Politecnico di Torino

Department of Applied Science and Technology (DISAT)

Master's Degree Thesis

Cybersecurity Risk Assessment in Chemical Process Control Systems

Cybersecurity and control of process and energy plants

Supervisors

Prof. Demichela Micaela

Prof. Davide Fissore

Candidate

Fakhriddin Abdullaev

Academic Year 2025–2026

Abstract

Industrial processes in modern chemical plants rely on interconnected control systems that combine physical equipment with digital technologies. While this integration improves efficiency and automation, it also introduces cybersecurity risks that can directly affect physical operations. Cyber attacks may manipulate control logic, alter sensor data, or disrupt communication between system components, potentially leading to unsafe conditions.

This thesis investigates how cybersecurity can be integrated into traditional risk-assessment methods used in process engineering. Classical techniques such as Hazard and Operability Study (HAZOP) and Fault Tree Analysis (FTA) are extended to include cyber-related causes, allowing both accidental failures and deliberate attacks to be analysed within a unified framework.

The methodology is applied to a case study of a three-phase separator system (VS003), demonstrating how cyber threats can impact key process variables such as pressure and liquid level. A particular focus is placed on common-cause failures in which a single spoofed measurement simultaneously defeats the basic process control system and the independent safety instrumented system — a dependency invisible to purely physical analysis.

Prevention and protection measures are then examined through a defence-in-depth lens structured around the NIST Cybersecurity Framework functions (Identify, Protect, Detect, Respond, Recover) and the IEC 62443 zone-and-conduit model. Layered controls including network segmentation, role-based access control, intrusion detection, offline backup and recovery, SIS isolation, and patch management are discussed and mapped to the VS003 architecture.

Finally, an AI-based operator-support concept is introduced, illustrating how large language models can assist operators in detecting anomalies and coordinating incident response while remaining outside the safety-critical control loop.

Contents

Abstract i

List of Figures iii

List of Tables vi

List of Abbreviations vii

1 Introduction 1
- 1.1 Background and Motivation . . . 1
- 1.2 Problem Statement . . . 1
- 1.3 Objectives of the Thesis . . . 2
- 1.4 Scope and Limitations . . . 3

2 Industrial Control Systems and Architecture 4
- 2.1 Components and Operation of Industrial Control Systems . . . 4
- 2.2 Layered Architecture of Industrial Control Systems . . . 5
- 2.3 Communication Networks and Industrial Protocols . . . 6
- 2.4 Cyber–Physical Systems in Industrial Environments . . . 7
- 2.5 Security Challenges in Industrial Control Systems . . . 7

3 Risk Assessment Methodology 9
- 3.1 Research Workflow and Methodological Approach . . . 9
- 3.2 From Process Safety to Cyber Risk Assessment . . . 11
- 3.3 Integration of Cybersecurity into Hazard Identification . . . 12
- 3.4 HAZOP Methodology for Cyber–Physical Systems . . . 12
- 3.5 Fault Tree Analysis for Cyber–Physical Risk Modeling . . . 13
- 3.6 Risk Evaluation and Mitigation Framework . . . 13

4 Hazard Identification Techniques 16
- 4.1 Approaches to Hazard Identification . . . 16
- 4.2 Hazard and Operability Study (HAZOP) . . . 16
- 4.3 Fault Tree Analysis (FTA) . . . 17
- 4.4 Comparison and Integration of HAZOP and FTA . . . 18

5 Vulnerabilities and Cyber Attacks in Industrial Control Systems 20
- 5.1 Sources of Vulnerability in Industrial Control Systems . . . 20
- 5.2 Vulnerabilities in PLC, SCADA, and HMI Systems . . . 20
- 5.3 Types of Cyber Attacks in Industrial Control Systems . . . 21
- 5.4 Impact of Cyber Attacks on Physical Processes . . . 22

5.5 Notable Real-World ICS Cyber Incidents 22

6 Prevention and Protection Measures **30**
6.1 A Layered Approach to Cybersecurity Protection 30
6.2 Network Segmentation and Secure Architecture 31
6.3 Access Control and Authentication Mechanisms 32
6.4 Monitoring, Detection, and Incident Response 33
6.5 Standards and Best Practices for Industrial Cybersecurity 33

7 Case Study: VS003 Separator System **36**
7.1 System Description and Architecture 36
7.2 Identification of Process Variables and Hazards 39
7.3 Cybersecurity Risk Analysis of the VS003 System 40
7.4 Mitigation Measures for the VS003 System 42

8 LLM-Based Operator Assistant **44**
8.1 Concept of AI-Assisted Monitoring in Industrial Control Systems 44

9 Conclusion **46**

List of Figures

2.1 Overview of the data flow in an industrial control system, from field sensors through the controller to the actuators, with the supervisory layer (SCADA/HMI) and the closed feedback loop. 4
2.2 Logic processing in a PLC: acquisition of input signals, execution of the control program against predefined setpoints, and output of commands to the actuators. 5
2.3 Purdue reference architecture, with the VS003 components mapped to each level. 6

3.1 Roadmap 1 — Integrated cyber–physical risk-assessment methodology adopted in this thesis. 10
3.2 Roadmap 2 — Control-Room (CR) Operator support system. Branch 2 (support design) is the work proposed in this thesis; Branches 1 and 3 (marked ⊗) were drawn for explanation and are not part of the proposal. . 11
3.3 Integrated cyber–physical risk-assessment methodology: HAZOP and FTA form the safety baseline, cyber causes are injected as additional initiating events, and the results feed risk evaluation and mitigation in an iterative cycle. 15

5.1 **Technical attack lifecycle of Stuxnet (2010).** 23
5.2 **Technical attack lifecycle of the German Steel Mill incident (2014).** The attack highlights the dangers of insufficient network segmentation between IT and OT environments and the compromise of safety-critical systems. 24
5.3 **Evolutionary attack lifecycle of the Ukraine Power Grid incidents (2015 vs. 2016).** The layout contrasts the transition from manual, credential-driven infrastructure manipulation in 2015 to automated, protocol-specific malware architectures in 2016. 25
5.4 **Technical attack lifecycle of the TRITON / TRISIS incident (2017).** The diagram outlines the escalation sequence from network exploitation to industrial safety controller reprogramming, accentuating why physical key positioning and decoupled safety zoning remain critical structural defenses. 26
5.5 **Technical attack lifecycle of the Oldsmar Water-Treatment Plant incident (2021).** The diagram maps the infiltration vector from unauthenticated remote boundary exposure to manual dosing alteration, validating the role of live operator validation and chemical process constraints. 27

5.6 **Technical threat lifecycle of the Colonial Pipeline incident (2021).** The schematic highlights how structural disruption confined entirely to enterprise IT networks can induce complete administrative blindness, triggering a precautionary operational physical shutdown. 28

6.1 Defense-in-depth: the layered protection and prevention measures applied to the VS003 system. No single control is sufficient; together they reduce both the likelihood and the impact of a cyber attack. The boxes are coloured by their role in the NIST cybersecurity framework: protect (teal), detect (amber), respond (red) and recover (green). 31
6.2 Secure network architecture for the VS003 system. Zones are separated by firewalls with deny-by-default rules and host/protocol whitelists; remote access enters through a VPN with multi-factor authentication into a DMZ. Monitoring data flows upward (read-only) so the remote operator can *view* sensor status, while control commands originate only in the control-room zone and flow down to the field — there is no command path from the remote/DMZ side into the process. 32
6.3 The core functions of the NIST Cybersecurity Framework, applied as a continuous improvement cycle to the industrial control system. 34

7.1 Control architecture of the VS003 separator: the basic process control system (PLC/BPCS) and the independent safety instrumented system (SIS) act on the gas (PV003) and liquid (LV003) valves using the pressure (PT003) and level (LT004) measurements. 37
7.2 Laboratory-scale separation-process setup presented during a research meeting organized by Prof. Micaela Demichela. The setup belongs to a collaborating researcher and is reproduced with permission for illustrative and educational purposes. The image is included to support understanding of the process architecture discussed in this thesis and does not represent an experimental setup developed within the present work. 38
7.3 Process scheme presented during a research meeting organized by Prof. Micaela Demichela and reproduced with permission from the original researcher. The figure is included as a reference example to facilitate understanding of the process and does not represent a process design developed within this thesis. 39
7.4 Fault tree for separator overpressure. The shaded cyber basic events appear alongside the classical physical faults; the right branch shows the common cause, where a spoofed PT003 signal defeats both the control and the safety layer. 42

8.1 Architecture of the LLM-based operator assistant. The assistant correlates process data, alarm logs, and network monitoring to detect anomalies and explain them to the operator, but it only *advises*; all control decisions remain with the human operator, and the assistant never actuates the safety-critical control loop directly. 45

8.2 Core functions of the LLM-based operator assistant. The top row lists what the assistant does — correlating data, detecting anomalies, explaining them, and recommending responses — while the bottom row marks the human-centered boundaries: the operator remains in control and the assistant never acts on the safety-critical loop. The capabilities are shown in blue and the safety boundaries in green. 45

List of Tables

3.1 Likelihood scale used in the risk assessment. For cyber-initiating events the level is judged qualitatively from exposure, attacker capability, and exploitability; for physical-initiating events it may be anchored to order-of-magnitude failure-rate data from reliability sources (e.g. OREDA, IEC 61511). 14
3.2 Severity scale, defined by the consequence on the physical process. 14
3.3 Risk matrix combining likelihood and severity. Cells give the risk level: L = Low, M = Medium, H = High, VH = Very High. 15

4.1 Standard HAZOP guidewords and their application to process parameters [1]. 17
4.2 Main elements used in Fault Tree Analysis [2]. 18
4.3 Complementary characteristics of HAZOP and FTA, and the role of each in this thesis. 19

5.1 Attack classes affecting the control system, mapped to MITRE ATT&CK for ICS techniques [3]. Technique identifiers verified against attack.mitre.org. 22
5.2 Summary of the ICS cyber incidents and the protection measure most relevant to the VS003 system. 29

6.1 Security Levels defined by IEC 62443, based on the capability of the threat they are designed to resist [4]. 34
6.2 The nine defence-in-depth layers proposed for VS003, mapped to their NIST Cybersecurity Framework function and to the threat each addresses. 35

7.1 Cyber-extended HAZOP worksheet for the VS003 separator (node: separator vessel). 40
7.2 Risk assessment and reduction for the VS003 system: likelihood (L), severity (S), inherent risk before mitigation, the mitigation applied, and the residual risk. Levels follow the matrix of Table 3.3. 43

List of Abbreviations

Abbreviation	Meaning
BPCS	Basic Process Control System
DCS	Distributed Control System
DoS	Denial of Service
FTA	Fault Tree Analysis
HAZOP	Hazard and Operability Study
HMI	Human–Machine Interface
ICS	Industrial Control System
IT	Information Technology
MFA	Multi-Factor Authentication
OT	Operational Technology
PLC	Programmable Logic Controller
PSV	Pressure Safety (Relief) Valve
SCADA	Supervisory Control and Data Acquisition
SIS	Safety Instrumented System

Chapter 1

Introduction

1.1 Background and Motivation

Over the past decades, chemical process industries have undergone a profound digital transformation. Plants that were once supervised mechanically and manually have been progressively replaced by industrial control systems (ICS) that combine physical equipment—sensors, valves, pumps, and vessels—with digital controllers and communication networks.

This transformation has improved efficiency, automation, and real-time monitoring capabilities. Operators are now able to supervise entire processes from centralized control rooms, while automated control loops maintain stable operating conditions with minimal manual intervention. However, this increasing level of connectivity has introduced new challenges that were not present in earlier industrial environments.

One of the most important challenges is cybersecurity. As industrial systems become interconnected and, in some cases, linked to corporate networks or external services, they are exposed to cyber threats. Unlike traditional information systems, where security breaches mainly affect data, attacks on industrial control systems can directly influence physical processes [5]. A cyber incident may therefore lead to equipment malfunction, production downtime, or unsafe operating conditions. This concern is no longer theoretical: recent industry analyses report a sharp rise in attacks against operational technology, with ransomware targeting industrial organizations increasing markedly year over year and critical infrastructure identified as a prime target [6, 7]. The 2021 Colonial Pipeline incident, although it originated in the corporate network, illustrates how a cyber event can force the shutdown of a large industrial operation [8].

The growing convergence between information technology (IT) and operational technology (OT) has changed the nature of risk in industrial environments. Traditional safety approaches, which focus primarily on equipment failures and human errors, are no longer sufficient on their own. There is a clear need to treat cybersecurity as an integral part of process safety, especially in critical sectors such as chemical processing, where failures can have severe consequences.

1.2 Problem Statement

Chemical process industries operate under strict safety and reliability requirements, where even small deviations in process variables such as pressure, temperature, or liquid level can lead to hazardous conditions. To manage these risks, established methodologies such

as the Hazard and Operability Study (HAZOP) and Fault Tree Analysis (FTA) are widely used to identify potential failures and evaluate their consequences [1, 2].

However, these traditional approaches were designed primarily for accidental events—equipment malfunction, process deviations, and human error. They generally do not consider intentional actions, such as cyber attacks, that deliberately target control systems and communication networks. With the increasing digitalization of industrial environments, this limitation has become more significant.

Modern industrial control systems rely on interconnected components, including programmable logic controllers (PLCs), supervisory systems, and communication infrastructures. These components may contain vulnerabilities that attackers can exploit to manipulate system behavior. For example, unauthorized access to a controller may allow modification of control logic, while falsified sensor data may lead an automated controller to take incorrect actions.

The key problem addressed in this thesis is the limited integration between cybersecurity analysis and traditional process-safety methodologies. Existing risk-assessment techniques do not fully capture the impact of cyber threats on physical industrial processes. A unified approach is therefore needed that can identify, analyze, and evaluate both conventional hazards and cyber-induced risks within industrial control systems.

1.3 Objectives of the Thesis

The main objective of this thesis is to develop a structured approach for assessing cybersecurity risks in industrial control systems used in chemical process environments. The work aims to bridge the gap between traditional process-safety analysis and modern cybersecurity practice by integrating both perspectives into a unified methodology.

To achieve this goal, several specific objectives are defined. First, the thesis analyzes the architecture and components of industrial control systems, including programmable logic controllers, supervisory systems, human–machine interfaces, and communication networks. Understanding the role and interaction of these components is essential for identifying potential points of vulnerability.

Second, the study investigates common vulnerabilities and attack vectors affecting industrial systems, examining how cyber threats can target different system components and propagate through the control architecture.

Third, the thesis extends established hazard-identification techniques, namely HAZOP and FTA, to incorporate cyber-related causes. This allows a more comprehensive evaluation of risks that includes both accidental failures and intentional malicious actions.

Fourth, the work proposes a set of prevention and protection measures, covering both technical solutions and organizational practices, aimed at improving the security and resilience of industrial control systems.

Finally, the proposed methodology is applied to a case study of a three-phase separator control system (VS003), demonstrating how cyber risks can influence process behavior. The thesis also explores the concept of an artificial-intelligence-based operator support tool to assist in detecting abnormal conditions and potential cyber incidents.

1.4 Scope and Limitations

This thesis focuses on cybersecurity risk assessment within industrial control systems used in chemical process environments. The analysis considers typical components of such systems, including field devices, control units, supervisory systems, and communication networks. The objective is to evaluate how cyber threats can affect the behavior of physical processes and introduce additional risks beyond traditional failure scenarios.

The study is centered on a specific case study, the VS003 three-phase separator system. This system is used as a representative example to demonstrate the application of the proposed methodology. While the analysis provides insights relevant to similar industrial systems, it does not aim to cover all possible configurations or industrial sectors.

The work is primarily analytical and conceptual. It focuses on identifying risks, modeling their impact, and proposing mitigation strategies. The thesis does not include experimental implementation of cyber attacks or real-time deployment of detection systems. Furthermore, the proposed artificial-intelligence-based operator assistant is presented as a conceptual approach rather than a fully developed system.

Chapter 2

Industrial Control Systems and Architecture

2.1 Components and Operation of Industrial Control Systems

An industrial control system (ICS) is the layer that lets a computer run a physical plant. In a chemical facility it reads the state of the process—pressures, temperatures, flows, levels—and converts those readings into the valve and pump commands that keep the plant inside its safe operating envelope [5]. The VS003 separator examined in this thesis is exactly such a process: its behaviour is governed entirely by the sensors, controllers, and actuators described in this chapter, and that is what makes it a meaningful target for a combined safety-and-cyber analysis.

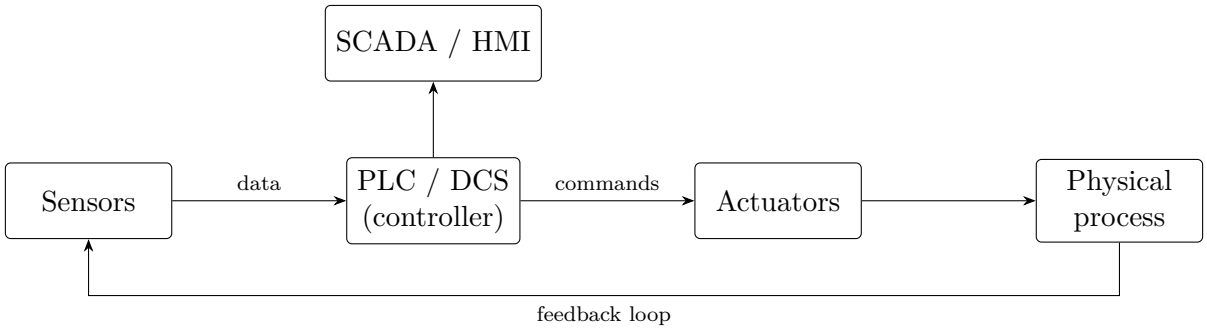


Figure 2.1: Overview of the data flow in an industrial control system, from field sensors through the controller to the actuators, with the supervisory layer (SCADA/HMI) and the closed feedback loop.

Three kinds of device sit closest to the process. Sensors measure what the process is doing—in VS003, the pressure transmitter PT003 and the level transmitter LT004. Controllers read those measurements and decide what to do. Actuators then carry the decision out: the control valves PV003 and LV003 modulate to hold pressure and liquid level at their setpoints. Together they close a loop that runs continuously, without an operator in it for routine adjustments.

The workhorse of the control layer is the programmable logic controller (PLC). A PLC executes its control logic in a fixed, repeating scan—read inputs, evaluate logic,

write outputs—and is engineered to keep doing so for years in conditions that would not suit ordinary computing hardware. Where a single controller is not enough, for instance a plant with many units spread across a site, a distributed control system (DCS) coordinates several controllers at once.

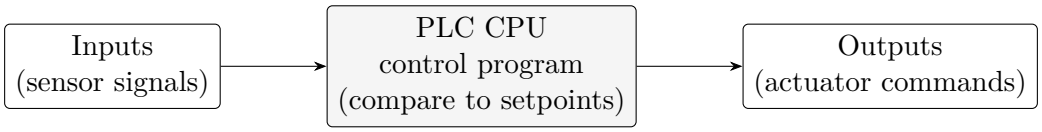


Figure 2.2: Logic processing in a PLC: acquisition of input signals, execution of the control program against predefined setpoints, and output of commands to the actuators.

Above the controllers sits the supervisory layer, which is where operators actually see the plant. A Supervisory Control and Data Acquisition (SCADA) system gathers data from the controllers, stores it, and renders it on screen [5]. The Human–Machine Interface (HMI) is what the operator watches: it shows process values, raises alarms, and is the point from which a manual command is issued when something looks wrong.

Networks tie these layers together using industrial protocols built for fast, predictable data exchange. As plants modernise, those networks no longer stop at the control system—they reach up into corporate IT for reporting and remote access. That connectivity is convenient, and, as later chapters show, it is also where much of the cyber risk enters.

What separates an ICS from ordinary IT is precisely this: its outputs are physical. A wrong command here does not corrupt a file, it moves a valve. That coupling is the reason cybersecurity in a chemical plant cannot be treated separately from process safety—a theme the rest of the thesis develops around the VS003 case.

2.2 Layered Architecture of Industrial Control Systems

Control systems are usually described as a stack of layers, each with its own responsibility, running from the metal of the process at the bottom to the business network at the top. The reference model for this is the Purdue Enterprise Reference Architecture [5]. It is useful here for two reasons at once: it organises how the plant operates, and it shows where the security boundaries naturally belong.

At the bottom are the physical process and the field devices wired to it—the sensors and actuators. This is the only level that touches the physical world; everything above it is moving numbers. In VS003 this is where PT003, LT004, and the control valves live.

Above the field level is the control layer, home to the PLCs and distributed control systems. A controller takes the sensor inputs, evaluates them against its programmed logic, and issues commands to the actuators. Its job is to hold the process stable and safe in real time, scan after scan.

The supervisory layer provides monitoring and coordination. SCADA servers and HMI stations live here, letting operators visualise process data, follow performance, and acknowledge alarms. Historical data is archived at this level and later used to investigate events and tune operation.

Higher still is the operations or site-management layer—production planning, optimisation, maintenance scheduling—often hosting the data historians and application servers

that aggregate information from the control system. At the top, the enterprise layer holds the corporate IT systems: business applications, resource-planning software, and the connections out to external or cloud services.

The reason this separation matters is not tidiness. Each layer carries a different level of exposure: the enterprise network talks to the outside world, the field layer must never have to. Keeping the boundaries between them clear is what stops an intrusion in a business system from reaching a controller—and the tighter plants are interconnected, the more that depends on deliberate design rather than on the old assumption of physical isolation.

In this thesis the VS003 separator maps onto the model cleanly: PT003, LT004, and the control valves at the field level; the PLC and the safety instrumented system at the control level; the SCADA and HMI stations at the supervisory level. That mapping is the scaffold for everything that follows—the hazard analysis, the attack scenarios, and the segmentation and access-control measures proposed in later chapters.

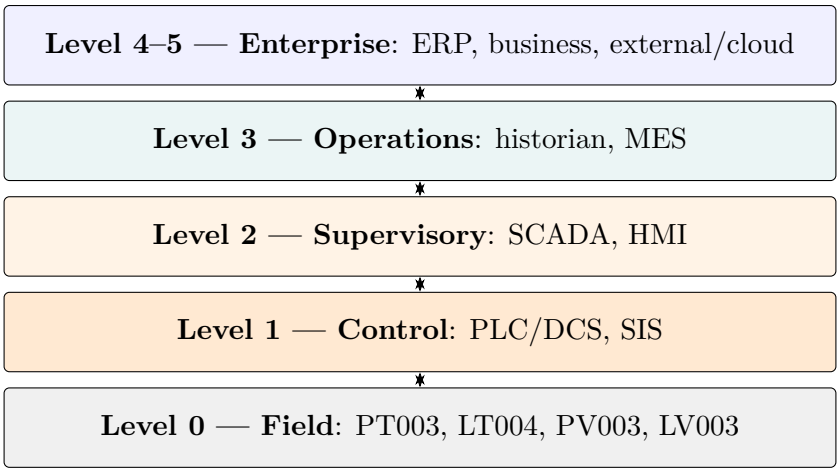


Figure 2.3: Purdue reference architecture, with the VS003 components mapped to each level.

2.3 Communication Networks and Industrial Protocols

Nothing in a control system works without the network that connects it. Sensor readings have to reach the controller, and the controller's commands have to reach the valves, on time, every cycle. A late or lost message is not a nuisance here; it can destabilise the process.

That requirement shapes the protocols. Industrial communication is tuned for real-time, deterministic delivery, and historically that tuning came at the expense of security. Modbus, one of the most common protocols, is a clear example: it is simple and fast, and it sends commands with no authentication, no encryption, and no integrity check [5]. PROFINET serves high-speed automation traffic. OPC Unified Architecture (OPC UA) is the more recent answer to the security gap, building authentication and encryption into the standard itself rather than leaving them to the operator.

The legacy of insecure-by-design protocols is the problem. On a protocol like Modbus, an attacker who reaches the network can read traffic, alter packets, or inject commands

of their own, and the controller has no way to tell those commands from legitimate ones. This is the assumption behind the spoofing scenario analysed later for VS003.

The exposure has widened because control networks rarely stay isolated. They now interconnect with corporate IT and with remote-access services, which improves operability but enlarges the attack surface. A single compromised connection can become the foothold from which an attacker moves laterally toward the controllers—which is why the network is not just plumbing but one of the main routes an attack actually travels.

2.4 Cyber–Physical Systems in Industrial Environments

A modern control system is best understood as a cyber–physical system [5]: computation and physical process bound together so tightly that neither can be reasoned about alone. The digital side is not observing the plant from a distance—it is steering it, continuously.

The loop is simple to state. Sensors read the process—pressure, temperature, flow, level—and pass the values to the controller; the controller decides; the actuators change the physical state; and the new state is measured again. Cyber and physical alternate around this loop many times a second.

Because the loop depends on both halves, a fault on the cyber side lands on the physical side. A bad sensor value produces a wrong control decision. Tampered logic can stop a safety function from acting when it should. The failure is digital in origin but physical in consequence.

That is what makes the risk different from IT. A breach in an office network costs data or uptime; a breach in a cyber–physical system can drive the plant into an unsafe state, damage equipment, or release hazardous material. In a chemical separator, where pressure and level have hard limits, the gap between a digital fault and a physical hazard can be very short.

So cybersecurity has to live inside the process-safety analysis, not beside it. The risk assessment has to account not only for components failing by accident but for an adversary deliberately exploiting the cyber side—and it is that requirement, applied to VS003, that the next chapter turns into a method.

2.5 Security Challenges in Industrial Control Systems

Securing a control system is not the same problem as securing an IT network [4]. The difficulty is that the obvious security measures often collide with the things a plant cannot give up: continuous availability, real-time response, and equipment that must run untouched for years.

The first obstacle is history. Many control systems were built for isolated environments, where being physically cut off was the security model. Their protocols and devices were never given encryption, authentication, or access control, because nothing on the wire was supposed to be hostile. Connect those same systems to a wider network and the missing protections become the weak point.

Equipment lifetime makes this worse. A PLC may be specified to run for decades, and taking it offline to patch or replace is expensive and disruptive. The practical result is

that known-vulnerable devices stay in service far longer than any IT asset would, because the cost of change is measured in lost production.

Availability pulls in the same direction. An office machine can be rebooted for an update; a running separator usually cannot. Security controls therefore have to be introduced without interrupting the process or destabilising it—a constraint that rules out many measures that would be routine in IT.

Connectivity adds the modern twist. Linking operational technology to IT brings real benefits in monitoring and efficiency, but every new link is also a possible way in. Remote-access services, flat or misconfigured networks, and unprotected interfaces are the openings attackers look for first.

And the system includes people. Operators interact with it through interfaces and judgement calls, and a lapse—an unmanaged credential, a missed warning, simple lack of awareness—can open a door that no protocol weakness would have.

Taken together, these constraints rule out the simple answer of dropping IT security tools onto a plant and calling it done. Security here has to be designed around the process, aware of its limits, and integrated with safety rather than layered on top—which is the starting point for the risk-assessment method developed next.

Chapter 3

Risk Assessment Methodology

3.1 Research Workflow and Methodological Approach

The work in this thesis followed a workflow agreed with the supervisors at the start of the project. The aim was concrete: to build a cyber–physical risk-assessment method that an engineer could actually apply to an industrial control system in a chemical plant, and then to apply it to one.

It began with a literature review across the areas the problem touches—cyber–physical systems, industrial control systems, cybersecurity fundamentals, risk-assessment methods, HAZOP, Fault Tree Analysis, and hazard-identification techniques including the Japanese hazard-prediction method Kiken Yochi Training (KYT). The point of that reading was not coverage for its own sake but to see where process safety and cybersecurity overlap, and where the standard safety tools stop short.

That groundwork was then put to work on the VS003 three-phase separator. Hazard identification combined the usual process-safety reasoning with cyber considerations; each scenario was assessed for risk and traced back to the cyber actions that could trigger it. The final stage moved from analysis to response: mitigation measures, and the outline of an AI-based operator-support concept meant to help the control-room operator make sense of abnormal conditions rather than to act on the plant itself.

This methodology is summarized in Figure 3.1. Figure 3.2 presents the broader control-room operator support framework outlined during the planning phase; within it, the design of operator-support measures (Branch 2) is addressed conceptually in this thesis, while the experimental activities and machine-learning components are identified as directions for future work.

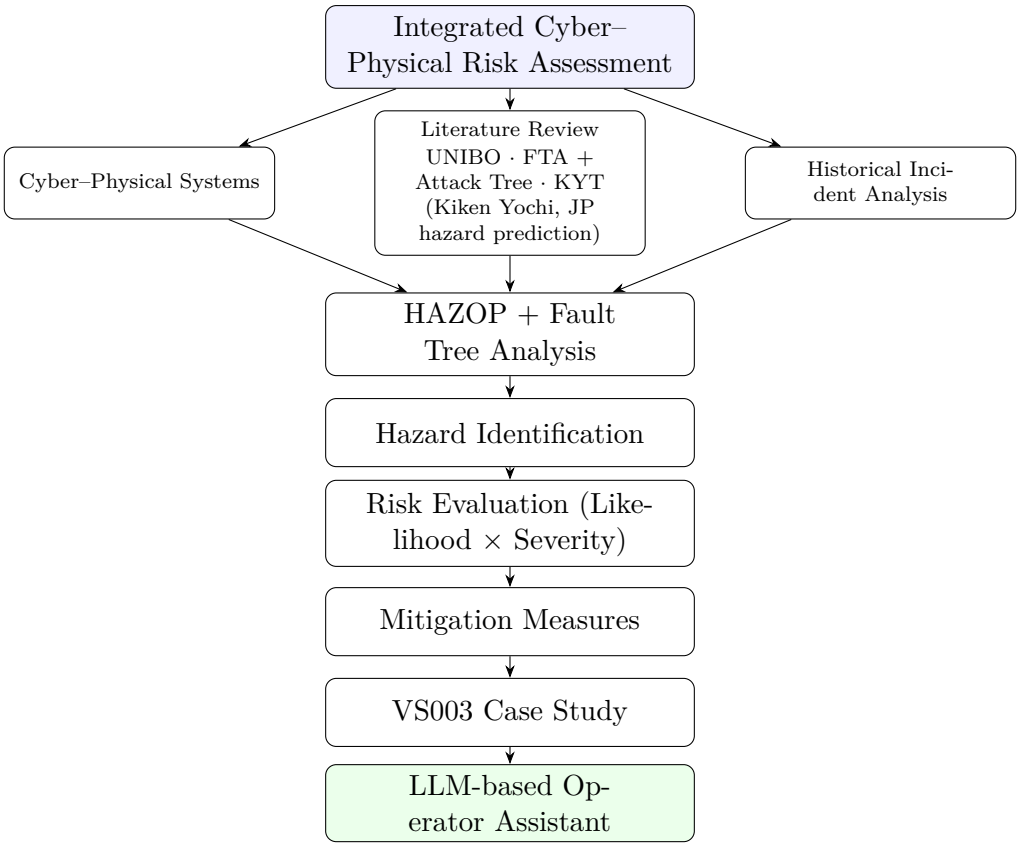


Figure 3.1: Roadmap 1 — Integrated cyber–physical risk-assessment methodology adopted in this thesis.

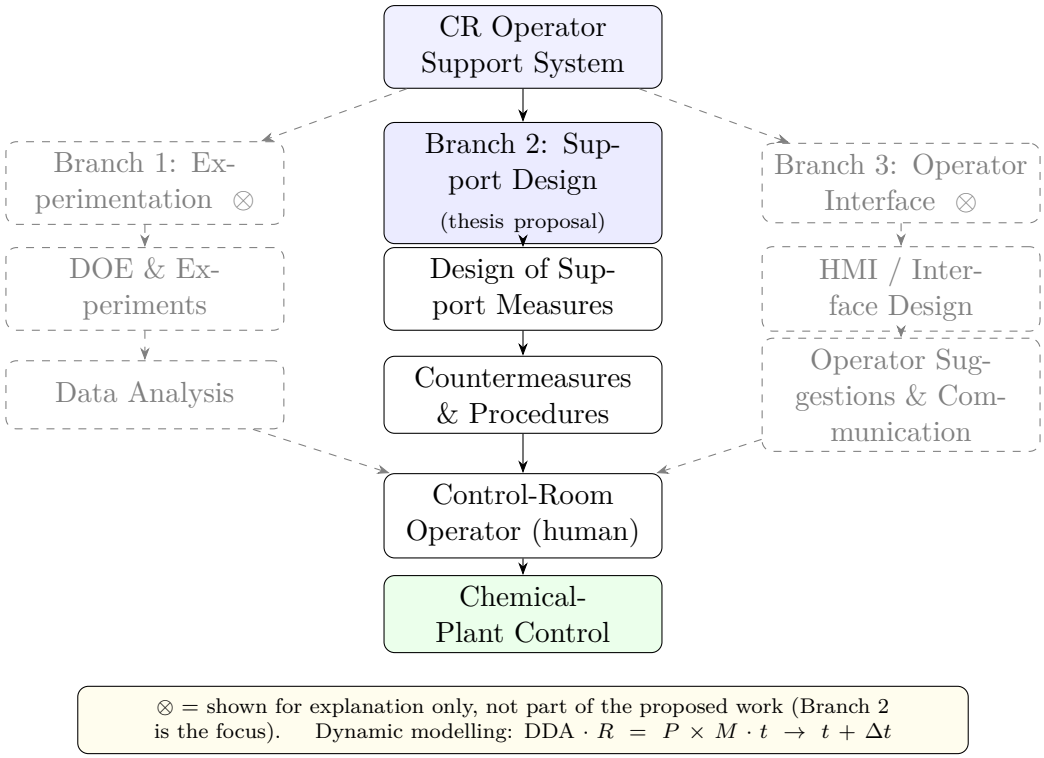


Figure 3.2: Roadmap 2 — Control-Room (CR) Operator support system. Branch 2 (support design) is the work proposed in this thesis; Branches 1 and 3 (marked ⊗) were drawn for explanation and are not part of the proposal.

3.2 From Process Safety to Cyber Risk Assessment

Risk assessment is common ground between safety engineering and cybersecurity: in both, it is the structured way of asking what can go wrong, how likely it is, and how bad it would be [9]. In a plant, that question is what keeps the process inside its limits instead of drifting toward them.

Process engineering has long answered it for accidental events—equipment failures, process upsets, human error—using HAZOP and Fault Tree Analysis to find hazards and trace their consequences [1, 2]. Those tools are good at physical failure. What they were never built to handle is an intelligent adversary.

That gap matters now because control systems are networked. A cyber threat can exploit a controller, a protocol, or a piece of software to force an action on the process, and unlike a random failure it is deliberate—chosen to slip past defences or to stay hidden. Assessing it means bringing in things a HAZOP sheet never had a column for: who the attacker is, how they get in, which weakness they use, and what it does to both the digital and the physical side of the plant.

The stakes are higher in an ICS than in an office. Falsified sensor data, altered control logic, or a jammed communication link does not just lose information—it makes the process behave wrongly, and in a chemical plant that shows up as pressure past its limit, flow out of control, or a safety function that fails to act.

The response taken here is to stop treating the two analyses as separate. Combining engineering risk assessment with cyber analysis is what lets a single picture cover both the accidental and the malicious, and it is the perspective the rest of this methodology is

built on.

3.3 Integration of Cybersecurity into Hazard Identification

Classical hazard identification looks for deviations caused by failures, disturbances, or operator error. HAZOP, for instance, walks through process parameters with guidewords to surface abnormal conditions and their consequences. It does this well—but it has no place in it for a deliberate cyber cause.

The move this thesis makes is to treat a cyber action as just another way to initiate a deviation. Unauthorised access, spoofed sensor data, or altered control logic can produce abnormal conditions that are, to the process, indistinguishable from a conventional fault. The difference is intent: the fault was chosen, and chosen to exploit a specific weakness.

In practice this means widening each deviation to carry both kinds of cause. "High pressure" in the separator might come from a mechanical fault—but equally from a falsified PT003 reading or a tampered controller setpoint. "No flow" or "wrong level" might come from a stuck valve—or from a denial-of-service attack on the controller link, a spoofed signal, or an unauthorised configuration change.

Reading deviations this way does two useful things. It shows how a cyber event travels through the system into physical behaviour, and it exposes dependencies between digital components and safety functions that a purely physical study leaves invisible—most importantly, the shared-signal dependency that the VS003 case turns on. The result is a single hazard picture that holds accidental and malicious causes together, which is the shift the methodology depends on.

3.4 HAZOP Methodology for Cyber–Physical Systems

HAZOP is a structured hazard- and operability-review technique [1]: it examines deviations of key process parameters—flow, pressure, temperature, level—and, for each, asks what could cause it, what would follow, and what protects against it.

The procedure is systematic. The process is divided into nodes, usually from the P&ID, and for each node guidewords such as "no," "more," "less," and "reverse" are applied to the relevant parameters to generate deviations. Each deviation is then worked through for causes, consequences, and safeguards.

Extending it to a cyber–physical system does not change the procedure; it changes the cause list. For every deviation, the analysis now asks not only which equipment fault could produce it but which cyber event could produce the same effect.

Take "no flow." Conventionally that is a closed valve or a failed pump. In a networked plant it can equally be a denial-of-service attack cutting the controller-to-actuator link, or modified logic that simply stops the actuator responding. Take "more pressure": a process disturbance, yes—but also a falsified PT003 reading that convinces the controller to drive pressure past its safe limit. The cyber cause sits in the worksheet beside the physical one, on equal terms.

Carrying both causes through the same review is what lets it catch the scenarios that only appear at the boundary between the digital and physical sides—where a fault in one

shows up as a hazard in the other. Done this way, a familiar engineering method becomes a structured way to surface cyber–physical risk, and a starting point for deciding what to do about it.

3.5 Fault Tree Analysis for Cyber–Physical Risk Modeling

Where HAZOP fans out from deviations, Fault Tree Analysis (FTA) works downward from a single bad outcome and asks what combinations of failures produce it. It is a logical map from causes to one defined consequence.

It starts from a top event—a system-level hazard such as vessel overpressure, loss of containment, or a shutdown. From there the analysis traces every contributing cause that could lead to it, building the tree downward.

Those causes are joined by logic gates. An AND gate means several conditions must hold together; an OR gate means any one of them is enough. That structure breaks an otherwise tangled failure scenario into parts that can be reasoned about one at a time.

For a cyber–physical system, the tree carries cyber causes alongside the hardware faults, degradation, and human errors a conventional fault tree already holds. They are not a separate diagram—they are additional branches in the same tree.

Consider "high pressure in the separator." One branch is a mechanical fault, say a relief valve that fails to lift. Another is a cyber action—falsified sensor data, or an unauthorised change to controller parameters—entering through its own gate. The crucial case is when one cyber cause feeds more than one branch at once: a single falsified reading can mislead the basic control and, if the safety system reads the same signal, defeat both at the same time. That common-cause behaviour is exactly what the VS003 case study analyses in detail.

Putting cyber causes into the tree makes it possible to see how an attack interacts with the physical process to reach a hazard, and to find the points where a single failure or action does the most damage—which is where mitigation effort is worth spending first. Used together, FTA's downward logic and HAZOP's deviation sweep cover the same system from two directions.

3.6 Risk Evaluation and Mitigation Framework

Once hazards are identified and their causes modelled, the question becomes how much risk each one carries and what to do about it. Risk here is the familiar pairing of likelihood and consequence—but in an ICS with cyber threats in play, both terms are harder to pin down than in a conventional safety study.

Likelihood is the harder one. A random failure has a rate; a cyber event does not. Its probability depends on whether a capable, motivated attacker exists, on the exposure of the interface, and on how easily the weakness can be exploited. Estimating it therefore means weighing the threat landscape, not just the equipment.

Consequence is judged, as in any process study, by the effect on the plant—equipment damage, lost production, environmental harm, risk to people. The cyber twist is timing: an attack can sit undetected, so the realistic consequence has to include effects that surface late, after the damage is done.

To act on this, the work uses a function-based control framework—identify, protect, detect, respond, recover [9]—as the backbone for selecting measures and, later, for organising the defence-in-depth proposed for VS003. The measures themselves fall into familiar groups: technical (segmentation, secure protocols, authentication, intrusion detection), organisational (access policy, training, incident-response planning), and physical (protecting the critical hardware itself).

The framework's real value is forcing prioritisation. Everything cannot be fixed at once, so effort goes where likelihood and consequence are highest—and pairing the HAZOP results with the fault tree is what points to those critical components and scenarios rather than leaving the choice to intuition.

Together, evaluation and mitigation close the loop: hazards are not just analysed but answered with concrete measures. The scales below make this repeatable. Table 3.1 and Table 3.2 fix what each likelihood and severity level means, and Table 3.3 combines them into the risk rating used when the method is applied to VS003 in the case study.

Table 3.1: Likelihood scale used in the risk assessment. For cyber-initiating events the level is judged qualitatively from exposure, attacker capability, and exploitability; for physical-initiating events it may be anchored to order-of-magnitude failure-rate data from reliability sources (e.g. OREDA, IEC 61511).

Level	**Descriptor**	**Interpretation**
L1	Rare	Isolated or strongly protected; no plausible path; very low failure rate.
L2	Unlikely	Requires significant attacker capability or privileged access.
L3	Possible	Feasible with known techniques and moderate access.
L4	Likely	Exposed interface, weak controls, known and easy exploit.
L5	Almost certain	Trivially exploitable / actively exploited, or frequent failure.

Table 3.2: Severity scale, defined by the consequence on the physical process.

Level	**Descriptor**	**Consequence**
S1	Negligible	No safety impact; minor deviation, automatically corrected.
S2	Minor	Short process upset; no release; limited downtime.
S3	Moderate	Process trip or shutdown; limited equipment stress.
S4	Major	Equipment damage, small release, potential injury.
S5	Catastrophic	Loss of containment, rupture or explosion; serious injury or environmental harm.

Table 3.3: Risk matrix combining likelihood and severity. Cells give the risk level: L = Low, M = Medium, H = High, VH = Very High.

	S1	**S2**	**S3**	**S4**	**S5**
L5	M	H	VH	VH	VH
L4	M	M	H	VH	VH
L3	L	M	H	H	VH
L2	L	L	M	H	H
L1	L	L	L	M	H

As an illustration of how the scales combine, take the spoofed-pressure scenario analysed later in the case study: a falsified PT003 reading that drives the separator toward overpressure. Its likelihood is judged Possible to Likely (L3–L4)—reaching the control network is required, but once that access exists no exotic capability is—while its severity is Major to Catastrophic (S4–S5), since the end state is loss of containment. Read off Table 3.3, that pairing sits in the High-to-Very-High band. The point the matrix alone does not show is that the same falsified signal can also defeat the safety layer, so the scenario cannot be down-rated by assuming the SIS will catch it—which is precisely why it is treated as a priority for the measures in the prevention chapter.

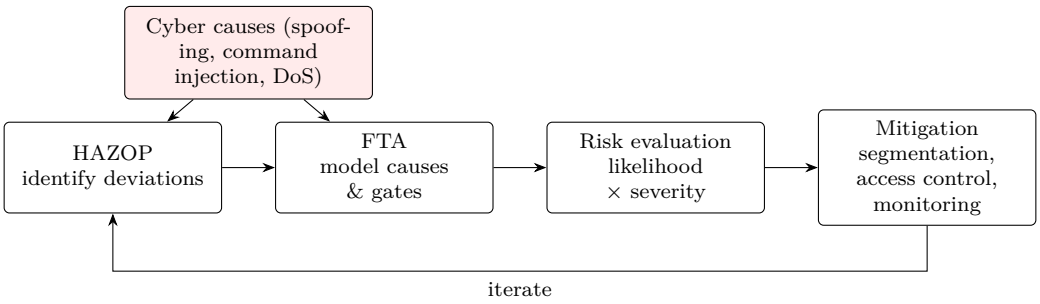


Figure 3.3: Integrated cyber–physical risk-assessment methodology: HAZOP and FTA form the safety baseline, cyber causes are injected as additional initiating events, and the results feed risk evaluation and mitigation in an iterative cycle.

Chapter 4

Hazard Identification Techniques

4.1 Approaches to Hazard Identification

The previous chapter set out the integrated risk-assessment methodology; this one looks more closely at the two techniques it rests on, HAZOP and Fault Tree Analysis, and at where each fits.

Hazard identification is the first real step of any risk assessment: the systematic search for the conditions, events, or behaviours that could push a system into an unsafe state. In a chemical plant the value of doing it well is simply that hazards get caught on paper, early, rather than in operation.

No single technique covers every case. The methods that have grown up over the years differ in detail, in the data they need, and in how they reason—some qualitative, leaning on structured expert discussion, others quantitative, built on probability and explicit models. HAZOP and Fault Tree Analysis are the two most established in the process industries [1, 2], and they sit at opposite ends of that spread: HAZOP is qualitative and works forward from process parameters, while FTA is deductive and works backward from a defined failure. Others—FMEA, Event Tree Analysis, simple checklists—each give a different cut at the problem and are chosen to suit the system and the question being asked [2].

What none of the classical methods originally accounted for is the cyber dimension. As the earlier chapters argued, a cyber event can initiate a process deviation just as a physical fault can, so it has to enter the hazard analysis on the same footing—which means adapting these techniques to the interaction between the digital and physical sides rather than inventing separate ones. The two sections that follow develop HAZOP and FTA in that light; the final section weighs their strengths against each other.

4.2 Hazard and Operability Study (HAZOP)

HAZOP is the process industry's standard technique for finding hazards [1]: a structured, systematic search for deviations from normal operation, together with their causes and consequences. It is most often run during design, but it works equally well on an existing process when the goal is to improve safety or operability.

It is a team method by design. Experts from different disciplines work through the system together, which is part of why it catches things a single reviewer would miss. The process is first broken into nodes—sections where parameters such as flow, pressure, temperature, or level can be examined meaningfully.

For each node, the team takes a parameter and applies guidewords to it to force out deviations. The standard set—"no," "more," "less," "as well as," "part of," "reverse," "other than"—turns a parameter into specific failures: "no flow," "more pressure," "reverse flow." Each deviation that survives is then worked through for its causes (equipment failure, operator error, disturbance), its consequences for safety and performance, and the safeguards already in place, with recommendations where a gap shows.

The cyber extension slots into the same step. "Incorrect level measurement" need not come from a failing LT004—it can come from someone altering the signal. "No response from the actuator" can be a communication fault, or an attack on the control network that strands the command. Reading each deviation with that second possibility in mind lets a single HAZOP hold both ordinary process hazards and cyber ones, which is what makes it fit a plant where the digital and physical sides are inseparable. Table 4.1 lists the standard guidewords and the kind of deviation each produces.

Table 4.1: Standard HAZOP guidewords and their application to process parameters [1].

Guideword	**Meaning**	**Example deviation**
No / None	Complete negation of the design intent	No flow
More	Quantitative increase	More pressure, more flow
Less	Quantitative decrease	Less flow, low level
As well as	Qualitative increase (additional effect)	Contamination, extra phase
Part of	Qualitative decrease (partial effect)	Wrong concentration
Reverse	Logical opposite of the intent	Reverse flow
Other than	Complete substitution	Wrong material

4.3 Fault Tree Analysis (FTA)

Fault Tree Analysis is the deductive counterpart to HAZOP [2]. It starts from a single undesired event—the top event—and reasons backward to the combinations of causes that could produce it. Where HAZOP sweeps forward across deviations, FTA drills down from one outcome.

The top event is defined first: in a process plant, something like overpressure, a shutdown, or loss of containment. From there the tree is developed downward, identifying the intermediate and basic events that could lead to it.

The events are joined by logic gates, mainly AND and OR. An OR gate fires if any one input occurs; an AND gate fires only if all of them occur together. That distinction is what the whole method turns on, because it separates a system with redundancy from one without: an outcome behind an AND gate needs several things to fail at once, while one behind an OR gate needs only the weakest link.

At the bottom sit the basic events—component failures, human errors, environmental conditions—each requiring no further breakdown. When failure rates are attached to them, the tree yields a probability for the top event, which is what makes FTA quantitative as well as explanatory.

The cyber extension adds new basic events rather than a new structure. A top event like "failure to maintain safe pressure" can be fed not only by a mechanical fault but by a manipulated control signal, an unauthorised logic change, or a disrupted network link, each entering through its own gate. Tracing those through the tree shows how an attack reaches a physical hazard and, just as usefully, which components carry the most

weight toward the top event—the natural targets for mitigation. Table 4.2 summarises the elements used to build the tree.

Table 4.2: Main elements used in Fault Tree Analysis [2].

Element	**Meaning**
Top event	The undesired system-level event being analyzed (e.g. separator overpressure).
Intermediate event	A fault developed further through lower-level gates.
Basic event	A primary fault requiring no further development (component failure, cyber action).
OR gate	The output event occurs if *any* input event occurs.
AND gate	The output event occurs only if *all* input events occur simultaneously.

4.4 Comparison and Integration of HAZOP and FTA

HAZOP and FTA pull in opposite directions, and that is precisely why they pair well. One works forward from causes to consequences, the other backward from a consequence to its causes; using only one leaves a blind spot the other covers.

HAZOP's strength is breadth. Applying guidewords across every node turns up a wide range of deviations, including ones nobody set out looking for, which is what makes it so useful early in design. Its weakness is that it lists causes without modelling how they combine—it tells you what can go wrong, not which coincidences matter.

FTA has the opposite profile. It models the logical combination of causes precisely, and can put numbers on them, but only for a top event chosen in advance—it goes deep on a question it has to be handed. On its own it can miss a hazard that was never nominated as a top event.

Used together, each supplies what the other lacks. HAZOP does the wide search and surfaces the deviations and their cyber-related causes—data manipulation, a communication failure, a spoofed reading. The serious ones are then promoted to top events and developed as fault trees, where the logic that links them is made explicit and the dangerous combinations stand out. Breadth from one, depth from the other.

For a cyber–physical plant this pairing is more than convenient. It is what lets the analysis hold the full range of hazards and the structure underneath the critical ones at the same time, covering the physical and digital sides together—and in the VS003 case study it is exactly this combination that exposes the shared-signal common cause, a dependency that neither method, used alone, would have made obvious.

Table 4.3 sets the two methods side by side and shows where each is used in this thesis.

Table 4.3: Complementary characteristics of HAZOP and FTA, and the role of each in this thesis.

	HAZOP	**FTA**
Direction	Forward: from causes to consequences	Backward: from a top event to its causes
Reasoning	Inductive, exploratory	Deductive, structured
Starting point	Process nodes and parameters	A single predefined top event
Nature	Qualitative	Logical; quantitative when failure data exist
Main output	A broad list of deviations, causes, consequences and safeguards	A logical model of how causes combine
Strength	Breadth—surfaces hazards not anticipated in advance	Depth—exposes combinations and common causes
Limitation	Does not model how causes combine	Only as good as the chosen top event
Role here	Wide sweep of cyber and physical deviations on VS003	Model of separator overpressure, including the spoofed-PT003 common cause

Chapter 5

Vulnerabilities and Cyber Attacks in Industrial Control Systems

5.1 Sources of Vulnerability in Industrial Control Systems

A control system is what makes a modern chemical plant run, and it is also where the plant is most exposed. The vulnerabilities are not exotic; they come from three ordinary facts—equipment designed before security mattered, kit that stays in service for decades, and networks that used to be isolated and no longer are [5, 10].

The deepest of these is design history. Most control equipment was built for a world where the network was trusted because it was physically closed off, so protections that are routine in IT—encryption, authentication, access control—were never added to the protocols, the PLCs, or the supervisory software [4]. Connect that equipment to anything, and the missing protections become the opening.

Lifetime makes it worse. A PLC may run for twenty years, and taking it down to patch or replace costs production, so devices with known weaknesses stay in service long after a fix exists.

Connectivity is the modern accelerant. Control networks now reach corporate IT and remote-access services for monitoring and maintenance; each of those links is useful, and each is also a way in.

And then there are people. A weak password, a misconfigured device, an access right that was never revoked—attackers exploit these at least as often as a technical flaw, because they are easier. On top of it all sits sheer complexity: a system with this many interacting parts hides dependencies nobody fully maps, and a hidden dependency is a vulnerability waiting to be found. That is exactly what the VS003 analysis later uncovers in the shared pressure signal.

5.2 Vulnerabilities in PLC, SCADA, and HMI Systems

Three components carry the control function—the PLC, the SCADA system, and the HMI—and each is exposed in its own way [5].

The PLC is the one wired to the process; in VS003 it is what reads PT003 and LT004

and drives the valves. It is built for reliability and real-time response, not security, and it shows: many PLCs accept logic changes from anyone who can reach them on the network, and firmware fixes are rarely applied on schedule. An attacker who gets that far can rewrite the control program, stall the process, or move the plant deliberately.

SCADA sits above it, gathering data from the controllers and presenting it to the operator. Because it is the system most likely to be networked outward, it is the most exposed: weak authentication, plaintext protocols, or a flat network turn a SCADA server into a single point from which an attacker can watch everything and command everything.

The HMI is where operator and system meet—process values, alarms, manual adjustments. Compromise it through stale software, missing access control, or misconfiguration, and the attacker controls what the operator sees. Falsified readings, suppressed alarms, or injected commands do not just affect the machine; they push the human into the wrong decision, which is precisely what happened at Oldsmar.

The components are also each other's weak points. A compromised HMI is a route to the PLCs; a compromised SCADA server reaches every device at once. Those interdependencies are what turn a single foothold into a path through the whole architecture—and they are why the security of these three decides the security of the plant.

5.3 Types of Cyber Attacks in Industrial Control Systems

The attacks that reach control systems vary in method and aim, but each leans on a weakness in a component, a protocol, or a person [11].

The simplest is unauthorised access—through a weak or stolen credential, or an exposed remote-access service. Once inside, an attacker maps the network, reads what they should not, and changes control parameters. A denial-of-service attack takes the opposite tack: instead of reading the system it floods or disables it, and in a plant that can mean a controller losing contact with its field devices and the process running blind.

Malware is the heavier tool, from ordinary ransomware to code written for control systems specifically [12]; it arrives on a USB stick, in an update, or over a connection, then spreads, disrupts, or in the worst case manipulates the process directly. Data-manipulation and spoofing attacks are subtler and, for a plant, more dangerous: a falsified reading or an injected command can steer the process while every screen still looks normal. Man-in-the-middle attacks do the same by sitting on the wire between components and altering what passes. Behind all of these sit insider threats—an authorised person, careless or malicious, who already holds the access an outsider has to fight for.

Table 5.1 maps these classes onto the MITRE ATT&CK for ICS framework, the common vocabulary for describing exactly how each is carried out.

Table 5.1: Attack classes affecting the control system, mapped to MITRE ATT&CK for ICS techniques [3]. Technique identifiers verified against attack.mitre.org.

Attack class	**ATT&CK for ICS technique**	**ID**
Unauthorized command injection	Unauthorized Message: Command Message	T1692.001
Sensor-data spoofing / falsification	Unauthorized Message: Reporting Message	T1692.002
Modification of setpoints/parameters	Modify Parameter	T0836
Process manipulation	Manipulation of Control	T0831
Disabling / overriding safety	Loss of Protection	T0837
Remote-access entry point	Remote Services	T0886
HMI manipulation	Graphical User Interface	T0823

5.4 Impact of Cyber Attacks on Physical Processes

What sets an ICS attack apart is that it does not stop at the screen—it reaches the process. That is the line between industrial cybersecurity and ordinary IT security, where a breach costs data or uptime but never containment.

In a chemical plant the control system holds variables like pressure, temperature, flow, and level inside hard limits. An attack that alters logic, falsifies a reading, or cuts communication can push any of them past those limits. Falsify a sensor and the controller acts on a lie—raising pressure that is already high; change a parameter and an automatic trip may simply not fire when it should. Cut the controller-to-actuator link and the process stops responding to its own conditions, which in a separator means losing level or pressure control with nothing acting to recover it.

The serious scenarios combine these. Pair data manipulation with a denial-of-service and the attacker both misleads the operator and blocks the correction—the coordinated form that made Stuxnet and the Ukraine grid attacks so effective. The end results are physical: equipment damage, downtime, environmental release, danger to people, and in the worst cases a major accident [8]. This is the whole reason cyber causes belong inside the process-safety analysis—and it is the lens the VS003 case study applies next.

5.5 Notable Real-World ICS Cyber Incidents

The risks discussed in this chapter are not hypothetical. Over the past fifteen years, several cyber attacks have targeted industrial control systems and, in a number of cases, caused physical damage. The most significant incidents are described below, each accompanied by a diagram of its attack lifecycle that highlights how it was carried out and the measures that could have prevented it.

Stuxnet (2010). Stuxnet is widely regarded as the first malware designed to cause physical damage [12]. It spread through removable media and exploited several previously unknown vulnerabilities to reach the Siemens programmable logic controllers that governed uranium-enrichment centrifuges, altering their rotational speed while reporting normal values to the operators. The incident demonstrated that even an air-gapped system is not immune, and it motivated controls on removable media, integrity checking of controller logic, and behaviour-based anomaly detection.

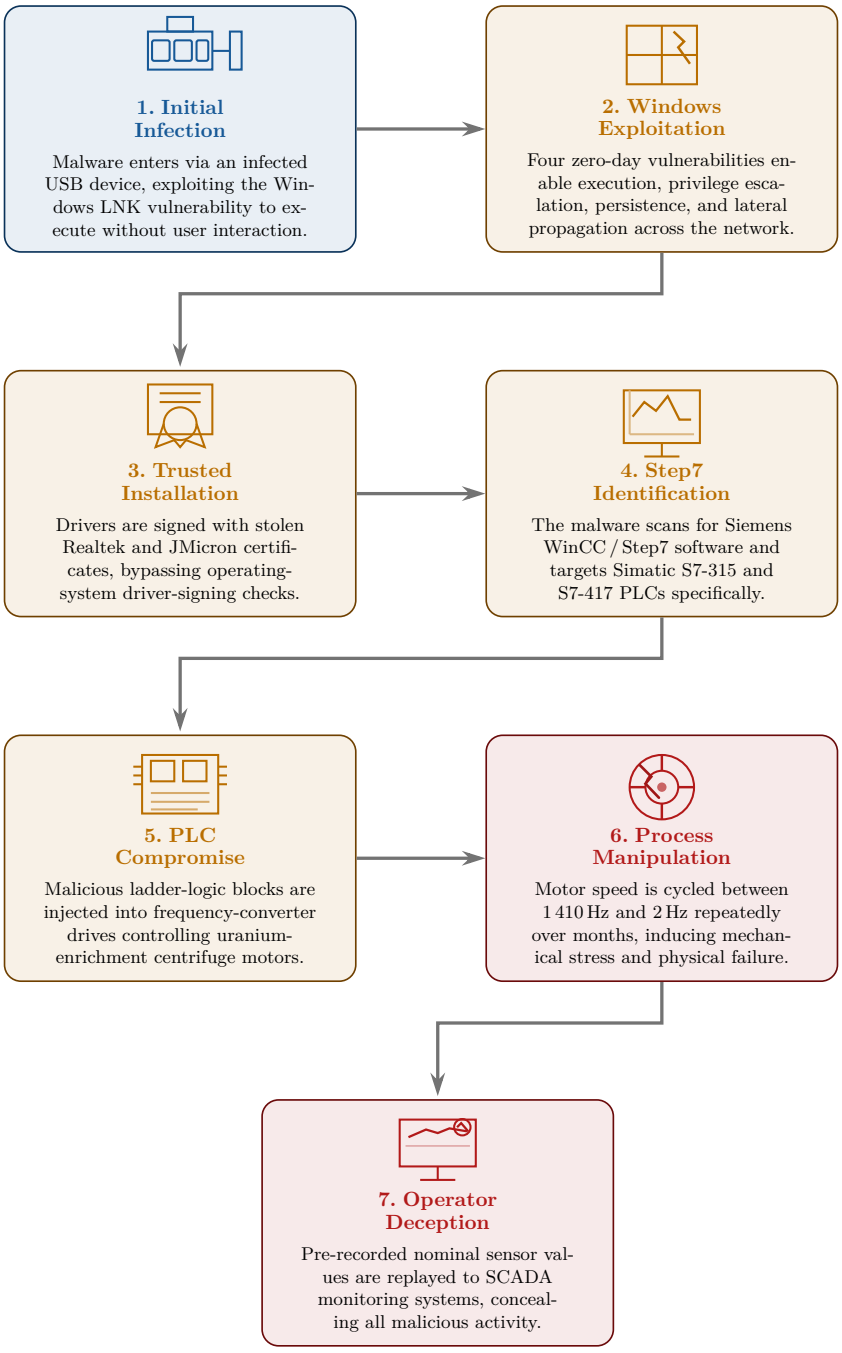


Figure 5.1: **Technical attack lifecycle of Stuxnet (2010).**

German steel mill (2014). Attackers used spear-phishing to obtain credentials to the plant's office network and then moved laterally into the production network [13]. They interfered with the control system in such a way that a blast furnace could not be shut down in a controlled manner, causing massive physical damage. This case underlines the importance of strict separation between IT and OT networks and of keeping safety-shutdown functions independent of the control path that may be compromised.

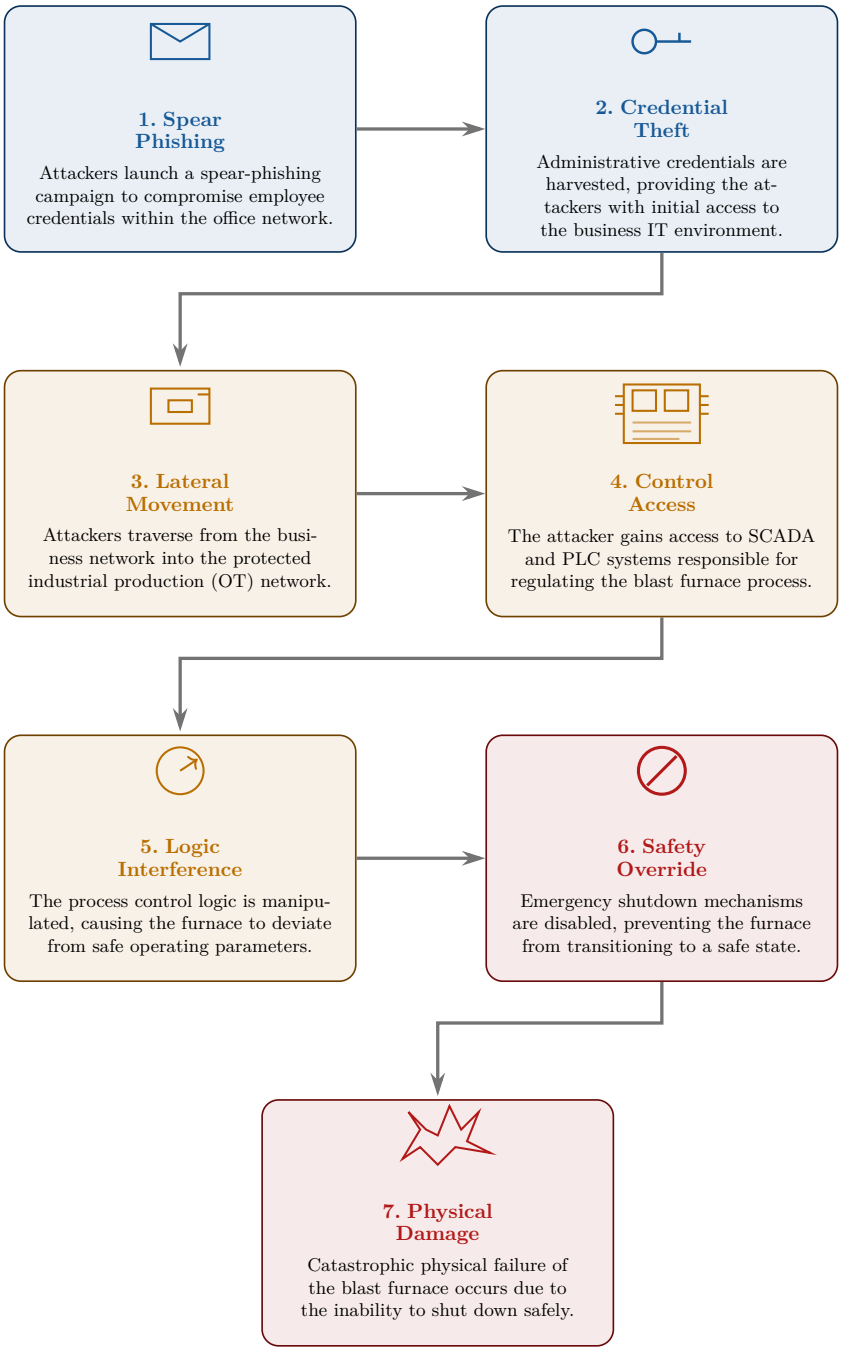


Figure 5.2: **Technical attack lifecycle of the German Steel Mill incident (2014).** The attack highlights the dangers of insufficient network segmentation between IT and OT environments and the compromise of safety-critical systems.

Ukraine power grid (2015 and 2016). In December 2015, the BlackEnergy malware, delivered through phishing, allowed attackers to steal operator credentials and remotely open circuit breakers, leaving around 230,000 people without electricity; a wiper component then erased affected systems [11]. A year later, the Industroyer (CrashOverride) malware automated a comparable attack by abusing the IEC-104 control protocol [14]. These incidents highlight the need for multi-factor authentication on remote access, network segmentation, and the ability to operate the process manually.

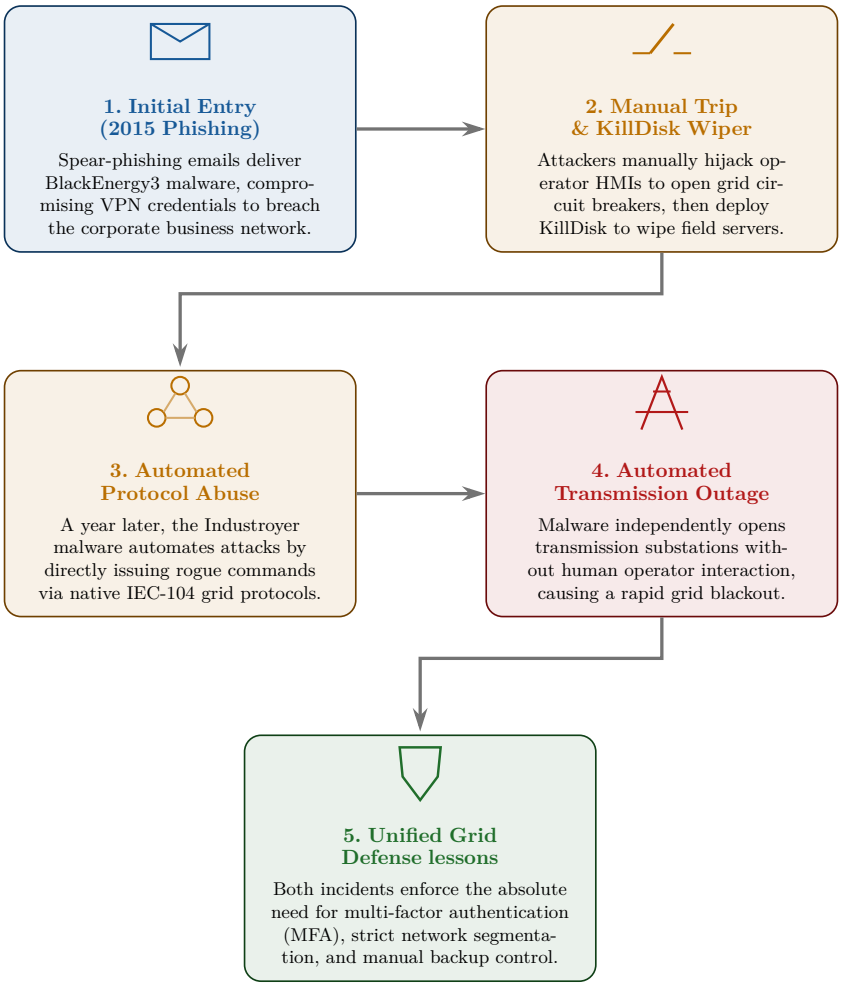


Figure 5.3: **Evolutionary attack lifecycle of the Ukraine Power Grid incidents (2015 vs. 2016).** The layout contrasts the transition from manual, credential-driven infrastructure manipulation in 2015 to automated, protocol-specific malware architectures in 2016.

TRITON / TRISIS (2017). TRITON, also called TRISIS, was the first malware specifically designed to target a safety instrumented system [15]. It reached the Triconex SIS controllers of a petrochemical plant and attempted to reprogram their safety logic; the attack was discovered when it inadvertently triggered a plant trip. Because the VS003 system studied in this thesis relies on an independent SIS, this incident is directly relevant: it shows why the SIS must reside in its own security zone, why its controllers should be left in the RUN (non-programmable) position during operation, and why engineering access to the safety logic must be tightly controlled and monitored.

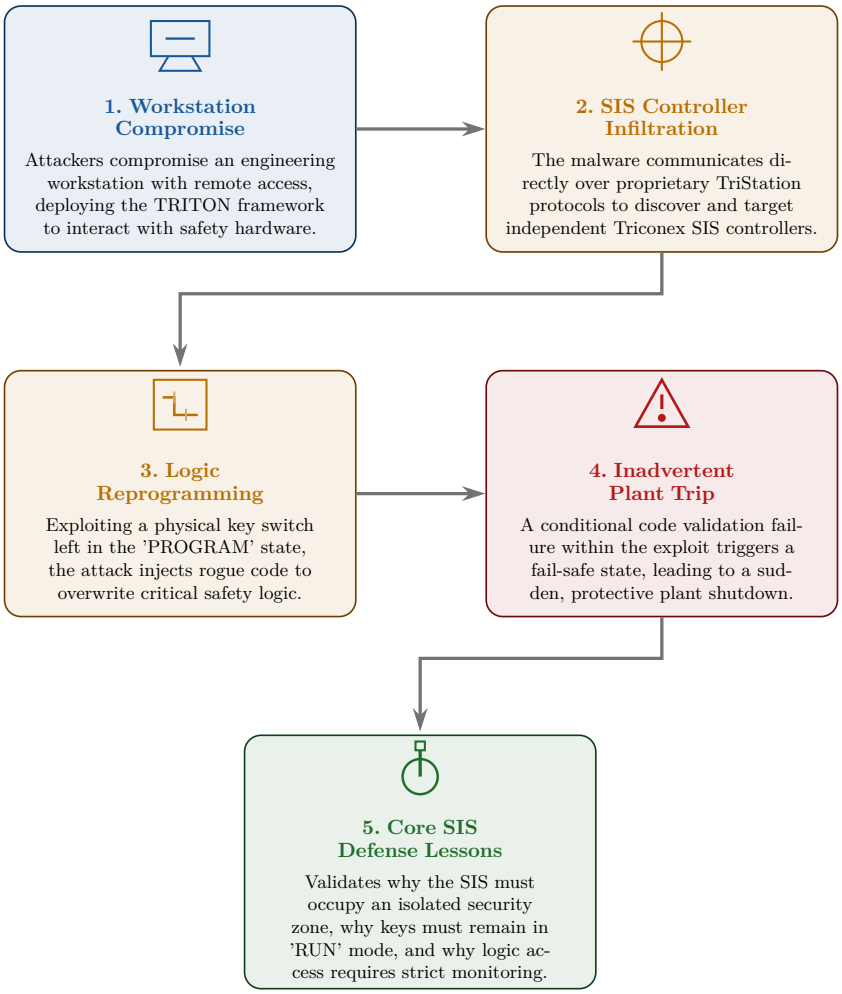


Figure 5.4: **Technical attack lifecycle of the TRITON / TRISIS incident (2017).** The diagram outlines the escalation sequence from network exploitation to industrial safety controller reprogramming, accentuating why physical key positioning and decoupled safety zoning remain critical structural defenses.

Oldsmar water-treatment plant (2021). At a water-treatment facility in Oldsmar, Florida, an intruder gained access to an operator interface through an unsecured remote-access tool and tried to raise the dosing of sodium hydroxide to dangerous levels [16]. An operator noticed the change and reverted it before any harm occurred. The case shows how unsecured remote access to a chemical-dosing process can directly threaten safety, and why such access must be removed or protected with multi-factor authentication and supported by alarms and operator supervision.

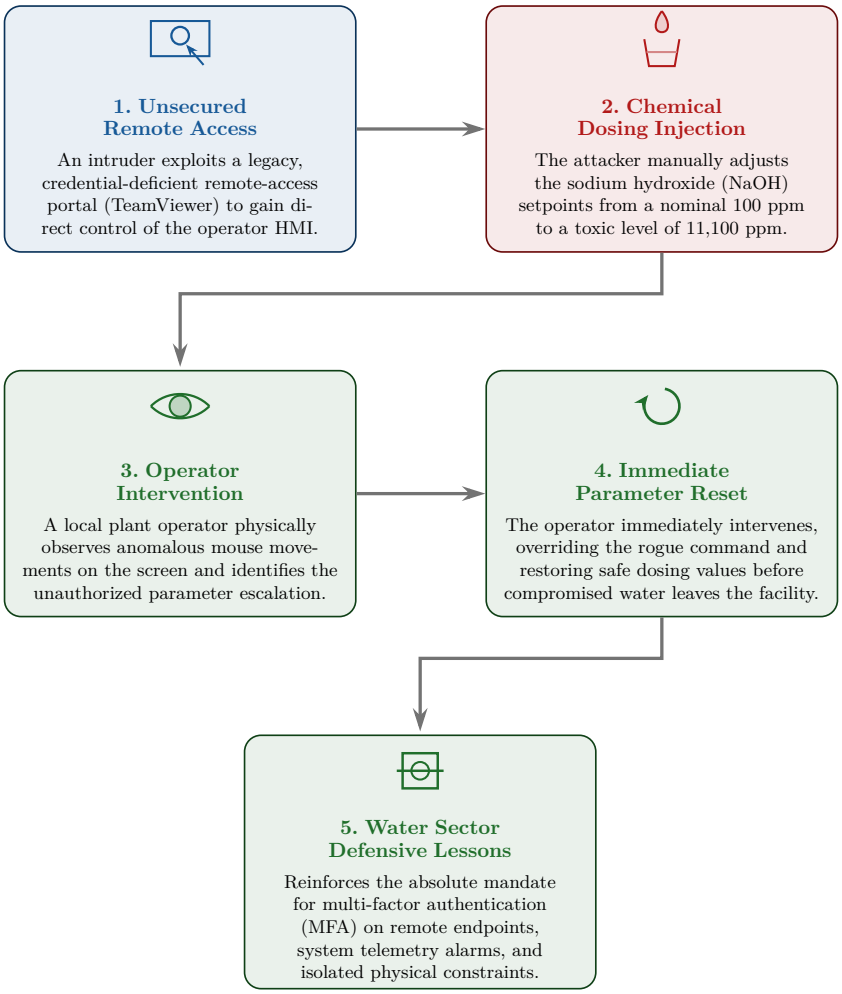


Figure 5.5: **Technical attack lifecycle of the Oldsmar Water-Treatment Plant incident (2021).** The diagram maps the infiltration vector from unauthenticated remote boundary exposure to manual dosing alteration, validating the role of live operator validation and chemical process constraints.

Colonial Pipeline (2021). The Colonial Pipeline incident was a ransomware attack on the company's business network, reportedly entering through a virtual private network account that lacked multi-factor authentication [8]. Although the control system itself was not manipulated, the operator shut the pipeline down as a precaution, causing widespread fuel shortages. It illustrates how an attack confined to the IT side can still halt a physical process, reinforcing the value of segmentation, multi-factor authentication, and tested backups.

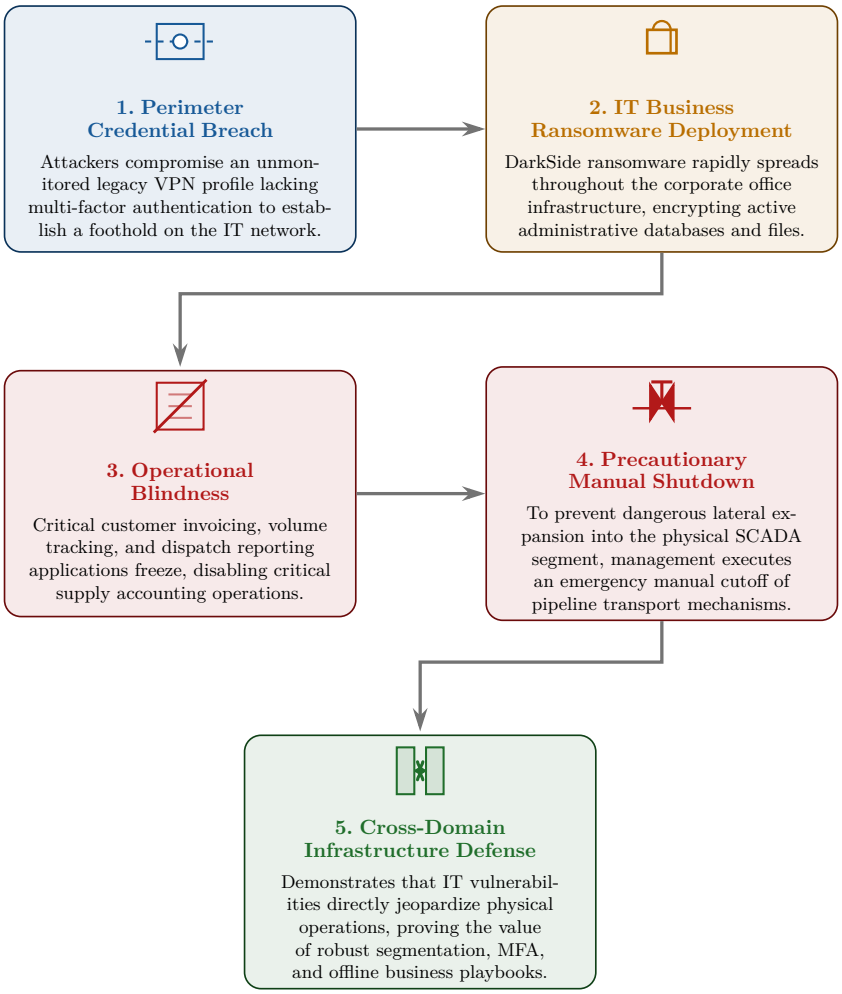


Figure 5.6: **Technical threat lifecycle of the Colonial Pipeline incident (2021).** The schematic highlights how structural disruption confined entirely to enterprise IT networks can induce complete administrative blindness, triggering a precautionary operational physical shutdown.

The six incidents are drawn together in Table 5.2, which records what each one ultimately exploited and the control most directly relevant to the VS003 system.

Table 5.2: Summary of the ICS cyber incidents and the protection measure most relevant to the VS003 system.

Incident (year)	**What it exploited**	**Most relevant control for VS003**
Stuxnet (2010)	Removable media into an air-gapped network; covert PLC logic change	Removable-media control; controller-logic integrity checks; anomaly detection
German steel mill (2014)	Phishing into IT, lateral move into OT; blocked safe shutdown	Strict IT/OT separation; safe shutdown independent of the control path
Ukraine grid (2015–2016)	Stolen remote credentials; breaker operation; protocol abuse	Multi-factor remote access; segmentation; manual fallback operation
TRITON / TRISIS (2017)	Reprogramming of a safety instrumented system	SIS in its own zone; key-switch in RUN; controlled, monitored engineering access
Oldsmar (2021)	Unsecured remote access to an operator HMI; dosing change	Remove or secure remote access with MFA; alarms and operator supervision
Colonial Pipeline (2021)	Ransomware on IT via a VPN account without MFA	MFA on all remote access; IT/OT segmentation; tested offline backups

Chapter 6

Prevention and Protection Measures

6.1 A Layered Approach to Cybersecurity Protection

Protecting a control system is not the same job as protecting an IT network, and that difference shapes every measure in this chapter. The plant cannot trade away availability or real-time response for security, so a control that protects the process by interrupting it is no protection at all. The measures have to fit around the process, not the other way round.

The organising idea is defence in depth [4]: not one strong barrier but several independent layers, so that defeating one still leaves the attacker facing the next, and no single layer is trusted to hold alone. Figure 6.1 collects the layers proposed for VS003 and groups them by their role in the NIST functions—protect, detect, respond, recover.

Those layers fall into three kinds. Technical controls secure the components and the network: authentication, encryption, segmentation, intrusion detection. Organisational controls are the policies, procedures, and training that decide how people use the system. Physical controls keep hands off the hardware in the first place. The sections that follow take the technical ones in turn—segmentation, then access control, then monitoring and response—and close with the standards that tie them into one architecture.

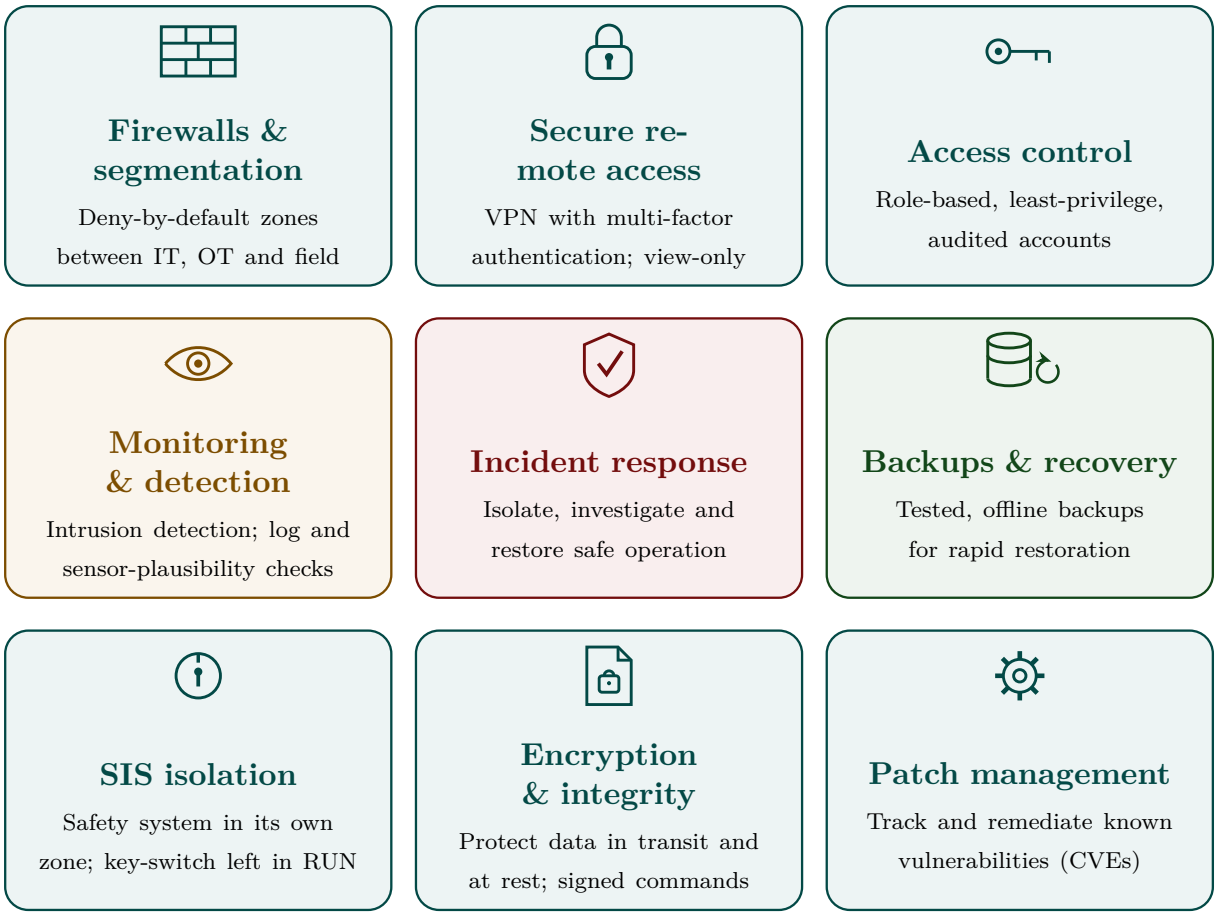


Figure 6.1: Defense-in-depth: the layered protection and prevention measures applied to the VS003 system. No single control is sufficient; together they reduce both the likelihood and the impact of a cyber attack. The boxes are coloured by their role in the NIST cybersecurity framework: protect (teal), detect (amber), respond (red) and recover (green).

6.2 Network Segmentation and Secure Architecture

Segmentation is the most effective single defence a control system has, and the reasoning is plain: most of the damage in an incident comes from the attacker moving sideways after getting in, and segmentation is what stops the movement. Split the network into zones by function and exposure, and a foothold in one stops being a foothold in all.

In practice the split follows the layered architecture. The OT network—controllers, sensors, actuators—is separated from corporate IT by firewalls, secure gateways, and a DMZ that is the only controlled crossing between them. IEC 62443 gives this its vocabulary: assets with the same security needs sit in a *zone*, and every path between zones is a *conduit* that can be filtered and watched [4].

The segmentation is layered, not single. Enterprise is split from control; within control, supervisory is split from the field; and the most critical components sit in the most restricted zones. In VS003 the safety instrumented system is kept in its own zone, apart from the basic process control system, with its engineering key-switch left in the RUN position during operation so the safety logic cannot be remotely reprogrammed—the direct lesson of the TRITON incident from the previous chapter.

Firewalls enforce the zone rules, passing only the traffic explicitly allowed, while in-

trusion detection watches the conduits for anything that should not be on them. Segmentation also buys least privilege almost for free: restrict the communication paths to what the process actually needs and the attack surface shrinks with them. Remote access, where maintenance requires it, comes in the same controlled way—through a VPN with multi-factor authentication—rather than around the side. The payoff is twofold: an intrusion is both less likely to spread and, when it does, held to the zone it began in. This is what keeps the VS003 field devices reachable only from inside the control room, never from the enterprise side or the internet.

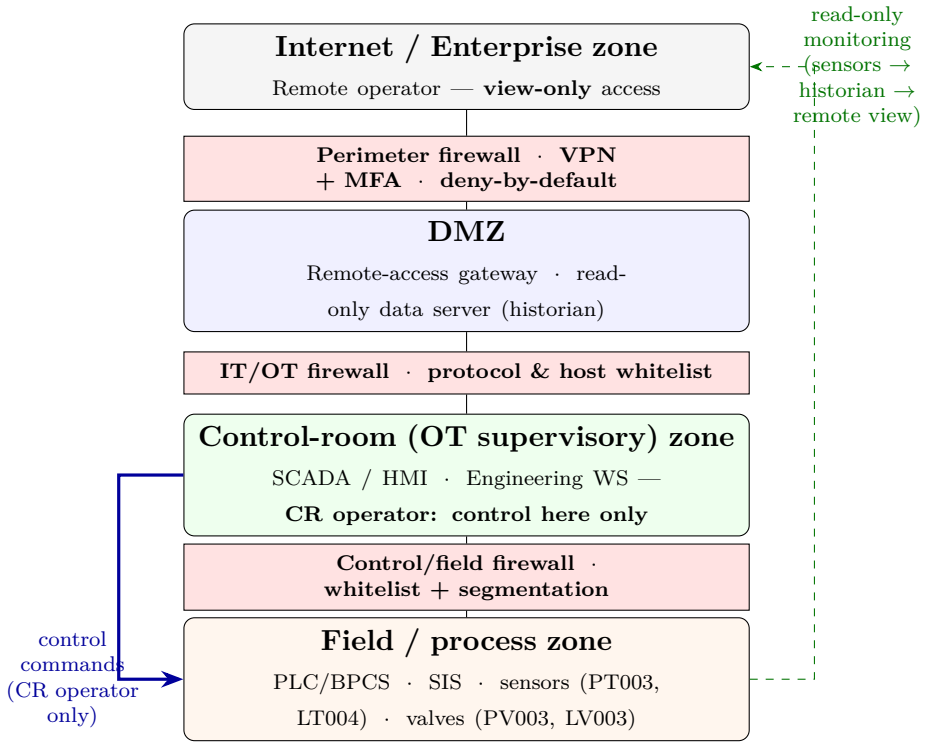


Figure 6.2: Secure network architecture for the VS003 system. Zones are separated by firewalls with deny-by-default rules and host/protocol whitelists; remote access enters through a VPN with multi-factor authentication into a DMZ. Monitoring data flows upward (read-only) so the remote operator can *view* sensor status, while control commands originate only in the control-room zone and flow down to the field — there is no command path from the remote/DMZ side into the process.

6.3 Access Control and Authentication Mechanisms

Access control and authentication decide who and what may touch the system, and in a plant where that touch moves valves, the decision is a safety matter as much as a security one [5].

Authentication is the first question: is this user or device who it claims to be? Username-and-password is still the norm in industrial systems, and it is also the weak link, since reused or guessable credentials are among the most common ways in. Multi-factor authentication closes much of that gap by demanding a second proof—a one-time code, a biometric—and it is now the baseline for any access that reaches the control network, including the remote access of the previous section.

Authorisation is the second question: what may this verified user actually do? Role-based access control answers it by tying permissions to a role—operator, engineer, administrator—so each person holds what the job needs and no more. An operator has no reason to reprogram a controller, and under least privilege simply cannot.

Devices deserve the same scrutiny as people. Only trusted devices should be allowed to communicate, with digital certificates proving identity, so a rogue device cannot just join the network and start injecting commands. Identity alone is not enough, though: the traffic itself should be protected—sensitive data encrypted, critical commands signed—so that a tampered or spoofed message, the very thing the VS003 attack relies on, is rejected before it reaches a controller.

None of this holds without credential hygiene. Passwords have to be governed, rotated, and—most easily forgotten—revoked the moment someone changes role or leaves, because a live credential for a job no longer held is a door left open. And since even legitimate users can do harm, by accident or intent, activity is logged, duties separated, and audit trails kept. Together these keep every other layer's assumption about *who is acting* actually true.

6.4 Monitoring, Detection, and Incident Response

Prevention is never complete, so a control system also has to see an attack in progress and act on it. Monitoring and detection are the eyes; incident response is the hands.

Monitoring in a plant has an advantage IT lacks: there are two independent things to watch. On the cyber side, unusual traffic or access attempts; on the physical side, pressure, level, or flow drifting in ways the process should not produce on its own. Correlating the two is what makes it powerful—a command that looks valid on the network but drives a physical variable the wrong way is precisely the signature of the VS003 spoofing scenario, and neither view alone would catch it.

Intrusion detection systems do the network watching, either by signature, matching known attack patterns, or by anomaly, flagging departures from a learned baseline. For control systems the anomaly approach matters most, because the threats that reach a plant are often novel enough that no signature exists yet. Logging sits under both: a record of actions and events that feeds both the real-time alert and the later investigation.

Detection is only worth anything if a response follows. A structured sequence—identify, contain, eradicate, recover [9]—gives that response shape, but in a plant every step carries a constraint the IT version does not: it must not itself create an unsafe state. Isolating a compromised controller can be the right move and still be the wrong one if it drops the process into an uncontrolled condition, so containment has to be planned against the physics, not just the network. That planning is as much organisational as technical—operators, engineers, and security staff working to one agreed plan—and it is what turns a detected incident into a contained one rather than a worse one.

6.5 Standards and Best Practices for Industrial Cybersecurity

Security measures hold together better when they follow a recognised framework than when they are chosen ad hoc, and two frameworks in particular shape industrial practice.

IEC 62443 is the standard written specifically for industrial automation and control [4]. Its central idea is the one segmentation already used: divide the system into zones and conduits, then assign each a security level according to the strength of attacker it must withstand. Those levels, set out in Table 6.1, turn a vague intention to be secure into a concrete target a zone either meets or does not.

The NIST framework supplies the other axis [9]. Where IEC 62443 structures the architecture, NIST structures the lifecycle, through five functions—identify, protect, detect, respond, recover—that recur across this chapter and organise the defence-in-depth figure. Taken together, one says how to build the zones and the other how to keep them defended over time rather than secured once and forgotten.

Best practices in industrial cybersecurity also emphasize the importance of defense-in-depth, regular risk assessment, and continuous monitoring. Organizations are encouraged to conduct periodic security evaluations, update system configurations, and apply patches where possible. Patch and vulnerability management deserves particular attention: known vulnerabilities — for example those catalogued as CVEs — should be tracked, assessed for their relevance to the installed devices, and remediated through tested updates during scheduled maintenance windows, since unpatched controllers and engineering workstations remain one of the most common entry points for attackers. In addition, training and awareness programs are essential to ensure that personnel understand their role in maintaining system security.

Compliance with standards and best practices not only improves system security but also supports regulatory requirements and industry certification processes. It demonstrates that the organization has implemented appropriate measures to manage cybersecurity risks and protect critical infrastructure.

For VS003, then, IEC 62443 and the NIST functions are what make the measures in this chapter a coherent architecture rather than a list: the first decides how the zones are drawn and how strongly each must be defended, the second keeps that defence under continuous review.

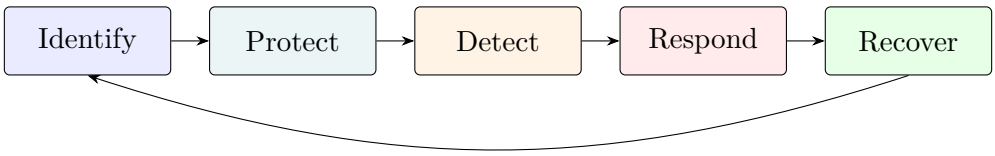


Figure 6.3: The core functions of the NIST Cybersecurity Framework, applied as a continuous improvement cycle to the industrial control system.

Table 6.1: Security Levels defined by IEC 62443, based on the capability of the threat they are designed to resist [4].

Level	**Protection against**
SL 1	Casual or coincidental violation.
SL 2	Intentional violation using simple means, with low resources, generic skills, and low motivation.
SL 3	Intentional violation using sophisticated means, with moderate resources, IACS-specific skills, and moderate motivation.
SL 4	Intentional violation using sophisticated means, with extended resources, IACS-specific skills, and high motivation.

Table 6.2 brings the chapter together: it lists the nine layers of Figure 6.1, the NIST Cybersecurity Framework function each one serves, and what it is intended to counter on the VS003 system.

Table 6.2: The nine defence-in-depth layers proposed for VS003, mapped to their NIST Cybersecurity Framework function and to the threat each addresses.

Layer	**NIST function**	**What it counters on VS003**
Firewalls & segmentation	Protect	Lateral movement from IT/enterprise toward the controllers
Secure remote access (VPN + MFA)	Protect	Credential-based entry through remote services (cf. Oldsmar, Colonial)
Role-based access control	Protect	Misuse of privilege; unnecessary access to control logic
Encryption & integrity	Protect	Spoofed or tampered commands reaching a controller
SIS isolation (key-switch in RUN)	Protect	Remote reprogramming of the safety logic (cf. TRITON)
Patch & vulnerability management	Protect	Exploitation of known CVEs on controllers and workstations
Monitoring & intrusion detection	Detect	A spoofed signal whose physical effect contradicts the network view
Incident response	Respond	Uncontained spread; unsafe ad-hoc reactions during an attack
Tested offline backups	Recover	Ransomware or wiper damage to engineering and supervisory systems

Chapter 7

Case Study: VS003 Separator System

7.1 System Description and Architecture

The VS003 separator system is a representative example of a cyber–physical industrial process used in chemical plants [5]. Its primary function is to separate a mixture of fluids into different phases, typically gas, oil, and water, based on differences in density and physical properties. This process is essential for ensuring the quality and stability of downstream operations.

The operation of the separator relies on maintaining specific process conditions, particularly pressure and liquid level. These variables are measured continuously: a pressure transmitter (PT003) monitors the pressure inside the vessel, and a level transmitter (LT004) monitors the liquid level. Both provide real-time data to the control system.

The control system is centered on a programmable logic controller (PLC), which acts as the basic process control system (BPCS). The PLC receives the signals from PT003 and LT004 and executes the control logic that regulates the process. Based on the measured values, it sends commands to control valves: a gas-outlet valve (PV003) and a liquid-outlet valve (LV003). These valves adjust the flow of fluids leaving the separator, keeping the process variables within safe operating limits.

In addition to the basic control system, an independent safety instrumented system (SIS) provides an additional layer of protection. The SIS operates separately from the BPCS and is designed to bring the system to a safe state under critical conditions, for example by acting on a pressure relief valve (PSV) or an emergency shutdown valve when pressure or level exceeds safe thresholds.

Supervisory systems, namely SCADA and human–machine interfaces (HMIs), allow operators to monitor the status of the separator and interact with the control system. Operators can observe process variables, receive alarms, and intervene when necessary.

The architecture of the VS003 system follows the typical structure of industrial control systems, comprising field devices, control units, supervisory systems, and communication networks. These components are interconnected through industrial communication protocols, forming a cyber–physical system in which digital control directly influences physical behavior.

Understanding the structure and operation of the VS003 separator is essential for identifying potential hazards, vulnerabilities, and cyber risks. This system serves as the basis for applying the risk-assessment methodology developed in the previous chapters.

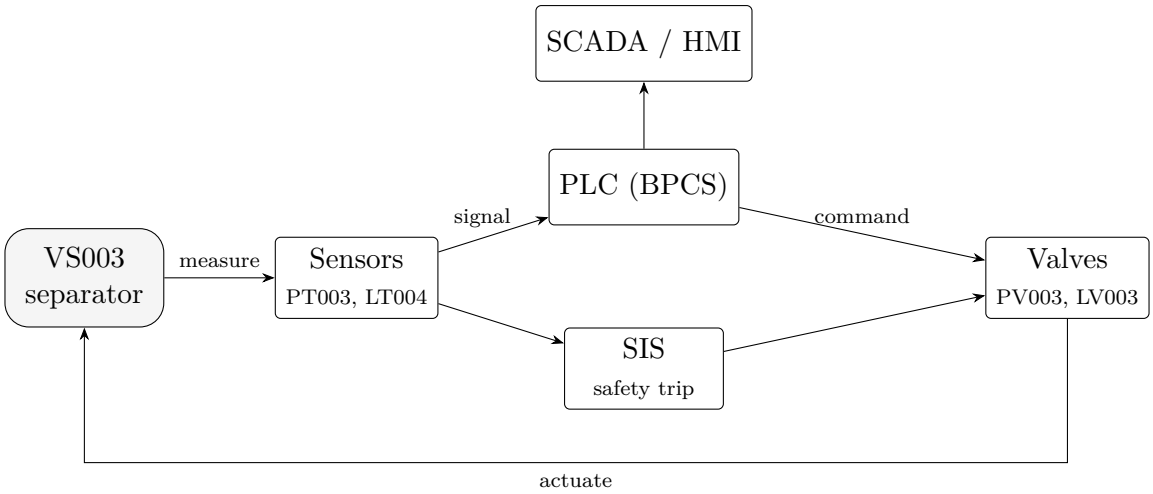


Figure 7.1: Control architecture of the VS003 separator: the basic process control system (PLC/BPCS) and the independent safety instrumented system (SIS) act on the gas (PV003) and liquid (LV003) valves using the pressure (PT003) and level (LT004) measurements.

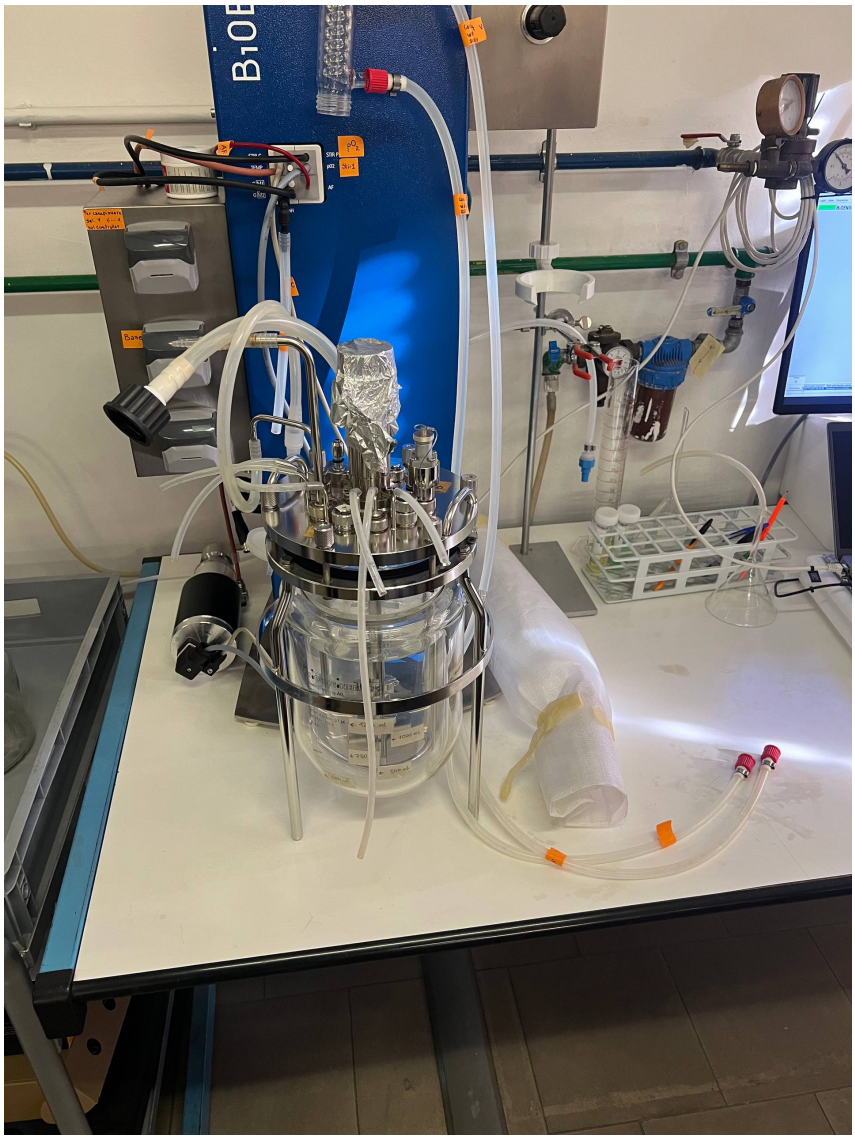


Figure 7.2: Laboratory-scale separation-process setup presented during a research meeting organized by Prof. Micaela Demichela. The setup belongs to a collaborating researcher and is reproduced with permission for illustrative and educational purposes. The image is included to support understanding of the process architecture discussed in this thesis and does not represent an experimental setup developed within the present work.

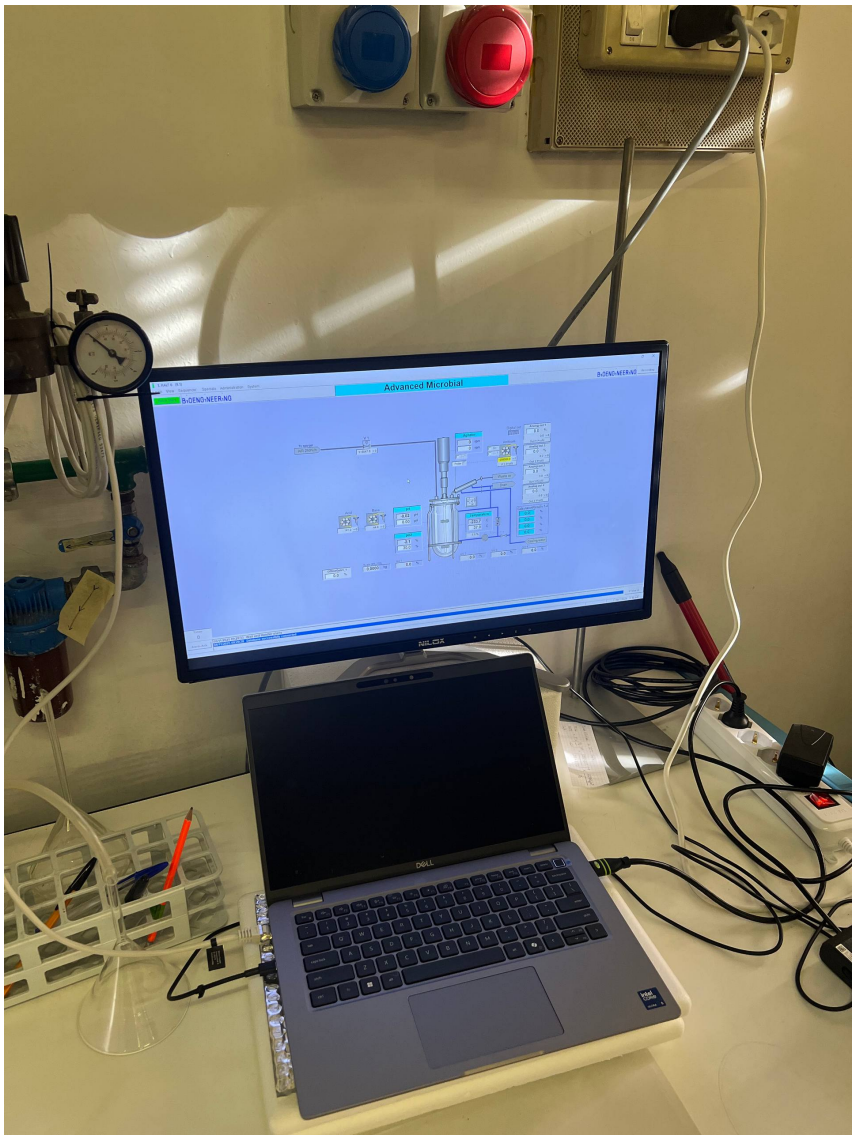


Figure 7.3: Process scheme presented during a research meeting organized by Prof. Micaela Demichela and reproduced with permission from the original researcher. The figure is included as a reference example to facilitate understanding of the process and does not represent a process design developed within this thesis.

7.2 Identification of Process Variables and Hazards

The safe operation of the VS003 separator depends on maintaining key process variables within predefined limits. The most critical variables are pressure and liquid level, since deviations in these parameters can lead to unsafe operating conditions or system failure.

Pressure within the separator must be carefully controlled to prevent overpressure conditions that could damage equipment or lead to loss of containment. Pressure is regulated through the control valves and monitored using the pressure transmitter PT003. If the pressure exceeds safe limits, safety mechanisms such as the relief valve or an emergency shutdown may be activated by the SIS.

Liquid-level control is equally important, as improper levels can affect the efficiency of phase separation and create hazardous situations. For example, an excessively high liquid

level may lead to carry-over into the gas outlet, while a low level may allow gas to enter the liquid outlet, disrupting downstream processes. The level transmitter LT004 provides continuous measurements, and the control valves adjust the inflow and outflow of fluids to maintain stable conditions.

Using a HAZOP-based approach, potential deviations in these variables can be systematically identified [1]. For pressure, deviations such as "more pressure" or "less pressure" are analyzed, while for liquid level, deviations such as "high level" or "low level" are considered. Each deviation is associated with possible causes, including equipment failure, process disturbances, and cyber-related events.

For instance, a "high pressure" deviation may result from a blocked outlet, a malfunctioning control valve, or incorrect control logic. In a cyber context, it may also be caused by falsified sensor data or unauthorized changes to controller parameters. Similarly, a "high level" condition may be caused by excessive inflow, valve malfunction, or manipulated level measurements.

The consequences of these deviations can be significant. High pressure may lead to mechanical stress or failure of the separator, while improper liquid levels can result in inefficient separation or process instability. In extreme cases, these conditions may escalate into safety incidents.

Identifying these variables and their associated hazards is a critical step in the risk-assessment process. It provides the foundation for analyzing vulnerabilities, modeling failure scenarios, and evaluating the impact of cyber threats on the physical system.

Table 7.1: Cyber-extended HAZOP worksheet for the VS003 separator (node: separator vessel).

Deviation	**Physical causes**	**Cyber causes**	**Consequence / safeguard**
Pressure — MORE	Blocked outlet; PV003 fails closed; logic error	Falsified low PT003 reading; PLC setpoint change; MITM; SIS bypass	Overpressure, rupture. *PSV; SIS high-P trip; redundant PT*
Pressure — LESS	Leak; valve stuck open	Falsified high reading → over-venting	Vacuum / implosion. *Low-P alarm*
Level — MORE	Excess inflow; LV003 fails closed; LT004 drift	Spoofed low-level → outflow reduced	Carry-over to gas line. *High-level alarm + SIS*
Level — LESS	Low inflow; valve stuck open	Spoofed high-level → outflow increased	Gas blow-by. *Low-level trip*

7.3 Cybersecurity Risk Analysis of the VS003 System

The VS003 separator, as a cyber–physical industrial process, is exposed to various cybersecurity risks [11] that can affect both its digital components and its physical behavior. These risks arise from vulnerabilities in system components, communication networks, and user interactions. A detailed analysis of these risks is essential to understand how cyber threats can influence process safety.

One of the primary targets in the system is the programmable logic controller, which executes the control logic. If an attacker gains unauthorized access to the PLC, it may be

possible to modify the control algorithms or setpoints. Such modifications could lead to incorrect control actions, for example increasing pressure beyond safe limits or disabling safety interlocks.

Sensors also represent a critical point of vulnerability. The pressure and level transmitters provide the data required for control decisions. If sensor data is manipulated or falsified, the control system operates on incorrect information. For instance, a falsified low-pressure reading may cause the system to increase pressure unnecessarily, potentially creating hazardous conditions.

The communication networks connecting the system components are another important attack surface. A man-in-the-middle attack could alter the communication between the PLC and the field devices, resulting in incorrect control actions without immediately alerting operators.

Supervisory systems, including the SCADA and HMI interfaces, provide high-level access to system operations. If compromised, they can be used to monitor process data, issue commands, or disable alarms. An attacker controlling the HMI may manipulate the displayed information, preventing operators from detecting abnormal conditions.

Based on these vulnerabilities, several attack scenarios can be defined. The first involves unauthorized modification of the PLC logic, leading to incorrect valve control and a pressure increase. The second involves spoofing of sensor data, causing the system to misinterpret process conditions. The third involves network-based attacks that disrupt communication or inject false commands. A particularly severe case occurs when a falsified PT003 reading is used both by the BPCS and by the SIS: the same spoofed signal can defeat the control loop and the safety layer simultaneously, acting as a common cause that a purely physical analysis would treat as two independent protections.

The impact of these scenarios depends on how the system responds to abnormal conditions. Where safety systems are bypassed or disabled, the consequences may include overpressure, overflow, or process instability, which in turn can lead to equipment damage, production loss, or safety hazards.

The cybersecurity risk analysis of the VS003 system demonstrates how digital vulnerabilities can translate into physical risks. It highlights the importance of treating cyber threats as part of the overall risk assessment and provides a basis for defining appropriate mitigation measures.

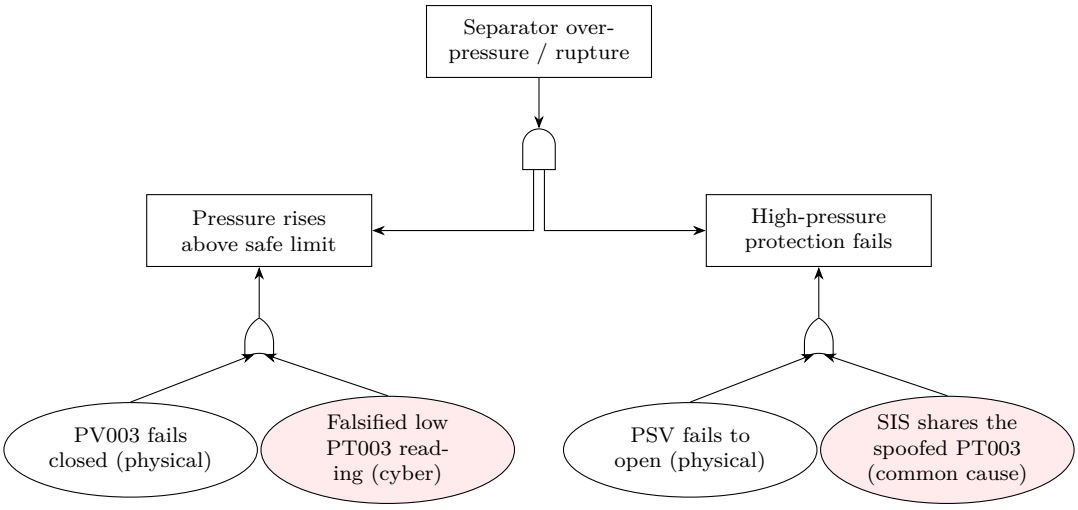


Figure 7.4: Fault tree for separator overpressure. The shaded cyber basic events appear alongside the classical physical faults; the right branch shows the common cause, where a spoofed PT003 signal defeats both the control and the safety layer.

7.4 Mitigation Measures for the VS003 System

Based on the identified vulnerabilities and potential attack scenarios, a set of mitigation measures can be proposed to enhance the security and resilience of the VS003 separator. These measures aim to reduce the likelihood of cyber attacks and limit their impact on the physical process.

One of the most important measures is network segmentation [4]. The control network of the VS003 system should be isolated from external and corporate networks using firewalls and secure gateways. Critical components, such as the PLC and the safety instrumented system, should be placed in protected zones with restricted access; in particular, the SIS should reside in its own zone so that a compromise of the BPCS cannot reach it. This reduces the risk of unauthorized access and limits the propagation of cyber threats within the system.

Access-control mechanisms should also be strengthened. Strong authentication methods, including multi-factor authentication, should be used for accessing the control systems and supervisory interfaces. Role-based access control can ensure that users have only the permissions necessary for their tasks. In addition, regular review and management of user accounts are essential to prevent unauthorized access.

To address vulnerabilities in the communication networks, secure communication protocols should be implemented where possible. Encryption and integrity checks can help protect data from interception and manipulation. Network monitoring tools, such as intrusion detection systems, should be deployed to identify suspicious activity and potential attacks in real time.

Protection of sensor data is particularly important in the VS003 system, as control decisions depend on accurate measurements. Redundant sensors and validation mechanisms can be used to detect inconsistencies; for example, comparing readings from multiple sensors, or cross-checking the BPCS and SIS measurements, can help identify falsified or abnormal values and counter the common-cause spoofing scenario described above.

The PLC should be protected against unauthorized modification by restricting access to the control logic and ensuring that firmware updates are securely managed. Regular

backups of the control programs should be maintained to allow rapid recovery in case of compromise.

Supervisory systems, including the SCADA and HMI interfaces, should be secured through proper configuration, regular software updates, and strict access-control policies. Alarm systems should be protected to ensure that operators receive accurate and timely information about system conditions.

In addition to technical measures, organizational practices play a key role in system security [9]. Personnel should be trained in cybersecurity awareness, and clear procedures should be established for incident response. Regular security assessments and audits can help identify weaknesses and ensure that protection measures remain effective.

By implementing these mitigation measures, the VS003 separator can achieve a higher level of resilience against cyber threats. These measures demonstrate how the principles discussed in previous chapters can be applied in a practical industrial context.

Table 7.2: Risk assessment and reduction for the VS003 system: likelihood (L), severity (S), inherent risk before mitigation, the mitigation applied, and the residual risk. Levels follow the matrix of Table 3.3.

Attack scenario	**L**	**S**	**Inherent risk**	**Mitigation applied**	**Residual risk**
PT003 pressure spoofing — common cause defeating BPCS and SIS	L3	S5	**Very High**	Independent/redundant transmitter; plausibility check; BPCS–SIS separation	Medium
Unauthorized setpoint or logic change on the PLC	L3	S4	High	Write-protection; role-based access control	Medium
LT004 level spoofing	L3	S4	High	Level cross-check; high/low-level alarm with SIS	Medium
Malicious command injection to a valve	L3	S4	High	Command authentication; network segmentation	Medium
Man-in-the-middle on the PLC–field link	L2	S4	High	Encryption and integrity checks; segmentation	Medium
Denial of service on the control network	L3	S3	High	Redundant communication; rate limiting; monitoring	Medium
HMI compromise (alarm suppression)	L3	S4	High	HMI hardening; multi-factor authentication	Medium
SIS safety-logic reprogramming (TRITON-style)	L2	S5	High	Separate safety zone; key-switch in RUN; strict engineering access	Low
Remote-access entry point without MFA	L3	S4	High	Remove or secure remote access; MFA; VPN hardening	Low

Chapter 8

LLM-Based Operator Assistant

8.1 Concept of AI-Assisted Monitoring in Industrial Control Systems

The increasing complexity of industrial control systems and the growing number of cyber threats make it more challenging for operators to detect abnormal conditions in real time. Traditional monitoring approaches rely on predefined alarms and rule-based systems, which may not always capture complex or subtle anomalies, especially those caused by sophisticated cyber attacks [5]. As a result, there is a need for more advanced tools that can support operators in understanding system behavior and responding to potential incidents.

Recent developments in artificial intelligence, particularly large language models (LLMs), offer new possibilities for enhancing monitoring and decision support in industrial environments [17]. These models can process large amounts of data, identify patterns, and generate human-readable explanations. When integrated into the supervisory layer of an industrial control system, they can assist operators by providing insights into system behavior and highlighting potential risks.

An LLM-based operator assistant can be designed to analyze multiple sources of information, including process data, alarm logs, and network activity. By correlating these inputs, the system can highlight unusual patterns that may indicate a cyber incident or an abnormal process condition [11]. For example, a sudden change in sensor readings combined with unusual network traffic may suggest a data-manipulation attack.

In addition to detection, the assistant can support decision-making by providing contextual information and recommendations [9]. Instead of presenting raw data, the system can generate explanations, such as identifying possible causes of an anomaly and suggesting appropriate actions. This can help operators respond more effectively, especially in complex situations where multiple factors are involved.

Another potential advantage of using LLMs is their ability to adapt to different scenarios. Unlike rule-based systems, which require predefined conditions, learning-based systems can be tuned on historical data and updated as conditions change. This makes them attractive for dynamic industrial environments.

However, the integration of AI into industrial control systems also introduces important challenges, including ensuring the reliability of the AI system, protecting it from manipulation (for example adversarial inputs or prompt-injection attacks), and maintaining transparency in its decision-making. For these reasons, the assistant is positioned as a decision-support tool that operates alongside, and outside of, the safety-critical control

loop: it informs the operator but does not directly actuate the process and does not replace the safety instrumented system. Because large language models are non-deterministic and may produce incorrect outputs, they must not be used as a safety function; the human operator remains responsible for all control decisions.

The concept of an LLM-based operator assistant therefore represents a forward-looking, human-centered approach to industrial cybersecurity. It combines advanced data analysis with a clear separation from safety-critical functions, aiming to improve situational awareness and to enhance the ability of operators to detect and respond to cyber threats. Figure 8.2 summarises the assistant's core functions together with the human-centered boundaries that keep it outside the safety-critical loop.

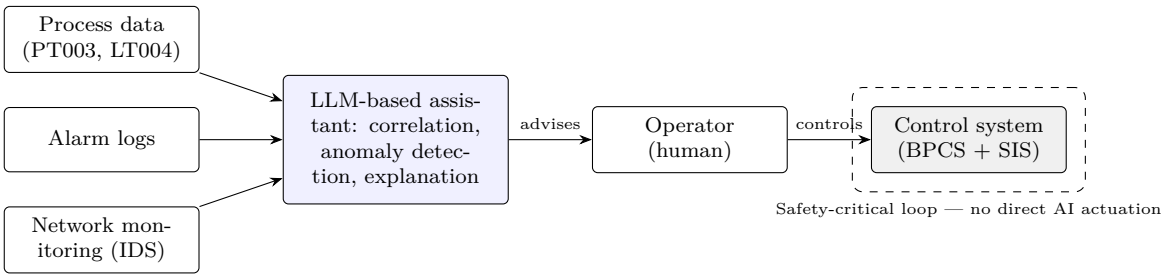


Figure 8.1: Architecture of the LLM-based operator assistant. The assistant correlates process data, alarm logs, and network monitoring to detect anomalies and explain them to the operator, but it only *advises*; all control decisions remain with the human operator, and the assistant never actuates the safety-critical control loop directly.

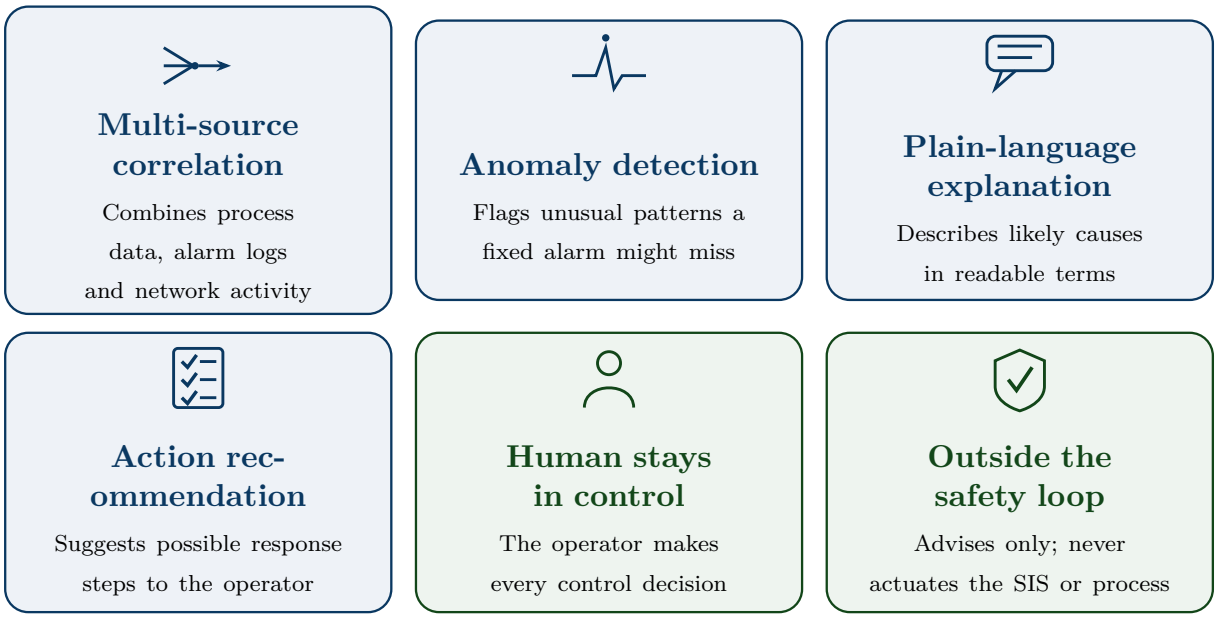


Figure 8.2: Core functions of the LLM-based operator assistant. The top row lists what the assistant does — correlating data, detecting anomalies, explaining them, and recommending responses — while the bottom row marks the human-centered boundaries: the operator remains in control and the assistant never acts on the safety-critical loop. The capabilities are shown in blue and the safety boundaries in green.

Chapter 9

Conclusion

This thesis has addressed the problem of cybersecurity risk assessment in industrial control systems used in chemical process environments. The increasing integration of digital technologies into industrial operations has created new challenges in which cyber threats can directly influence physical processes and introduce risks beyond traditional safety considerations.

A structured methodology was developed to integrate cybersecurity into established risk-assessment approaches. Traditional techniques such as the Hazard and Operability Study (HAZOP) and Fault Tree Analysis (FTA) were extended to include cyber-related causes, allowing both accidental failures and intentional attacks to be analysed within a unified framework. This integration provides a more comprehensive understanding of system behaviour and of the potential risk scenarios in cyber–physical environments.

The analysis of industrial control system architecture and communication networks highlighted vulnerabilities in key components such as programmable logic controllers, supervisory systems, and human–machine interfaces. These vulnerabilities can be exploited through various types of cyber attack, including unauthorised access, data manipulation, and network-based disruptions. The study showed how such attacks can propagate through the system and affect critical process variables.

A case study of the VS003 separator system was used to apply the proposed methodology in a practical context. The analysis identified the key process variables, the associated hazards, and realistic cyber-attack scenarios, and demonstrated how cyber events can lead to unsafe operating conditions such as incorrect pressure or liquid-level control. In particular, it showed how a single cyber cause — such as a spoofed pressure measurement — can act as a common cause that defeats both the basic process control system and the independent safety instrumented system, a dependency that a purely physical analysis would overlook. Based on these findings, a set of mitigation measures was proposed, including network segmentation, access-control mechanisms, sensor-integrity checks, monitoring systems, and organisational practices [4].

The thesis subsequently examined prevention and protection measures in depth. A defence-in-depth architecture was proposed for the VS003 system, combining nine complementary control layers: perimeter firewalls and network segmentation between IT, OT, and field zones; secure remote access via VPN with multi-factor authentication; role-based, least-privilege access control; continuous monitoring and intrusion detection; a structured incident-response plan; tested offline backups for rapid recovery; isolation of the safety instrumented system in its own network zone with its engineering key-switch held in the RUN position during normal operation; encryption and integrity protection for data in

transit and at rest; and systematic patch and vulnerability management [4, 5]. These layers are mutually reinforcing: no single control is sufficient on its own, but together they substantially reduce both the likelihood and the impact of a successful attack.

The secure network architecture places the VS003 field devices and SIS in a dedicated field zone separated from the OT supervisory zone by a control/field firewall, with the supervisory zone itself separated from the enterprise network and any remote-access connection through a DMZ. This arrangement ensures that monitoring data can flow upward to remote viewers in read-only form, while control commands can only originate in the control-room zone — eliminating any command path from the remote or DMZ side into the process.

The applicability of recognised standards to the VS003 system was also discussed. The IEC 62443 zone-and-conduit model provides a structured basis for network segmentation and the assignment of security levels to system components [4]. The NIST Cybersecurity Framework offers a complementary lifecycle view through its five core functions — Identify, Protect, Detect, Respond, and Recover — which together support continuous improvement of cybersecurity posture rather than one-time compliance. Applied together, these frameworks confirm that the protection measures selected for VS003 are both technically sound and aligned with industry best practice.

In addition to the core analysis, the thesis introduced the concept of an artificial-intelligence-based operator assistant. This concept illustrates how advanced data-analysis techniques, such as large language models, can support operators in detecting anomalies and responding to potential cyber incidents, while remaining outside the safety-critical control loop [17].

The results of this work emphasise the importance of integrating cybersecurity into process-safety analysis. Industrial systems should no longer be evaluated solely on the basis of physical failures, but also in terms of their exposure to cyber threats. A combined approach that considers both aspects is essential for ensuring safe and reliable operation in modern industrial environments.

No study of this kind is without limitations, and several should be kept in view when reading its results. The risk assessment is qualitative: cyber-initiating events are rated from exposure, attacker capability, and exploitability rather than from computed probabilities, because the failure-rate data that underpin physical reliability analysis do not exist for deliberate attacks. The fault tree is used accordingly as a logical model of how causes combine rather than as a fully quantified one; attaching meaningful probabilities to the cyber branches would require attack data beyond the scope of this work.

The analysis also rests on a single case. VS003 was chosen as a representative three-phase separator, and the shared-signal common cause it reveals is general in nature, but the specific deviations, scenarios, and ratings would have to be re-derived for a different process and instrumentation layout. The conclusions are tied in the same way to an assumed threat model: the scenarios consider an adversary able to reach the control network and manipulate signals or logic, and they hold relative to that assumption rather than to an exhaustive enumeration of every possible attacker or entry point.

Finally, the operator assistant of Chapter 8 is a concept rather than an implementation. It is specified at the level of architecture, information flow, and its deliberate separation from the safety-critical loop, but it has not been built, trained on plant data, or validated, and its practical value remains to be demonstrated.

Future work may focus on the practical implementation of the proposed methodology, including the development of automated tools for risk assessment, the quantification of

cyber-induced risk, and the integration of AI-based monitoring systems into real industrial control environments. Further research may also explore advanced detection techniques, improved methods for securing communication networks and control devices, and formal verification approaches for safety instrumented systems operating under combined physical and cyber threat models.

Bibliography

[1] T. A. Kletz, *HAZOP and HAZAN: Identifying and Assessing Process Industry Hazards*. Institution of Chemical Engineers, 1999.

[2] C. A. Ericson, *Fault Tree Analysis Primer*. CreateSpace, 2011.

[3] MITRE Corporation, "Att&ck for ics," 2026, https://attack.mitre.org/matrices/ics/.

[4] International Electrotechnical Commission, "Industrial communication networks – network and system security (iec 62443)," 2018.

[5] National Institute of Standards and Technology, "Guide to industrial control systems (ics) security," NIST, Tech. Rep. SP 800-82 Rev. 2, 2015.

[6] Dragos, Inc., "Ot/ics cybersecurity year in review," Dragos, Tech. Rep., 2026.

[7] European Union Agency for Cybersecurity (ENISA), "Enisa threat landscape 2025," ENISA, Tech. Rep., 2025.

[8] Cybersecurity and Infrastructure Security Agency (CISA) and FBI, "Darkside ransomware: Best practices for preventing business disruption from ransomware attacks," Joint Cybersecurity Advisory AA21-131A, 2021.

[9] National Institute of Standards and Technology, "Risk management framework for information systems and organizations," NIST, Tech. Rep. SP 800-37 Rev. 2, 2018.

[10] MITRE Corporation, "Common vulnerabilities and exposures (cve)," 2024, https://cve.mitre.org.

[11] R. M. Lee, M. J. Assante, and T. Conway, "Analysis of the cyber attack on the ukrainian power grid," *SANS ICS*, 2016.

[12] R. Langner, "Stuxnet: Dissecting a cyberwarfare weapon," *IEEE Security & Privacy*, vol. 9, no. 3, pp. 49–51, 2011.

[13] Bundesamt fuer Sicherheit in der Informationstechnik (BSI), "Die lage der it-sicherheit in deutschland 2014," BSI, Tech. Rep., 2014.

[14] ESET, "Industroyer: Biggest threat to industrial control systems since stuxnet," 2017.

[15] Dragos, Inc., "Trisis malware: Analysis of safety system targeted malware," 2017.

[16] Cybersecurity and Infrastructure Security Agency (CISA), "Compromise of u.s. water treatment facility," 2021.

[17] National Institute of Standards and Technology, “Nist ai risk management framework,” NIST, Tech. Rep., 2023.