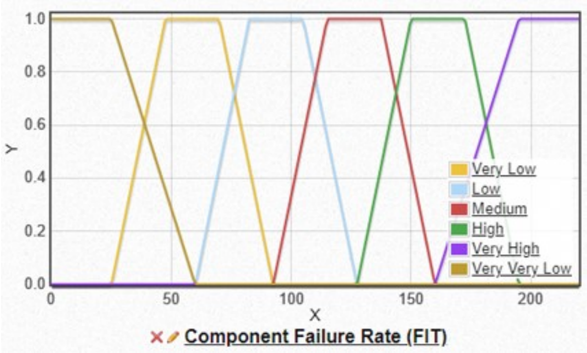


Figure 3.3: Fuzzy set with six trapezoidal membership functions

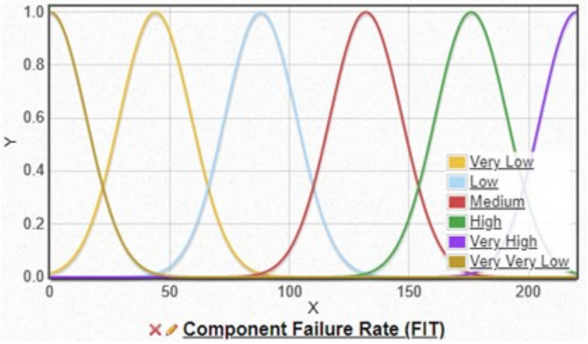


Figure 3.4: Fuzzy set with six Gaussian membership functions

From the literature review the most commonly used membership functions in FMEA applications is the triangular type [127], [126, 128–130], [131–133], [115, 117–125, 134–136] followed by the trapezoidal type [126, 133, 136, 137, 139]. The Gaussian fuzzy set type was the least utilised [138–141].

3.4.3 Universe of Discourse for Membership Functions

The universe of discourse for the membership functions was determined based on industry data collected from FMEDA experts. Specifically, the universe of discourse for the component failure rate reflects the range of failure rates identified in this expert data. Failure rate data ranged from 1 to 220 FIT, so the universe of discourse is 0 to 220 FIT, this range is also applied to the output membership function, the Weighted revealed failure rate. The universe of discourse for the Probability of revealing the failure is based on both full stroke test and partial stroke test applications. Typically, the probability of revealing a dangerous failure for a full stroke test is 1. The dangerous failures related to an ESDV being Fail to Close (FTC), Delayed Operation (DOP) and Leak in the Closed Position (LCP). There may be an argument made for the reliability of the means of revealing the failure, such as instrumentation, but generally a probability of 1 is acceptable. Conversely, the revealability of these failures is not 100% for a partial stroke test, but an estimate of probability can be determined.

A partial stroke test is undertaken by moving the ESDV from its fully open position to, typically, 20% of its closed position, hence we are only potentially revealing part of the Fail to Close (FTC) and Delayed Operation (DOP) dangerous failure modes. The FTC failure mode consists of two sub-failures, i.e. stuck open and fail to fully close [55, 67]. Therefore, a probability of revealing the failure can be assigned to both of these sub-failures, for example, stuck open can be revealed if the ESDV is moved 20%. Equally, the DOP failure mode is made up of two sub-failure, for example, fail to close 20% in the specified time and fail to fully close in the specified time. Table 3.2 illustrates that dividing the failure modes equally, each sub-failure is assigned a probability of 0.5. The probability of revealing a stuck open failure mode is considered conservative at 0.5, like wise the probability of revealing fail to close 20% in the specified time. The universe of discourse for the probability of revealing a failure for both full stroke test and partial stroke test was therefore assigned a range 0.5 to 1.

Table 3.2: Partial stroke test sub-failures

Failure mode	Sub-failure mode	Probability of revealing the sub-failure mode
Fail to fully close (FTC)	Stuck open	0.5
	Fail to fully close	0 to 0.5
Delayed Operation (DOP)	Fail to close 20% in the specified time	0.5
	Fail to fully close in the specified time	0 to 0.5

Table 3.3 summaries the universe of discourse for the input and output membership functions:

Table 3.3: Input and output membership function - universe of discourse

Membership function	Type	Universe of discourse
Component failure rate	Input	0 to 220 FIT
Probability of revealing a failure	Input	0.5 to 1
Weighted revealed failure rate	Output	0 to 220 FIT

3.4.4 Partitioning of Fuzzy Sets and the Application of Linguistic Variables

With regard to partitioning of fuzzy sets, the number of membership functions varies from three to ten, with the majority employing fuzzy inference systems with five membership functions [129–132, 138, 140, 141]. The greater number of membership functions improves the accuracy of the FIS but with added complexity, for example a large number of membership functions such as 10 will significantly increase the size of the rule base. Equally spaced partitioning of membership functions is a common feature of FIS in literature, which also provides a sensible starting point in the design, Adjustment of partitioning can be carried out if necessary after testing and validation.

The linguistic variables used in the design are therefore Very Low, Low, Medium, High and Very High, and equally spaced.

3.4.5 Overlap of Membership Functions

The fuzzy sets are initially equally partitioned and implemented with an approximate 50% overlap between membership functions. Overlap between membership functions provides smooth transition from one membership function to another. The recommended degree of membership function overlap is from approximately 10% to 50% [142]. Shouldered membership functions are employed on first and last membership functions to allow low and high end values of the universe of discourse to be processed. An example of a triangular 5 membership function fuzzy set, equally spaced with shouldered membership functions is shown in Figure 3.5.

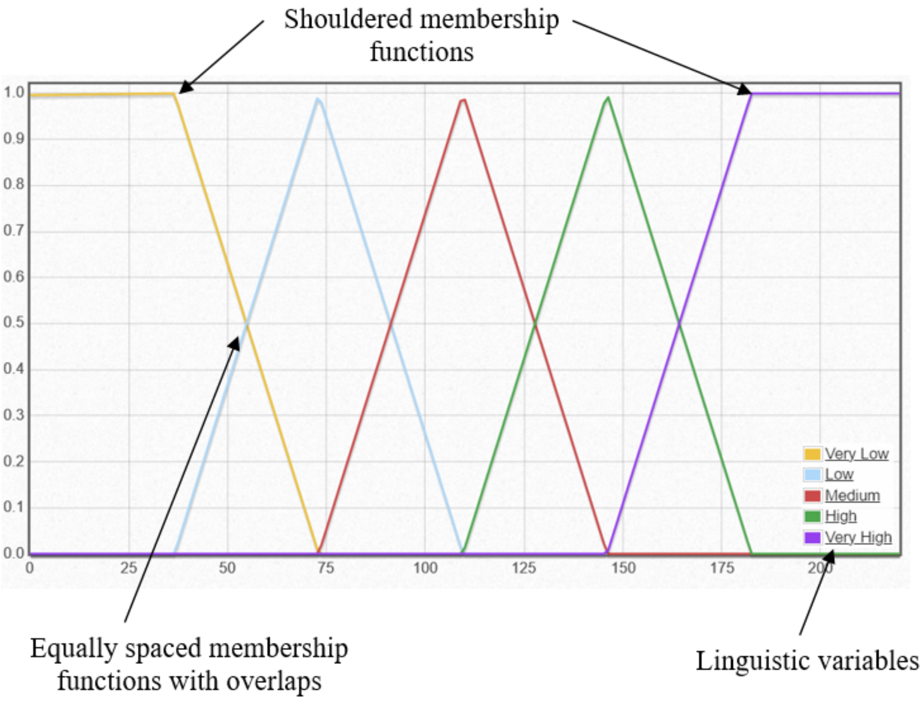


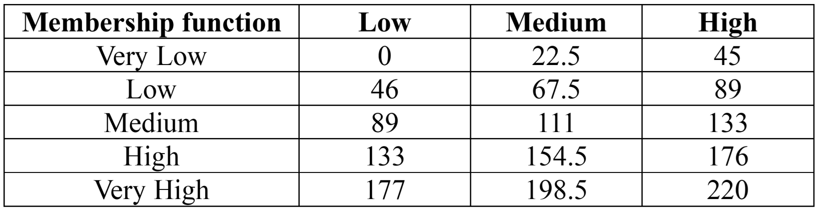
Figure 3.5: Example fuzzy set showing equally spaced membership functions with overlaps, shouldered membership functions and linguistic variables

3.4.6 Rule Base

The IF THEN rule base was developed by mapping different combinations of the inputs to the output. A set of rules were developed based on five membership functions per input (5x5) which results a total number of 25 rules. The rules were developed as detailed in the following discussion.

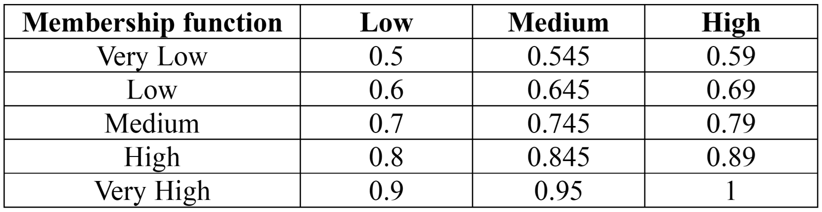
The calculation of Weighted revealed failure rate is the product of the two inputs, Component failure rate and the Probability of revealing a failure. As discussed previously the universe of discourse for component failure rate is from 0 to 220 FIT and dividing into equally between five membership functions provides a range of approximately 45 FIT per membership function. The same process was applied to the probability of revealing a failure, where the universe of discourse ranges from 0.5 to 1. By dividing this range equally into five membership functions, each function covers a range of 0.1. Then taking the low, medium and high ranges of the fuzzy sets the following two tables were developed for the component failure rate membership functions (Table 3.4) and probability of revealing a failure membership functions (Table 3.5):

Table 3.4: Component failure rate (FIT) membership function ranges



Membership function	**Low**	**Medium**	**High**
Very Low	0	22.5	45
Low	46	67.5	89
Medium	89	111	133
High	133	154.5	176
Very High	177	198.5	220

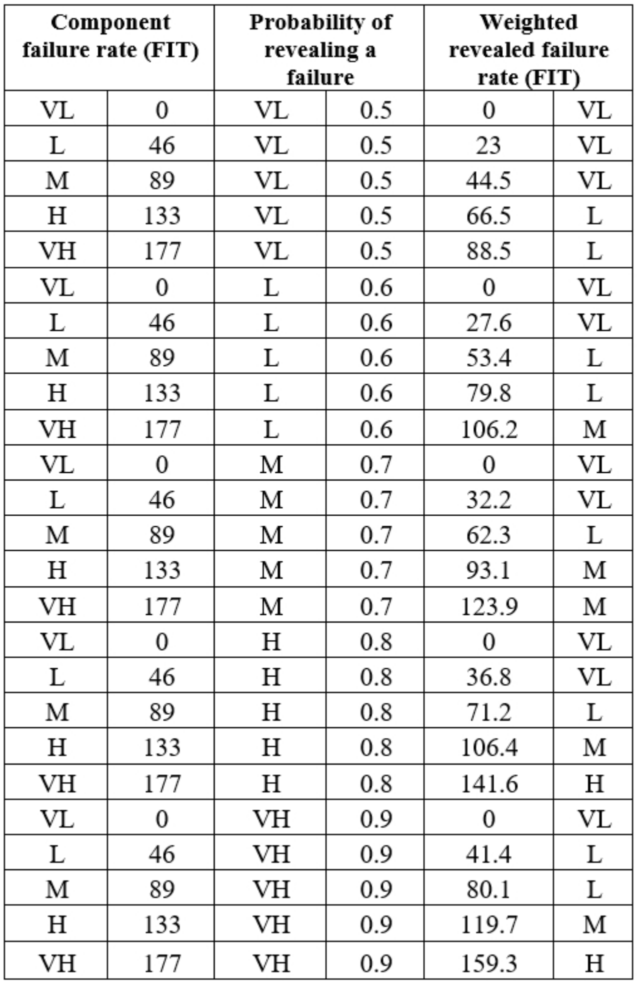
Table 3.5: Probability of revealing a failure membership function ranges



Membership function	**Low**	**Medium**	**High**
Very Low	0.5	0.545	0.59
Low	0.6	0.645	0.69
Medium	0.7	0.745	0.79
High	0.8	0.845	0.89
Very High	0.9	0.95	1

The various ranges of each membership function were then tabulated to calculate the Weighted revealed failure rates for the Low, Medium and high ranges for each fuzzy set. The low range values are shown in Table 3.6:

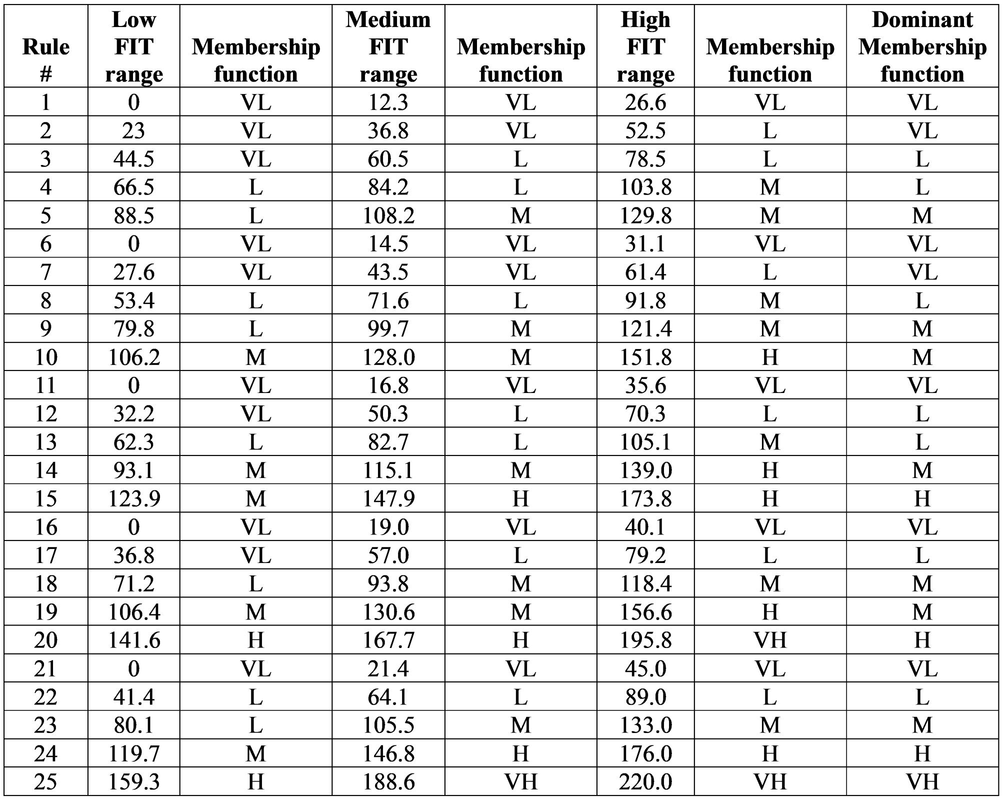
Table 3.6: Example of the determination of linguistic variables for the Low range of component failure rate (VL = Very Low, L = Low, M = Medium, H = High, VH = Very High)



Component failure rate (FIT)		Probability of revealing a failure		Weighted revealed failure rate (FIT)	
VL	0	VL	0.5	0	VL
L	46	VL	0.5	23	VL
M	89	VL	0.5	44.5	VL
H	133	VL	0.5	66.5	L
VH	177	VL	0.5	88.5	L
VL	0	L	0.6	0	VL
L	46	L	0.6	27.6	VL
M	89	L	0.6	53.4	L
H	133	L	0.6	79.8	L
VH	177	L	0.6	106.2	M
VL	0	M	0.7	0	VL
L	46	M	0.7	32.2	VL
M	89	M	0.7	62.3	L
H	133	M	0.7	93.1	M
VH	177	M	0.7	123.9	M
VL	0	H	0.8	0	VL
L	46	H	0.8	36.8	VL
M	89	H	0.8	71.2	L
H	133	H	0.8	106.4	M
VH	177	H	0.8	141.6	H
VL	0	VH	0.9	0	VL
L	46	VH	0.9	41.4	L
M	89	VH	0.9	80.1	L
H	133	VH	0.9	119.7	M
VH	177	VH	0.9	159.3	H

The Weighted revealed failure rate for Low, Medium and High ranges were then compared and the dominant membership function linguistic terms identified, as shown in Table 3.7.

Table 3.7: Weighted revealed failure rates for Low, Medium and High ranges of membership values (VL = Very Low, L = Low, M = Medium, H = High, VH = Very High)



Rule #	Low FIT range	Membership function	Medium FIT range	Membership function	High FIT range	Membership function	Dominant Membership function
1	0	VL	12.3	VL	26.6	VL	VL
2	23	VL	36.8	VL	52.5	L	VL
3	44.5	VL	60.5	L	78.5	L	L
4	66.5	L	84.2	L	103.8	M	L
5	88.5	L	108.2	M	129.8	M	M
6	0	VL	14.5	VL	31.1	VL	VL
7	27.6	VL	43.5	VL	61.4	L	VL
8	53.4	L	71.6	L	91.8	M	L
9	79.8	L	99.7	M	121.4	M	M
10	106.2	M	128.0	M	151.8	H	M
11	0	VL	16.8	VL	35.6	VL	VL
12	32.2	VL	50.3	L	70.3	L	L
13	62.3	L	82.7	L	105.1	M	L
14	93.1	M	115.1	M	139.0	H	M
15	123.9	M	147.9	H	173.8	H	H
16	0	VL	19.0	VL	40.1	VL	VL
17	36.8	VL	57.0	L	79.2	L	L
18	71.2	L	93.8	M	118.4	M	M
19	106.4	M	130.6	M	156.6	H	M
20	141.6	H	167.7	H	195.8	VH	H
21	0	VL	21.4	VL	45.0	VL	VL
22	41.4	L	64.1	L	89.0	L	L
23	80.1	L	105.5	M	133.0	M	M
24	119.7	M	146.8	H	176.0	H	H
25	159.3	H	188.6	VH	220.0	VH	VH

The dominant weighted revealed failure rate linguistic terms were then combined with the input linguistic terms to create a rule set, with the inputs and outputs shown in Table 3.8.

Table 3.8: IF-THEN Rule Base

Rule #	Component failure rate (FIT)	Probability of revealing a failure	Weighted revealed failure rate (FIT)
1	Very Low	Very Low	Very Low
2	Low	Very Low	Very Low
3	Medium	Very Low	Low
4	High	Very Low	Low
5	Very High	Very Low	Medium
6	Very Low	Low	Very Low
7	Low	Low	Very Low
8	Medium	Low	Low
9	High	Low	Medium
10	Very High	Low	Medium
11	Very Low	Medium	Very Low
12	Low	Medium	Low
13	Medium	Medium	Low
14	High	Medium	Medium
15	Very High	Medium	High
16	Very Low	High	Very Low
17	Low	High	Low
18	Medium	High	Medium
19	High	High	Medium
20	Very High	High	High
21	Very Low	Very High	Very Low
22	Low	Very High	Low
23	Medium	Very High	Medium
24	High	Very High	High
25	Very High	Very High	Very High

Examples of the rules:

Rule #1 = IF **Component failure rate** is *Very Low* AND the **Probability of revealing the failure** is *Very Low* THEN the **Weighted revealed failure rate** is *Very Low*
Rule #5 = IF **Component failure rate** is *Very High* AND the **Probability of revealing the failure** is *Very Low* THEN the **Weighted revealed failure rate** is *Medium*

3.4.7 Method of Defuzzification

Common defuzzification methods include centroid (centre of area or centre of gravity), weighted average and mean-max membership. The most commonly used methods of defuzzification in literature were centroid [126–128, 130, 131, 134, 138, 140, 141] and mean-max membership [136, 137, 139] and weighted average [132, 135].

The next chapter focusses on estimating proof test coverage using ESDV maintenance and test data collected from industry; these results are later used to compare with the findings from the fuzzy inference system method.

Chapter 4

Method 1 - Estimating PTC using Equipment Failure Data

4.1 Introduction

This chapter presents a method for estimating ESDV proof test coverage using equipment maintenance and test data gathered from industry. The findings of this work serve to validate the accuracy of both the FMEDA and fuzzy inference system approaches for determining proof test coverage for a full stroke test. Figure 4.1 provides an overview of the structure of this chapter.

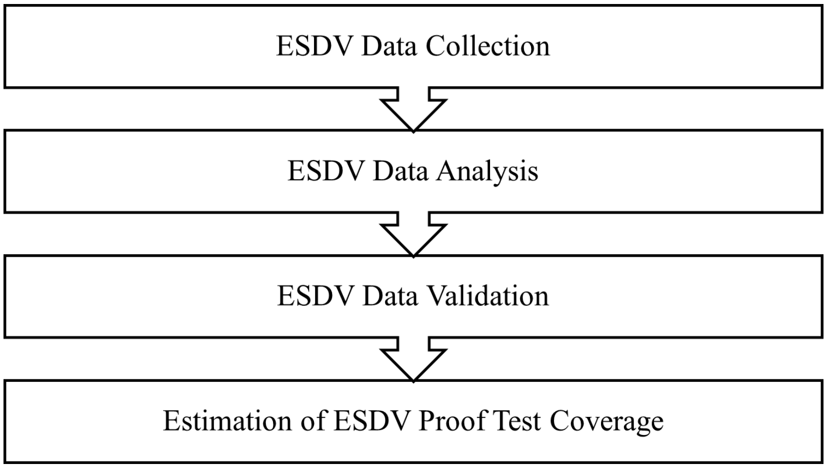


Figure 4.1: Chapter Four Structure

4.2 Data Collection

A study was conducted with the objective of using Emergency Shutdown Valve (ESDV) data collected from industry to calculate proof test coverage using failure modes, effects, and diagnostic analysis (FMEDA). Several operating companies including Liquefied Natural Gas (LNG) terminals, a Liquefied Petroleum Gas (LPG) terminal, a nickel refinery, onshore gas terminals, offshore gas facilities, natural gas networks, and oil refineries were exclusively contacted for this project with mixed success. Numerous valve and actuator manufacturers and valve service companies were also contacted regarding failure data. ESDV maintenance and test data (secondary data) were supplied by four organisations, but only two, which remain anonymous, provided data suitable for analysis. The majority of companies do not undertake ESDV seat leakage tests; therefore, their maintenance and test data would not necessarily include the Leak in the Closed Position (LCP) failure mode. The purpose of collecting these data was to establish the cause of failure, the failure modes, and the failure rates of the components to be used as input to an FMEDA. The maintenance and test data was recorded in their Computerised Maintenance Management Systems (CMMS) and access to the data for this study was provided in the form of data downloads from these systems into spreadsheets. The study concentrated on ball valves as these are typically used for emergency shutdown (ESD) applications [30] and the focus of this research.

The recorded data was found to be at the equipment level rather than the component level, making it unsuitable for input into an FMEDA. It lacked sufficient detail to identify which ESDV component had failed and the cause of the failure. Further investigation would be required to obtain this information, which was not feasible during the period of this investigation. The study was then reconsidered and the equipment level data was used to assess the proof test coverage based on equipment failure modes rather than component failure modes. The data enabled the calculation of proof test coverage for a full stroke test, which would later prove useful for comparison with the results from the FMEDA and fuzzy inference systems. Aspects of the data we also used to determine the probability of revealing a failure for the Fail to Close (FTC) failure mode. Data were reviewed to determine dangerous failure modes following the guidance in standard BS EN 14224: Petroleum, Petrochemical and Natural Gas Industries - Collection and Exchange of Reliability and Maintenance [143].

The use of FMEDA to determine proof test coverage is the focus of the next chapter through which an assessment is performed using component failure rate data from an industry database.

4.3 ESDV failure modes

The BS EN 14224 standard was originally introduced in 1999 [144] (the most recent revision released in 2016) with the objective of improving the way the petroleum, petrochemical, and natural gas industries collect data so that it can be exchanged to improve equipment safety, availability, reliability, and maintainability. The standard covers various industrial equipment, which includes numerous types of valve, and for each piece of equipment the potential failure modes and associated failure codes. Industry data handbooks such as Offshore and Onshore Reliability Equipment Data

(OREDA) [26] and the Foundation for Industrial and Technical Research (SINTEF) [27] identify equipment failure modes as prescribed in this standard. The boundary of the ESDV assembly identified in BS EN 14224 is shown in Figure 4.2. The focus of this research is the solenoid valve, actuator, and valve.

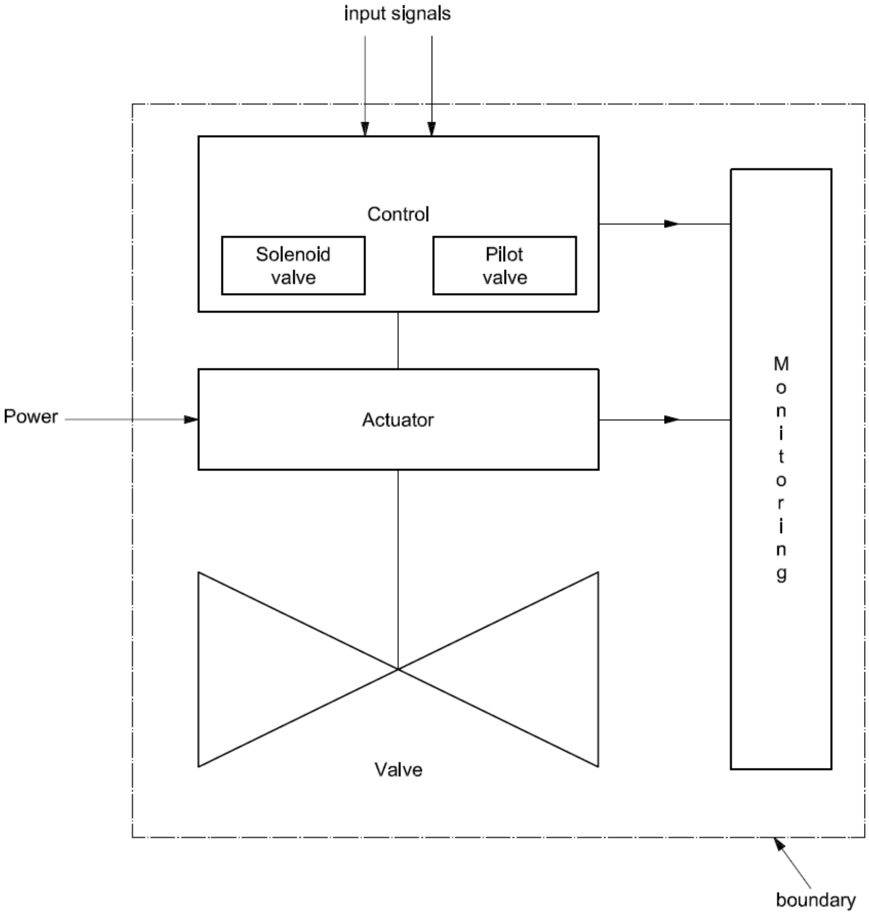


Figure 4.2: ESDV Assembly Boundary [143]

Table 4.1 shows the failure modes and descriptions of valve assemblies, including the universally accepted failure mode code.

Table 4.1: ESDV assembly dangerous failure modes [143]

Failure mode code	Description	Examples of failure
AIR	Abnormal instrument reading	False alarm, faulty instrument indication
DOP	Delayed operation	Opening/closing time below spec.
ELP	External leakage - process medium	Oil, gas, condensate, water
ELU	External leakage - utility medium	Hydraulic oil, lubrication oil, barrier oil, coolant, water, etc.
FTC	Failure to close on demand	Does not close on demand
FTO	Failure to open on demand	Does not open on demand, stuck closed or fail to open fully
HIO	High output	Overspeed/output above acceptance
INL	Internal leakage	Leakage internally of process or utility fluids
LCP	Leakage in closed position	Leak through valve in closed position
LOO	Low output	Delivery/output below acceptance
NOI	Noise	Abnormal or excessive noise
OTH	Other	Failure modes not covered above
PLU	Plugged/ choked	Partial or full flow restriction
SER	Minor in-service problems	Loose items, discoloration, dirt
SPO	Spurious operation	Fails to operate as demanded, false alarm, premature closure/stop, unexpected operation/fails to operate as demanded
STD	Structural deficiency	Material damages (cracks, wear, fracture, corrosion), reduced integrity
UNK	Unknown	Too little information to define a failure mode

As discussed previously, the failure modes of interest from a proof test coverage perspective are failures categorised as dangerous, as highlighted in Table 4.2.

Table 4.2: ESDV assembly dangerous failure modes [143]

Code	Definition	Description
FTC	Fail to close on demand	Stuck open or fail to close fully
DOP	Delayed operation	Closing time different from specification
LCP	Leakage in closed position	Leak-through valve in closed position

As proof test coverage concerns only dangerous failures, safe failures such as internal leak (INL), external leak process (ELP), and external leak utilities (ELU) were excluded from this study. The purpose of proof testing is to reveal dangerous failures, and the proof test coverage is a function of the number of dangerous failures revealed during testing and the total number of dangerous failures.

4.4 ESDV Data Details

Details of the datasets provided by companies are discussed in the following.

4.4.1 Dataset 'A'

Dataset 'A' included maintenance and test data for the offshore platform Riser ESDV (RESDV). RESDV isolates the topside process from the well and the subsea pipeline in the event of an emergency shutdown, to prevent loss of containment [145]. RESDVs are an essential risk reduction measure for offshore installations and are a legal requirement under the Pipelines Safety Regulations 1996 [146]. Data included Solenoid-Operated Valve (SOV), actuator (hydraulic), and valve failures. The data included the following information:

- Date.
- Equipment ID (tag number).
- Work order number.
- Location.
- Description of failure.
- Failure mode (as per BS EN 14224).

A sample of the data collected is shown in Appendix B.1.

4.4.2 Dataset 'B'

Dataset 'B' included maintenance and test data for Process ESDVs. Process ESDVs are used to protect the facility from excursions from normal operating conditions. Data included Solenoid-Operated Valve (SOV), actuator (pneumatic), and valve failures. The data included the following information:

- Date.
- Equipment ID (tag number).
- Location.
- Application.
- Description of failure.
- Failure mode (as per BS EN 14224).

4.4.3 Dataset 'C'

Dataset 'C' included maintenance and test data for over pressure protection valves (actuated ball valves). These devices are activated to close on detection of excessive pipeline pressure to protect the pipeline from loss of containment. The data includes actuator (pneumatic) and valve failures. The data included the following information:

- Date.
- Equipment ID (tag number).
- Activity to rectify.
- Defective part.
- Consequence of failure.
- Fault type (in-house fault codes assigned).

A sample of the data collected is shown in Appendix B.2.

4.5 ESDV Data Analysis

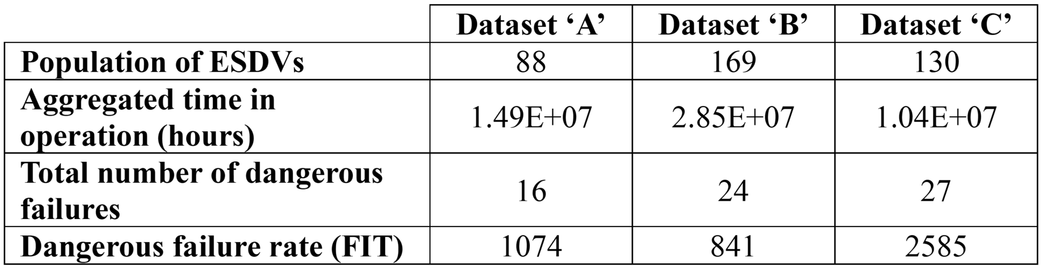
A review of the datasets included confirming the assigned failure mode codes aligned with the description of failures.

The company that provided Dataset 'C' used its own in-house fault type coding which did not conform to BS EN 14224 failure mode coding, which meant additional work was required to determine appropriate failure mode codes that could be used in this study.

4.5.1 ESDV Dangerous Failure Rates

Table 4.3 shows, for each dataset, the population of ESDVs, aggregated time the ESDVs were in operation, total number of failures, and the overall dangerous failure rate for the ESDVs. A Failure in Time (FIT) is defined as one failure in one billion operating hours.

Table 4.3: ESDV population, aggregated time in operation, total number of failures and dangerous failure rate



	Dataset ‘A’	Dataset ‘B’	Dataset ‘C’
Population of ESDVs	88	169	130
Aggregated time in operation (hours)	1.49E+07	2.85E+07	1.04E+07
Total number of dangerous failures	16	24	27
Dangerous failure rate (FIT)	1074	841	2585

The aggregated time is the product of the population of ESDVs and the time the ESDVs were in operation. The dangerous failure rate (λ_D) is calculated using the following equation:

$$\lambda_D = \frac{k}{t} \tag{4.1}$$

Where:
k = total number of dangerous failures
t = aggregated time in operation.

For example, Dataset ’C’ includes 130 ESDVs and data is collected over a period of 3,348 days (80,352 hours); therefore, the aggregated time in operation is 10,445,760 hours. Thus, the total dangerous failure rate is

$$\lambda_D = \frac{27}{10,445,760} = 2585FIT$$

4.5.2 ESDV Failure Modes, Number of Failures and Failure Mode Distribution

Table 4.4 shows, for each dataset, the number of ESDV failures per failure mode and the failure mode distribution.

Table 4.4: The number of ESDV failures per failure mode and the failure mode distribution

	Dataset ‘A’		**Dataset ‘B’**		**Dataset ‘C’**	
Failure Mode	**No. of Failures**	**Failure Mode Distribution**	**No. of Failures**	**Failure Mode Distribution**	**No. of Failures**	**Failure Mode Distribution**
FTC	9	56%	17	71%	13	48%
DOP	4	25%	4	16%	10	37%
LCP	3	19%	3	13%	4	15%
Total	16	100%	24	100%	27	100%

As can be seen from the table, the dominant failure mode is the Fail to Close (FTC) failure mode followed by Delayed Operation (DOP). Dataset ’B’ has the highest percentage of FTC failures at 71% and also the lowest percentage of Leak in the Closed Position (LCP) failures at 13%.

The FTC failure mode is defined as ”stuck open or fail to fully close”. ”Stuck open” failure mode is easily verified through testing. ”fail to fully close” not so, this failure mode assumes that if the valve indicates fully closed via local position indicator or limit switches [27] that the valve obturator has fully closed - but in fact, without positive confirmation the obturator is in the closed position the revealability of FTC remains an assumption.

4.6 Validation of Failure Rate Data against Industry Databases

Validating failure rate data is crucial as it ensures that the data are credible and meaningful. The datasets were validated against established industry equipment failure databases, namely OREDA [26], SINTEF [27], FARADIP.FOUR [29] and exida [78]. The analysis of the datasets, considering overall equipment failure rates and the failure mode distribution, indicated good correlation with that of industry equipment failure databases, as discussed in the following.

4.6.1 Failure Rate Comparison

Table 4.5 shows the dangerous failure rates of the datasets compared to the figures of the industry database.

Table 4.5: Dangerous Failure Rate Comparison for ESDV assemblies (OREDA [26, p. 453], SINTEF [27, p. 96], FARDIP.FOUR [29], exida [78]

Data source	Dangerous failure rate (FIT)
OREDA	1380
SINTEF	2500
FARADIP.FOUR	1000
exida	800 to 2570
Dataset 'A'	1074 ($\lambda_{90\%}$ = 675 to 1636)
Dataset 'B'	841 ($\lambda_{90\%}$ = 580 to 1183)
Dataset 'C'	2585 ($\lambda_{90\%}$ = 1824 to 3565)

FARADIP.FOUR geometric mean is 2000 FIT. Noting that the FARADIP.FOUR failure rate range is not specific to a type of failure, hence the high upper value. The data suggest that the dangerous failure rates of all datasets are comparable with the ranges stated in the industry databases. The Chi-Square (χ^2) 90% confidence interval (upper and lower) failure rates are generally comparable, with the exception of Dataset 'C', where the upper value is notably higher.

4.6.2 Failure Mode Distribution Comparison

Figure 4.3 indicates that the failure mode distribution of all datasets align with that of OREDA [26] and SINTEF [27] industry databases, i.e., the dominant failure mode is Fail to Close (FTC), followed by Delayed Operation (DOP) and then Leak in the Closed Position (LCP).

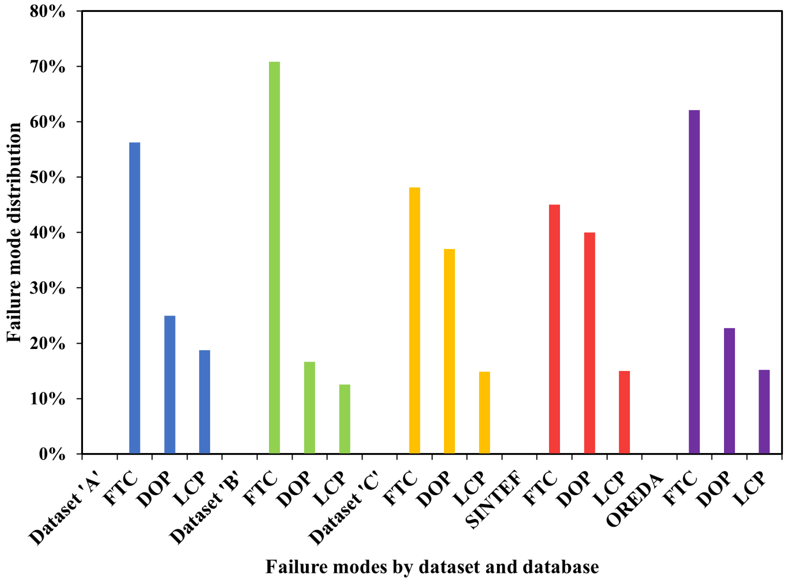


Figure 4.3: Failure Mode Distribution Comparison of Datasets and Industry Databases

4.6.3 Dataset ’A’ Failure Mode Distribution Compared with HSE COIN Database

The UK Health and Safety Executive (HSE) Corporate Operational Information system (COIN) database is a central repository used to manage information related to workplace health and safety inspections, investigations, and enforcement actions. The database holds information on a wide range of incidents, safety violations, risk assessments, and compliance reports from various industries throughout the UK.

The database is used primarily by HSE inspectors and officials to track regulatory compliance, monitor workplace safety, and support enforcement decisions. It helps HSE in maintaining records of inspections and incidents and can be used to analyse trends and inform future safety initiatives. Access to the database is limited to HSE officials and personnel and is not publicly available for general use. However, some information obtained from the database can be referenced in reports or publications issued by the HSE.

The HSE Data Mining team performed a database search to identify records of Riser ESDV (RESDV) related failures, and identified 179 cases over a period of 8 years (2006 to 2014). The 179 cases were reviewed and, where possible, immediate or underlying causes of failure were identified. Twenty-nine duty holders were contacted with surveys on RESDV failures, of these twenty-two responded covering 117 of the 179 identified incidents. A total of 104 reports were submitted by duty holders, although some of the incidents initially identified were not caused by RESDV failures.

Figure 4.4 shows the failure modes of RESDVs taken from the HSE’s COIN database [147]. As can be seen, the dominant failure mode is Fail to Close (FTC) with 109 recorded incidents. The failure modes FTC, Delayed Operation (DOP), and Leak in the Closed Position (LCP) account for approximately 93% of the failure modes.

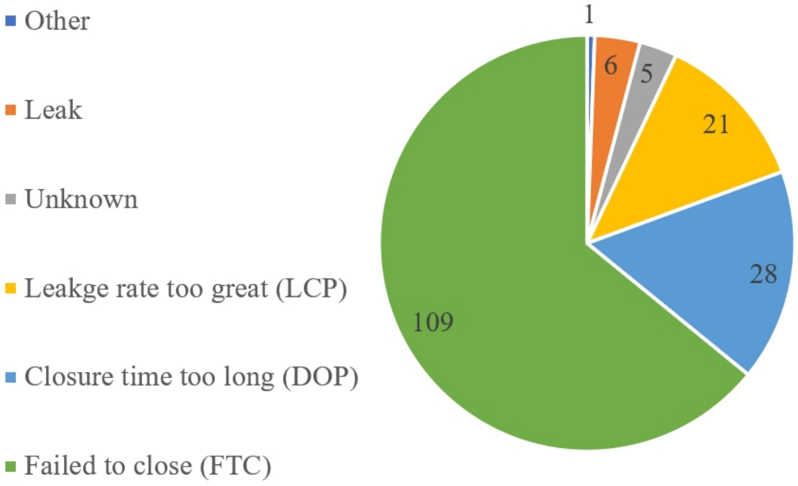


Figure 4.4: HSE COIN RESDV Failure Modes (approximate numbers) [147]

Figure 4.5 shows a comparison between Dataset #1 RESDV failure mode distribution and that of the HSE Corporate Operational Information (COIN) database [147]. The chart illustrates good correlation between the two sets of failure mode distributions, with the FTC failure mode being dominant, followed by DOP and then LCP.

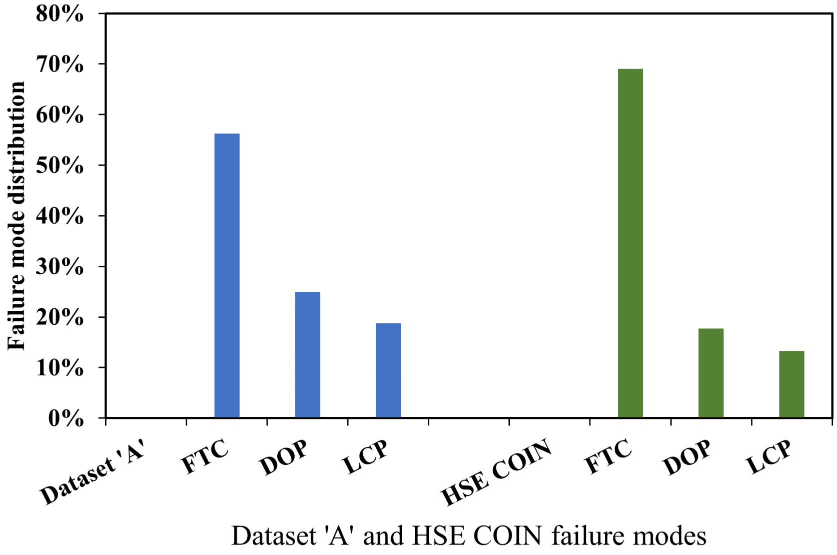


Figure 4.5: Dataset 'A' and HSE COIN database Failure Rate Distribution Comparison

4.7 Estimation of Proof Test Coverage from ESDV Failure Mode Distribution

The study included a total of 387 ESDVs over an aggregated operating period of approximately 54 million hours. Validation checks suggest that valve failure rates and failure mode distributions align with those of industry databases and therefore are meaningful. The Proof Test Coverage (PTC) was assessed for each valve type based on the following ESDV full stroke proof test and the field data analysed from Section 4.5:

- Offline
- Timed closure
- Full stroke of the valve (from open to closed)
- No seat leakage test (or confirmation of obturator closure)

Implementing the above proof test would suggest that the Fail to Close (FTC) and Delayed Operation (DOP) would be revealed with a high level of confidence, hence the probability of revealing these failure modes would be assumed to be 1. On that basis, it can be assumed that the estimate of PTC is the sum of the two revealable failure modes, FTC and DOP [27, 75]. Table 4.6 shows the estimated PTC values for a full stroke test using the failure mode distribution of the three datasets, indicating a PTC range of 81% to 87%.

Table 4.6: Estimated PTC values based on failure mode distribution

	Dataset ‘A’	**Dataset ‘B’**	**Dataset ‘C’**
Failure Mode	**Failure Mode Distribution**	**Failure Mode Distribution**	**Failure Mode Distribution**
FTC	56%	71%	48%
DOP	25%	16%	37%
LCP	19%	13%	15%
PTC	**81%**	**87%**	**85%**

4.8 Discussion and Conclusions

For a full stroke test, the Fail to Close (FTC) failure mode is inferred through verification by a local indicator or limit switches. If the ESDV is constructed with materials suitable for its service conditions, the actuator is appropriately sized to match the valve stem, considering the Maximum Allowable Stem Torque (MAST), and the local indicator and limit switches are deemed reliable, the probability of false verification will be significantly reduced. Monitoring of valve stem movement during testing, if possible, should be considered as an additional means of verifying FTC.

Fail to Close (FTC) failure mode can be broken down into sub-failures, i.e., stuck open and fail to fully close. From Datasets ’A’ and ’B’ it was possible to identify the breakdown of these failure modes by the description of the failure. The percentage of stuck open failure modes was approximately 70%. This value can be used as part of the calculation for partial stroke testing proof test coverage, as discussed in more detail in the following chapters.

Typically, proof tests which include seat leakage tests to confirm LCP failure mode are executed offline at pressures far lower than operating pressures, therefore the probability of a lower failure rate is expected, as opposed to testing at worst-case process conditions i.e. at maximum operating pressure. Hence, the PTC for an offline test is likely to be lower than that of a test carried out at process conditions.

The ESDV assemblies in Dataset ’A’ (RESDVs) are equipped with hydraulic actuators as opposed to Dataset ’B’ and Dataset ’C’ which are pneumatic. From a failure rate perspective, these have comparable dangerous failure rate ranges, with the hydraulic type having a range of 325 FIT to 1100 FIT and the pneumatic equivalent 300 FIT to 1200 FIT [78]. The difference in the failure rates is marginal and therefore

should not have a detrimental impact on the data analysis aspect of this study.
Proof test coverage assumes that the proof tests are performed perfectly by the proof tester, without human error. Therefore, it may be necessary to examine the potential impact of human error to determine the effect on the overall PTC. A Human Reliability Assessment (HRA) focusses on evaluating and improving the reliability of human performance in complex systems, particularly in high-risk industries such as nuclear power, oil & gas, and chemical processing. The goal is to determine, quantify, and mitigate the likelihood of human errors that could lead to system failures or accidents. Techniques such as the Human Error Assessment and Reduction Technique (HEART) [148] and Standardised Plant Analysis Risk-Human Reliability Analysis (SPAR-H) [149] are examples of HRA techniques.

The proof test coverage results from this study indicate that the proof test coverage for an ESDV full stroke test based on industry equipment maintenance and test records is approximately 81% to 87% and comparable with published PTC estimates. The PTC estimates align with the upper range of published values, resulting in favourable $\mathrm{PFD}_{\mathrm{avg}}$ outcomes and potentially allowing reduced proof test frequencies compared to estimates at the lower end of the scale. Hence, a reduced maintenance overhead can be recommended for the operating company.

Figure 4.6 shows the published PTC estimates (blue) and the results of this analysis based on the equipment failure rate data (green).

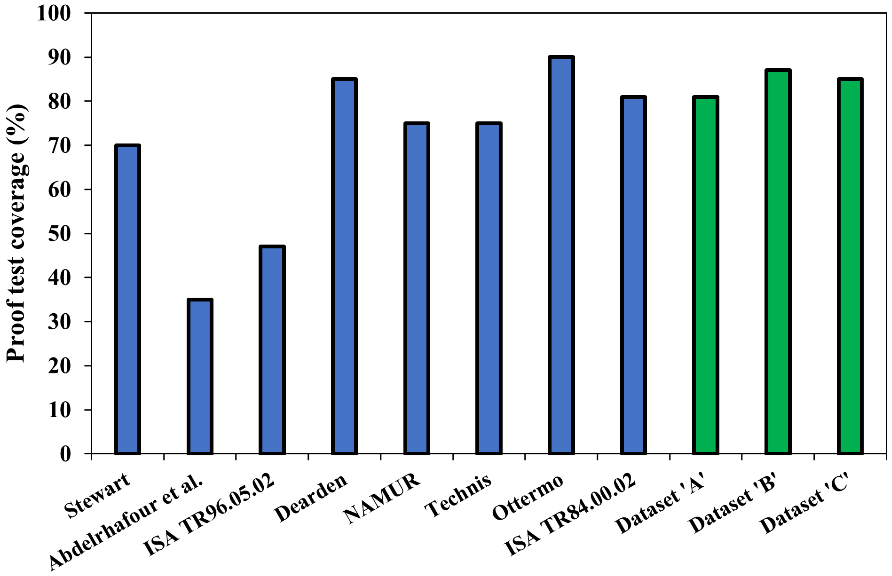


Figure 4.6: Published PTC estimates (blue) and results from analysis (green). Stewart [47], Abdelrhafour et al. [75], ISA TR96 [79], Dearden [49], NAMUR [48], Technis [69], Ottermo [27], ISA TR84 [80]

Although upper and lower Chi-square confidence levels (χ^2) were applied to the data, additional data collection efforts would provide a broader sample that could be more representative of the industry. However, a significant challenge in this study was acquiring equipment failure data from organisations, as it is considered commercially sensitive.

Using equipment failure data collected from industry offers several advantages in estimating proof test coverage. Since the data originates from actual operational contexts, it reflects real-world conditions, including wear, maintenance practices, and environmental factors. This specificity provides tailored insights into proof test coverage for a particular plant or system, rather than relying on generic assumptions. In addition, industry data captures recent trends and evolving equipment performance, enabling dynamic adjustments to maintenance strategies. Using this approach, plant operators can achieve more accurate and context-specific optimisation of proof test intervals.

However, relying on industry data also has limitations. The quality and completeness of the data often depend on how consistently maintenance and testing records are kept, and this can vary widely between organisations. Access to comprehensive datasets may be restricted due to commercial sensitivity or confidentiality concerns, leading to smaller sample sizes and reduced statistical robustness. Furthermore, results derived from industry data can be biased or highly variable between different sites, potentially limiting their generalisability. Despite these challenges, industry data remains a valuable resource when available, especially when used in combination with other methods to ensure a holistic understanding of proof test coverage.

4.8.1 Summary of Findings

The key findings of this chapter can be briefly summarised in the following points:

- Using equipment maintenance and test data, the proof test coverage for an ESDV full stroke test can be estimated by identifying dangerous failure modes in accordance with BS EN 14224.
- This study found that the PTC results for a full stroke test were consistent with the upper values of the published estimates.
- Access to comprehensive datasets was restricted due to commercial sensitivity, leading to a small sample size.

This chapter discussed a method of estimating proof test coverage using equipment maintenance and test data obtained from industry. The quality of maintenance and test data can vary as it is based on the precision and diligence of the individual conducting and recording the activities. Access to comprehensive datasets has been shown to be challenging, limited by commercial sensitivity and confidentiality constraints, resulting in smaller sample sizes and decreased statistical reliability. To enhance the accuracy of the PTC estimates, it is proposed that the analysis be focused at the component level, utilising data from component failure rate databases and other relevant sources. As identified in the literature review, failure modes, effects, and diagnostics analysis

(FMEDA) is the preferred method to estimate the Proof Test Coverage (PTC) of an Emergency Shutdown Valve (ESDV).

The objective of the next chapter is to utilise a recognised component failure database and other relevant sources to perform a Failure Modes, Effects, and Diagnostic Analysis (FMEDA) to estimate the proof test coverage for an ESDV.

Chapter 5

Method 2 - Estimating PTC using FMEDA and Component Failure Rate Data

5.1 Introduction

In the previous chapter, proof test coverage for a full stroke test was estimated based on ESDV maintenance and test data, in which failure data were focused at the equipment level. The quality and completeness of industry data often depend on the consistency of maintenance and testing record keeping, which can vary significantly between organisations. Furthermore, access to comprehensive datasets is frequently restricted due to commercial sensitivity or confidentiality concerns, resulting in smaller sample sizes and reduced statistical reliability. In addition, results derived from such data may exhibit bias or significant variability between different sites, potentially limiting their applicability and generalisability.

In addition to drawing on my prior industrial experience, I undertook a number of industry-based technical training courses during the course of this research to reinforce and broaden my understanding of key subject areas. These courses were chosen not only to deepen my technical insight into areas such as Failure Modes, Effects, and Diagnostic Analysis (FMEDA), but also to provide opportunities to engage with other professionals and draw on their experiences related to the challenges I was investigating. Certificates for the courses attended are included in Appendix G.

This chapter discusses the FMEDA technique and the process of estimating proof test coverage based on ESDV component failure rate data. FMEDA offers a structured and systematic approach to estimating proof test coverage by leveraging established component failure rate databases and reliability data. This method ensures consistency and repeatability in the analysis, as it relies on standardised failure modes and probabilities rather than site-specific variations. Figure 5.1 provides an overview of the structure of this chapter.

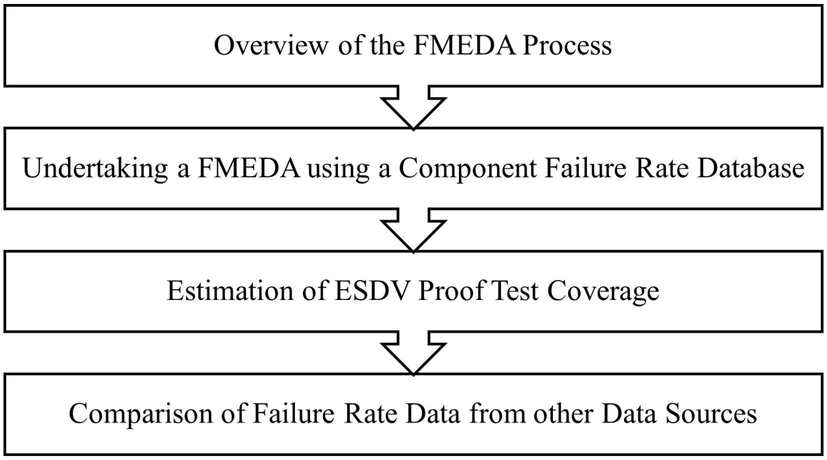

Overview of the FMEDA Process
Undertaking a FMEDA using a Component Failure Rate Database
Estimation of ESDV Proof Test Coverage
Comparison of Failure Rate Data from other Data Sources

Figure 5.1: Chapter Five Structure

5.2 The FMEDA process

FMEDA is a systematic approach to the analysis of equipment failures. The analysis is carried out at the component level through which failure modes are identified and the effects of the failure on the component and system are assessed. Component failure rates are then assigned to failure modes. Part of the assessment is to determine whether a failure can be detected by on-board diagnostics, and if not, the probability of revealing a failure during the proof test. The FMEDA methodology is based on the Failure Modes and Effects (FMEA) technique, a technique that is well established in the field of reliability engineering. Other types of failure modes analysis such as Failure Modes, Effects and Criticality Analysis (FMECA) is also derived from the FMEA technique. To perform an FMEDA, several inputs are necessary; typically this information can be obtained from associated drawings and documentation. The inputs to an FMEDA include the following:

- Equipment data-sheets
- Safety Requirements Specifications (SRS)
- Schematic and/or general arrangement diagrams
- Component failure rate data

The equipment data sheets will provide specification information regarding each device of the ESDV assembly, including the device model and manufacturer. A typical ESDV data sheet is shown in Appendix C.1. Before starting an FMEDA it is important to identify the assessment criteria. Typically, this information can be found in the safety requirements specification (SRS) and includes:

- Safe state of the ESDV - fail safe state, open or closed.
- ESDV proof test requirements - Full Stroke Test (FST) or Partial Stroke Test (PST). The method of verifying the proof test, i.e., locally to the ESDV or remotely via the control system. The type of partial stroke test facilities. Whether the test is online or off-line and whether it will be possible to verify valve closure by means of instrumentation for downstream pressure and/or flow measurement.
- ESDV seat leakage requirements - seat leakage class, typically a maximum allowable leakage rate as opposed to Tight Shut-Off (TSO).

The FMEDA process begins with subdividing the device into its individual elements or parts, in the case of an ESDV that is the Solenoid-Operated Valve (SOV), actuator and valve, as shown in Figure 5.2. The schematic or general arrangement diagrams for each device will be required to identify the individual component parts.

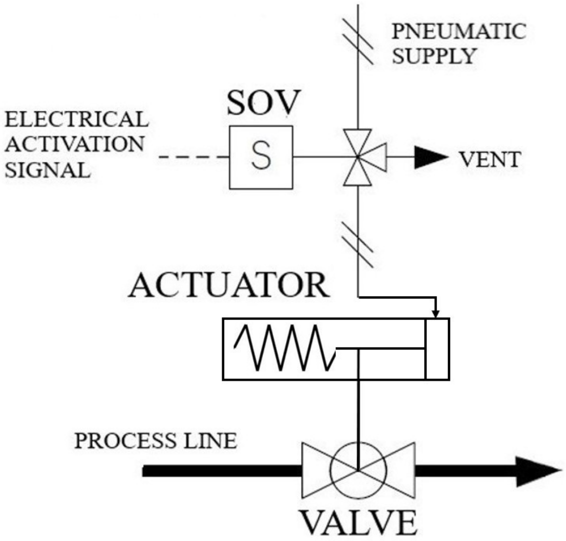


Figure 5.2: An ESDV assembly schematic showing SOV, actuator and valve.

The parts are then subdivided into subassemblies or components. The block diagram (Figure 5.3) shows the individual parts and components of an ESDV assembly. The appropriate level of detail for the analysis should be determined by the context and the specific desired outcomes. In general, greater detail in the level of subdivision of the subject of the FMEDA provides an equivalent level of detail on possible failure modes and effects, hence a more accurate evaluation[31].

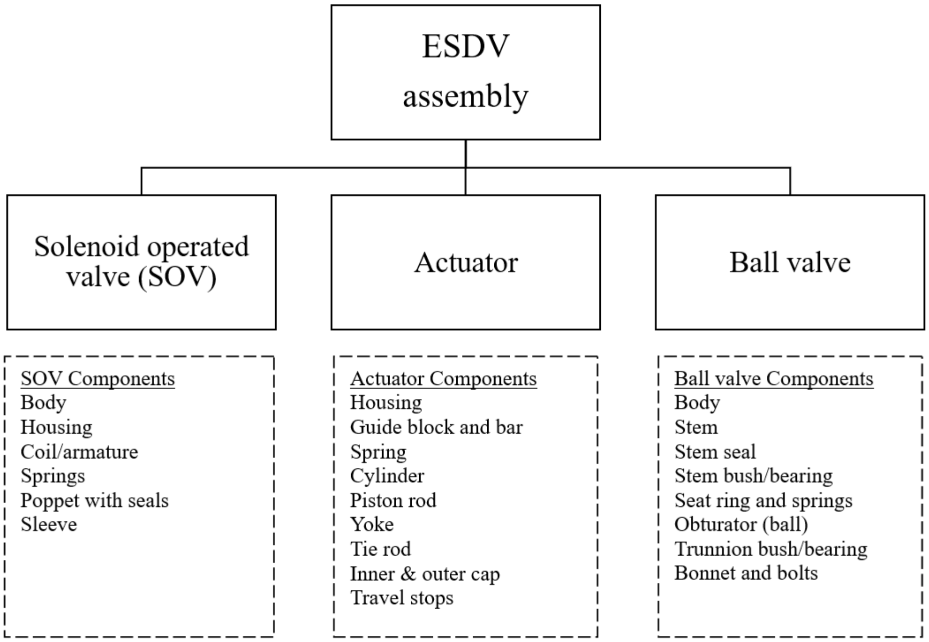


Figure 5.3: Block diagram of an ESDV assembly showing its individual parts and components.

As discussed, it will be necessary to understand the safe and dangerous failure modes of the equipment and specifications such as the required seat leakage requirements. Safe and dangerous failure modes are classified as such based on the context of the application of the ESDV. It is important at this point to clarify the definition of a dangerous failure, as it may be mistakenly assumed that a failure mode, such as a leak to the environment caused by failed body seals, would be classified as dangerous. However, in the context of a Safety Instrumented Function (SIF) ESDV, such a failure would typically be considered a safe failure, despite its potential hazard. As discussed previously, SIFs are designed for specific potentially hazardous events, for example overfilling a storage vessel or overpressurising a process pipeline. Therefore, it is imperative, when performing an FMEDA, that the assessment is undertaken within the context of the application of the ESDV. The application of the ESDV will be provided in the Safety Requirements Specification (SRS).

Each component part is assessed to determine how it can fail (failure mode) and what impact the failure (failure effect) has on the operation of the component, and on the system, i.e., will the failure effect be Safe (S), Dangerous (D), No Effect (NE) or No Part (NP). A safe failure occurs when a component involved in executing the safety function malfunctions, causing the safety function to activate unexpectedly, either bringing the plant to a safe state or maintaining it in a safe state [32]. In the case of an ESDV, the safe state is the closed position (for a fail-safe closed ESDV). A dangerous failure is the failure of a component that plays a role in the implementation of the safety function that prevents it from operating when required so that the plant is placed in a hazardous or potentially hazardous state [32]. A "no effect" failure is a failure of a component that plays a part in implementing the safety function but has no direct effect on the safety function [32]. A "no-part" failure classification refers to the failure of a component that does not contribute to the execution of the safety function [32].

A failure rate is assigned to the failure mode of the component, which is typically a distributed value derived from the overall failure rate of the component, as there is usually more than one type of failure mode. For example, a compression spring shown in Figure 5.4 has two failure modes, i.e., settle and break. The final part of the evaluation includes consideration of whether the failure mode will be revealed during the proof test and the reliability of the proof test verification method, as described in [67]:

- Revealability - to what degree is the failure mode revealable during a proof test, e.g., a Partial Stroke Test (PST) might only reveal a relatively small percentage of the Fail to Close (FTC) and the Delayed Operation (DOP) failure modes.

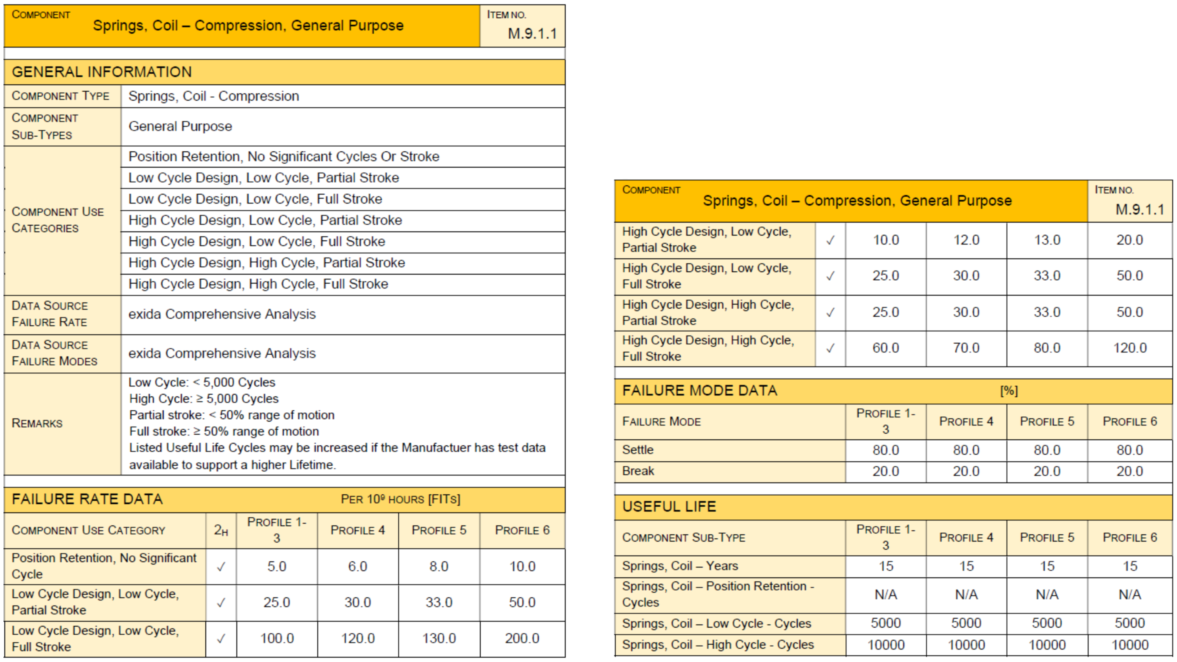
- Reliability - to what degree are the proof test results dependable, such that the revealed results reflect the valve condition, e.g. the reliability of the limit switches indicating Fail to Close (FTC) failure mode (assuming instrumentation is used to verify the operation of the valve). Instrumentation such as limit switches will typically have a high reliability, i.e., greater than 0.95, which will not have a significant impact on the assessment.

The product of the component failure rate (for each particular failure mode), its revealibility and the test reliability provides a 'weighted' revealed dangerous undetected failure rate. The resulting PTC is calculated from the sum of the weighted revealed dangerous undetected failure rates (*Weighted Revealed* λ_{DU}) and the sum of the total dangerous undetected failure rates (*Total* λ_{DU}) using the following PTC equation.

$$PTC = \frac{Weighted\ Revealed\ \lambda_{DU}}{Total\ \lambda_{DU}} \tag{5.1}$$

5.2.1 Component Failure Rate Data

The initial FMEDA assessment was performed using the exida's Component Reliability Database (CRD) handbook [28]. The exida component reliability database (CRD) provides a list of valve assembly components, such as Solenoid-Operated Valves (SOVs), actuators, and valves. Potential failure modes, an overall component failure rate, and a failure rate distribution for specific failure modes are also provided for each component. Component failure rates are detailed with variations based on environmental conditions. The exida failure modes and failure rates were used as a basis for the FMEDA. Figure 5.4 illustrates an example of the entry in the CRD database for a compression spring.



Component	Springs, Coil – Compression, General Purpose	Item No. M.9.1.1

General Information	
Component Type	Springs, Coil - Compression
Component Sub-Types	General Purpose
Component Use Categories	Position Retention, No Significant Cycles Or Stroke Low Cycle Design, Low Cycle, Partial Stroke Low Cycle Design, Low Cycle, Full Stroke High Cycle Design, Low Cycle, Partial Stroke High Cycle Design, Low Cycle, Full Stroke High Cycle Design, High Cycle, Partial Stroke High Cycle Design, High Cycle, Full Stroke
Data Source Failure Rate	exida Comprehensive Analysis
Data Source Failure Modes	exida Comprehensive Analysis
Remarks	Low Cycle: < 5,000 Cycles High Cycle: ≥ 5,000 Cycles Partial stroke: < 50% range of motion Full stroke: ≥ 50% range of motion Listed Useful Life Cycles may be increased if the Manufactuer has test data available to support a higher Lifetime.

Failure Rate Data — Per 10^9 hours [FITs]

Component Use Category	2_H	Profile 1-3	Profile 4	Profile 5	Profile 6
Position Retention, No Significant Cycle	✓	5.0	6.0	8.0	10.0
Low Cycle Design, Low Cycle, Partial Stroke	✓	25.0	30.0	33.0	50.0
Low Cycle Design, Low Cycle, Full Stroke	✓	100.0	120.0	130.0	200.0
High Cycle Design, Low Cycle, Partial Stroke	✓	10.0	12.0	13.0	20.0
High Cycle Design, Low Cycle, Full Stroke	✓	25.0	30.0	33.0	50.0
High Cycle Design, High Cycle, Partial Stroke	✓	25.0	30.0	33.0	50.0
High Cycle Design, High Cycle, Full Stroke	✓	60.0	70.0	80.0	120.0

Failure Mode Data [%]

Failure Mode	Profile 1-3	Profile 4	Profile 5	Profile 6
Settle	80.0	80.0	80.0	80.0
Break	20.0	20.0	20.0	20.0

Useful Life

Component Sub-Type	Profile 1-3	Profile 4	Profile 5	Profile 6
Springs, Coil – Years	15	15	15	15
Springs, Coil – Position Retention - Cycles	N/A	N/A	N/A	N/A
Springs, Coil – Low Cycle - Cycles	5000	5000	5000	5000
Springs, Coil – High Cycle - Cycles	10000	10000	10000	10000

Figure 5.4: exida CRD data for a compression spring [28]

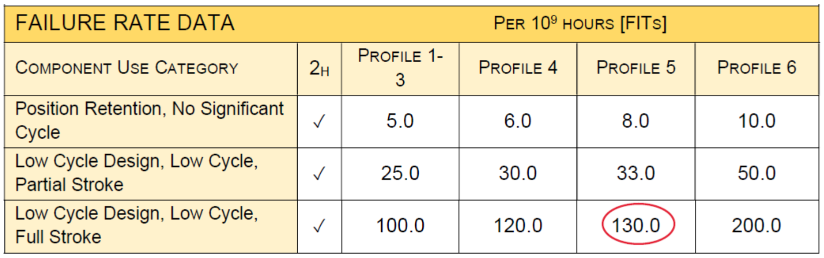
5.2.2 FMEDA Worksheet

Figure 5.5 the FMEDA worksheet that was developed and used in this research with an example entry.

Item #	Component type	Failure mode	Failure effect on component	Failure effect on assembly	Failure type (S, D, NE, NP)	Failure mode BS EN 14224	Failure rate (FIT)	Distribution of failure rate (%)	Failure revealed by diagnostics (Y, N)	Total distributed failure rate (FIT)	Probability of revealing failure	Weighted revealed failure rate (FIT)
A	B	C	D	E	F	G	H	I	J	K	L	M
1	SOV spring	Break	SOV fails to return to safe position	ESDV does not close on demand	Dangerous	FTC	130	20	N	26	1	26

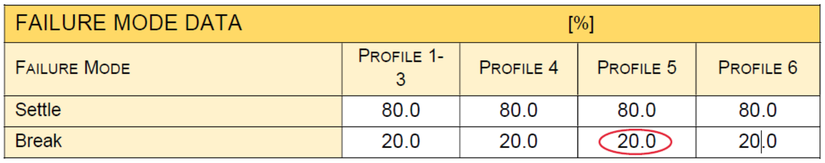
Figure 5.5: FMEDA worksheet with example entry

The example entry describes the breaking (C) of a SOV spring (B) which cause the valve to not return to its safe position (D) and will result in the ESDV not closing on demand (E), which is classed as a dangerous failure (F) and is assigned the failure mode FTC (G) according to BS EN 14224. As a spring has two failure modes (break and settle), the failure rate (H) is distributed between the two, with break (C) assigned 20% (I) with the resultant total distributed failure rate being 26 FIT (K). Figure 5.6 highlights the failure rate (H) and Figure 5.7 highlights the failure distribution (I) of the exida CRD database.



FAILURE RATE DATA		Per 10^9 hours [FITs]			
Component Use Category	2_H	Profile 1-3	Profile 4	Profile 5	Profile 6
Position Retention, No Significant Cycle	✓	5.0	6.0	8.0	10.0
Low Cycle Design, Low Cycle, Partial Stroke	✓	25.0	30.0	33.0	50.0
Low Cycle Design, Low Cycle, Full Stroke	✓	100.0	120.0	130.0	200.0

Figure 5.6: exida CRD data for compression spring failure rate



FAILURE MODE DATA	[%]			
Failure Mode	Profile 1-3	Profile 4	Profile 5	Profile 6
Settle	80.0	80.0	80.0	80.0
Break	20.0	20.0	20.0	20.0

Figure 5.7: exida CRD data for compression spring distribution of failure for the 'break' failure mode

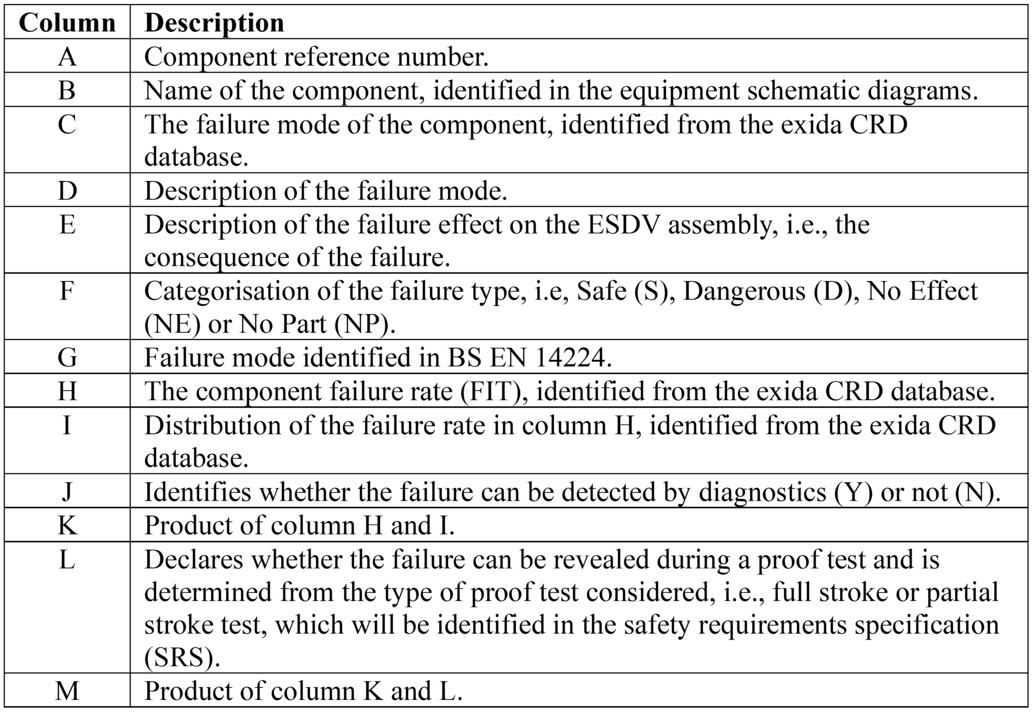
Note that the "Profile" in the component failure rates (Figure 5.6) refers to the environmental conditions of the location of the equipment, with profile 5, for example, representing the offshore environment. Profile 5 has been chosen in this study as it is the worst case environmental conditions. Table 5.1 shows the various environment profiles.

Table 5.1: Environmental profiles [28]

	Profile					
	1	**2**	**3**	**4**	**5**	**6**
Location	Cabinet mounted/ climate controlled	General field mounted (30ºC average internal product temperature)	General field mounted (40ºC average internal product temperature)	Subsea	Offshore	Process wetted

As this is assessment is for a full stroke test the probability of revealing the failure (L) is 1 (i.e. 100%) because the if the valve will not close in the test, it can be assumed it will fail in service. The weighted revealed failure rate (M) for the spring break failure mode is the product of the total distributed failure rate (K) and the probability of revealing the failure (L), i.e. 26 FIT. Table 5.2 provides an explanation of the worksheet entries.

Table 5.2: Explanation of FMEDA entries (exida CRD [28], BS EN 14224 [143])



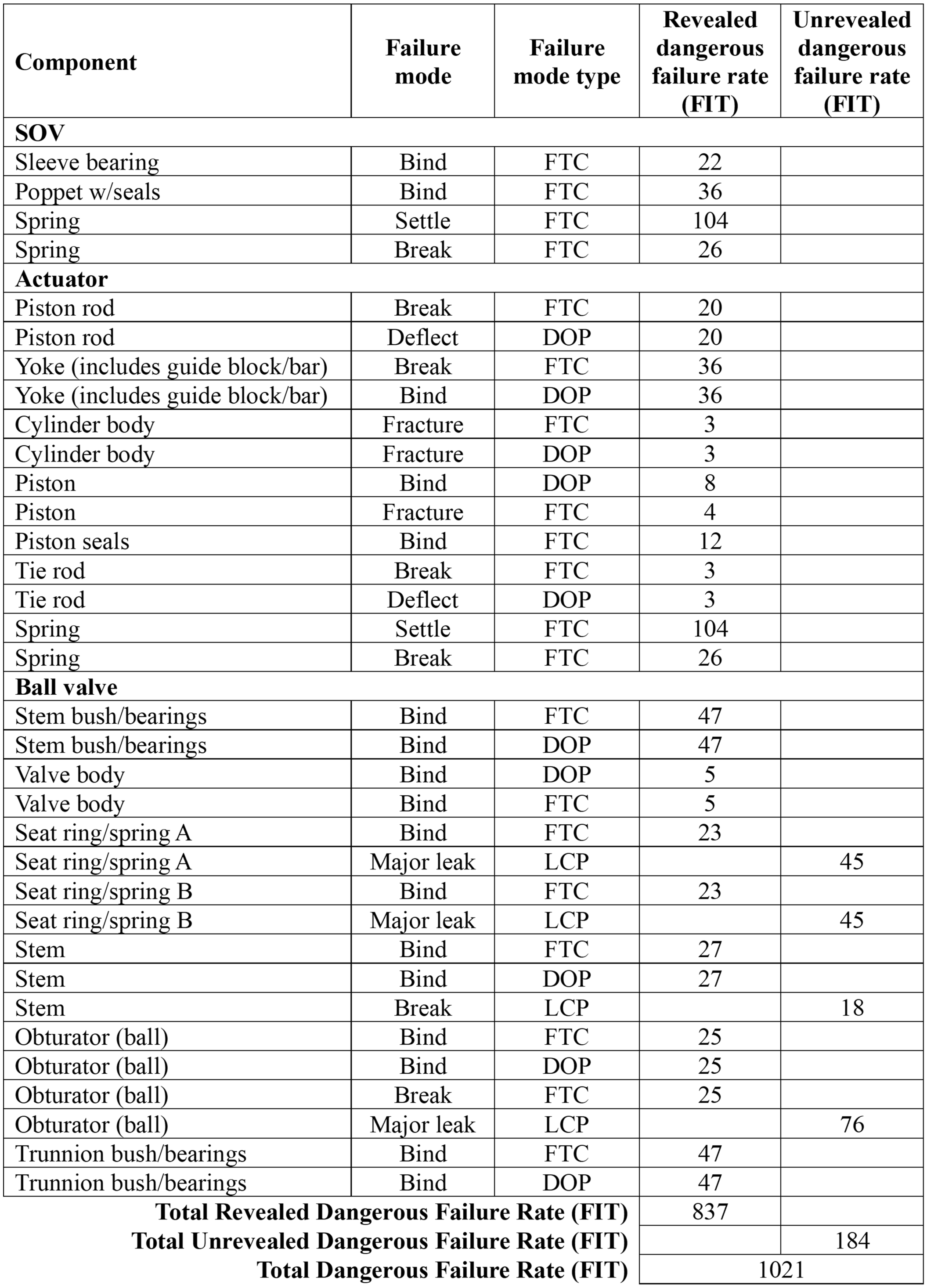
Column	Description
A	Component reference number.
B	Name of the component, identified in the equipment schematic diagrams.
C	The failure mode of the component, identified from the exida CRD database.
D	Description of the failure mode.
E	Description of the failure effect on the ESDV assembly, i.e., the consequence of the failure.
F	Categorisation of the failure type, i.e, Safe (S), Dangerous (D), No Effect (NE) or No Part (NP).
G	Failure mode identified in BS EN 14224.
H	The component failure rate (FIT), identified from the exida CRD database.
I	Distribution of the failure rate in column H, identified from the exida CRD database.
J	Identifies whether the failure can be detected by diagnostics (Y) or not (N).
K	Product of column H and I.
L	Declares whether the failure can be revealed during a proof test and is determined from the type of proof test considered, i.e., full stroke or partial stroke test, which will be identified in the safety requirements specification (SRS).
M	Product of column K and L.

The total dangerous failure rate is calculated from the sum of entries in column K (note: only items declared dangerous in column F). The total weighted revealed dangerous failure rate is calculated from the sum of entries under column M (note: only items declared as dangerous in column F).

5.3 FMEDA Results

Table 5.3 shows the results of the FMEDA carried out using exida CRD. The tables show the components, failure modes and types, identified dangerous revealed and dangerous unrevealed failures and associated failure rates. Safe failures are not included as the calculation of proof test coverage is based on dangerous failures only. The full FMEDA worksheet showing dangerous, safe and no effect failures is shown in Appendix C.2.

Table 5.3: FMEDA results - showing failure rates for components for an ESDV assembly



Component	**Failure mode**	**Failure mode type**	**Revealed dangerous failure rate (FIT)**	**Unrevealed dangerous failure rate (FIT)**
SOV				
Sleeve bearing	Bind	FTC	22	
Poppet w/seals	Bind	FTC	36	
Spring	Settle	FTC	104	
Spring	Break	FTC	26	
Actuator				
Piston rod	Break	FTC	20	
Piston rod	Deflect	DOP	20	
Yoke (includes guide block/bar)	Break	FTC	36	
Yoke (includes guide block/bar)	Bind	DOP	36	
Cylinder body	Fracture	FTC	3	
Cylinder body	Fracture	DOP	3	
Piston	Bind	DOP	8	
Piston	Fracture	FTC	4	
Piston seals	Bind	FTC	12	
Tie rod	Break	FTC	3	
Tie rod	Deflect	DOP	3	
Spring	Settle	FTC	104	
Spring	Break	FTC	26	
Ball valve				
Stem bush/bearings	Bind	FTC	47	
Stem bush/bearings	Bind	DOP	47	
Valve body	Bind	DOP	5	
Valve body	Bind	FTC	5	
Seat ring/spring A	Bind	FTC	23	
Seat ring/spring A	Major leak	LCP		45
Seat ring/spring B	Bind	FTC	23	
Seat ring/spring B	Major leak	LCP		45
Stem	Bind	FTC	27	
Stem	Bind	DOP	27	
Stem	Break	LCP		18
Obturator (ball)	Bind	FTC	25	
Obturator (ball)	Bind	DOP	25	
Obturator (ball)	Break	FTC	25	
Obturator (ball)	Major leak	LCP		76
Trunnion bush/bearings	Bind	FTC	47	
Trunnion bush/bearings	Bind	DOP	47	
		Total Revealed Dangerous Failure Rate (FIT)	837	
		Total Unrevealed Dangerous Failure Rate (FIT)		184
		Total Dangerous Failure Rate (FIT)	1021	

The total dangerous failure rate for the ESDV is 1021 FIT which is in good agreement with industry equipment databases such as OREDA [26] (1380 FIT), SINTEF [27] (2500 FIT) and FARADIP.FOUR [29] (1000 FIT). As discussed previously, the total dangerous failure rate refers to the number of failures that inhibit the ESDV from performing its intended function, which can result in a dangerous event. For the ESDVs examined in this research, dangerous failures include fail to close (FTC), delayed operation (DOP), and leak in the closed position (LCP). There are no unrevealed dangerous failures for the SOV or actuator, only for the valve. As it is a full stroke test without seat leakage verification, it is possible to reveal all the failure types associated with the SOV and actuator, i.e., fail to close (FTC) and delayed operation (DOP). The distribution of failure modes is shown in Table 5.4, which aligns with the expected distribution hierarchy, i.e., FTC highest failure mode percentage followed by delayed operation (DOP) and then leak in the closed position (LCP). Therefore, the proof test coverage estimate for a full stroke test based on exida CRD data was 82%.

Table 5.4: exida CRD FMEDA distribution of failure modes

Failure Mode	Failure Mode Distribution
FTC	60%
DOP	22%
LCP	18%

5.4 FMEDA Using Other Data Sources

Additional FMEDA assessments were undertaken by others and a comparison of the results was made. Incorporating expert judgement posed a challenge in this research, as only a small number of experts were able to provide input. The exida CRD failure modes were used as a basis for the FMEDA but the failure rate data were sourced elsewhere. The failure rate data sources included those from industry equipment failure rate databases, British Standards, and certified equipment reports. Six additional FMEDA assessments were made using the FMEDA worksheet shown previously in Figure 5.5.

Details of the FMEDA Datasets and the Experts involved in the assessments are provided as follows.

- Dataset #1 - the failure modes, component failure rates and failure rate distribution are based on exida CRD [28], and the failure types are based on the judgement of Expert #1, #2 and #3.

- Dataset #2 - the failure modes are based on exida CRD [28], component failure rates are based on FARADIP.FOUR [29] and the judgement of Expert #2, the failure rate distribution is based on the judgement of Expert #2, and the failure types are based on the judgement of Expert #1, #2 and #3.

- Dataset #3 - the failure modes are based on exida CRD [28], component failure rates are based on FARADIP.FOUR [29] and the judgement of Expert #3, the failure rate distribution is based on the judgement of Expert #3, and the failure types are based on the opinions of Expert #1, #2 and #3.

- Dataset #4 - the failure modes are based on exida CRD [28], component failure rates are based on data from BS EN 61508 [32] certified equipment reports and the judgement of Expert #1, the failure rate distribution is based on the judgement of Expert #1, and the failure types are based on the opinions of Expert #1, #2 and #3.

- Dataset #5 - the failure modes are based on exida CRD [28], component failure rates are based on data from BS EN 61508 [32] certified equipment reports, the failure rate distribution is based on the judgement of Expert #1, and the failure types are based on the opinions of Expert #1, #2 and #3.

- Dataset #6 - the failure modes are based on exida CRD [28], component failure rates are based on data from BS EN 61508 [32] certified equipment reports, the failure rate distribution is based on the judgement of Expert #1, and the failure types are based on the judgement of Expert #1, #2 and #3.

- Dataset #7 - the failure modes are based on exida CRD [28], component failure rates and the failure rate distribution are based on data from BS EN 17955 [150], and the failure types are based on the judgement of Expert #1, #2 and #3.

Experts #1, #2 and #3 are registered with the UK Engineering Council as Chartered Engineers and have extensive experience in functional safety within the process industry sector. All experts have more than five years of experience in the analysis of failure modes and effects related to electromechanical equipment.

The exida Component Reliability Database (CRD) [28] provides a comprehensive set of component failure rate, failure mode and useful life data. The datasets have been derived for many engineering applications and provide all the data needed for a realistic FMEDA prediction, a device failure rate prediction method that requires a full set of data. The fifth edition was published in 2021 in a three volume set (Mechanical, Electrical, Electrical Sensor). The data sets are provided for individual components for electrical, mechanical and sensor. The mechanical volume includes contributions from thirteen FMEDA experts.

The FARADIP database [29] is provided by Technis consultancy. FARADIP is a unique failure rate and failure mode data bank, based on over 1000 published data sources together with Technis's own collection of reliability data. More than 50 versions since 1987. FARADIP is widely used as a data reference and provides failure rate data ranges for a nested hierarchy of items that cover electrical, electronic, mechanical, pneumatic, instrumentation and protective devices. Failure mode percentages are included, together with an FMEDA package that calculates the Safe Failure Fraction.

Vendor claims of so-called "data" and FMEDA predicted values, i.e., non-empirical field data, are explicitly excluded from FARADIP. A long-standing Technis study has shown that vendor-based claims of failure rates exhibit over half an order of magnitude of optimism.

Certified equipment reports based on BS EN 61508 [32] are issued by certifying organisations for manufacturers of equipment used in functional safety applications. The reports provide details of the quantitative and qualitative assessments carried out by functional safety experts. Reports can be obtained for all safety instrumented function elements, that is, sensors, logic solvers, and final elements, such as ESDVs. Typically, the components that make up an ESDV, such as SOVs, actuators, and valves, are certified individually. The equipment is assessed from both a hardware safety integrity and a systematic safety integrity perspective, enabling end users to evaluate its viability for use. Hardware safety integrity refers to the ability of hardware components within a safety-related system to perform their intended safety functions reliably, especially in preventing failures that could lead to hazardous situations. It is a measure of the reliability and fault tolerance of the physical components of a system, such as sensors, logic solvers, and final elements, to maintain functional safety over time. Systematic safety integrity refers to the ability of a safety-related system to perform its intended function reliably and consistently, accounting for potential systematic failures. Systematic failures arise from errors in the design and development process, and are not due to random hardware faults. Typically, a report will include a summary of the assessment with failure modes and associated failure rates for the equipment. The reports are easily accessible in the public domain.

BS EN 17955 [150] is a European standard that establishes requirements for evaluating mechanical devices within automated industrial valve assemblies used as final elements in safety instrumented systems (SIS). It also provides a database of mechanical component failure rates for commonly used industrial valve assemblies.

5.4.1 Comparison of Dataset Component Failure Rates

Table 5.5 shows the results of the FMEDA assessments. Each component is identified in conjunction with a specific failure mode and its corresponding failure mode type. For each component failure mode, a failure rate was assigned. The table has been colour coded to provide a 'heat map' to emphasise the difference in the failure rate data. Green indicates the lowest values, yellow the midrange values, and red the highest values. Many components have a wide range of failure rates, e.g., Datasets #2 and Dataset #5 have significant differences for the SOV components, in particular the sleeve bearing, poppet and spring. Similarly, the failure rates for the actuator piston given by Dataset #1 and Dataset #7 are significantly different from all other datasets.

Table 5.5: Variation in component failure rate data (FIT). Green indicates the lowest values, yellow the midrange values, and red the highest values

Component	Failure mode	Failure mode type	Dataset #1	Dataset #2	Dataset #3	Dataset #4	Dataset #5	Dataset #6	Dataset #7
Sleeve bearing	Bind	FTC	22	110	50	56	2	40	10
Poppet w/seals	Bind	FTC	36	220	30	45	1	32	6
Spring	Settle	FTC	104	165	40	45	1	32	16
Spring	Break	FTC	26	55	40	45	1	32	24
Piston rod	Break	FTC	20	19	30	13	11	13	5
Piston rod	Deflect	DOP	20	38	130	26	21	26	5
Yoke (includes guide block/bar)	Break	FTC	36	19	30	13	11	13	15
Yoke (includes guide block/bar)	Bind	DOP	36	150	70	52	43	52	15
Cylinder body	Fracture	FTC	3	19	30	13	11	13	12
Cylinder body	Fracture	DOP	3	19	20	13	11	13	12
Piston	Bind	DOP	8	150	160	104	86	104	16
Piston	Fracture	FTC	4	19	90	13	11	13	8
Piston seals	Bind	FTC	12	75	70	52	43	52	16
Tie rod	Break	FTC	3	19	30	13	11	13	5
Tie rod	Deflect	DOP	3	37	70	26	21	26	5
Spring	Settle	FTC	104	57	70	39	32	39	16
Spring	Break	FTC	26	19	30	39	32	39	24
Stem bush/bearings	Bind	FTC	47	73	43	24	22	7	45
Stem bush/bearings	Bind	DOP	47	73	43	48	44	15	45
Valve body	Bind	DOP	5	19	40	12	11	4	12
Valve body	Bind	FTC	5	19	20	12	11	4	12
Seat ring/spring A	Bind	FTC	23	55	40	24	22	7	6
Seat ring/spring A	Major leak	LCP	45	91	105	121	110	37	15
Seat ring/spring B	Bind	FTC	23	55	40	24	22	7	6
Seat ring/spring B	Major leak	LCP	45	91	105	121	110	37	15
Stem	Bind	FTC	27	73	20	73	66	22	15
Stem	Bind	DOP	27	73	20	73	66	22	15
Stem	Break	LCP	18	73	65	12	11	4	8
Obturator (ball)	Bind	FTC	25	19	43	12	11	4	6
Obturator (ball)	Bind	DOP	25	37	43	24	22	7	6
Obturator (ball)	Break	FTC	25	19	43	12	11	4	6
Obturator (ball)	Major leak	LCP	76	145	65	73	66	22	39
Trunnion bush/bearings	Bind	FTC	47	73	43	48	44	15	15
Trunnion bush/bearings	Bind	DOP	47	73	43	48	44	15	15
Total Dangerous Failure Rate			1021	2248	1808	1369	1040	786	491

Figure 5.8 shows the minimum and maximum failure rates for each component across all datasets. There are a number of component failure rates that differ by two orders of magnitude. The most notable difference in failure rates is the SOV poppet with seals, with a variance of 219 FIT. The significant difference is attributed to Dataset #2, which has a notably high failure rate of 220 FIT, compared to all other datasets, and Dataset #5 which has a notably low failure rate of 1 FIT. The smallest difference in failure rates between datasets is observed in the valve body, with a variation of 16 FIT.

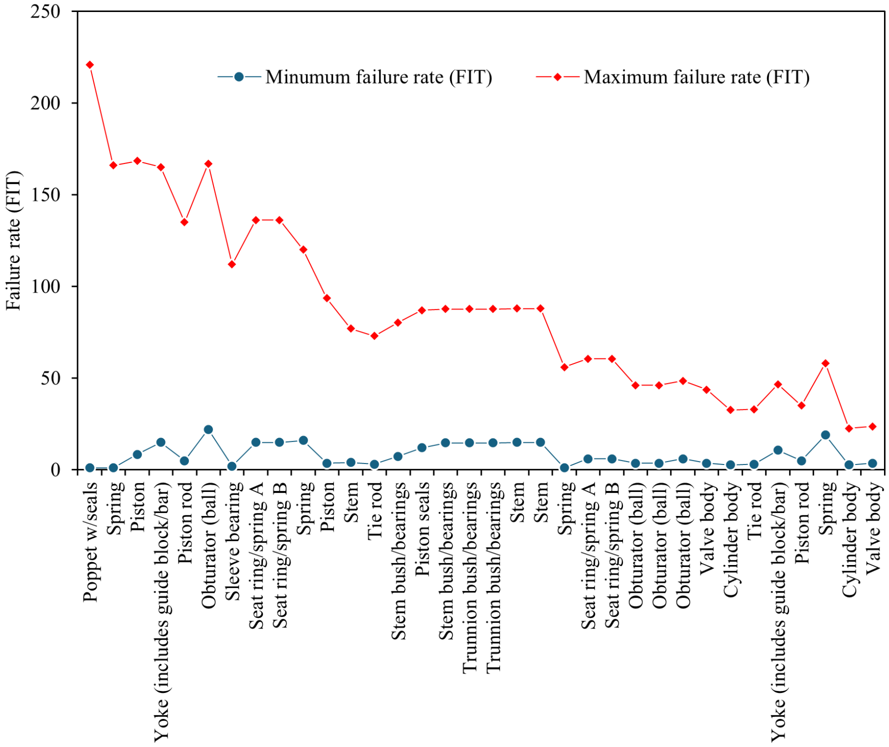
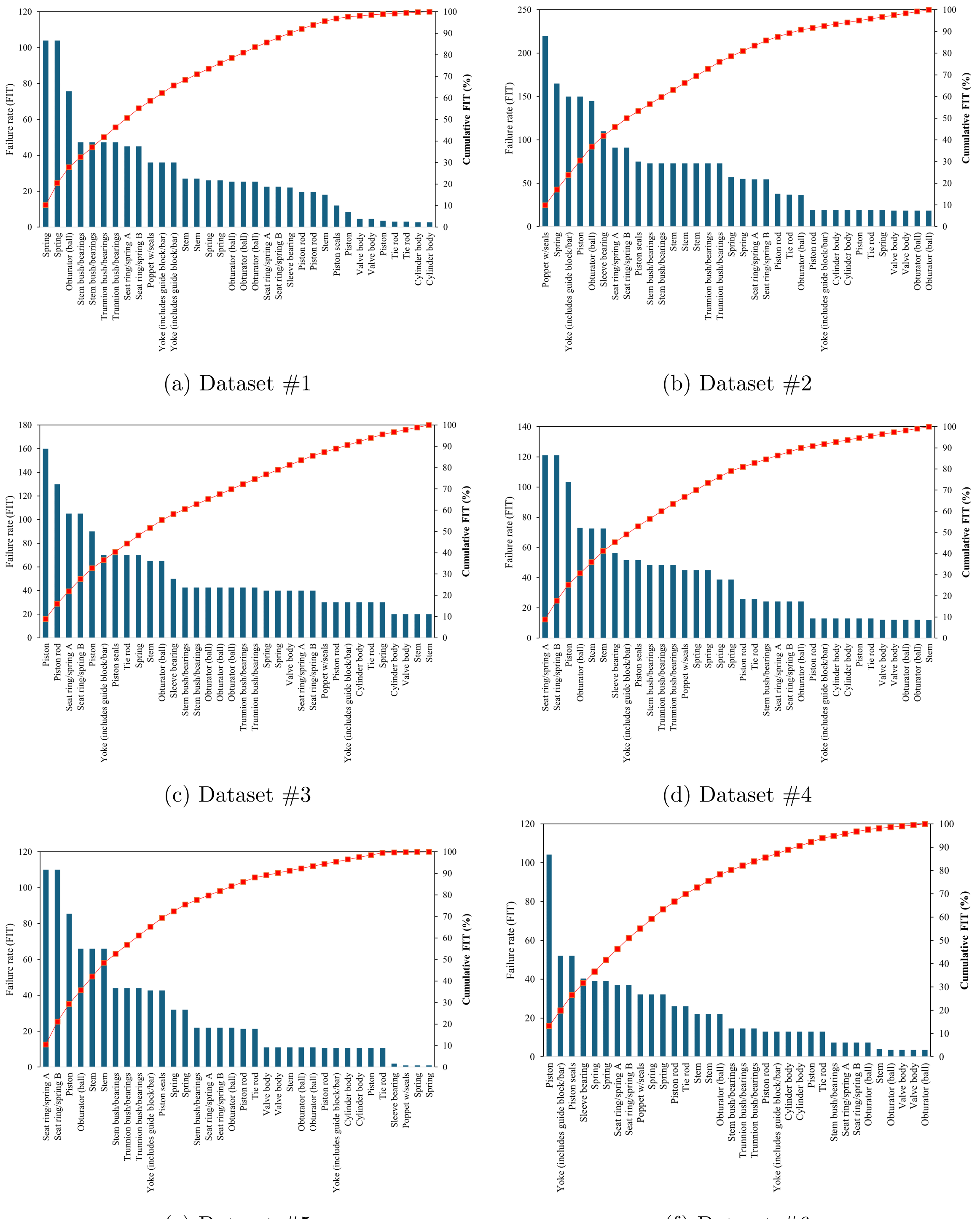


Figure 5.8: Component failures rates - minimum and maximum values across all datasets

Figures 5.9 and Figure 5.10 illustrate the variance of failure rates between datasets for ESDV assembly components in the form of Pareto charts. The charts emphasise the extent of the variation between datasets, including the hierarchy of component failures, as well as the actual component failure rates. There is some agreement, for example, Datasets #4 and #5 identify that the seat ring is the most likely component to fail and Datasets #3 and #4 identify that the valve stem is the least likely component to fail.



(a) Dataset #1

(b) Dataset #2

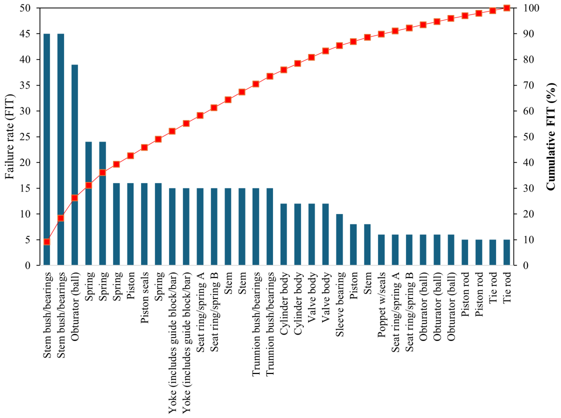
(c) Dataset #3

(d) Dataset #4

(e) Dataset #5

(f) Dataset #6

Figure 5.9: Pareto charts of assumed failure rates for ESDV components - Datasets #1 to #6



(a) Dataset #7

Figure 5.10: Pareto chart of assumed failure rates for ESDV components - Dataset #7

As previously discussed, significant differences exist between the datasets used to perform an FMEDA, which can result in considerable variation in the overall failure rate and proof test coverage outcomes. Tables 5.6, 5.7, and 5.8 present comparisons of expert-derived failure rate datasets, with Table 5.6 using regression analysis (R^2), Table 5.7 employing the Mean Square Error (MSE) and Table 5.8 using the Root Mean Square Error (RMSE) for evaluation.

Table 5.6 indicates that the majority (95%) of the R^2 values are less than 0.5, which indicates a very poor correlation between the data from different experts.

Table 5.6: Regression analysis - comparing component failure rate datasets

Dataset #1 v Dataset #2	**Dataset #1 v Dataset #3**	**Dataset #1 v Dataset #4**	**Dataset #1 v Dataset #5**	**Dataset #1 v Dataset #6**	**Dataset #1 v Dataset #7**
0.228	0.001	0.110	0.046	0.014	0.186
Dataset #2 v Dataset #3	**Dataset #2 v Dataset #4**	**Dataset #2 v Dataset #5**	**Dataset #2 v Dataset #6**	**Dataset #2 v Dataset #7**	
0.075	0.342	0.092	0.315	0.054	
Dataset #3 v Dataset #4	**Dataset #3 v Dataset #5**	**Dataset #3 v Dataset #6**	**Dataset #3 v Dataset #7**		
0.258	0.263	0.410	0.001		
Dataset #4 v Dataset #5	**Dataset #4 v Dataset #6**	**Dataset #4 v Dataset #7**			
0.788	0.432	0.083			
Dataset #5 v Dataset #6	**Dataset #5 v Dataset #7**				
0.224	0.081				
Dataset #6 v Dataset #7					
0.011					

Table 5.7 indicates that the majority (95%) of MSE values are greater than 500 indicating large deviations between datasets.

Table 5.7: MSE - comparing component failure rate datasets

Dataset #1 v Dataset #2	Dataset #1 v Dataset #3	Dataset #1 v Dataset #4	Dataset #1 v Dataset #5	Dataset #1 v Dataset #6	Dataset #1 v Dataset #7
3231	2135	1139	1150	950	750
Dataset #2 v Dataset #3	**Dataset #2 v Dataset #4**	**Dataset #2 v Dataset #5**	**Dataset #2 v Dataset #6**	**Dataset #2 v Dataset #7**	
2815	2312	3719	3626	5033	
Dataset #3 v Dataset #4	**Dataset #3 v Dataset #5**	**Dataset #3 v Dataset #6**	**Dataset #3 v Dataset #7**		
1120	1421	1507	2645		
Dataset #4 v Dataset #5	**Dataset #4 v Dataset #6**	**Dataset #4 v Dataset #7**			
291	811	1504			
Dataset #5 v Dataset #6	**Dataset #5 v Dataset #7**				
743	1033				
Dataset #6 v Dataset #7					
532					

Table 5.8 shows that the actual mean error between the expert datasets ranges from 17 FIT to 71 FIT, confirming the poor relationships between the expert datasets.

Table 5.8: RMSE (FIT) - comparing component failure rate datasets

Dataset #1 v Dataset #2	Dataset #1 v Dataset #3	Dataset #1 v Dataset #4	Dataset #1 v Dataset #5	Dataset #1 v Dataset #6	Dataset #1 v Dataset #7
57	46	34	34	31	27
Dataset #2 v Dataset #3	**Dataset #2 v Dataset #4**	**Dataset #2 v Dataset #5**	**Dataset #2 v Dataset #6**	**Dataset #2 v Dataset #7**	
53	48	61	60	71	
Dataset #3 v Dataset #4	**Dataset #3 v Dataset #5**	**Dataset #3 v Dataset #6**	**Dataset #3 v Dataset #7**		
33	38	39	51		
Dataset #4 v Dataset #5	**Dataset #4 v Dataset #6**	**Dataset #4 v Dataset #7**			
17	28	39			
Dataset #5 v Dataset #6	**Dataset #5 v Dataset #7**				
27	32				
Dataset #6 v Dataset #7					
23					

The above metrics indicate very poor regression and substantial deviations between datasets. Table 5.9 shows the average and geometric mean values for the datasets.

Table 5.9: Average and geometric mean values of datasets

	R^2	MSE	RMSE
Average	0.191	1832	40
Geometric mean	0.081	1455	38

Figure 5.11 shows the variance in the datasets, the ESDV failure rates that range from approximately 500 to 2200 FIT.

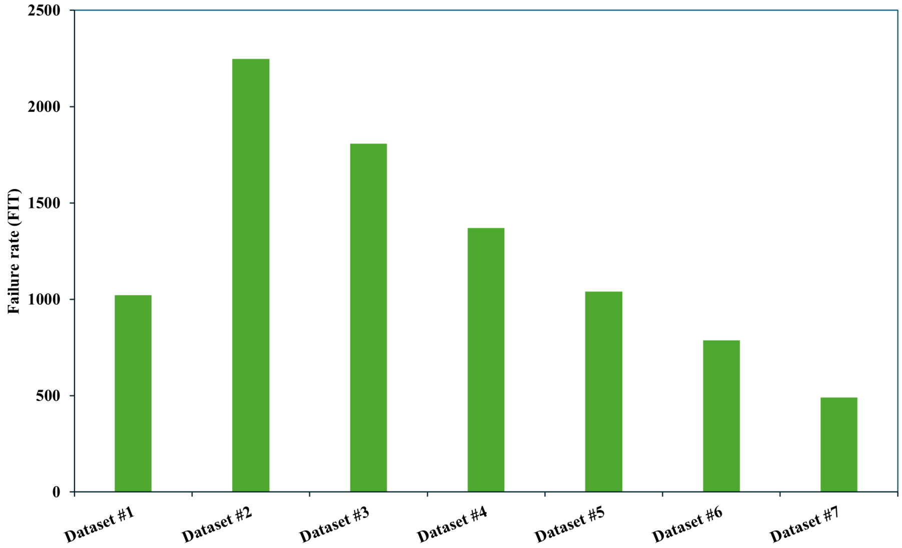


Figure 5.11: ESDV total failure rate for each dataset

Table 5.10 shows the variation in proof test coverage calculated from the weighted revealed failure rates and total failure rates, with a maximum difference being 16% between Dataset #5 and Dataset #6.

Table 5.10: Proof test coverage estimates using expert datasets

Dataset #1	Dataset #2	Dataset #3	Dataset #4	Dataset #5	Dataset #6	Dataset #7
82%	82%	81%	77%	71%	87%	84%

5.5 Discussion and Conclusions

The Proof Test Coverage (PTC) estimates derived from FMEDA show significantly smaller differences between the datasets compared to the variations in failure rates. As discussed previously, PTC represents the percentage of dangerous failures that are detected relative to the total dangerous failure rate. Therefore, regardless of the general magnitude of the failure rate, the PTC is determined by the proportion of failures revealed within a specific dataset. Thus, it is possible for a dataset with a significantly higher overall failure rate to have a comparable PTC estimate as a dataset with a much lower overall failure rate, as the PTC is determined by the proportion of revealed failures, not the total failure rate itself. An example of this can be seen with Dataset #2 and Dataset #7. Furthermore, it is possible for two datasets with comparable total failure rates to have different PTC estimates, as seen with Dataset #1 and Dataset #5. This variation occurs because the PTC depends on the proportion of revealed dangerous failures, not just the total failure rate.

The impact of the findings is significant if we compare the PTC results with those of published data, in particular the FMEDA PTC estimate of 70% provided by Stewart [47]. Dataset #5 provides a comparable result, while Dataset #6 shows an approximately 17% increase, indicating a more favourable PFD_{avg} and potentially leading to less frequent proof testing, thus reducing maintenance overhead.

The FMEDA results indicate that the only part of the ESDV assembly that has unrevealed failures is the ball valve itself, suggesting that the SOV and actuator can be excluded when estimating proof test coverage for a full stroke test. Leak in the Closed Position (LCP) being the only dangerous failure mode that is not revealed during a full stroke proof test and the failure causes are attributed to the ball valve itself.

The scope of FMEDA assessments is limited to the primary elements of an ESDV and typically excludes ancillary equipment, such as pressure filter regulators, pilot valves, quick exhaust valves, and flow regulators, even though these components are often specified in ESDV designs. Ancillary equipment has the potential to introduce additional dangerous failure modes. For example, failure of a pressure-filter regulator could lead to excessive pressure being applied to the actuator, thereby increasing actuator torque and potentially over-stressing the valve stem. This over-stress could result in the stem twisting or even shearing. However, actuator specifications typically take into account the valve's Maximum Allowable Stem Torque (MAST), which is determined based on the regulated air pressure supplied to the actuator, typically 4 barg, regulated down from an instrument air supply of 8.5 barg. MAST is defined as the maximum torque/thrust that can be applied to the valve stem without risk of damage [88].

5.5.1 Summary of Findings

The key findings of this chapter can be briefly summarised in the following points:

- The FMEDA process is complex, requiring numerous inputs such as component failure rate data and failure modes, which are challenging to acquire due to the lack of a standardised database.

- The component failure rates vary significantly across datasets showing poor regression, Mean Square Error (MSE), and Root Mean Square Error (RMSE) metrics when compared against each other. The actual mean error between expert datasets ranges from 17 FIT to 71 FIT.

- The corresponding proof test coverage estimates also showed considerable variation between the datasets, although not as pronounced as the failure rate data, ranging from 71% to 87% for a full stroke test.

- The FMEDA PTC estimates are comparable to the estimates obtained from industry data in Chapter 4. The average estimates of PTC for FMEDA and industry data are 80% and 84%, respectively.

The results of this chapter demonstrate that conducting an FMEDA of an ESDV assembly reveals differing opinions on component failure rates among industry experts and data sources. The differences in the datasets cause significant variation in the estimation of proof test coverage. To address the differences in opinion, additional work is required to reconcile the experts' judgements, which is the subject of the next chapter. The objective of the next chapter is to develop a Fuzzy Inference System (FIS) that consolidates expert judgement, validated using datasets from various sources. The performance of the FIS will be evaluated through statistical measures and metrics, including linear regression analysis (R^2), Mean Square Error (MSE), and Root Mean Square Error (RMSE).

Chapter 6

Method 3 - Estimating PTC using a Fuzzy Inference System

6.1 Introduction

As discussed in the previous chapter, the input data used in an ESDV Failure Modes, Effects, and Diagnostic Analysis (FMEDA) is subjective, derived from the interpretation of experts on the failure of ESDV components, resulting in a variety of different assessment outcomes. The subjective data relates both to component failure rates and the probability of revealing a failure during a proof test. The latter is particularly significant when assessing proof test coverage for a partial stroke test, as it relies on assumptions about the full movement of a valve based only on a partial stroke, making it less definitive as a test. As identified in the literature review Fuzzy Inference Systems (FIS) are suitable for applications where data is subjective and have been used in Failure Modes and Effects Analysis (FMEA) type applications such as Failure Modes, Effects, and Criticality Analysis (FMECA) but not yet with the use of FMEDA.

This chapter discusses an approach to handling subjective data, such as component failure rate and probability of revealing a failure, using fuzzy logic, and the methodology used to develop a suitable FIS to infer a weighted revealed failure rate to allow calculation of proof test coverage. The objective is to design an FIS that models the real world; in this case an FMEDA based on its input data, i.e., component failure rate and probability of revealing a failure. Figure 6.1 provides an overview of the structure of this chapter.

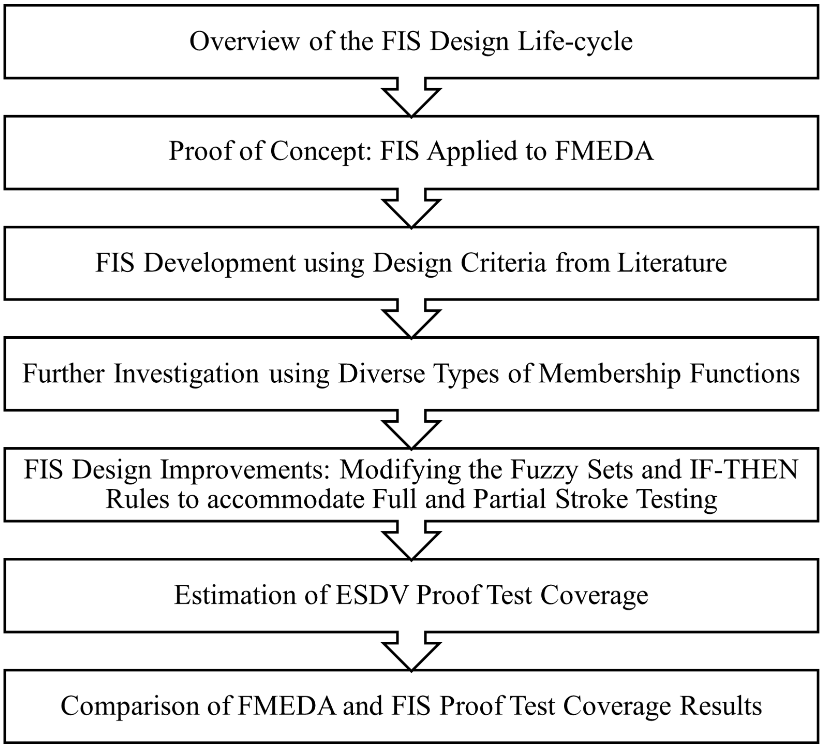

Overview of the FIS Design Life-cycle
Proof of Concept: FIS Applied to FMEDA
FIS Development using Design Criteria from Literature
Further Investigation using Diverse Types of Membership Functions
FIS Design Improvements: Modifying the Fuzzy Sets and IF-THEN Rules to accommodate Full and Partial Stroke Testing
Estimation of ESDV Proof Test Coverage
Comparison of FMEDA and FIS Proof Test Coverage Results

Figure 6.1: Chapter Six Structure

6.2 FIS Design Life-cycle

6.2.1 Development Phases

The fuzzy inference system was designed using five phases of development. Figure 6.2 shows the fuzzy inference system design life-cycle and the development phases from conceptualisation through to validation.

The first phase is conceptualisation, where the problem is acknowledged and the data, along with opportunities for improvement in the FMEDA technique, are identified. The theoretical framework of the potential solution is developed and a literature review is undertaken to identify a fuzzy inference design criteria.

The second phase is the development of the fuzzy inference system design based on the findings of the literature review, highlighting the most common FIS design criteria. Criteria such as input and output membership function types, universe of discourse (based on expert data), partitioning, fuzzy set overlap, rule base, and method of defuzzification. FMEDA data is also collected from experts during this phase which provides a beneficial input to the design development.

The third phase is the realisation of the FIS design in software. Initial FIS designs were developed using JuzzyOnline [151], an open source online toolkit for fuzzy inference systems. MATLAB software (Fuzzy Logic Toolbox) was used in the latter stages of the life-cycle to refine the FIS design.

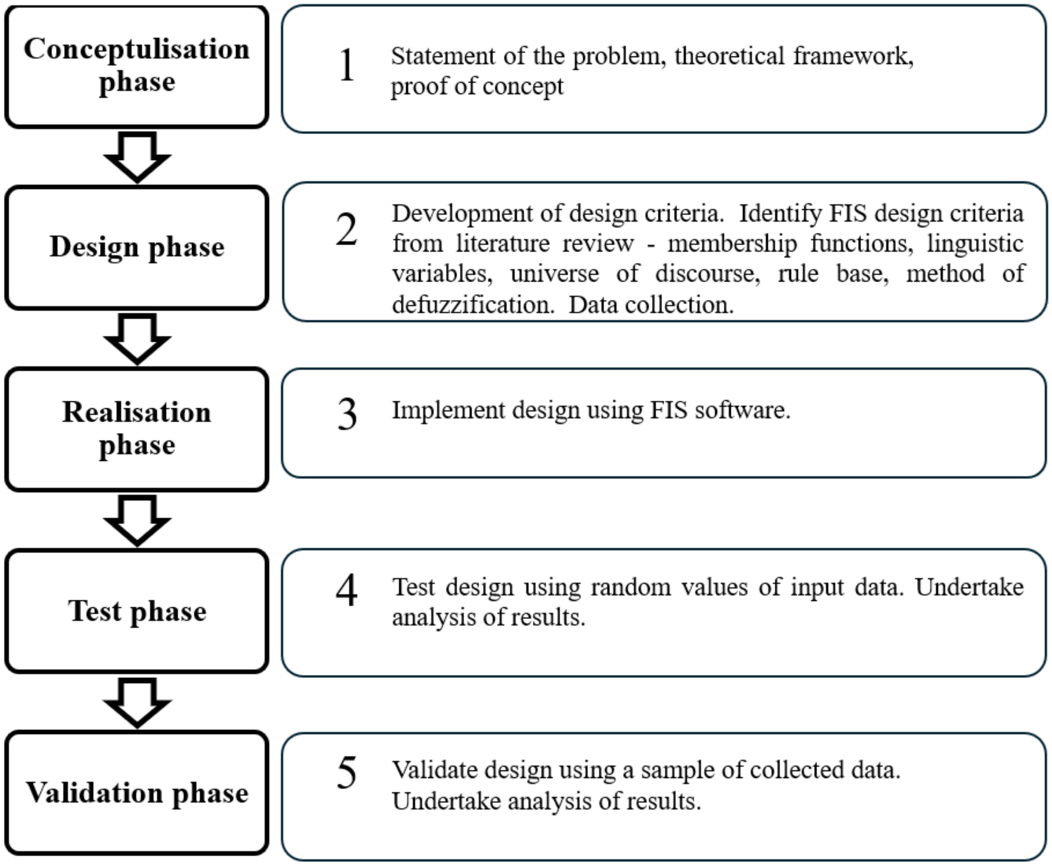

Conceptulisation phase
1
Statement of the problem, theoretical framework, proof of concept
Design phase
2
Development of design criteria. Identify FIS design criteria from literature review - membership functions, linguistic variables, universe of discourse, rule base, method of defuzzification. Data collection.
Realisation phase
3
Implement design using FIS software.
Test phase
4
Test design using random values of input data. Undertake analysis of results.
Validation phase
5
Validate design using a sample of collected data. Undertake analysis of results.

Figure 6.2: FIS Design Life-cycle

The fourth and fifth phases are testing and validation, respectively. Tests were carried out using random number datasets for both inputs of the FIS, these were random numbers within the universe of discourse of the inputs. The validation was carried out using datasets from various sources and includes expert judgement. The goal was to identify an FIS that best represents the opinion of experts within the scope of this research [128].

6.2.2 Phase 1 – Conceptualisation

6.2.2.1 Proof of Concept - FIS Applied to FMEDA

A preliminary demonstration to verify the feasibility of using the FIS for use with FMEDA data was undertaken with the aims to show that the FIS can effectively address the problem before moving forward with full-scale development or implementation. Initial FIS designs were developed including setting up membership functions, input variables such as component failure rates and probabilities, and the rule base needed to perform inference. Testing of the FIS was undertaken on a limited dataset to demonstrate that it can provide reasonable outputs, to calculate proof test coverage estimates. The outputs of the FIS were then compared with FMEDA outputs to assess if it offers comparable or better results, to validate whether the FIS is a viable tool for the problem.

Fuzzy inference research papers relating to FMEA and FMECA were reviewed, noting that fuzzy inference applied to FMEDA is a novel idea, hence no such literature is available. Fuzzy inference systems have been successfully applied to FMEA and FMECA processes, making them valuable candidates for testing in FMEDA applications.

Initially, a fuzzy inference design was developed using JuzzyOnline [151], an open source online toolkit for developing fuzzy inference systems (FIS). The aim of the FIS design is to model the FMEDA process and its associated data input. The performance of the FIS is determined by a number of attributes, including the type and size of the membership function, the base of rules, and the method of defuzzification.

As discussed earlier, Mamdani Type-1 is the preferred FIS method over Mamdani Type-2 and TSK due to its simplicity. Mamdani Type-1 is the most commonly used type in the FMEA and FMECA applications and is the focus of this research. The FIS design was developed and tested against random numbers and validated against a single FMEDA dataset, experimenting with different fuzzy set shapes and configurations, a 'trial and error' method, to see what results could be obtained and if the idea was worth pursuing.

The initial design was based on three inputs; component failure rate, revealability (of the failure mode) and the reliability (of the failure mode detection hardware). The output was referred to as the ‘Proof Test Coverage (PTC) significance’. The design used the most simplistic membership function, the triangular type, with five, equally partitioned, fuzzy sets. A fuzzy set is defined by a membership function that maps each element in the universe of discourse to a degree of membership. The partitions were labelled Very Low (VL), Low (L), Medium (M), High (H), and Very High (VH). The input fuzzy sets are shown in Figure 6.3 and the output in Figure 6.4.

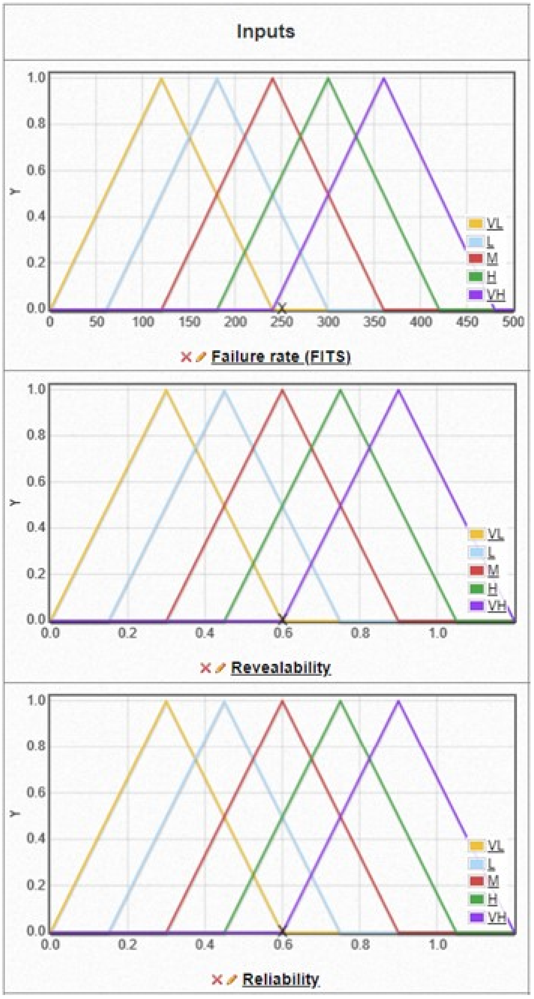


Figure 6.3: Initial FIS design - Input fuzzy sets

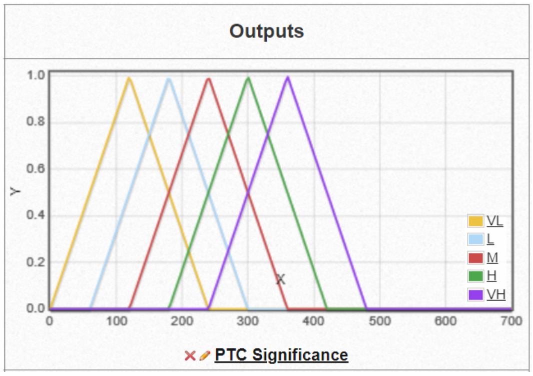
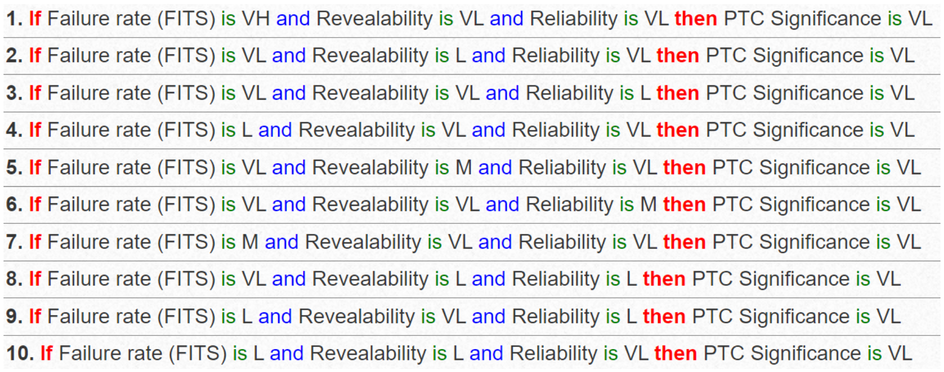


Figure 6.4: Initial FIS design - Output fuzzy set

As there were three input fuzzy sets with each having five membership functions, this generated an IF-THEN rule base of 125 rules. Figure 6.5 shows a sample of the rules.



1. If Failure rate (FITS) is VH and Revealability is VL and Reliability is VL then PTC Significance is VL
2. If Failure rate (FITS) is VL and Revealability is L and Reliability is VL then PTC Significance is VL
3. If Failure rate (FITS) is VL and Revealability is VL and Reliability is L then PTC Significance is VL
4. If Failure rate (FITS) is L and Revealability is VL and Reliability is VL then PTC Significance is VL
5. If Failure rate (FITS) is VL and Revealability is M and Reliability is VL then PTC Significance is VL
6. If Failure rate (FITS) is VL and Revealability is VL and Reliability is M then PTC Significance is VL
7. If Failure rate (FITS) is M and Revealability is VL and Reliability is VL then PTC Significance is VL
8. If Failure rate (FITS) is VL and Revealability is L and Reliability is L then PTC Significance is VL
9. If Failure rate (FITS) is L and Revealability is VL and Reliability is L then PTC Significance is VL
10. If Failure rate (FITS) is L and Revealability is L and Reliability is VL then PTC Significance is VL

Figure 6.5: Sample of IF-THEN rules

The design was tested using a set of random numbers, resulting in an R^2 value of approximately 0.7, a high MSE of 5956, and an RMSE in the region of 77 FIT. These metrics indicate that the model provides a poor fit to the data. Figure 6.6 shows a comparison of FMEDA output and FIS output values.

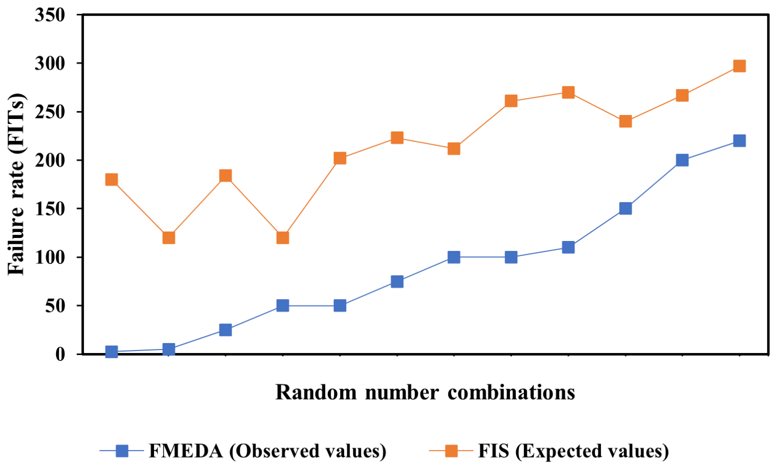


Figure 6.6: Initial FIS design - a comparison of FMEDA (Observed values) and FIS (Expected values) outputs

The partitioning of the fuzzy sets was modified in an attempt to improve the fit of the FIS to the FMEDA data. Figures 6.7 and Figure 6.8 show modifications of the input membership functions. Figure 6.7 shows unequal partitioning to improve the sensitivity at the lower end of the range.

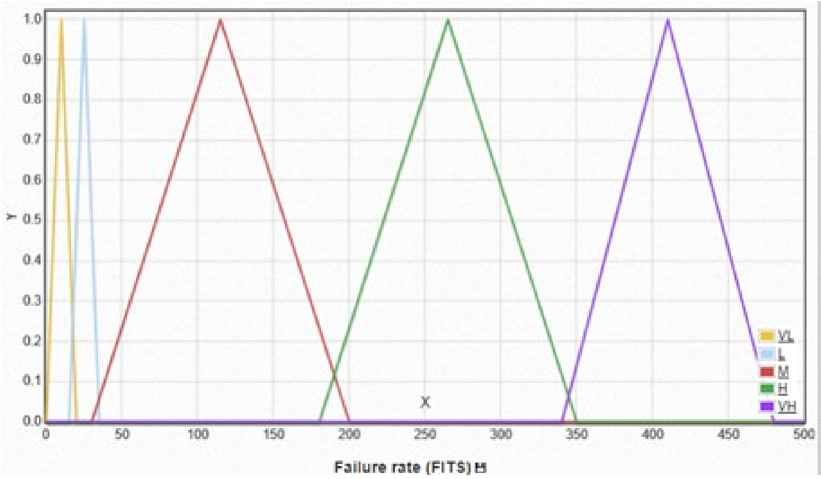


Figure 6.7: Modified input membership function (unequal partitioning to improve sensitivity at the lower end of the range)

Figure 6.8 shows fuzzy sets with reduced overlap that did not improve the performance of the system.

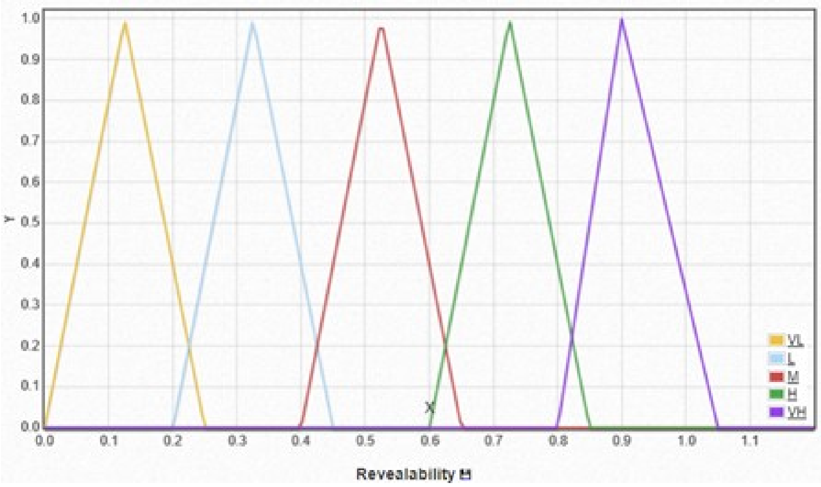


Figure 6.8: Modified input membership function (reduced overlap of fuzzy sets)

Modifications to the output membership functions were also experimented with as shown in Figure 6.9. This particular configuration did not improve the system performance.

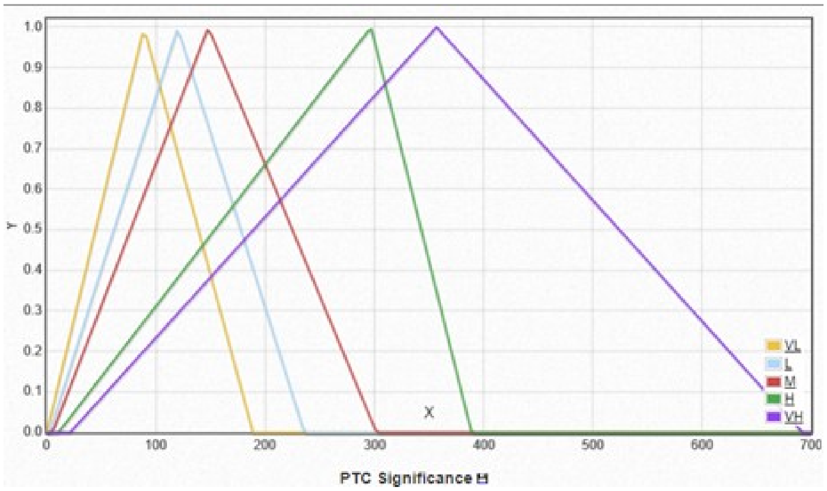


Figure 6.9: Modified output membership function (unequal partitioning)

Several iterations of the design improved the performance of the FIS. A key change was the number of inputs; this was decreased from three to two. Originally the inputs were component failure rate, revealability and reliability, but from testing it was evident that the system was too complex with three inputs which generated 125 rules. Therefore, revealability and reliability were combined to make one input called Detection later renamed as the Probability of revealing the failure. Figure 6.10 shows the modifications reduced the number of inputs, the change in the universe of discourse, and the shape of the fuzzy set, all improving the performance of the FIS.

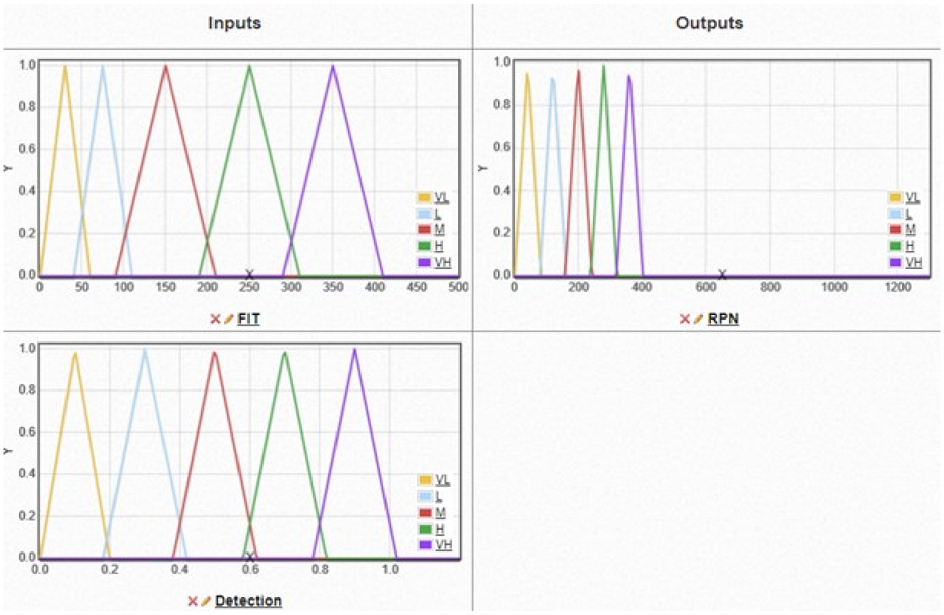


Figure 6.10: Reduced number of inputs and universe of discourse

Reducing the number of inputs meant that the rule base was much smaller with only 25 which was far more manageable. The universe of discourse was also modified to coincide with actual data set failure rate ranges, i.e., 1 to 220 FIT, mindful that the universe of discourse could change in the event of expanding the model with new data sets. Different shapes of fuzzy sets were used favouring trapezoidal for the probability of revealing failure input and weighted revealed failure rate (output).

The optimised proof-of-concept final design is shown in the following figures. The inputs, component failure rate and probability of revealing the failure and the output, weighted revealed failure rate, are shown in Figures 6.11, 6.12, and 6.13, respectively.

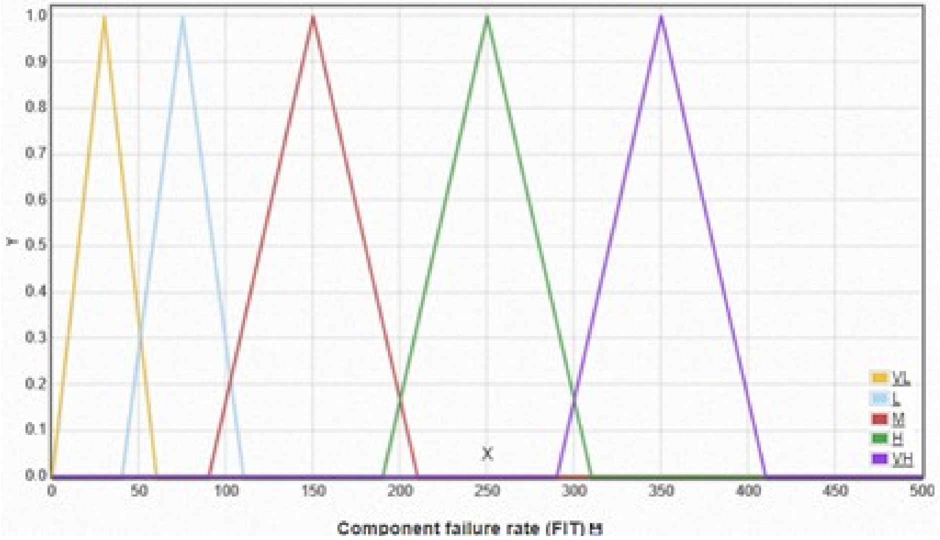


Figure 6.11: Component failure rate

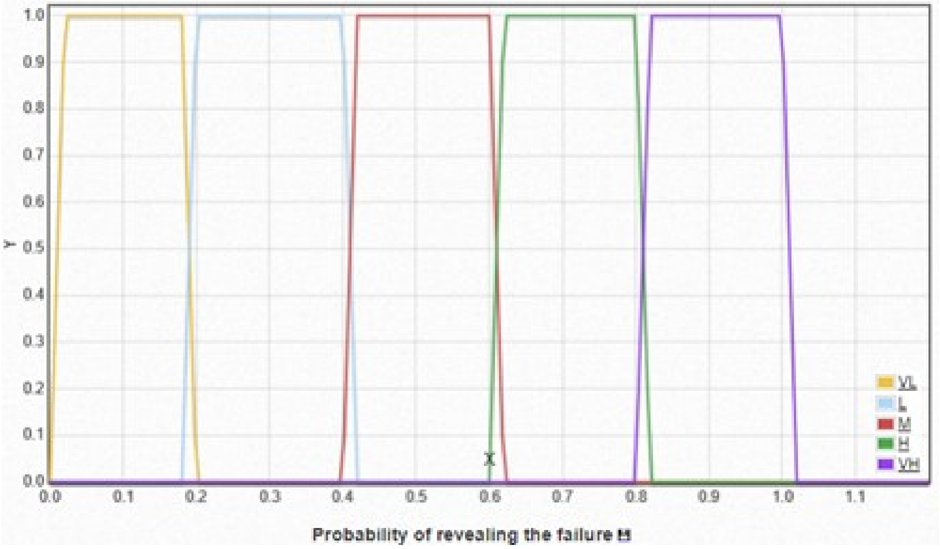


Figure 6.12: Probability of revealing the failure

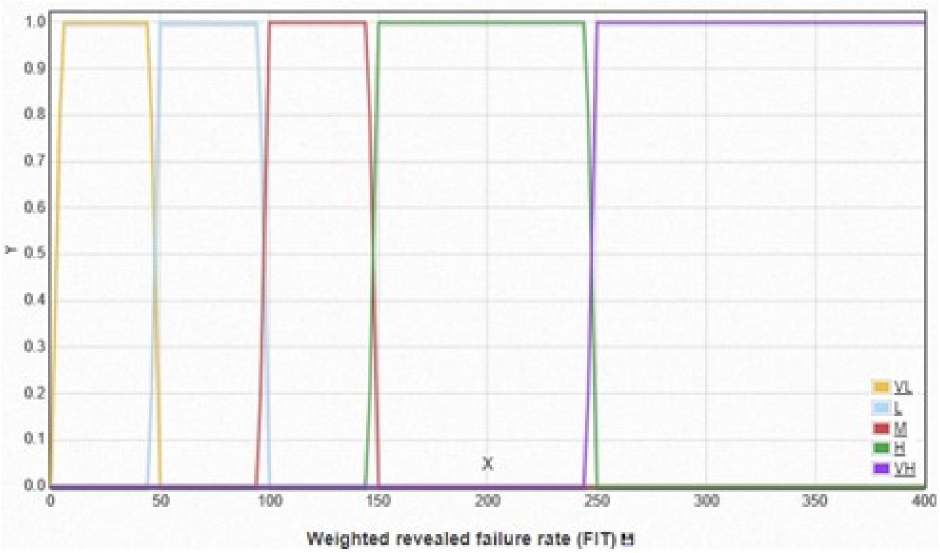
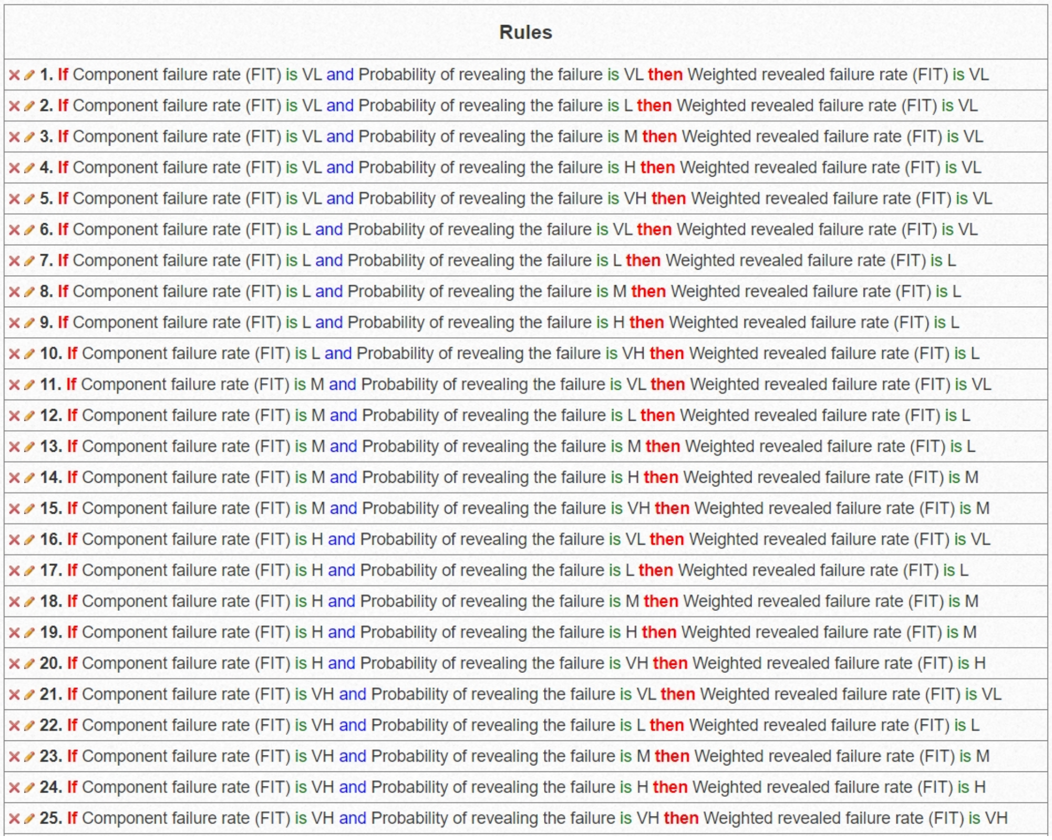


Figure 6.13: Weighted revealed failure rate

The IF-THEN rule base for the final design is shown in Figure 6.14.



Rules
1. If Component failure rate (FIT) is VL and Probability of revealing the failure is VL then Weighted revealed failure rate (FIT) is VL
2. If Component failure rate (FIT) is VL and Probability of revealing the failure is L then Weighted revealed failure rate (FIT) is VL
3. If Component failure rate (FIT) is VL and Probability of revealing the failure is M then Weighted revealed failure rate (FIT) is VL
4. If Component failure rate (FIT) is VL and Probability of revealing the failure is H then Weighted revealed failure rate (FIT) is VL
5. If Component failure rate (FIT) is VL and Probability of revealing the failure is VH then Weighted revealed failure rate (FIT) is VL
6. If Component failure rate (FIT) is L and Probability of revealing the failure is VL then Weighted revealed failure rate (FIT) is VL
7. If Component failure rate (FIT) is L and Probability of revealing the failure is L then Weighted revealed failure rate (FIT) is L
8. If Component failure rate (FIT) is L and Probability of revealing the failure is M then Weighted revealed failure rate (FIT) is L
9. If Component failure rate (FIT) is L and Probability of revealing the failure is H then Weighted revealed failure rate (FIT) is L
10. If Component failure rate (FIT) is L and Probability of revealing the failure is VH then Weighted revealed failure rate (FIT) is L
11. If Component failure rate (FIT) is M and Probability of revealing the failure is VL then Weighted revealed failure rate (FIT) is VL
12. If Component failure rate (FIT) is M and Probability of revealing the failure is L then Weighted revealed failure rate (FIT) is L
13. If Component failure rate (FIT) is M and Probability of revealing the failure is M then Weighted revealed failure rate (FIT) is L
14. If Component failure rate (FIT) is M and Probability of revealing the failure is H then Weighted revealed failure rate (FIT) is M
15. If Component failure rate (FIT) is M and Probability of revealing the failure is VH then Weighted revealed failure rate (FIT) is M
16. If Component failure rate (FIT) is H and Probability of revealing the failure is VL then Weighted revealed failure rate (FIT) is VL
17. If Component failure rate (FIT) is H and Probability of revealing the failure is L then Weighted revealed failure rate (FIT) is L
18. If Component failure rate (FIT) is H and Probability of revealing the failure is M then Weighted revealed failure rate (FIT) is M
19. If Component failure rate (FIT) is H and Probability of revealing the failure is H then Weighted revealed failure rate (FIT) is M
20. If Component failure rate (FIT) is H and Probability of revealing the failure is VH then Weighted revealed failure rate (FIT) is H
21. If Component failure rate (FIT) is VH and Probability of revealing the failure is VL then Weighted revealed failure rate (FIT) is VL
22. If Component failure rate (FIT) is VH and Probability of revealing the failure is L then Weighted revealed failure rate (FIT) is L
23. If Component failure rate (FIT) is VH and Probability of revealing the failure is M then Weighted revealed failure rate (FIT) is M
24. If Component failure rate (FIT) is VH and Probability of revealing the failure is H then Weighted revealed failure rate (FIT) is H
25. If Component failure rate (FIT) is VH and Probability of revealing the failure is VH then Weighted revealed failure rate (FIT) is VH

Figure 6.14: IF-THEN rule base

The results of the testing and validation are shown in Table 6.1:

Table 6.1: Results from test and validation

	Random Numbers	Expert Dataset
R^2	0.94	0.92
MSE	222	90
RMSE (FIT)	14.9	9.5

The optimised FIS design comparing the FMEDA and FIS outputs using random numbers is shown in Figure 6.15.

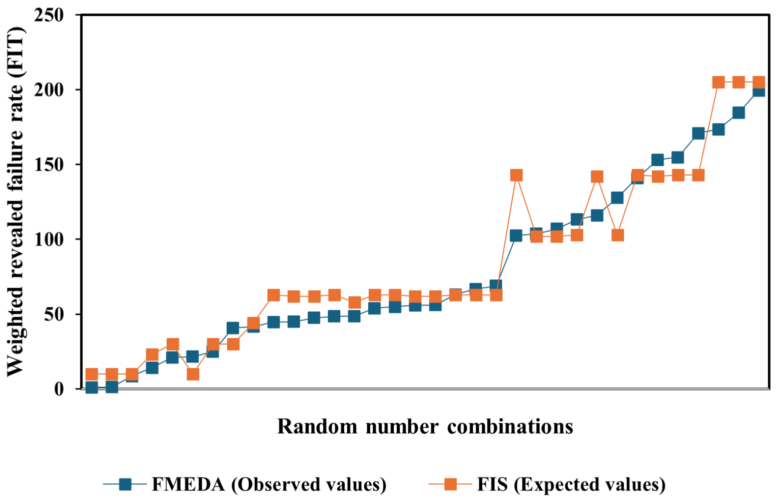


Figure 6.15: Optimised FIS design showing a comparison of FMEDA and FIS output values

The results of the proof-of-concept showed that it was feasible to design an FIS capable of modelling the FMEDA process. It is important to note that this initial FIS is trained on a single dataset and that any model's effectiveness is inherently dependent on the quality of the data it receives.

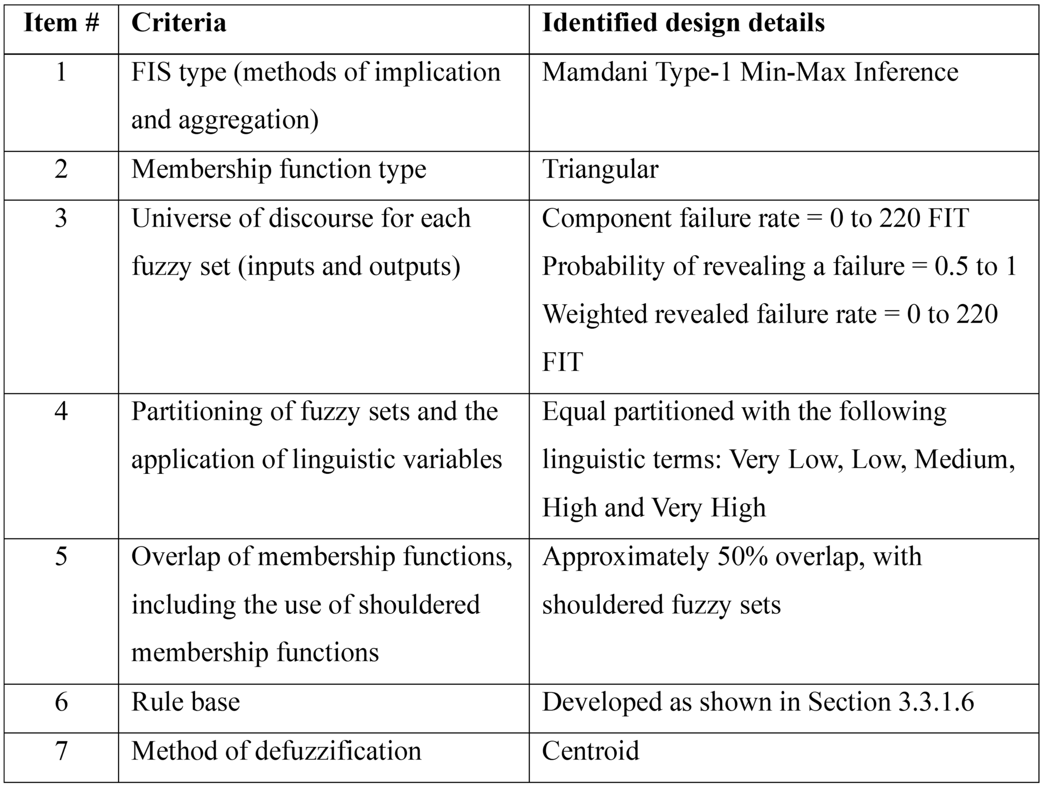
Given that expert opinions vary significantly, this fitting process is likely to perform well only within the limits of the specific dataset. To create a robust model that can be applied more broadly, a consolidation of opinions is necessary to ensure that it extends beyond the expert views presented here. This will be addressed in a later section, but the next step involves refining the model.

More work is required to enhance the performance of the FIS with a more in-depth review of relevant research papers. Following the proof-of-concept exercise an FIS design was developed based on scientific research papers relating to fuzzy inference systems applied to FMEA type applications such as FMECA. The design criteria from the literature review is discussed in the next section.

6.2.3 Phase 2 – Development of an FIS Design Criteria Based on Literature

Table 6.2 shows a summary of the FIS design identified from a review of literature from Section 3.4.

Table 6.2: Identified FIS design criteria from literature



Item #	Criteria	Identified design details
1	FIS type (methods of implication and aggregation)	Mamdani Type-1 Min-Max Inference
2	Membership function type	Triangular
3	Universe of discourse for each fuzzy set (inputs and outputs)	Component failure rate = 0 to 220 FIT Probability of revealing a failure = 0.5 to 1 Weighted revealed failure rate = 0 to 220 FIT
4	Partitioning of fuzzy sets and the application of linguistic variables	Equal partitioned with the following linguistic terms: Very Low, Low, Medium, High and Very High
5	Overlap of membership functions, including the use of shouldered membership functions	Approximately 50% overlap, with shouldered fuzzy sets
6	Rule base	Developed as shown in Section 3.3.1.6
7	Method of defuzzification	Centroid

The following sections refer to the revised FIS design as the Triangular 5x5 design.

6.2.4 Phase 3 – Implementation of the Revised Design Using JuzzyOnline

The fuzzy inference system was implemented using the JuzzyOnline FIS online toolbox [151]. The input and output membership functions, and the Rule base as shown in Figures 6.16, 6.17, Figures 6.18 and Figure 6.19 respectively.

6.2.4.1 Input membership functions:

1. **Component failure rate (FIT)**

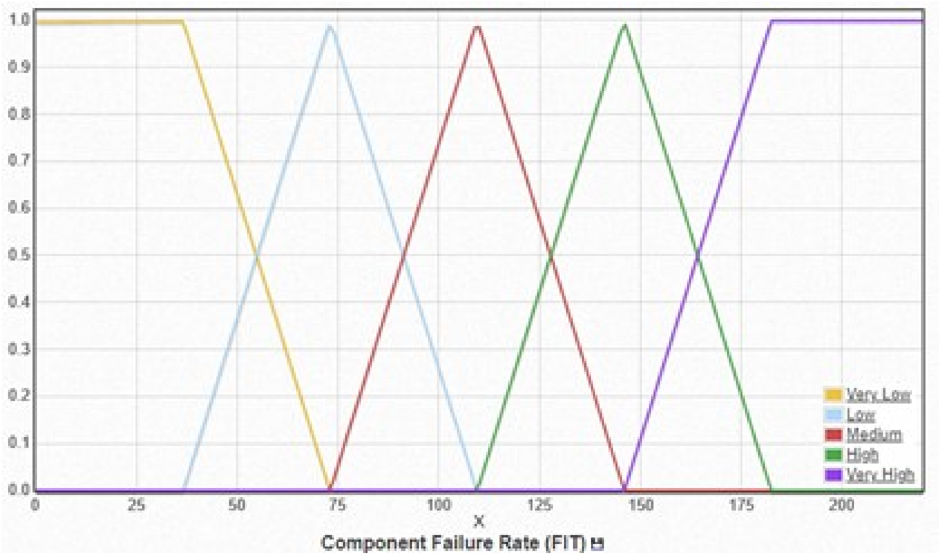


Figure 6.16: Component failure rate (FIT) membership function

2. **Probability of revealing a failure**

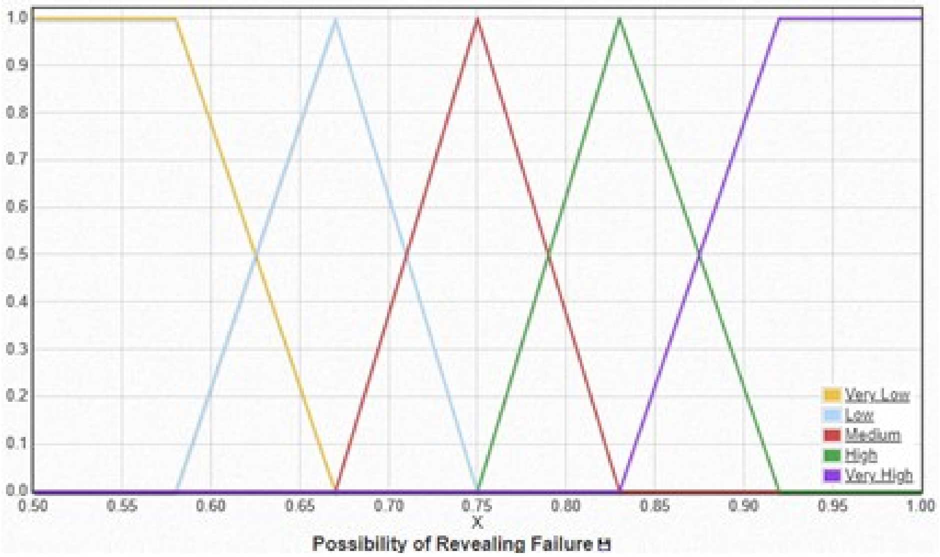


Figure 6.17: Probability of revealing a failure membership function

6.2.4.2 Output membership functions:

Weighted revealed failure rate (FIT)

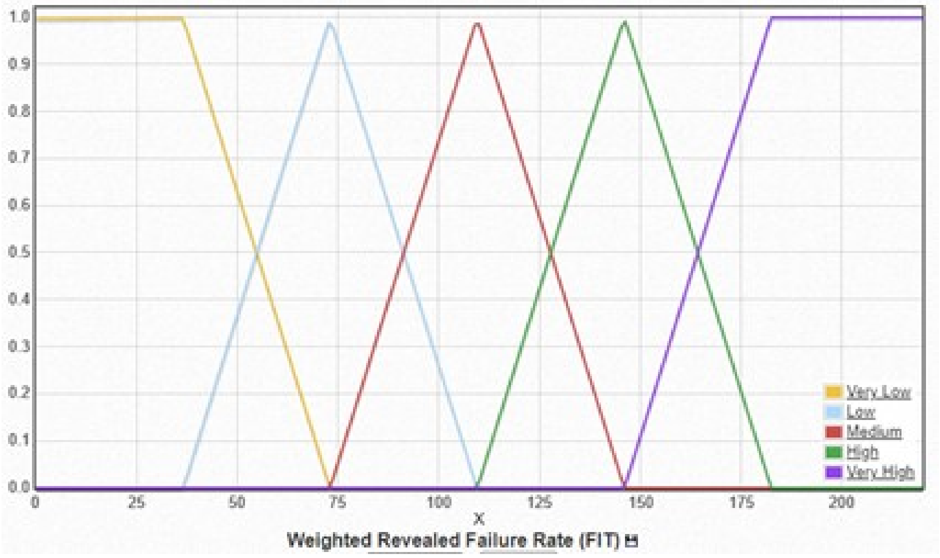
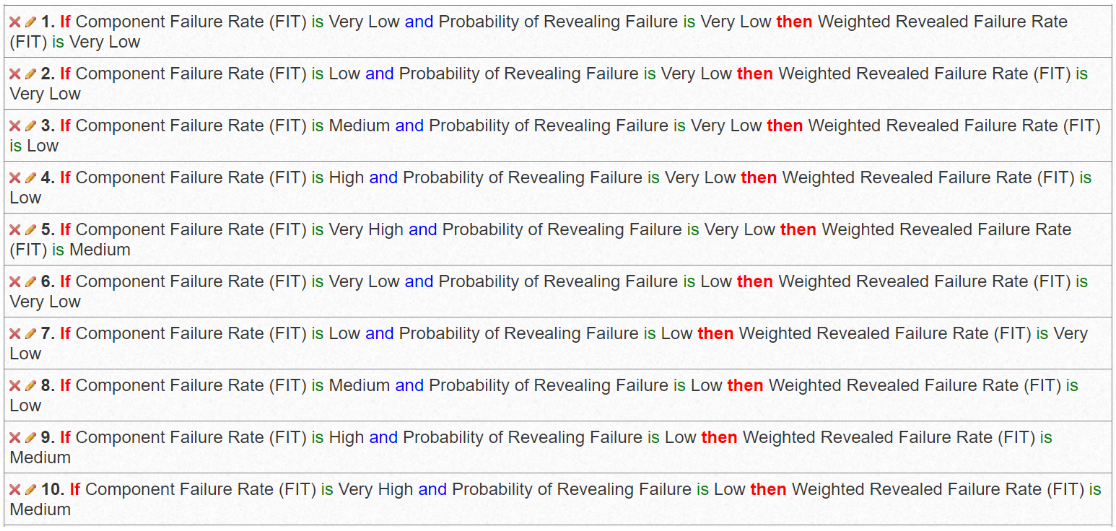


Figure 6.18: Weighted revealed failure rate (FIT) membership function

6.2.4.3 Rule base

Figure 6.19 shows a sample of the IF-THEN rule base, extracted from JuzzyOnline [151]. The full set of rules are shown in Appendix D.1:



✕✎ **1.** If Component Failure Rate (FIT) is Very Low and Probability of Revealing Failure is Very Low **then** Weighted Revealed Failure Rate (FIT) is Very Low

✕✎ **2.** If Component Failure Rate (FIT) is Low and Probability of Revealing Failure is Very Low **then** Weighted Revealed Failure Rate (FIT) is Very Low

✕✎ **3.** If Component Failure Rate (FIT) is Medium and Probability of Revealing Failure is Very Low **then** Weighted Revealed Failure Rate (FIT) is Low

✕✎ **4.** If Component Failure Rate (FIT) is High and Probability of Revealing Failure is Very Low **then** Weighted Revealed Failure Rate (FIT) is Low

✕✎ **5.** If Component Failure Rate (FIT) is Very High and Probability of Revealing Failure is Very Low **then** Weighted Revealed Failure Rate (FIT) is Medium

✕✎ **6.** If Component Failure Rate (FIT) is Very Low and Probability of Revealing Failure is Low **then** Weighted Revealed Failure Rate (FIT) is Very Low

✕✎ **7.** If Component Failure Rate (FIT) is Low and Probability of Revealing Failure is Low **then** Weighted Revealed Failure Rate (FIT) is Very Low

✕✎ **8.** If Component Failure Rate (FIT) is Medium and Probability of Revealing Failure is Low **then** Weighted Revealed Failure Rate (FIT) is Low

✕✎ **9.** If Component Failure Rate (FIT) is High and Probability of Revealing Failure is Low **then** Weighted Revealed Failure Rate (FIT) is Medium

✕✎ **10.** If Component Failure Rate (FIT) is Very High and Probability of Revealing Failure is Low **then** Weighted Revealed Failure Rate (FIT) is Medium

Figure 6.19: Sample of the IF-THEN rule base

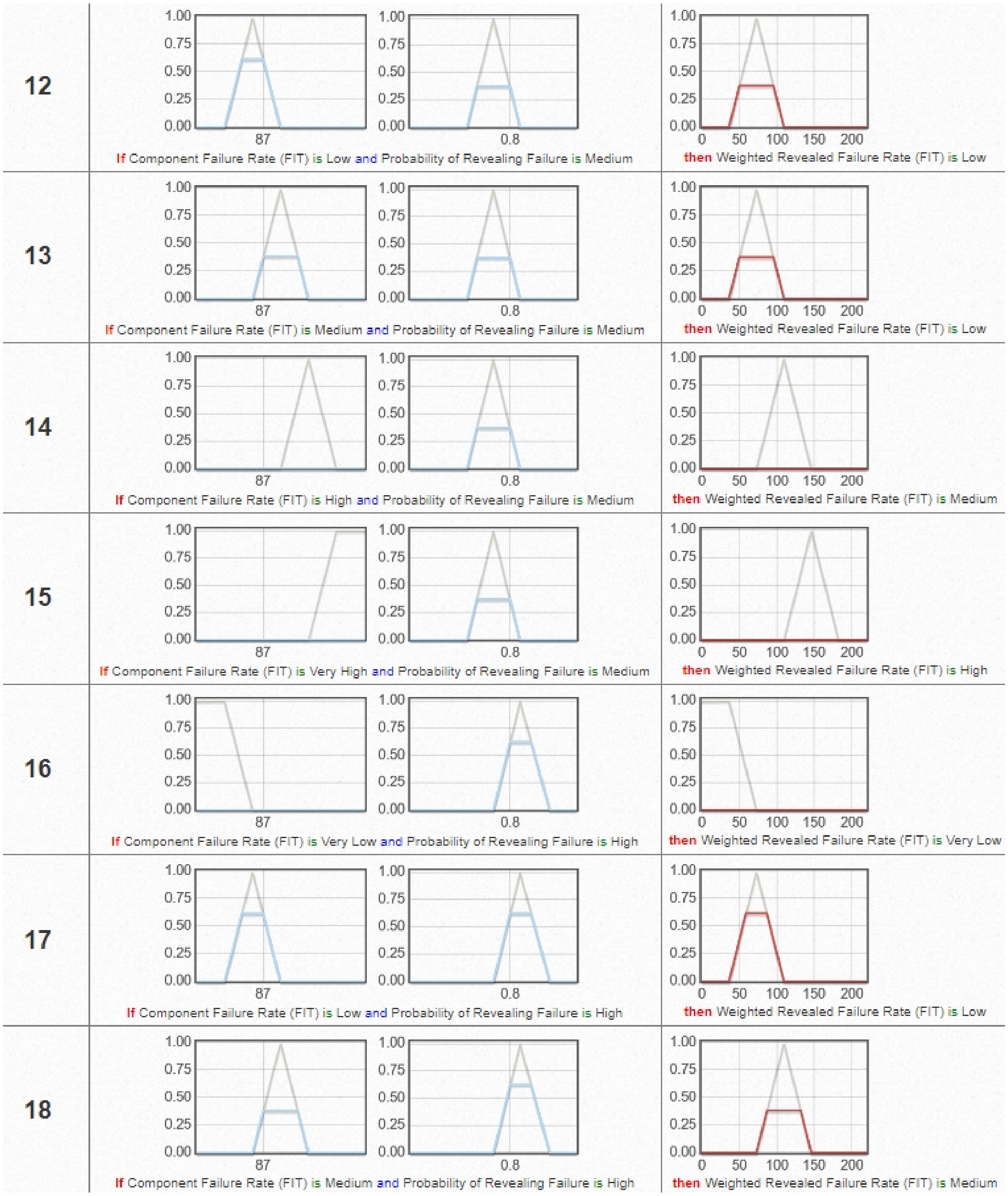
6.2.5 Phase 4 – Testing of the Revised FIS Design

The purpose of testing a fuzzy inference system is to evaluate its performance in making predictions based on its input data. The use of random numbers as input to the FIS is a method of stress testing the design allowing modifications to be made based on the test results prior to validation. Random numbers within the input fuzzy set's universe of discourse were generated using Excel, i.e., Component failure rate 1 to 220 FIT and Probability of revealing a failure 0.5 to 1. Random numbers were inputted into the FIS design using JuzzyOnline and the resulting values, i.e., the expected values, were recorded for analysis against the FMEDA values, i.e., the observed values.

The operation of the FIS is shown in the following example, with inputs:

- Component failure rate = 87 FIT
- Probability of revealing a failure = 0.8

As shown in Figure 6.20, Rules #12, #13, #17 and #18 are activated to produce an aggregated output.



12 If Component Failure Rate (FIT) is Low and Probability of Revealing Failure is Medium then Weighted Revealed Failure Rate (FIT) is Low

13 If Component Failure Rate (FIT) is Medium and Probability of Revealing Failure is Medium then Weighted Revealed Failure Rate (FIT) is Low

14 If Component Failure Rate (FIT) is High and Probability of Revealing Failure is Medium then Weighted Revealed Failure Rate (FIT) is Medium

15 If Component Failure Rate (FIT) is Very High and Probability of Revealing Failure is Medium then Weighted Revealed Failure Rate (FIT) is High

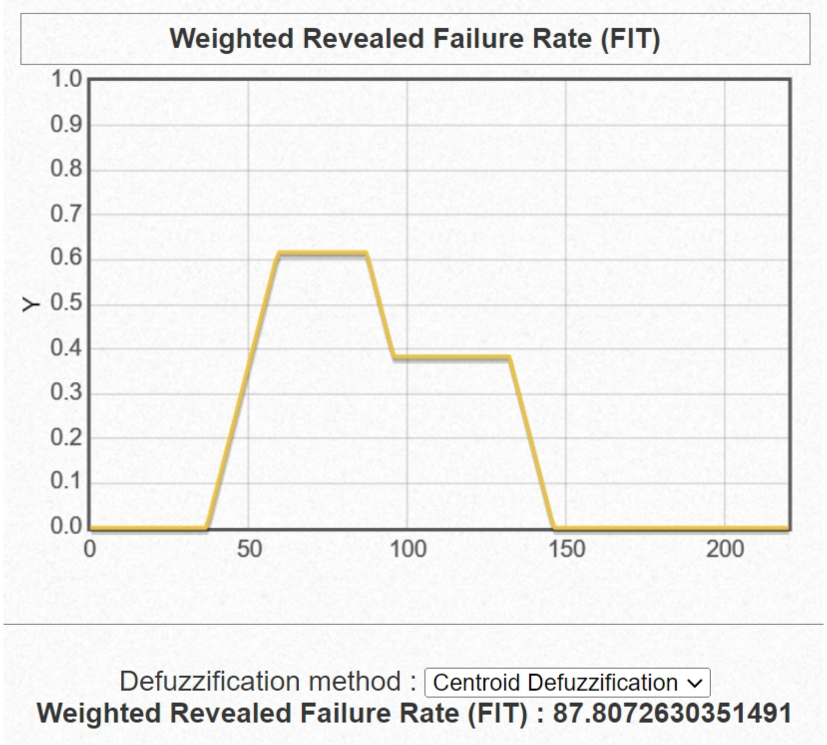
16 If Component Failure Rate (FIT) is Very Low and Probability of Revealing Failure is High then Weighted Revealed Failure Rate (FIT) is Very Low

17 If Component Failure Rate (FIT) is Low and Probability of Revealing Failure is High then Weighted Revealed Failure Rate (FIT) is Low

18 If Component Failure Rate (FIT) is Medium and Probability of Revealing Failure is High then Weighted Revealed Failure Rate (FIT) is Medium

Figure 6.20: Activated rules - Rule #12, #13, #17 and #18

The aggregated output with centroid defuzzification is shown in Figure 6.21. The observed value, the weighted revealed failure rate is 69.6 FIT and the expected output is 87.8 FIT.



Defuzzification method : Centroid Defuzzification

Weighted Revealed Failure Rate (FIT) : 87.8072630351491

Figure 6.21: Aggregated output with centroid defuzzification

6.2.5.1 Analysis of the FIS test results - a comparison of the expected and observed results

The regression trendline for the FMEDA output data (observed values) versus the FIS output data (expected values) based on random numbers is shown in Figure 6.22. The R^2 value is 0.9403 indicating that the data fit very well to the regression model. The difference (errors) between the FMEDA output data and the FIS output data is shown in Figure 6.23. The MSE value is 204 indicating a relatively low error between data points. The actual mean error (RMSE) is approximately 14 FIT.

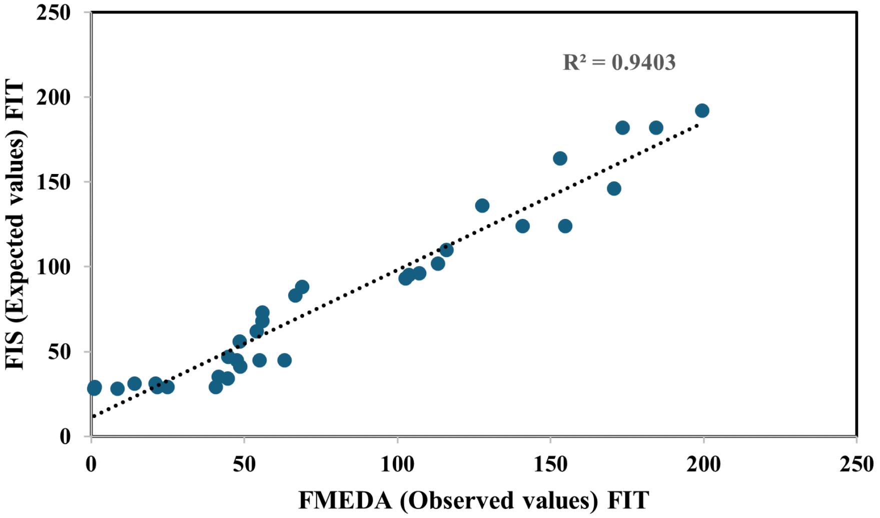


Figure 6.22: Triangular 5x5 FIS Regression

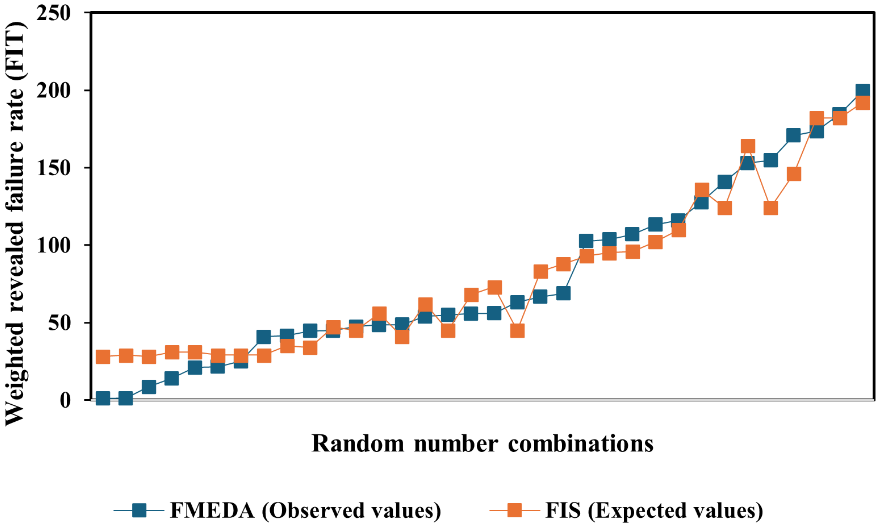


Figure 6.23: Triangular 5x5 FIS - FMEDA v FIS results

6.2.6 Phase 5 – Validation

The first version of the design was tested on data that was used to design the FIS, which can lead to 'over-fitting' and poor generalisation to unseen data. The purpose of validation is to ensure that the design can accommodate various datasets. The FIS design was validated using data sets provided by experts in FMEDA (see Section 5.4 for more details), including three synthetic datasets (Datasets #8, #9, and #10), which are highlighted in blue in Table 6.3.

Table 6.3: FMEDA expert data (FIT). Synthetic datasets are highlighted in blue.

	Datasets									
Component	**#1**	**#2**	**#3**	**#4**	**#5**	**#6**	**#7**	**#8**	**#9**	**#10**
Sleeve bearing	22	110	50	56	2	40	10	28	84	44
Poppet w/seals	36	220	30	45	1	32	6	23	68	72
Spring	104	165	40	45	1	32	16	23	68	208
Spring	26	55	40	45	1	32	24	23	68	52
Piston rod	20	19	30	13	11	13	5	6	19	39
Piston rod	20	38	130	26	21	26	5	13	39	39
Yoke	36	19	30	13	11	13	15	6	19	72
Yoke	36	150	70	52	43	52	15	26	78	72
Cylinder body	3	19	30	13	11	13	12	6	19	5
Cylinder body	3	19	20	13	11	13	12	6	19	5
Piston	8	150	160	104	86	104	16	52	155	17
Piston	4	19	90	13	11	13	8	6	19	7
Piston seals	12	75	70	52	43	52	16	26	78	24
Tie rod	3	19	30	13	11	13	5	6	19	6
Tie rod	3	37	70	26	21	26	5	13	39	6
Spring	104	57	70	39	32	39	16	19	58	208
Spring	26	19	30	39	32	39	24	19	58	52
Stem bush/bearings	47	73	43	24	22	7	45	24	73	95
Stem bush/bearings	47	73	43	48	44	15	45	24	73	95
Valve body	5	19	40	12	11	4	12	6	18	9
Valve body	5	19	20	12	11	4	12	6	18	9
Seat ring/spring A	23	55	40	24	22	7	6	12	36	45
Seat ring/spring A	45	91	105	121	110	37	15	60	180	90
Seat ring/spring B	23	55	40	24	22	7	6	12	36	45
Seat ring/spring B	45	91	105	121	110	37	15	60	180	90
Stem	27	73	20	73	66	22	15	36	109	54
Stem	27	73	20	73	66	22	15	36	109	54
Stem	18	73	65	12	11	4	8	6	18	36
Obturator (ball)	25	19	43	12	11	4	6	6	18	51
Obturator (ball)	25	37	43	24	22	7	6	12	36	51
Obturator (ball)	25	19	43	12	11	4	6	6	18	51
Obturator (ball)	76	145	65	73	66	22	39	36	108	152
Trunnion bush/bearings	47	73	43	48	44	15	15	24	73	95
Trunnion bush/bearings	47	73	43	48	44	15	15	24	73	95

Synthetic datasets (Datasets #8, #9 and #10) have been included in the validation phase as only a limited amount of data could be gathered in the time frame of this research. Synthetic data sets are derived from existing data (Datasets #1 and #4) by applying multiplying factors to represent a range of values, encompassing both optimistic and pessimistic scenarios, as described in the following.

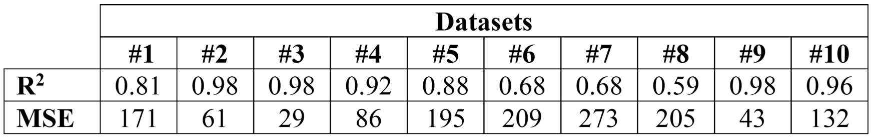
- Dataset #8 - based on Dataset #4 with a multiplying factor of 0.5.
- Dataset #9 - based on Dataset #4 with a multiplying factor of 1.5.
- Dataset #10 - based on Dataset #1 with a multiplying factor of 2.

Synthetic data sets have the same distribution characteristics as original data sets, that is, Pearson's correlation coefficient = 1 [152]. Careful consideration is given to these data when considering the overall results.

6.2.6.1 Analysis of the results from validation - a comparison of the expected and observed results

Table 6.4 shows the R^2 and MSE results from validation. There are a number of results that show low regression, such as Dataset #1, #6, #7 and #8. The MSE results indicates that the FIS for Dataset #3 has the smallest errors between data points and Dataset #7 the largest.

Table 6.4: Validation results R^2 and MSE



	Datasets									
	#1	**#2**	**#3**	**#4**	**#5**	**#6**	**#7**	**#8**	**#9**	**#10**
R^2	0.81	0.98	0.98	0.92	0.88	0.68	0.68	0.59	0.98	0.96
MSE	171	61	29	86	195	209	273	205	43	132

Figures 6.24 and 6.25 present a comparison of datasets with strong and weak regression, respectively.

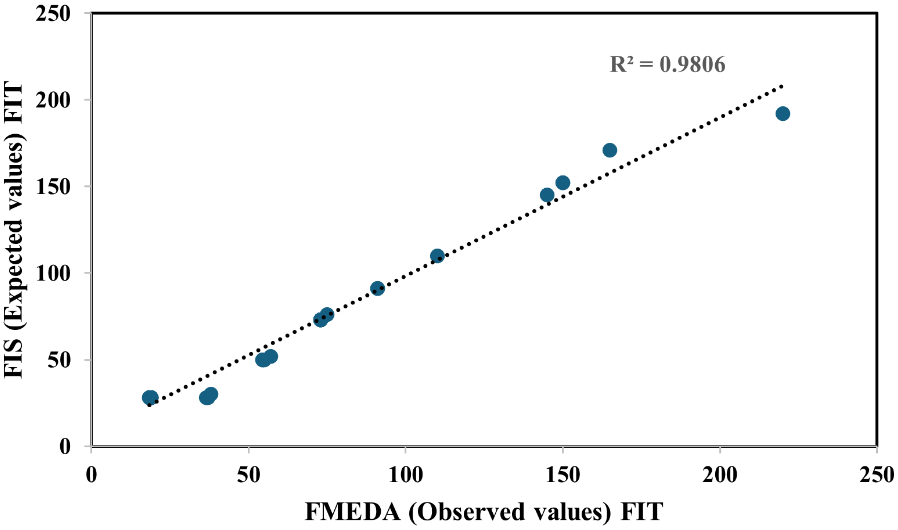


Figure 6.24: Dataset #2 indicating strong R^2

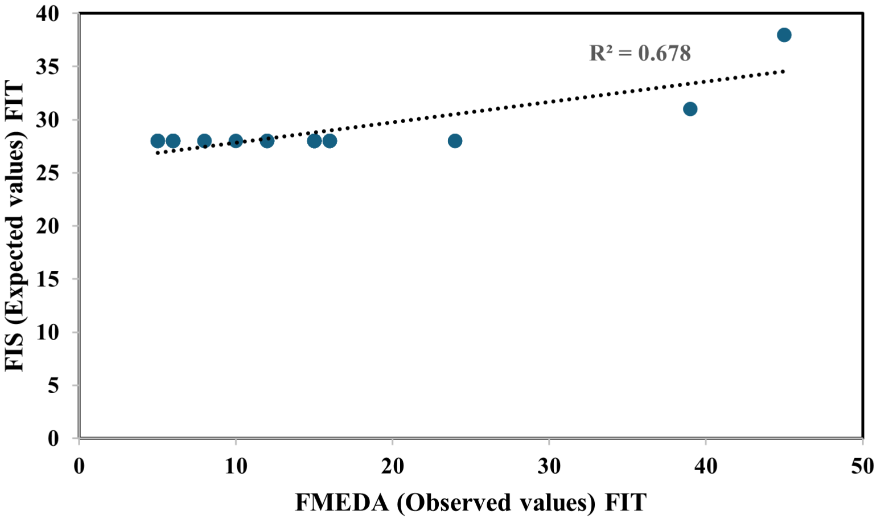


Figure 6.25: Dataset #7 indicating weak R^2

6.2.7 Discussion - Triangular 5x5 FIS Design

The test results (Section 6.2.5.1) using random data indicate that the Triangular 5x5 FIS design has a strong regression, i.e., R^2 = 0.9403, and the errors between FMEDA (observed values) and FIS (expected values) are relatively low, i.e. MSE = 204. In contrast, the validation results (Section 6.2.6.1) suggest that in some cases regression is weak, i.e, R^2 is approximately 0.6, for Dataset #6, Dataset #7 and Dataset #8.

Weak regression indicates that the model fails to accurately capture the relationships between inputs and outputs, i.e., a poor fit to the data. Causes of weak regression include an inadequate IF-THEN rule base or poorly defined membership functions.

The results from validation indicate that the Triangular 5x5 design is not suitable as the results from validation for four datasets indicate weak regression between observed and expected data; therefore, further effort is required to determine an appropriate FIS design. Using the methodology described earlier in Section 6.2 and returning to Phase 2 to determine a suitable FIS design. The following section describes the further research and analysis carried out to identify a suitable FIS design.

6.3 Refinement of Fuzzy Sets by Experimenting with Different Types of Membership Functions (Shapes and Quantities)

From the FMEA, FMECA, and FTA research papers, it is evident that the most commonly used fuzzy set type is triangular. Trapezoidal and Gaussian types also feature in the literature. This section examines FIS designs that include trapezoidal and Gaussian fuzzy set types.

The other significant design feature is the number of membership functions used in the fuzzy sets. The Triangular 5x5, which was based on the most common design criteria identified in the literature review, has already been investigated, albeit with unfavourable results. Additional commonly used variations in fuzzy inference design were considered in the following sections, and an analysis of the results was performed to determine the most suitable design. The variation in FIS design included:

- Triangular 3x3 membership functions.
- Triangular 6x6 membership functions.
- Trapezoidal 3x3 membership functions.
- Trapezoidal 5x5 membership functions.
- Trapezoidal 6x6 membership functions.
- Gaussian 3x3 membership functions.
- Gaussian 5x5 membership functions.
- Gaussian 6x6 membership functions.

An example of a Trapezoidal 6x6 FIS deign is shown in Figure 6.26 Component failure rate, Figure 6.27 Probability of revealing a failure and Figure 6.28 Weighted revealed failure rate.

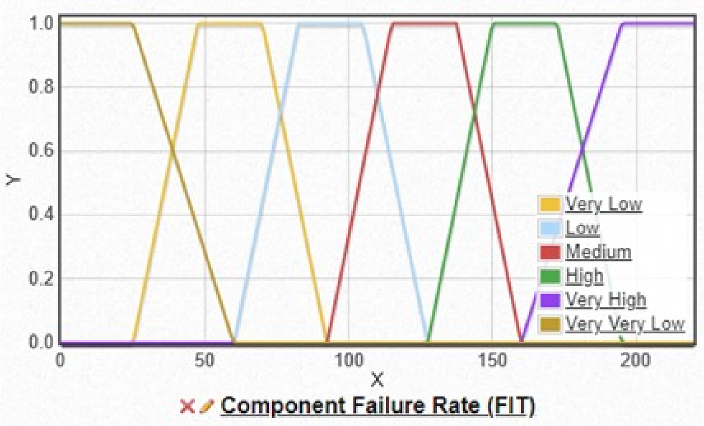


Figure 6.26: Trapezoidal 6x6 - Component failure rate fuzzy set

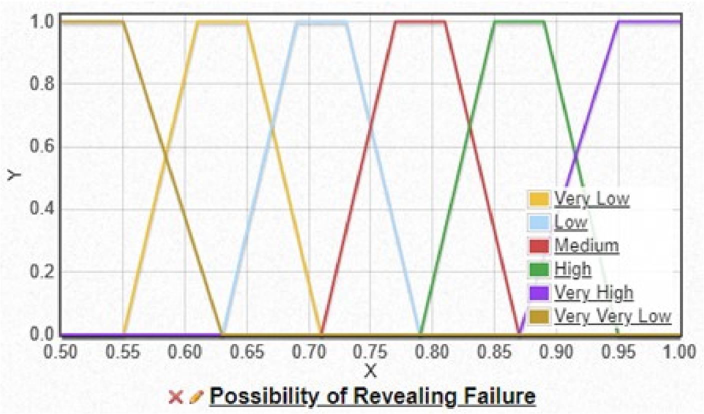


Figure 6.27: Trapezoidal 6x6 - Probability of revealing a failure fuzzy set

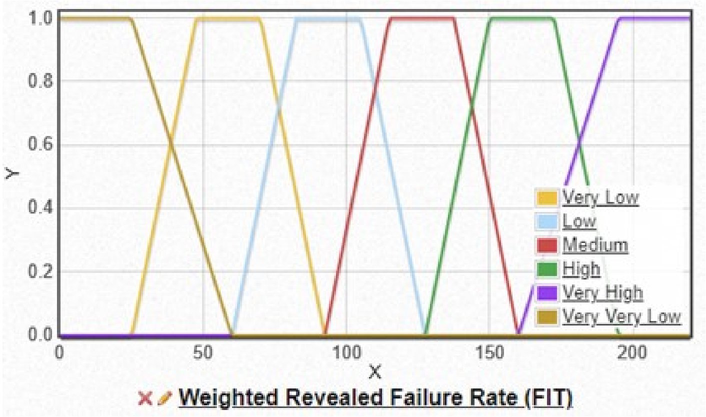


Figure 6.28: Trapezoidal 6x6 - Weighted revealed failure rate fuzzy set

An example of a Gaussian 6x6 FIS design is shown in Figure 6.29 Component failure rate, Figure 6.30 Probability of revealing a failure, and Figure 6.31 Weighted revealed failure rate.

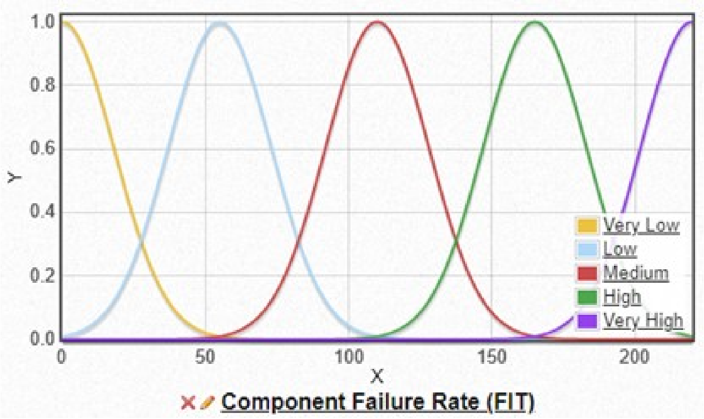


Figure 6.29: Trapezoidal 6x6 - Component failure rate fuzzy set

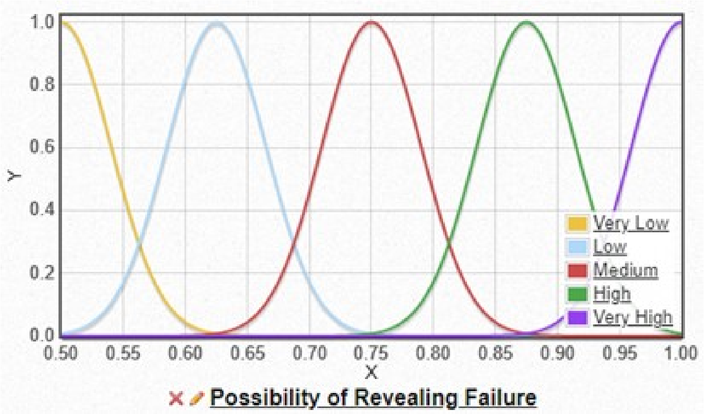


Figure 6.30: Trapezoidal 6x6 - Probability of revealing a failure fuzzy set

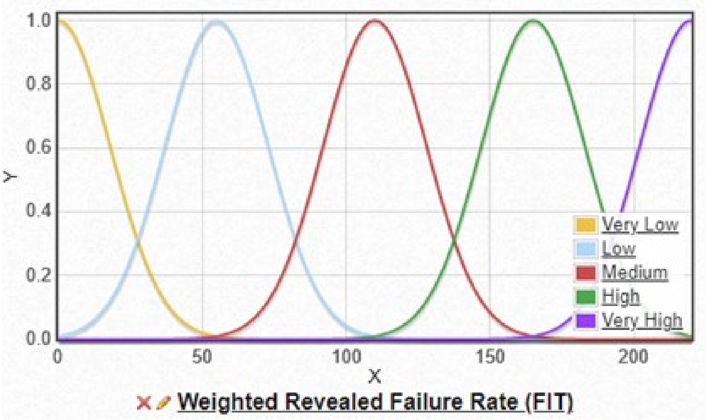


Figure 6.31: Trapezoidal 6x6 - Weighted revealed failure rate fuzzy set

6.3.1 Testing of Experimental FIS Designs

The FIS designs were tested using random data as per Section 6.2.5. Table 6.6 shows the R^2 values comparing FMEDA data (Observed values) with FIS data (expected values) for the different design configurations and Table 6.6 shows the associated MSE values.

Table 6.5: R^2 random number test results for different FIS designs

Membership functions	Triangular	Trapezoidal	Gaussian
3x3	0.878	0.874	0.661
5x5	0.940	0.959	0.943
6x6	0.974	0.971	0.955

Table 6.6: MSE random number test results for different FIS designs

Membership functions	Triangular	Trapezoidal	Gaussian
3x3	466	498	1821
5x5	204	181	230
6x6	105	122	152

The test results indicate that the best performing FIS design based on the R^2 and MSE metrics is the Triangular 6x6 design, at 0.974 and 105 respectively. The worst performing FIS design is the Gaussian 3x3 design, at 0.661 and 1821. Figures 6.32 and Figure 6.33 show the regression graphs for the best and worst FIS designs, respectively.

As can be seen in Figure 6.33, there is a greater spread of values around the regression line for the Gaussian 3x3 design.

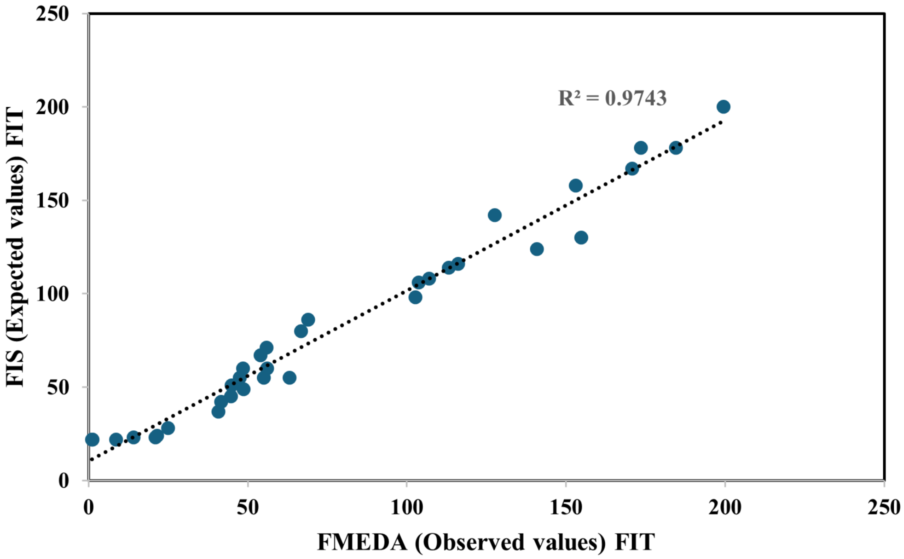


Figure 6.32: Triangular 6x6 FIS design indicating good R^2

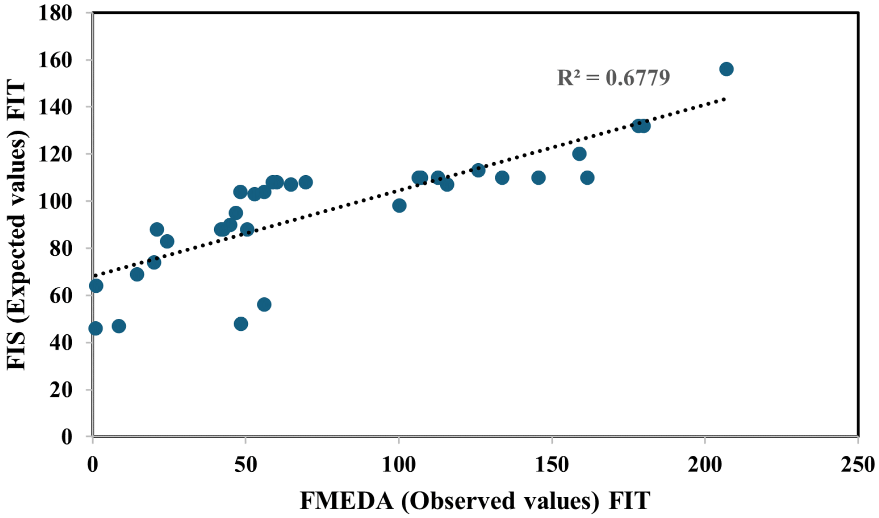


Figure 6.33: Gaussian 3x3 FIS design indicating poor R^2

6.3.2 Validation of Experimental FIS Designs

Validation was carried out using expert data, based on a full stroke test that has a probability of revealing dangerous failures = 1. Table 6.7 shows the R_2 values comparing FMEDA data with the FIS data for different design configurations and Table 6.8 shows the associated MSE values.

Table 6.7: R^2 expert data validation results for different FIS designs

<table>
<tr><th colspan="11">Triangular</th></tr>
<tr><th>Membership functions</th><th>Dataset #1</th><th>Dataset #2</th><th>Dataset #3</th><th>Dataset #4</th><th>Dataset #5</th><th>Dataset #6</th><th>Dataset #7</th><th>Dataset #8</th><th>Dataset #9</th><th>Dataset #10</th></tr>
<tr><td>3x3</td><td>0.689</td><td>0.943</td><td>0.934</td><td>0.776</td><td>0.777</td><td>0.542</td><td>0.046</td><td>0.516</td><td>0.951</td><td>0.906</td></tr>
<tr><td>5x5</td><td>0.808</td><td>0.981</td><td>0.978</td><td>0.918</td><td>0.876</td><td>0.679</td><td>0.677</td><td>0.591</td><td>0.981</td><td>0.957</td></tr>
<tr><td>6x6</td><td>0.904</td><td>0.997</td><td>0.999</td><td>0.980</td><td>0.951</td><td>0.887</td><td>0.755</td><td>0.772</td><td>0.999</td><td>0.989</td></tr>
<tr><th colspan="11">Trapezoidal</th></tr>
<tr><th>Membership functions</th><th>Dataset #1</th><th>Dataset #2</th><th>Dataset #3</th><th>Dataset #4</th><th>Dataset #5</th><th>Dataset #6</th><th>Dataset #7</th><th>Dataset #8</th><th>Dataset #9</th><th>Dataset #10</th></tr>
<tr><td>3x3</td><td>0.694</td><td>0.965</td><td>0.947</td><td>0.793</td><td>0.800</td><td>0.552</td><td>0.116</td><td>0.514</td><td>0.971</td><td>0.912</td></tr>
<tr><td>5x5</td><td>0.916</td><td>0.996</td><td>0.992</td><td>0.983</td><td>0.960</td><td>0.919</td><td>0.756</td><td>0.758</td><td>0.999</td><td>0.993</td></tr>
<tr><td>6x6</td><td>0.894</td><td>0.996</td><td>0.999</td><td>0.984</td><td>0.958</td><td>0.910</td><td>0.756</td><td>0.888</td><td>0.999</td><td>0.992</td></tr>
<tr><th colspan="11">Gaussian</th></tr>
<tr><th>Membership functions</th><th>Dataset #1</th><th>Dataset #2</th><th>Dataset #3</th><th>Dataset #4</th><th>Dataset #5</th><th>Dataset #6</th><th>Dataset #7</th><th>Dataset #8</th><th>Dataset #9</th><th>Dataset #10</th></tr>
<tr><td>3x3</td><td>0.936</td><td>0.878</td><td>0.844</td><td>0.878</td><td>0.932</td><td>0.936</td><td>0.997</td><td>0.997</td><td>0.890</td><td>0.934</td></tr>
<tr><td>5x5</td><td>0.972</td><td>0.973</td><td>0.939</td><td>0.967</td><td>0.967</td><td>0.966</td><td>0.981</td><td>0.892</td><td>0.971</td><td>0.986</td></tr>
<tr><td>6x6</td><td>0.960</td><td>0.988</td><td>0.976</td><td>0.962</td><td>0.963</td><td>0.959</td><td>0.928</td><td>0.894</td><td>0.992</td><td>0.993</td></tr>
</table>

Table 6.8: MSE expert data validation results for different FIS designs

<table>
<tr><th colspan="11">Triangular</th></tr>
<tr><th>Membership functions</th><th>Dataset #1</th><th>Dataset #2</th><th>Dataset #3</th><th>Dataset #4</th><th>Dataset #5</th><th>Dataset #6</th><th>Dataset #7</th><th>Dataset #8</th><th>Dataset #9</th><th>Dataset #10</th></tr>
<tr><td>3x3</td><td>489</td><td>220</td><td>81</td><td>314</td><td>533</td><td>594</td><td>818</td><td>625</td><td>167</td><td>338</td></tr>
<tr><td>5x5</td><td>171</td><td>61</td><td>29</td><td>86</td><td>195</td><td>209</td><td>273</td><td>205</td><td>43</td><td>132</td></tr>
<tr><td>6x6</td><td>84</td><td>15</td><td>1</td><td>30</td><td>91</td><td>94</td><td>123</td><td>95</td><td>4</td><td>54</td></tr>
<tr><th colspan="11">Trapezoidal</th></tr>
<tr><th>Membership functions</th><th>Dataset #1</th><th>Dataset #2</th><th>Dataset #3</th><th>Dataset #4</th><th>Dataset #5</th><th>Dataset #6</th><th>Dataset #7</th><th>Dataset #8</th><th>Dataset #9</th><th>Dataset #10</th></tr>
<tr><td>3x3</td><td>489</td><td>218</td><td>79</td><td>313</td><td>552</td><td>593</td><td>823</td><td>622</td><td>168</td><td>339</td></tr>
<tr><td>5x5</td><td>120</td><td>36</td><td>31</td><td>61</td><td>130</td><td>139</td><td>170</td><td>163</td><td>22</td><td>93</td></tr>
<tr><td>6x6</td><td>105</td><td>19</td><td>6</td><td>33</td><td>93</td><td>96</td><td>124</td><td>96</td><td>7</td><td>58</td></tr>
<tr><th colspan="11">Gaussian</th></tr>
<tr><th>Membership functions</th><th>Dataset #1</th><th>Dataset #2</th><th>Dataset #3</th><th>Dataset #4</th><th>Dataset #5</th><th>Dataset #6</th><th>Dataset #7</th><th>Dataset #8</th><th>Dataset #9</th><th>Dataset #10</th></tr>
<tr><td>3x3</td><td>2232</td><td>1647</td><td>1789</td><td>1982</td><td>2187</td><td>2392</td><td>2576</td><td>2462</td><td>1681</td><td>1827</td></tr>
<tr><td>5x5</td><td>183</td><td>107</td><td>159</td><td>113</td><td>138</td><td>173</td><td>134</td><td>177</td><td>111</td><td>103</td></tr>
<tr><td>6x6</td><td>123</td><td>85</td><td>61</td><td>72</td><td>100</td><td>108</td><td>109</td><td>128</td><td>71</td><td>45</td></tr>
</table>

From these metrics, the best performing FIS design can be identified for a full stroke test based on the datasets used for validation. The best performing FIS design from a regression perspective was both the Triangular and Trapezoidal 6x6 for Dataset #3 dataset at R^2 = 0.999. From an error perspective, the best performing FIS design was the Triangular 6x6 for Dataset #3 at MSE = 1. The FIS design with the lowest performance from a regression perspective was the Triangular 3x3 for Dataset #7 dataset at 0.046, and for the error it was the Gaussian 3x3 for Dataset #7 at MSE = 2576.

The results of validation confirm, in general, that the greater the number of fuzzy sets, the better the regression and error values. There are exceptions, such as Gaussian 3x3 for Dataset #7 has the best regression with R^2 = 0.997 but also the worst MSE value at 2576. The line chart, Figure 6.34, plots the Gaussian 3x3 FIS design for Dataset #7 showing the FMEDA output versus the FIS output. The RMSE value is 51, i.e., the mean error between data points is 51 FIT. This is a good example of under-fitting of the FIS design, indicating that the model is too simplistic.

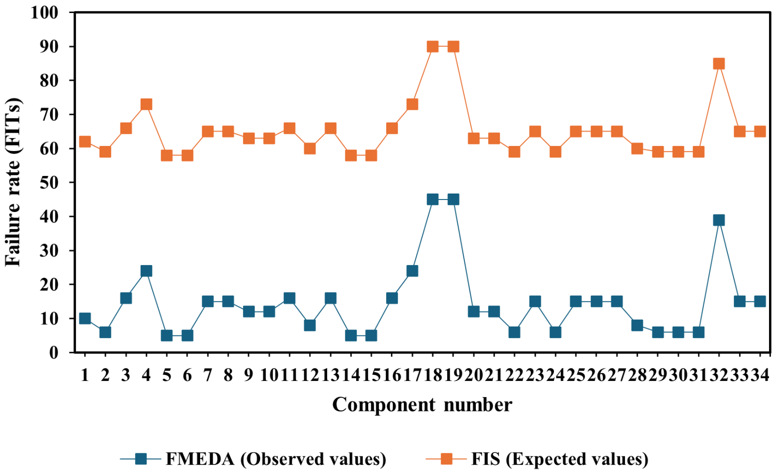


Figure 6.34: Gaussian 3x3 Dataset #7 - example of under-fitting

From the regression and error results it is evident that both the Trapezoidal and Gaussian 6x6 design are preferred designs for the current datasets used for validation. The Trapezoidal design producing a average $R^2 = 0.938$ and an average MSE = 64 and the Gaussian design 0.961 and 90. However, the Trapezoidal 6x6 design does not perform well for Dataset #7 with $R^2 = 0.756$. The Gaussian design is more consistent across all datasets and, therefore, the preferred design at this point.

6.3.3 An Assessment of FIS Designs Considering Manual Partial Stroke Testing (PST)

Thus far the FIS design has considered the datasets relating to a full stroke test, through which the probability of revealing a failure is assumed to be 1. Partial stroke testing, on the other hand, will have varying values of probability of revealing a failure as discussed, in Section 6.2.3 Table 3.2, from 0.5 to 1. Although testing has been carried out using random numbers in the range of 0.5 to 1 (representing probability of revealing a failure), validation tests are required to ensure that the FIS design is suitable for both Full Stroke Test (FST) and Partial Stroke Test (PST) proof test coverage assessments. Original component failure rate data provided by experts (Section 5.4) was used for the validation of the PST (see component failure rates in Section 6.2.6, Table 6.3). The probability of revealing a failure was determined by Expert #1 and Expert #2 for the revealable dangerous failure rates, Fail to Close (FTC), and Delayed Operation (DOP). The probabilities are as follows:

- FTC failure mode – a probability of 0.7 based on field data collection within Chapter 4.

- DOP failure mode - a probability 0.8 considering sub-failures shown in Table 6.9 and the following justification.

Table 6.9: Probability of revealing DOP sub-failures [67]

Failure Mode	Sub-failure Mode	Probability of Revealing the Sub-failure Mode
Delayed Operation (DOP)	Fail to close 20% in the specified time.	0.5
	Fail to fully close in the specified time.	0 to 0.5

A probability of 0.5 assigned to sub-failure mode, Fail to close 20% in the specified time, and an additional 0.3 based on the high likelihood that the ESDV will travel to its fully closed position if it can reach the 20% closed position within the specified time. Assuming the ESDV is still within its useful life period and the most recent proof test record indicates a successful full stroke test. Leaving a total probability of revealing DOP failures = 0.8.

The above probabilities can also be supported by considering the design of a Scotch-yoke actuator, the type considered in this study. The torque curves for a Scotch-yoke actuator are shown in Figure 6.35, in which the torque is greatest at the start and at the end of its travel. In addition, it is standard practice to include a safety factor of 1.5 to 2 in operating torque when specifying valve and actuator assemblies. These features provide confidence in the ESDV operation.

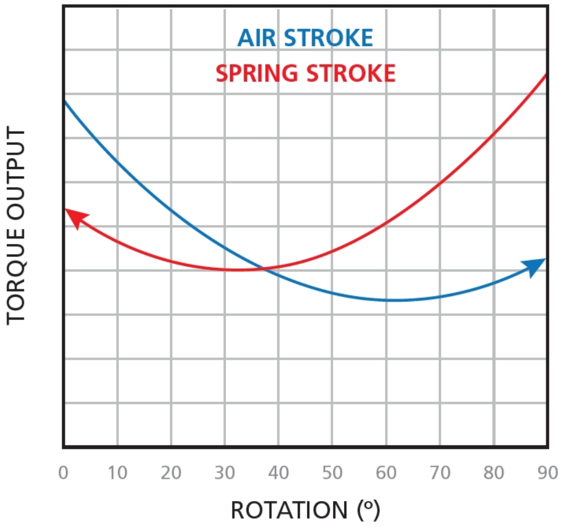


Figure 6.35: Scotch-yoke actuator torque curves [44]

Examples of the Gaussian 6x6 FIS design using the probability of revealing a failure = 0.7 for the FTC failure mode (Example 1) and 0.8 for the DOP failure mode (Example

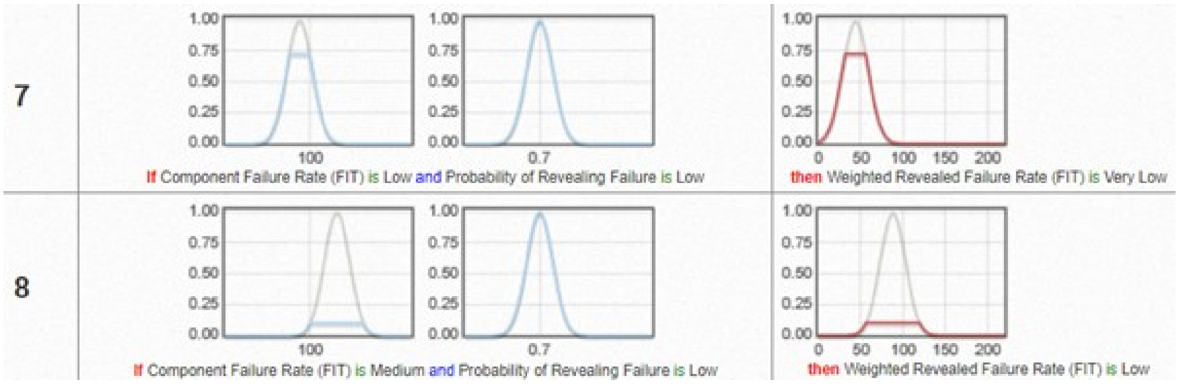
2) are presented in the following. These examples illustrate how the IF-THEN rules are activated and aggregated to form the output, the weighted revealed failure rate.

Example 1

Component failure rate = 100 FIT

Probability of revealing the failure = 0.7 (representing FTC failure mode)

Figure 6.36 illustrates the activation of Rules #7 and #8.



If Component Failure Rate (FIT) is Low and Probability of Revealing Failure is Low

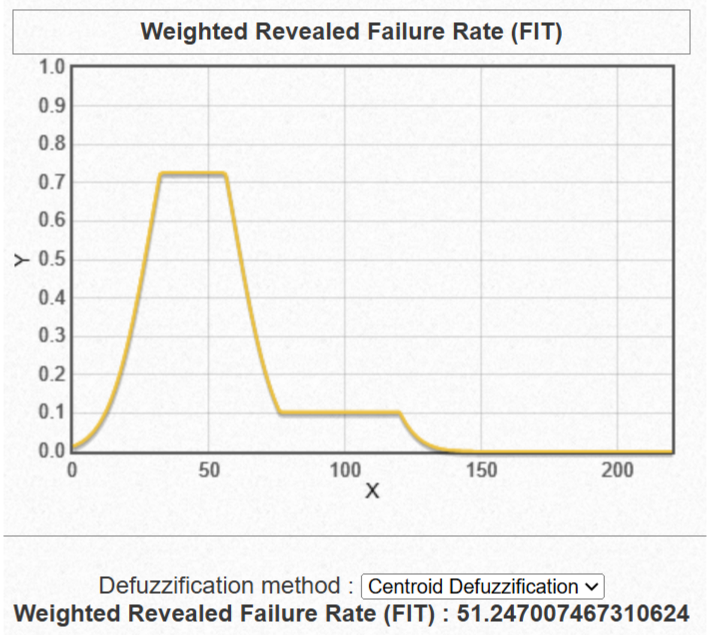
then Weighted Revealed Failure Rate (FIT) is Very Low

If Component Failure Rate (FIT) is Medium and Probability of Revealing Failure is Low

then Weighted Revealed Failure Rate (FIT) is Low

Figure 6.36: Activation of rules #7 and #8

Figure 6.37 shows the aggregated weighted revealed failure rate of 51.2 FIT.

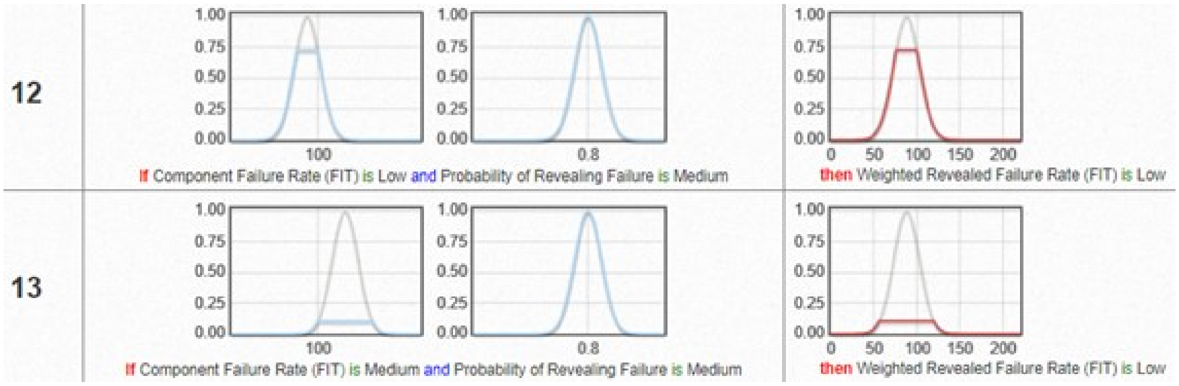


Defuzzification method : Centroid Defuzzification

Weighted Revealed Failure Rate (FIT) : 51.247007467310624

Figure 6.37: Aggregated output of the inputs, 100 FIT and 0.7 probability

Example 2
Component failure rate = 100 FIT
Probability of revealing the failure = 0.8 (representing DOP failure mode)
Figure 6.38 illustrates the activation of Rules #12 and #13.

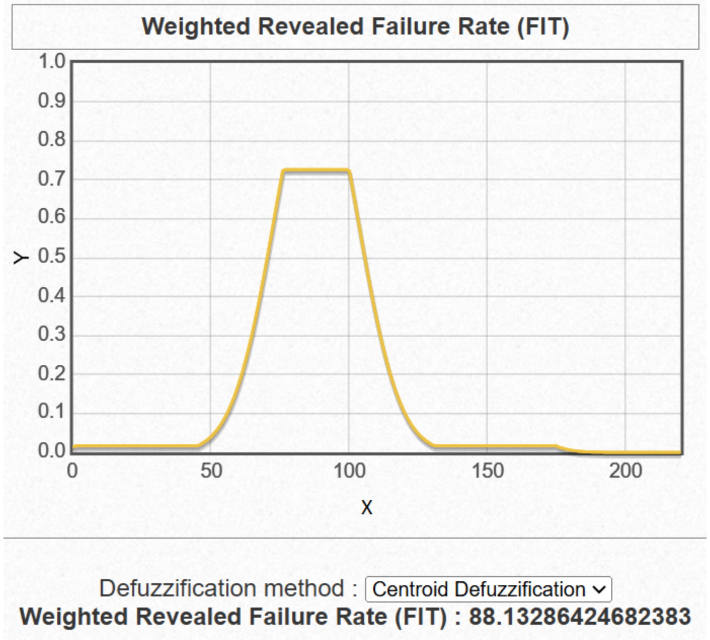


12: If Component Failure Rate (FIT) is Low and Probability of Revealing Failure is Medium then Weighted Revealed Failure Rate (FIT) is Low

13: If Component Failure Rate (FIT) is Medium and Probability of Revealing Failure is Medium then Weighted Revealed Failure Rate (FIT) is Low

Figure 6.38: Activation of rules #12 and #13

Figure 6.39 shows the aggregated weighted revealed failure rate of 88.1 FIT.



Defuzzification method : Centroid Defuzzification
Weighted Revealed Failure Rate (FIT) : 88.13286424682383

Figure 6.39: Aggregated output of the inputs, 100 FIT and 0.8 probability

6.3.3.1 Validation of FIS designs considering PST

As with FST validation, the following R^2 (Table 6.10) and MSE (Table 6.11) metrics were produced for the various configurations of FIS design for PST:

Table 6.10: Validation R^2 results considering PST

Triangular										
Membership functions	**Dataset #1**	**Dataset #2**	**Dataset #3**	**Dataset #4**	**Dataset #5**	**Dataset #6**	**Dataset #7**	**Dataset #8**	**Dataset #9**	**Dataset #10**
3x3	0.569	0.892	0.898	0.860	0.846	0.537	0.114	0.498	0.978	0.851
5x5	0.793	0.946	0.955	0.947	0.914	0.622	0.347	0.669	0.952	0.908
6x6	0.908	0.973	0.975	0.970	0.965	0.863	0.729	0.854	0.9867	0.966
Trapezoidal										
Membership functions	**Dataset #1**	**Dataset #2**	**Dataset #3**	**Dataset #4**	**Dataset #5**	**Dataset #6**	**Dataset #7**	**Dataset #8**	**Dataset #9**	**Dataset #10**
3x3	0.635	0.915	0.914	0.863	0.861	0.553	0.007	0.658	0.972	0.836
5x5	0.844	0.958	0.922	0.968	0.954	0.790	0.722	0.857	0.9635	0.949
6x6	0.906	0.971	0.961	0.962	0.964	0.888	0.757	0.859	0.9862	0.958
Gaussian										
Membership functions	**Dataset #1**	**Dataset #2**	**Dataset #3**	**Dataset #4**	**Dataset #5**	**Dataset #6**	**Dataset #7**	**Dataset #8**	**Dataset #9**	**Dataset #10**
3x3	0.943	0.694	0.759	0.754	0.847	0.777	0.934	0.958	0.777	0.679
5x5	0.757	0.933	0.826	0.944	0.952	0.850	0.917	0.809	0.939	0.958
6x6	0.778	0.938	0.907	0.896	0.962	0.927	0.866	0.786	0.9412	0.947

Table 6.11: Validation MSE results considering PST

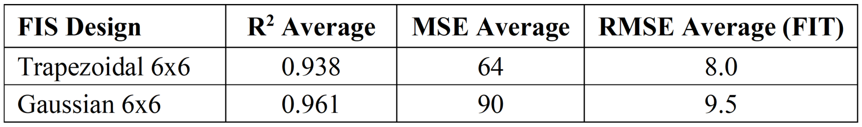
Triangular										
Membership functions	**Dataset #1**	**Dataset #2**	**Dataset #3**	**Dataset #4**	**Dataset #5**	**Dataset #6**	**Dataset #7**	**Dataset #8**	**Dataset #9**	**Dataset #10**
3x3	822	503	354	593	879	953	1193	1002	469	594
5x5	298	162	89	204	336	354	459	363	135	245
6x6	141	109	89	141	175	182	204	159	88	175
Trapezoidal										
Membership functions	**Dataset #1**	**Dataset #2**	**Dataset #3**	**Dataset #4**	**Dataset #5**	**Dataset #6**	**Dataset #7**	**Dataset #8**	**Dataset #9**	**Dataset #10**
3x3	656	465	291	473	714	751	961	789	472	542
5x5	270	163	228	221	298	377	329	274	152	256
6x6	144	100	133	150	174	180	186	145	85	198
Gaussian										
Membership functions	**Dataset #1**	**Dataset #2**	**Dataset #3**	**Dataset #4**	**Dataset #5**	**Dataset #6**	**Dataset #7**	**Dataset #8**	**Dataset #9**	**Dataset #10**
3x3	2868	2368	2661	2795	2818	3200	3017	2943	2593	2624
5x5	398	252	415	306	327	390	330	376	269	237
6x6	266	221	235	314	209	213	187	245	200	160

From these metrics, we can identify the best performing FIS design for a partial stroke test based on the datasets used for validation. The best performing FIS design from a regression perspective was the Triangular 6x6 for Expert #9 dataset at R2 = 0.987. From an error perspective, the best performing FIS design was the Trapezoidal 6x6 for Dataset #9 dataset at MSE = 85. The FIS design with the lowest performance from a regression perspective was the Triangular 3x3 for Expert #7 dataset at 0.007, and for the error it was the Gaussian 3x3 for Dataset #6 at MSE = 3200.

6.3.4 Comparison of FST and PST Validation Results

The results of the validation of FST (Section 6.3.2) indicate that the most suitable designs based on the average values of R^2, MSE and RMSE, considering each configuration and dataset, are the Trapezoidal and Gaussian 6x6. Table 6.12 presents the validation results for FST, highlighting that the Trapezoidal design demonstrates a lower MSE, while the Gaussian design achieves a higher R^2 value.

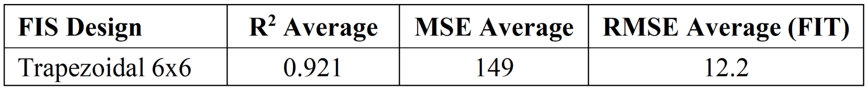
Table 6.12: Average values of R^2, MSE and RMSE for FST



FIS Design	R^2 Average	MSE Average	RMSE Average (FIT)
Trapezoidal 6x6	0.938	64	8.0
Gaussian 6x6	0.961	90	9.5

The results of the PST validation (Section 6.3.3.1) imply that the most suitable design based on average values of R^2, MSE and RMSE is the Trapezoidal. Table 6.13 presents the validation results for PST.

Table 6.13: Average values of R^2, MSE and RMSE for PST



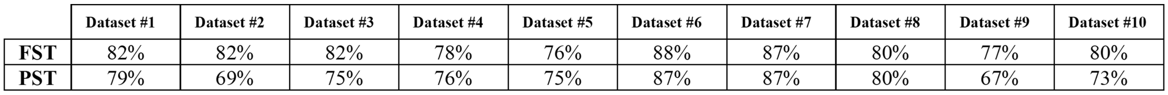
FIS Design	R^2 Average	MSE Average	RMSE Average (FIT)
Trapezoidal 6x6	0.921	149	12.2

Therefore, from the above metrics, in Table 6.12 and Table 6.13, it can be concluded that the Trapezoidal 6x6 design is the most suitable design for the evaluation of the coverage of the FST and PST proof test coverage assessments, based on the expert datasets used in this study.

6.3.5 Comparison of FST and PST Proof Test Coverage Values

The Proof Test Coverage (PTC) values were calculated for both the Full Stroke Test (FST) and the Partial Stroke Test (PST) using the Trapezoidal 6x6 FIS design. Table 6.14 shows that when comparing the PTC estimates derived from FST and PST, it is evident that the Trapezoidal design is limited as the results indicate that in some cases the value of PST PTC is equal to that of FST. For example, the PTC values are the same for Dataset #7 and #8. As an FST will reveal more dangerous failures than a PST, for the same ESDV, it should not be possible that the PST PTC values are equal or greater than the FST PTC values.

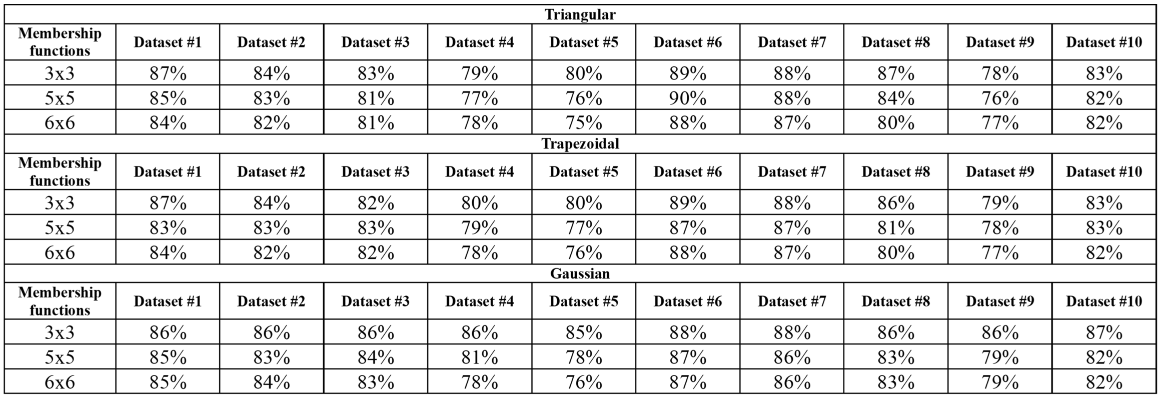
Table 6.14: Trapezoidal 6x6 - FST and PST PTC values



	Dataset #1	Dataset #2	Dataset #3	Dataset #4	Dataset #5	Dataset #6	Dataset #7	Dataset #8	Dataset #9	Dataset #10
FST	82%	82%	82%	78%	76%	88%	87%	80%	77%	80%
PST	79%	69%	75%	76%	75%	87%	87%	80%	67%	73%

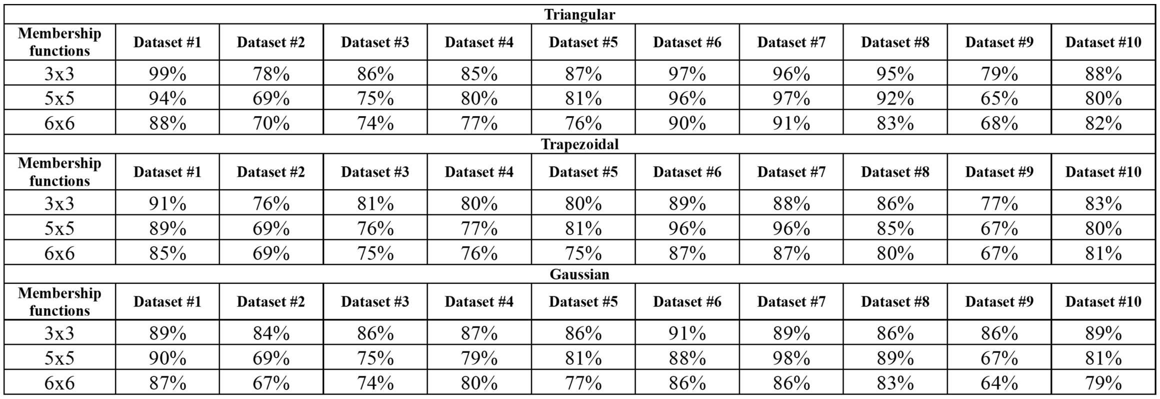
Revisiting the FIS designs and comparing the values of FST and PST PTC it is evident that all designs have the same problem as shown in Table 6.15 and Table 6.16, where in some cases PST PTC $\geq$ FST PTC, for example, Trapezoidal 6x6, Dataset #1 PST PTC (85%) > FST PTC (84%):

Table 6.15: FST PTC values



Triangular										
Membership functions	**Dataset #1**	**Dataset #2**	**Dataset #3**	**Dataset #4**	**Dataset #5**	**Dataset #6**	**Dataset #7**	**Dataset #8**	**Dataset #9**	**Dataset #10**
3x3	87%	84%	83%	79%	80%	89%	88%	87%	78%	83%
5x5	85%	83%	81%	77%	76%	90%	88%	84%	76%	82%
6x6	84%	82%	81%	78%	75%	88%	87%	80%	77%	82%
Trapezoidal										
Membership functions	**Dataset #1**	**Dataset #2**	**Dataset #3**	**Dataset #4**	**Dataset #5**	**Dataset #6**	**Dataset #7**	**Dataset #8**	**Dataset #9**	**Dataset #10**
3x3	87%	84%	82%	80%	80%	89%	88%	86%	79%	83%
5x5	83%	83%	83%	79%	77%	87%	87%	81%	78%	83%
6x6	84%	82%	82%	78%	76%	88%	87%	80%	77%	82%
Gaussian										
Membership functions	**Dataset #1**	**Dataset #2**	**Dataset #3**	**Dataset #4**	**Dataset #5**	**Dataset #6**	**Dataset #7**	**Dataset #8**	**Dataset #9**	**Dataset #10**
3x3	86%	86%	86%	86%	85%	88%	88%	86%	86%	87%
5x5	85%	83%	84%	81%	78%	87%	86%	83%	79%	82%
6x6	85%	84%	83%	78%	76%	87%	86%	83%	79%	82%

Table 6.16: PST PTC values



Triangular										
Membership functions	**Dataset #1**	**Dataset #2**	**Dataset #3**	**Dataset #4**	**Dataset #5**	**Dataset #6**	**Dataset #7**	**Dataset #8**	**Dataset #9**	**Dataset #10**
3x3	99%	78%	86%	85%	87%	97%	96%	95%	79%	88%
5x5	94%	69%	75%	80%	81%	96%	97%	92%	65%	80%
6x6	88%	70%	74%	77%	76%	90%	91%	83%	68%	82%
Trapezoidal										
Membership functions	**Dataset #1**	**Dataset #2**	**Dataset #3**	**Dataset #4**	**Dataset #5**	**Dataset #6**	**Dataset #7**	**Dataset #8**	**Dataset #9**	**Dataset #10**
3x3	91%	76%	81%	80%	80%	89%	88%	86%	77%	83%
5x5	89%	69%	76%	77%	81%	96%	96%	85%	67%	80%
6x6	85%	69%	75%	76%	75%	87%	87%	80%	67%	81%
Gaussian										
Membership functions	**Dataset #1**	**Dataset #2**	**Dataset #3**	**Dataset #4**	**Dataset #5**	**Dataset #6**	**Dataset #7**	**Dataset #8**	**Dataset #9**	**Dataset #10**
3x3	89%	84%	86%	87%	86%	91%	89%	86%	86%	89%
5x5	90%	69%	75%	79%	81%	88%	98%	89%	67%	81%
6x6	87%	67%	74%	80%	77%	86%	86%	83%	64%	79%

Table 6.17 highlights the FIS designs / Expert datasets where the FST PTC values > PST (green cells), and where the PST PTC ≥ FST PTC (red font).

Table 6.17: Difference between PTC values from FST and PST approaches

Triangular										
Membership functions	**Dataset #1**	**Dataset #2**	**Dataset #3**	**Dataset #4**	**Dataset #5**	**Dataset #6**	**Dataset #7**	**Dataset #8**	**Dataset #9**	**Dataset #10**
3x3	-12%	6%	-3%	-6%	-7%	-8%	-8%	-8%	-1%	-5%
5x5	-9%	14%	6%	-3%	-5%	-6%	-9%	-8%	11%	2%
6x6	-4%	12%	7%	1%	-1%	-2%	-4%	-3%	9%	0%
Trapezoidal										
Membership functions	**Dataset #1**	**Dataset #2**	**Dataset #3**	**Dataset #4**	**Dataset #5**	**Dataset #6**	**Dataset #7**	**Dataset #8**	**Dataset #9**	**Dataset #10**
3x3	-4%	8%	1%	0%	0%	0%	0%	0%	2%	0%
5x5	-6%	14%	7%	2%	-4%	-9%	-9%	-4%	11%	3%
6x6	-1%	13%	7%	2%	1%	1%	0%	0%	10%	1%
Gaussian										
Membership functions	**Dataset #1**	**Dataset #2**	**Dataset #3**	**Dataset #4**	**Dataset #5**	**Dataset #6**	**Dataset #7**	**Dataset #8**	**Dataset #9**	**Dataset #10**
3x3	-3%	2%	0%	-1%	-1%	-3%	-1%	0%	0%	-2%
5x5	-5%	14%	9%	2%	-3%	-1%	-12%	-6%	12%	1%
6x6	-2%	17%	9%	-2%	-1%	1%	0%	0%	15%	3%

The table indicates that the FIS designs, considering all expert datasets, fail to meet the following assumption: FST PTC > PST PTC. It is also evident that the FIS designs where expert datasets do not comply with this criteria are generally the datasets that are 1000 FIT or less (total dangerous failure rate of an ESDV assembly), i.e., Dataset #1, #5, #6, #7 and #8.

6.3.6 Analysis of the Expert Datasets to Determine Distribution of Failure Rates

Considering the PTC issue discussed above, the expert datasets were further scrutinised. The Pareto chart, Figure 6.40, shows the spread of component failure rate for Dataset #1 to Dataset #7, (noting that Datasets #8, #9 and #10 are synthetic). The graph indicates that more than 80% of the data lie within the 1 FIT to 58 FIT region, suggesting a skew towards the lower end of the distribution.

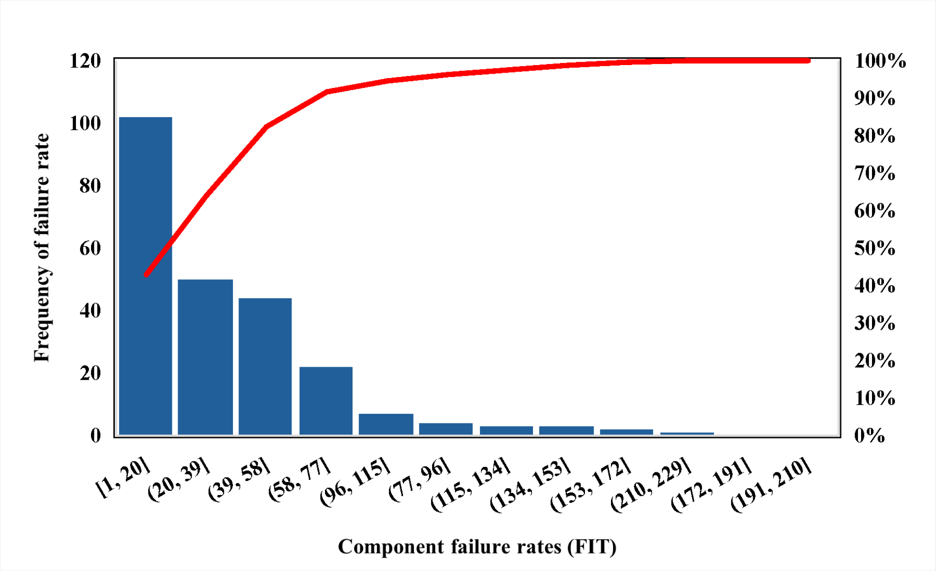


Figure 6.40: Distribution of component failure rates - Dataset #1 to #7

The distribution graph, Figure 6.41, shows the plot of the Cumulative Distribution Function (CDF) and the Probability Density Function (PDF) for the combined failure rates of the datasets. The CDF and PDF are two key tools in probability theory and statistics used to describe the distribution of continuous random variables. The graph indicates that the majority of the data lie towards the lower end of the range, which is characteristic of a right-skewed (positively skewed) distribution, with a mean of 36.1 FIT.

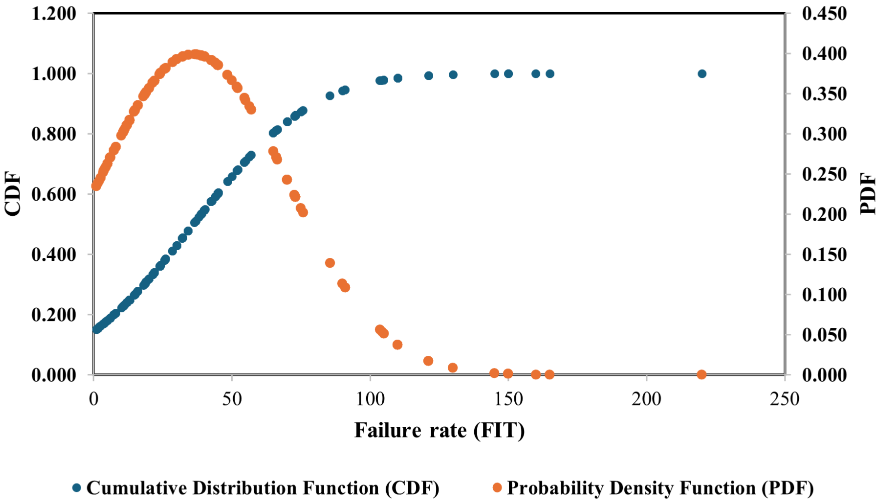


Figure 6.41: Data distribution – Component failure rate CDF and PDF illustrating right-skewed behaviour.

The current FIS designs employ equally spaced fuzzy sets within the membership functions. The above Pareto chart and distribution graph indicate that the data are not equally distributed; therefore, there is a potential for error in the lower end of the range. To improve Trapezoidal 6x6 FIS design, the membership function partitions were modified to accommodate the skewed failure rate data to allow a more representative inference.

The following section discusses these modifications implemented using MATLAB, a robust and versatile software platform widely used for engineering and scientific computations. MATLAB's advanced features and extensive toolbox for fuzzy logic and data analysis make it ideal for developing and fine-tuning the FIS.

6.4 Use of MATLAB for FIS Design Enhancements

MATLAB (version R2024a) was used to redesign the FIS. MATLAB includes a *Fuzzy Logic Toolbox* which is a more advanced application than the JuzzyOnline [151] toolbox and allows a more efficient process to optimise FIS designs. It is important to note that the MATLAB Fuzzy Logic Toolbox and the JuzzyOnline Toolbox have been shown to produce comparable results [153]. The Fuzzy Logic Toolbox also includes applications such as 'System Validation' and 'Error Data', which allow multiple inputs and outputs to be analysed simultaneously. The system validation application provides automatic charting for inputs and outputs, including a line chart showing errors between data points with an RMSE calculation. This allows for quick interpretation of the FIS design and identification of potential enhancement modifications. The following section describes the modification process and discusses the results of the enhanced FIS design.

6.4.1 FIS Design Enhancements to Improve Low Range Sensitivity

6.4.1.1 Membership functions

Figure 6.42 shows the histogram, based on which the partitioning of the Trapezoidal 6x6 membership functions was adjusted to accommodate the skewed failure rate data. Six bins were used to represent the six membership functions of the fuzzy sets of the Component failure rate input and the Weighted revealed failure rate output.

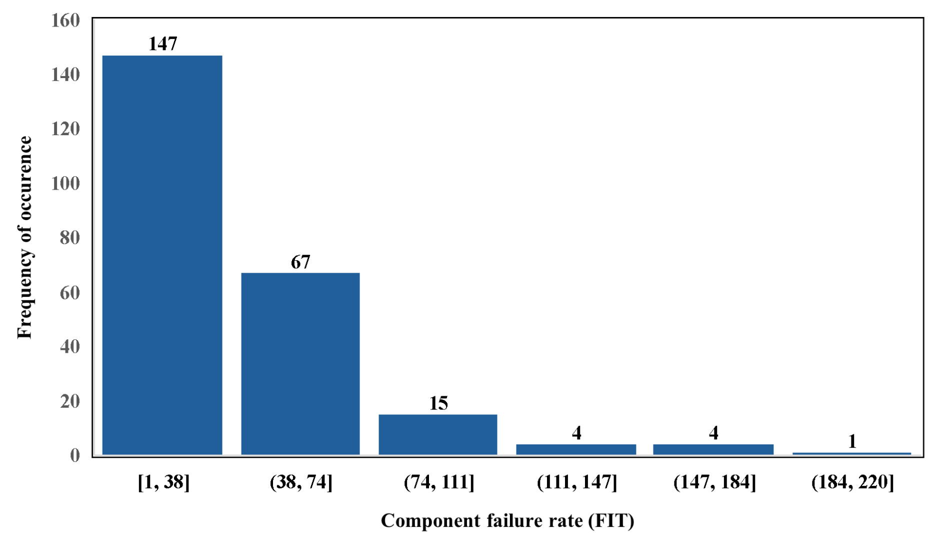


Figure 6.42: Histogram of failure rates

Table 6.18 shows the range of occurrences that were used to determine the approximate ranges of the membership functions, in which three membership functions occupy the first bin (Bin # 1) of the failure rates to improve the sensitivity at the lower end of the failure rate range.

Table 6.18: Membership function ranges

Membership function	**Range of failure rates (FIT)**	
Very Very Low	0	10
Very Low	11	20
Low	21	40
Medium	41	70
High	71	120
Very High	121	220

From the data in Table 6.18 it was then possible to create membership functions for both the component failure rate input and the weighted revealed failure rate output, as shown in Figure 6.43 and Figure 6.44, respectively.

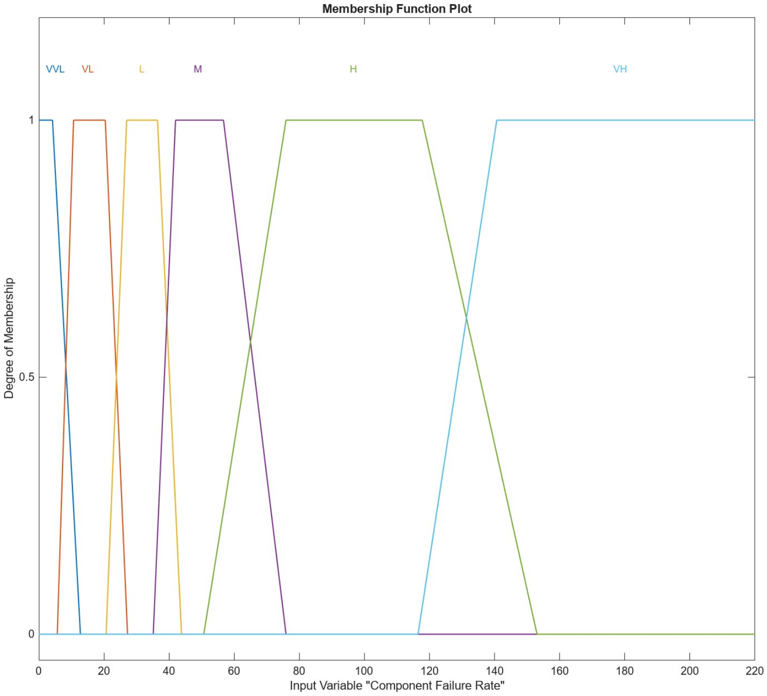


Figure 6.43: Modified Component failure rate membership function

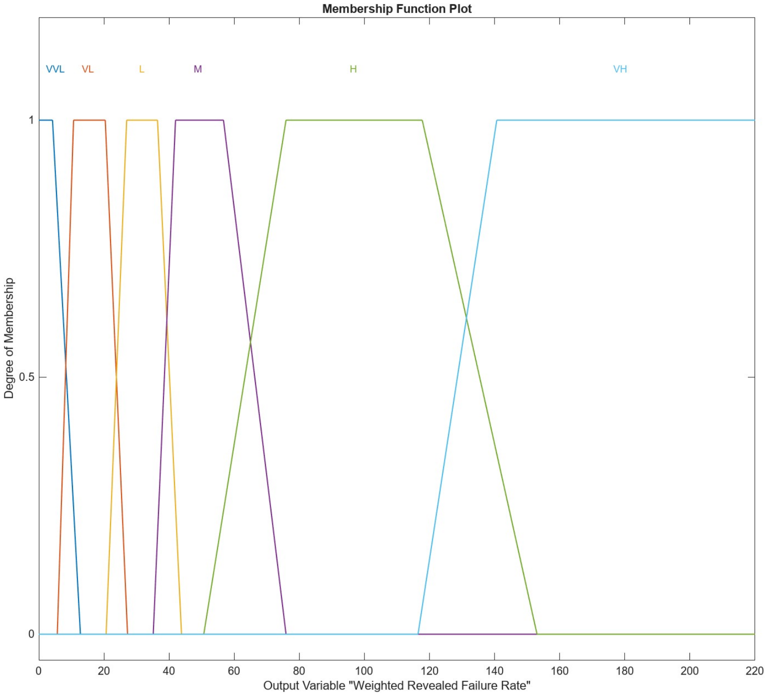


Figure 6.44: Modified Weighted revealed failure rate membership function

Figure 6.45 illustrates that the partitioning of the probability of revealing a failure fuzzy sets remained unchanged.

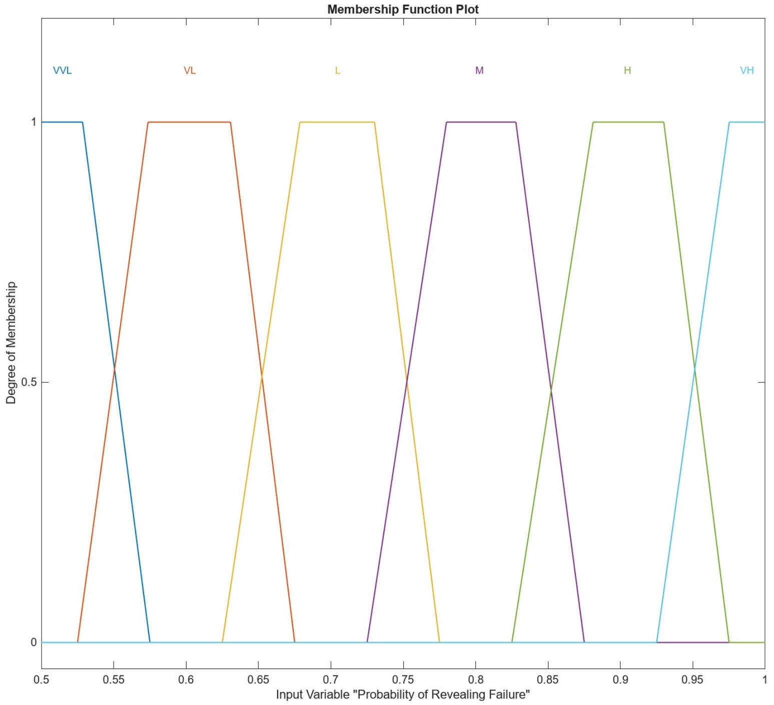


Figure 6.45: Unmodified Probability of revealing the failure membership function

6.4.1.2 Development of rule base using new membership function ranges

The IF-THEN rule base was developed using the mid-range values of each membership function as shown in Table 6.19 and Table 6.20.

Table 6.19: Component failure rate (FIT) membership function ranges

Membership function	Low	High	Mid
Very Very Low	0	10	5.0
Very Low	11	20	15.5
Low	21	40	30.5
Medium	41	70	55.5
High	71	120	95.5
Very High	121	220	170.5

Table 6.20: Probability of revealing a failure membership function ranges

Membership function	**Low**	**High**	**Mid**
Very Very Low	0.5	0.58	0.54
Very Low	0.59	0.67	0.63
Low	0.68	0.76	0.72
Medium	0.77	0.85	0.81
High	0.86	0.94	0.90
Very High	0.95	1	0.98

Table 6.21 shows the mid-range values that were used to calculate the output values to obtain the equivalent output fuzzy set.

Table 6.21: Determination of the weighted revealed failure rates (FIT) (VVL = Very Very Low, VL = Very Low, L = Low, M = Medium, H = High, VH = Very High)

Rule #	Component failure rate (FIT)		Probability of revealing failure		Weighted revealed failure rate (FIT)	
1	VVL	5	VVL	0.54	3	VVL
2	VL	15.5	VVL	0.54	8	VVL
3	L	30.5	VVL	0.54	16	VL
4	M	55.5	VVL	0.54	30	L
5	H	95.5	VVL	0.54	52	M
6	VH	170.5	VVL	0.54	92	H
7	VVL	5	VL	0.63	3	VVL
8	VL	15.5	VL	0.63	10	VVL
9	L	30.5	VL	0.63	19	VL
10	M	55.5	VL	0.63	35	L
11	H	95.5	VL	0.63	60	M
12	VH	170.5	VL	0.63	107	H
13	VVL	5	L	0.72	4	VVL
14	VL	15.5	L	0.72	11	VL
15	L	30.5	L	0.72	22	L
16	M	55.5	L	0.72	40	L
17	H	95.5	L	0.72	69	M
18	VH	170.5	L	0.72	123	VH
19	VVL	5	M	0.81	4	VVL
20	VL	15.5	M	0.81	13	VL
21	L	30.5	M	0.81	25	L
22	M	55.5	M	0.81	45	M
23	H	95.5	M	0.81	77	H
24	VH	170.5	M	0.81	138	VH
25	VVL	5	H	0.9	5	VVL
26	VL	15.5	H	0.9	14	VL
27	L	30.5	H	0.9	27	L
28	M	55.5	H	0.9	50	M
29	H	95.5	H	0.9	86	H
30	VH	170.5	H	0.9	153	VH
31	VVL	5	VH	0.975	5	VVL
32	VL	15.5	VH	0.975	15	VL
33	L	30.5	VH	0.975	30	L
34	M	55.5	VH	0.975	54	M
35	H	95.5	VH	0.975	93	H
36	VH	170.5	VH	0.975	166	VH

From Table 6.21, 36 IF-THEN rules were obtained. Examples of the rules are presented in the following.

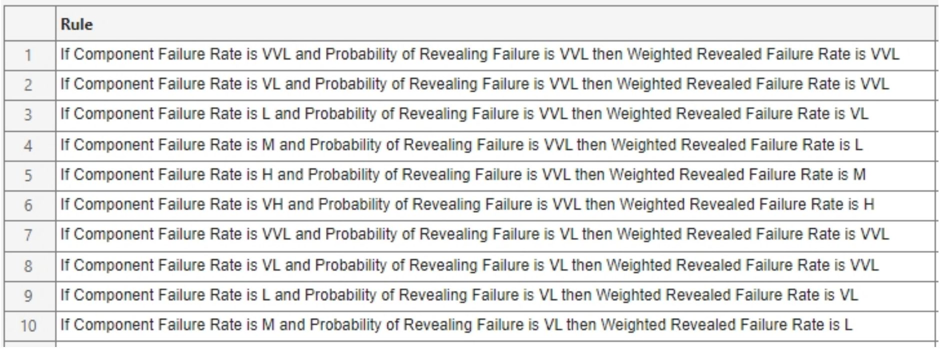
Rule #8 = IF **Component failure rate** is *Very Low* AND the **Probability of revealing the failure** is *Very Low* THEN the **Weighted revealed failure rate** is

Very Very Low

Rule #12 = IF **Component failure rate** is *Very High* AND the **Probability of revealing the failure** is *Very Low* THEN the **Weighted revealed failure rate** is *High*

Table 6.22 shows an excerpt of the rule base developed in MATLAB.

Table 6.22: Example of the rule base developed in MATLAB



	Rule
1	If Component Failure Rate is VVL and Probability of Revealing Failure is VVL then Weighted Revealed Failure Rate is VVL
2	If Component Failure Rate is VL and Probability of Revealing Failure is VVL then Weighted Revealed Failure Rate is VVL
3	If Component Failure Rate is L and Probability of Revealing Failure is VVL then Weighted Revealed Failure Rate is VL
4	If Component Failure Rate is M and Probability of Revealing Failure is VVL then Weighted Revealed Failure Rate is L
5	If Component Failure Rate is H and Probability of Revealing Failure is VVL then Weighted Revealed Failure Rate is M
6	If Component Failure Rate is VH and Probability of Revealing Failure is VVL then Weighted Revealed Failure Rate is H
7	If Component Failure Rate is VVL and Probability of Revealing Failure is VL then Weighted Revealed Failure Rate is VVL
8	If Component Failure Rate is VL and Probability of Revealing Failure is VL then Weighted Revealed Failure Rate is VVL
9	If Component Failure Rate is L and Probability of Revealing Failure is VL then Weighted Revealed Failure Rate is VL
10	If Component Failure Rate is M and Probability of Revealing Failure is VL then Weighted Revealed Failure Rate is L

The overall MATLAB FIS design showing the inputs and outputs is shown in Figure 6.46.

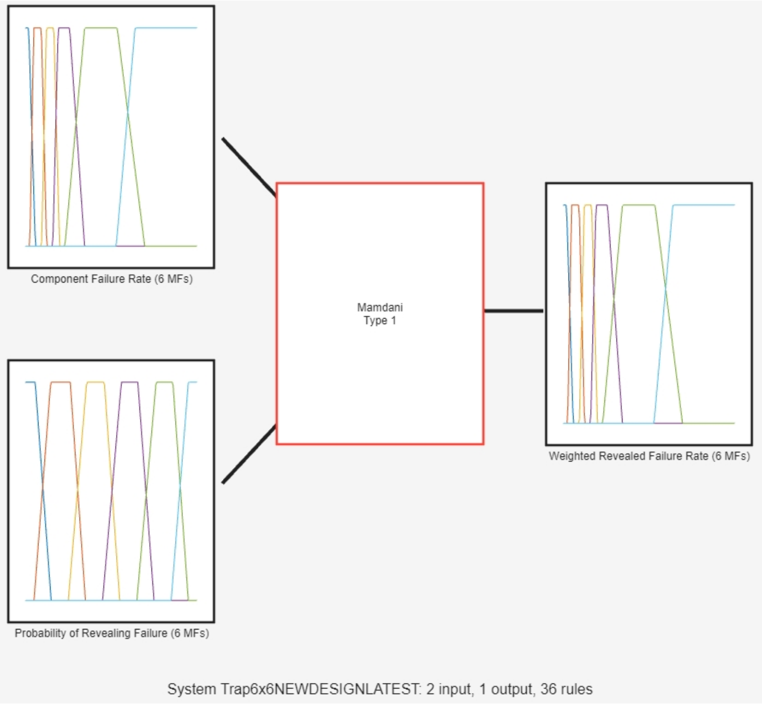

Component Failure Rate (6 MFs)
Mamdani
Type 1
Weighted Revealed Failure Rate (6 MFs)
Probability of Revealing Failure (6 MFs)
System Trap6x6NEWDESIGNLATEST: 2 input, 1 output, 36 rules

Figure 6.46: Overview of enhanced FIS design developed in MATLAB

6.4.1.3 Testing of the enhanced FIS design using random numbers

As discussed earlier, testing was undertaken using random numbers. Random numbers within the input membership function universe of discourse were generated using Excel, i.e., Component failure rate 1 to 220 FIT and Probability of revealing a failure 0.5 to 1. The results of the testing showed that the RMSE value was relatively high at 20.925 (MSE = 438). Figure 6.47 shows the inputs of random numbers and the associated outputs. The first chart 'Input(1)' shows the random data for the Component Failure Rate (FIT). The second chart 'Input(2)' shows the random data for the Probability of revealing the failure. The third chart 'Output (1)' shows the output of the FMEDA identified by Output(1)[Reference] (blue line) and the equivalent FIS output identified by Output(1)[Trap6x6NEWDESIGNLATEST] (orange line). The fourth chart 'Output(1)Error' shows the error between the data points of the outputs of the FMEDA (observed values) and the FIS (expected values). The actual RMSE value is also shown in this chart.

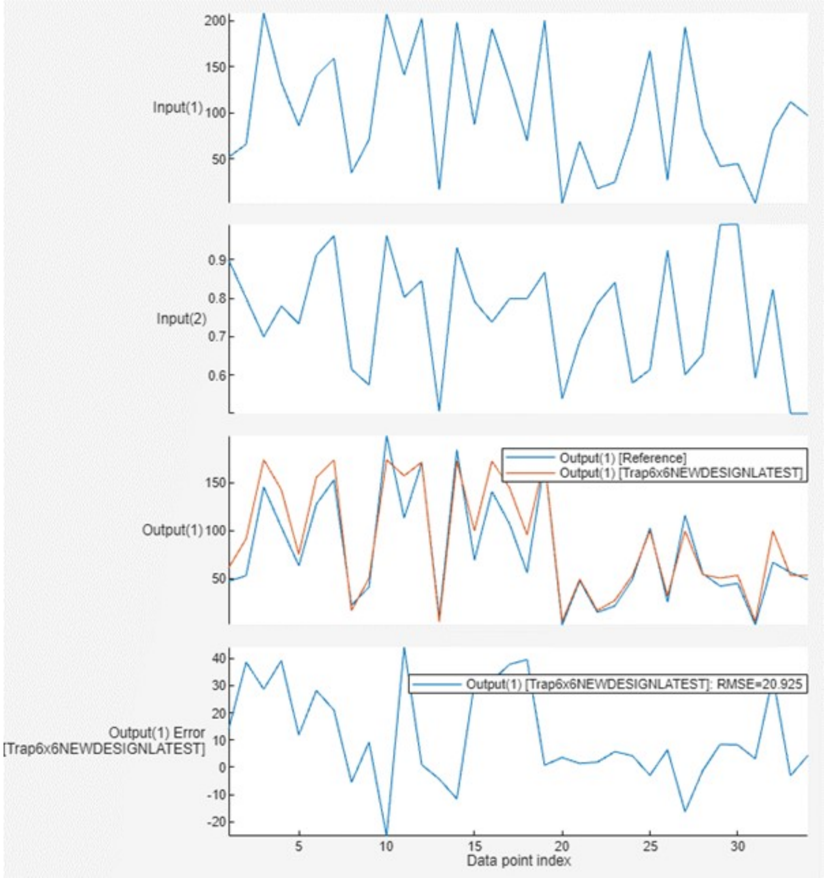


Figure 6.47: MATLAB FIS test charts showing random data points for component failure rate and probability of revealing the failure (Input 1 and 2), the FMEDA and FIS outputs (Output 1), and the RMSE (Output 1 Error).

From the FMEDA v FIS output chart, Figure 6.48, it is evident that the FIS design is limited on the mid to high end and causing the relatively high RMSE. The red arrow indicates the different values between the FMEDA and FIS outputs, 103.7 FIT and 142.8 FIT respectively, an error of 39.1 FIT.

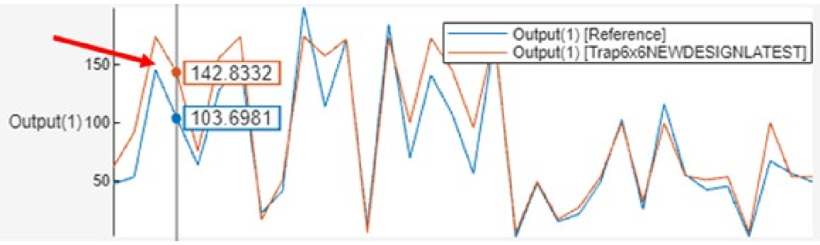


Figure 6.48: MATLAB FMEDA and FIS outputs chart showing error in mid to high FIT values

6.4.2 FIS Design Enhancements to Improve Mid to High Range Sensitivity

To improve the sensitivity of the medium to high range, the fuzzy set of component failure rates with the Medium (M) and High (H) membership functions were modified to cover a greater range of component failure rates and the Very High (VH) membership function was reduced in size, as shown in Figure 6.49.

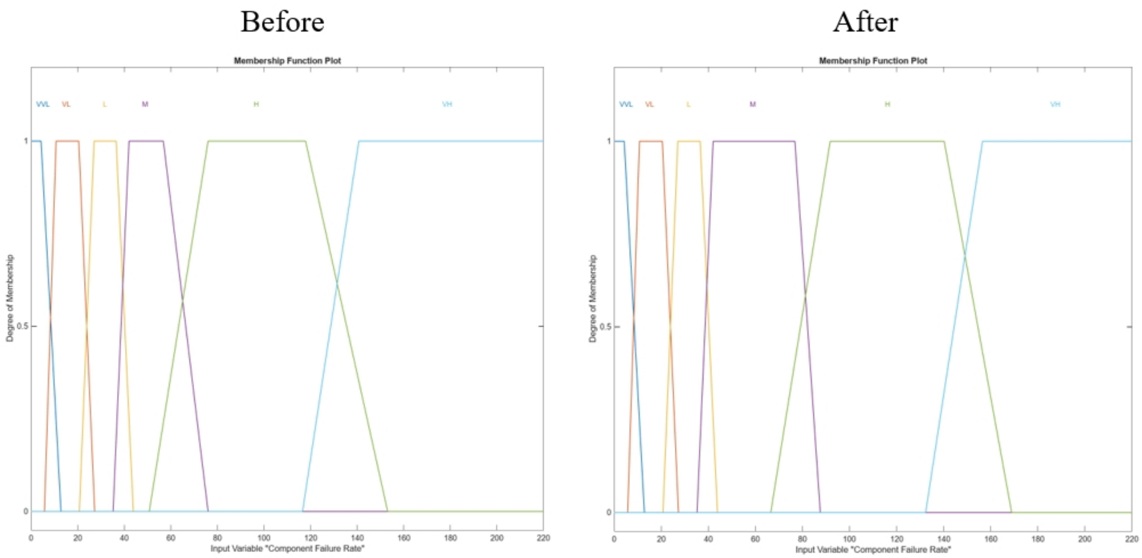


Figure 6.49: Modification of the Component failure rate Medium, High and Very High fuzzy sets

The FIS was retested using random numbers and the results showed less error in the mid to high range, as shown in Figure 6.50.

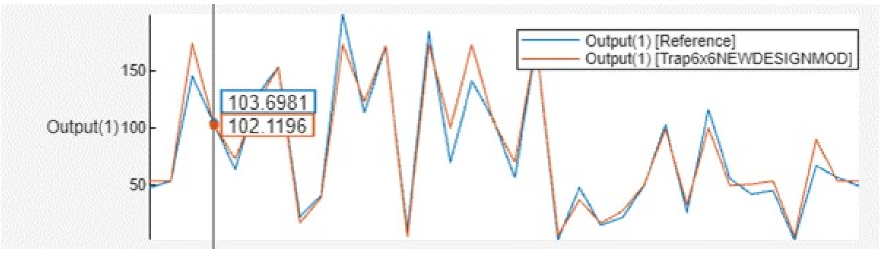


Figure 6.50: MATLAB FMEDA and FIS outputs chart after modification

The RMSE value improved from the original 20.952 FIT to 12.538 FIT as shown in Figure 6.51.

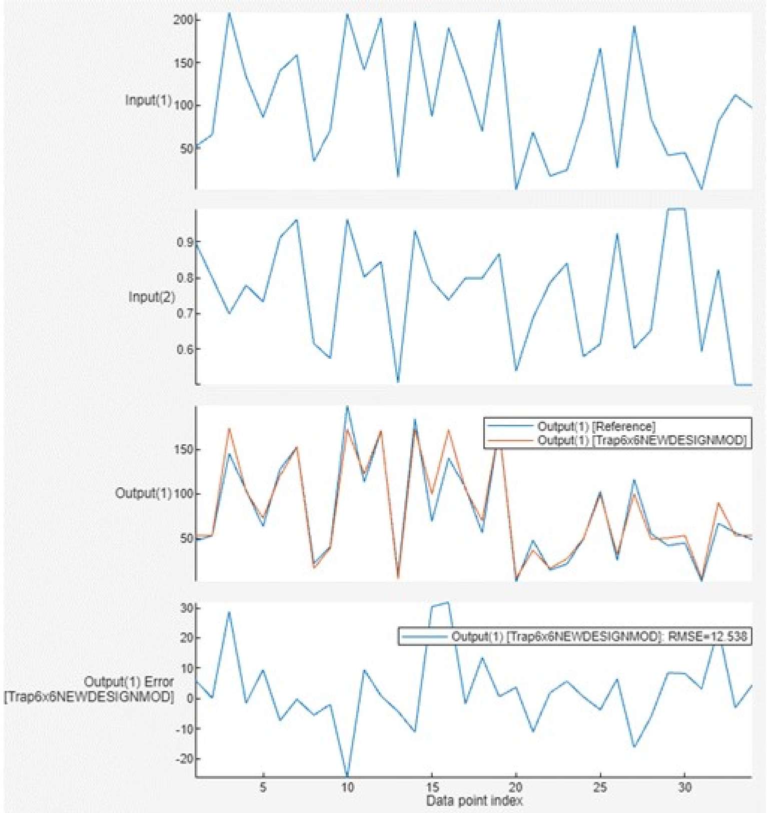
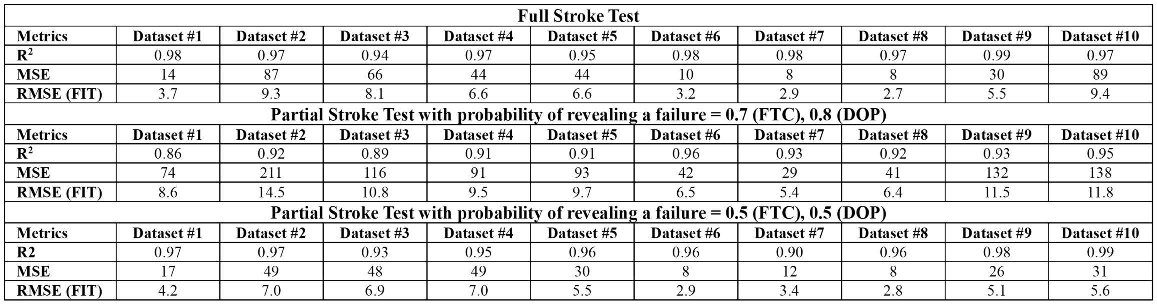


Figure 6.51: MATLAB FMEDA and FIS outputs RMSE chart

6.4.3 Validation of the Enhanced FIS Design Using Expert Datasets

Validation was carried out using expert data, based on a Full Stroke (FST) and Partial Stroke Tests (PST). The results of the validation are presented in the following. Table 6.23 shows the analysis metrics for FST and PST with probability of revealing a failure of 0.7 (FTC) and 0.8 (DOP) as per section above, and metrics for PST for the low end of the probability of revealing a failure range, i.e., 0.5 for both FTC and DOP failure modes.

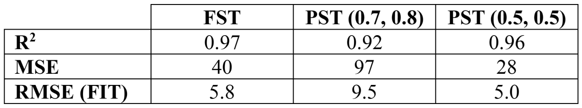
Table 6.23: Full stroke test and Partial stroke test validation metrics



Full Stroke Test										
Metrics	**Dataset #1**	**Dataset #2**	**Dataset #3**	**Dataset #4**	**Dataset #5**	**Dataset #6**	**Dataset #7**	**Dataset #8**	**Dataset #9**	**Dataset #10**
R^2	0.98	0.97	0.94	0.97	0.95	0.98	0.98	0.97	0.99	0.97
MSE	14	87	66	44	44	10	8	8	30	89
RMSE (FIT)	3.7	9.3	8.1	6.6	6.6	3.2	2.9	2.7	5.5	9.4
Partial Stroke Test with probability of revealing a failure = 0.7 (FTC), 0.8 (DOP)										
Metrics	**Dataset #1**	**Dataset #2**	**Dataset #3**	**Dataset #4**	**Dataset #5**	**Dataset #6**	**Dataset #7**	**Dataset #8**	**Dataset #9**	**Dataset #10**
R^2	0.86	0.92	0.89	0.91	0.91	0.96	0.93	0.92	0.93	0.95
MSE	74	211	116	91	93	42	29	41	132	138
RMSE (FIT)	8.6	14.5	10.8	9.5	9.7	6.5	5.4	6.4	11.5	11.8
Partial Stroke Test with probability of revealing a failure = 0.5 (FTC), 0.5 (DOP)										
Metrics	**Dataset #1**	**Dataset #2**	**Dataset #3**	**Dataset #4**	**Dataset #5**	**Dataset #6**	**Dataset #7**	**Dataset #8**	**Dataset #9**	**Dataset #10**
R2	0.97	0.97	0.93	0.95	0.96	0.96	0.90	0.96	0.98	0.99
MSE	17	49	48	49	30	8	12	8	26	31
RMSE (FIT)	4.2	7.0	6.9	7.0	5.5	2.9	3.4	2.8	5.1	5.6

The validation results indicate high regression and low error values for full and partial testing. Table 6.24 shows the average values of the metrics.

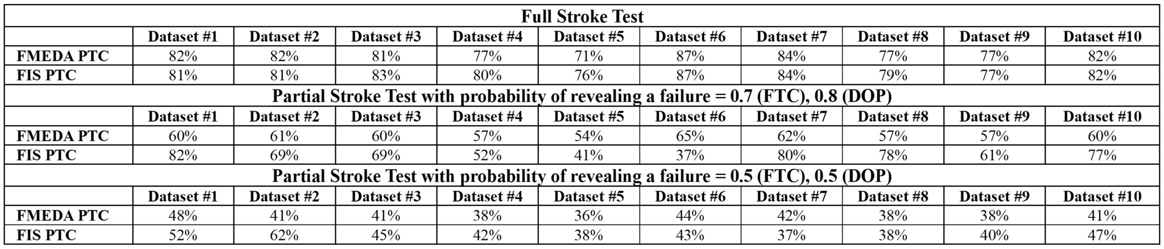
Table 6.24: Validation metrics average values



	FST	**PST (0.7, 0.8)**	**PST (0.5, 0.5)**
R^2	0.97	0.92	0.96
MSE	40	97	28
RMSE (FIT)	5.8	9.5	5.0

Table 6.25 shows the proof test coverage results for FMEDA and FIS for full stroke test (FST) and partial stroke test (PST) variants.

Table 6.25: FMEDA and FIS Full stroke test and Partial stroke test PTC values



Full Stroke Test										
	Dataset #1	**Dataset #2**	**Dataset #3**	**Dataset #4**	**Dataset #5**	**Dataset #6**	**Dataset #7**	**Dataset #8**	**Dataset #9**	**Dataset #10**
FMEDA PTC	82%	82%	81%	77%	71%	87%	84%	77%	77%	82%
FIS PTC	81%	81%	83%	80%	76%	87%	84%	79%	77%	82%
Partial Stroke Test with probability of revealing a failure = 0.7 (FTC), 0.8 (DOP)										
	Dataset #1	**Dataset #2**	**Dataset #3**	**Dataset #4**	**Dataset #5**	**Dataset #6**	**Dataset #7**	**Dataset #8**	**Dataset #9**	**Dataset #10**
FMEDA PTC	60%	61%	60%	57%	54%	65%	62%	57%	57%	60%
FIS PTC	82%	69%	69%	52%	41%	37%	80%	78%	61%	77%
Partial Stroke Test with probability of revealing a failure = 0.5 (FTC), 0.5 (DOP)										
	Dataset #1	**Dataset #2**	**Dataset #3**	**Dataset #4**	**Dataset #5**	**Dataset #6**	**Dataset #7**	**Dataset #8**	**Dataset #9**	**Dataset #10**
FMEDA PTC	48%	41%	41%	38%	36%	44%	42%	38%	38%	41%
FIS PTC	52%	62%	45%	42%	38%	43%	37%	38%	40%	47%

The proof test coverage results for the full stroke test show a maximum difference of approximately 5% between FMEDA and FIS. The partial stroke test with 0.7 and 0.8 probability of revealing a failure resulted in a maximum difference between FMEDA and FIS of 28% and the partial stroke test with 0.5 probability of a maximum difference of 21%. The results of the partial stroke test using the probability of 0.7 and 0.8 of revealing a failure have PTC values, in some cases, similar to those of FST. On closer inspection, it appeared that the FIS did not differentiate between the probability of revealing failure values from 0.8 to 1, which was the main cause of such optimistic PTC values. In order to improve the accuracy in this range, the design was modified further.

6.4.4 Final FIS Design Enhancement to Improve Sensitivity of the Probability of Revealing a Failure

The modified design included an additional membership function at the top end of the discourse universe of the failure rate of the component to create a 7x6 FIS, shown in Figure 6.52, which included 42 rules as opposed to 35 in the previous design. The rules were developed as previously described.

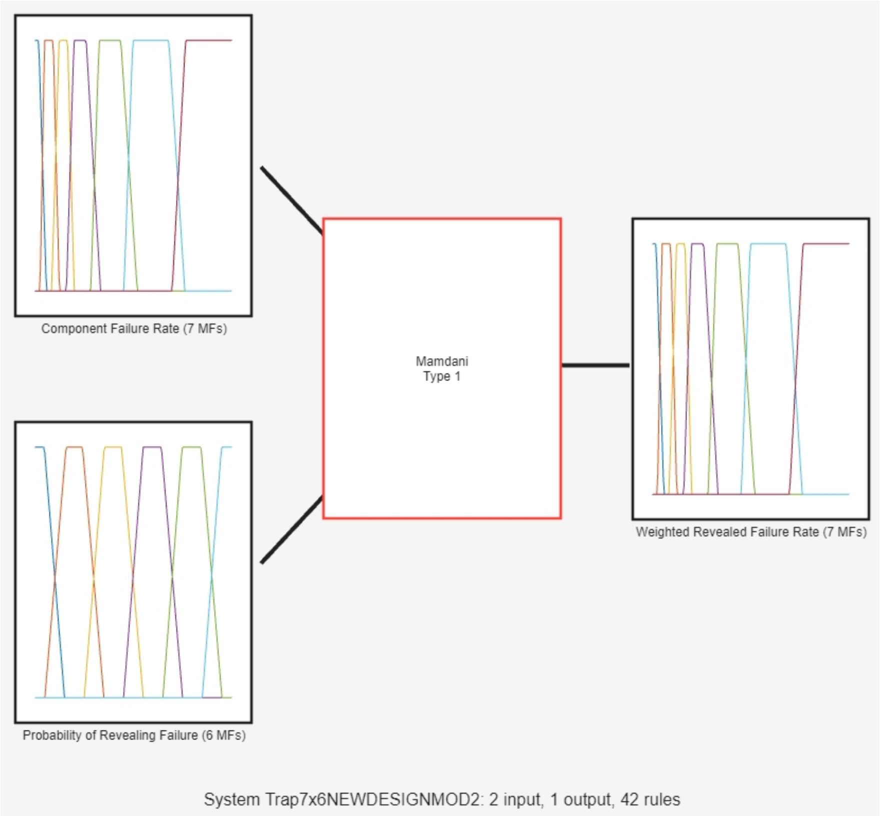

Component Failure Rate (7 MFs)
Mamdani
Type 1
Weighted Revealed Failure Rate (7 MFs)
Probability of Revealing Failure (6 MFs)
System Trap7x6NEWDESIGNMOD2: 2 input, 1 output, 42 rules

Figure 6.52: FIS overview of the 7x6 membership function design

Figure 6.53 shows the output of the FIS design that was tested using random numbers. The chart indicates a small improvement in RMSE compared to the previous 6x6 design. The 6x6 design RMSE = 12.5 FIT, compared to the 7x6 design RMSE = 11.1 FIT.

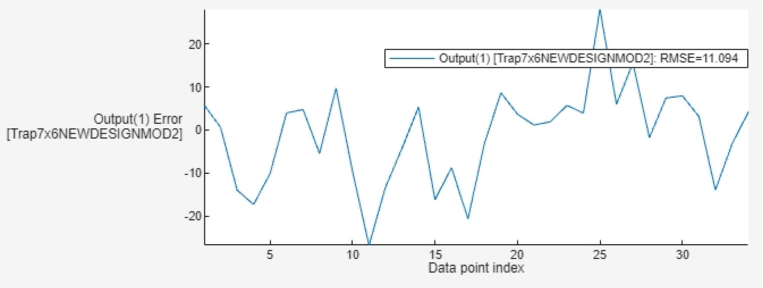
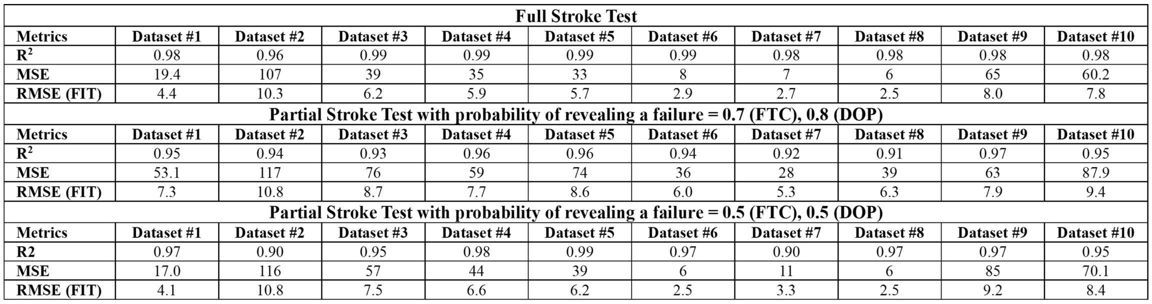


Figure 6.53: FIS 7x6 Random Numbers RMSE chart

Table 6.26 shows the results of the validation of the 7x6 FIS design. The R^2 values range from 0.9 to 0.99, indicating a strong fit between the datasets. The MSE values range from 6 to 117, while the RMSE values vary between 2.5 and 10.8, reflecting differences in prediction accuracy.

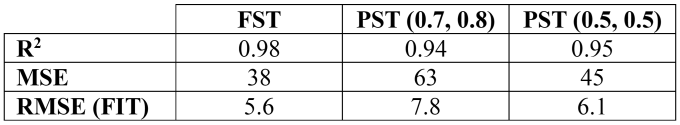
Table 6.26: Full stroke test (7x6 design) validation metrics



Full Stroke Test										
Metrics	**Dataset #1**	**Dataset #2**	**Dataset #3**	**Dataset #4**	**Dataset #5**	**Dataset #6**	**Dataset #7**	**Dataset #8**	**Dataset #9**	**Dataset #10**
R^2	0.98	0.96	0.99	0.99	0.99	0.99	0.98	0.98	0.98	0.98
MSE	19.4	107	39	35	33	8	7	6	65	60.2
RMSE (FIT)	4.4	10.3	6.2	5.9	5.7	2.9	2.7	2.5	8.0	7.8
Partial Stroke Test with probability of revealing a failure = 0.7 (FTC), 0.8 (DOP)										
Metrics	**Dataset #1**	**Dataset #2**	**Dataset #3**	**Dataset #4**	**Dataset #5**	**Dataset #6**	**Dataset #7**	**Dataset #8**	**Dataset #9**	**Dataset #10**
R^2	0.95	0.94	0.93	0.96	0.96	0.94	0.92	0.91	0.97	0.95
MSE	53.1	117	76	59	74	36	28	39	63	87.9
RMSE (FIT)	7.3	10.8	8.7	7.7	8.6	6.0	5.3	6.3	7.9	9.4
Partial Stroke Test with probability of revealing a failure = 0.5 (FTC), 0.5 (DOP)										
Metrics	**Dataset #1**	**Dataset #2**	**Dataset #3**	**Dataset #4**	**Dataset #5**	**Dataset #6**	**Dataset #7**	**Dataset #8**	**Dataset #9**	**Dataset #10**
R2	0.97	0.90	0.95	0.98	0.99	0.97	0.90	0.97	0.97	0.95
MSE	17.0	116	57	44	39	6	11	6	85	70.1
RMSE (FIT)	4.1	10.8	7.5	6.6	6.2	2.5	3.3	2.5	9.2	8.4

The validation results indicate improved R^2, MSE and RMSE values for the full and partial tests. Table 6.27 shows the average values of the metrics:

Table 6.27: FIS (7x6 design) validation metrics average values



	FST	**PST (0.7, 0.8)**	**PST (0.5, 0.5)**
R^2	0.98	0.94	0.95
MSE	38	63	45
RMSE (FIT)	5.6	7.8	6.1

Table 6.28 shows the Proof Test Coverage (PTC) results for the Full Stroke Test (FST) and Partial Stroke Test (PST) variants. The FST values show that the maximum

difference between FMEDA and FIS PTC is 3% for Dataset #8. For PST with a probability of 0.7 (FTC) and 0.8 (DOP), the maximum difference between FMEDA and FIS PTC is 22% also for Dataset #8. In contrast, for PST with a probability of 0.5 (FTC & DOP), the maximum difference is 6% for both Dataset #7 and Dataset #10.

Table 6.28: Full stroke test and Partial stroke test (7x6 design) PTC values

Full Stroke Test										
	Dataset #1	Dataset #2	Dataset #3	Dataset #4	Dataset #5	Dataset #6	Dataset #7	Dataset #8	Dataset #9	Dataset #10
FMEDA PTC	82%	82%	81%	77%	71%	87%	84%	77%	77%	82%
FIS PTC	81%	83%	81%	76%	72%	88%	84%	80%	77%	83%
Partial Stroke Test with probability of revealing a failure = 0.7 (FTC), 0.8 (DOP)										
	Dataset #1	Dataset #2	Dataset #3	Dataset #4	Dataset #5	Dataset #6	Dataset #7	Dataset #8	Dataset #9	Dataset #10
FMEDA PTC	60%	61%	60%	57%	54%	65%	62%	57%	57%	60%
FIS PTC	70%	58%	61%	59%	61%	76%	80%	79%	54%	57%
Partial Stroke Test with probability of revealing a failure = 0.5 (FTC), 0.5 (DOP)										
	Dataset #1	Dataset #2	Dataset #3	Dataset #4	Dataset #5	Dataset #6	Dataset #7	Dataset #8	Dataset #9	Dataset #10
FMEDA PTC	41%	41%	41%	38%	36%	44%	42%	38%	38%	41%
FIS PTC	43%	41%	44%	40%	37%	42%	36%	38%	40%	47%

Figure 6.54 shows the absolute differences between the FMEDA and FIS PTC results, indicating that the FST results have the smallest differences and the PST (0.7, 0.8) have the greatest differences.

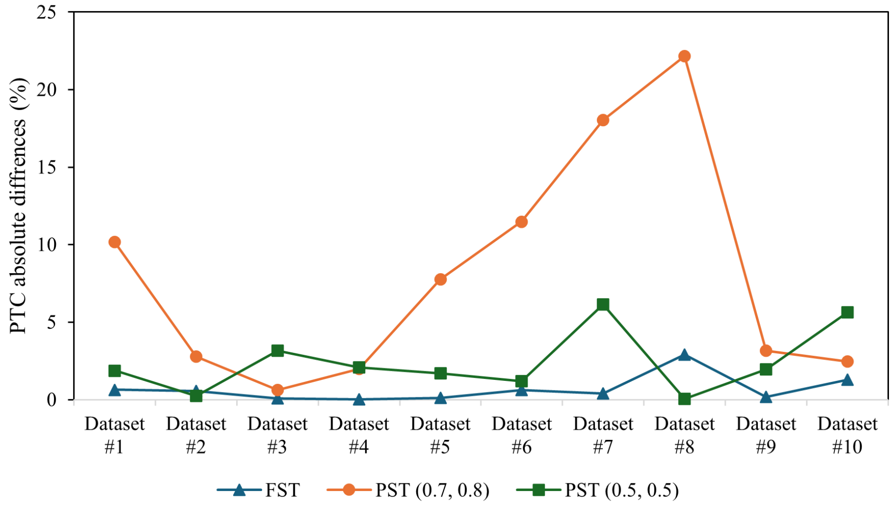


Figure 6.54: Absolute differences between FMEDA and FIS PTC results

6.4.5 Comparison of the Trapezoidal 6x6 and 7x6 Designs

A control surface diagram is a graphical representation that shows how the input variables in a Fuzzy Inference System (FIS) are mapped to the output based on the rule base. This diagram provides a comprehensive overview of the system and helps visualise the relationships between inputs and outputs, offering valuable insights into the system's behaviour. The inputs are the Component failure rate (*x-axis*) and the Probability of revealing a failure (*z-axis*) and the output is the weighted revealed failure rate (*y-axis*).

Figures 6.55 and Figure 6.56 show the control surface diagrams for the FIS 6x6 and 7x6 designs, respectively. The most prominent observations between the two diagrams are the higher end values of the weighted revealed failure rate (highlighted in yellow) and the smoother slope (more steps) of the weighted revealed failure rate range. The 6x6 design (Figure 6.55) shows minimal variation in the probability of revealing failure values of 0.7 to 1, while 7x6 has improved this area.

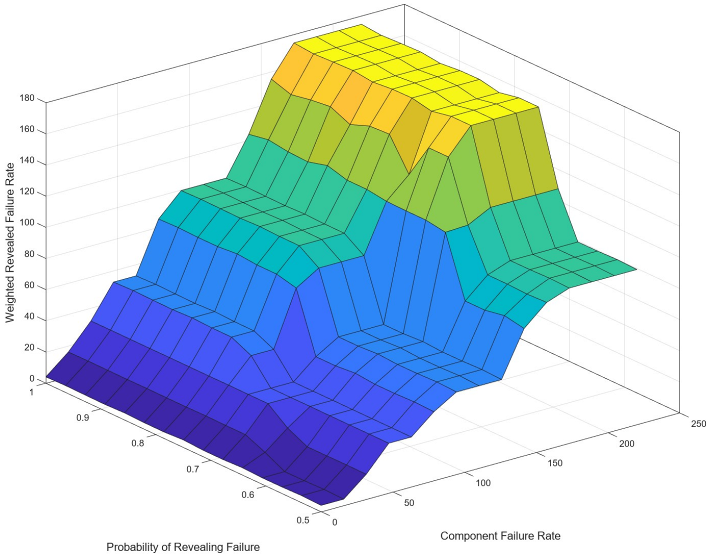


Figure 6.55: FIS 6x6 design - control surface diagram

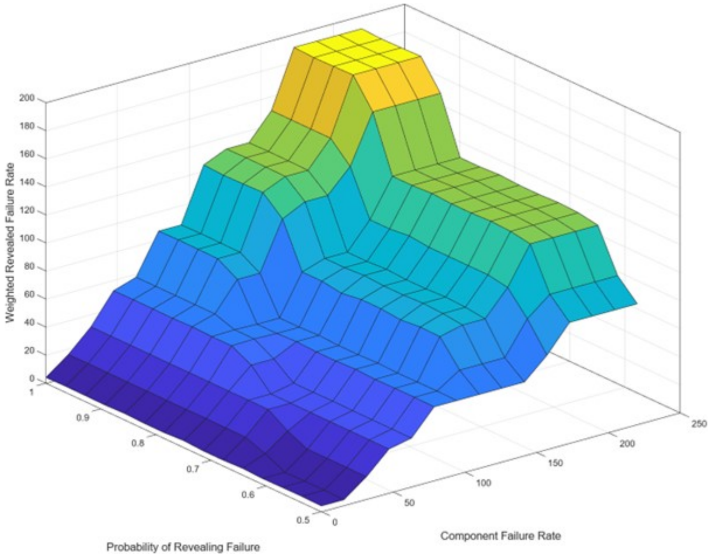


Figure 6.56: FIS 7x6 design - control surface diagram

The aim is to design an FIS that has smooth contours, where the transition from one value to the next is seamless, as idealised (with red lines) in the control surface diagram (Figure 6.57). The issue is that the effort required and the complexity of the design to attain such performance would not be justified.

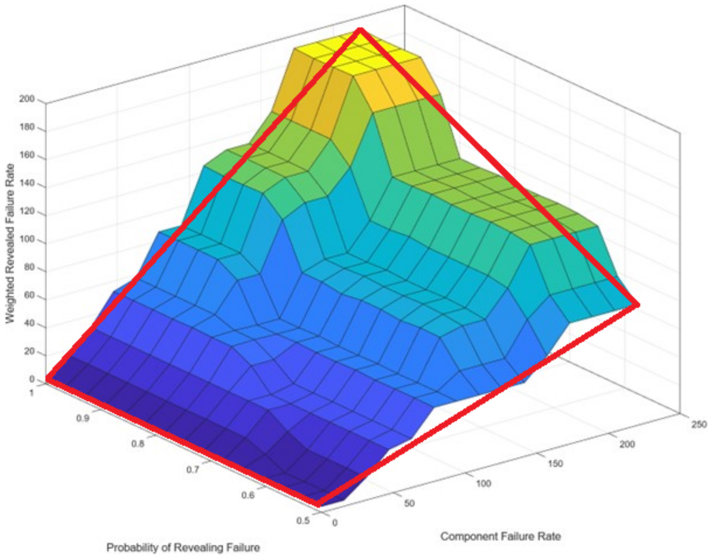


Figure 6.57: FIS design - control surface diagram with idealised contours shown in red.

Figures 6.58 and Figure 6.59 show the output and RMSE charts for the FIS 6x6 and 7x6, respectively, with extended testing using 10,000 random numbers. The random numbers are within the universe of discourse for both FMEDA and FIS inputs, i.e., the component failure rate and the probability of revealing a failure. The figures have two charts; the first chart (a) 'Output(1)', this shows the FMEDA output (observed values) which is the blue line, and the FIS output (expected values) is the orange line. The second chart (b) Output(1) Error shows the mean error (RMSE) between the observed values of FMEDA and the expected values of FIS.

The charts indicate, from extended testing, that the 7x6 design has an improved RMSE value of 12.615 FIT compared to the 6x6 design, which is 16.322. These charts also highlight the limitations of the designs, with the 6x6 design output (represented by the orange line) capped at around 170 FIT, whereas the 7x6 design extends to approximately 190 FIT. This further confirms the effectiveness of modifying the design to enhance high-end sensitivity by increasing the number of membership functions from 6 to 7, which subsequently increases the number of IF-THEN rules from 36 to 42.

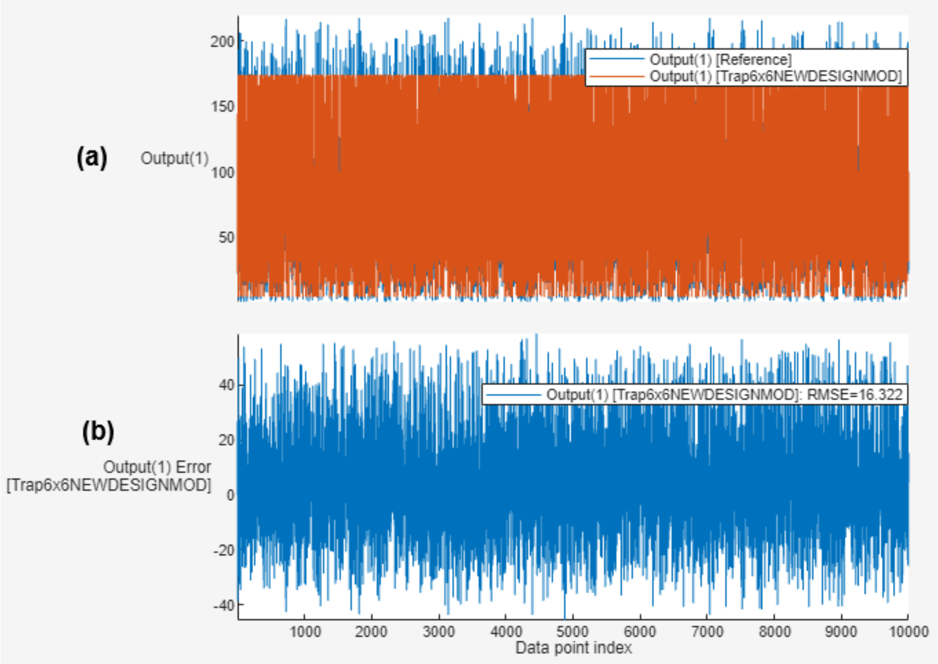


Figure 6.58: FIS 6x6 design with 10,000 random numbers test, (a) shows the FMEDA and FIS outputs and (b) shows the error (RMSE) between the FMEDA and FIS outputs

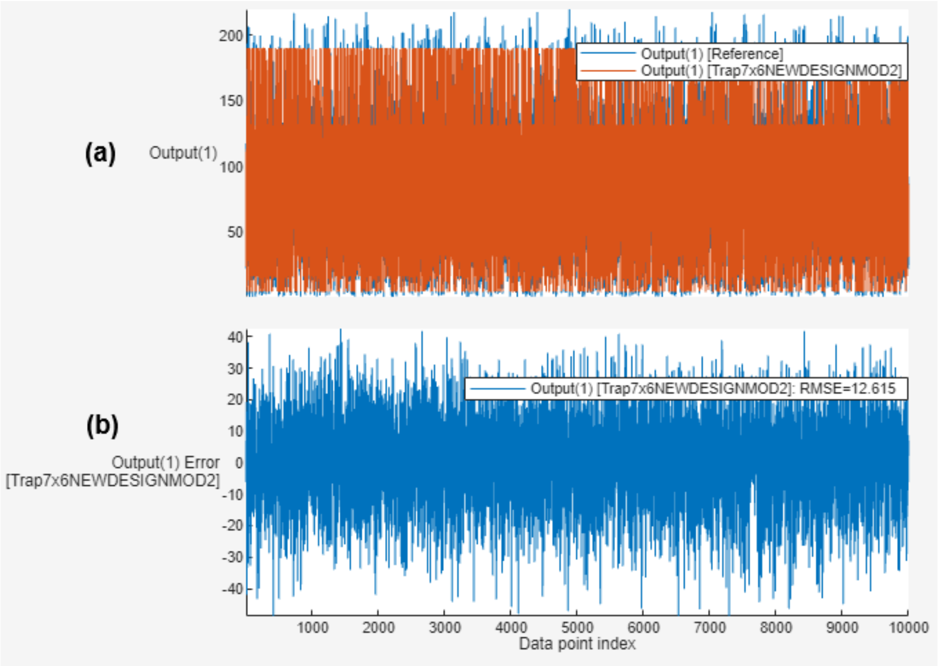


Figure 6.59: FIS 7x6 design with 10,000 random numbers test, (a) shows the FMEDA and FIS outputs and (b) shows the error (RMSE) between the FMEDA and FIS outputs

6.5 Discussion and Conclusions

This chapter proposed the use of a Fuzzy Inference System (FIS)to manage the uncertainty of data related to estimating proof test coverage using a Failure Modes, Effects, and Diagnostics Analysis (FMEDA). Various designs were tested using both Juzzy-Online's toolbox and MATLAB's Fuzzy Logic Toolbox. The final design of the fuzzy inference system, the Trapezoidal 7x6, metrics indicate that the model is very close to the expert data sets with an average R^2 value for the Full Stroke Test (FST) and Partial Stroke Test (PST) data of 0.96 with an average MSE of 49. The average RMSE value is approximately 7 FIT, which is less than 1% error when taking the average ESDV assembly failure rate used in this research (1358 FIT). Although it may be possible to improve the accuracy of the FIS model through further modifications, this would increase its complexity, and the effort required would likely outweigh the benefits achieved.

The Proof Test Coverage (PTC) for the FST suggests that there is only a 3% difference between the FMEDA and FIS versions for the data sets within this research (Table 6.28). This would likely have minimal impact on the PFD_{avg} or the proof test interval. However, since the validation data in this research are limited, specifically the number of expert datasets, it cannot be ruled out that other FMEDA data could result in FIS values where the PTC values differ significantly from those of the FMEDA.

In contrast, the results of the PST proof test coverage test indicate a significant difference between the FMEDA and FIS. The PST results that use probabilities of 0.7 and 0.8 to reveal a failure seem overly optimistic, with PTC estimates within 1% of the FST, implying that the PST could be considered an equivalent test to the FST. The PST results with a probability of 0.5 of revealing a failure were more conservative compared to other studies, with the maximum differences between FMEDA and FIS PTC being approximately 6% for Dataset #7 and Dataset #10.

The results of the PTC of the partial stroke test of FIS, in some cases, were 22% greater than the FMEDA, as shown in Table 6.28, probability of revealing failure = 0.7 (FTC) and 0.8 (DOP), for Dataset #8. The significance of this is that applying these findings to the PFD_{avg} calculations could lead to a reduction in the frequency of partial stroke testing and/or allow for an extension of the full stroke test interval. Although PST can improve the testing strategy and improve operational efficiency, it should not typically replace FST in safety-critical applications. Instead, it is often used in conjunction with FST to provide a more comprehensive assessment of valve health and functionality. The potential outcome of these results is a reduction in ESDV proof testing overhead, leading to maintenance cost savings for operating companies. For example, using the FIS PTC estimates for Datasets #5, #6 and #7 it is possible to extend the FST interval from one year to two years. Similarly, for Dataset #8 it is possible to extend both the FST and PST intervals, from six months to one year and from one year to two years, respectively. This suggests that ESDV proof test intervals can be extended by up to 100% when applying the FIS PTC estimates from this research.

The limitations of the FIS design include a component failure rate range constrained to 220 FIT and a probability of revealing a failure that spans from 0.5 to 1. The implications are that if component failure rates exceed 220 FIT, the FIS design will require modification, specifically by extending the universe of discourse to accommodate these higher rates. The distribution of failure rates, which is utilised to establish the spread of the membership functions within the fuzzy sets, must also be taken into account, as it could potentially influence the IF-THEN rules. The same applies to the probability of revealing a failure, as any probability exceeding the current range of 0.5 to 1 will require adjustments to the FIS design to accommodate these values.

The FIS full stroke test PTC results are in good agreement with published PTC estimates ranging from 72% to 88%. The FIS PST results with probabilities of 0.7 and 0.8 range from 54% to 80%, which is relatively high, especially considering that the maximum published PTC estimate is 70%. The FIS PST results with a probability of 0.5 in the range of 36% to 44%. Although this is relatively low, it remains within the range of published PTC estimates. The FIS full stroke test PTC results also align well with the findings of the PTC estimates derived from the ESDV data in Chapter 4, indicating the accuracy of the FIS method.

As discussed previously, the assessment of the ESDV equipment in this research was limited to the solenoid valve, actuator, and ball valve. In order to improve the accuracy of the PTC assessment, consideration should be given to all auxiliary device failures that could lead to a dangerous failure of an ESDV. This is particularly relevant for an ESDV designed to facilitate partial stroke testing, as the number of auxiliary devices required for such testing is significantly greater than that of the basic components of an ESDV, as illustrated in Appendix A.1. The inclusion of auxiliary devices in the assessment will not only increase the total dangerous failure rate of the ESDV, but is also likely to affect the PTC estimates.

If a fuzzy inference system were to be designed again for similar purposes, it would be beneficial to (1) use MATLAB for ease of implementation and design analysis capabilities (2) use trapezoidal membership functions for the initial design (3) analyse the component failure rate data to understand its distribution, establish the universe of discourse, and determine the partitioning of fuzzy sets. and (4) consider the number of fuzzy sets per input to be based on the results of testing and validation.

This chapter demonstrated that a fuzzy inference system can effectively address the subjectivity associated with FMEDA data, yielding some intriguing results despite the limited datasets used to design and validate the model. This work represents the preliminary steps in developing a tool for determining ESDV proof test coverage for both full and partial stroke testing. Additional work is necessary to improve the design by obtaining additional FMEDA datasets for validation. Subsequently, the FIS design can be modified as needed based on the analysis results.

6.5.1 Summary of Findings

The key findings of this chapter can be briefly summarised in the following points:

- Trapezoidal fuzzy sets were employed for the fuzzy inference system using seven membership functions for component failure rates and six for probability of revealing a failure. This FIS configuration produced the most favourable results, with R^2, MSE and RMSE values of 0.96, 49, and 7, respectively.

- The variation in proof test coverage between FMEDA and the FIS results were observed to be up to 3% for the full stroke test and as much as 22% for the partial stroke test. It is important to note that these results were based on a limited number of datasets.

- The generally higher proof test coverage values for partial stroke testing suggest opportunities for optimising proof testing schedules, which could lead to a reduction in maintenance overhead.

- In some cases, the FIS results demonstrate a 100% improvement in test intervals, potentially extending the typical partial stroke test interval from six months to one year and the full stroke test interval from one year to two years.

- MATLAB's Fuzzy Logic Toolbox is a powerful and efficient tool for designing fuzzy inference systems, offering a dynamic user interface and integrated tools for performance evaluation.

In summary, this chapter introduces a novel Fuzzy-FMEDA framework that successfully models subjective expert data, which are challenging for traditional FMEDA techniques to capture. Although no studies have directly applied fuzzy logic to FMEDA, related research in the use of FMEA and FMECA supports the potential of this approach to address key limitations in traditional FMEDA methods. This unique contribution provides a new approach to integrate expert judgement into safety-related assessments, paving the way for future studies to further validate and refine the model.

Human errors, such as incorrect execution of the proof test procedure, can lead to false positives or false negatives. This could result in dangerous failures being undetected, causing proof test coverage to be overestimated, giving a false sense of safety. The next chapter discusses the impact of human error on proof testing and the potential implications on proof test coverage.

Chapter 7

The Impact of Human Error on Proof Testing

7.1 Introduction

The previous chapter examined the use of Fuzzy Inference Systems (FIS) to address the uncertainty in data related to Failure Modes, Effects, and Diagnostics Analysis (FMEDA) for estimating the Proof Test Coverage (PTC) of an Emergency Shutdown Valve (ESDV). By utilising expert datasets and employing measures and metrics such as R^2, MSE, and RMSE, an FIS was developed to estimate PTC for both full stroke and partial stroke tests, effectively addressing data subjectivity. However, this approach did not account for the influence of human error during the execution of an ESDV proof test and its potential effect on PTC.

This chapter examines the impact of human error on ESDV proof testing and its effect on proof test coverage. It reviews existing methodologies for quantifying Human Error Probability (HEP) and evaluates the extent to which human error can influence the results of the proof test. Figure 7.1 provides an overview of the structure of this chapter.

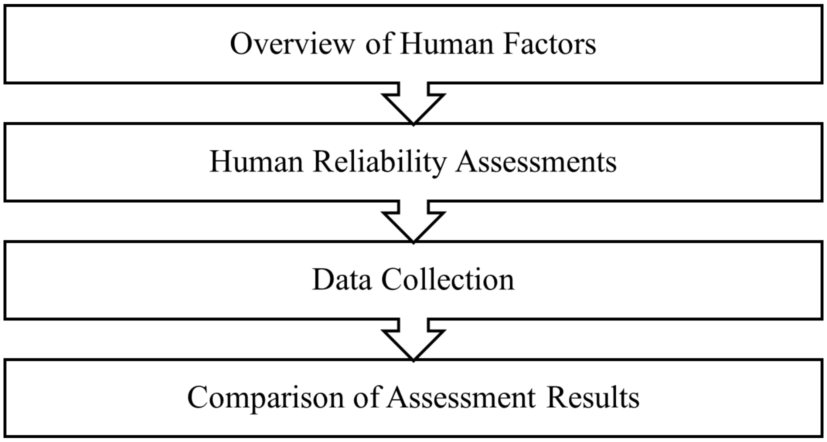

Overview of Human Factors
Human Reliability Assessments
Data Collection
Comparison of Assessment Results

Figure 7.1: Chapter Seven Structure

7.2 Overview of Human Factors

Reducing Error and Influencing Behaviour (HSG48) [154] is the key document to understand the approach of the UK Health and Safety Executive's (HSE) to human factors. HSG48 provides a simple introduction to generic industry guidance on human factors, which is defined as:

"Human factors refer to environmental, organisational and job factors, and human and individual characteristics, which influence behaviour at work in a way which can affect health and safety" [154, p. 5].

Human factors focus on what tasks individuals are being asked to perform (the task and its characteristics), who is performing them (the individual and their competence), and the environment in which they are working (the organisation and its attributes).

When assessing the effectiveness of a proof test, it is essential to account for the possibility that the tester could make an omission, an oversight or other error that may result in failing to detect a hardware fault. It is important to quantify the likelihood of this possibility for comparison with other factors discussed in this thesis that contribute to the ineffectiveness of the proof test. Human Reliability Assessments (HRA) provide a means of quantifying task-related human error as discussed in the next section.

7.2.1 Overview of Human Reliability Assessments

Two of the most widely used Human Reliability Assessment (HRA) techniques are the Standardised Plant Analysis Risk-Human Reliability Assessment (SPAR-H) method [155] and the Human Error Assessment and Reduction Technique (HEART) [156]. Consequently, these methods were selected as the focus for evaluating HEP in this study. The SPAR-H and HEART HRA models adopt a similar approach, which is to select a task type and then to apply modifying factors based on the stresses and environmental factors to which the proof tester is subjected.

7.2.1.1 SPAR-H

The SPAR-H method begins with identification of the generic task HEP which is based on whether the task is diagnostic or action-based. The method is based on eight Performance Shaping Factors (PSFs) that cover the majority of contributors to human error [157]. The PSFs include: Available Time, Stress/Stressors, Complexity, Experience/Training, Procedures, Ergonomics/Human Machine Interface (HMI), Fitness for Duty and Work Processes. Each PSF has various levels that reflect the degree of association with the task being assessed. The products of all the PSFs are then combined with the generic task HEP to calculate the HEP for the task. Figure 7.2 shows an overview of the SPAR-H process when evaluating a Human Failure Event (HFE).

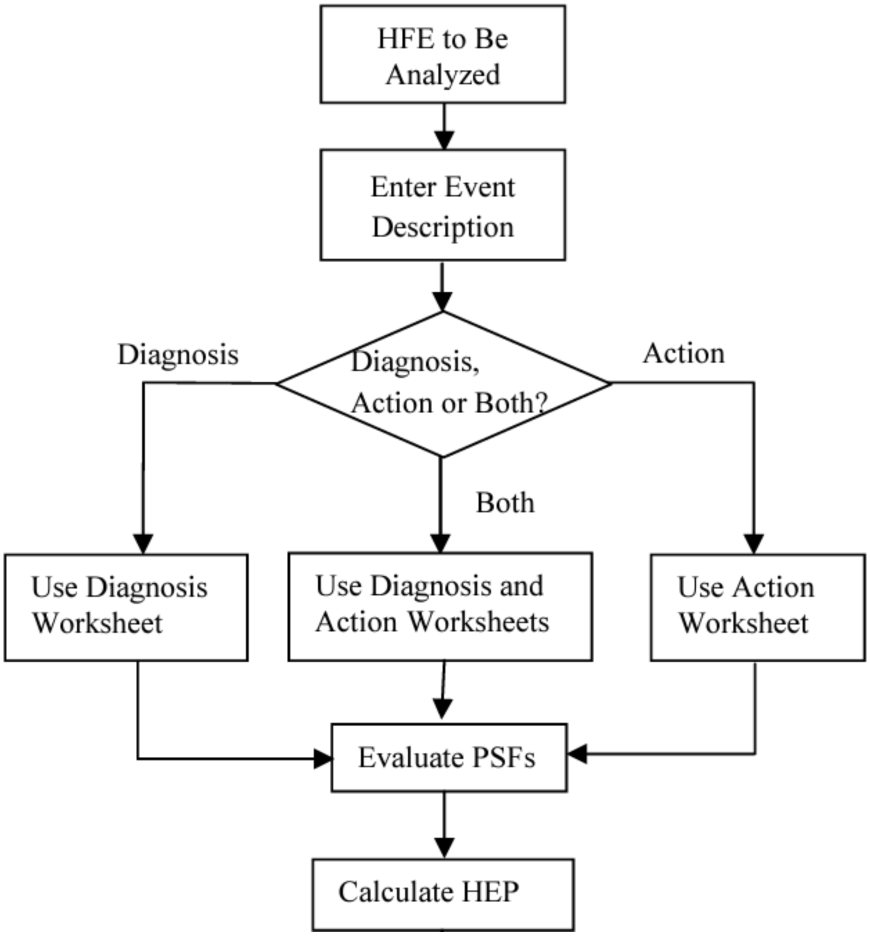


Figure 7.2: SPAR-H flowchart [158]

7.2.1.2 HEART

Similarly, the HEART method begins the quantification process by determining which of a number of generic HEPs is appropriate for the task under investigation [159]. The PSF effects on the task are then considered, though these are called Error Producing Conditions (EPCs) in HEART terminology. After selecting EPCs from a range of 38 options, each is weighted based on the maximum impact they can have on performance, the extent to which that maximum effect should be applied in the specific scenario is then determined. The product of the effects is then combined with the generic task HEP to calculate the HEP for the task. Figure 7.3 shows an overview of the HEART process with an example.

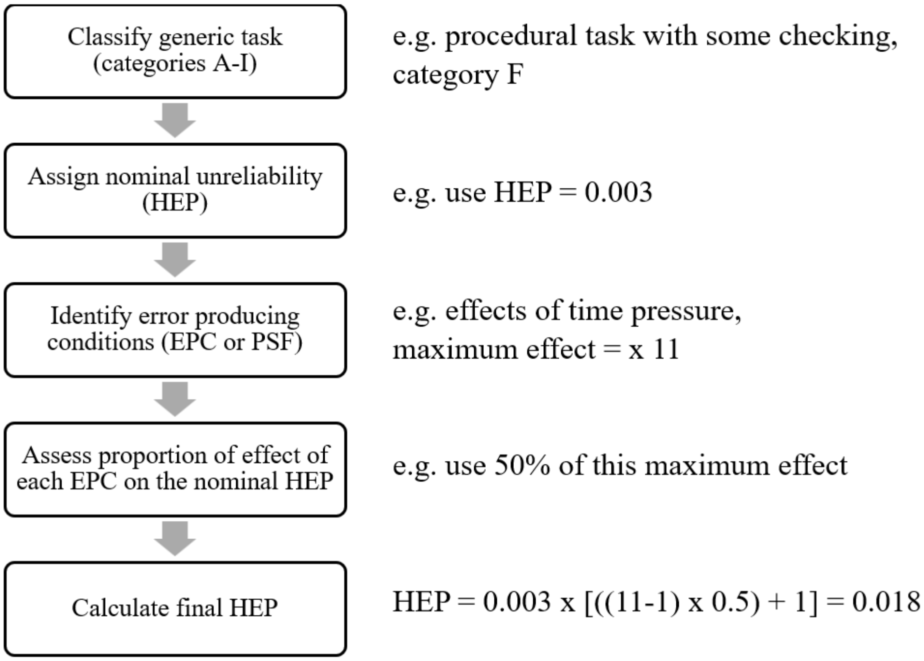

Classify generic task (categories A-I)
e.g. procedural task with some checking, category F
Assign nominal unreliability (HEP)
e.g. use HEP = 0.003
Identify error producing conditions (EPC or PSF)
e.g. effects of time pressure, maximum effect = x 11
Assess proportion of effect of each EPC on the nominal HEP
e.g. use 50% of this maximum effect
Calculate final HEP
HEP = 0.003 x [((11-1) x 0.5) + 1] = 0.018

Figure 7.3: HEART flowchart [159]

7.3 Human Reliability Assessment Data Collection

To evaluate the impact of human error on ESDV proof testing, HRA data were gathered from four human factors experts. The experts are UK Engineering Council registered (chartered engineers) with extensive experience of functional safety within the process industry sector. All experts have more than five years of experience in HRAs.
Each expert performed a SPAR-H and HEART assessment based the following *favourable* and *unfavourable* task definitions:

- *Undertake an ESDV proof test to a detailed procedure, under normal environmental conditions and with adequate competency.*
- *Undertake an ESDV proof test to a poorly written procedure, unfavourable environmental conditions and inadequate competency.*

In this chapter, it is pessimistically assumed that all human errors during preventive maintenance will lead to a failure to diagnose. A comparison of the HRA results is discussed in the next section.

7.4 Comparison of Human Reliability Assessment Results

Table 7.1 shows a summary of the HRA results (SPAR-H and HEART) carried out by experts, considering both favourable and unfavourable maintenance factors. Refer to Appendix E for HRA assessment worksheets for Expert ’A’.

Table 7.1: SPAR-H and HEART Assessments - Human Error Probabilities (HEPs)

HRA type / Assessor	**Favourable maintenance factors**	**Unfavourable maintenance factors**
SPAR-H / Expert A	5.03E-03	8.35E-01
SPAR-H / Expert B	1.20E-02	2.44E-01
SPAR-H / Expert C	3.99E-03	7.50E-01
SPAR-H / Expert D	1.00E-02	5.03E-01
HEART / Expert A	7.85E-02	2.52E-01
HEART / Expert B	1.42E-02	1.92E-01
HEART / Expert C	5.16E-02	2.58E-01
HEART / Expert D	9.75E-03	2.17E-01

Figure 7.4 shows a comparison of HRA results, indicating a significant difference in human error probabilities between favourable and unfavourable maintenance factors related to ESDV proof testing. A logarithmic scale is used to handle the wide range of HEP values.

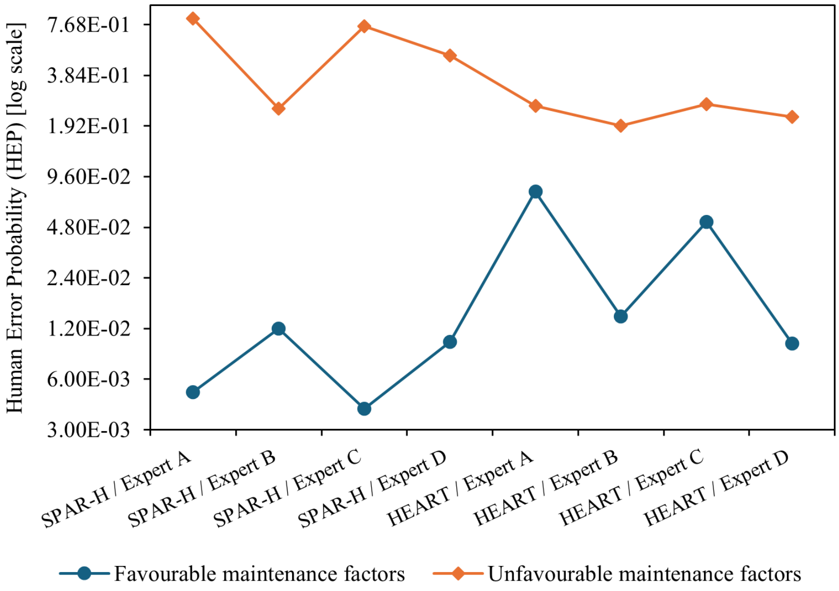


Figure 7.4: SPAR-H and HEART - Favourable and Unfavourable Maintenance Factors

The HRA results indicate that for a proof test with an adequate proof test procedure, appropriate environmental conditions, and a competent tester, an HEP of 1.38E-02 (geometric mean) with a range of 1.00E-3 to 7.85E-02 is achieved. The proof test with unfavourable factors indicates a HEP of 3.47E-01 (geometric mean) with a range of 1.92E-01 to 8.35E-01.

7.5 Discussion and Conclusions

This chapter examined the impact of human error on ESDV proof testing and reviewed two commonly used Human Reliability Assessments (HRAs) to determine Human Error Probabilities (HEPs) for the task. From the HRA results, it is evident that human error can negatively impact the quality of proof testing if the proof test procedure, environmental and training factors are inadequate. The results indicate that there is an approximately 35% chance of an error when considering unfavourable proof test factors as opposed to approximately 1% for favourable factors. Compared to the work by Aalipour et al [157], the results show a similar trend, with a significant difference in HEP between favourable and unfavourable maintenance tasks for both SPAR-H and HEART. This suggests that the reliability of the proof test process itself is a key factor preventing the system from reaching its maximum potential proof test coverage. Thus, attention to proof test procedures, environmental conditions, and competency of the tester is paramount.

When designing a safety instrumented function, consideration should be given to ensuring that proof testing can be performed effectively, with adequate proof test facilities available. The design should also account for the location of the installation, ensuring that access to allow maintenance and testing is sufficient.

A written proof test procedure is a key tool in conducting a proof test effectively and should describe every step that is to be performed by the proof tester, hence reducing the opportunity for error. The procedures should be developed by a qualified engineer familiar with the equipment and test methods. Procedures should be reviewed and validated in the field by a competent proof tester, typically an instrument technician, to ensure accuracy before they are issued for use.

The ideal proof tester is preferably a time-served apprentice with extensive experience in maintaining and testing safety instrumented systems, as well as familiarity with proof testing ESDVs. A strong understanding of the fundamentals of functional safety and the necessity and importance of proof testing is also highly desirable.

Planning activities such as proof testing is essential to ensure that the stress placed on the proof tester is at a minimum. Allow ample time to carry out the test and report the findings. Planning will also seek to gain co-operation from the operations team to provide access to the equipment, i.e., taking the equipment out of service, if required.

Incorporating these factors will significantly reduce the likelihood of human error during the execution of an ESDV proof test procedure.

7.5.1 Summary of Findings

The key findings of this chapter can be briefly summarised in the following points:

- Using HRA techniques such as SPAR-H and HEART, it is possible to quantitatively demonstrate that human error can significantly impact the outcome of a proof test.
- The HRA results of the experts demonstrate consensus in their evaluations of the proof tests, considering both favourable and unfavourable factors.
- The HRA results indicate a roughly one in three chance of error during the proof test, which limits the ability to achieve the maximum potential coverage for the proof test.
- Key areas for reducing human error include the proof test procedure, the competency of the proof tester, and environmental factors such as time constraints.

Chapter 8

Conclusions and Recommendations

An Emergency Shutdown Valve (ESDV) is a critical element of a safety instrumented function with limited online diagnostics; consequently, effective proof testing and test interval are imperative to ensure the desired functionality and hence the safe operation of a high-hazard process facility. There is much debate regarding the estimates of Proof Test Coverage (PTC), in particular for an emergency shutdown valve, with great variability evident in the published literature. The objective of this research was to address this problem by developing a robust assessment method that can be used to estimate proof test coverage.

Three methods were investigated to estimate the coverage of the proof test: using ESDV maintenance and test data from industry, performing a Failure Modes, Effects, and Diagnostics Analysis (FMEDA) based on an industry component failure rate database and additional sources, and combining FMEDA with a fuzzy inference system (Fuzzy-FMEDA). As human factors are essential when considering aspects of process safety, it was necessary to also investigate the impact of human error on proof testing. The focus of the research was on the application of the FMEDA technique, combined with an inference system to address the inherent uncertainty in the FMEDA data.
This chapter consolidates the key findings from this research, focussing on the Fuzzy Inference System (FIS) modelling of the uncertainty of the FMEDA data within safety critical systems. The chapter outlines the main conclusions drawn from the research objectives and discusses the broader implications of these findings. In addition, it provides recommendations for future research and practical applications to enhance understanding and management of proof test coverage.

8.1 Conclusions

Four primary investigations were conducted, which led to several conclusions. The key findings are summarised and discussed in detail in the following sections.

- With sufficient and correct maintenance and test data, it is possible to estimate proof test coverage by analysing the dangerous failure modes of ESDVs and reviewing how often those failure modes are detected during proof testing.

- The FMEDA approach to estimating PTC provided a systematic means to analysing failure modes of an ESDV but the subjectivity of the input data presented variance in results when estimating proof test coverage.
- To overcome the uncertainty in the FMEDA data, a fuzzy inference system was developed, which demonstrated that it is possible to design such a system to model the FMEDA process.
- The Fuzzy-FMEDA results indicated that the model could be integrated into an ESDV maintenance strategy to help optimise the proof test intervals effectively.
- Human error, assessed through standard human reliability evaluation methods, was found to significantly impact proof testing outcomes.

8.1.1 Estimating PTC using equipment failure data

Maintenance and test data were collected from industry and using BS EN 14224, dangerous failure modes were identified. Based on these findings, proof test coverage was estimated for a full stroke test. The analysis was constrained by the limited availability of data, as collecting this information from organisations proved challenging due to its commercial sensitivity, despite the presence of nondisclosure agreements (NDA). The collected data lacked sufficient detail, making it impossible to pinpoint the specific ESDV component that had failed and the reason for the failure. Instead, only the failure mode of the equipment could be identified. Analysis of three different datasets revealed that the proof test coverage for a full stroke test was estimated to range from 81% to 87%. Historically, a meaningful record of maintenance and testing findings in industry has generally been lacking, influenced by factors such as company policies and the limited knowledge and experience of staff. Ultimately, ESDV failures would undergo root cause analysis (RCA) to identify the cause of failure; however, these analyses can be a costly exercise for an operating company.

8.1.2 Estimating PTC using FMEDA and component failure rate data

A failure modes, effect and diagnostic analysis was identified as the preferred method of conducting an evaluation of proof test coverage for an ESDV. An FMEDA was performed using an industry component failure rate database, and proof test coverage for a full stroke test was estimated to be 82%, which was in good agreement with published estimates. Other data sources were also collected for comparison, which highlighted variability in component failure rate data resulting in a variety of different PTC estimates, ranging from 71% to 87% for a full stroke test. Due to the absence of a standardised component failure rate database, various organisations have developed their own, resulting in significant variance across datasets. Ideally, collaboration among key industry players would yield a shared data collection that could be widely accepted and utilised for estimating proof test coverage and making informed decisions regarding equipment failure.

8.1.3 Estimating PTC Using a Fuzzy Inference System

A fuzzy inference system was introduced as a method to manage uncertainty in the FMEDA data. The results of this research demonstrate that a fuzzy inference system can be effectively used to estimate Proof Test Coverage (PTC), while accounting for uncertainty in the input data. The analysis highlighted that trapezoidal membership functions were more effective for the FIS modelling within the FMEDA context, as indicated by high R^2 values (ranging from 0.9 to 0.99) and low RMSE (minimum of 2.5). This finding suggests that trapezoidal membership functions provide a robust balance between model accuracy and interpretability, contributing to enhanced reliability in PTC estimates. Additionally, the observed difference between FMEDA and FIS generated PTC values was 3% for full stroke tests and up to 22% for partial stroke tests. For partial stroke tests with a relatively low probability of revealing a failure, the variation in PTC was up to 6%.

The research results suggest that optimal outcomes would be achieved through a combination of full and partial proof tests. The partial stroke test proof test coverage results suggest the potential for extended full proof test intervals, thereby reducing the need to take an ESDV out of service for testing. The benefit of this approach is that proof test intervals can be optimised, leading to a more refined maintenance strategy. This, in turn, improves plant uptime and provides greater confidence in the reliability and operation of safety instrumented functions.

The FIS design was limited by the availability of data, as FMEDA expertise is difficult to obtain due to the specialised nature of this engineering field.

This research has explored an innovative Fuzzy-FMEDA framework that effectively models subjective expert data, which conventional FMEDA techniques struggle to capture. Although no prior studies have directly applied fuzzy logic to FMEDA, related work on FMEA and FMECA underscores the promise of this approach in addressing critical limitations of traditional FMEDA methods. The Fuzzy-FMEDA approach allows for more accurate proof test coverage results based on expert knowledge, allowing for optimised maintenance planning, ultimately reducing cost to operators and providing confidence in the reliability of safety instrumented systems. This novel contribution introduces a new method for incorporating expert judgement into safety-related assessments, establishing a foundation for future studies to validate and enhance the model further.

8.1.4 The Impact of Human Error on Proof Testing

From the results of the Human Reliability Assessments (HRA), SPAR-H, and HEART, which were performed by several experts, it was evident that human factors can significantly impact the result of proof tests. The HRA results indicate that for a proof test with an adequate proof test procedure, appropriate environmental conditions, and a competent tester, a Human Error Probability (HEP) of 1.38E-02 (geometric mean) is achieved in the range of 1.00E-3 to 7.85E-02 is achieved. The proof test with unfavourable factors indicates a HEP of 3.47E-01 (geometric mean) with a range of 1.92E-01 to 8.35E-01. The results of the human error investigation and its impact on proof testing highlighted the importance of ensuring that the test procedures, environmental factors, and personnel competency levels are adequate.

8.2 Further Work and Recommendations

This thesis has highlighted several areas where further investigation could be valuable, including both extensions of current work and new directions inspired by noteworthy findings.

The accuracy of the FIS model is based on the FMEDA data provided by experts. An additional effort to collect such data will enable model refinement and improved predication of proof test coverage. New data have the potential to extend the universe of discourse of the component failure rate variable and also affect the distribution of failure rates, suggesting a change in partitioning of fuzzy sets and therefore a change in IF-THEN rules. Special attention should be paid to the data collected on the probability of revealing a failure for partial stroke tests, as these data have been the most limited and have been shown to have the greatest influence on proof test coverage results.

The participation of industry experts is vital for the future success of any follow-up work in Fuzzy-FMEDA research. One of the main challenges in this field is the difficulty in collecting high-quality data due to its commercial sensitivity. Many organisations are reluctant to share detailed operational data for confidentiality reasons, which limits the availability of robust datasets. In such cases, reliance on experts becomes essential, as their implicit knowledge and practical experience can provide valuable insights that compensate for the lack of accessible quantitative field data. Beyond addressing data limitations, the involvement of industry experts also enhances the credibility and validity of the research. By incorporating the knowledge of industry experts, the assumptions and outcomes of the Fuzzy-FMEDA model are more likely to reflect actual industry practices, increasing its applicability. In addition, engaging with experts helps to build trust within the industry, facilitating broader acceptance and adoption of the research findings. Experts can also highlight emerging trends and potential challenges, helping to future proof the methodology and maintain its relevance. Therefore, collaborating with industry experts not only helps overcome immediate data challenges but also enhances the overall relevance and effectiveness of the research.

Given the safety-critical nature of Emergency Shutdown Valves (ESDV) and their role in protecting both personnel and the environment, ensuring the availability of reliable and consistent component failure rate data is essential for accurate and effective FMEDA. A lack of such data compromises the reliability of proof test coverage estimates, potentially leading to unsafe operational conditions. Intervention by the UK Health and Safety Executive (HSE) with a legislative effort to address the issue could help standardise the collection, sharing, and protection of sensitive failure rate data, encouraging greater transparency within the industry. Legislative action could include establishing guidelines or incentives for companies to share anonymised failure rate data while protecting commercially sensitive information. It could also encourage the development of industry-wide databases, where data can be collated and accessed by organisations conducting such assessments. Such an initiative would not only improve the quality of FMEDA, but also enhance overall safety standards. The involvement of the HSE could provide a regulatory framework that ensures that data are collected in a consistent, accurate, and secure manner, ultimately improving decision making about proof test coverage and reliability. Furthermore, it could encourage industry-wide collaboration, benefiting not just individual companies but the sector as a whole.

Other areas that could be investigated are alternate membership functions. So far triangular, trapezoidal, and Gaussian membership functions have been experimented with but there are other, not so common types, that could result in more efficient and effective fuzzy inference designs. These include the sigmoid and generalised bell membership functions [160]. The use of such membership functions could reduce the number of fuzzy sets used and hence the number of IF-THEN rules, allowing for an optimised fuzzy inference system.

Although the Mamdani Type-1 fuzzy inference system has been shown to be the most commonly used type for applications regarding failure modes and effects analysis, further experiments using Mamdani Type-2 fuzzy inference systems can be investigated. A Mamdani Type-2 fuzzy inference system extends the classical Mamdani Type-1 FIS by handling higher degrees of uncertainty in complex systems. Unlike Type-1 FIS, which uses fixed membership functions for each fuzzy set, Type-2 fuzzy inference systems employ fuzzy membership functions, where each membership degree is itself a fuzzy set rather than a single value. This results in a *fuzzy footprint* or Footprint of Uncertainty (FOU) around each membership function, allowing it to capture uncertainties more flexibly [161]. By capturing finer levels of uncertainty, Type-2 FIS can provide more precise outcomes in systems that involve expert judgement and vague or ambiguous input data. The MATLAB Fuzzy Logic Toolbox includes functionality for converting Mamdani Type-1 fuzzy inference systems to Type-2 systems, enabling efficient comparison of results.

This research has investigated the two most commonly used proof test types, i.e., the full stroke test and the partial stroke test. Several other proof tests are available as discussed in Chapter 2, which are variants of the full stroke test. The full stroke test investigated in this research being an offline test therefore, is not subjected to process conditions such as pressure, temperature, and flow. Further work should be undertaken to determine the proof test coverage for a full stroke test under process conditions, i.e. with the ESDV online. Similarly, further investigation should also be carried out to determine the proof test coverage when ESDV seat leakage tests are available. These additional investigations will enable a comprehensive range of proof test types and proof test coverages, providing plant operators with the full scope of ESDV testing.

Another promising area for future research is the potential integration of Artificial Intelligence (AI) with the Fuzzy-FMEDA approach. Several methods are available that could potentially optimise the fuzzy inference system. For example, *automated parameter optimization* which in Machine Learning (ML) algorithms could help fine-tune the parameters of the fuzzy inference system [162]. Traditional machine learning algorithms include Support Vector Machine (SVM), Logistic Regression (LR), Decision Tree (DT), and k-Nearest Neighbour (kNN) [163]. Deep Neural Networks (DNNs), which encompass Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) [162], can also be used. These methods can optimise the type of membership functions, their positions, and the number of fuzzy sets based on the FMEDA data to improve model accuracy and robustness.

The Adaptive Neuro Fuzzy Inference System (ANFIS) could enable the fuzzy inference system to adapt its behaviour over time based on new data [164]. Reinforcement learning, for example, could allow the fuzzy model to adjust its inference rules as new operational data and expert feedback is collected, making it more effective in dynamic environments.

Data-driven expert systems such as Natural Language Processing (NLP) could analyse expert reports, maintenance and test records, or equipment failure reports to extract relevant information for FMEDA [165]. An AI-enhanced expert system could then automatically feed this data into the Fuzzy-FMEDA model, ensuring it remains up-to-date and aligned with real-world observations.

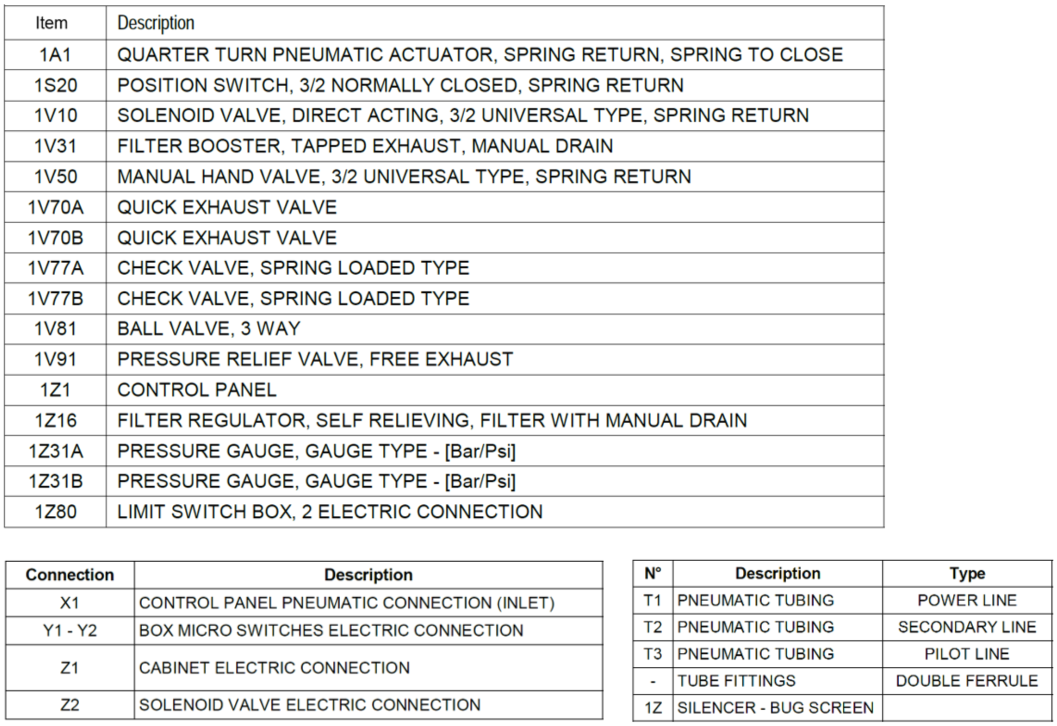
Combining fuzzy systems with other AI approaches, that is, hybrid AI models, such as Support Vector Machine (SVM), Bayesian Network (BN) and Artificial Neural Network (ANN), could help capture complex nonlinear relationships in safety-related assessments [166]. This could improve system performance by leveraging the strengths of both fuzzy logic and data-driven AI models.

Appendix A

ESDV Actuator Configuration

A.1 ESDV Actuator Pneumatic Schematic showing Partial Stroke Test Functionality

With reference to Figure A.1 and Figure A.2. When the actuator is energised (ESDV fully open) through SOV 1V10, a partial stroke test can be achieved by operating the 1V50 hand switch. The actuator will start to move (closing the valve) until the limit of travel is reached (cam and limit switch 1S20). Releasing the 1V50 hand switch will return the actuator to its pre-test position, i.e. fully open.



Item	Description
1A1	QUARTER TURN PNEUMATIC ACTUATOR, SPRING RETURN, SPRING TO CLOSE
1S20	POSITION SWITCH, 3/2 NORMALLY CLOSED, SPRING RETURN
1V10	SOLENOID VALVE, DIRECT ACTING, 3/2 UNIVERSAL TYPE, SPRING RETURN
1V31	FILTER BOOSTER, TAPPED EXHAUST, MANUAL DRAIN
1V50	MANUAL HAND VALVE, 3/2 UNIVERSAL TYPE, SPRING RETURN
1V70A	QUICK EXHAUST VALVE
1V70B	QUICK EXHAUST VALVE
1V77A	CHECK VALVE, SPRING LOADED TYPE
1V77B	CHECK VALVE, SPRING LOADED TYPE
1V81	BALL VALVE, 3 WAY
1V91	PRESSURE RELIEF VALVE, FREE EXHAUST
1Z1	CONTROL PANEL
1Z16	FILTER REGULATOR, SELF RELIEVING, FILTER WITH MANUAL DRAIN
1Z31A	PRESSURE GAUGE, GAUGE TYPE - [Bar/Psi]
1Z31B	PRESSURE GAUGE, GAUGE TYPE - [Bar/Psi]
1Z80	LIMIT SWITCH BOX, 2 ELECTRIC CONNECTION

Connection	Description
X1	CONTROL PANEL PNEUMATIC CONNECTION (INLET)
Y1 - Y2	BOX MICRO SWITCHES ELECTRIC CONNECTION
Z1	CABINET ELECTRIC CONNECTION
Z2	SOLENOID VALVE ELECTRIC CONNECTION

N°	Description	Type
T1	PNEUMATIC TUBING	POWER LINE
T2	PNEUMATIC TUBING	SECONDARY LINE
T3	PNEUMATIC TUBING	PILOT LINE
-	TUBE FITTINGS	DOUBLE FERRULE
1Z	SILENCER - BUG SCREEN	

Figure A.1: ESDV actuator pneumatic schematic key [44]

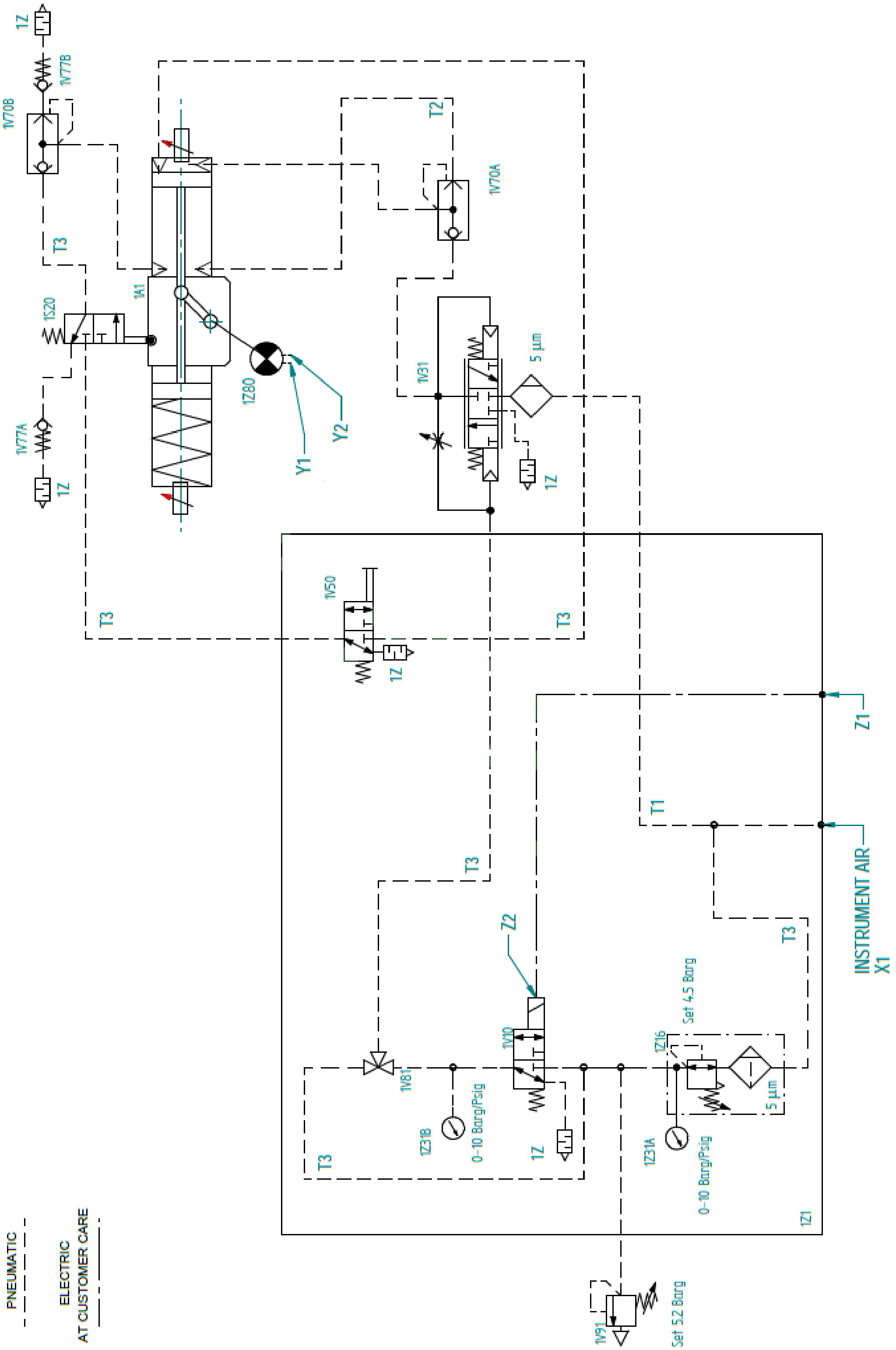
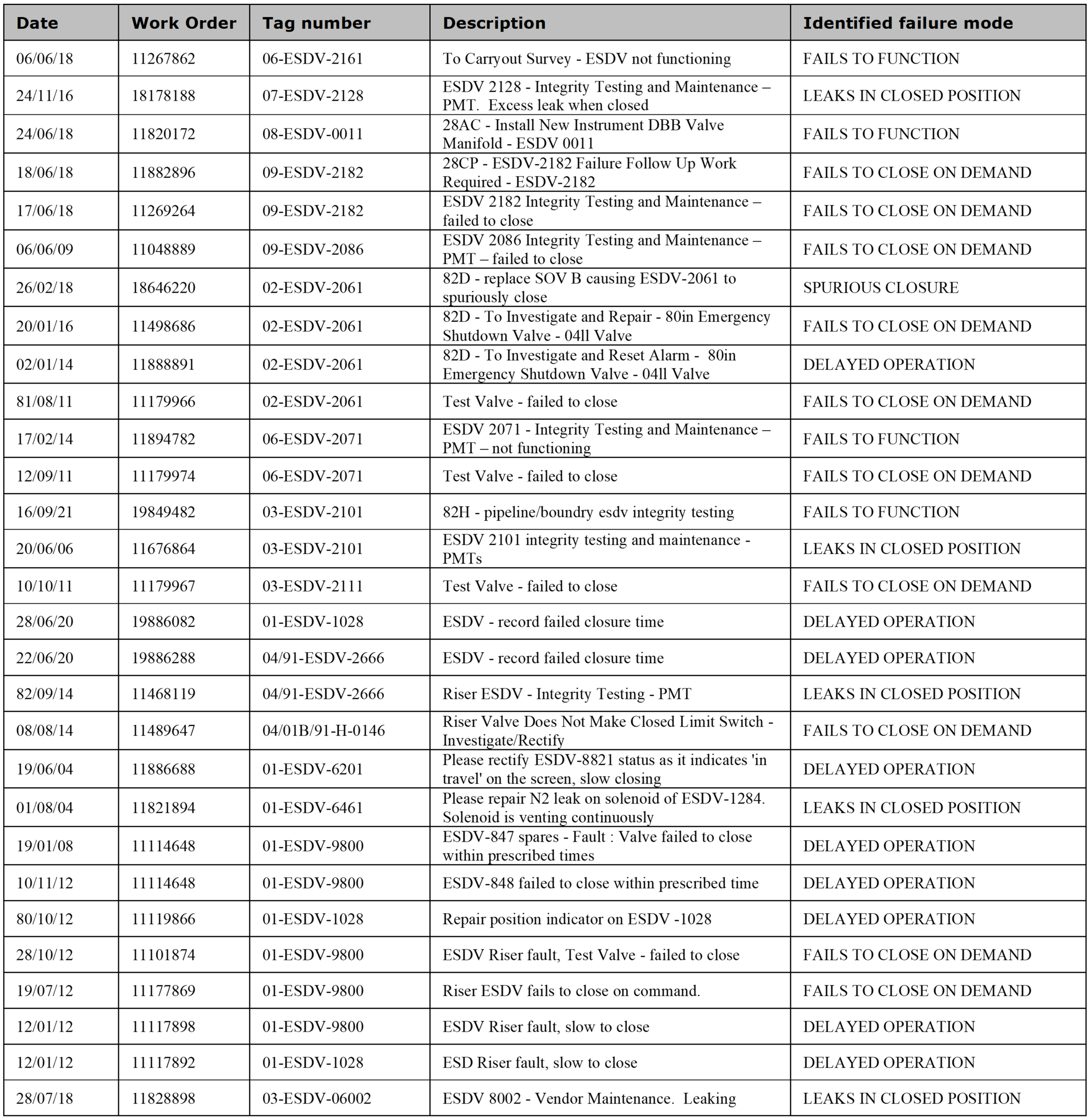


Figure A.2: ESDV actuator pneumatic schematic [44]

Appendix B

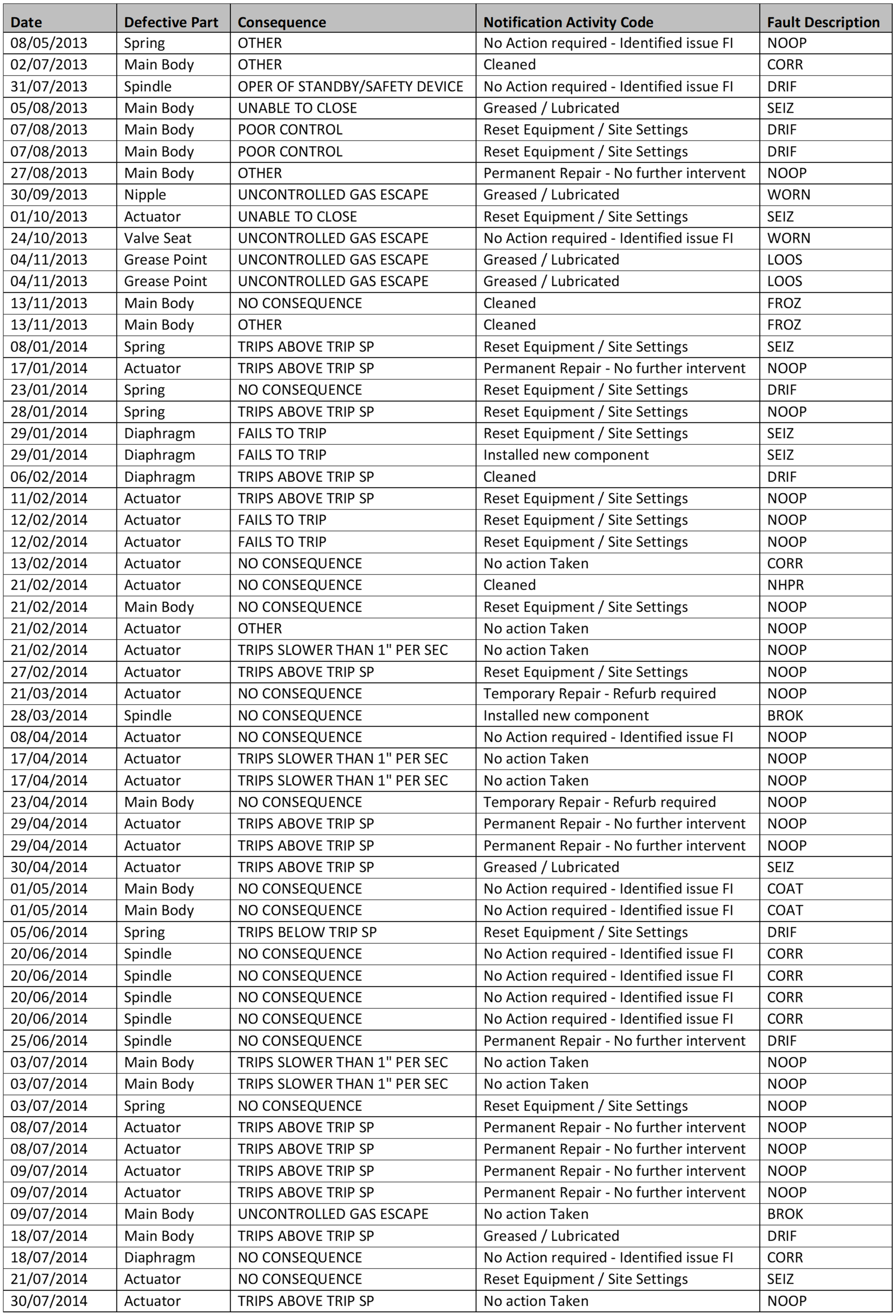
Equipment Maintenance and Test Data

B.1 Sample 1 - Data Collected from Industry (Anonymised)



Date	Work Order	Tag number	Description	Identified failure mode
06/06/18	11267862	06-ESDV-2161	To Carryout Survey - ESDV not functioning	FAILS TO FUNCTION
24/11/16	18178188	07-ESDV-2128	ESDV 2128 - Integrity Testing and Maintenance – PMT. Excess leak when closed	LEAKS IN CLOSED POSITION
24/06/18	11820172	08-ESDV-0011	28AC - Install New Instrument DBB Valve Manifold - ESDV 0011	FAILS TO FUNCTION
18/06/18	11882896	09-ESDV-2182	28CP - ESDV-2182 Failure Follow Up Work Required - ESDV-2182	FAILS TO CLOSE ON DEMAND
17/06/18	11269264	09-ESDV-2182	ESDV 2182 Integrity Testing and Maintenance – failed to close	FAILS TO CLOSE ON DEMAND
06/06/09	11048889	09-ESDV-2086	ESDV 2086 Integrity Testing and Maintenance – PMT – failed to close	FAILS TO CLOSE ON DEMAND
26/02/18	18646220	02-ESDV-2061	82D - replace SOV B causing ESDV-2061 to spuriously close	SPURIOUS CLOSURE
20/01/16	11498686	02-ESDV-2061	82D - To Investigate and Repair - 80in Emergency Shutdown Valve - 04ll Valve	FAILS TO CLOSE ON DEMAND
02/01/14	11888891	02-ESDV-2061	82D - To Investigate and Reset Alarm - 80in Emergency Shutdown Valve - 04ll Valve	DELAYED OPERATION
81/08/11	11179966	02-ESDV-2061	Test Valve - failed to close	FAILS TO CLOSE ON DEMAND
17/02/14	11894782	06-ESDV-2071	ESDV 2071 - Integrity Testing and Maintenance – PMT – not functioning	FAILS TO FUNCTION
12/09/11	11179974	06-ESDV-2071	Test Valve - failed to close	FAILS TO CLOSE ON DEMAND
16/09/21	19849482	03-ESDV-2101	82H - pipeline/boundry esdv integrity testing	FAILS TO FUNCTION
20/06/06	11676864	03-ESDV-2101	ESDV 2101 integrity testing and maintenance - PMTs	LEAKS IN CLOSED POSITION
10/10/11	11179967	03-ESDV-2111	Test Valve - failed to close	FAILS TO CLOSE ON DEMAND
28/06/20	19886082	01-ESDV-1028	ESDV - record failed closure time	DELAYED OPERATION
22/06/20	19886288	04/91-ESDV-2666	ESDV - record failed closure time	DELAYED OPERATION
82/09/14	11468119	04/91-ESDV-2666	Riser ESDV - Integrity Testing - PMT	LEAKS IN CLOSED POSITION
08/08/14	11489647	04/01B/91-H-0146	Riser Valve Does Not Make Closed Limit Switch - Investigate/Rectify	FAILS TO CLOSE ON DEMAND
19/06/04	11886688	01-ESDV-6201	Please rectify ESDV-8821 status as it indicates 'in travel' on the screen, slow closing	DELAYED OPERATION
01/08/04	11821894	01-ESDV-6461	Please repair N2 leak on solenoid of ESDV-1284. Solenoid is venting continuously	LEAKS IN CLOSED POSITION
19/01/08	11114648	01-ESDV-9800	ESDV-847 spares - Fault : Valve failed to close within prescribed times	DELAYED OPERATION
10/11/12	11114648	01-ESDV-9800	ESDV-848 failed to close within prescribed time	DELAYED OPERATION
80/10/12	11119866	01-ESDV-1028	Repair position indicator on ESDV -1028	DELAYED OPERATION
28/10/12	11101874	01-ESDV-9800	ESDV Riser fault, Test Valve - failed to close	FAILS TO CLOSE ON DEMAND
19/07/12	11177869	01-ESDV-9800	Riser ESDV fails to close on command.	FAILS TO CLOSE ON DEMAND
12/01/12	11117898	01-ESDV-9800	ESDV Riser fault, slow to close	DELAYED OPERATION
12/01/12	11117892	01-ESDV-1028	ESD Riser fault, slow to close	DELAYED OPERATION
28/07/18	11828898	03-ESDV-06002	ESDV 8002 - Vendor Maintenance. Leaking	LEAKS IN CLOSED POSITION

B.2 Sample 2 - Data Collected from Industry (Anonymised)

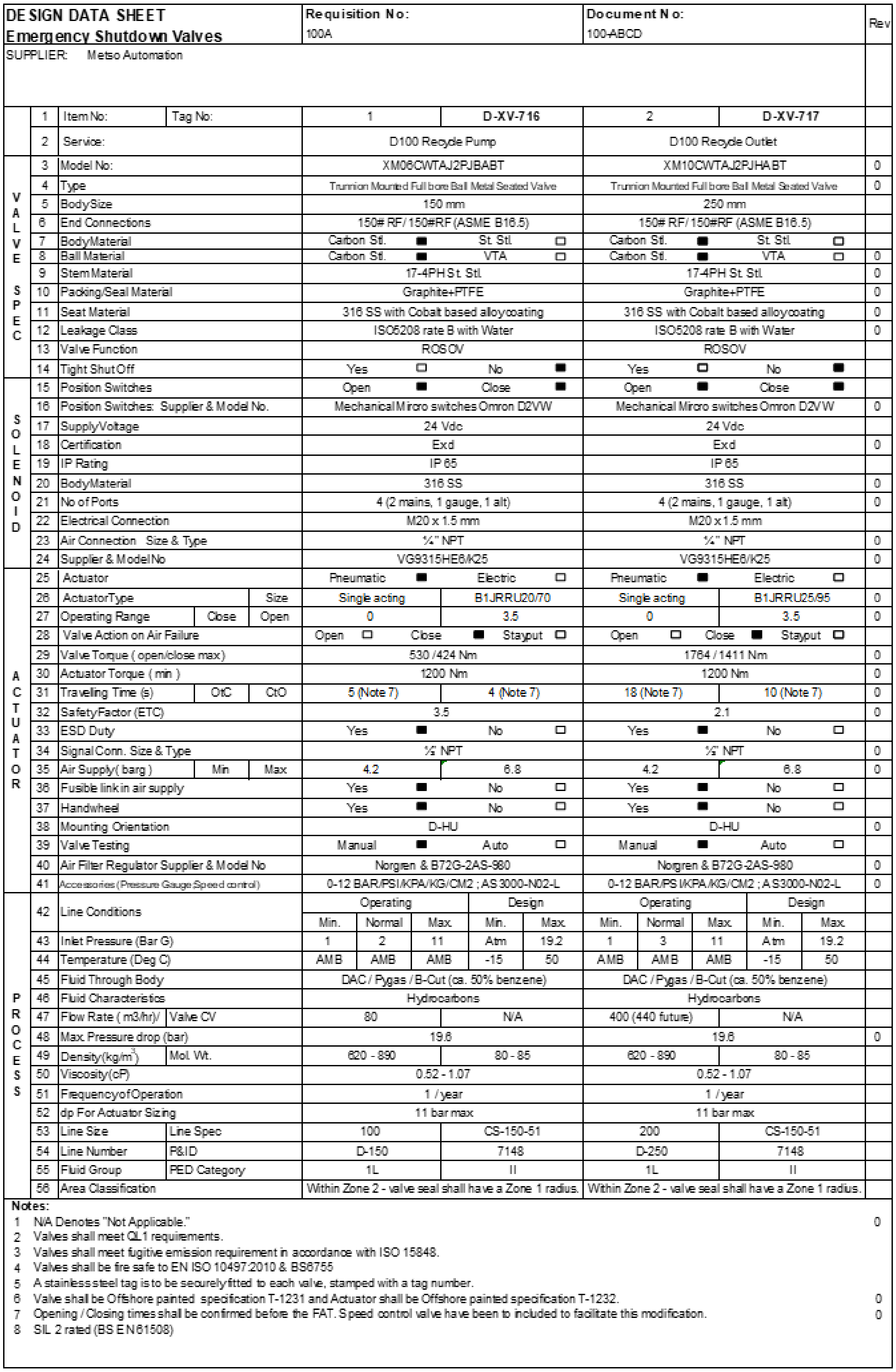


Date	Defective Part	Consequence	Notification Activity Code	Fault Description
08/05/2013	Spring	OTHER	No Action required - Identified issue FI	NOOP
02/07/2013	Main Body	OTHER	Cleaned	CORR
31/07/2013	Spindle	OPER OF STANDBY/SAFETY DEVICE	No Action required - Identified issue FI	DRIF
05/08/2013	Main Body	UNABLE TO CLOSE	Greased / Lubricated	SEIZ
07/08/2013	Main Body	POOR CONTROL	Reset Equipment / Site Settings	DRIF
07/08/2013	Main Body	POOR CONTROL	Reset Equipment / Site Settings	DRIF
27/08/2013	Main Body	OTHER	Permanent Repair - No further intervent	NOOP
30/09/2013	Nipple	UNCONTROLLED GAS ESCAPE	Greased / Lubricated	WORN
01/10/2013	Actuator	UNABLE TO CLOSE	Reset Equipment / Site Settings	SEIZ
24/10/2013	Valve Seat	UNCONTROLLED GAS ESCAPE	No Action required - Identified issue FI	WORN
04/11/2013	Grease Point	UNCONTROLLED GAS ESCAPE	Greased / Lubricated	LOOS
04/11/2013	Grease Point	UNCONTROLLED GAS ESCAPE	Greased / Lubricated	LOOS
13/11/2013	Main Body	NO CONSEQUENCE	Cleaned	FROZ
13/11/2013	Main Body	OTHER	Cleaned	FROZ
08/01/2014	Spring	TRIPS ABOVE TRIP SP	Reset Equipment / Site Settings	SEIZ
17/01/2014	Actuator	TRIPS ABOVE TRIP SP	Permanent Repair - No further intervent	NOOP
23/01/2014	Spring	NO CONSEQUENCE	Reset Equipment / Site Settings	DRIF
28/01/2014	Spring	TRIPS ABOVE TRIP SP	Reset Equipment / Site Settings	NOOP
29/01/2014	Diaphragm	FAILS TO TRIP	Reset Equipment / Site Settings	SEIZ
29/01/2014	Diaphragm	FAILS TO TRIP	Installed new component	SEIZ
06/02/2014	Diaphragm	TRIPS ABOVE TRIP SP	Cleaned	DRIF
11/02/2014	Actuator	TRIPS ABOVE TRIP SP	Reset Equipment / Site Settings	NOOP
12/02/2014	Actuator	FAILS TO TRIP	Reset Equipment / Site Settings	NOOP
12/02/2014	Actuator	FAILS TO TRIP	Reset Equipment / Site Settings	NOOP
13/02/2014	Actuator	NO CONSEQUENCE	No action Taken	CORR
21/02/2014	Actuator	NO CONSEQUENCE	Cleaned	NHPR
21/02/2014	Main Body	NO CONSEQUENCE	Reset Equipment / Site Settings	NOOP
21/02/2014	Actuator	OTHER	No action Taken	NOOP
21/02/2014	Actuator	TRIPS SLOWER THAN 1" PER SEC	No action Taken	NOOP
27/02/2014	Actuator	TRIPS ABOVE TRIP SP	Reset Equipment / Site Settings	NOOP
21/03/2014	Actuator	NO CONSEQUENCE	Temporary Repair - Refurb required	NOOP
28/03/2014	Spindle	NO CONSEQUENCE	Installed new component	BROK
08/04/2014	Actuator	NO CONSEQUENCE	No Action required - Identified issue FI	NOOP
17/04/2014	Actuator	TRIPS SLOWER THAN 1" PER SEC	No action Taken	NOOP
17/04/2014	Actuator	TRIPS SLOWER THAN 1" PER SEC	No action Taken	NOOP
23/04/2014	Main Body	NO CONSEQUENCE	Temporary Repair - Refurb required	NOOP
29/04/2014	Actuator	TRIPS ABOVE TRIP SP	Permanent Repair - No further intervent	NOOP
29/04/2014	Actuator	TRIPS ABOVE TRIP SP	Permanent Repair - No further intervent	NOOP
30/04/2014	Actuator	TRIPS ABOVE TRIP SP	Greased / Lubricated	SEIZ
01/05/2014	Main Body	NO CONSEQUENCE	No Action required - Identified issue FI	COAT
01/05/2014	Main Body	NO CONSEQUENCE	No Action required - Identified issue FI	COAT
05/06/2014	Spring	TRIPS BELOW TRIP SP	Reset Equipment / Site Settings	DRIF
20/06/2014	Spindle	NO CONSEQUENCE	No Action required - Identified issue FI	CORR
20/06/2014	Spindle	NO CONSEQUENCE	No Action required - Identified issue FI	CORR
20/06/2014	Spindle	NO CONSEQUENCE	No Action required - Identified issue FI	CORR
20/06/2014	Spindle	NO CONSEQUENCE	No Action required - Identified issue FI	CORR
25/06/2014	Spindle	NO CONSEQUENCE	Permanent Repair - No further intervent	DRIF
03/07/2014	Main Body	TRIPS SLOWER THAN 1" PER SEC	No action Taken	NOOP
03/07/2014	Main Body	TRIPS SLOWER THAN 1" PER SEC	No action Taken	NOOP
03/07/2014	Spring	NO CONSEQUENCE	Reset Equipment / Site Settings	NOOP
08/07/2014	Actuator	TRIPS ABOVE TRIP SP	Permanent Repair - No further intervent	NOOP
08/07/2014	Actuator	TRIPS ABOVE TRIP SP	Permanent Repair - No further intervent	NOOP
09/07/2014	Actuator	TRIPS ABOVE TRIP SP	Permanent Repair - No further intervent	NOOP
09/07/2014	Actuator	TRIPS ABOVE TRIP SP	Permanent Repair - No further intervent	NOOP
09/07/2014	Main Body	UNCONTROLLED GAS ESCAPE	No action Taken	BROK
18/07/2014	Main Body	TRIPS ABOVE TRIP SP	Greased / Lubricated	DRIF
18/07/2014	Diaphragm	NO CONSEQUENCE	No Action required - Identified issue FI	CORR
21/07/2014	Actuator	NO CONSEQUENCE	Reset Equipment / Site Settings	SEIZ
30/07/2014	Actuator	TRIPS ABOVE TRIP SP	No action Taken	NOOP

Appendix C

FMEDA Data

C.1 Example ESDV Datasheet



DESIGN DATA SHEET
Emergency Shutdown Valves

Requisition No: 100A

Document No: 100-ABCD

Rev

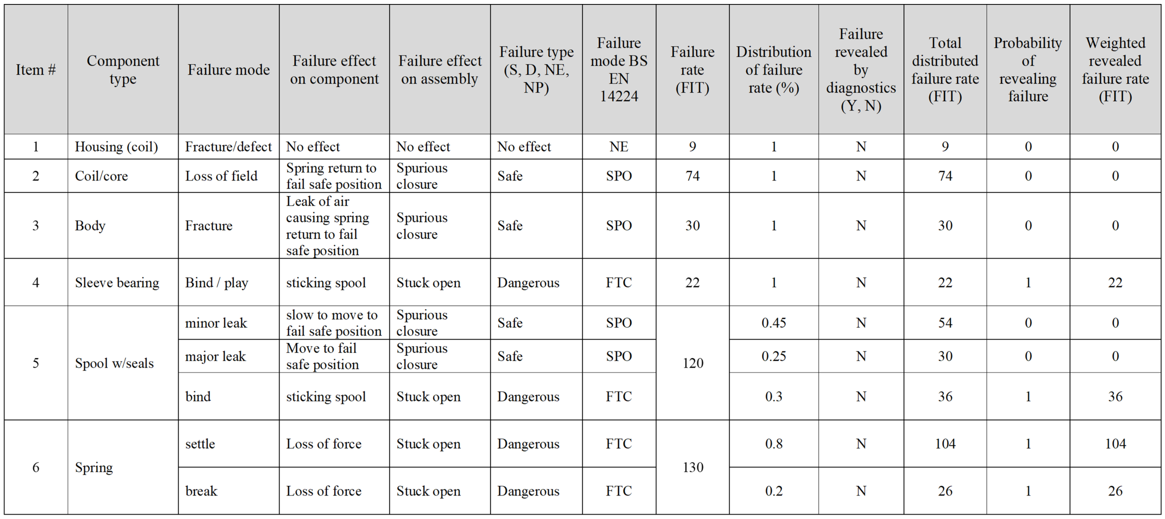
SUPPLIER: Metso Automation

	1	Item No:	Tag No:		1	**D-XV-716**				2	**D-XV-717**				
	2	Service:			D100 Recycle Pump					D100 Recycle Outlet					
VALVE SPEC	3	Model No:			XM06CWTAJ2PJBABT					XM10CWTAJ2PJHABT					0
	4	Type			Trunnion Mounted Full bore Ball Metal Seated Valve					Trunnion Mounted Full bore Ball Metal Seated Valve					0
	5	BodySize			150 mm					250 mm					
	6	End Connections			150# RF/ 150#RF (ASME B16.5)					150# RF/ 150#RF (ASME B16.5)					
	7	BodyMaterial			Carbon Stl. ■	St. Stl ☐				Carbon Stl. ■	St. Stl ☐				
	8	Ball Material			Carbon Stl. ■	VTA ☐				Carbon Stl. ■	VTA ☐				0
	9	Stem Material			17-4PH St. Stl.					17-4PH St. Stl.					0
	10	Packing/Seal Material			Graphite+PTFE					Graphite+PTFE					0
	11	Seat Material			316 SS with Cobalt based alloycoating					316 SS with Cobalt based alloycoating					0
	12	Leakage Class			ISO5208 rate B with Water					ISO5208 rate B with Water					0
	13	Valve Function			ROSOV					ROSOV					
	14	Tight Shut Off			Yes ☐	No ■				Yes ☐	No ■				
SOLENOID	15	Position Switches			Open ■	Close ■				Open ■	Close ■				
	16	Position Switches: Supplier & Model No.			Mechanical Mircro switches Omron D2VW					Mechanical Mirco switches Omron D2VW					0
	17	SupplyVoltage			24 Vdc					24 Vdc					
	18	Certification			Exd					Exd					0
	19	IP Rating			IP 65					IP 65					
	20	BodyMaterial			316 SS					316 SS					0
	21	No of Ports			4 (2 mains, 1 gauge, 1 alt)					4 (2 mains, 1 gauge, 1 alt)					0
	22	Electrical Connection			M20 x 1.5 mm					M20 x 1.5 mm					
	23	Air Connection Size & Type			¼" NPT					¼" NPT					0
	24	Supplier & Model No			VG9315HE6/K25					VG9315HE6/K25					0
ACTUATOR	25	Actuator			Pneumatic ■	Electric ☐				Pneumatic ■	Electric ☐				
	26	ActuatorType		Size	Single acting	B1JRRU20/70				Single acting	B1JRRU25/95				0
	27	Operating Range	Close	Open	0	3.5				0	3.5				0
	28	Valve Action on Air Failure			Open ☐	Close ■	Stayput ☐			Open ☐	Close ■	Stayput ☐			
	29	Valve Torque (open/close max)			530 /424 Nm					1764 / 1411 Nm					0
	30	Actuator Torque (min)			1200 Nm					1200 Nm					0
	31	Traveling Time (s)	OtC	CtO	5 (Note 7)	4 (Note 7)				18 (Note 7)	10 (Note 7)				0
	32	SafetyFactor (ETC)			3.5					2.1					0
	33	ESD Duty			Yes ■	No ☐				Yes ■	No ☐				
	34	Signal Conn. Size & Type			½" NPT					½" NPT					0
	35	Air Supply(barg)	Min	Max	4.2	6.8				4.2	6.8				0
	36	Fusible link in air supply			Yes ■	No ☐				Yes ■	No ☐				
	37	Handwheel			Yes ■	No ☐				Yes ■	No ☐				
	38	Mounting Orientation			D-HU					D-HU					0
	39	Valve Testing			Manual ■	Auto ☐				Manual ■	Auto ☐				
	40	Air Filter Regulator Supplier & Model No			Norgren & B72G-2AS-980					Norgren & B72G-2AS-980					0
	41	Accessories (Pressure Gauge Speed control)			0-12 BAR/PSI/KPA/KG/CM2 ; AS 3000-N02-L					0-12 BAR/PSI/KPA/KG/CM2 ; AS3000-N02-L					0
PROCESS	42	Line Conditions			Operating			Design		Operating			Design		
					Min.	Normal	Max.	Min.	Max.	Min.	Normal	Max.	Min.	Max.	
	43	Inlet Pressure (Bar G)			1	2	11	Atm	19.2	1	3	11	Atm	19.2	
	44	Temperature (Deg C)			AMB	AMB	AMB	-15	50	AMB	AMB	AMB	-15	50	
	45	Fluid Through Body			DAC / Pygas / B-Cut (ca. 50% benzene)					DAC / Pygas / B-Cut (ca. 50% benzene)					
	46	Fluid Characteristics			Hydrocarbons					Hydrocarbons					
	47	Flow Rate (m3/hr)/	Valve CV		80	N/A				400 (440 future)	N/A				
	48	Max. Pressure drop (bar)			19.6					19.6					0
	49	Density(kg/m^3)	Mol. Wt.		620 - 890	80 - 85				620 - 890	80 - 85				
	50	Viscosity(cP)			0.52 - 1.07					0.52 - 1.07					
	51	FrequencyofOperation			1 / year					1 / year					
	52	dp For Actuator Sizing			11 bar max					11 bar max					
	53	Line Size	Line Spec		100	CS-150-51				200	CS-150-51				
	54	Line Number	P&ID		D-150	7148				D-250	7148				
	55	Fluid Group	PED Category		1L	II				1L	II				
	56	Area Classification			Within Zone 2 - valve seal shall have a Zone 1 radius.					Within Zone 2 - valve seal shall have a Zone 1 radius.					

Notes:

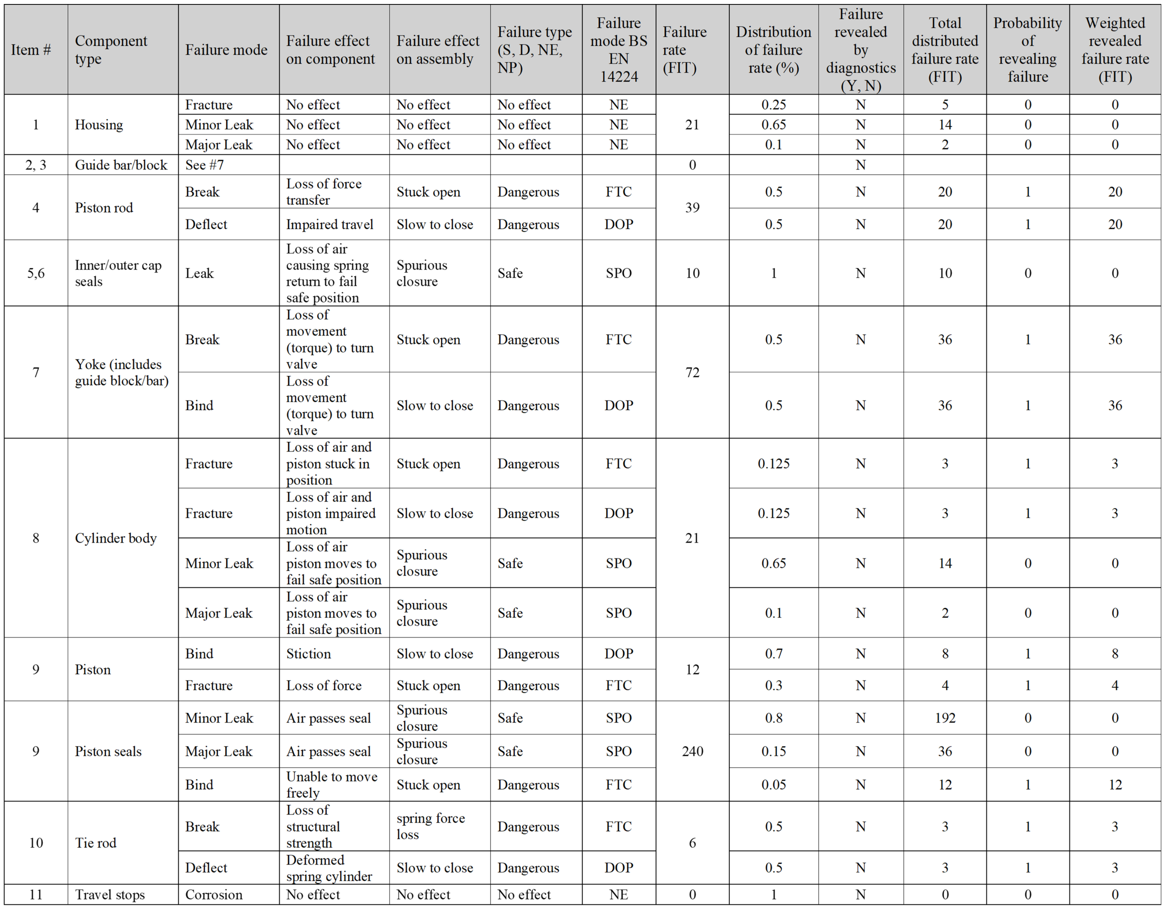
1 N/A Denotes "Not Applicable." 0
2 Valves shall meet QL1 requirements.
3 Valves shall meet fugitive emission requirement in accordance with ISO 15848.
4 Valves shall be fire safe to EN ISO 10497:2010 & BS6755
5 A stainless steel tag is to be securely fitted to each valve, stamped with a tag number.
6 Valve shall be Offshore painted specification T-1231 and Actuator shall be Offshore painted specification T-1232. 0
7 Opening / Closing times shall be confirmed before the FAT. Speed control valve have been to included to facilitate this modification. 0
8 SIL 2 rated (BS EN61508)

C.2 FMEDA Worksheets



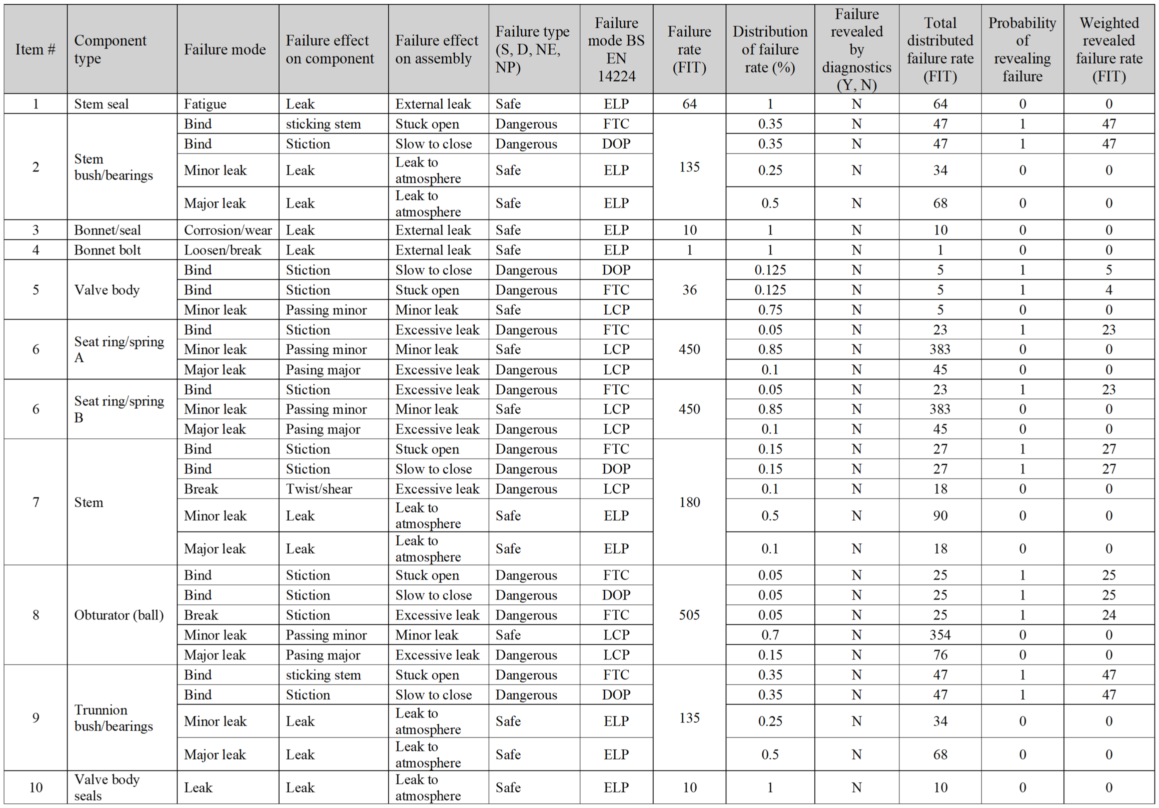
Item #	Component type	Failure mode	Failure effect on component	Failure effect on assembly	Failure type (S, D, NE, NP)	Failure mode BS EN 14224	Failure rate (FIT)	Distribution of failure rate (%)	Failure revealed by diagnostics (Y, N)	Total distributed failure rate (FIT)	Probability of revealing failure	Weighted revealed failure rate (FIT)
1	Housing (coil)	Fracture/defect	No effect	No effect	No effect	NE	9	1	N	9	0	0
2	Coil/core	Loss of field	Spring return to fail safe position	Spurious closure	Safe	SPO	74	1	N	74	0	0
3	Body	Fracture	Leak of air causing spring return to fail safe position	Spurious closure	Safe	SPO	30	1	N	30	0	0
4	Sleeve bearing	Bind / play	sticking spool	Stuck open	Dangerous	FTC	22	1	N	22	1	22
5	Spool w/seals	minor leak	slow to move to fail safe position	Spurious closure	Safe	SPO	120	0.45	N	54	0	0
		major leak	Move to fail safe position	Spurious closure	Safe	SPO		0.25	N	30	0	0
		bind	sticking spool	Stuck open	Dangerous	FTC		0.3	N	36	1	36
6	Spring	settle	Loss of force	Stuck open	Dangerous	FTC	130	0.8	N	104	1	104
		break	Loss of force	Stuck open	Dangerous	FTC		0.2	N	26	1	26

Figure C.1: Solenoid Operated Valve FMEDA Worksheet



Item #	Component type	Failure mode	Failure effect on component	Failure effect on assembly	Failure type (S, D, NE, NP)	Failure mode BS EN 14224	Failure rate (FIT)	Distribution of failure rate (%)	Failure revealed by diagnostics (Y, N)	Total distributed failure rate (FIT)	Probability of revealing failure	Weighted revealed failure rate (FIT)
1	Housing	Fracture	No effect	No effect	No effect	NE	21	0.25	N	5	0	0
		Minor Leak	No effect	No effect	No effect	NE		0.65	N	14	0	0
		Major Leak	No effect	No effect	No effect	NE		0.1	N	2	0	0
2, 3	Guide bar/block	See #7					0		N			
4	Piston rod	Break	Loss of force transfer	Stuck open	Dangerous	FTC	39	0.5	N	20	1	20
		Deflect	Impaired travel	Slow to close	Dangerous	DOP		0.5	N	20	1	20
5,6	Inner/outer cap seals	Leak	Loss of air causing spring return to fail safe position	Spurious closure	Safe	SPO	10	1	N	10	0	0
7	Yoke (includes guide block/bar)	Break	Loss of movement (torque) to turn valve	Stuck open	Dangerous	FTC	72	0.5	N	36	1	36
		Bind	Loss of movement (torque) to turn valve	Slow to close	Dangerous	DOP		0.5	N	36	1	36
8	Cylinder body	Fracture	Loss of air and piston stuck in position	Stuck open	Dangerous	FTC	21	0.125	N	3	1	3
		Fracture	Loss of air and piston impaired motion	Slow to close	Dangerous	DOP		0.125	N	3	1	3
		Minor Leak	Loss of air piston moves to fail safe position	Spurious closure	Safe	SPO		0.65	N	14	0	0
		Major Leak	Loss of air piston moves to fail safe position	Spurious closure	Safe	SPO		0.1	N	2	0	0
9	Piston	Bind	Stiction	Slow to close	Dangerous	DOP	12	0.7	N	8	1	8
		Fracture	Loss of force	Stuck open	Dangerous	FTC		0.3	N	4	1	4
9	Piston seals	Minor Leak	Air passes seal	Spurious closure	Safe	SPO	240	0.8	N	192	0	0
		Major Leak	Air passes seal	Spurious closure	Safe	SPO		0.15	N	36	0	0
		Bind	Unable to move freely	Stuck open	Dangerous	FTC		0.05	N	12	1	12
10	Tie rod	Break	Loss of structural strength	spring force loss	Dangerous	FTC	6	0.5	N	3	1	3
		Deflect	Deformed spring cylinder	Slow to close	Dangerous	DOP		0.5	N	3	1	3
11	Travel stops	Corrosion	No effect	No effect	No effect	NE	0	1	N	0	0	0

Figure C.2: Actuator FMEDA Worksheet



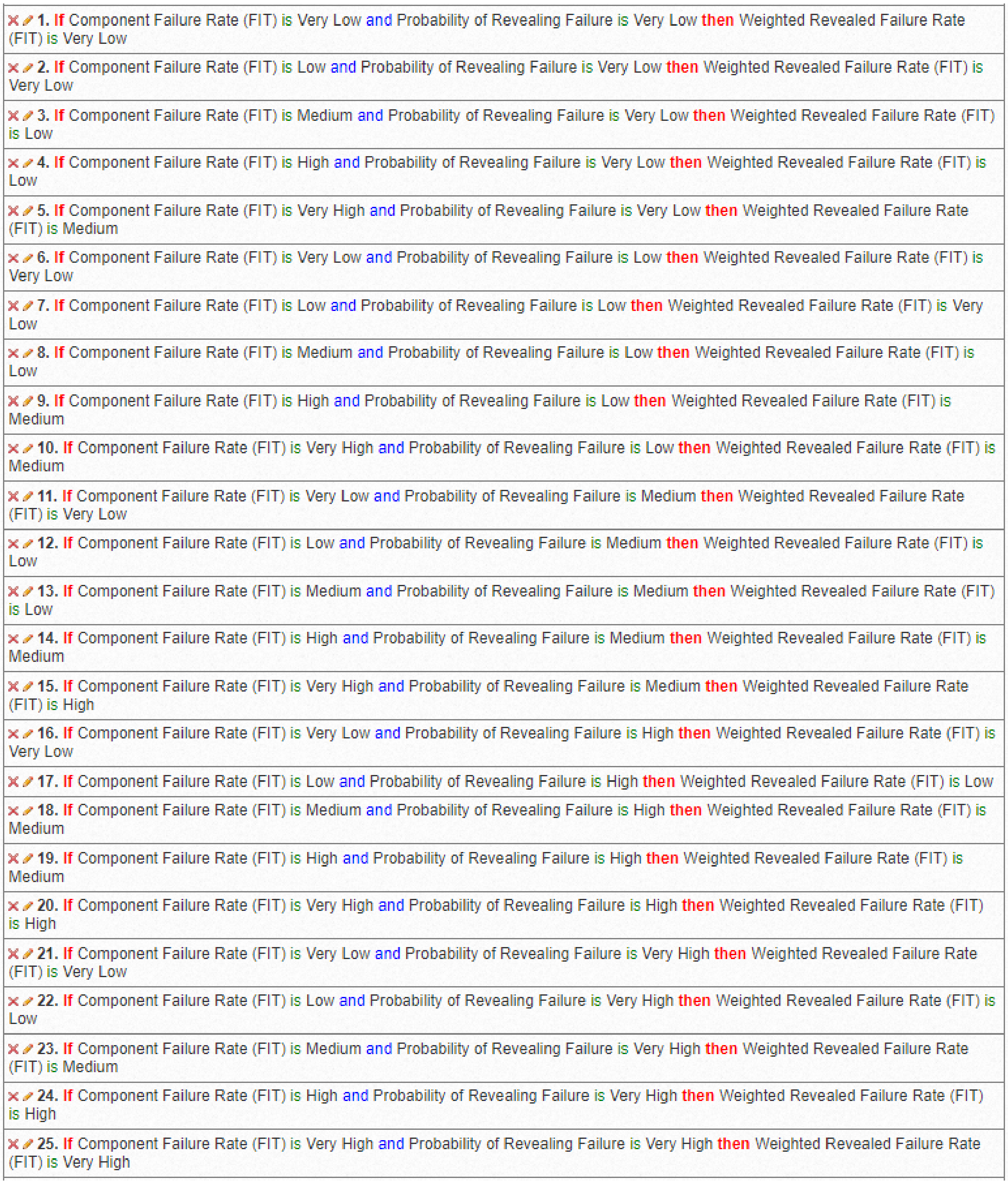
Item #	Component type	Failure mode	Failure effect on component	Failure effect on assembly	Failure type (S, D, NE, NP)	Failure mode BS EN 14224	Failure rate (FIT)	Distribution of failure rate (%)	Failure revealed by diagnostics (Y, N)	Total distributed failure rate (FIT)	Probability of revealing failure	Weighted revealed failure rate (FIT)
1	Stem seal	Fatigue	Leak	External leak	Safe	ELP	64	1	N	64	0	0
2	Stem bush/bearings	Bind	sticking stem	Stuck open	Dangerous	FTC	135	0.35	N	47	1	47
		Bind	Stiction	Slow to close	Dangerous	DOP		0.35	N	47	1	47
		Minor leak	Leak	Leak to atmosphere	Safe	ELP		0.25	N	34	0	0
		Major leak	Leak	Leak to atmosphere	Safe	ELP		0.5	N	68	0	0
3	Bonnet/seal	Corrosion/wear	Leak	External leak	Safe	ELP	10	1	N	10	0	0
4	Bonnet bolt	Loosen/break	Leak	External leak	Safe	ELP	1	1	N	1	0	0
5	Valve body	Bind	Stiction	Slow to close	Dangerous	DOP	36	0.125	N	5	1	5
		Bind	Stiction	Stuck open	Dangerous	FTC		0.125	N	5	1	4
		Minor leak	Passing minor	Minor leak	Safe	LCP		0.75	N	5	0	0
6	Seat ring/spring A	Bind	Stiction	Excessive leak	Dangerous	FTC	450	0.05	N	23	1	23
		Minor leak	Passing minor	Minor leak	Safe	LCP		0.85	N	383	0	0
		Major leak	Pasing major	Excessive leak	Dangerous	LCP		0.1	N	45	0	0
6	Seat ring/spring B	Bind	Stiction	Excessive leak	Dangerous	FTC	450	0.05	N	23	1	23
		Minor leak	Passing minor	Minor leak	Safe	LCP		0.85	N	383	0	0
		Major leak	Pasing major	Excessive leak	Dangerous	LCP		0.1	N	45	0	0
7	Stem	Bind	Stiction	Stuck open	Dangerous	FTC	180	0.15	N	27	1	27
		Bind	Stiction	Slow to close	Dangerous	DOP		0.15	N	27	1	27
		Break	Twist/shear	Excessive leak	Dangerous	LCP		0.1	N	18	0	0
		Minor leak	Leak	Leak to atmosphere	Safe	ELP		0.5	N	90	0	0
		Major leak	Leak	Leak to atmosphere	Safe	ELP		0.1	N	18	0	0
8	Obturator (ball)	Bind	Stiction	Stuck open	Dangerous	FTC	505	0.05	N	25	1	25
		Bind	Stiction	Slow to close	Dangerous	DOP		0.05	N	25	1	25
		Break	Stiction	Excessive leak	Dangerous	FTC		0.05	N	25	1	24
		Minor leak	Passing minor	Minor leak	Safe	LCP		0.7	N	354	0	0
		Major leak	Pasing major	Excessive leak	Dangerous	LCP		0.15	N	76	0	0
9	Trunnion bush/bearings	Bind	sticking stem	Stuck open	Dangerous	FTC	135	0.35	N	47	1	47
		Bind	Stiction	Slow to close	Dangerous	DOP		0.35	N	47	1	47
		Minor leak	Leak	Leak to atmosphere	Safe	ELP		0.25	N	34	0	0
		Major leak	Leak	Leak to atmosphere	Safe	ELP		0.5	N	68	0	0
10	Valve body seals	Leak	Leak	Leak to atmosphere	Safe	ELP	10	1	N	10	0	0

Figure C.3: Ball valve FMEDA Worksheet

Appendix D

Fuzzy Inference System Data

D.1 FIS Triangular 5x5 IF-THEN Rule Base



✕ ✎ **1.** If Component Failure Rate (FIT) is Very Low and Probability of Revealing Failure is Very Low then Weighted Revealed Failure Rate (FIT) is Very Low

✕ ✎ **2.** If Component Failure Rate (FIT) is Low and Probability of Revealing Failure is Very Low then Weighted Revealed Failure Rate (FIT) is Very Low

✕ ✎ **3.** If Component Failure Rate (FIT) is Medium and Probability of Revealing Failure is Very Low then Weighted Revealed Failure Rate (FIT) is Low

✕ ✎ **4.** If Component Failure Rate (FIT) is High and Probability of Revealing Failure is Very Low then Weighted Revealed Failure Rate (FIT) is Low

✕ ✎ **5.** If Component Failure Rate (FIT) is Very High and Probability of Revealing Failure is Very Low then Weighted Revealed Failure Rate (FIT) is Medium

✕ ✎ **6.** If Component Failure Rate (FIT) is Very Low and Probability of Revealing Failure is Low then Weighted Revealed Failure Rate (FIT) is Very Low

✕ ✎ **7.** If Component Failure Rate (FIT) is Low and Probability of Revealing Failure is Low then Weighted Revealed Failure Rate (FIT) is Very Low

✕ ✎ **8.** If Component Failure Rate (FIT) is Medium and Probability of Revealing Failure is Low then Weighted Revealed Failure Rate (FIT) is Low

✕ ✎ **9.** If Component Failure Rate (FIT) is High and Probability of Revealing Failure is Low then Weighted Revealed Failure Rate (FIT) is Medium

✕ ✎ **10.** If Component Failure Rate (FIT) is Very High and Probability of Revealing Failure is Low then Weighted Revealed Failure Rate (FIT) is Medium

✕ ✎ **11.** If Component Failure Rate (FIT) is Very Low and Probability of Revealing Failure is Medium then Weighted Revealed Failure Rate (FIT) is Very Low

✕ ✎ **12.** If Component Failure Rate (FIT) is Low and Probability of Revealing Failure is Medium then Weighted Revealed Failure Rate (FIT) is Low

✕ ✎ **13.** If Component Failure Rate (FIT) is Medium and Probability of Revealing Failure is Medium then Weighted Revealed Failure Rate (FIT) is Low

✕ ✎ **14.** If Component Failure Rate (FIT) is High and Probability of Revealing Failure is Medium then Weighted Revealed Failure Rate (FIT) is Medium

✕ ✎ **15.** If Component Failure Rate (FIT) is Very High and Probability of Revealing Failure is Medium then Weighted Revealed Failure Rate (FIT) is High

✕ ✎ **16.** If Component Failure Rate (FIT) is Very Low and Probability of Revealing Failure is High then Weighted Revealed Failure Rate (FIT) is Very Low

✕ ✎ **17.** If Component Failure Rate (FIT) is Low and Probability of Revealing Failure is High then Weighted Revealed Failure Rate (FIT) is Low

✕ ✎ **18.** If Component Failure Rate (FIT) is Medium and Probability of Revealing Failure is High then Weighted Revealed Failure Rate (FIT) is Medium

✕ ✎ **19.** If Component Failure Rate (FIT) is High and Probability of Revealing Failure is High then Weighted Revealed Failure Rate (FIT) is Medium

✕ ✎ **20.** If Component Failure Rate (FIT) is Very High and Probability of Revealing Failure is High then Weighted Revealed Failure Rate (FIT) is High

✕ ✎ **21.** If Component Failure Rate (FIT) is Very Low and Probability of Revealing Failure is Very High then Weighted Revealed Failure Rate (FIT) is Very Low

✕ ✎ **22.** If Component Failure Rate (FIT) is Low and Probability of Revealing Failure is Very High then Weighted Revealed Failure Rate (FIT) is Low

✕ ✎ **23.** If Component Failure Rate (FIT) is Medium and Probability of Revealing Failure is Very High then Weighted Revealed Failure Rate (FIT) is Medium

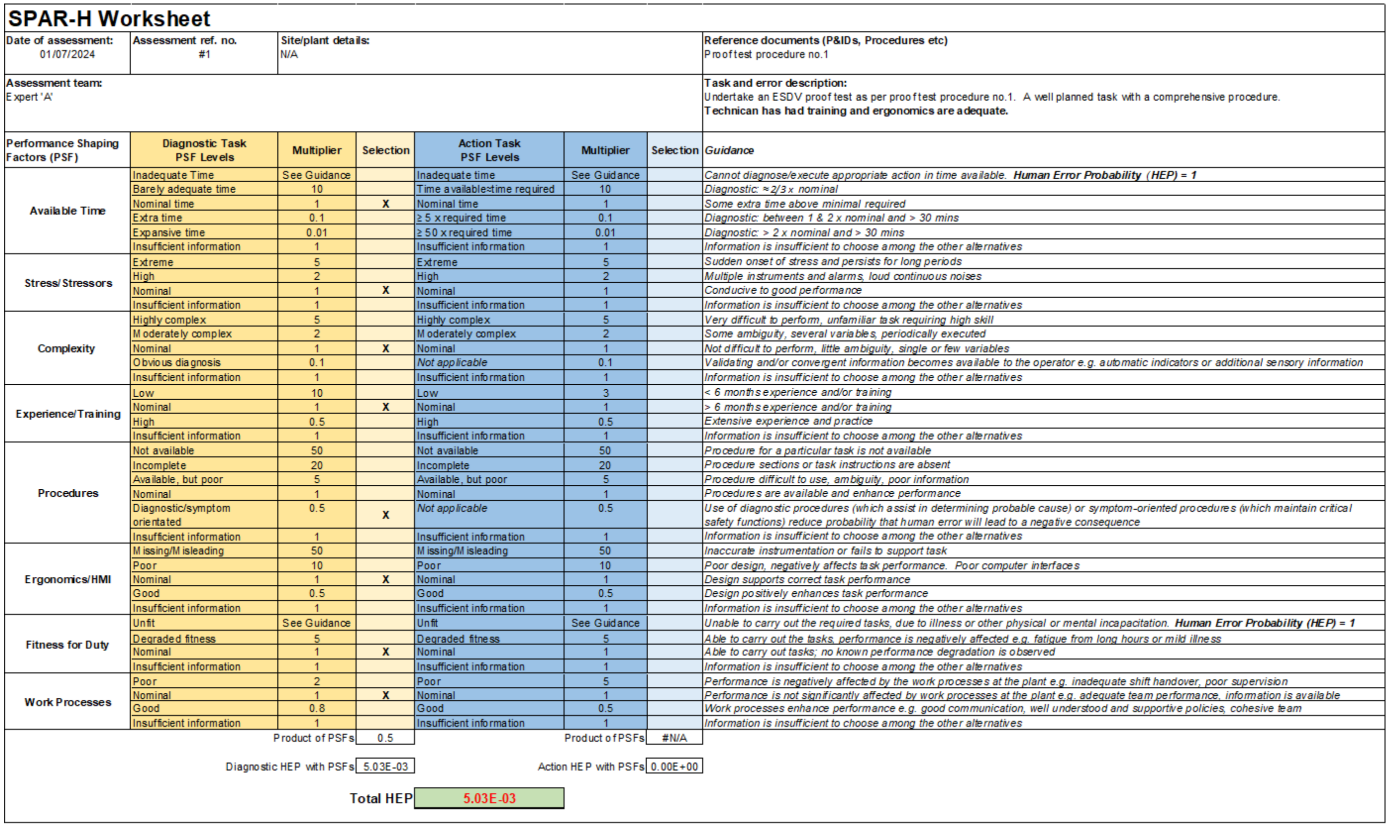
✕ ✎ **24.** If Component Failure Rate (FIT) is High and Probability of Revealing Failure is Very High then Weighted Revealed Failure Rate (FIT) is High

✕ ✎ **25.** If Component Failure Rate (FIT) is Very High and Probability of Revealing Failure is Very High then Weighted Revealed Failure Rate (FIT) is Very High

Appendix E

Human Reliability Assessments

E.1 Expert 'A' - Human Reliability Assessments

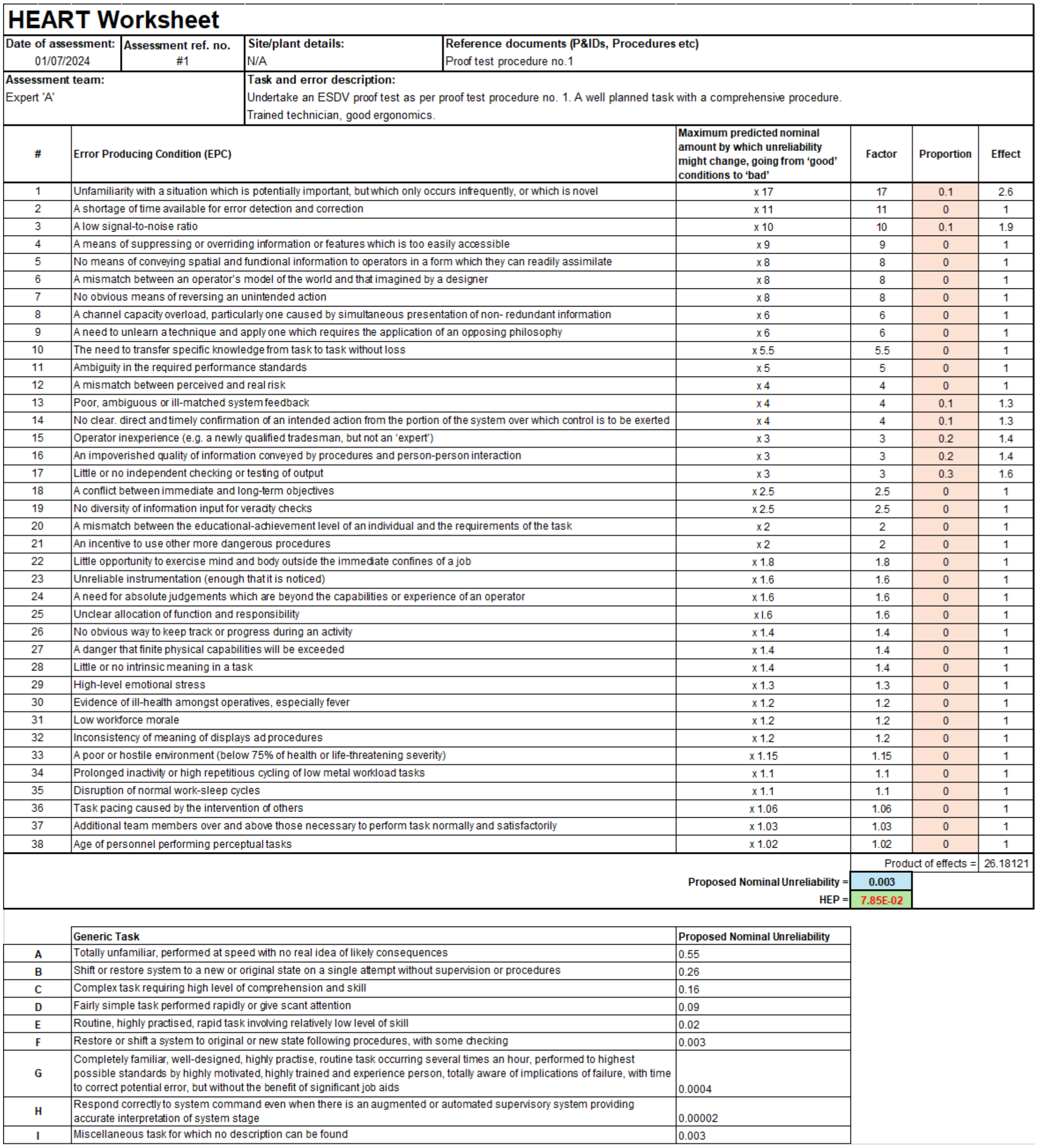


SPAR-H Worksheet

Date of assessment: 01/07/2024	**Assessment ref. no.** #1	**Site/plant details:** N/A	**Reference documents (P&IDs, Procedures etc)** Prooftest procedure no.1
Assessment team: Expert 'A'			**Task and error description:** Undertake an ESDV proof test as per prooftest procedure no.1. A well planned task with a comprehensive procedure. **Technican has had training and ergonomics are adequate.**

Performance Shaping Factors (PSF)	Diagnostic Task PSF Levels	Multiplier	Selection	Action Task PSF Levels	Multiplier	Selection	*Guidance*
Available Time	Inadequate Time	See Guidance		Inadequate time	See Guidance		*Cannot diagnose/execute appropriate action in time available.* ***Human Error Probability (HEP) = 1***
	Barely adequate time	10		Time available≈time required	10		*Diagnostic: ≈2/3 x nominal*
	Nominal time	1	X	Nominal time	1		*Some extra time above minimal required*
	Extra time	0.1		≥ 5 x required time	0.1		*Diagnostic: between 1 & 2 x nominal and > 30 mins*
	Expansive time	0.01		≥ 50 x required time	0.01		*Diagnostic: > 2 x nominal and > 30 mins*
	Insufficient information	1		Insufficient information	1		*Information is insufficient to choose among the other alternatives*
Stress/Stressors	Extreme	5		Extreme	5		*Sudden onset of stress and persists for long periods*
	High	2		High	2		*Multiple instruments and alarms, loud continuous noises*
	Nominal	1	X	Nominal	1		*Conducive to good performance*
	Insufficient information	1		Insufficient information	1		*Information is insufficient to choose among the other alternatives*
Complexity	Highly complex	5		Highly complex	5		*Very difficult to perform, unfamiliar task requiring high skill*
	Moderately complex	2		Moderately complex	2		*Some ambiguity, several variables, periodically executed*
	Nominal	1	X	Nominal	1		*Not difficult to perform, little ambiguity, single or few variables*
	Obvious diagnosis	0.1		*Not applicable*	0.1		*Validating and/or convergent information becomes available to the operator e.g. automatic indicators or additional sensory information*
	Insufficient information	1		Insufficient information	1		*Information is insufficient to choose among the other alternatives*
Experience/Training	Low	10		Low	3		*< 6 months experience and/or training*
	Nominal	1	X	Nominal	1		*> 6 months experience and/or training*
	High	0.5		High	0.5		*Extensive experience and practice*
	Insufficient information	1		Insufficient information	1		*Information is insufficient to choose among the other alternatives*
Procedures	Not available	50		Not available	50		*Procedure for a particular task is not available*
	Incomplete	20		Incomplete	20		*Procedure sections or task instructions are absent*
	Available, but poor	5		Available, but poor	5		*Procedure difficult to use, ambiguity, poor information*
	Nominal	1		Nominal	1		*Procedures are available and enhance performance*
	Diagnostic/symptom orientated	0.5	X	*Not applicable*	0.5		*Use of diagnostic procedures (which assist in determining probable cause) or symptom-oriented procedures (which maintain critical safety functions) reduce probability that human error will lead to a negative consequence*
	Insufficient information	1		Insufficient information	1		*Information is insufficient to choose among the other alternatives*
Ergonomics/HMI	Missing/Misleading	50		Missing/Misleading	50		*Inaccurate instrumentation or fails to support task*
	Poor	10		Poor	10		*Poor design, negatively affects task performance. Poor computer interfaces*
	Nominal	1	X	Nominal	1		*Design supports correct task performance*
	Good	0.5		Good	0.5		*Design positively enhances task performance*
	Insufficient information	1		Insufficient information	1		*Information is insufficient to choose among the other alternatives*
Fitness for Duty	Unfit	See Guidance		Unfit	See Guidance		*Unable to carry out the required tasks, due to illness or other physical or mental incapacitation.* ***Human Error Probability (HEP) = 1***
	Degraded fitness	5		Degraded fitness	5		*Able to carry out the tasks, performance is negatively affected e.g. fatigue from long hours or mild illness*
	Nominal	1	X	Nominal	1		*Able to carry out tasks; no known performance degradation is observed*
	Insufficient information	1		Insufficient information	1		*Information is insufficient to choose among the other alternatives*
Work Processes	Poor	2		Poor	5		*Performance is negatively affected by the work processes at the plant e.g. inadequate shift handover, poor supervision*
	Nominal	1	X	Nominal	1		*Performance is not significantly affected by work processes at the plant e.g. adequate team performance, information is available*
	Good	0.8		Good	0.5		*Work processes enhance performance e.g. good communication, well understood and supportive policies, cohesive team*
	Insufficient information	1		Insufficient information	1		*Information is insufficient to choose among the other alternatives*
		Product of PSFs	0.5		Product of PSFs	#N/A	
		Diagnostic HEP with PSFs	5.03E-03		Action HEP with PSFs	0.00E+00	
		Total HEP		5.03E-03			

Figure E.1: SPAR-H - Favourable Factors



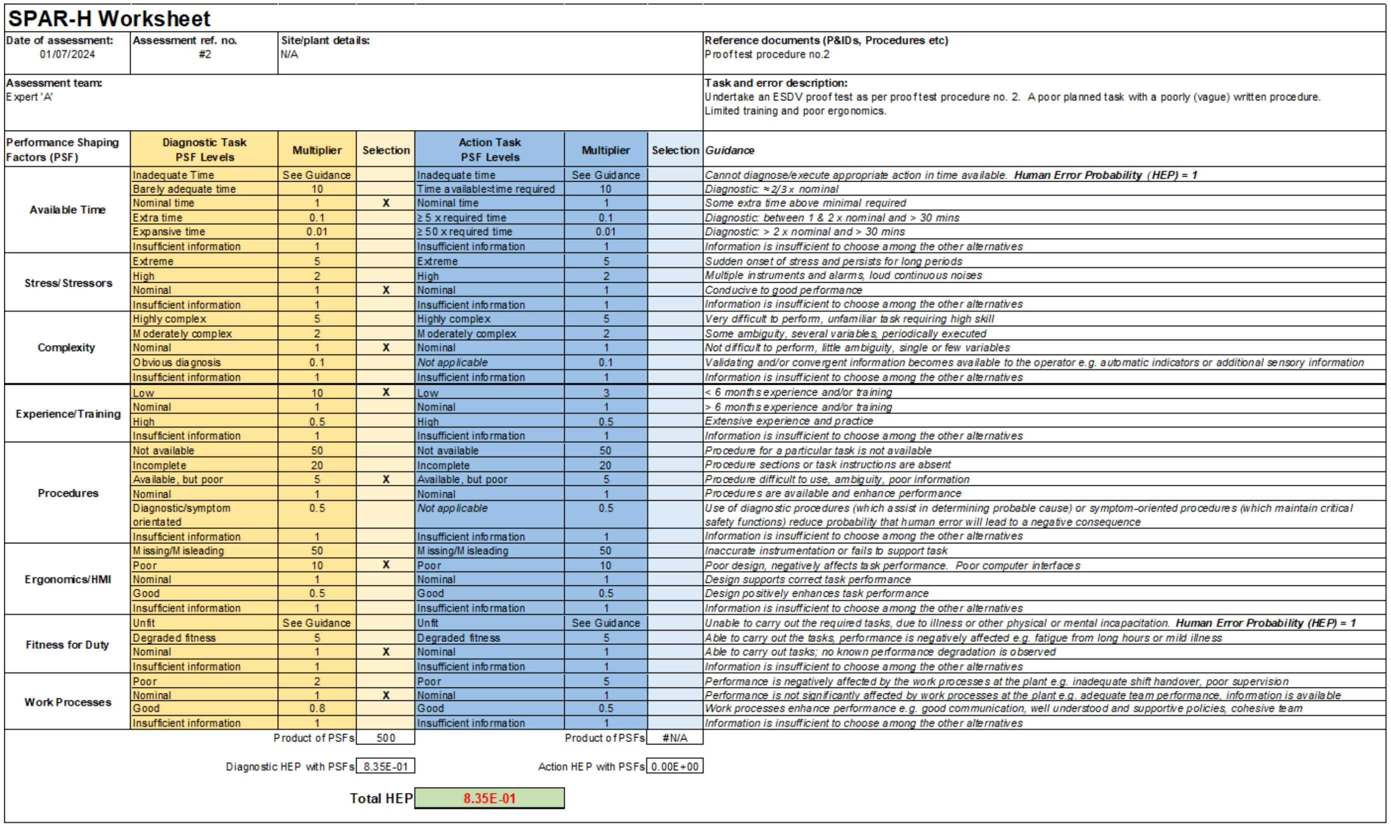
HEART Worksheet

Date of assessment: 01/07/2024	**Assessment ref. no.** #1	**Site/plant details:** N/A	**Reference documents (P&IDs, Procedures etc)** Proof test procedure no.1
Assessment team: Expert 'A'		**Task and error description:** Undertake an ESDV proof test as per proof test procedure no. 1. A well planned task with a comprehensive procedure. Trained technician, good ergonomics.	

#	Error Producing Condition (EPC)	Maximum predicted nominal amount by which unreliability might change, going from 'good' conditions to 'bad'	Factor	Proportion	Effect
1	Unfamiliarity with a situation which is potentially important, but which only occurs infrequently, or which is novel	x 17	17	0.1	2.6
2	A shortage of time available for error detection and correction	x 11	11	0	1
3	A low signal-to-noise ratio	x 10	10	0.1	1.9
4	A means of suppressing or overriding information or features which is too easily accessible	x 9	9	0	1
5	No means of conveying spatial and functional information to operators in a form which they can readily assimilate	x 8	8	0	1
6	A mismatch between an operator's model of the world and that imagined by a designer	x 8	8	0	1
7	No obvious means of reversing an unintended action	x 8	8	0	1
8	A channel capacity overload, particularly one caused by simultaneous presentation of non- redundant information	x 6	6	0	1
9	A need to unlearn a technique and apply one which requires the application of an opposing philosophy	x 6	6	0	1
10	The need to transfer specific knowledge from task to task without loss	x 5.5	5.5	0	1
11	Ambiguity in the required performance standards	x 5	5	0	1
12	A mismatch between perceived and real risk	x 4	4	0	1
13	Poor, ambiguous or ill-matched system feedback	x 4	4	0.1	1.3
14	No clear. direct and timely confirmation of an intended action from the portion of the system over which control is to be exerted	x 4	4	0.1	1.3
15	Operator inexperience (e.g. a newly qualified tradesman, but not an 'expert')	x 3	3	0.2	1.4
16	An impoverished quality of information conveyed by procedures and person-person interaction	x 3	3	0.2	1.4
17	Little or no independent checking or testing of output	x 3	3	0.3	1.6
18	A conflict between immediate and long-term objectives	x 2.5	2.5	0	1
19	No diversity of information input for veracity checks	x 2.5	2.5	0	1
20	A mismatch between the educational-achievement level of an individual and the requirements of the task	x 2	2	0	1
21	An incentive to use other more dangerous procedures	x 2	2	0	1
22	Little opportunity to exercise mind and body outside the immediate confines of a job	x 1.8	1.8	0	1
23	Unreliable instrumentation (enough that it is noticed)	x 1.6	1.6	0	1
24	A need for absolute judgements which are beyond the capabilities or experience of an operator	x 1.6	1.6	0	1
25	Unclear allocation of function and responsibility	x l.6	1.6	0	1
26	No obvious way to keep track or progress during an activity	x 1.4	1.4	0	1
27	A danger that finite physical capabilities will be exceeded	x 1.4	1.4	0	1
28	Little or no intrinsic meaning in a task	x 1.4	1.4	0	1
29	High-level emotional stress	x 1.3	1.3	0	1
30	Evidence of ill-health amongst operatives, especially fever	x 1.2	1.2	0	1
31	Low workforce morale	x 1.2	1.2	0	1
32	Inconsistency of meaning of displays ad procedures	x 1.2	1.2	0	1
33	A poor or hostile environment (below 75% of health or life-threatening severity)	x 1.15	1.15	0	1
34	Prolonged inactivity or high repetitious cycling of low metal workload tasks	x 1.1	1.1	0	1
35	Disruption of normal work-sleep cycles	x 1.1	1.1	0	1
36	Task pacing caused by the intervention of others	x 1.06	1.06	0	1
37	Additional team members over and above those necessary to perform task normally and satisfactorily	x 1.03	1.03	0	1
38	Age of personnel performing perceptual tasks	x 1.02	1.02	0	1
				Product of effects =	26.18121
		Proposed Nominal Unreliability =	**0.003**		
		HEP =	**7.85E-02**		

	Generic Task	Proposed Nominal Unreliability
A	Totally unfamiliar, performed at speed with no real idea of likely consequences	0.55
B	Shift or restore system to a new or original state on a single attempt without supervision or procedures	0.26
C	Complex task requiring high level of comprehension and skill	0.16
D	Fairly simple task performed rapidly or give scant attention	0.09
E	Routine, highly practised, rapid task involving relatively low level of skill	0.02
F	Restore or shift a system to original or new state following procedures, with some checking	0.003
G	Completely familiar, well-designed, highly practise, routine task occurring several times an hour, performed to highest possible standards by highly motivated, highly trained and experience person, totally aware of implications of failure, with time to correct potential error, but without the benefit of significant job aids	0.0004
H	Respond correctly to system command even when there is an augmented or automated supervisory system providing accurate interpretation of system stage	0.00002
I	Miscellaneous task for which no description can be found	0.003

Figure E.2: HEART - Favourable Factors



SPAR-H Worksheet

Date of assessment:	Assessment ref. no.	Site/plant details:	Reference documents (P&IDs, Procedures etc)
01/07/2024	#2	N/A	Prooftest procedure no.2

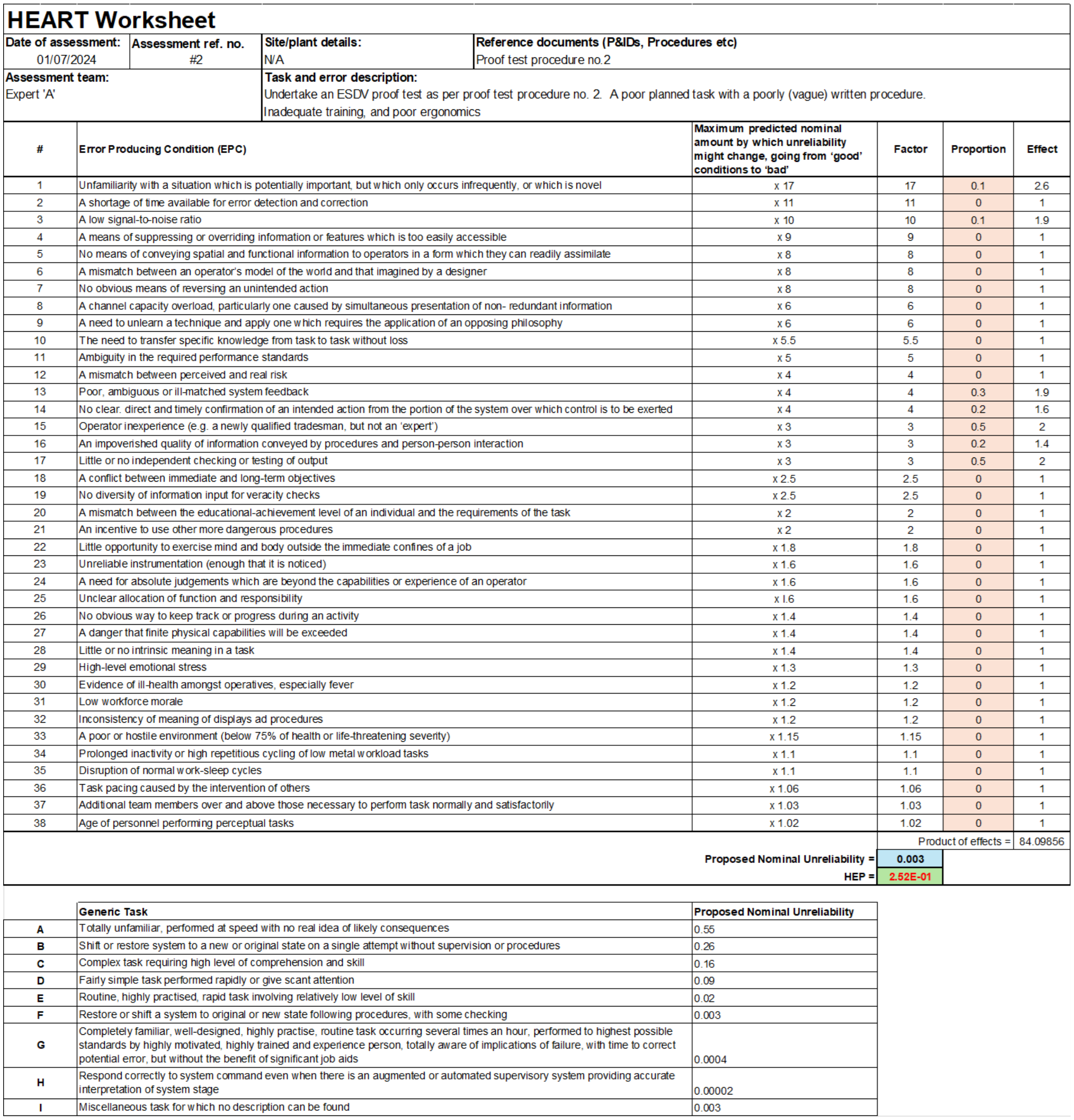
Assessment team:
Expert 'A'

Task and error description:
Undertake an ESDV proof test as per prooftest procedure no. 2. A poor planned task with a poorly (vague) written procedure. Limited training and poor ergonomics.

Performance Shaping Factors (PSF)	Diagnostic Task PSF Levels	Multiplier	Selection	Action Task PSF Levels	Multiplier	Selection	*Guidance*
Available Time	Inadequate Time	See Guidance		Inadequate time	See Guidance		*Cannot diagnose/execute appropriate action in time available.* ***Human Error Probability (HEP) = 1***
	Barely adequate time	10		Time available≈time required	10		*Diagnostic: ≈2/3 x nominal*
	Nominal time	1	X	Nominal time	1		*Some extra time above minimal required*
	Extra time	0.1		≥ 5 x required time	0.1		*Diagnostic: between 1 & 2 x nominal and > 30 mins*
	Expansive time	0.01		≥ 50 x required time	0.01		*Diagnostic: > 2 x nominal and > 30 mins*
	Insufficient information	1		Insufficient information	1		*Information is insufficient to choose among the other alternatives*
Stress/Stressors	Extreme	5		Extreme	5		*Sudden onset of stress and persists for long periods*
	High	2		High	2		*Multiple instruments and alarms, loud continuous noises*
	Nominal	1	X	Nominal	1		*Conducive to good performance*
	Insufficient information	1		Insufficient information	1		*Information is insufficient to choose among the other alternatives*
Complexity	Highly complex	5		Highly complex	5		*Very difficult to perform, unfamiliar task requiring high skill*
	Moderately complex	2		Moderately complex	2		*Some ambiguity, several variables, periodically executed*
	Nominal	1	X	Nominal	1		*Not difficult to perform, little ambiguity, single or few variables*
	Obvious diagnosis	0.1		*Not applicable*	0.1		*Validating and/or convergent information becomes available to the operator e.g. automatic indicators or additional sensory information*
	Insufficient information	1		Insufficient information	1		*Information is insufficient to choose among the other alternatives*
Experience/Training	Low	10	X	Low	3		*< 6 months experience and/or training*
	Nominal	1		Nominal	1		*> 6 months experience and/or training*
	High	0.5		High	0.5		*Extensive experience and practice*
	Insufficient information	1		Insufficient information	1		*Information is insufficient to choose among the other alternatives*
Procedures	Not available	50		Not available	50		*Procedure for a particular task is not available*
	Incomplete	20		Incomplete	20		*Procedure sections or task instructions are absent*
	Available, but poor	5	X	Available, but poor	5		*Procedure difficult to use, ambiguity, poor information*
	Nominal	1		Nominal	1		*Procedures are available and enhance performance*
	Diagnostic/symptom orientated	0.5		*Not applicable*	0.5		*Use of diagnostic procedures (which assist in determining probable cause) or symptom-oriented procedures (which maintain critical safety functions) reduce probability that human error will lead to a negative consequence*
	Insufficient information	1		Insufficient information	1		*Information is insufficient to choose among the other alternatives*
Ergonomics/HMI	Missing/Misleading	50		Missing/Misleading	50		*Inaccurate instrumentation or fails to support task*
	Poor	10	X	Poor	10		*Poor design, negatively affects task performance. Poor computer interfaces*
	Nominal	1		Nominal	1		*Design supports correct task performance*
	Good	0.5		Good	0.5		*Design positively enhances task performance*
	Insufficient information	1		Insufficient information	1		*Information is insufficient to choose among the other alternatives*
Fitness for Duty	Unfit	See Guidance		Unfit	See Guidance		*Unable to carry out the required tasks, due to illness or other physical or mental incapacitation.* ***Human Error Probability (HEP) = 1***
	Degraded fitness	5		Degraded fitness	5		*Able to carry out the tasks, performance is negatively affected e.g. fatigue from long hours or mild illness*
	Nominal	1	X	Nominal	1		*Able to carry out tasks; no known performance degradation is observed*
	Insufficient information	1		Insufficient information	1		*Information is insufficient to choose among the other alternatives*
Work Processes	Poor	2		Poor	5		*Performance is negatively affected by the work processes at the plant e.g. inadequate shift handover, poor supervision*
	Nominal	1	X	Nominal	1		*Performance is not significantly affected by work processes at the plant e.g. adequate team performance, information is available*
	Good	0.8		Good	0.5		*Work processes enhance performance e.g. good communication, well understood and supportive policies, cohesive team*
	Insufficient information	1		Insufficient information	1		*Information is insufficient to choose among the other alternatives*
		Product of PSFs	500		Product of PSFs	#N/A	
		Diagnostic HEP with PSFs	8.35E-01		Action HEP with PSFs	0.00E+00	

Total HEP 8.35E-01

Figure E.3: SPAR-H - Unfavourable Factors



HEART Worksheet

Date of assessment: 01/07/2024	**Assessment ref. no.** #2	**Site/plant details:** N/A	**Reference documents (P&IDs, Procedures etc)** Proof test procedure no.2
Assessment team: Expert 'A'		**Task and error description:** Undertake an ESDV proof test as per proof test procedure no. 2. A poor planned task with a poorly (vague) written procedure. Inadequate training, and poor ergonomics	

#	Error Producing Condition (EPC)	Maximum predicted nominal amount by which unreliability might change, going from 'good' conditions to 'bad'	Factor	Proportion	Effect
1	Unfamiliarity with a situation which is potentially important, but which only occurs infrequently, or which is novel	x 17	17	0.1	2.6
2	A shortage of time available for error detection and correction	x 11	11	0	1
3	A low signal-to-noise ratio	x 10	10	0.1	1.9
4	A means of suppressing or overriding information or features which is too easily accessible	x 9	9	0	1
5	No means of conveying spatial and functional information to operators in a form which they can readily assimilate	x 8	8	0	1
6	A mismatch between an operator's model of the world and that imagined by a designer	x 8	8	0	1
7	No obvious means of reversing an unintended action	x 8	8	0	1
8	A channel capacity overload, particularly one caused by simultaneous presentation of non- redundant information	x 6	6	0	1
9	A need to unlearn a technique and apply one which requires the application of an opposing philosophy	x 6	6	0	1
10	The need to transfer specific knowledge from task to task without loss	x 5.5	5.5	0	1
11	Ambiguity in the required performance standards	x 5	5	0	1
12	A mismatch between perceived and real risk	x 4	4	0	1
13	Poor, ambiguous or ill-matched system feedback	x 4	4	0.3	1.9
14	No clear. direct and timely confirmation of an intended action from the portion of the system over which control is to be exerted	x 4	4	0.2	1.6
15	Operator inexperience (e.g. a newly qualified tradesman, but not an 'expert')	x 3	3	0.5	2
16	An impoverished quality of information conveyed by procedures and person-person interaction	x 3	3	0.2	1.4
17	Little or no independent checking or testing of output	x 3	3	0.5	2
18	A conflict between immediate and long-term objectives	x 2.5	2.5	0	1
19	No diversity of information input for veracity checks	x 2.5	2.5	0	1
20	A mismatch between the educational-achievement level of an individual and the requirements of the task	x 2	2	0	1
21	An incentive to use other more dangerous procedures	x 2	2	0	1
22	Little opportunity to exercise mind and body outside the immediate confines of a job	x 1.8	1.8	0	1
23	Unreliable instrumentation (enough that it is noticed)	x 1.6	1.6	0	1
24	A need for absolute judgements which are beyond the capabilities or experience of an operator	x 1.6	1.6	0	1
25	Unclear allocation of function and responsibility	x l.6	1.6	0	1
26	No obvious way to keep track or progress during an activity	x 1.4	1.4	0	1
27	A danger that finite physical capabilities will be exceeded	x 1.4	1.4	0	1
28	Little or no intrinsic meaning in a task	x 1.4	1.4	0	1
29	High-level emotional stress	x 1.3	1.3	0	1
30	Evidence of ill-health amongst operatives, especially fever	x 1.2	1.2	0	1
31	Low workforce morale	x 1.2	1.2	0	1
32	Inconsistency of meaning of displays ad procedures	x 1.2	1.2	0	1
33	A poor or hostile environment (below 75% of health or life-threatening severity)	x 1.15	1.15	0	1
34	Prolonged inactivity or high repetitious cycling of low metal workload tasks	x 1.1	1.1	0	1
35	Disruption of normal work-sleep cycles	x 1.1	1.1	0	1
36	Task pacing caused by the intervention of others	x 1.06	1.06	0	1
37	Additional team members over and above those necessary to perform task normally and satisfactorily	x 1.03	1.03	0	1
38	Age of personnel performing perceptual tasks	x 1.02	1.02	0	1
				Product of effects =	84.09856
		Proposed Nominal Unreliability =	**0.003**		
		HEP =	**2.52E-01**		

	Generic Task	Proposed Nominal Unreliability
A	Totally unfamiliar, performed at speed with no real idea of likely consequences	0.55
B	Shift or restore system to a new or original state on a single attempt without supervision or procedures	0.26
C	Complex task requiring high level of comprehension and skill	0.16
D	Fairly simple task performed rapidly or give scant attention	0.09
E	Routine, highly practised, rapid task involving relatively low level of skill	0.02
F	Restore or shift a system to original or new state following procedures, with some checking	0.003
G	Completely familiar, well-designed, highly practise, routine task occurring several times an hour, performed to highest possible standards by highly motivated, highly trained and experience person, totally aware of implications of failure, with time to correct potential error, but without the benefit of significant job aids	0.0004
H	Respond correctly to system command even when there is an augmented or automated supervisory system providing accurate interpretation of system stage	0.00002
I	Miscellaneous task for which no description can be found	0.003

Figure E.4: HEART - Unfavourable Factors

Appendix F

Conference Paper

A conference paper was presented at the Industrial Valve Summit (IVS), Bergamo, Italy, 2024.

Kriescher et al. Estimation of Proof Test Coverage for Emergency Shutdown Valves. Presented at: IVS2024 – Industrial Valve Summit; 2024 May; Bergamo, Italy. Available from: https://www.valvecampus.com/our-services/knowledge-platform/

Conference programme:
https://industrialvalvesummit.com/conference-programme-2024/

A copy of the paper is included in the following pages.

Estimation of Proof Test Coverage for Emergency Shut Down Valves

Steve Kriescher[a,b] ([redacted]),
[redacted], Roderick Thomas[c]
([redacted]), Chris Phillips[b] ([redacted]),
Neil Mac Parthaláin[d] ([redacted]) and David J Smith[e]
([redacted])

[a]Instrument and Control Solutions, Neath, SA10 7PF. Wales; [b]Department of Chemical Engineering, Faculty of Science and Engineering, Swansea University, Swansea, SA1 8EN. Wales; [c]School of Management, Swansea University, Swansea, SA1 8EN. Wales; [d]Department of Computer Science, Aberystwyth University, Aberystwyth, Ceredigion, SY23 3DB. Wales; [e]Technis, Tonbridge, TN10 4LG. England.

IVS 2024 Conference, Bergamo, Italy. 15th to 16th May 2024

ABSTRACT

Published proof test coverage (PTC) estimates for emergency shut down valves (ESDVs) are only in moderate agreement with each other and are largely opinion based. A study was undertaken to determine the effectiveness of ESDV proof testing using data collected from the field, and use of a failure mode and effects based methodology. ESDV maintenance and test data were collected from a number of operating facilities in the oil and gas industry and subsequently analysed to estimate PTC. In parallel, a Failure Modes, Effects and Diagnostics Analysis (FMEDA) was undertaken using published component failure rate data to predict PTC for a specific type of ESDV proof test. Due to the subjective nature of the FMEDA technique, in particular the selection of component failure rates and the determination of the probability of revealing failure modes, a Fuzzy Inference System (FIS) was proposed to handle the data in recognition of the implicit uncertainties. The results from the field data, the FMEDA approach and the FIS suggested PTC values for a full stroke test (FST) ranged from 73% to 87%, which is in agreement with the majority of published estimates. The estimates for the FMEDA and Fuzzy-FMEDA approaches were 73% and 82%, respectively, indicating a more optimistic result for the Fuzzy-FMEDA approach which also aligned with the results obtained from the field data analysis. The study concluded that the PTC for an FST is approximately 80% to 85%.

KEYWORDS

SIS, ESDV, Proof Test Coverage, FMEDA, Fuzzy Inference System Uncertainty

1. Introduction

Functional safety guidance, such as IEC 61508 [1] and IEC 61511 [2], are used for bench-marking within the Control of Major Accident Hazards (COMAH) regulations [3], and has evolved in order to encourage the application of appropriate safety engineering activities. A safety instrumented system (SIS) has a number of individual safety instrumented functions (SIFs) which protect a process plant/facility from potentially hazardous events. SIFs are designed to protect against specific hazardous events for example, overfilling a storage tank or over pressurising a pipeline, hence preventing a loss of containment. A typical SIS architecture consists of sensor(s), logic solver(s) and final element(s) as shown in Figure 1 [2]:

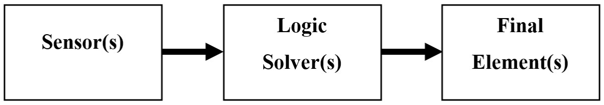


Figure 1.: SIS Architecture

Final elements are typically emergency shut down valves (ESDVs). The purpose of proof testing is to reveal potentially dangerous failures, which are not revealed by equipment on-board diagnostics. In the case of ESDVs, diagnostics are limited or even non-existent and therefore reliance on effective proof testing is essential. The effectiveness of proof testing is also known as proof test coverage (PTC), which is the percentage of dangerous failures that are revealed during a proof test, given by the following equation. Whereby λ_{DU} is the dangerous undetected failure rate:

$$PTC = \frac{Revealed\lambda_{\mathrm{DU}}}{Total\lambda_{\mathrm{DU}}} \tag{1}$$

There are a number of different ESDV proof test regimes, as shown below [4], which will have varying PTC:

(1) Partial valve stroke test – the valve moved typically 5-10%, online (at process operating conditions)
(2) Full valve stroke test – the valve moved to its closed (or open) position, offline (not at process operating conditions) – this test was the focus of the study.
(3) Full valve stroke test at process operating conditions – as per test (2) with the valve online
(4) Full valve stroke test and leak test – with the valve offline
(5) Full valve stroke test at process operating conditions and leak test – as per test (4) with the valve online

PTC is parameter used when calculating a SIFs Probability of Failure on Demand (PFD_{avg}) and hence determining proof test interval. Estimating credible values of PTC is essential as poor prediction of PTC can potentially mean that test frequency is, on one hand, inadequate (compromising safety) or, on the other, excessive (a cost penalty). A commonly used proof test regime is the full valve stroke test as described in (2) above. A number of industry experts and organisations have provided estimates of PTC for this type of test, these are presented in Figure 2 below. This shows the extent of the variability and lack of consensus with wide ranging estimates of PTC for a typical proof test, from 35 to 90%.

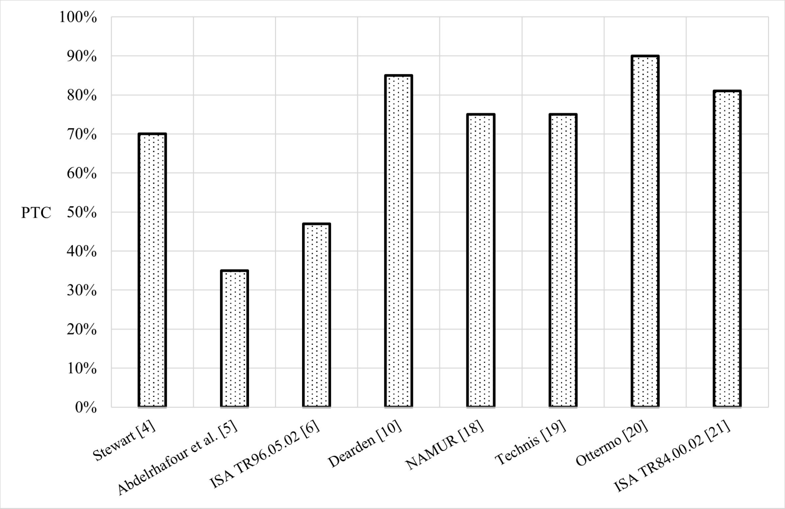


Figure 2.: PTC estimates from literature

The PTC variance poses a significant challenge for the practitioner to decide which PTC value to select when calculating PFD_{avg}. In light of the differing PTC estimations, the following section investigates estimating PTC, using the following methods:

(1) Field data analysis
(2) Failure modes, effects and diagnostics analysis (FMEDA)
(3) FMEDA using a Fuzzy Inference System (FIS)

2. Methodology

The focus of the methods outlined in this section are to estimate PTC for an ESDV proof test as described in test (2) above, i.e., a full stroke test (FST). The ESDV under investigation is fail safe, close on demand. For the application considered minor seat leakage is acceptable to achieve a safe state, i.e., there is no requirement for tight shut-off (TSO). The ESDV assembly considered in this study consists of the following devices:

- Solenoid Operated Valve (SOV) - A pneumatic 3/2 direct acting SOV. Configured as fail safe closed, i.e., on loss of signal the SOV returns to its vent (safe) position.
- Actuator - A pneumatic scotch yoke, spring return type actuator. Configured as fail safe, i.e., on loss of air supply the actuator will return to its safe state, hence returning the valve to its closed position.
- Ball valve - A full bore, trunnion mounted ball valve with static seat rings and springs.

2.1. PTC Determined from Field Data

ESDV maintenance and test data was provided by a number of oil and gas industry companies. The companies collected the data using their computerised maintenance and management systems (CMMS) and access to the data for this research was provided in the form of data downloads from these systems into spreadsheets. The data was then reviewed to determine dangerous failure modes following the guidance in standard BS EN ISO 14224 *Petroleum, petrochemical and natural gas industries - Collection and exchange of reliability and maintenance* [7]. This aim of this standard is to improve the way the industry collects data so that it can be exchanged to enhance safety, availability, reliability and maintainability of equipment. The standard covers various industrial equipment including numerous types of valves, and for each piece of equipment the potential failure modes and associated failure codes.

2.1.1. Data Analysis

The collected data was analysed to determine failure modes of the equipment, where possible. As previously stated, BS EN ISO 14224 provides failure modes for specific pieces of equipment. Proof test coverage is based on revealed and unrevealed dangerous failure modes. The dangerous failure modes for a fail safe closed EDSV are shown below:

- Fail to close on demand (FTC), i.e., stuck open or fail to close fully
- Delayed operation (DOP), i.e., slow to close
- Leak in closed position (LCP), i.e., excessive leak in closed position

Noting that the LCP failure mode would not be revealed when considering a full stroke test.

2.2. Prediction of PTC using Failure Modes, Effects and Diagnostics Analysis (FMEDA)

FMEDA is a technique originally developed to determine diagnostic coverage of electronic equipment, by identifying electronic component failure modes their effects on the system and whether the failures are detected or not by diagnostics [8]. The technique requires component failure rates to be assigned to failure modes which allows the distribution of failures and the overall diagnostic coverage to be determined. The FMEDA technique as described by Stewart [4] can also be utilised to determine proof test coverage for mechanical systems such as ESDVs. Each component part is assessed to determine how it can fail (failure mode) and what impact the failure (failure effect) has on the operation of the system, i.e., will the failure effect be safe, dangerous or have no effect. A failure rate is then assigned to the failure mode of the component, which is likely to be a distributed value from the overall component failure rate as there are usually more than one type of failure mechanism. The component failure modes, component failure rates and failure rate distributions were taken from exida's Component Reliability Data (CRD) handbook [11]. The final part of the assessments includes consideration of whether the failure mode will be revealed during the proof test and the reliability of the verification method of the proof test, as described below [9]:

- Revealability - to what degree is the failure mode reveal-able during a proof test, e.g., a partial stroke test (PST) will only reveal a relatively small percentage of the delayed operation (DOP) failure mode.
- Reliability - to what degree are the proof test results dependable, such that the revealed results reflect the valve condition, e.g. the reliability of the limit switches indicating fail to fully close (FTC) failure mode.

The product of the component failure rate (for each particular failure mode), its revealibility and the test reliability provides a 'weighted' revealed dangerous undetected failure rate. The resultant PTC is calculated, as given in equation (1) above, from the sum of the weighted revealed dangerous undetected failure rates ($Revealed_{\lambda}DU$) and the sum of the total dangerous undetected failure rates ($Total_{\lambda}DU$).

2.3. Prediction of PTC using Fuzzy-FMEDA

The FMEDA technique requires the human judgement of the assessor, in particular the selection of appropriate component failure rate and the identification of failure modes. As highlighted by Dearden [10], there are also uncertainties in the failure rates as well as the identification of the failure modes in any given context, which makes any attempt to determine precise percentages of PTC very challenging. Likewise, research undertaken by Isenburg [12] suggests that the lack of mechanical component failure rate databases leads to subjectivity which ultimately impacts the outcome of the assessment.

Fuzzy reasoning is based upon fuzzy set theory [16] which is a generalisation of classical set theory. In fuzzy set theory, membership can be defined in the interval [0,1]. Fuzzy logic is a many-valued logic and an extension of classical logic and is built on fuzzy set theory. Fuzzy reasoning offers the ability to reason using vague, human-defined concepts, e.g., 'fairly warm weather', and therefore is often referred to as 'computing with words', hence can be utilised for the sorts of imprecise descriptions applied to valve failure. It has had successful application to a wide variety of problems relating to: control, knowledge representation and modelling, and more generally to decision systems that can handle vagueness in data. A fuzzy inference system (FIS) uses the process of fuzzification to model the input domain of the problem, under consideration, thus mapping crisp values from the real-world into linguistic variables.

2.3.1. Fuzzy Inference Systems

Fuzzy inference is the process of mapping from a given set of inputs to an output using fuzzy logic as the inference step. This mapping provides a basis from which decisions can be made, or by which patterns can be recognised by firing various associated rules. The actual process of fuzzy inference involves several components that are described in the membership functions, linguistic variables, and the rule base.

Mamdani inference was initially proposed [14] as an approach to create a control system by synthesising a set of linguistic control rules derived from human expert opinion. In such an approach, the output of each rule is a fuzzy set. As Mamdani systems have more intuitive and easy-to-understand rule bases, they lend themselves well to application in the area of expert systems, where the rules are derived from human expert knowledge, e.g., medical diagnosis. The general approach to the Mamdani inference system is described below and is summarised in Figure 3.

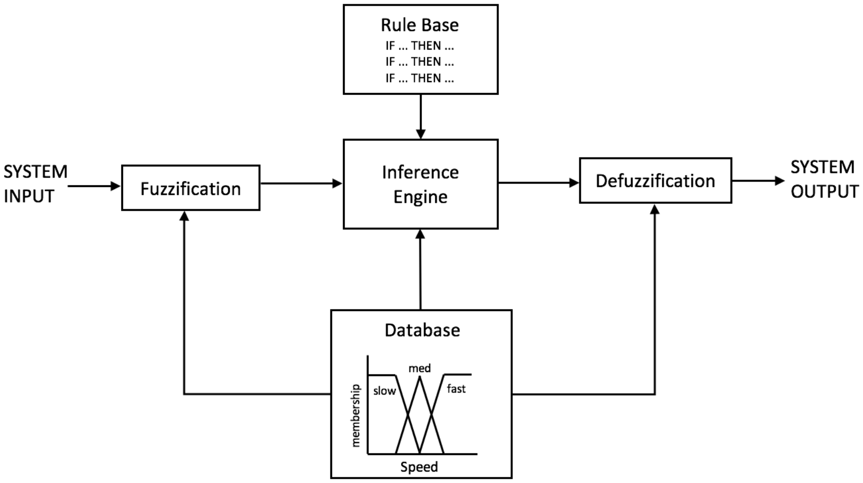


Figure 3.: Fuzzy Inference System

Broadly speaking there are three primary components to any FIS:

- Fuzzification - the process of mapping crisp real-world values onto fuzzy linguistic variables.
- Inference engine - fires the fuzzy if-then rules according to the fuzzy input in order to derive a consequent (output). In particular, the fuzzy if-then rules are used to evaluate linguistic values and map them to an output fuzzy set using the firing strength.
- Defuzzification - this step converts the output or consequent of the inference engine to a crisp value.

2.3.2. Fuzzification

The FIS was developed using Juzzy Online [23], an online fuzzy logic toolkit. In order to apply FIS to the FMEDA problem, the weighted revealed failure rate of a component needs to be determined. The following inputs of each dangerous failure are required in order to do so:

- Component failure rate (distributed)
- Possibility of revealing the failure

The fuzzification process maps the component failure rate and the possibility of revealing the failure inputs into their respective fuzzy subsets. The linguistic variables for these inputs are defined based upon the values of the input domains. These were obtained by partitioning the input domains, component failure rate and possibility of revealing the failure, as shown in Tables 1 and 2. The possibility of revealing the failure was also based on the FMEDA data and partitioned in equal increments from 0 to a possibility of 1. The overlaps between adjacent membership functions permit smooth interpolation of the inputs across membership functions. FIT is the number of failures that can be expected in one billion (10^9) device hours of operation.

Linguistic terms	Component failure rate (FIT)
Very Low	0 - 60
Low	40 - 110
Medium	90 - 210
High	190 - 310
Very High	290 - 410

Table 1.: Component Failure Rate Partitioning

Linguistic terms	Possibility of revealing the failure
Very Low	0 - 0.21
Low	0.2 - 0.41
Medium	0.4 - 0.61
High	0.6 - 0.81
Very High	0.8 - 1

Table 2.: Possibility of Revealing Failure Partitioning

Using the information from Tables 1 and 2, the fuzzification step can be implemented by expressing each partition as a fuzzy subset with the associated linguistic variable. Note that it is possible to employ many different types of fuzzy membership function, however, for real-world applications, it has been shown that triangular and trapezoidal functions are sufficient for most applications [22]. In terms of Fuzzy-FMEDA, for component failure rate, triangular memberships (Figure 4) are employed, while for Possibility of revealing the failure, trapezoidal memberships (Figure 5) are employed to describe the aforementioned partitions. The scaling describes the range of input values and their corresponding membership to each of the fuzzy linguistic labels and their respective fuzzy subsets. The y-axis represents the degree of membership to a fuzzy subset, so an input value that is in the middle of a subset has full membership of that category.

The input values in Figures 4 and 5 can be determined either quantitatively or qualitatively. For example, the frequency of occurrence for a failure mode may be derived from a detailed reliability analysis of the system or from failure data collected on similar designs. On the other hand, the possibility of revealing the failure ranking may be based entirely on the vague and imprecise expert opinion of the assessors experience and judgement. Any input can be fuzzified by mapping the component failure rate value and the possibility of revealing the failure assessment value to the corresponding antecedent input. The membership degrees for the corresponding fuzzy sets can then be ascertained.

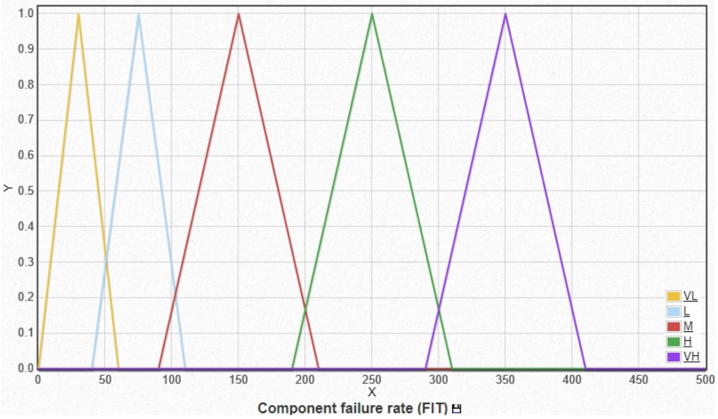


Figure 4.: Input Membership Functions: Component failure rate

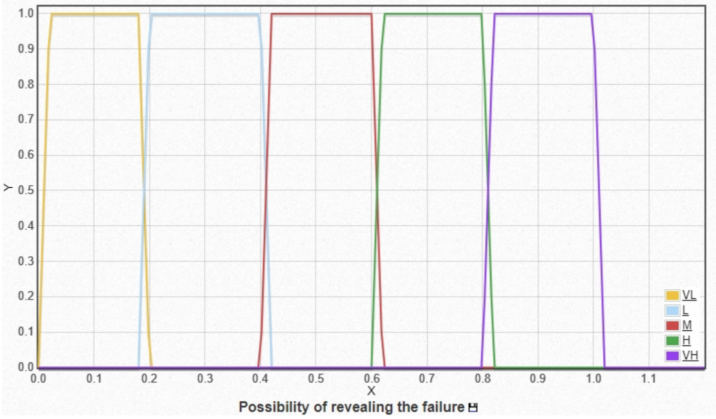


Figure 5.: Input Membership Functions: Possibility of revealing the failure

2.3.3. Rule base

25 different rules were developed based on the observations from the data, examples of which are shown below:

- Rule 10 - IF *Component failure rate (FIT)* is *Low* AND *Possibility of revealing the failure* is *Very High* THEN *Weighted revealed failure rate (FIT)* is *Low*
- Rule 11 - IF *Component failure rate (FIT)* is *Medium* AND *Possibility of revealing the failure* is *Very Low* THEN *Weighted revealed failure rate (FIT)* is *Very Low*

2.3.4. Partitioning of the Output and Defuzzification

The output membership partition and linguistic terms are shown in Table 3. The corresponding set of consequents and the membership functions for the output are shown in Figure 6 below. Defuzzification of the output is conducted using most commonly used method, the centre-of-gravity method (COG). This method determines the centre of area of fuzzy output set and returns the corresponding crisp value.

Linguistic term	Weighted revealed failure rate (FIT)
Very Low	0 - 50
Low	45 - 100
Medium	95 - 150
High	145 - 250
Very High	240 - 450

Table 3.: Output Membership Partitioning

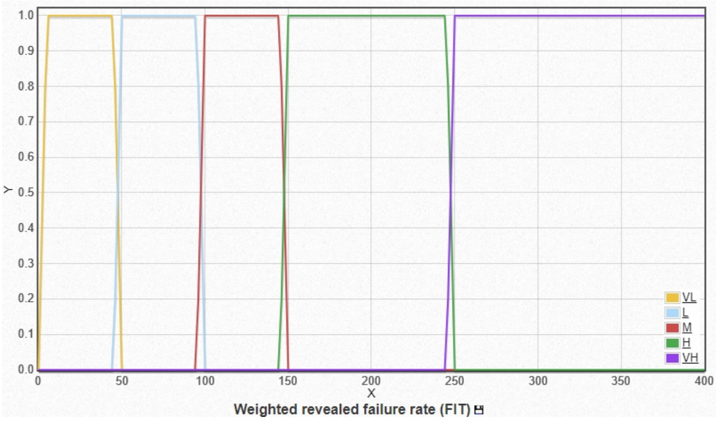


Figure 6.: Output Membership Functions - Weighted revealed failure rate

3. Results

3.1. Results from Field Data

The analysis included a total population 387 ESDVs with an aggregated operational time of 53.9 million hours. The overall dangerous failure rate was calculated to be 1243 Failures In Time (FIT). Note: FIT is the number of failures that can be expected in one billion (10^9) device hours of operation.
The overall dangerous failure rate of 1243 FIT falls within the range quoted by exida [15], and in good agreement with OREDA [16], SINTEF [20] and Technis [17].

The results from the field data analysis are shown below in Table 4. PTC for a full stroke test was estimated at 81% , 87% and 85% from Data Source A, B and C, respectively.

Failure Mode	No. of Failures (A)	Failure Distribution (A)	No. of Failures (B)	Failure Distribution (B)	No. of Failures (C)	Failure Distribution (C)
FTC	9	56%	17	71%	13	48%
DOP	4	25%	4	16%	10	37%
LCP	3	19%	3	13%	4	15%
Total	16	100%	24	100%	27	100%

Table 4.: Results from Data Source A, B and C

3.2. Results from FMEDA and Fuzzy-FMEDA

Table 5 shows a comparison of the outputs from the FMEDA and Fuzzy-FMEDA techniques giving the total dangerous failures for an ESDV assembly and the total number of dangerous failures expected to be revealed during a full stroke proof test.

As a means to validate the results from the FMEDA and Fuzzy-FMEDA techniques comparisons were made against two industry data sources. The industry data sources suggest that the dangerous failure rate for the type of ESDV assessed in this study is in the region of 700 to 2700 FIT [15] and predominantly 1000 FIT [17]. Both sets of results fall within the range quoted by exida [15], albeit on the lower end of the scale at 784 and 970 FIT, and in good agreement with Technis [17].

	FMEDA	Fuzzy-FMEDA
Total dangerous failures (FIT)	784	970
Total dangerous failures revealed (FIT)	571	796
Resultant PTC (%)	73	82

Table 5.: FMEDA and Fuzzy-FMEDA Results

3.3. Overall Comparison

The PTC values determined from the field data, FMEDA and Fuzzy-FMEDA are shown below in Table 6.

Method	Data	PTC %
Field Data Analysis (ISO 14224)	Data Source A	81
Field Data Analysis (ISO 14224)	Data Source B	87
Field Data Analysis (ISO 14224)	Data Source C	85
FMEDA	exida CRD	73
Fuzzy-FMEDA	FMEDA Data	82

Table 6.: Summary of PTC results from field data, FMEDA and Fuzzy-FMEDA approaches

The results suggest that the FMEDA technique aligns with the PTC estimates provided by NAMUR [18], Stewart [4] and Technis [19], in the 70 to 80% region. While the Fuzzy-FMEDA result aligns with the PTC estimates provided by Dearden [10], Ottermo [20], ISA TR84.00.02 [21] and the PTC estimates determined from the field data (Section 3.1), in the 80 to 90% region.

The Fuzzy-FMEDA estimates are optimistic in comparison to the FMEDA, the difference is primarily due to the FIS being insensitive to lower failure rates, owing to the manner of how the membership functions are grouped. It can be concluded that the difference in PTC is an indication of the inherent uncertainty with the FMEDA data.

4. Conclusions

The results of this study indicate that the PTC for a typical proof test (the full stroke timed test) based on the methods adopted is in the region of 73 to 87%, which represents agreement with the majority of the published estimates.

This paper presented a novel implementation of a fuzzy inference system in order to handle the implicit uncertainty in the FMEDA approach. Since the Fuzzy-FMEDA method aligns well with the PTC results obtained from the field data analysis it is reasonable to infer, from this initial work, that the PTC for a FST is approximately 80% to 85%.

The Fuzzy-FMEDA approach outlined in this paper has the potential to resolve some of the problems associated with the FMEDA technique. As well as quantitative data, a fuzzy inference system can also handle qualitative, ambiguous and imprecise or vague data. The only limitation of the Fuzzy-FMEDA approach is an FIS has to be developed using a specific piece of software and it can be time consuming to design and test the system.

This investigation, however, forms the basis for a more extensive study and future work. An optimised FIS should be designed by a group of experts as this will provide much richer and more representative expert judgement, models and ultimately results.

Acknowledgements
SK would like to thank Dr. William Goble (exida) for kindly providing a copy of the Component Reliability Database (CRD) Handbook and the Failure Modes, Effects and Diagnostic Analysis (FMEDA) training in support of this research.

References

[1] International Electrotechnical Commission. (2010). *Functional safety of electrical/electronic/programmable electronic safety-related systems - Part 1: General requirements* (IEC Standard No. 61508:2010). Retrieved from https://webstore.iec.ch/publication/5515

[2] International Electrotechnical Commission. (2016). *Functional safety - Safety instrumented systems for the process industry sector - Part 1: Framework, definitions, system, hardware and application programming requirements* (IEC Standard No. 61511:2016). Retrieved from https://webstore.iec.ch/publication/61289

[3] HSE. (2015). *A guide to the Control of Major Accident Hazards Regulations (COMAH) 2015 (L111)*. HSE. Retrieved from https://www.hse.gov.uk/pubns/priced/l111.pdf

[4] Stewart, L. (2019, September 22-26). *How Final Element Proof Test Can Affect Your SIF* [Conference Paper]. Proceedings of the 29th European Safety and Reliability Conference, Hannover, Germany. https://doi.org/10.3850/978-981-11-0745-0-0478-cd

[5] Abdelrhafour, M., Bajaj, N. and Boily, S. (2012, November 27-29). *Proof Test Procedure Effectiveness on Safety Instrumented Systems.*[Conference Paper]. 2012 Safety Control Systems Conference IDC Technologies. Edmonton, Canada. Retrieved from https://vdocuments.mx/idc-conference-2012-proof-test-procedure-effectiveness-on-safety-instrumented.html

[6] International Society of Automation. (2022). *In-Situ Proof Testing of Automated Valves* (ISA Standard No. ISA-TR96.05.02-2022). Retrieved from https://www.isa.org/products/isa-tr96-05-02-2022-in-situ-proof-testing-of-autom

[7] British Standards Institution. (2016). BS EN ISO 14224: Petroleum, petrochemical and natural gas industries. Collection and exchange of reliability and maintenance data for equipment. Retrieved from https://www.standardsuk.com/products/BS-EN-ISO-14224-2016

[8] Goble, W. M., and Brombacher, A. C. (1999). Using a failure modes, effects and diagnostic analysis (FMEDA) to measure diagnostic coverage in programmable electronic systems. *Reliability Engineering and System Safety (1999)*, Volume 66, Issue 2, November 1999,145-148. https://doi.org/10.1016/S0951-8320(99)00031-9

[9] Lundteigen, M. A., and Rausand, M. (2008). Partial stroke testing of process shutdown valves: How to determine the test coverage. *Journal of Loss Prevention in the Process Industries*, 21(6), 579–588. Retrieved from https://doi.org/10.1016/j.jlp.2008.04.007

[10] Dearden, H.T. (2020). *Functional Safety in Practice* (3rd ed.). CreateSpace.

[11] exida (2021). *Component Reliability Database (CRD) Handbook Volume 2 - Mechanical Components*. (5th Edition). Signature Book Printing.

[12] Isenburg, J. (2022). Mechanical components for functional safety applications (SIL): Standardisation urgently needed. *Industrial Valve Summit 2022 Conference*. Retrieved from https://www.valvecampus.com/abstract/mechanical-components-for-functional-safety-applications-sil-standardization-urgently-needed/

[13] Zadeh, L. A. (1968). Fuzzy Algorithms. *Information and Control*. Volume 12, Issue 2, February 1968, Pages 94-102. https://doi.org/10.1016/S0019-9958(68)90211-8

[14] Mamdani, E. H. (1974). Application of Fuzzy Algorithms for Control of Simple Dynamic Plant. *Proceedings of the Institution of Electrical Engineers*. 121 (12): 1585–1588. doi:10.1049/iee.1974.0328.

[15] exida (2023). *SIL Safe Data*. exida. Retrieved from http://silsafedata.com/

[16] SINTEF and NTNU (2015). *Offshore and Onshore Reliability Data Handbook (OREDA). Volume 1 - Topside Equipment*. 6th Edition.

[17] Technis. (2023). *FARADIP.FOUR Database*. Technis. Retrieved from https://www.technis.org.uk

[18] NAMUR. (2018). Flexible proof testing of field devices in safety instrumented systems (NA 106). NAMUR. Retrieved from https://infostore.saiglobal.com/en-us/standards/NAMUR-NA-106-2018-1135947_SAIG_NAMUR_NAMUR_2683920/

[19] Technis. (2020). *Guidelines on Assessing Proof Test Coverage* (T996). Technis. Retrieved from https://www.technis.org.uk/guidelines.html

[20] Ottermo, M., Hauge, S., and Håbrekke, S. (2021). *Reliability Data for Safety Instrumented Systems PDS Data Handbook*. SINTEF

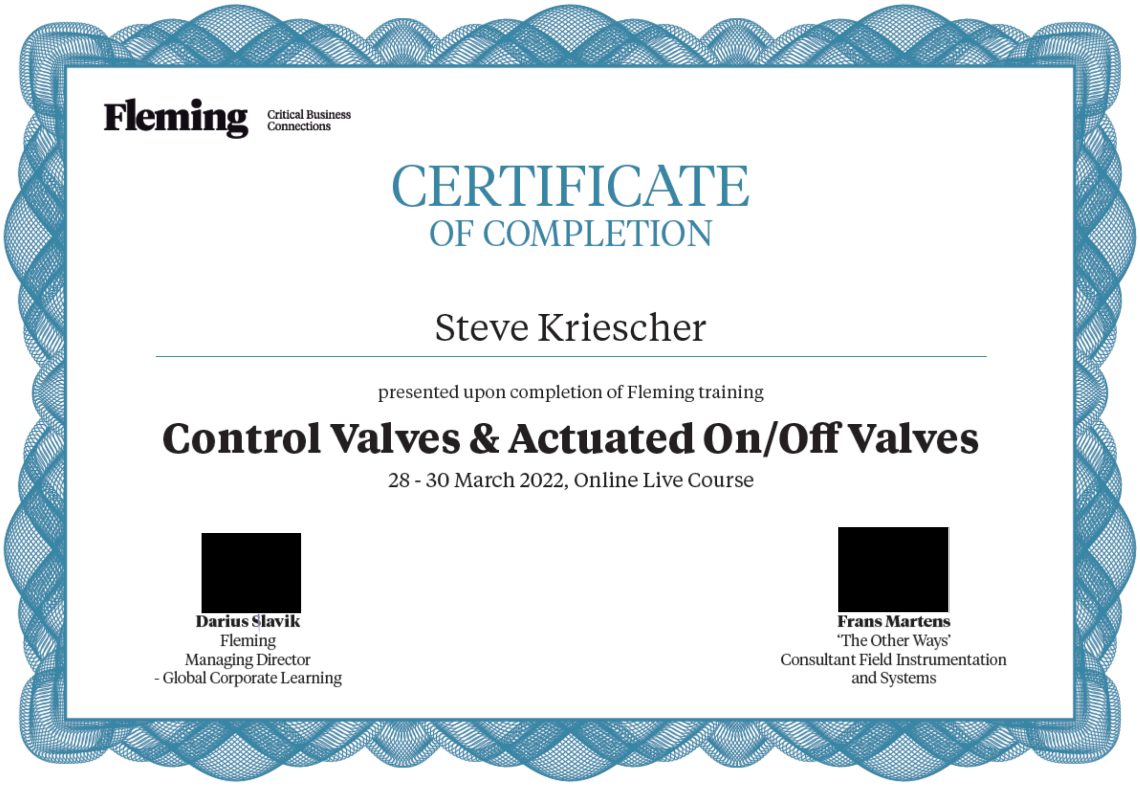
[21] International Society of Automation. (2022). *Safety Integrity Level (SIL) Verification of Safety Instrumented Functions* (ISA Standard No. ISA-84.00.02-2022). Retrieved from https://www.isa.org/products/isa-tr84-00-02-2022-safety-integrity-level-sil-ver

[22] Ali, O.A.M., Ali, A.Y. and Sumait, B.S. (2015). Comparison between the effects of different types of membership functions on fuzzy logic controller performance. *International Journal*, 76, pp.76-83. Retrieved from https://www.researchgate.net/publication/282506091_Comparison_between_the_Effects_of_Different_Types_of_Membership_Functions_on_Fuzzy_Logic_Controller_Performance

[23] Wagner, C., Pierfitt, M. and McCulloch, J. (2014). *Juzzy online: An online toolkit for the design, implementation, execution and sharing of Type-1 and Type-2 fuzzy logic systems.* IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), Beijing, China, pp. 2321-2328, doi: 10.1109/FUZZ-

Appendix G

Technical Training Course Certificates



Fleming Critical Business Connections

CERTIFICATE
OF COMPLETION

Steve Kriescher

presented upon completion of Fleming training

Control Valves & Actuated On/Off Valves

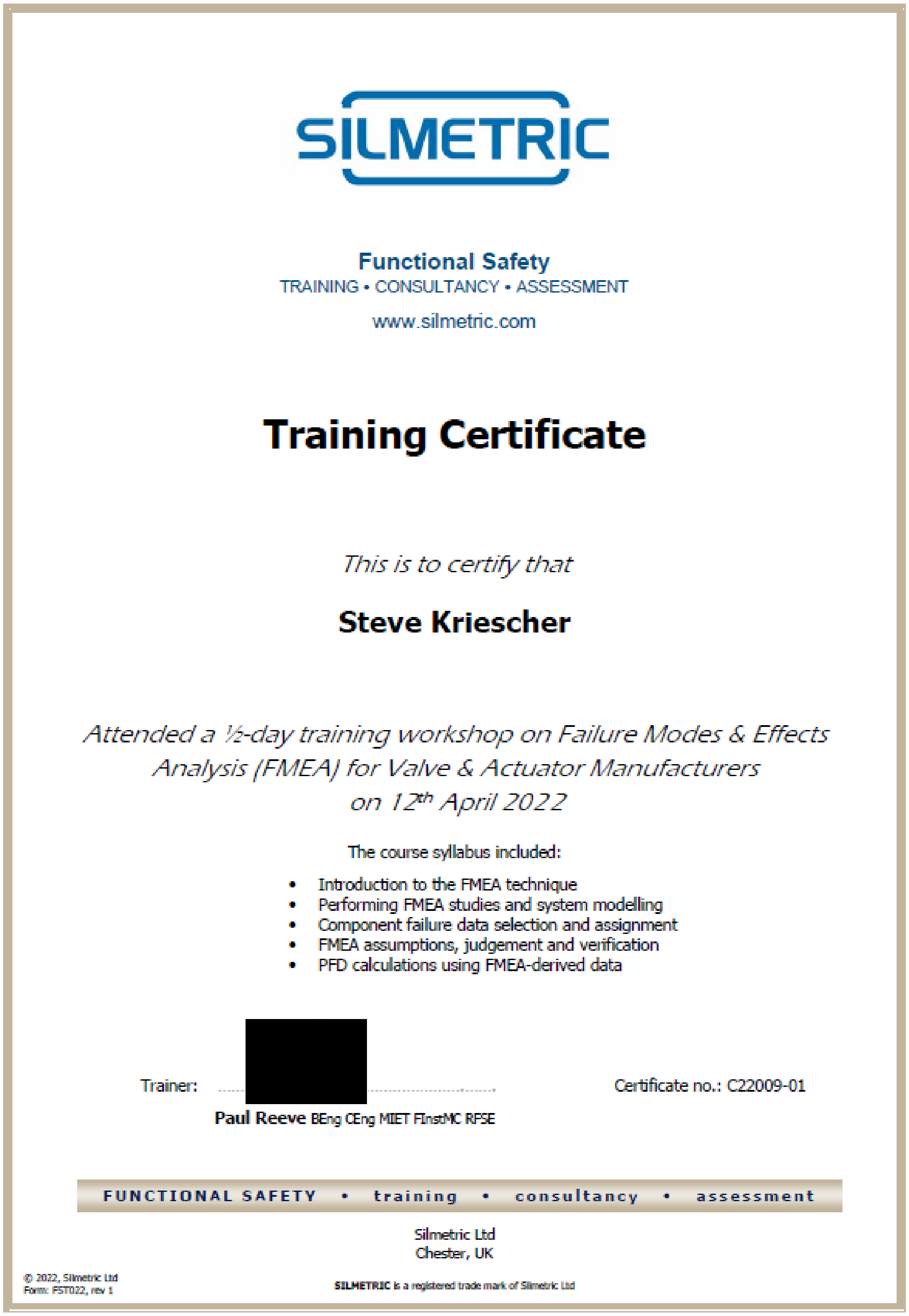
28 - 30 March 2022, Online Live Course

[redacted]

Darius Slavik
Fleming
Managing Director
- Global Corporate Learning

[redacted]

Frans Martens
'The Other Ways'
Consultant Field Instrumentation
and Systems



SILMETRIC

Functional Safety
TRAINING • CONSULTANCY • ASSESSMENT

www.silmetric.com

Training Certificate

This is to certify that

Steve Kriescher

Attended a ½-day training workshop on Failure Modes & Effects Analysis (FMEA) for Valve & Actuator Manufacturers on 12th April 2022

The course syllabus included:

- Introduction to the FMEA technique
- Performing FMEA studies and system modelling
- Component failure data selection and assignment
- FMEA assumptions, judgement and verification
- PFD calculations using FMEA-derived data

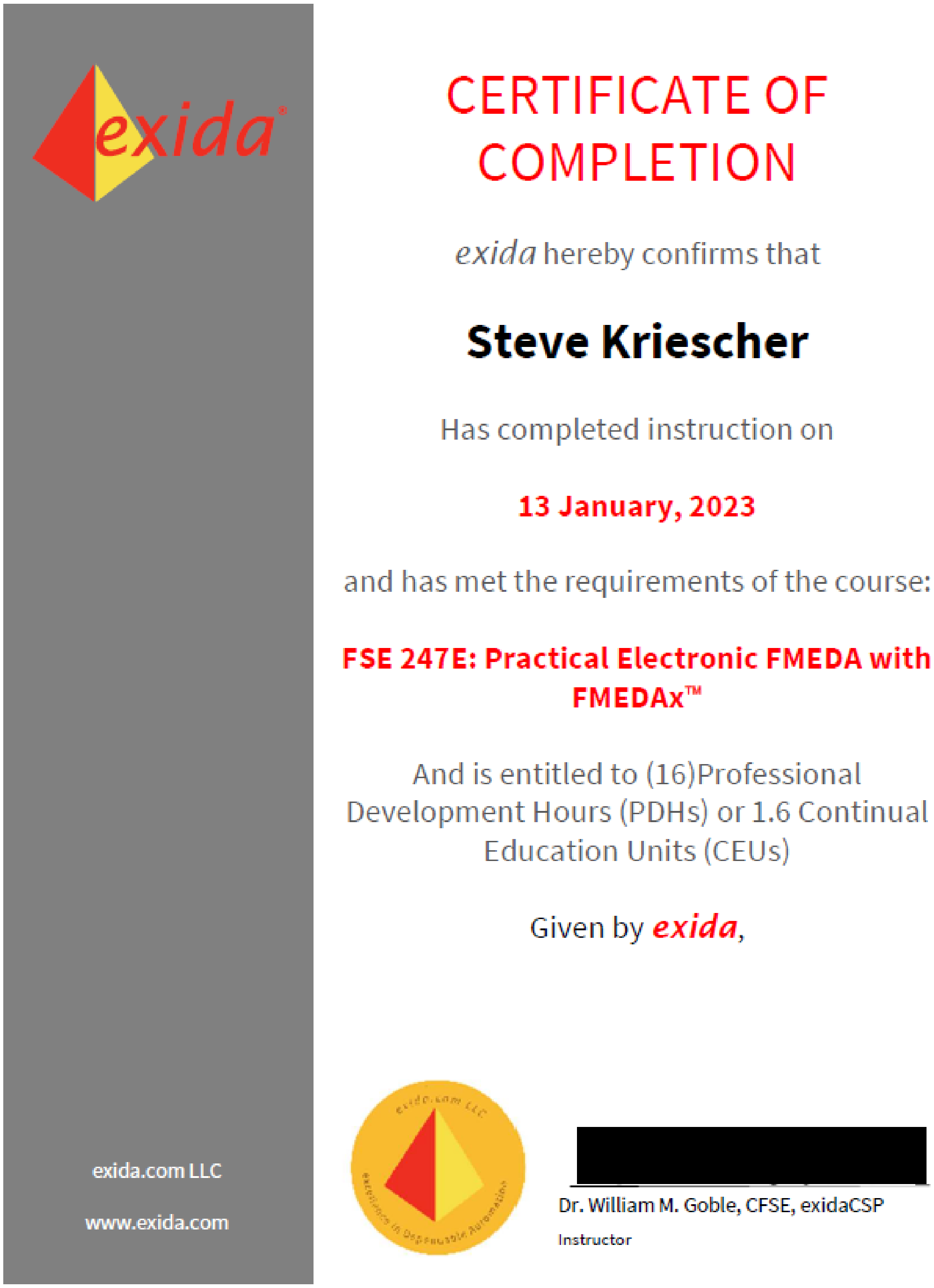
Trainer: [redacted]
Paul Reeve BEng CEng MIET FInstMC RFSE

Certificate no.: C22009-01

FUNCTIONAL SAFETY • training • consultancy • assessment

Silmetric Ltd
Chester, UK

Form: FST022, rev 1



exida

exida.com LLC

www.exida.com

CERTIFICATE OF COMPLETION

exida hereby confirms that

Steve Kriescher

Has completed instruction on

13 January, 2023

and has met the requirements of the course:

FSE 247E: Practical Electronic FMEDA with FMEDAx™

And is entitled to (16)Professional Development Hours (PDHs) or 1.6 Continual Education Units (CEUs)

Given by ***exida***,

exida.com LLC

excellence in Dependable Automation

[redacted]

Dr. William M. Goble, CFSE, exidaCSP

Instructor

References

[1] Okoh, P. and Hagen. Stein. Maintenance-related major accidents: Classification of causes and case study. *Journal of Loss Prevention in the Process Industries.* 26. 2013, p 1060-1070. http://dx.doi.org/10.1016/j.jlp.2013.04.002

[2] CCPS. *Centre for Chemical Process Safety, Guidelines for process safety metrics.* John Wiley & Sons, Hoboken, New Jersey, 2010.

[3] Tokarski. E. *The Safety Professional's Role: In Support of Industrial Facilities Operations and Maintenance.* Xlibris. 2013.

[4] U.S. Environmental Protection Agency. *Deepwater Horizon – BP Gulf of Mexico Oil Spill.* Accessed June 2024 https://www.epa.gov/enforcement/deepwater-horizon-bp-gulf-mexico-oil-spill

[5] Reuters. *BP Deepwater Horizon costs balloon to $65 billion.* Accessed June 2024 https://www.reuters.com/article/world/bp-deepwater-horizon-costs-balloon-to-65-billion-idUSKBN1F50O5/

[6] Gurung. K,, Jayadeep. L,, Siwek. J,, Vora. S and Zhou. D,. Texas City Refinery explosion — safety out of focus. *IChemE Loss Prevention Bulletin 275.* 2020. Accessed June 2024 Retrieved from http://dx.doi.org/10.1016/j.jlp.2013.04.002

[7] HSE. *The Explosion and Fires at the Texaco Refinery (ISBN 0717614131.* HSE. 1997. Retrieved from https://www.jesip.org.uk/wp-content/uploads/2022/03/Texaco-Refinery-Explosion.pdf

[8] Safety4Sea. *Learn from the past: Deepwater Horizon oil spill* Accessed June 2024 https://safety4sea.com/cm-learn-from-the-past-deepwater-horizon-oil-spill/

[9] HSE. *A guide to the Control of Major Accident Hazards Regulations (COMAH) 2015.* Series code L111. HSE. 2005. Retrieved from https://www.hse.gov.uk/pubns/priced/l111.pdf

[10] HSE. *The Offshore Installations (Offshore Safety Directive) (Safety Case etc) Regulations 2015.* Series code L154. HSE. 2015. Retrieved from https://www.hse.gov.uk/pubns/priced/l154.pdf

[11] HSE. *Introduction to the Seveso III Directive.* 2012. https://www.hse.gov.uk/seveso/introduction.htm Accessed: October 2024

[12] Occupational Safety and Health Administration. 1992. *29 CFR 1910.119 – Process Safety Management of Highly Hazardous Chemicals.* U.S. Department of Labor; 1992. Available from: https://www.osha.gov/laws-regs/regulations/standardnumber/1910/1910.119

[13] Occupational Safety and Health Administration. 2000. *Process Safety Management: Guidelines for Compliance (OSHA 3133).* U.S. Department of Labor. https://www.osha.gov/sites/default/files/publications/osha3133.pdf

[14] HSE. *Reducing risks, protecting people.* Series code R2P2. HSE. 2001. Retrived from https://www.hse.gov.uk/enforce/assets/docs/r2p2.pdf

[15] Baybutt, P. The ALARP principle in process safety. *Proc. Safety Prog.* 2014, 33: 36-40. https://doi.org/10.1002/prs.11599

[16] Timms, C. R. IEC 61511/an aid to COMAH and safety case regulations compliance. *Measurement and Control.* 2004, 37(4), 115-122.

[17] International Electrotechnical Commission. *Functional safety - Safety instrumented systems for the process industry sector - Part 1: Framework, definitions, system, hardware and application programming requirements* (IEC Standard No. 61511:2016). Retrieved from https://webstore.iec.ch/publication/61289

[18] Willey, R. J. *Layer of Protection Analysis*, Procedia Engineering, Volume 84, 2014, Pages 12-22,ISSN 1877-7058, https://doi.org/10.1016/j.proeng.2014.10.405.

[19] Stauffer. T., Sands. N. P., Strobhar. D. *Plugging the Holes in the Swiss Cheese Model* [Conference Paper]. Proceedings of the AIChE 2017 Spring Meeting and 13th Global Congress on Process Safety, San Antonio, TX. 2017. Retrieved from https://www.aiche.org/resources/publications/cep/2017/september/plug-holes-swiss-cheese-model

[20] BS ISO. *Asset management - Overview, principles and terminology* (BS ISO Standard No. 55000:2014). Retrieved from https://knowledge.bsigroup.com/products/asset-management-overview-principles-and-terminology?version=standard

[21] Niu, G., Yang, B., Pecht. M. Development of an optimized condition-based maintenance system by data fusion and reliability-centered maintenance. *Reliability Engineering and System Safety (2010)*. 2010. https://doi:10.1016/j.ress.2010.02.016

[22] Thomas. R. and Rees. D. J. Progress in Predictive Asset Maintenance Management. *International Journal of Condition Monitoring and Diagnostic Engineering Management (COMDEM)*. 2021. ISSN 1363 – 7681

[23] Reliableplant. *Condition-based Maintenance: A Complete Guide* Accessed June 2024 https://www.reliableplant.com/condition-based-maintenance-31823

[24] LLoyds Register. *Reducing the maintenance burden across oil and gas operations* Accessed June 2024 https://www.opuskinetic.com/wp-content/uploads/2019/06/36.-Lloyds-Register-Reducing-the-maintenance-burden-across-oil-and-gas-operations.pdf

[25] Green, D. and Bell, R. *Proof Testing... A key performance indicator for designers and end users of Safety Instrumented Systems* [Conference Paper]. Symposium Series No 162, Hazards 27, Birmingham, UK. 2017. Retrieved from https://www.icheme.org/media/15519/paper-67.pdf

[26] OREDA. *Offshore and Onshore Reliability Data (Volume 1)*. SINTEF & NTNU. 2015. https://www.oreda.com/handbook

[27] Ottermo, M., Hauge, S., and Håbrekke, S. *Reliability Data for Safety Instrumented Systems PDS Data Handbook*. SINTEF. 2021.

[28] exida. *Component Reliability Database (CRD) Handbook Volume 2 - Mechanical Components*. (5th Edition). Signature Book Printing. 2021.

[29] Technis. *FARADIP.FOUR Database*. Technis. 2023. Retrieved from https://www.technis.org.uk

[30] Cameron. K., Lewis. A., Montalvão. D., Reza Herfatmanesh. M. In-service performance of emergency shutdown valves and dependent operational relationships in the offshore oil and gas industry. *Petroleum*. Volume 9, Issue 4, 2023, p 613-620. https://doi.org/10.1016/j.petlm.2023.06.004.

[31] International Electrotechnical Commission. *Failure modes and effects analysis (FMEA and FMECA)* (IEC Standard No. 60812:2018). Retrieved from https://webstore.iec.ch/publication/26359

[32] International Electrotechnical Commission. *Functional safety of electrical/electronic/programmable electronic safety-related systems - Part 1: General requirements* (IEC Standard No. 61508:2010). Retrieved from https://webstore.iec.ch/publication/5515

[33] Summers, A. E. IEC 61511 and the capital project process — A protective management system approach. *Journal of Hazardous Materials.* 2006, Volume 130, Issues 1–2, Pages 28-32. https://doi.org/10.1016/j.jhazmat.2005.07.063.

[34] International Electrotechnical Commission. *Hazard and operability studies (HAZOP studies) - Application guide* (IEC Standard No. 61882:2016).

[35] Wang J., Hu D., Peng C., Zhi H., Wu L. Safety assessment through HAZOP-LOPA-SIL analysis implementation in the DPA demulsifier production process. *Process Saf Prog.* 2023, 42(1): 38-47. doi:10.1002/prs.12414

[36] Baum, D., Faulk, N., & Perez, P. J. Improved integration of LOPA with HAZOP analyses. *Process Safety Progress.* 2009, 28(4), 308-311. https://doi.org/10.1002/prs.10350

[37] Babouri, K., & Bendib, R. Assessment of Safety Integrity Requirements for Fired Heater System in Accordance with IEC 61508. *International Journal of Latest Engineering Science (IJLES).* 2(5). 2019.

[38] Jayanthi Vijay Sarathy. *Understanding High Integrity Pressure Protection Systems.* Engineering Practice Magazine. 2018. https://www.researchgate.net/publication/346419109_Engineering_Practice_Magazine_July_2018

[39] Park, J. H. The Improvement of SIL Calculation Methodology. *International Conference on Probabilistic Safety Assessment and Management.* 2016, 11-15.

[40] Kallambettu, J., & Viswanathan, V. Application of functional safety to electrical power equipment and systems in process industries.*Journal of Loss Prevention in the Process Industries.* 2018, 56, 155-161.

[41] Smith, D. J. *Reliability, maintainability and risk: practical methods for engineers.* Butterworth-Heinemann. 2021.

[42] Instrumentation Tools. *What is a Shutdown Valve?* Accessed June 2024 https://instrumentationtools.com/shutdown-valve

[43] RS. *MAXSEAL IC04 SERIES* Accessed June 2024 https://us.rs-online.com/product/norgren/y413aa1h1bs/72054748/

[44] Rotork. *Fluid Power Actuators Explained* Accessed June 2024 https://www.rotork.com/uploads/documents-versions/24733/1/pub010-024-00_0516.pdf

[45] American Petroleum Institute. *Specification for Pipeline Valves - API 6D.* 2008. Retrieved from https://law.resource.org/pub/us/cfr/ibr/002/api.6d.2008.pdf

[46] The Process Piping. *Introduction to Ball Valves* Accessed June 2024 https://www.theprocesspiping.com/introduction-to-ball-valve/

[47] Stewart, L. *How Final Element Proof Test Can Affect Your SIF* [Conference Paper]. Proceedings of the 29th European Safety and Reliability Conference, Hannover, Germany. 2019. https://doi.org/10.3850/978-981-11-0745-0-0478-cd

[48] NAMUR. Flexible proof testing of field devices in safety instrumented systems (NA 106). NAMUR. 2018. Retrieved from https://infostore.saiglobal.com/en-us/standards/NAMUR-NA-106-2018-1135947_SAIG_NAMUR_NAMUR_2683920/

[49] Dearden, H. T. *Functional Safety in Practice* (3rd ed.). CreateSpace. 2020.

[50] Wood, J. What is Good Practice for the Proof Testing of Safety Instrumented Systems of Low Safety Integrity? [Conference Paper]. *IChemE - Symposium Series No 159, Hazards 24* Edinburgh, Scotland. 2014. https://www.icheme.org/media/8926/xxiv-paper-31.pdf

[51] Brissaud, F., Barros, A., & Bérenguer, C. Probability of failure on demand of safety systems: Impact of partial test distribution. *Proceedings of the Institution of Mechanical Engineers: Part O: Journal of Risk and Reliability.* 2012, 226(4), p 426–436. https://doi.org/10.1177/1748006X12448142

[52] Gruhn. P., Pittman. J., Susan Wiley. S. and LeBlanc. T. *ISA Transactions.* Quantifying the impact of partial stroke valve testing of safety instrumented systems. Volume 37, Issue 2, 1998, p 87-94 https://doi.org/10.1016/S0019-0578(98)00009-3

[53] Doddi. S. Understand partial-stroke testing. Control Engineering Journal. 2020. https://www.controleng.com/articles/understand-partial-stroke-testing/

[54] Safety and SIS. *Partial Valve Stroke Test (PVST).* 2020. https://safetyandsis.com/partial-valve-stroke-test/

[55] Lundteigen, M. A., & Rausand, M. The effect of partial stroke testing on the reliability of safety valves. *ESREL'07.* 2007.

[56] Angela E. Summers, A., E. *Partial Stroke Testing of Block Valves*, Instrument Engineers Handbook, Volume 4, Chapter 6.9, 2006.

[57] Sato, Y. Introduction to partial stroke testing. *2008 SICE Annual Conference.* Chofu, Japan, 2008, p 2754-2758. doi: 10.1109/SICE.2008.4655133.

[58] International Society of Automation. *Partial Stroke Testing of Automated Valves* (ISA Standard No. ISA-TR96.05.01-2017). Retrieved from https://www.isa.org/products/isa-tr96-05-01-2017-partial-stroke-testing-of-auto

[59] Gubert, A., Basco, J., & Serral, D. Simplified vs rigorous techniques for safety integrity level verification in safety instrumented systems including proof test coverage [Conference Paper]. *7th European Meeting on Chemical Industry and Environment.* Tarragona, Spain. 2015. https://www.researchgate.net/publication/237842588

[60] exida. *The Key Variables Needed for PFD_{avg} Calculation.* [White paper]. exida. 2018. https://www.exida.com.sg/wp-content/uploads/2022/04/Key-Variables-Needed-for-PFDavg-Calculation.pdf. Retrieved September 2024

[61] Nunns, S. Health and Safety Executive & ABB Ltd. *Principles for proof testing of safety instrumented systems in the chemical industry.* HSE Books. 2002.

[62] Hokstad, P., Flotten, P, Holmstrom, S., Mckenna, F. & Onshus, T. A reliability model for optimization of test schemes for fire and gas detectors. *In Reliability Engineering and System Safety.* 1995, (Vol. 47), 15-25. https://doi.org/10.1016/0951-8320(94)00038-P

[63] Abonyi, J., Baradits, G. & Madár, J. Novel Model of Proof Test Coverage Factor. *Proceedings of the 10th International Symposium of Hungarian Researchers on Computational Intelligence and Informatics*. Hungary, 2009, p 209-301. https://www.researchgate.net/publication/237842588_Novel_Model_of_Proof_Test_Coverage_Factor

[64] Rielly, R. Assuring the Probability of Failure on Demand of a Safety Instrumented System without Full Proof Testing. *Measurement and Control*. 2016, Vol. 49(9), p 279–285. https://doi.org/10.1177/0020294016663975

[65] Jin, J., Pang, L., Hu, B., & Wang, X. Impact of proof test interval and coverage on probability of failure of safety instrumented function. *Annals of Nuclear Energy*. 2016, 87, 537–540. https://doi.org/10.1016/j.anucene.2015.09.028

[66] Hauge, S., Kråkenes, T., Hokstad, P., Håbrekke, S., & Jin, H. *Prediction Method for Safety Instrumented Systems PDS Method Handbook*. SINTEF. 2013.

[67] Lundteigen, M. A., and Rausand, M. Partial stroke testing of process shutdown valves: How to determine the test coverage. *Journal of Loss Prevention in the Process Industries*, 21(6), 2008, p 579–588. Retrieved from https://doi.org/10.1016/j.jlp.2008.04.007

[68] Jin, H., Lundteigen, M. A., & Rausand, M. Uncertainty assessment of Reliability estimates for safety-instrumented systems. *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*. 2012, 226(6), 646–655. https://doi.org/10.1177/1748006X12462780

[69] Smith, D.J. *Reliability, Maintainability and Risk*. 10th ed. Butterworth-Heinemann. 2021. ISBN-10 0323912613.

[70] Goble, W. M., and Brombacher, A. C. Using a failure modes, effects and diagnostic analysis (FMEDA) to measure diagnostic coverage in programmable electronic systems. *Reliability Engineering and System Safety (1999)*, Volume 66, Issue 2, 1999, p 145-148. https://doi.org/10.1016/S0951-8320(99)00031-9

[71] Easton, B. J. Impact of Imperfect Proof Testing on the Performance of Safety Instrumented Functions. *Proceedings of the 31st European Safety and Reliability Conference*. 2021, p 744-751. Retrieved from https://doi.org/10.3850/978-981-18-2016-8

[72] Bukowski J.V. and van Beurden, I. Impact of proof test effectiveness on safety instrumented system performance. *Proceedings of the 2009 Annual Reliability and Maintainability Symposium*, 2009, p 157-163, Retrieved from https://doi.org/10.1109/RAMS.2009.4914668.

[73] O'Brien, C., Stewart, L., Bredemeyer, L. *Final Elements in Safety Instrumented Systems*. (1st ed.). Signature Book Printing. 2018. ISBN-13: 978-193497718-7.

[74] Stewart. L. *Evaluating and Proving SIS Safety Levels*. The valve manufacturers association of America. 2014. https://www.valvemagazine.com/articles/evaluating-and-proving-sis-safety-levels

[75] Abdelrhafour, M., Bajaj, N. and Boily, S. *Proof Test Procedure Effectiveness on Safety Instrumented Systems*.[Conference Paper]. 2012 Safety Control Systems Conference IDC Technologies. Edmonton, Canada. 2012. Retrieved from https://vdocuments.mx/idc-conference-2012-proof-test-procedure-effectiveness-on-safety-instrumented.html

[76] Technis. (2020). *Guidelines on Assessing Proof Test Coverage* (T996). Technis. Retrieved from https://www.technis.org.uk/guidelines.html

[77] Isenburg, J. Mechanical components for functional safety applications (SIL): Standardisation urgently needed. *Industrial Valve Summit 2022 Conference*. 2022. Retrieved from https://www.valvecampus.com/abstract/mechanical-components-for-functional-safety-applications-sil-standardization-urgently-needed/

[78] Exida. *SIL Safe Data*. Exida. 2023. Retrieved from http://silsafedata.com/

[79] International Society of Automation. *In-Situ Proof Testing of Automated Valves* (ISA Standard No. ISA-TR96.05.02-2022). Retrieved from https://www.isa.org/products/isa-tr96-05-02-2022-in-situ-proof-testing-of-autom

[80] International Society of Automation. *Safety Integrity Level (SIL) Verification of Safety Instrumented Functions* (ISA Standard No. ISA-84.00.02-2022). Retrieved from https://www.isa.org/products/isa-tr84-00-02-2022-safety-integrity-level-sil-ver

[81] ANSI/FCI 70-2. *Standard for Control Valve Seat Leakage Testing*. 2021. https://www.intertekinform.com/en-gb/standards/ansi-fci-70-2-2021-509586_saig_fci_fci_2918128/

[82] Pate´-Cornell. M. Uncertainties in Risk Analysis: Six Levels of Treatment. *Journal Reliability Eng. and System Safety*. Vol. 54, 1996, p 95-111.

[83] Salahdine. F., Kaabouch. N., Hassan El Ghazi. H. Techniques for Dealing with Uncertainty in Cognitive Radio Networks. *IEEE 7th annual computing and communication workshop and conference (CCWC)*. Piscataway: IEEE. 2017, p 1–6.

[84] Yiping Li, Jianwen Chen, and Ling Feng. Dealing with Uncertainty: A Survey of Theories and Practices. *IEEE Transactions on Knowledge and Data Engineering*. 2013, Vol. 25, NO. 11.

[85] Luengo, D., Martino, L., Bugallo, M. et al. A survey of Monte Carlo methods for parameter estimation. *EURASIP J. Adv. Signal Process*. 2020. https://doi.org/10.1186/s13634-020-00675-6

[86] Van de Schoot, R., Depaoli, S., King, R. et al. Bayesian statistics and modelling. *Nat Rev Methods Primers 1*. 2021. https://doi.org/10.1038/s43586-020-00001-2

[87] Zadeh, L. A. Fuzzy sets. *Information and Control*. Volume 8, Issue 3, 1965, p 338-353, https://doi.org/10.1016/S0019-9958(65)90241-X.

[88] Sotoodeh, K. *The importance of maximum allowable stem torque in valves*. SN Appl. Sci. 1, 433, 2019. https://doi.org/10.1007/s42452-019-0445-0

[89] Kabir, S. and Papadopoulos, Y. A review of applications of fuzzy sets to safety and reliability engineering. *International Journal of Approximate Reasoning*. Volume 100, 2018, Pages 29-55. https://doi.org/10.1016/j.ijar.2018.05.005.

[90] Dempster, A. P. Upper and Lower Probabilities Induced by a Multivalued Mapping. *The Annals of Statistics 28*. 1967, p 325-339.

[91] Shafer, G. *A Mathematical Theory of Evidence* Vol. 42, 1976. Princeton, NJ, Princeton University Press.

[92] Sentz, K. and Ferson, S. Combination of Evidence in Dempster-Shafer Theory. Sandia National Laboratories. 2002. https://doi.org/10.2172/800792

[93] Destercke. S., and Dubois. D. and Chojnacki. E. Possibilistic information fusion using maximal coherent subsets. *IEEE Transactions on Fuzzy Systems.* 2009. DOI: 10.1109/TFUZZ.2008.2005731 Source: IEEE Xplore

[94] Dubois, D., Lang, J., Prade, H. A brief overview of possibilistic logic. In: Kruse, R., Siegel, P. (eds) Symbolic and Quantitative Approaches to Uncertainty. ECSQARU 1991. *Lecture Notes in Computer Science.* vol 548. 1991. Springer, Berlin, Heidelberg. https://doi.org/10.1007/3-540-54659-6_65

[95] Božić, V. *Fuzzy Approach to Risk Management: Enhancing Decision-Making Under Uncertainty.* 2023. DOI: 10.13140/RG.2.2.13517.82405

[96] Sharma, R. K., Kumar, D. and Kumar, P. Systematic failure mode effect analysis (FMEA) using fuzzy linguistic modelling. *International Journal of Quality & Reliability Management.* Vol. 22 No. 9, 2005, p 986-1004. DOI 10.1108/02656710510625248

[97] Baraldi, P., Podofillini, L., Mkrtchyan, L., Zio, E. and Dang. V.N. Comparing the treatment of uncertainty in Bayesian networks and fuzzy expert systems used for a human reliability analysis application. *Reliability Engineering & System Safety.* Volume 138, 2015, p 176-193. https://doi.org/10.1016/j.ress.2015.01.016

[98] Cherkassky, V. Fuzzy Inference Systems: A Critical Review. *Computational Intelligence: Soft Computing and Fuzzy-Neuro Integration with Applications. NATO ASI Series.* vol 162. 1998. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-58930-0_10

[99] Smets, P. "What is Dempster-Shafer's model." *Advances in the Dempster-Shafer theory of evidence (1994).* 1994, p 5-34.

[100] Kulkarni. N. S. and Johnson. A. R. Dempster-Shafer Theory Approach to FMEA. *Proceedings of the 2012 Industrial and Systems Engineering Research Conference.* 2012.

[101] Thomas, S., L. Hierarchical Bayesian Approach to IEC 61511 Prior Use. *American Institute of Chemical Engineers 2018 and 14th Global Congress on Process Safety.* 2018.

[102] Watkar V.B. Fuzzy Inference Systems: Types & Applications. Journal of Emerging Technologies and Innovative Research (JETIR). 2022.

[103] Zhang, G., Patuwo, B. E., & Hu, M. Y. Forecasting with artificial neural networks: The state of the art. *International Journal of Forecasting*, 1998. 14(1), 35–62. https://doi.org/10.1016/S0169-2070(97)00044-7

[104] Dietterich, T. G. Ensemble methods in machine learning. *In Multiple Classifier Systems.* 2000. (pp. 1–15). Springer. https://doi.org/10.1007/3-540-45014-9_1

[105] Kaelbling, L. P., Littman, M. L., & Moore, A. W. Reinforcement learning: A survey. Journal of Artificial Intelligence Research, 1996, 4, 237–285. https://doi.org/10.1613/jair.301

[106] Deng, J. Control problems of grey systems. Systems & Control Letters, 1982, 1(5), 288–294. https://doi.org/10.1016/S0167-6911(82)80025-X

[107] Lin, Y., and Liu, S. A historical introduction to grey systems theory, *2004 IEEE International Conference on Systems*, Man and Cybernetics (IEEE Cat. No.04CH37583), 2004, pp. 2403-2408 vol.3, doi: 10.1109/ICSMC.2004.1400689.

[108] Tao, F., Zhang, H., Liu, A., & Nee, A. Y. C. (2019). Digital twin in industry: State-of-the-art. *IEEE Transactions on Industrial Informatics*, 15(4), 2405–2415. https://doi.org/10.1109/TII.2018.2873186

[109] Mamdani, E. H. Application of Fuzzy Algorithms for Control of Simple Dynamic Plant. *Proceedings of the Institution of Electrical Engineers.* 121 (12): 1974, p 1585–1588. doi:10.1049/iee.1974.0328.

[110] Takagi, T., and Sugeno, M. Fuzzy identification of Systems and its Applications to Modeling and Control. *Systems, Man and Cybernetics.* IEEE Transactions on, SMC-15(1):116–132, 1985. doi: 10.1109/TSMC.1985.6313399

[111] Soteris A. Kalogirou. *Solar Energy Engineering* (Second Edition). Academic Press. 2014, p 583-699, ISBN 9780123972705, https://doi.org/10.1016/B978-0-12-397270-5.00011-X.

[112] Mendel, J., Hagras, H., Tan, W. W., Melek, W. W. and Ying, H. Introduction to type-2 fuzzy logic control: theory and applications. John Wiley & Sons. 2014.

[113] Omar Adil M. Ali, Aous Y. Ali, Balasem Salem Sumait. Comparison between the Effects of Different Types of Membership Functions on Fuzzy Logic Controller Performance. *International Journal of Emerging Engineering Research and Technology.* Volume 3, Issue 3, 2015, p 76-83

[114] Naseer Sabri, S. A. Aljunid, M. S. Salim, R. B. Badlishah, R. Kamaruddin, M. F. Abd Malek. Fuzzy Inference System: Short Review and Design. *International Review of Automatic Control (I.RE.A.CO.)* Vol. 6, N. 4. 2013. ISSN 1974-6059

[115] Safavi, H. R., Alijanian, M. A., & Golmohammadi, M. H. Consideration of climate conditions in reservoir operation using fuzzy inference system (FIS). *British Journal of Environment and Climate Change.* 3(3), 2013, p 444-463.

[116] Guillaume, S. Designing fuzzy inference systems from data: an interpretability-oriented review. *IEEE Transactions on Fuzzy Systems.* 2001, p 426-443. 10.1109/91.928739

[117] Ye. B. and Z. Wei. Efficiency increment of CFD modeling by using ANFIS artificial intelligence for thermal-based separation modeling. *Case Studies in Thermal Engineering.* Volume 60. 2024. https://doi.org/10.1016/j.csite.2024.104820

[118] Tatachar, A. V. Comparative assessment of regression models based on model evaluation metrics. *International Journal of Innovative Technology and Exploring Engineering.* 8(9), 2021, p 853-860.

[119] Chicco, D., Warrens, M. J., & Jurman, G. The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. *Peerj computer science.* 7, 2021.

[120] Tan W.W., Chua T.W. Uncertain rule-based fuzzy logic systems: introduction and new directions (Mendel, JM; 2001)[book review]. *IEEE Computational intelligence magazine.* 2007 Feb;2(1):72-3.

[121] Kim Y.J. Monte Carlo vs. fuzzy Monte Carlo simulation for uncertainty and global sensitivity analysis. *Sustainability.* 2017 Mar 31;9(4):539.

[122] Dubois D, Prade H. The three semantics of fuzzy sets. *Fuzzy sets and systems.* 1997 Sep 1;90(2):141-50.

[123] Gacto MJ, Alcalá R, Herrera F. Interpretability of linguistic fuzzy rule-based systems: An overview of interpretability measures. *Information Sciences*. 2011 Oct 15;181(20):4340-60.

[124] Li, Y., Li, D., Pedrycz, W. and Wu, J. An approach to measure the robustness of fuzzy reasoning. *Int. J. Intell. Syst.*, 2005. 20: 393-413. https://doi.org/10.1002/int.20072

[125] Javidan, M. M., & Kim, J. Fuzzy Sensitivity Analysis of Structural Performance. *Sustainability*, 2022, 14(19), 11974. https://doi.org/10.3390/su141911974

[126] José Roberto Ribas, Juliana Crenitte Ribas Severo, Luciana Fernandes Guimarães, Kim Parente Currlin Perpetuo. A fuzzy FMEA assessment of hydroelectric earth dam failure modes: A case study in Central Brazil. *Energy Reports 7*. 2021, p 4412–4424

[127] Kwai-Sang Chin & Allen Chan & Jian-Bo Yang. Development of a Fuzzy FMEA based product design system. *The International Journal of Advanced Manufacturing Technology*. 36: 2008, p 633–649. DOI 10.1007/s00170-006-0898-3

[128] José Jovani Cardiel-Ortega and Roberto Baeza-Serrato. Failure Mode and Effect Analysis with a Fuzzy Logic Approach. *Systems*. 2023. https://doi.org/10.3390/systems11070348

[129] Shamsu Hassan, Jin Wang, Christos Kontovas, Musa Bashir. Modified FMEA hazard identification for cross-country petroleum pipeline using Fuzzy Rule Base and approximate reasoning. *Journal of Loss Prevention in the Process Industries 74*. 2022.

[130] G. Gupta and R. P. Mishra. A Failure Mode Effect and Criticality Analysis of Conventional Milling Machine Using Fuzzy Logic: Case Study of RCM. *Quality and Reliability Engineering*. 2016. DOI: 10.1002/qre.2011

[131] P. A. A. Garcia, R. Schirru & P. F. Frutuoso E Melo. A fuzzy data envelopment analysis approach for FMEA. *Progress in Nuclear Energy*. Vol. 46, No. 3-4, 2005, p 359-373.

[132] Matteo Mario Savino, Alessandro Brun, Carlo Riccio. Integrated system for maintenance and safety management through FMECA principles and fuzzy inference engine. *European Journal Industrial Engineering*. Vol. 5, No. 2, 2011.

[133] K Xu, L.C Tang, M Xie, S.L Ho, M.L Zhu. Fuzzy assessment of FMEA for engine systems. *Reliability Engineering & System Safety* Volume 75, Issue 1, 2002, p 17-29.

[134] Sivaprakasam Rajakarunakaran, A. Maniram Kumar, V. Arumuga Prabhu. Applications of fuzzy faulty tree analysis and expert elicitation for evaluation of risks in LPG refuelling station. *Journal of Loss Prevention in the Process Industries* 33, 2015.

[135] Refaul Ferdous, Faisal Khan,, Rehan Sadiq, Paul Amyotte, and Brian Veitch. Fault and Event Tree Analyses for Process Systems Risk Analysis: Uncertainty Handling Formulations. *Risk Analysis* Vol. 31, No. 1, 2011. DOI: 10.1111/j.1539-6924.2010.01475.x

[136] Shuen-Tai Ung and Wei-Min Shen. A Novel Human Error Probability Assessment Using Fuzzy Modelling. *Risk Analysis* Vol. 31, No. 5, 2011. DOI: 10.1111/j.1539-6924.2010.01536.x

[137] John B. Bowles & C. Enrique Peldez. Fuzzy logic prioritization of failures in a system failure mode, effects and criticality analysis. *Reliability Engineering and System Safety*. 50, 1995, p 203-213.

[138] Jakkula Balaraju, Mandela Govinda Raj, Chivukula Suryanarayana Murthy. Fuzzy-FMEA risk evaluation approach for LHD machine – A case study. *Journal of Sustainable Mining* 18, 2019, p 257–268.

[139] Joel John George, V.R. Renjith, Priscilla George, Amal S. George. Application of fuzzy failure mode effect and criticality analysis on unloading facility of LNG terminal. *Journal of Loss Prevention in the Process Industries.* 61, 2019, p 104–113

[140] Grzegorz Filo, Joanna Fabiś-Domagała, Mariusz Domagała, Edward Lisowski, and Hassan Momeni. The idea of fuzzy logic usage in a sheet-based FMEA analysis of mechanical systems*MATEC Web of Conferences* 183, 03009. 2018. https://doi.org/10.1051/matecconf/201818303009

[141] V.R. Renjith, Manoj Jose Kalathil, P. Haresh Kumar, Dilip Madhavan. Fuzzy FMECA (failure mode effect and criticality analysis) of LNG storage facility. *Journal of Loss Prevention in the Process Industries.* 56, 2018, p 537–547.

[142] Earl Cox. *The Fuzzy Systems Handbook. A Practitioners Guide to Building, Using, and Maintaining Fuzzy Systems.* Academic Press Ltd. 1994. ISBN 0-12-194270-8

[143] British Standards Institution. *Petroleum, petrochemical and natural gas industries. Collection and exchange of reliability and maintenance data for equipment.* (BS EN ISO 14224: 2016). https://www.standardsuk.com/products/BS-EN-ISO-14224-2016

[144] British Standards Institution. *Petroleum, petrochemical and natural gas industries. Collection and exchange of reliability and maintenance data for equipment.* (BS ISO 14224: 1999). https://knowledge.bsigroup.com/products/petroleum-and-natural-gas-industries collection-and-exchange-of-reliability-and-maintenance-data-for-equipment/standard

[145] Goff, R. J. Learning from the Causes of Failures of Offshore Riser Emergency Shutdown Valves. *Symposium Series No 161*, Hazards 26, 2016. Retrieved from https://www.icheme.org/media/11788/hazards-26-paper-51-learning-from-the-causes-of-failures-of-offshore-riser-emergency-shutdown-valves.pdf

[146] HSE. *A guide to the Pipelines Safety Regulations 1996 (L82).* HSE. 1996. Retrieved from https://www.hse.gov.uk/pubns/priced/l82.pdf

[147] Goff, R., Kay, J. Investigations into the immediate and underlying causes of failures of offshore riser emergency shutdown valves. HSE Books, 2015. (RR1072) http://www.hse.gov.uk/research/rrhtm/rr1072.htm

[148] Stanton, N. A., Salmon, P. M., Rafferty, L. A., Walker, G. H., Baber, C., & Jenkins, D. P. *Human Factors Methods: A Practical Guide for Engineering and Design (2nd ed.).* CRC Press. 2013. https://doi.org/10.1201/9781315587394

[149] Gould, K. S., Ringstad, A. J., & van de Merwe, K. *Human Reliability Analysis in Major Accident Risk Analyses in the Norwegian Petroleum Industry.* Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 56(1), 2012, 2016-2020. https://doi.org/10.1177/1071181312561421

[150] British Standards. *Industrial valves. Functional safety of safety-related automated valves* (BS EN standards no. 17955:2024). Retrieved from https://bsol.bsigroup.com/

[151] Wagner, C., Pierfitt, M. and McCulloch, J. *Juzzy online: An online toolkit for the design, implementation, execution and sharing of Type-1 and Type-2 fuzzy logic systems.* IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), Beijing, China, 2014, p 2321-2328, doi: 10.1109/FUZZ-IEEE.2014.6891548.

[152] Markus Hittmeir, Andreas Ekelhart, and Rudolf Mayer. On the Utility of Synthetic Data: An Empirical Evaluation on Machine Learning Tasks *In Proceedings of the 14th International Conference on Availability, Reliability and Security.* (ARES 2019) (ARES '19), 2019, Canterbury, United Kingdom. ACM, New York, NY, USA, 6 pages.

[153] Valdez, R., Maldonado, Y., Quevedo, J.A. Fuzzy Hardware Tool:An Adaptable Tool to Facilitate the Implementation of Fuzzy Inference Systems in Hardware. *Electronics.* 2023, 12, 2853. https://doi.org/10.3390/electronics12132853

[154] HSE. *Reducing error and influencing behaviour (HSG48).* 1999. https://www.hse.gov.uk/pubns/books/hsg48.htm Accessed: October 2024

[155] Gertman, D., Blackman, H., Marble, J., Byers, J. and Smith, C. *The SPAR-H Human Reliability Analysis Method.* NUREG/CR-6883. 2004. https://www.nrc.gov/reading-rm/doc-collections/nuregs/contract/cr6883/cr6883.pdf Accessed October 2024

[156] Williams, J. C. HEART – A Proposed Method for Achieving High Reliability in Process Operation by means of Human Factors Engineering Technology. *In Proceedings of a Symposium on the Achievement of Reliability in Operating Plant.* Safety and Reliability Society, Elsevier. 1985, p 87-109.

[157] Aalipour, M., Ayele, Y.Z. & Barabadi, A. *Human reliability assessment (HRA) in maintenance of production process: a case study.* International Journal of System Assurance Engineering Management 7, 2016, p 229–238. https://doi.org/10.1007/s13198-016-0453-z

[158] Wang, Y. The Human Reliability Analysis in Level 2 PSA Using SPAR-H Method. *Advanced Materials Research.* Vols. 608-609, 2013, p 848-853. doi:10.4028/www.scientific.net/AMR

[159] Kirwan, B. The validation of three Human Reliability Quantification techniques - THERP, HEART and JHEDI: Part 1 - technique descriptions and validation issues. *Applied Ergonomics.* Vol 27. No. 6. 1996, p 359-373. https://doi.org/10.1016/S0003-6870(96)00044-0

[160] Thoppil, N. M., Vasu, V., & Rao, C. S. P. On the criticality analysis of computer numerical control lathe subsystems for predictive maintenance. *Arabian Journal for Science and Engineering*, 45, 2020, p 5259-5271.

[161] Mo, H., Wang, F. Y., Zhou, M., Li, R., & Xiao, Z. Footprint of uncertainty for type-2 fuzzy sets. *Information Sciences*, 272, 2014, p 96-110.

[162] ul Hassan, F., Nguyen, T., Le, T., & Le, C. Automated prioritization of construction project requirements using machine learning and fuzzy Failure Mode and Effects Analysis (FMEA). *Automation in Construction*, 154, 2023.

[163] Bansal, M., Goyal, A., & Choudhary, A. A comparative analysis of K-nearest neighbour, genetic, support vector machine, decision tree, and long short term memory algorithms in machine learning. *Decision Analytics Journal*, 3, 2022.

[164] Ivančan, J., Lisjak, D., Pavletić, D., & Kolar, D. Improvement of Failure Mode and Effects Analysis Using Fuzzy and Adaptive Neuro-Fuzzy Inference System. *Machines.* 2023, 11, 739. https://doi.org/10.3390/machines11070739

[165] Wu, Z., Liu, W., & Nie, W. Literature review and prospect of the development and application of FMEA in manufacturing industry. *The International Journal of Advanced Manufacturing Technology*, 112, 2021, p 1409-1436.

[166] Zhang, S., Zhang, A., Chen, P., Li, H., Zeng, X., Chen, S., & Zhou, Q. Application of artificial intelligence hybrid models in safety assessment of submarine pipelines: Principles and methods. Ocean Engineering, 312, 2024.