

Ph.D. Program in Civil, Chemical and Environmental Engineering
Curriculum in Chemical, Materials and Process Engineering



Department of Civil, Chemical and Environmental Engineering
Polytechnic School, University of Genoa, Italy.



DARMS - Dynamic Asset integrity and Risk Management System

How Machine Learning and Systems Engineering cooperate to enhance the resilience of complex systems

Tomaso Vairo

“Randomness reflects incomplete information”

[N.N. Taleb]

DARMS
DYNAMIC ASSET INTEGRITY
AND RISK MANAGEMENT SYSTEM
HOW MACHINE LEARNING AND SYSTEMS ENGINEERING COOPERATE TO ENHANCE
THE RESILIENCE OF COMPLEX SYSTEMS

BY

TOMASO VAIRO

*Dissertation discussed in partial fulfillment of
the requirements for the Degree of*

DOCTOR OF PHILOSOPHY

*Civil, Chemical and Environmental Engineering
curriculum in Chemical, Materials and Process Engineering,
Department of Civil, Chemical and Environmental Engineering, University of Genoa, Italy*



December 2021

Supervisor:

Prof. Bruno Fabiano

Curriculum Coordinator

Prof. Attilio Converti

Coordinator of the PhD program

Prof. Roberta Massabò

External Reviewers:

Prof. Micaela Demichela, Dipartimento di Scienza Applicata e Tecnologia, Politecnico di Torino

Prof. Elpida Piperopoulos, Dipartimento di Ingegneria, Università di Messina

Examination Committee:

Prof. Elisabetta Finocchio, DICCA, Università di Genova

Prof. Maurizio Masi – Dipartimento di Chimica, Materiali e Ingegneria Chimica, Politecnico di Milano

Prof. Roberta Bongiovanni – Dipartimento di Scienza Applicata e Tecnologia, Politecnico di Torino

Prof. Ales Bernatik – Faculty of Safety Engineering, Technical University of Ostrava

Ph.D. program in Civil, Chemical and Environmental Engineering

Curriculum in Chemical, Materials and Process Engineering

Cycle XXXIV

ABSTRACT

“Static, incomplete, superficial, wrong”. The traditional approach to risk analysis, as applied in the process industries, has been largely criticized in response to recent major accidents. Since it was first proposed, modifications and improvements have been made, and a formal accepted approach is included in several regulations and standards (as the recent development of guidelines for the ageing management in SEVESO installations). Quantitative Risk Assessment (QRA) is based on consolidated procedures.

Nevertheless, the need of safety improvement asks for more advanced tools for hazard identification and risk evaluation. Besides considering technical aspects (e.g., malfunctions and process upsets), operational errors, organizational aspects, such as lack of attention and motivation to the safety culture, may lead to risk increment in terms of likelihood of undesired failures. Not all those aspects may be investigated with conventional QRA techniques, which have also the disadvantage of being intrinsically static and failing to capture risk variations during the lifecycle of a plant or production site. Despite their proved effectiveness, many hazards identification and risk assessment techniques lack the dynamic dimension, which is the ability to learn from new risk notions, experience, and early warnings. Now’s the time to go beyond the limits of conventional static methods for hazard identification and risk assessment; the risk assessment is, indeed, a very useful approach in support of this change but at the same time it is not exhaustive to capture also the possible “failure” in the interface/interaction among the several single components of a complex system beside their specific failures.

This research work discusses a novel approach for dynamizing the risk assessment process, integrating measured process data, asset integrity and operative conditions.

In the first part of the thesis, the inferential process and the application of Machine Learning to inference is discussed, and various applications of standard, and tailored, machine learning algorithms to industrial and environmental risks are detailed as case studies.

The second part is focused on the resilience engineering. The resilience paradigm is discussed, as well as the concept of emerging properties of complex systems. It will be shown how real-time data analytics, through appropriate AI models, combined with the expert knowledge of process engineering, constitute the fundamental technological key to pursue the resilience of plants and processes.

The third section integrates the aforementioned concepts within the wide framework of Systems Engineering. Accordingly, a dynamic and systemic model is presented, to address the significant shortcomings of the current risk analysis models. The Dynamic Asset-integrity and Risk Management System (DARMS) is designed starting from the Bow-tie technique, integrated with improved Machine Learning algorithms, to overcome the epistemic uncertainty in the prior probabilities and likelihoods of escalation factors and barriers. Subsequently, a Hidden Markov Model (HMM), based on Bayesian Inference, is developed to analyze real-time risk, and produce reliable predictions on the state of the whole

system during the operations. The application of the proposed model is demonstrated on an Oil and Gas terminal under Seveso legislation. The results of the case study provide a better understanding of the advanced Data Driven modeling of accident scenarios. The proposed model will serve as a useful tool for the operational safety management of complex systems.

INDEX

INTRODUCTION	5
I.1. Focus on the Resilience.....	5
I.1.1. The evolution of risk assessment	6
I.2. Machine Learning and Analytical Models	11
I.2.1. Learning from data.....	11
I.3. Systems Engineering.....	17
I.3.1. Systematic and Systemic approaches.....	19
OPEN RESEARCH QUESTIONS	22
1. INFERENCE AND MACHINE LEARNING.....	26
1.1. Inferential Statistics	26
1.1.1. White, Grey and Black swans	28
1.2. The Bayesian approach to Abduction	30
1.2.1. Markov Chain Monte Carlo.....	32
1.3. Dynamizing the QRA.....	33
1.3.1. Fault Tree and Bayesian Networks.....	34
1.3.2. Consequence analysis and Bayesian Networks.....	38
1.3.3. Consequence analysis and Decision Trees.....	41
1.4. Process Variables Prediction and Optimization.....	41
1.4.1. Predicting critical variables deviations.....	42
1.5. Conclusion of Section 1	46
2. THE RESILIENCE OF COMPLEX SYSTEMS	49
2.1. (mis)Understanding Resilience.....	49
2.1.1. Antifragility and resilience	50
2.1.2. Defining Resilience Engineering.....	51
2.2. A Framework for Resilience Assessment	52
2.2.1. Hidden Markov Model (HMM).....	54
2.3. An application of the RA Framework.....	56
2.3.1. Fault Tree Analysis	57
2.3.2. Development of the predictive model	58
2.3.3. Definition of a resilience score.....	61
2.4. Conclusion of Section 2	62

3. ENGINEERING SYSTEMS RESILIENCE PROPERTIES.....	66
3.1. Systemic Resilience.....	66
3.1.1. Defining the SE context for resilience.....	67
3.2. Integrating the RA Framework in the SE Lifecycle.....	73
3.2.1. Defining a Resilience metric	74
3.2.2. Dynamic Risk Index	75
3.3. Conclusion of Section 3	75
4. AN APPLICATION OF THE DYNAMIC ASSET-INTEGRITY AND RISK MANAGEMENT SYSTEM (DARMS) TO A LNG BUNKERING FACILITY	78
4.1. DARMS Architecture.....	78
4.2. Case Study.....	80
4.2.1. QRA of the bunkering operation	81
4.2.2. Model development	82
4.3. Conclusion of Section 4	85
CONCLUSION	88
PUBLICATIONS LIST.....	92
LIST OF FIGURES.....	94
LIST OF TABLES.....	96
REFERENCES	97
ACKNOWLEDGEMENTS.....	104

INTRODUCTION

I.1. FOCUS ON THE RESILIENCE

Research into system safety can be divided in two major aspects: understanding of how accidents occur and understanding of how risks can be adequately assessed and reduced. While there have been significant developments in the first field, there has been no comparable developments in the second.

A system is safe if it can absorb perturbations and the identification and assessment of plausible risks is therefore an essential prerequisite for system safety. Since accidents and risk assessment are two sides of the same coin, and since both are constrained in the same way by the underlying models and theories, it would be reasonable to assume that developments in system safety had matched developments in accident analysis. Risk assessment responds to the need to have an etiology of accidents, i.e. a study of possible causes or origins of accidents, but also an etiology of safety is needed, i.e. the understanding of what safety is and of how it may be endangered.

Such etiology of accidents is the crucial point of system safety and resilience engineering, and its development has been lacking.

The two pioneers of resilience engineering and system safety research are undoubtedly Prof. Erik Hollnagel and Prof. Nancy Leveson.

Prof. Hollnagel defines the concept of “Safety-II”, the basis of resilient performance. Safety-II implies a mindset change: Safety-II is ensuring systems performances under changing conditions, absorbing fluctuations without disruption. This concept is in contrast with the so called “Safety-I”, which is described as avoiding undesired conditions [1], [2].

The crux of Hollnagel theory is that fluctuations in the operation of the components of a complex system are unavoidable, thus they must be absorbed into a resilient system.

The practical implication is the need to understand and describe the systems everyday functional performance and its variability as well. The target implied in Safety-II is to “make things go right” but with the essential acknowledgment that performance variability is ineluctable; therefore, it is necessary as well to find ways to monitor and control this variability [1]. Traditional safety analysis and a risk assessment might not be appropriate anymore since the complexity of systems implies a significant difficulty to consider all the situations where safety can be endangered.

Prof. Hollnagel developed the Functional Resonance Analysis Method [4], which provides a way to describe outcomes using the idea of “resonance” arising from the variability of everyday performance. For a system, its essential functions are described using six basic characteristics. The potential functional variability is coupled with the functional resonance, described as a complement of causality. Functions are not defined a priori, nor hierarchically ordered; instead, they are described individually, and the relations between them are defined by empirically established functional dependencies.

The concept of “Safety-II” is very appealing, and worth of investigation. In Hollnagel, this powerful concept, is mainly related to workplaces safety, human factor, and organizational aspects, and without any doubt, there’s field for expanding “Safety-II” in the process industry world. Most of the ideas behind “Safety-II” are already implemented in the process industry, even if the shortcomings in defining likelihoods and interdependencies remain. FRAM does not appear to be a “safety” or risk analysis method.

Those remarks on the Hollnagel’s definition of safety are also reflected in a recent paper by Prof. Nancy Leveson [3]. Prof. Leveson underlines those goals such as resilience, flexibility, and adaptability are important, but they are much more likely to be achieved using approaches other than “Safety-II”. These properties must be built into the system as a whole, they are not a function simply of the behavior of human operators, which seems to be the almost total emphasis in “Safety-II”. Moreover, Prof. Leveson analyzed the FRAM, and stated that is not clear what can be learned from them through analysis. Every analysis method related to safety must be based on some causality model or assumptions about how accidents occur. As there is no description of the underlying causality model on which FRAM is based, it is unclear what this causality model exactly is. Prof. Hollnagel defines FRAM as a “method without a model rather than a model together with a method”. Prof. Leveson underlines that an analysis method is always based on some model or assumptions about system behavior. In fact, a hazard analysis model should include assumptions about how accidents occur, whether that model is conscious or subconscious. FRAM therefore is not a way to analyze a system to determine its properties, but it can be used for the creation of a specification of a particular system.

The successful use of a different approach to safety, called “System Safety”, is explained by Leveson. System Safety integrates the evolution (and sometimes revolutionary change) of engineering practice over time; these changes include greatly increasing complexity, the extensive and growing use of computers and other forms of new technology, and a changing role of humans in complex systems. This general approach is called “Safety-III”, just to put it into the Hollnagel context. Prof. Leveson extends the concept of System Safety and developed the Systems-Theoretic Accident Model and Processes (STAMP - Leveson 2012), which expands linear models to more complex causes of accidents. STAMP is a theoretical accident causality model; there are several general analysis methods built on STAMP, for example, STPA (Systems-Theoretic Process Analysis) is used for general hazard analysis and CAST (Casual Analysis using System Theory) is used for retrospective accident analysis. STAMP represents a fundamental change in the etiology of accidents, providing better and less subjective understanding on why accidents occur and how to prevent future ones [3].

1.1.1. The evolution of risk assessment

Humans have always been concerned about their safety. Prior to the industrial age, natural disasters provided the biggest challenge. Things started to change in the early part of the Industrial Revolution in Europe and the United States. Working in factories was extremely dangerous. When people accepted

employment, they also accepted the risks involved in the job and should be smart enough to avoid danger. Factories were plenty of potentially dangerous equipment, such as unguarded machines, flying shuttles, and open belt drives, as well as unsafe conditions, such as open holes. There were no fire escapes, and the lighting was inadequate. So, it is not surprising that the first attempt to analyse safety is related to workplaces. Accident investigation was actually a very small part of safety engineering in the past. Most of the effort was placed on identifying hazards and designing safety into products from the beginning of development.

Around 1930, workplace and system safety (including process safety) started to develop separately. On the workplace safety side, Heinrich (Industrial Accident Prevention, 1931) claimed that workplace accidents result from unsafe actions and unsafe conditions but suggested that people cause far more accidents than do unsafe conditions. Seizing on these arguments, opponents of mechanical solutions to workplace safety began to direct attention away from unsafe machinery and toward unsafe user acts.

Engineers, instead, focus on eliminating, preventing, and responding to hazards rather than focusing on the behaviour or actions of humans. An accident may involve conditions over which the engineer or designer of the system has no control, for example when, in a process plant, chemicals are inadvertently released, and atmospheric conditions are such that the chemicals are transferred to a place where humans are present.

The substances in the process industry are inherently hazardous; hazards are states of the system that can lead to losses of containment, and from here comes the concept of “Loss prevention”. In chemical and petrochemicals, the three major hazards are fire, explosion, and toxic release; that’s why much of the emphasis in loss prevention is on avoiding the escape of flammable or toxic substances through leaks, ruptures, explosions, ...

Design and operating procedures to eliminate or control these hazards have been included into procedures and standards produced by industrial sector and authorities. This approach was sufficient as long as the scale of process industry was relatively small, and development was slow enough to learn by experience, but after World War II the process industry began to grow in complexity, size, and technology at a very high rate, and the major accident hazard grew at a corresponding rate. The operation of chemical plants increased in difficulty and complexity, and the consequences of accidents, and the environmental concerns, increase. The control of hazards became more difficult, and the opportunity to learn by trial and error became almost impossible. At the same time, the social context has been changing. Major accidents (above all the Seveso accident, 1976, which resulted in the highest known exposure to TCDD in residential populations) had an enormous impact on the people, which generated political pressure for legislation to prevent major industrial accidents. Most of this legislation requires a hazard analysis, including identification of hazards and their causes and modelling of the most significant accident scenarios. After a relatively short time, the accidental release of MIC at Bhopal (1982) demonstrated that technical hazard

analysis by itself is not enough and that the implementation of a safety management system may be even more important [3]. The evolution of risk assessment is shown in Fig.1.

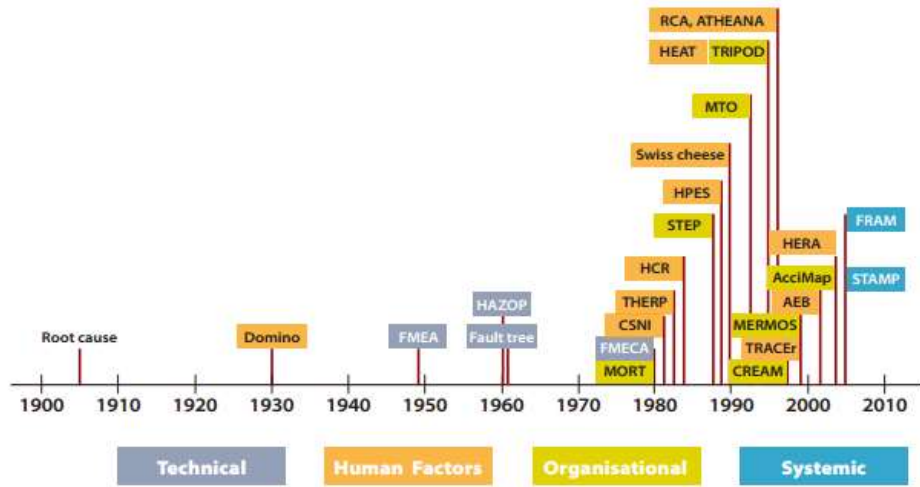


Figure 1: Timeline of Risk assessment methods

The archetype of a simple linear model is Heinrich's (1931) Domino model [5], which explains accidents as the linear propagation of a chain of causes and effects (Fig. 2).



Figure 2: Heinrich's Domino model

This model, which is related to workplace safety, was associated with one of the earliest attempts of formulating a complete theory of safety, expressed in terms of ten axioms of industrial safety [5]. The first of these axioms reads as follows: "The occurrence of an injury invariably results from a completed sequence of factors – the last one of these being the accident itself". The accident in turn is invariably caused or permitted directly by the unsafe act of a person and/or a mechanical or physical hazard. According to this view, an accident is basically a disturbance inflicted on an otherwise stable system.

Bird and Loftus [6] updated the domino sequence (Fig. 3) to reflect the management's relationship with the causes and effects of all incidents.

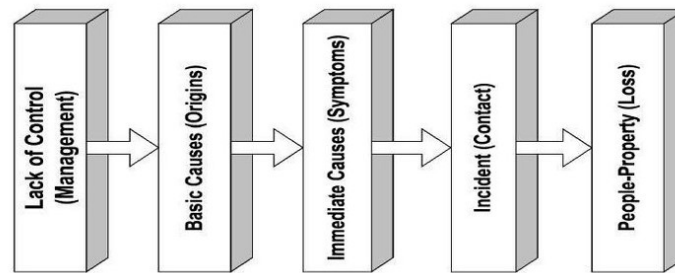


Figure 3: Bird and Loftus' Domino model

Bird and Loftus' Domino theory uses five dominos that represent the following events involved in all incidents:

1. *Lack of Control - Management*: Control in this instance refers to the functions of a manager: planning, organizing, leading, and controlling.
2. *Basic Cause(s) - Origin(s)*: The basic causes are frequently classified into two groups:
 - Personal factors such as lack of knowledge or skill, improper motivation, and/or physical or mental problems.
 - Job factors including inadequate work standards, inadequate design or maintenance, normal tool or equipment wear and tear, and/or abnormal tool.
3. *Immediate Causes(s) - Symptoms*: The primary symptoms of all incidents are unsafe acts and unsafe conditions.
4. *Incident - Contact*: An undesired event that could or does contact a source of energy above the threshold limit of body or structure.
5. *People – Property - Loss*: Loss refers to the adverse results of the accident. It is often evaluated in terms of property damage, as well as the effects upon humans, such as injuries and the working environment.

The central point in this theory is that management is responsible for the safety and health of the employees. Like Heinrich's theory, the Bird and Loftus domino theory emphasizes that contact incidents can be avoided if unsafe acts and conditions are prevented.

Although the domino model has been highly useful by providing a concrete approach to understanding accidents, it has reinforced the misunderstanding that accidents have a root cause and that this root cause can be found by searching backwards from the event through the chain of causes that preceded it.

A hazard analysis technique, developed in the mid-1960s, for and used primarily in the process industries is the Hazards and Operability Analysis (HAZOP) [7]. This technique is a structured and systematic examination of a complex planned or existing process or operation to determine causes and consequences of deviations from normal operating conditions. The obtained output is used to make changes in design, operating, and maintenance procedures.

The first complex linear model is the well-known Swiss cheese model proposed by Reason (1990) [8]. According to this, accidents can be seen as the result of interrelations between real time ‘unsafe acts’ by front-line operators and latent conditions such as weakened barriers and defenses, represented by the holes in the slices of ‘cheese’ (Fig. 4).

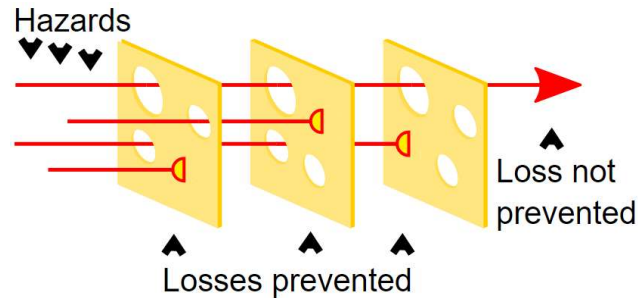


Figure 4: Reason's Swiss Cheese model

Although this model is technically more complex than the domino model, the focus remains on structures or components and the functions associated with these, rather than on the functions of the overall system as such. The Swiss cheese model comprises several identifiable components and related failures. Although causality is no longer a single linear propagation of effects, an accident is still the result of a combination of events, and the failure of a barrier is still the failure of a single component. The idea of a complex linear model such as this one, is to describe how coincidences occur.

The well-known risk assessment techniques, Fault Tree Analysis (FTA) and Event Tree Analysis (ETA), are respectively based on the Swiss Cheese model and on the Domino model. Some attempts to make FTA and ETA closer to the real operating conditions have been made by applying the fuzzy logic [9], which substantially extends the Boolean logic into a polyvalent logic, inserting a degree of truth (belonging function) to the properties of interest. However, the representation of the system remains still a picture, and the safety analysis cannot move from static relations between the system components.

Since some accidents are not completely understandable with even complex linear models, different solutions may be useful to achieve better safety performances. Real accidents involve unexpected combinations or aggregations of conditions or events [10]. The concept for expressing this situation is *concurrency*, meaning the temporal property of two (or more) events affecting each other. This has led to consider accidents as non-linear phenomena that emerge in a complex system, and the corresponding models are systemic accident models, or dynamic risk models. This view recognizes that complex system performance is always fluctuating, both because of the variability of the environment and the variability of the components.

Resilience has been defined in literature as “the ability of the systems to adapt to changing conditions in order to maintain a system property” [11]. In other words, a system is resilient if it can adjust its

functioning prior to, during, or following events (changes, disturbances, and opportunities), and thereby sustain required operations under both expected and unexpected conditions [2].

I.2. MACHINE LEARNING AND ANALYTICAL MODELS

Artificial intelligence is defined as “a branch of computer science dealing with the simulation of intelligent behavior in computers; the capability of a machine to imitate intelligent human behavior” [12].

Machine learning (ML) is the study of computer algorithms that improve automatically through experience. It is seen as a subset of artificial intelligence. Machine learning algorithms build a model based on sample data (training data), in order to make predictions or decisions without being explicitly programmed to do so. The study of mathematical optimization delivers methods, theory, and application domains to the field of machine learning. Machine learning involves computers discovering how they can perform tasks; it involves computers learning from data provided, so that they carry out certain tasks.

For simple tasks assigned to computers, it is possible to program algorithms telling the machine how to execute all steps required to solve the problem at hand; on the computer's part, no learning is needed. For more advanced tasks, it can be challenging for a human to manually create the needed algorithms. In practice, it can turn out to be more effective to help the machine develop its own algorithm, rather than having human programmers specify every needed step.

I.2.1. Learning from data

Each algorithm has an input and an output. the data is entered into the computer, the algorithm does what it must do, and in the end the result is obtained. Machine Learning turns everything upside down: we start from the data and the desired result, and we arrive at the algorithm that passes from one to the other.

Learning algorithms, or learners, are algorithms that create other algorithms. A core objective of a learner is to generalize from its experience. Generalization in this context is the ability of a learning machine to perform accurately on new, unseen examples/tasks after having experienced a learning data set. The training examples come from some generally unknown probability distribution (considered representative of the space of occurrences) and the learner has to build a general model about this space that enables it to produce sufficiently accurate predictions in new cases. Because training sets are finite and the future is uncertain, learning theory usually does not yield guarantees of the performance of algorithms. Instead, probabilistic bounds on the performance are quite common. The bias–variance decomposition [13] is one way to quantify generalization error (Fig. 5).

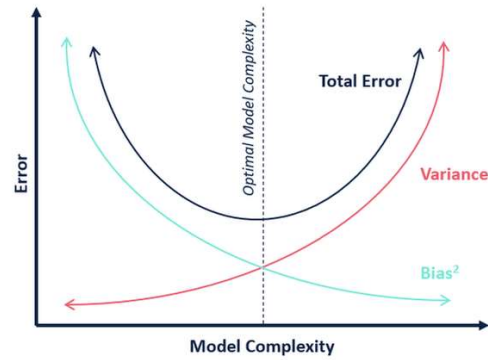


Figure 5: Bias - Variance tradeoff

For the best performance in the context of generalization, the complexity of the hypothesis should match the complexity of the function underlying the data. If the hypothesis is less complex than the function, then the model has under fitted the data. If the complexity of the model is increased in response, then the training error decreases. But if the hypothesis is too complex, then the model is subject to overfitting and generalization will be poorer (Fig. 6).

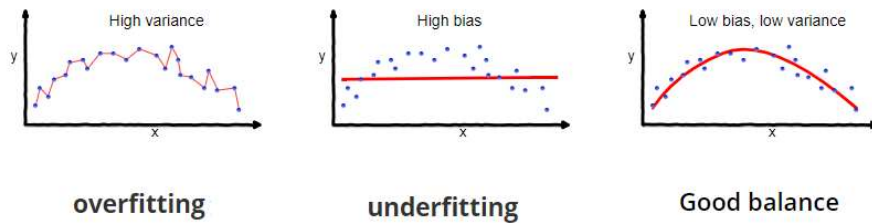


Figure 6: Overfitting and Underfitting

Mathematically: let the variable we are trying to predict as Y and other covariates as X . We assume there is a relationship between the two such that:

$$Y = f(X) + e \quad (1)$$

Where e is the error term, and it is normally distributed with a mean of 0. $f^{\wedge}(X)$ is a model of $f(X)$, obtained by linear regression or any other modeling technique. So, the expected squared error at a point x is:

$$Err(x) = E \left[(y - f^{\wedge}(x))^2 \right] \quad (2)$$

The $Err(x)$ can be further decomposed as

$$Err(x) = \underbrace{(E[f^{\wedge}(x)] - f(x))^2}_{\text{i.e. Bias}^2} + \underbrace{E[(f^{\wedge}(x) - E[f^{\wedge}(x)])^2]}_{\text{Variance}} + \underbrace{\sigma_e^2}_{\text{Irreducible Error}} \quad (3)$$

$\text{Err}(x)$ is the sum of Bias², variance and the irreducible error. Irreducible error is the error that cannot be reduced by creating good models. It is a measure of the amount of noise in the data.

Underfitting happens when a model unable to capture the underlying pattern of the data. These models usually have high bias and low variance. It happens when the amount of data to build an accurate model is too small, or when an improper model is selected, for example trying to build a linear model with a nonlinear data. These kinds of models are too simple to capture the complex patterns in data.

Overfitting happens when a model captures the noise along with the underlying pattern in data. It happens when the model is trained with an over noisy dataset. These models have low bias and high variance. These models are very complex like Decision trees which are prone to overfitting.

If the model is too simple and has very few parameters, then it may have high bias and low variance. On the other hand, if the model has large number of parameters, then it is going to have high variance and low bias. So, it is crucial to find the right balance without overfitting and underfitting the data. This tradeoff in complexity is why there is a tradeoff between bias and variance. An algorithm cannot be more complex and less complex at the same time. An optimal balance of bias and variance would never overfit or underfit the model. Therefore, understanding bias and variance is critical for understanding the behavior of prediction models.

In addition to performance bounds, learning theorists study the time complexity and feasibility of learning. In computational learning theory, a computation is considered feasible if it can be done in polynomial time. There are two kinds of time complexity results. Positive results show that a certain class of functions can be learned in polynomial time. Negative results show that certain classes cannot be learned in polynomial time.

Machine learning approaches are traditionally divided into three broad categories, depending on the nature of the "signal" or "feedback" available to the learning system:

- Supervised learning: The computer is presented with example inputs and their desired outputs, given by a "teacher", and the goal is to learn a general rule that maps inputs to outputs.
- Unsupervised learning: No labels are given to the learning algorithm, leaving it on its own to find structure in its input. Unsupervised learning can be a goal, i.e. discovering hidden patterns in data, or a means towards an end, as feature learning.
- Reinforcement learning: A computer program interacts with a dynamic environment in which it must perform a certain goal, taking the most appropriate decision (driving a vehicle, playing a game against an opponent, ...). As it navigates its problem space, the program is provided feedback that is analogous to rewards, which it tries to maximize.

As of 2020, deep learning has become the dominant approach for much ongoing work in the field of machine learning.

Learning algorithms are often grouped by similarity in terms of their function.

Regression Algorithms

Regression is concerned with modeling the relationship between variables that is iteratively refined using a measure of error in the predictions made by the model. Regression methods are a workhorse of statistics and have been co-opted into statistical machine learning. Regression algorithms predict the output values based on input features from the data fed in the system. The go-to methodology is the algorithm builds a model on the features of training data and using the model to predict the value for new data. Regression models have many applications, particularly in financial forecasting, trend analysis, marketing, time series predictions. The most popular regression algorithms are:

- Ordinary Least Squares Regression (OLSR)
- Linear Regression
- Logistic Regression
- Stepwise Regression
- Multivariate Adaptive Regression Splines (MARS)
- Locally Estimated Scatterplot Smoothing (LOESS)

Instance-based Algorithms

Instance-based learning model is a decision problem with instances or examples of training data that are deemed important or required to the model. Such methods typically build up a database of example data and compare new data to the database using a similarity measure in order to find the best match and make a prediction. For this reason, instance-based methods are also called winner-take-all methods and memory-based learning. Focus is put on the representation of the stored instances and similarity measures used between instances. The most popular instance-based algorithms are:

- k-Nearest Neighbor (kNN)
- Learning Vector Quantization (LVQ)
- Self-Organizing Map (SOM)
- Locally Weighted Learning (LWL)
- Support Vector Machines (SVM)

Regularization Algorithms

An extension made to another method (typically regression methods) that penalizes models based on their complexity, favoring simpler models that are also better at generalizing. The most popular regularization algorithms are:

- Ridge Regression
- Least Absolute Shrinkage and Selection Operator (LASSO)
- Elastic Net
- Least-Angle Regression (LARS)

Decision Tree Algorithms

Decision tree methods construct a model of decisions made based on actual values of attributes in the data. Decisions fork in tree structures until a prediction decision is made for a given record. Decision trees are trained on data for classification and regression problems. Decision trees are often fast and accurate and a big favorite in machine learning. The most popular decision tree algorithms are:

- Classification and Regression Tree (CART)
- Iterative Dichotomiser 3 (ID3)
- C4.5 and C5.0 (different versions of a powerful approach)
- Chi-squared Automatic Interaction Detection (CHAID)
- Decision Stump
- M5
- Conditional Decision Trees

Bayesian Algorithms

Bayesian methods are those that explicitly apply Bayes' Theorem for problems such as classification and regression. The most popular Bayesian algorithms are:

- Naive Bayes
- Gaussian Naive Bayes
- Multinomial Naive Bayes
- Averaged One-Dependence Estimators (AODE)
- Bayesian Belief Network (BBN)
- Bayesian Network (BN)

Clustering Algorithms

Clustering, like regression, describes the class of problem and the class of methods. Clustering methods are typically organized by the modeling approaches such as centroid-based and hierarchal. All methods are concerned with using the inherent structures in the data to best organize the data into groups of maximum commonalities. The most popular clustering algorithms are:

- k-Means
- k-Medians
- Expectation Maximization (EM)
- Hierarchical Clustering

Artificial Neural Network Algorithms

Artificial Neural Networks are models that are inspired by the structure and/or function of biological neural networks. They are a class of pattern matching that are commonly used for regression and classification problems but are really an enormous subfield comprised of hundreds of algorithms and

variations for all manner of problem types. The most popular classical artificial neural network algorithms are:

- Perceptron
- Multilayer Perceptron (MLP)
- Back-Propagation
- Stochastic Gradient Descent
- Hopfield Network
- Radial Basis Function Network (RBFN)

Deep Learning Algorithms

Deep Learning methods are a modern update to Artificial Neural Networks that exploit abundant cheap computation. They are concerned with building much larger and more complex neural networks and many methods are concerned with very large datasets of labelled analog data, such as image, text, audio, and video. The most popular deep learning algorithms are:

- Convolutional Neural Network (CNN)
- Recurrent Neural Networks (RNNs)
- Long Short-Term Memory Networks (LSTMs)
- Stacked Auto-Encoders
- Deep Boltzmann Machine (DBM)
- Deep Belief Networks (DBN)

Dimensionality Reduction Algorithms

Like clustering methods, dimensionality reduction seeks and exploit the inherent structure in the data, but in this case in an unsupervised manner or order to summarize or describe data using less information. This can be useful to visualize dimensional data or to simplify data which can then be used in a supervised learning method. Many of these methods can be adapted for use in classification and regression.

- Principal Component Analysis (PCA)
- Principal Component Regression (PCR)
- Partial Least Squares Regression (PLSR)
- Sammon Mapping
- Multidimensional Scaling (MDS)
- Projection Pursuit
- Linear Discriminant Analysis (LDA)
- Mixture Discriminant Analysis (MDA)
- Quadratic Discriminant Analysis (QDA)
- Flexible Discriminant Analysis (FDA)

Ensemble Algorithms

Ensemble methods are models composed of multiple weaker models that are independently trained and whose predictions are combined in some way to make the overall prediction. Much effort is put into what types of weak learners to combine and the ways in which to combine them. This is a very powerful class of techniques and as such is very popular.

- Boosting
- Bootstrapped Aggregation (Bagging)
- AdaBoost
- Weighted Average (Blending)
- Stacked Generalization (Stacking)
- Gradient Boosting Machines (GBM)
- Gradient Boosted Regression Trees (GBRT)
- Random Forest

Instance-based learning model is a decision problem with instances or examples of training data that are deemed important or required.

I.3. SYSTEMS ENGINEERING

Systems engineering (SE) is a transdisciplinary approach and means to enable the realization of successful systems. Successful systems must satisfy the needs of their customers, users, and other stakeholders. The term system means a collection of technical, natural, and social elements. SE considers systems as part of its foundations, and introduces two related definitions [14]:

- An engineered system is a technical or socio-technical systems system which is the subject of a SE life cycle. It is a system designed or adapted to interact with an anticipated operational environment to achieve one or more intended purposes while complying with applicable constraints.
- An engineered system context centers around an engineered system but also includes its relationships other engineered, social, or natural systems in one or more defined environments.

In the SE framework, the following elements are connected. *Systems Thinking* is the core integrative element of the framework. It binds the foundations, theories, and representations of systems science together with the hard, soft, and pragmatic approaches of systems practice. There is constant interplay between theories and practice, with theory informing practice and outcomes from practice informing theory. Systems thinking is the ongoing activity of assessing the system context and guiding appropriate adaptation.

Integrative Systems Science has a very wide scope and is grouped into three broad areas:

- Foundations, which help to organize knowledge and promote learning and discovery including meta-theories of methodology, ontology, epistemology, axiology, praxeology (theory of effective action), teleology, semiotics and semiosis, category theory, etc.
- Theories pertaining to systems are abstracted from domains and specialties, so as to be universally applicable: general system theory, systems pathology, complexity, anticipatory systems, cybernetics, autopoiesis, living systems, science of generic design, organization theory, etc.
- Representations and corresponding theories describe, explore, analyze, and make predictions about systems and their wider contexts, whether in terms of models, dynamics, networks, cellular auto-mate, life cycles, queues, graphs, rich pictures, narratives, games and dramas, agent-based simulations, etc.

Systems Approaches to Practice aim to act on real world experiences to produce desired outcomes without adverse, unintended consequences; practice needs to draw on the wide range of knowledge appropriate to the system-of-interest and its wider context. Traditional systems approaches are often described to be either hard or soft:

- Hard approaches are suited to solving well-defined problems with reliable data and clear goals, using analytical methods and quantitative techniques. They focus on technical systems, objective complexity, and optimization to achieve desired combinations of emergent properties.
- Soft approaches are suited to structuring problems involving incomplete data, unclear goals, and open inquiries, using a “learning system” metaphor, focus on communication, intersubjective complexity, interpretations and roles, and draw on subjective and “humanist” philosophies with constructivist and interpretivist foundations.

Pluralist approaches select an appropriate set of tools and patterns that will give sufficient and appropriate insights to manage the issue at hand, by applying multiple methodologies drawn from different foundations as appropriate to the situation.

Heuristics, boundary critiques, model unfolding, etc., enable the understanding of assumptions, contexts, and constraints, including complexity due to different stakeholders’ values and valuations. An appropriate mix of “hard”, “soft”, and custom methods draws on both systems and domain-specific traditions. Systems may be viewed as networks. Systems practice depends on measured data and specified metrics relevant to the problem situation and domain, the solicitation of local values and knowledge, and the pragmatic integration of experience, legacy practices, and discipline knowledge.

Integrative Systems Science allows us to identify, explore, and understand patterns of complexity through contributions from the foundations, theories, and representations of systems science and other disciplines relevant to the issue.

Systems Approaches to Practice address complex problems and opportunities using methods, tools, frameworks, patterns, etc., drawn from the knowledge of integrative systems science, while the observation of the results of systems practice enhances the body of theory.

Systems Thinking binds the two together.

Emergence is “the principle that entities exhibit properties which are meaningful only when attributed to the whole, not to its parts.” [15]. Emergent system behavior can be viewed as a consequence of the interactions and relationships between system elements rather than the behavior of individual elements. It emerges from a combination of the behavior and properties of the system elements and the systems structure or allowable interactions between the elements, and maybe triggered or influenced by a stimulus from the systems environment.

1.3.1. Systematic and Systemic approaches

A tension has existed throughout the history of Western thought around whether to focus on parts or the whole. The practice that springs from this history carries the same tension. This tension has been particularly visible within science and philosophy for a long time, and it gives rise to different approaches. Emphasizing the parts has been called mechanistic, reductionist or atomistic. An emphasis on the whole has been called holistic, organismic or ecological. As Fritjof Capra [16] notes: “In twentieth century science the holistic perspective has become known as *systemic* and the way of thinking it implies as *systems thinking*”. Capra also claims systems thinking is *contextual* thinking; and since explaining things in their context means explaining them in relation to their environment.

Two adjectives arise from the word “system”. *Systemic thinking*, thinking in terms of wholes, may be contrasted with *systematic thinking*, which is linear, step-by-step thinking. Likewise, it is possible to recognize systemic practice and systematic practice. Table 1 summarizes some of the characteristics that distinguish between systemic and systematic thinking and action [17].

Table 1: Systematic and Systemic thinking and action

Systematic thinking	Systemic thinking
The whole can be understood by considering just the parts through linear cause-effect mechanisms.	Properties of the whole differ; they are said to emerge from their parts, e.g. the wetness of water cannot be understood in terms of hydrogen and oxygen.
Systems exist as concrete entities; there is a correspondence between the description and the described phenomenon.	Boundaries of systems are determined by the perspectives of those who participate in formulating them. The result is a system of interest.

Perspective is not important.	Individuals hold partial perspectives of the whole; when combined, these provide multiple partial perspectives.
Analysis is linear.	Systems are characterized by feedback; may be negative, i.e. compensatory or balancing; or positive, i.e. exaggerating or reinforcing.
A situation can be understood by step-by-step analysis followed by evaluation and repetition of the original analysis.	Systems cannot be understood by analysis of the component parts. The properties of the parts are not intrinsic properties but can be understood only within the context of the larger whole through studying the interconnections.
Concentrates on basic building blocks. There is a foundation on which the parts can be understood.	Concentrates on basic principles of organization. Systems are nested within other systems – they are multi-layered and interconnect to form networks.
Analytical.	Contextual.
Concerned with entities and properties.	Concerned with process.
The system can be reconstructed after studying the components.	The properties of the whole system are destroyed when the system is dissected, either physically or theoretically, into isolated elements.
Systematic action	Systemic action
The espoused role of the decision-maker is that of participant-observer. In practice, however, the decision-maker claims to be objective and thus remains 'outside' the system being studied.	The espoused role and the action of the decision-maker is very much part of an interacting ecology of systems. How the researcher perceives the situation is critical to the system being studied. The role is that of participant-conceptualizer.
Ethics and values are not addressed as a central theme. They are not integrated into the change process; the researcher takes an objective stance.	Ethics are perceived as being multi-levelled as are the levels of systems themselves. What might be good at one level might be bad at another. Responsibility replaces objectivity in whole-systems ethics.
The system being studied is seen as distinct from its environment. It may be spoken of in open system terms, but intervention is performed as though it were a closed system.	It is the interaction of the practitioner and a system of interest with its context (its environment) that is the focus of exploration and change.
Perception and action are based on a belief in a 'real world'; a world of discrete entities that have meaning in and of themselves.	Perception and action are based on experience of the world, especially on the experience of patterns that connect entities and the meaning generated by viewing events in their contexts.
Traditions of understanding may not be	

questioned although the method of analysis may be evaluated.

There is an attempt to stand back and explore the traditions of understanding in which the practitioner is immersed.

Both systematic thinking and systemic thinking have their place; they are not in opposition but can rather be complementary in dealing with complex situations. Systemic thinking provides the context for systematic thinking and action. It is thus crucial to distinguish between systemic and systematic thinking and to embody these distinctions in practice.

OPEN RESEARCH QUESTIONS

Question I

The main objective of risk assessment within industrial settings is the minimization of accident probability or, at least, the preservation of this probability below an acceptable value. As commented by Genserik and Pasman [18], the risk picture obtainable from a QRA is static, being fully developed at the design stage.

Risk assessment needs a more detailed information on the current risk during operations, or better, on the ongoing safety of the whole system moment by moment. The quality of information, related with the analytical complexity, is shown in fig. 7.

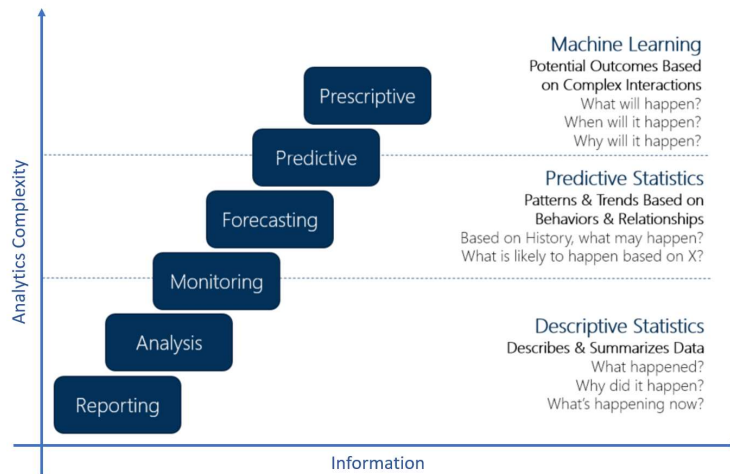


Figure 7: Analytical models

The Bayesian approach is currently widely recognized as a proper framework for analyzing risk in industrial plants [19], [20], [21], however is, in his standard implementation, unable to keep memory of the previous states of the plant components and thus is unable to catch the transition from “safe” to “unsafe” states, identifying the trend exclusively on the basis of the current state of the system.

On these grounds, several studies were performed focusing on a dynamic risk assessment by use of the BN in the process industries. As amply reported, even when performing an accurate risk analysis, it is not possible to rule out uncertainty completely, mainly due to lack of knowledge about the system and the physical variability of a system response. From those remarks, Research Question 1 arises:

How can real time measurements of process variables and Boolean failures be integrated into a dynamic risk model, for predicting the safety state transitions of a complex system?

Question II

Hollnagel [22] proposed the following four needs for resilient performance:

- The ability to respond.
- The ability to monitor.
- The ability to learn.
- The ability to anticipate.

The focal point to assess the resilience of a system is the identification of precursor events, which refers to early detection of “weak” signals from the system during the operations [22]. To identify the precursor events and thus maintain stability by applying appropriate adjustments, the analysis of a large amount of data is needed. By the data analysis, it is possible to predict the behavior of the system, thus catching the resilient performance according to the above mentioned 4 guidewords.

During the last decade, the so-called data-driven models are becoming more and more widespread. These models rely upon the methods of computational intelligence and machine learning, and thus assume the presence of a considerable amount of data describing the modelled system’s physics [24]. Those considerations lead to the Research Question II:

Can Data-Driven models be considered an appropriate approach to Resilience Assessment, complementing the Knowledge-Driven models describing systems physical behavior?

Question III

The definition of resilience includes the ability to maintain capability in the face of a disruption. Resilience engineering has to do with the resilience of the whole system, including organizations that design and operate engineered systems. That is why a systemic vision is a fundamental need.

To follow this way of thinking, SE also needs to consider all those organizational and technical systems which enable the systems resilience across the entire life cycle.

Resilience Engineering (RE) aims at providing tools to proactively manage risk, acknowledging the inherent complexity of system functioning and the correspondent need for performance variability [17] (Fig. 7 shows how the focal point of systems safety has changed in the past decades).

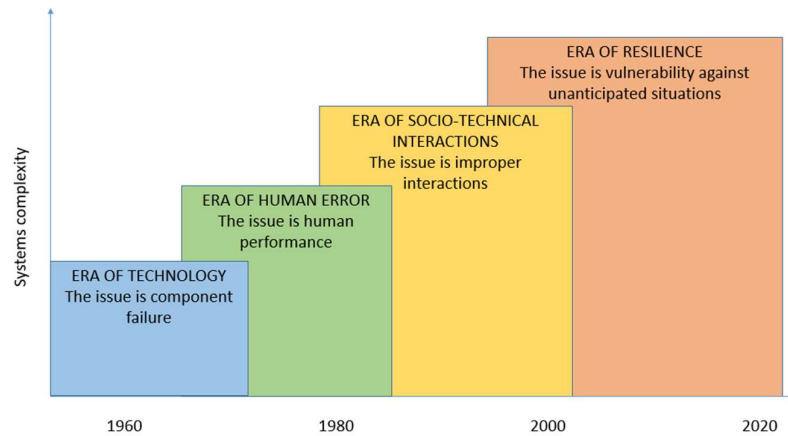


Figure 8: changing of systems safety focal point.

So, the concept of “resilience” has lately attracted widespread interest in systems safety. The term means the ability of a socio-technical system to adapt to disturbances and maintain its normal function.

If we want to face up to unanticipated situations, in complex systems, we need to establish a new research field, which integrates the Systems Engineering approach as a framework for developing resilient systems, in which, safety is an emerging property. This need for a new research field is the heart of the Research Question III:

How to integrate the Resilience Assessment into the Systems Engineering framework?

SECTION 1

1. INFERENCE AND MACHINE LEARNING

1.1. INFERENCE STATISTICS

The two main branches in statistics are called Descriptive and Inferential.

Descriptive statistics is the term given to the data analysis that helps to describe, show, or summarize data in a meaningful way such that patterns might emerge from it. However, this branch of statistics does not allow making conclusions beyond the analyzed data or reach conclusions regarding any hypotheses that might have been made. With descriptive statistics there is no uncertainty, because only the items that have been measured can be described, without trying to infer properties about a larger population [25].

This type of statistics usually describes data with two main tools:

- Central tendency, which is a way of describing the central position of a probability density distribution for a data set.
- Spread or dispersion, which represents the deviation from the most probable values.

Often, it is not possible to have access to the whole investigated population, but only a limited number of data is instead available and therefore a smaller sample must be considered. This is thought to be representative of the larger population.

Inferences are the steps of reasoning. Inferential statistics takes data from a sample and makes inferences about the larger population from which the sample was drawn. This sample needs to accurately reflect the population, but it does not have to be convenient. To be more arbitrary as possible, it is recommended to use a random sampling method. When using the inferential statistics, there will always be an error between the properties of the global population and the sample's ones. This error is always included in the results and an interval of confidence is outlined.

Within the inferential statistics there are two main approaches: frequentist and Bayesian inference.

- Frequentist inference calibrates the plausibility of propositions by considering repeated sampling of a population distribution to produce datasets. By considering the dataset's characteristics under repeated sampling, the frequentist properties of a statistical proposition can be quantified as fixed values.
- Bayesian inference preserves uncertainty. It is based on Bayes' theorem and updates an initial guess on the probability based on evidence.

Reasoning, indeed, is the process of using existing knowledge to draw conclusions, make predictions, or construct explanations. Three methods of reasoning are the deductive, inductive, and abductive approaches.

- *Deductive reasoning: conclusion guaranteed.*

Deduction is generally defined as "the deriving of a conclusion by reasoning." Its specific meaning in logic is "inference in which the conclusion about particulars follows necessarily from general or universal premises." Simply put, deduction - or the process of deducing - is the formation of a conclusion based on generally accepted statements or facts. Deductive reasoning always follows necessarily from general or universal premises. Therefore, while with deductive reasoning we can make observations and expand implications, we cannot make predictions about future or otherwise non-observed phenomena.

- *Inductive Reasoning: conclusion merely likely.*

Whereas in deduction the truth of the conclusion is guaranteed by the truth of the statements or facts, induction is a method of reasoning involving an element of probability. In logic, induction refers specifically to "inference of a generalized conclusion from particular instances." Inductive reasoning begins with observations that are specific and limited in scope, and proceeds to a generalized conclusion that is likely, but not certain, considering accumulated evidence. Inductive reasoning moves from the specific to the general. Much scientific research is carried out by the inductive method: gathering evidence, seeking patterns, and forming a hypothesis or theory to explain what is seen.

- *Abductive Reasoning: taking your best shot.*

The third method of reasoning, abduction, is defined as "a syllogism in which the major premise is evident but the minor premise, and therefore the conclusion, only probable." Basically, it involves forming a conclusion from the information that is known. Abductive reasoning typically begins with an incomplete set of observations and proceeds to the likeliest possible explanation for the set. Abductive reasoning yields the kind of daily decision-making that does its best with the information at hand, which often is incomplete. The abductive process can be creative, intuitive, even revolutionary. Einstein's work, for example, was not just inductive and deductive, but involved a creative leap of imagination and visualization that scarcely seemed warranted by the mere observation of moving trains and falling elevators. In fact, so much of Einstein's work was done as a "thought experiment", that some of his peers discredited it as too fanciful. Nevertheless, he appears to have been right-until now his remarkable conclusions about space-time continue to be verified experientially.

All three words are based on Latin *ducere*, meaning "to lead." The prefix *de-* means "from," and deduction derives from generally accepted statements or facts. The prefix *in-* means "to" or "toward," and

induction leads you to a generalization. The prefix ab- means "away," and you take away the best explanation in abduction [26].

The approach used in the thesis relies on Bayesian inference, which is the core of the implemented model. Bayesian statistics uses probability to quantify uncertainty or degree of belief, and probability distributions are used to represent the states of belief and is widely used in predictive models [20], [27].

1.1.1. White, Grey and Black swans

This is the well-known Rumsfeld's statement [28] that brought much fame and public attention to the concepts of managing risks. The concept is taken from the Johari window, a technique that helps people better understand their relationship with themselves and others. It was created by psychologists Joseph Luft and Harrington Ingham in 1955 [29].

This concept can be applied in risk management of complex systems. Considering the level of knowledge on the state of the system (or sub-system), and the level of knowledge on the behavior of the system (or sub-system), a matrix, identifying four categories, can be built.

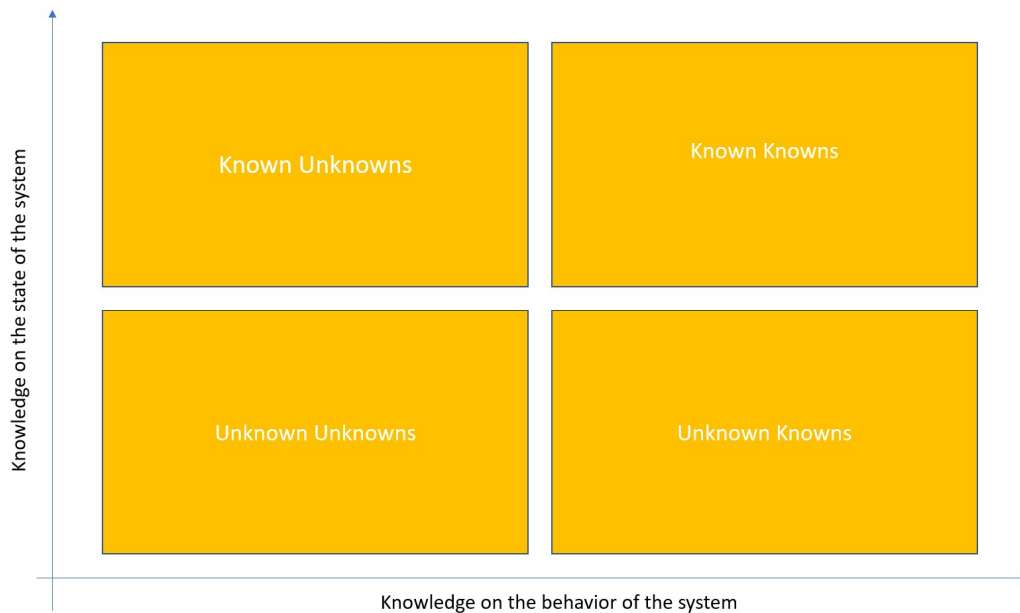


Figure 9: Known Unknowns matrix

These four categories are crucial to understand and assess risks.

Known Known Risks

Known knowns are the easiest type of risks. One 'known' stands for the fact that the organization is aware of the state of the system, thus it is aware that such a risk exists. The other 'known' is for the fact that the organization is aware of the behavior of the system, so the risk can be measured, and its effects can be quantified.

These types of risks are easiest to manage because the probability of them occurring as well as their impact is known.

It is thus possible to establish a relation with the inference models. There is no doubt the assessment of ‘known knowns’ risk belongs to the realm of *deduction*.

Known Unknown Risks

These risks are called known unknowns because the organization is aware of the existence of such a risk, knowing the state of the system. However, at the same time, the organization is not aware of the behavior of the system in that state, so they are unable to quantify probability and impacts that these risks will have on their system.

Unknown Known Risks

These are risks that are generally created due to the negligence of the organization. The behavior of the system when endangered by certain risks is known, but the organization decides not to manage it, consciously or unconsciously. ‘Known unknowns’ and ‘unknown knowns’ risk assessment falls in the field of *induction*.

Unknown Unknown Risks

These are the most dangerous type of risks which an organization faces. One unknown stand for the fact that the company is not even aware of the existence of such a risk, not properly knowing the state of the system. The other unknown goes without saying. These risks typically tend to have a very high impact and endanger the very existence of the organization. These are the so-called *black swans*. It is extremely difficult to predict the existence and impact of this type of risk with any degree of accuracy. This is where all the mathematical models of risk management begin to fail.

Taleb, in [49] distinguishes between three types of unexpected event:

White Swans

Result from the ‘Normal’ randomness of a Gaussian Distribution circumstance. This is the land of the *known unknown*. Its main features are:

- Averages have an empirical reality.
- The likelihood of an extreme event reduces at a faster and faster rate as you move away from the average.
- This effectively puts an upper and lower limit on the scale of an event (people can only be so tall or so short).
- Outliers are therefore extremely unlikely.
- No single event significantly affects the aggregate.

Because probabilities are known, the chance or risk of an unexpected event can be calculated precisely.

Grey Swans

Result from circumstances where Power Law (Fractal or Fat Tailed) randomness occurs. This is the province of the *somewhat known unknown*. Its main features are:

- Although probability of events follows a pattern this does not give precise predictability (think of forecasting the weather).
- Averages are meaningless because the range of possible events is so large.
- The likelihood of an extreme event reduces more and more slowly as you move away from the most frequent event.
- This means there is no upper limit.
- Outliers are therefore much more likely than you might think.
- One single event can change the whole picture.

Examples are magnitude of earthquakes, extinctions, frequency of words in a text, epidemics.

Black Swans

Result from all those forms of nonlinear randomness that we do not know about. This is the land of the *unknown unknown*. We have no means of calculating the likelihood of an event. Events appear without any probability distributions to describe them (however, as it will be shown in the following chapters, a best guess is always possible!).

A Black Swan:

- Is an outlier that lies outside of regular expectations (because nothing in the past can convincingly point to its possibility).
- Carries extreme impact.
- Gets explained after the event, therefore making it look like we could have predicted it (“if only ...”).

Managing unknown unknowns requires a deep change of paradigm in approaching uncertainties, and this is the field of *abduction*, where the prediction changes over time, being always the best guess based on evidence. Assessing and managing such risks is the purpose of the present work.

1.2. THE BAYESIAN APPROACH TO ABDUCTION

Abduction derives the best explanations for our observation. Statistical abduction attempts to define a probability distribution over explanations and to evaluate them by their probabilities. The framework of statistical abduction is general since many well-known probabilistic models, i.e., Bayesian Networks, Hidden Markov Models and Tree-based probabilistic models, are formulated as statistical abduction [30].

Most previous approaches to abductive reasoning have been based on first-order logic and determine a small set of assumptions sufficient to deduce the observations [26], [31]. This kind of reasoning can be represented by the following inference rule:

$$\frac{\psi \rightarrow \omega, \omega}{\psi} \quad (1)$$

i.e., if we observe ω and we have the rule $\psi \rightarrow \omega$, then we can infer that ψ is a plausible hypothesis (or explanation) for the occurrence of ω .

In general, there are several possible abductive hypotheses, and it is necessary to choose among them. To select the best explanations from the generated set, two kinds of criteria are used:

- metrics-based criteria (probability, weight, ...).
- simplicity criteria (the preferred explanation is the simplest available hypothesis).

In the context of probabilistic reasoning, abductive inference corresponds to finding the *maximum a posteriori* (MAP) probability state of the system variables, given some evidence (observed variables). In principle, we can solve this problem by "simply" generating the joint distribution, and then taking it as our starting point, to search for the configuration with maximum probability. However, this way to proceed is intractable even for problems with a small number of variables.

Recently, there has been a proliferation of new formalisms for integrating first-order logic and probabilistic graphical models to develop statistical relational models for knowledge representation, inference, and learning [32]. Of these formalisms, Markov Logic Networks [33], which combine first-order logic and undirected graphical models (Markov nets) have been used for abductive plan recognition.

In a logical framework, abduction, is usually defined as follows [26]:

- **Given:** Background knowledge B and observations O , both represented as sets of formulae in first-order logic, where O is typically restricted to a conjunction of ground literals.
- **Find:** A hypothesis H , also a set of logical formulae, such that $B \cup H \not\models \perp$ and $B \cup H \models O$.

Here \models means logical entailment and \perp means false, i.e., *find a set of assumptions that is consistent with the background theory and explains the observations*. There are generally many hypotheses H that explain a particular set of observations O . The best hypothesis is typically selected based on the size (simplicity) of H , following Occam's Razor. Often the background knowledge B is restricted to a set of Horn clauses and the hypothesis H is restricted to a set of ground atoms, resulting in abductive logic programming [34]. Several researchers have applied logical abduction to tasks like plan recognition and diagnosis, such as Ng and Mooney [35].

However, most of existing Logic Based Probabilistic Models impose restrictions on explanations (logical formulas) to realize efficient probability computation and learning.

1.2.1. Markov Chain Monte Carlo

To relax the previous mentioned restrictions, models based on MCMC (Markov chain Monte Carlo) for Bayesian inference can be adopted. The main advantage of MCMC over existing methods is that it has no restriction on formulas.

In the context of statistical abduction with Bayesian inference, whereas our deterministic knowledge can be described by logical formulas as rules and facts, our non-deterministic knowledge like frequency and preference can be reflected in a prior distribution in Bayesian inference.

The above-mentioned framework of statistical abduction can be formulated as inference on probabilistic models [36].

Let $\theta_j \equiv \{\theta_{jv}\}_{v=1}^{M_j}$ ($0 \leq \theta_{jv} \leq 1, \sum_{v=1}^{M_j} \theta_{jv} = 1$) be a parameter vector of a *categorical distribution* $Cat(\theta_j)$, corresponding to an M_j -sided dice, and also let $x_i \equiv \{x_{iv}\}_{v=1}^{N_i}$ ($x_{iv} \in \{0,1\}, \sum_{v=1}^{N_i} x_{iv} = 1$) be a value vector drawn from $Cat(\theta_{j_i})$, where j_i is the index of the categorical distribution which generates x_i (So, $N_i = M_{j_i}$ holds for each i). We use v_i ($1 \leq v_i \leq N_i$) to denote v such that $x_{iv} = 1$. Then a probability $p(x_i | \theta_{j_i})$ is equal to $\theta_{j_i v_i}$. Let θ and x be $\{\theta_j\}_{j=1}^M$ and $\{x_i\}_{i=1}^N$ respectively. Then a joint distribution $p(x | \theta)$ is computed as follow:

$$p(x | \theta) = \prod_{j=1}^M \prod_{i=1}^N \theta_{jv}^{\sigma_{jv}(x)}, \quad \sigma_{jv}(x) \equiv \sum_{i: j_i=j} x_{iv} \quad (1)$$

A function $f(x)$ is now introduced, which is a function of x such that $f(x) \in \{0, 1\}$, and use f (resp. $\neg f$) to denote the value of $f(x)$ is 1 (resp. 0). Then, the probability $p(f | \theta)$ is defined as follows:

$$p(f | x) \equiv f(x), \quad p(f | \theta) = \sum_x p(f, x | \theta) = \sum_x f(x) p(x | \theta) \quad (2)$$

Suppose the value and the definition of $f(x)$ are given as an observation O and knowledge base KB , respectively. Then, computing the most probable x given them on the joint distribution $p(f, x | \theta)$ is equivalent to performing statistical abduction, that is, finding the explanation E which has the highest probability from all possible x (a search space).

Starting from the logical framework for abduction, it is possible to transpose it into a Bayesian approach. Given an observation f and its rule in b , it is possible to perform Bayesian inference to infer the most probable x . In Bayesian inference, we assume a parameter θ as a random variable and introduce a *prior distribution* $p(\theta | \alpha)$ ($\alpha \equiv \{\alpha_k\}_{k=1}^L$) defined as:

$$p(\theta | \alpha) = \prod_{j=1}^M p(\theta_j | \alpha_{k_j}), \quad p(\theta_j | \alpha_{k_j}) = \frac{1}{Z(\alpha_{k_j})} \prod_{v=1}^{M_j} \theta_{jv}^{\alpha_{k_j v} - 1}, \quad Z(\alpha_k) \equiv \frac{\prod_{v=1}^{L_k} \Gamma(\alpha_{kv})}{\Gamma(\sum_{v=1}^{L_k} \alpha_{kv})} \quad (3)$$

Where $\{\alpha_{kv}\}_{v=1}^{L_k}$ ($\alpha_{kv} > 0$) is a parameter of a *Dirichlet distribution* $Dir(\alpha_k)$ and k_j denotes the index of the Dirichlet distribution which generates θ_j . Since Dirichlet distributions are conjugate to categorical distributions, the posterior distribution $p(\theta | x, \alpha)$, which is the modified distribution of θ by a given x , is also a product of Dirichlet distributions as follows:

$$\begin{aligned}
p(\theta | x, \alpha) &= \frac{p(x | \theta)p(\theta | \alpha)}{p(x | \alpha)}, p(x | \theta)p(\theta | \alpha) = \prod_{j=1}^M \frac{1}{Z(\alpha_{kj})} \prod_{v=1}^{M_j} \theta_{jv}^{\alpha_{kjv} + \sigma_{jv}(x) - 1}, \\
p(x | \alpha) &= \int p(x | \theta) p(\theta | \alpha) d\theta = \prod_{j=1}^M \frac{Z(\alpha_{kj} + \sigma_j(x))}{Z(\alpha_{kj})}, \quad \sigma_j(x) = \{\sigma_{jv}(x)\}_{v=1}^{M_j}
\end{aligned} \tag{3}$$

The most probable x given f is the one that maximizes $p(x | f, \alpha)$ computed as

$$p(x | f, \alpha) = \frac{f(x)p(x | \alpha)}{p(f | \alpha)}, \quad p(f | \alpha) = \sum_x f(x)p(x | \alpha) \tag{4}$$

Where $p(f | \alpha)$ is called marginal likelihood. Unfortunately, computing $p(f | \alpha)$ and $\operatorname{argmax}_x p(x | f, \alpha)$ involve evaluating $p(x | \alpha)$ on the large discrete search space. For avoiding this issue, the trick is switching from computing $\operatorname{argmax}_x p(x | f, \alpha)$ to sampling x from $p(x | f, \alpha)$.

Taking K samples $\{x_{(k)}\}_{k=1}^K$ from $p(x | f, \alpha)$, the one which maximize $p(x_{(k)} | \alpha)$ is the most probable explanation in the sample.

In addition, we can also approximate the marginal likelihood $p(f | \alpha)$ using the samples and a particular θ^* as follows:

$$p(f | \alpha) = \frac{p(\theta^* | \alpha)p(f | \theta^*)}{\sum_x p(\theta^* | x, \alpha)p(x | f, \alpha)}, \quad \sum_x p(\theta^* | x, \alpha)p(x | f, \alpha) \cong \frac{1}{K} \sum_{k=1}^K p(\theta^* | x_{(k)}, \alpha) \tag{5}$$

So, MCMC allows us to draw samples from a distribution even if we can't compute it. It can be used to sample from the posterior distribution over parameters.

A powerful specific implementation of MCMC sampling is Metropolis-Hastings (M-H). This technique requires a simple distribution called the proposal to help draw samples from an intractable posterior distribution. Metropolis-Hastings uses the proposal to randomly walk in the distribution space, accepting or rejecting jumps to new positions based on how likely the sample is [41].

M-H samples directly from $p(x | f, \alpha)$, by employing $p(x | f, \theta^*)$ as a proposal distribution, where θ^* is called a *production probability*. A candidate x^* is taken from $p(x | f, \theta^*)$. The candidate x^* is accepted with probability $A(x^*, x)$ computed by the following $R(x^*, x)$:

$$R(x^*, x) = \frac{p(x^* | f, \alpha)p(x | f, \theta^*)}{p(x | f, \alpha)p(x^* | f, \theta^*)} = \frac{p(x^* | \alpha)p(x | \theta^*)}{p(x | \alpha)p(x^* | \theta^*)} \tag{6}$$

The point here is that computing the marginal likelihood $p(f | \alpha)$, which is intractable but required to compute the target distribution $p(x | f, \alpha)$, is not required in the above computation.

1.3. DYNAMIZING THE QRA

As already mentioned, the traditional approach of Risk Assessment is affected by several limitations, essentially related to their static nature.

Bayesian inference represents a natural extension and refinement of Fault Tree Analysis (FTA) and Event Tree Analysis (ETA) frequentist approach [37]. Moving from frequencies to probability distributions allows incorporating local dependence between events and enabling both predictive and inference analysis.

The main uncertainties associated to the bow-tie analysis are on the likelihood, and interdependence, of root risk events in FT and events in ET, and this is due to insufficient statistical data and knowledge. Consequently, such an analysis may lead to exact, but often unrealistic, results. As shown in [20] and [39] a Bayesian hierarchical modelling technique, can be applied in order to get a better appreciation of the uncertainties in the whole bow-tie structure. In this way, the failure/occurrence probability will be no more defined by an exact number, but rather by a Probability Density Function (PDF).

As widely acknowledged, Bayesian hierarchical modelling is a statistical model written in multiple levels (hierarchical form) that estimates the parameters of the posterior distribution using the Bayesian method. Amin et al. [39] outline such an approach in the process sector. The sub-models combine to form the hierarchical model, and Bayes theorem is used to integrate them with the observed data and account for all the uncertainty that is present. The result of this integration is the posterior distribution, also known as the updated probability estimate, as additional evidence on the prior distribution is acquired [40]. Bayesian network established from the Bow-tie diagram can be used as a mapping algorithm, to identify critical influencing factors [41].

A conceptual illustration of a dynamic risk models depicted in fig. 10.

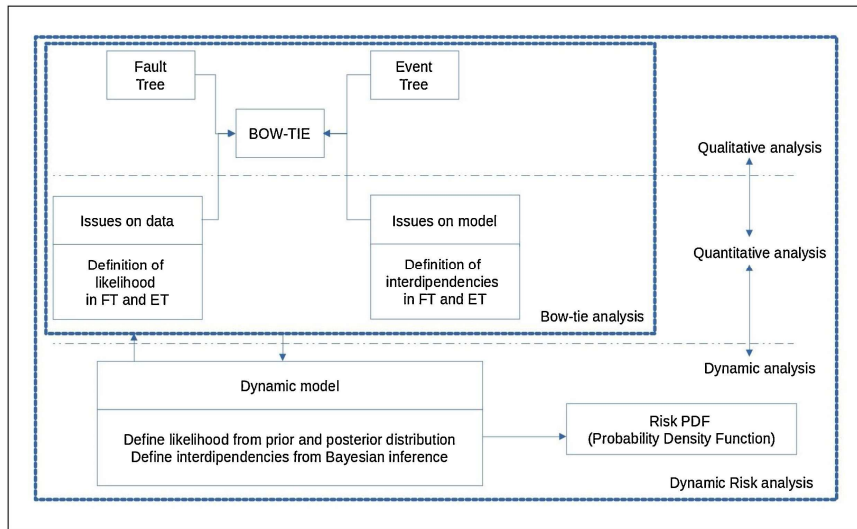


Figure 10: Dynamic Risk Assessment framework

1.3.1. Fault Tree and Bayesian Networks

Fault tree analysis (FTA) is a top-down, deductive failure analysis in which an undesired state of a system is analyzed using Boolean logic to combine a series of lower-level events. This analysis method is mainly used in safety engineering and reliability engineering to understand how systems can fail, to identify the best ways to reduce risk and to determine (or get a feeling for) event rates of a safety accident or a particular system level (functional) failure.

The Bayesian networks are constructed from FTA, where the elemental failure rates represent the *a priori* probabilities. The modelling design provides a set of independent nodes (root elements of FTA, i.e., critical items) and intermediate events for the top event. The network training is performed by using historical reliability data, near-miss and accidents data collected on the real system.

FTA encapsulate the concept that the failure of a (sub)system can be caused by the failure of lower-level (sub)systems, so that the minimal cut sets can be identified. Given a FTA, failure rates and failure distributions can be associated with each leaf in a Bayesian Network (BN) and then consolidated into a failure distribution for an entire system. In practice, quantifying leaf failures for an entire system is difficult or, in many cases, impossible. Bayesian networks allow the quantification by accepting evidence for the failure rate of any node and then, using Bayes' theorem to calculate the *a posteriori* probability of the failure rates of the sub-elements [19].

The Bayes' theorem can be used to update the prior knowledge of the *failure rate* λ with the information gained during an inspection. λ is not considered as a single value variable, but as a random variable expressed in the form of a probability density function (pdf). Then a posterior pdf given a newly observed evidence E is determined. This posterior distribution can be derived from the product of a prior distribution of failure rate values, $f(\lambda)$, and the new information as a likelihood function, $L(E | \lambda)$:

$$f(\lambda | E) = \frac{f(\lambda) L(E | \lambda)}{\int_0^{\infty} f(\lambda) L(E | \lambda)} \quad (7)$$

The likelihood function represents the probability that E is observed given a value of λ . With little information to start with as a prior distribution, a uniform distribution can be assumed, while for a likelihood function, in case of a constant failure rate, a Poisson distribution is suited.

In [19], an application of the outlined methodology to a petroleum products distribution terminal is described. The following fig. 11 represents the FT for a leakage in the Vapor Recovery Unit (VRU), and the fig. 12, the designed BN for the same event.

Leaking from the vapor recovery system and losses from components can occur both under random conditions and due to overpressures due to maneuvering errors on board (e.g., failure to open the valve of a tank during loading). The failure rates for the FTA are coming from literature.

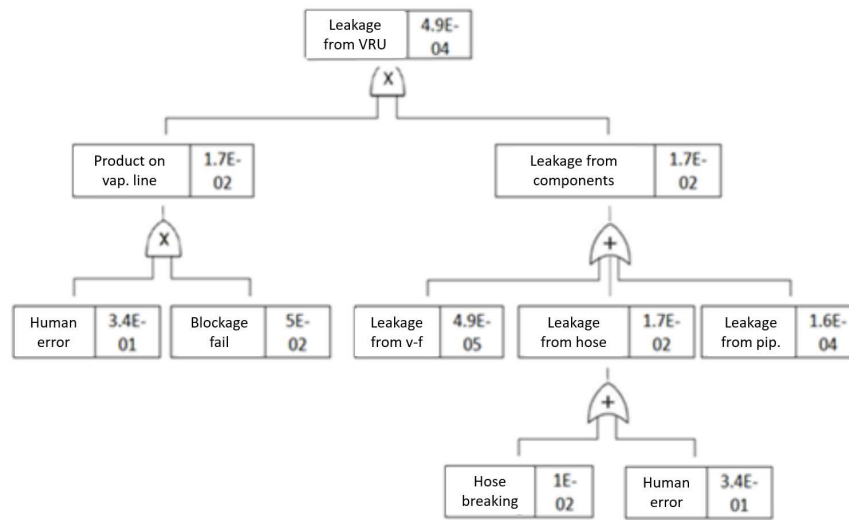


Figure 11: FTA for leakage on the VRU

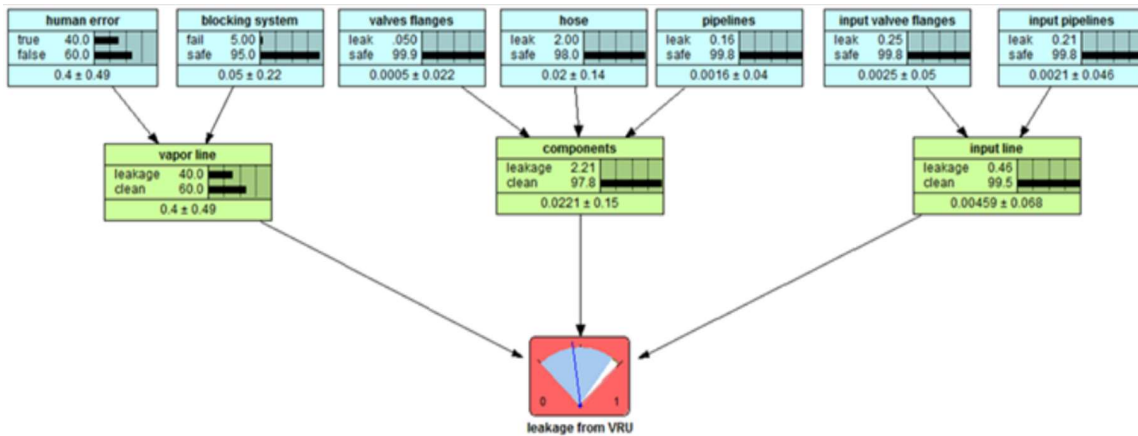


Figure 12: BN for leakage on the VRU

When the BN is designed, a sensitivity analysis can be empirically performed by altering each of the parameters of the query nodes and observing the consequent variation in the posterior probabilities of the query node (such as the endpoint).

Table 2 show the results of networks analysis performed by MCMC (Markov Chain Monte Carlo) sampling, evidencing the sensitivity of the Top Event nodes to findings in the other nodes and suggesting the possibility of setting up key management priorities, upon further validation with significant field data information.

Table 2: Sensitivity analysis for leakage on the VRU

Node	Belief Variance	Mutual Info
VRU	0.0010385	0.01180
Vapor line	0.0006889	0.00728
Flanges-Valves	0.0003494	0.00456
Input line	0.0003394	0.00354
Blocking System	0.0001996	0.00346
Human	0.0001497	0.00201
Hose	9.92 E-5	0.00150
Pipelines	9.01 E-5	0.00081
Input valves	7.71 E-5	0.00079

The networks are then updated, on one hand, with the measured reliability parameters (hard evidence) and, on the other one, with the evidence of the ongoing activities, which represent the likelihood coefficients (soft evidence). By considering a product transfer lasting nearly ten hours, soft evidence included, among others, the absence/presence of failures or losses in the early hours of the operation, which are the most critical. This evidence increases the likelihood coefficient related to the absence/presence of failures during the operation. Other evidence is, for instance, the absence/presence of failures, or malfunctions, at the start of VRU, or the absence of human error during the operations. From this, the trend of the probability of occurrence of the top events can be obtained, in relation to what is happening in the plant. Fig. 13 represent the posterior distribution for a leakage on the VRU, related to the collected evidence (randomly sampled from a categorical distribution).

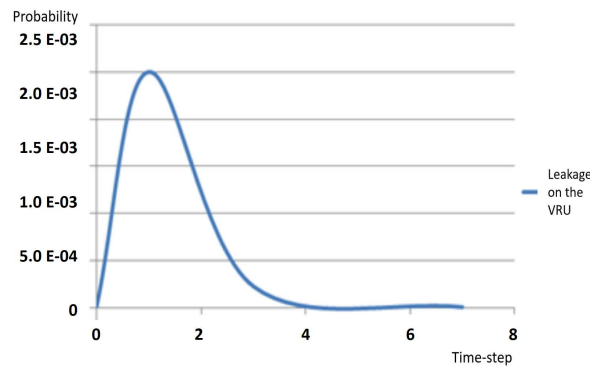


Figure 13: Posterior probability density function for a leakage on the VRU

The above-mentioned application demonstrate that BNs are a robust and flexible method for dynamizing the Fault Tree approach. The on-site evidence, collected during the operations, can be incorporated in a quantitative on-time safety analysis.

In the following paragraph, an application of the Bayesian approach to consequence evaluation is described.

1.3.2. Consequence analysis and Bayesian Networks

When dealing with accidental hazardous releases and consequences evaluation, the most used modelling approach is based on numerical simulations for evaluating atmospheric dispersion and transport starting from emission source characterization and boundary conditions. It relies on numerical solutions of the Navier-Stokes's equations (a simplified form of them, known as RANS – Reynold-Averaged Navier-Stokes). The main issue with this approach is that the quantification of uncertainty of modelling results, and the optimal selection of the atmospheric dispersion model, relies on an accurate evaluation on how model uncertainty in inputs affects the outputs [43]. In case of complex geometries, such as urban areas additional limitations arise [44].

Process-based numerical models attempt to simulate real-world processes in a virtual environment which can be easily manipulated and studied. Conceptually, the experimental design of these simulation studies broadly falls in one of three categories: predictive modelling, explanatory modelling, and exploratory modelling. However, the epistemologies of these three modes of modelling are yet incomplete and not fully understood. Not only do the three modes of modelling have different underlying assumptions, but they also have different criteria to establish validity and different limitations on the interpretations and inferences that can be made.

The dispersion calculations are a possible source of error, and such errors or uncertainties need to be quantified. An important source of uncertainty is the meteorological data used in the calculations. Such data may be less than ideal because of constraints imposed by both availability and by the variances associated with population from which the data are obtained [46].

Analytical models are mathematical models that have a closed-form solution, i.e., the solution to the equations used to describe changes in a system can be expressed by a mathematical analytical function [45]. The purpose of the analysis is to understand something about the nature of the real world by studying the observed data in relation to their context. Fig. 14 depicts a schematic view of analytical modelling: X-axis reflects the purpose of modelling, ranging from association/correlation to causality. Y- axis represents the source of the model conceptually ranging from theory to data. The theory (parametric) is labelled as Human Intelligence, suggesting the origin. At the opposite end of the Y-axis, data are commonly associated with Machine Learning and Artificial Intelligence.

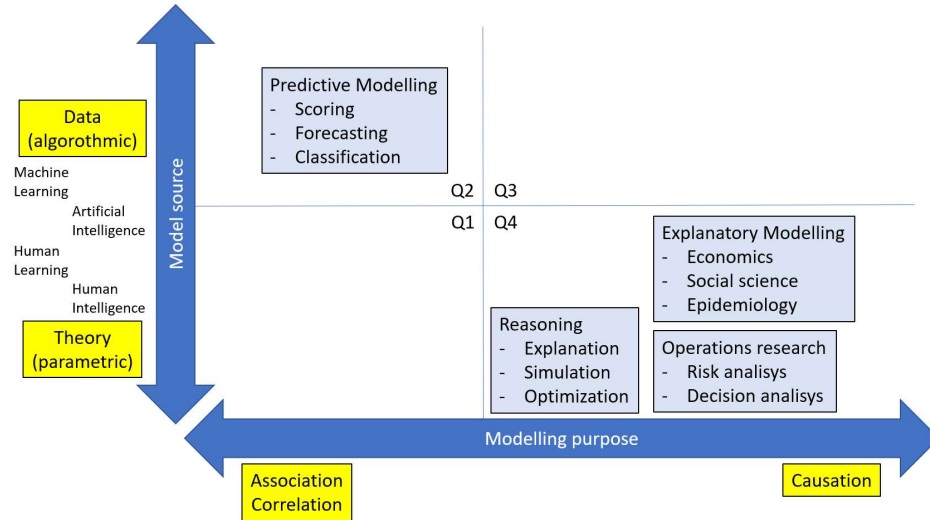


Figure 14: Map of analytical modelling

The Bayesian approach is based upon the probability assigned to an event as a consequence of the current knowledge (inference process). A Bayesian network provides a complete description of the application domain in the form of a conditional probability distribution and is a directed acyclic graph in which each link is directed from a parent node to its child. Each node represents a variable of the domain, and the links represent causal dependencies among the variables. The possible values of each variable are divided into some intervals (of a given width) or are considered in a Gaussian distribution. Evidence is called the a-priori information about the degree of certainty assigned to the possible states of a variable. Each variable without parents is characterized by an a-priori probability table, while each variable A with some parents $B_1 \dots B_n$ has a conditional probability table that expresses the joint probability. A generic element of the conditional probability table is the probability that occurs as a combination of two specific values of the variables. More in general, it is possible to express it in the form: $P(X_1 = x_1 \mid \dots \mid X_n = x_n)$ and consequently obtain Eq. 8.

$$P(x_1, \dots, x_n) = \prod_{i=1}^n P(x_i \mid Parents(x_i)) \quad (8)$$

The construction of a Bayesian network can be performed according to the following phases:

1. A set of variables A_i ($i=1, \dots, n$) is chosen to describe the system.
2. To each variable A_i is associated a node of the network.
3. The set $Parents(A_i)$ of the nodes parents of A_i is fixed.
4. The table of the conditioned probabilities for A_i is defined by the learning of a set of cases.

To test the described approach, it has been applied to the prediction of PM10 concentration in an urban area [27]. A two-stage modelling approach was adopted: The Bayesian network was manually established by expert knowledge, and then the previously obtained Bayesian network model is corrected by learning the interrelations from the data, thus performing an automatic construction of BBN.

The defined Bayesian Network relies on acquisition and elaboration of forecasting data of temperature T, wind W, humidity H and rain R, from a LAM meteorological model (2013-2016), and PM10 concentration data (2013-2016) from three stations in the urban monitoring network of Genova (S1, S2, S3), as daily average concentration.

The logic steps of the analytical modeling can be summarized as follows:

- Data extraction from the meteorological model, processing and labeling suitable for the construction of the network.
- Collection of concentration data from the control units, processing and labeling suitable for the construction of the network.
- Preliminary statistical analysis.
- Development of the network with different learning algorithms (hill climbing, Grow-Shrink, Incremental Association, Taboo search).
- Validation of the consistency of the network.

The network has been built with continuous values, deriving appropriate distributions from the input data set. The structure of the network derives from the data analytics (clusters analysis and correlations).

As depicted in Fig. 15, the automatically built structure shows analytical dependencies between the nodes that are often not intuitive, but that, with the use and validation of the model, are proved statistically significant.

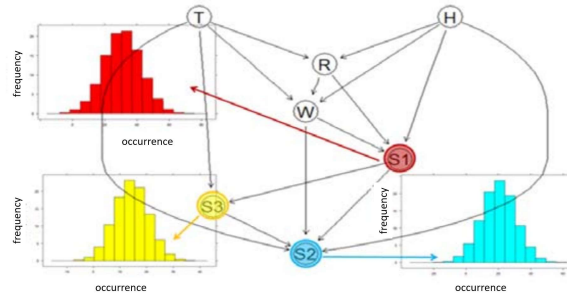


Figure 15: BBN with evidence of the interdependencies between the nodes.

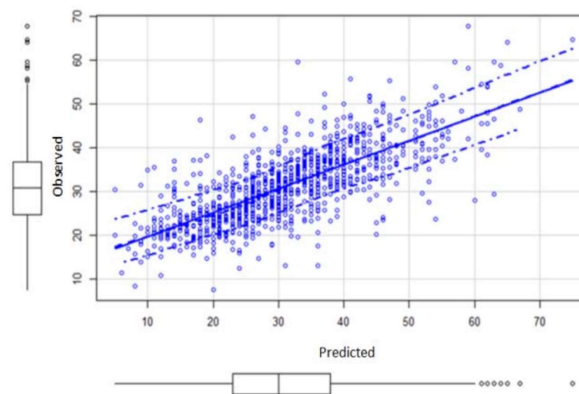


Figure 16: scatterplot observed vs. predicted.

As shown in Fig. 16, the BN shows a remarkable prediction ability. The most notable errors are on the extreme values (low and high concentration). To fix those errors, a different approach, described in the following paragraph, was tested.

1.3.3. Consequence analysis and Decision Trees

The above-described probabilistic inferential approach is heavily dependent on the prior probabilities. That's why, in the depicted case study, some errors on the extreme values arise.

In [47], a different Machine Learning modelling approach was used. The proposed framework relies on an algorithm performing gradient boosting on decision trees (LightGBM). Light GBM is based on a highly optimized library that performs very well in structured/tabular data problems, capable of gracefully managing a mix of scalar and categorical variables [48]. The predictive model was trained with meteorological data, ozone measurements in three urban areas, and time variables, all suitably pretreated. After a reliable cross-validation strategy, to balance bias and variance in the prediction results, and thus avoid situations of under-specification and over-specification, the model showed excellent results, as demonstrated in Fig. 17.

The most critical issue for obtaining reliable prediction is the data engineering phase. This topic is covered in the next Chapter.

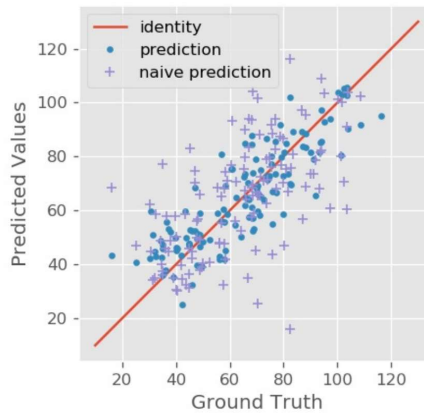


Figure 17: scatterplot observed vs. predicted – LightGBM.

1.4. PROCESS VARIABLES PREDICTION AND OPTIMIZATION

The need for extracting useful and valuable information from historical data in the process industry has led in the evolution of machine learning and data driven models. As shown in the previous sections, data driven models can improve the risk analysis process by dynamizing it, and therefore making it more consistent with what is happening in the plant. The Bayesian approach is a proper framework for analysing risk in industrial plants [19], [20], [21], however, it is unable to keep memory of the previous states of the

plant components and thus is unable to catch the transition from “safe” to “unsafe” states, identifying the trend exclusively based on the current state of the system.

1.4.1. Predicting critical variables deviations

In designing and implementing reliable operational control systems, based on data-driven models and machine learning for predicting the system behaviour, one of the critical issues to deal with is the coexistence of Boolean elements (e.g., failure of instruments) and analogical elements (deviation of process variables). The connection with the process variables may be difficult, in fact, the Bayesian approach is tailored for predicting Boolean events, such as failures, malfunctions, unavailability, etc., but when dealing with analogical variables, the predictive capability decreases.

Soft Sensors established themselves as a valuable alternative to the traditional means for the acquisition of critical process variables, process monitoring and other tasks which are related to process control [50]. The machine learning application acquires knowledge based on historical data of the industrial plant, and based on that data, it can predict, classify, and cluster new process outputs.

To solve this issue, in [51] a hybrid system based on a deep learning approach, coupled with a hierarchical Bayesian structure, is presented. The system started from the risk assessment outcome, which leads to the identification of the precursors; then, a predictive soft sensor, based on deep learning algorithms, was developed for predicting target variables and improve the decision-making process for hazardous condition prevention.

The hybrid system represents a novel process-monitoring framework based on machine-learning for identifying the different abnormal events of the system considering safety-relevant deviations leading to possible faults.

The idea behind the hybrid model, for integrating continuous variables and Boolean events is depicted in fig. 18.

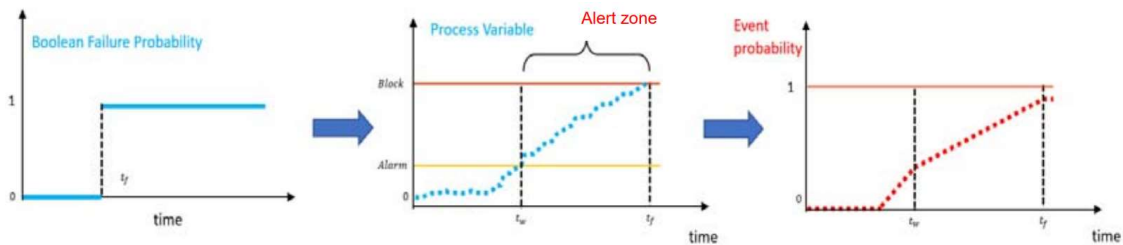


Figure 18: Hybrid model concept.

The proposed model was applied to the same plant component presented in § 1.3.1, related to a coastal storage facility located in Northern Italy. The plant is equipped with a real-time monitoring system properly updated to transfer actual process parameters values to the designed predictive system, with focus on discovering unanticipated behaviours and creating reliable alarms, according to early warning pillar of resilience. Figure 19 schematically depicts the VRU section and the main lines including mail vapour, vapour recovery, regeneration and absorption streams, vacuum pump suction circuit and water drainage.

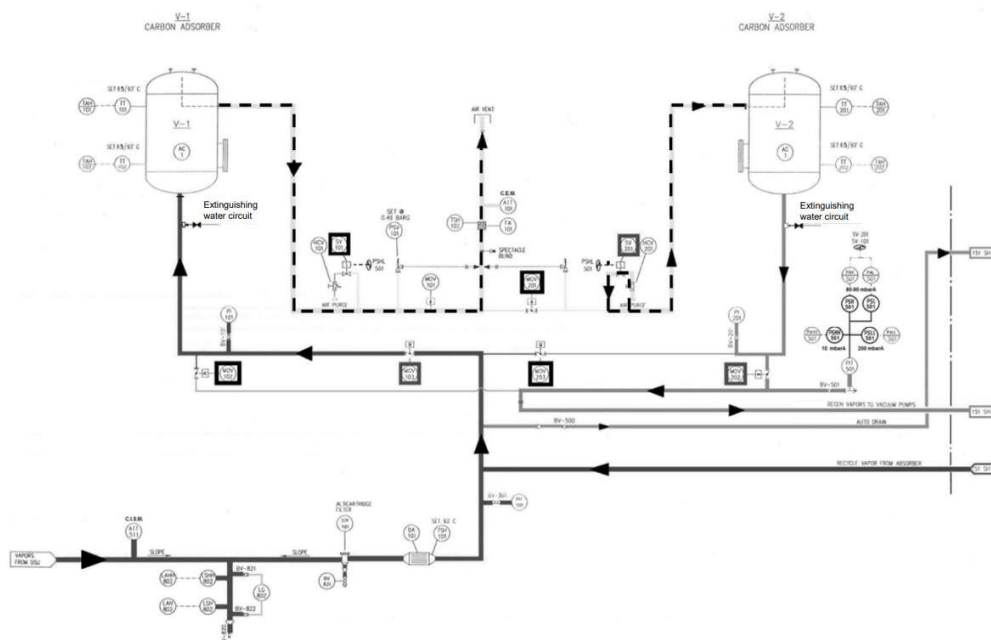


Figure 19: Schematic diagram of the Vapor Recovery Unit (VRU).

Given the core activity of the storage site, in case of accidental HC releases the most probable accidental scenarios include either atmospheric release, sea spills, liable to cause environmental damage in different matrixes, or fire/explosion scenarios due to ignition source presence. The core of VRU is the activated carbon absorption; the process is based on the Pressure Swing Absorption (PSA) with two absorbers operating alternatively, one active and the other one at regeneration stage. As in many process engineering tasks, accurate temperature prediction and estimate of heat transfer coefficient are required to guarantee the optimal operative and safety performance [52]. The activated carbon bed temperature represents the control parameter for the unit physical state and the regeneration step sequence. Additionally, it is selected as critical parameter for dynamic early warning of possible hazardous scenario, with set points respectively at $T= 70^{\circ}\text{C}$ for alarm activation and $T= 93^{\circ}\text{C}$ for emergency shutdown.

Table 3 summarizes the real-time monitored process variables in the VRU section of the plant, with a frequency of 5 sec. The whole dataset adopted in the study includes observations over one-year time span.

Table 3: Hardware sensors and relevant process parameters monitoring VRU process section.

Monitoring system	Aim
TT101	Temperature sensor in the upper area of the absorbent filter V-1
TT102	Temperature sensor in the lower area of the absorbent filter V-1
TT201	Temperature sensor in the upper area of the absorbent filter V-2
TT202	Temperature sensor in the lower area of the absorbent filter V-2
PIT101	Inlet pressure transmitter for filter line V-1 and V-2

PIT501	Pressure transmitter on the vacuum line that manages SV-101/201 and SV-501
VOC_INLET	It is the concentration of vapors entering the system
CIM_FLOW	CIM flow rate inside the plant.

The model is designed with 2 DNN (Deep Neural Networks) based on the Resilient back propagation algorithm (schematically represented in fig. 20) with a structure having the following characteristics as resulting after the validation phase:

- hidden layers: 6,
- neurons in each hidden layer: 24,
- learning rate: 1E-7,
- step max: 1E8,
- activation function: Tanh,
- error function: SSE,
- prediction time interval: 15 min.

The activation function performs data analysis and processing; when the sum of weighted inputs and biases exceeds a precise activation threshold, the activation function considers the argument valid and processable. Weight parameter quantifies the inputs importance. Weights and bias are corrected by learning algorithm for adapting the NN to the input dataset.

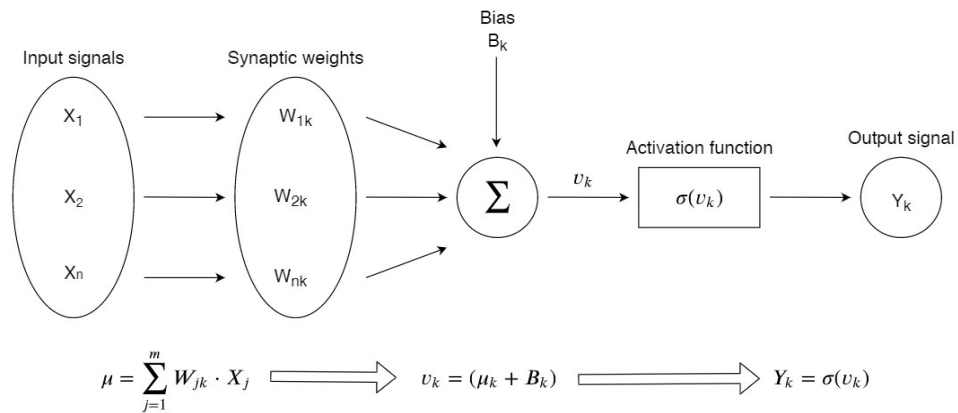


Figure 20: conceptual elements of a Backpropagation Neural Network.

Fig. 21 represents the scatterplots of predicted values vs. plant observed values (monitored empirical data).

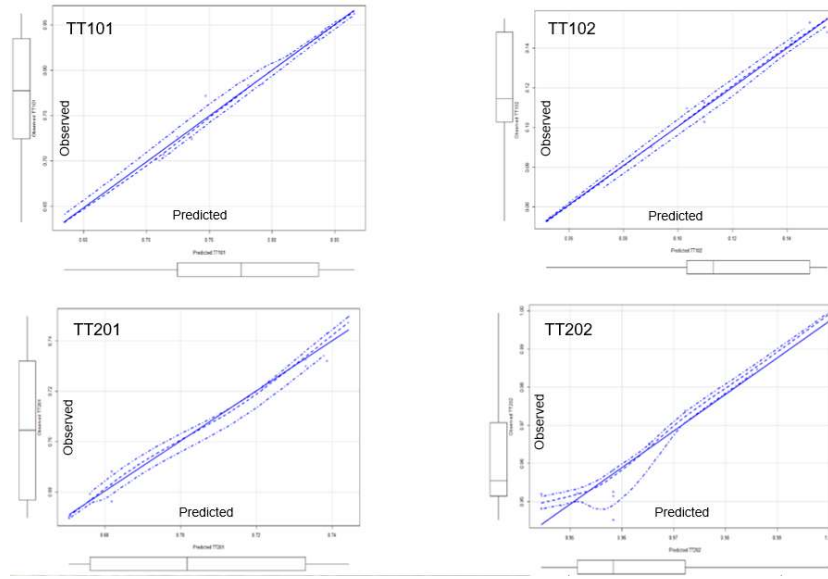


Figure 21: Scatterplots predicted values vs. ground truth for temperature sensors.

The robust and accurate prediction ability reflects on a clear-cut overlap between predicted and field observed data. The analogical variables are used in a HBN (Hierarchical Bayesian Network) to predict the posterior probability of failure of the component in accordance with the hybrid model, as shown in Figure 18. According to the outlined approach, by integrating the predictions of DNN-based soft sensors in a HBN, it is possible to obtain a continuous probability trend even for those elements characterized by a Boolean risk, such as failures of plant components.

The early warning principle of resilience is used to determine the variables set points. The HBN model relies on the probabilistic data and the data forecasted by the DNN. The Bayesian part of the hybrid model updates the risk probabilities by integrating the variables predictions time after time, considering the previous step (hierarchical structure). The algorithm has performed the prediction with remarkable accuracy and precision as demonstrated by standard statistical indicators shown in Table 4.

Table 4: Errors and accuracy.

	TT101	TT102	TT201	TT202
RMSE	0.008	0.003	0.003	0.004
MAE	0.005	0.002	0.003	0.003
Accuracy	0.999	0.999	0.999	0.999

The hybrid model gives a real time estimate of the components failures likelihoods by analysing the variables predictions, capturing the temporal and spatial dependencies of the relevant process parameters, and the interdependencies in the component failure analysis. It can provide robust performances thus representing a sharp jump towards early detection of systems weak signals and overall system resilience in

perspective resulting in a risk function characterized by predictive capabilities and the ability to be updated with time.

1.5. CONCLUSION OF SECTION 1

Risk assessment faces a series of overall challenges, partially related to technology advancements and increasing needs. The most pushing needs are the call for continuous risk assessment, improvement in learning past lessons and definition of techniques to extract information from relevant data, which are to be coupled with adequate capability to deal with unexpected events and provide the right support to enable risk management [53]. There are several reasons why ML can be an attractive solution:

- Do it *faster*: do something which is currently performed by risk analyst, but faster;
- Do it *better*: do something which is currently performed by one or more humans, but in a more consistent way;
- Do something *new*: do something which is currently infeasible.

The case studies results show a remarkable accuracy for the tested models, affordable predictions, and general suitability to meet risk assessment needs. Some issues should be considered. First of all, the importance of applying a proper ML pipeline (Fig. 22). A ML pipeline is a way to codify and automate the workflow it takes to produce a machine learning model. Machine learning pipelines consist of multiple sequential steps that do everything from data extraction and pre-processing to model training and deployment.

The importance of data pre-processing should be underlined. Data pre-processing is an important step to prepare the data to build a ML model. Data pre-processing is a data mining technique to turn the raw data gathered from diverse sources into cleaner information.

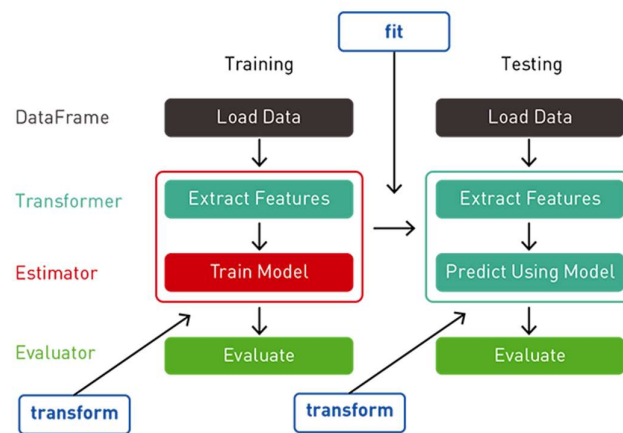


Figure 22: An example of ML pipeline.

In fact, the most common problems you can find with raw data can be divided into 3 groups:

- *Missing data*: this is inaccurate data since the information that isn't there creates gaps that might be relevant to the final analysis. Missing data often appears when there's a problem in the

collection phase, such as a glitch that caused a system's downtime, mistakes in data entry, or others.

- *Noisy data*: this group encompasses erroneous data and outliers that you can find in the data set but that is just meaningless information.
- *Inconsistent data*: inconsistencies happen when files with similar data in different formats and files are kept. Duplicates in different formats, mistakes in codes of names, or the absence of data constraints often lead to inconsistent data, that introduces deviations to deal with before analysis.

Without an appropriate data pre-processing, the final output would be plagued with faulty insights.

Another issue is represented by unbalanced data. Unbalanced datasets are prevalent in a multitude of fields and sectors, and of course, this includes industrial processes data, where the vast majority of data are representative for well-going operations. The challenge appears when machine learning algorithms try to identify the rare cases of accidents / near misses in rather big datasets. Due to the disparity of classes in the variables, the algorithm tends to categorize into the class with more instances, the majority class, while at the same time giving the false sense of a highly accurate model. Both the inability to predict rare events, the minority class, and the misleading accuracy detracts from the predictive models. The algorithm receives significantly more examples from one class, prompting it to be biased towards that class. It does not learn what makes the other class different and fails to understand the underlying patterns. For addressing this issue, the solutions are to resample the data for training, by selecting an appropriate amount of the less represented class, or introduce, in the model, penalties for misclassifications from the minority class more than the majority.

A general conclusion of the section 1, is the evidence that, in order to combine all of the relevant RA aspects, including Boolean events, continuous process variables, and interdependencies, in a dynamic model, a combination of different strategies is required:

- *Artificial Neural Networks (ANN)* for predicting the critical process variables;
- *Hierarchical Bayesian Networks (HBN)* for combining in a dynamic way all the Boolean RA elements, coming from Fault Trees and Event Trees;
- *Hidden Markov Models (HMM)* for combining, in an *abductive inferential reasoning system*, continuous, categorical, and Boolean elements, for obtaining an overall predictive model for the whole system.

The next step in designing the whole picture, is to integrate in the dynamic risk model the concept of resilience and investigate how the machine learning algorithms can be used in a complete resilience assessment.

SECTION 2

2. THE RESILIENCE OF COMPLEX SYSTEMS

2.1. (MIS)UNDERSTANDING RESILIENCE

Over the past decades, few concepts have gained such prominence as resilience. There has been an explosion of research and policies into ways to promote resilient systems, but the content has often lacked a clear definition of what resilience means, let alone how to apply resilience thinking. The concept of resilience has since taken many directions, reflecting the versatility of the term, and has been applied to ecological, psychological, social, socio-technical, organizational, and social-ecological systems.

Resilience is the capacity of a system to deal with change and continue to develop [54]. Resilience thinking is about generating increased knowledge about how we can strengthen the capacity to deal with changes. It is about finding ways to deal with unexpected events and crises and identifying sustainable ways for prospering within system's boundaries.

There are various aspects of the resilience theory that have been misrepresented and misunderstood in the literature and practically within organizations, and this leads to apparent divergences in safety management approaches.

Safety management, as it is frequently described in the literature and applied in practice, involves a strong focus on standardization and compliance, aiming to align individuals with organizational safety requirements and ideals. This safety management approach begins with central determinations of what is safe, and then works to implement mechanisms to align operational work with this plan through prescribed roles, requirements, and procedures. Safety management then focus on identifying deviations from prescribed work which need to be detected and eliminated. Over the past twenty years, this dominant views of safety within organisations have been increasingly challenged by the resilience theories. These theories suggest the importance of a focus on decentralization, or, more specifically, the capacity of organizations to guide adaptability of people and systems, through understanding and supporting how complex systems usually succeed, but sometimes fail. Organizational systems succeed despite the basic limits of predetermined plans, in a complex, interdependent and changing environment [58]. This is the heart of a resilient system. A resilient approach to safety focuses on looking for the different ways systems adapt to unexpected events.

Resilience engineering theory introduces the concept of *guided adaptability* [23], but often got misinterpreted as shown in the control-adapt paradox (§ I.3.1.). Guided adaptability is not about choosing between control or adaptation, but about *helping safe variations happen, and helping variations be safe* [58].

Adaptability is the primary, and defining, trait of resilience. If you are to provide specific operations, and your systems go down, there is no more room for operations.

Humans have long been the primary agent in making systems adapt. It has been people who are on-the-ready to investigate and get the system back up and running as quickly as possible, to make a system resilient to failure. Human attention has been required to ensure system resiliency. But human work is old-

school in the age of industry 4.0. According to [72], 50% of the choices have to do with human error or the necessity of human intervention.

So, part of the reason that resilience strategies aren't always effective is because they're coming from a place of misunderstanding. Just having strategy, policy and training doesn't cut it; knowledge doesn't necessarily lead to behavioural change. Resilience is about *culture* and *practice*, more than strategy and policies.

A well-known quote from Peter F. Drucker says that "*culture eats strategy for breakfast*", and this is the point! *Culture* basically means having a *Systemic approach*, thus catching the interdependencies among each system component (§ I.3.1.). Drucker pointed out the importance of including the human factor as one of the preeminent system components. The safety culture of an organisation is the *product of individual and group values, attitudes, perceptions, competencies, and patterns of behaviour that determine the commitment to, and the style and proficiency of, an organisation's health and safety management. Organisations with a positive safety culture are characterised by communications founded on mutual trust, by shared perceptions of the importance of safety and by confidence in the efficacy of preventive measures* [57].

2.1.1. Antifragility and resilience

Taleb, in [71], defines the concept of *antifragility*. For Taleb, the antifragile concept is a blueprint for prospering in a black swan environment (where surprising extreme events may occur). According to Taleb, prediction and risk management should be replaced by processes that allow the systems to move from fragility to antifragility.

In my opinion there is a kind of conflict in those concepts. First of all, prediction and risk assessment are obviously crucial for moving in the right direction toward antifragility, and, moreover, antifragility, following the same approach above depicted for safety, can be seen as an emerging property as well, where resilience still is one of the enabling properties. Antifragility represents a useful contribution to the practice of risk analysis. However, what is inaccurate in [71], is that the main goal of risk management is not to accurately estimate rare event probabilities, but to reveal and assess uncertainties, and make adequate decisions under uncertainty. Which is pretty the same concept Taleb associate to antifragility.

Similar concepts are reported in [72], where is underlined that, in the professional risk management context, it is more common to address the vulnerability and resilience of the system than its fragility. In fact, it's hard to see what the fragility concept covers that is not also captured by the vulnerability and resilience concepts. Vulnerability and resilience also relate to consequences of stresses. If a system is easily broken, it is vulnerable and not resilient. As it will be shown, the opposite also holds under some conditions. If the consequences of some stresses are (likely to be) severe, that is, the system is vulnerable, the system is also easily damaged from this stress; it is fragile with respect to this type of stress. Surprising events (the so called *black swan* type of events) may also occur in relation to the judgments made of robustness (and vulnerability and fragility). The point being made is that uncertainties and surprises need to be incorporated

in the concepts and measurements of fragility, vulnerability, and resilience to make them meaningful in a practical context.

Given this understanding of the antifragility concept, what does it add to risk analysis practice compared to vulnerability and resilience?

According to [72], the key contribution is the antifragility concept's idea of linking variation, uncertainties, and risk at the stress level to the positive and negative "risk" related to future performance. Robustness and resilience address both the stress dimension, but do not see these in relation to future developments that extend beyond established functions.

The development of the here presented system for Dynamic Asset-integrity and Risk Management (DARMS), fully integrates the concepts of dynamicity, resilience and antifragility.

2.1.2. Defining Resilience Engineering

There is no universally agreed definition of resilience engineering (RE); but it involves a collective aspect, is multifactorial, multilevel, and multidimensional; associated with the already mentioned four key principles (anticipation, response, learning and monitoring) and successful outcomes [55].

RE was introduced in the safety domain in 2003 as an alternative to explaining how organizational disasters such as Challenger and Columbia occurred repeatedly [56]. Unanticipated disasters underline the need of a new framework of systems safety that can cover unanticipated situations that spill out of the scope of conventional risk management.

The above examination suggests that a clear formulation of RE is lacking, and it is still in the "midst of defining itself" [59], or "a family of ideas" [60]. What is clear that it is a complex phenomenon, involves adjustments and adaptations, includes the concept of proactive management of safety risks, and is a feature that extends beyond individual components.

Facing the fact that complex systems are subject to failure, lead over time to the construction of multiple layers of defense against failure. These defenses include obvious technical components (e.g. backup systems, 'safety' features of equipment) and human components (e.g. training, knowledge) but also a variety of organizational, institutional, and regulatory defenses (e.g. policies and procedures, certification, work rules, team training). The effect of these measures is to provide a series of shields that normally divert operations away from accidents. Recognizing hazard and successfully manipulating system operations to remain inside the tolerable performance boundaries requires intimate contact with failure. More robust system performance is likely to arise in systems where it's possible *discern where system performance begins to deteriorate, becomes difficult to predict, or cannot be readily recovered* [73]. In intrinsically hazardous systems, operators are expected to encounter and appreciate hazards in ways that lead to overall performance that is desirable. Improved safety depends on providing operators with comprehensive views of the hazards.

The practice starts from accepting the reality that failures happen, and, through engineering, builds a way for the system to continue despite those failures. Good resilience engineering produces a system that can adapt. Resiliency can be built into any system, and it offers a lens to look at critical areas and operations. Resilience engineering is about readiness and adaptability. When a system is designed to be resilient, it means the system can encounter failures, and find a way to keep on keeping on.

RE represents a sophisticated way for managing safety; the sophistication not being in the technology, but in the way one thought about safety, accidents, and risks, and how these could be better managed using existing approaches, methods and approaches but in more innovative ways [54].

So, the following definitions can be proposed [42]: *Safety can be defined as an emergent property of complex socio-technical systems, where resilience is the key enabling property.* Starting from those concepts, a resilience assessment framework is defined.

2.2. A FRAMEWORK FOR RESILIENCE ASSESSMENT

All systems have an envelope of performance, or a range of adaptive behavior, due to finite resources and the inherent variability of its environment. Thus, there is a transition zone where systems shift regimes of performance when events push the system to edge of its envelope.

The real challenge for resilience assessment (RA), is to *understand, characterize, and model* the complexity of different interdependent components and sub-systems, which constitute a non-linear complex system in the *black swan* arena.

Resilience assessment fundamentally maintains much of the same philosophical background as traditional risk assessment. However, resilience analysis additionally delves into the *unknown, uncertain and unexpected at the scale of systems rather than individual components* [62]. Resilience thinking requires practitioners to ponder potential future threats to system stability and develop countermeasures or safeguards to prevent longstanding losses. Resilience analysis primarily focuses on outcomes: practitioners are directly concerned by the ability of the impacted organization, infrastructure, or environment to rebound from external shocks, recover and adapt to new conditions. In other words, where traditional risk assessment methods seek to harden a vulnerable component of the system based upon a snapshot in time, resilience analysis instead seeks to offer a ‘soft landing’ for the whole system.

The increasing interest about the resilience assessment is to be understood in the deep change of paradigm from the prescriptive approach to the performance-based one that continuous innovation in technology is asking for. Risk analysis in a process plant is a matter of hazard identification and assessment of possible upset scenarios consisting of chains of cause-and-effect events modelled by fault and event tree or bow tie. One of the reasons for the superior attention to resilience is the recent increased capability of data measurement/storage and relevant treatment for developing knowledge. The concept and the techniques to realize a resilient plant are still under investigation and are detailed in the seminal paper by

Jain et al. [61] presenting the large potential with issues such as error-tolerant equipment design, receptive to early warning signals during operations, with a ‘plasticity’ response and effective emergency response.

There is no single accepted set of components of resilience, so, the framework proposed in the present work, which is strictly related to the state-of-art of scientific literature, represents a robust approach to a systemic vision of safety management.

The focal point to assess the resilience of a system relies in the identification of precursor events, which refers to early detection of “weak” signals from the system during the operations [22]. To identify the precursor events and thus maintain stability by applying appropriate adjustments, the analysis of a large amount of data is needed. By the data analysis, it is possible to predict the behavior of the system, thus catching the *resilient performance* according to the above mentioned four guidewords. As commented by Sarkara et al. [64], accidents do not occur in a chaotic fashion, so underlying patterns and trends do exist and can be captured.

Data-driven modelling can be considered as an appropriate approach to resilience assessment that would complement the “knowledge-driven” models describing physical behavior. As already mentioned in (§ 1.3.1.), the Bayesian approach has been proven to be a robust probability reasoning method under uncertainty, providing a tool for incorporating evidence during operations. Hidden Markov Models (HMMs) seems to be one of the most promising and reliable approaches [65].

In the presented approach, HMMs are adopted as reference tool.

The logic diagram for the proposed resilience assessment framework, in terms of stepwise procedure, is depicted in Figure 23, where the previously recalled capabilities (monitor, learn, anticipate, and respond) identifying resilience performances can be pointed out.

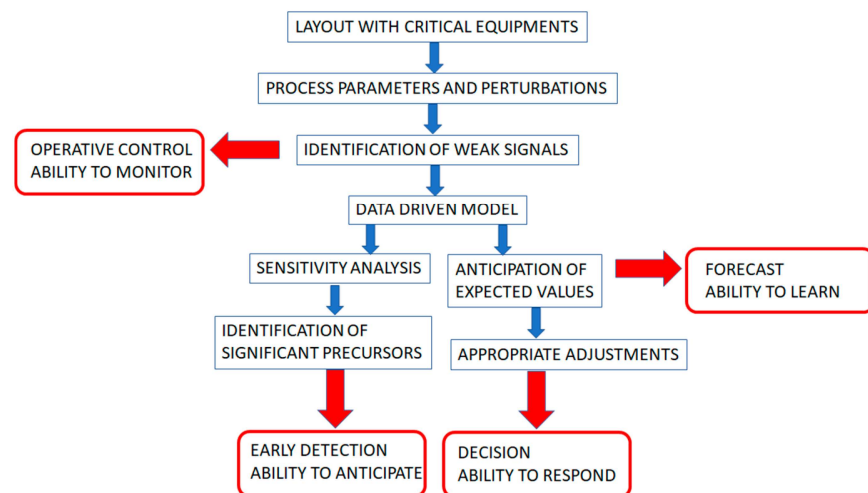


Figure 23: The resilience assessment framework

Starting from these premises, the integrated approach to carry out a resilience assessment in complex systems can summarized as follows.

Identification of Weak Signals

A weak signal indicates a possible degradation of the system's resilience and represents a decreased ability to cope with unexpected and unforeseen disruptions. They are seemingly random or disconnected pieces of information that at first appear to be irrelevant but can be recognized as part of a significant pattern by viewing them through a different frame or connecting them with other pieces of information. Weak signals can be identified starting from the risk assessment process, by analyzing how the process variables oscillate around the set points. Monitoring the critical process variables and their deviation from the set points allows establishing appropriate operational control strategies.

Data Driven Models and Precursors Identification

As widely discussed in Section 1, the focal point of ML models is the investigation of data. All the dependencies, correlations, inference statistics can be found in data, so it is crucial to build a good data-driven model for extracting all the information usually contained in the data. It is thus possible to identify the significant perturbations and, by training the model, anticipate the systems outcome, to improve decision-making and promptly choose the appropriate adjustments.

As previously anticipated, for analysing the significant parameters perturbations, and thus identifying and anticipating the weak signals, a Hidden Markov Model (HMM) can be applied.

2.2.1. Hidden Markov Model (HMM)

An HMM is a generative probabilistic model, in which a sequence of observable X variables is generated by a sequence of internal hidden states Z . The hidden states are not observed directly. The transitions between hidden states are assumed to have the form of a first order Markov chain. They can be specified by the start probability vector π and a transition probability matrix A . The emission probability of an observable can be any distribution with parameters θ conditioned on the current hidden state. The HMM is completely determined by π , A and θ .

In the present work, the hidden states are the states between a regular performance and a failure of a sub-system. The only known states are the first (the component is performing well) and the last (the component fails), and the hidden states in between may represent the precursors of accidental events. The emissions of the system are the process variable values.

Three possible implementations of the HMM are evaluated.

- In the first implementation, besides the observations, also the transition probabilities (derived a priori from FT, for the last state), and the emission probabilities (derived from expert knowledge) are inserted, in the form of a transition matrix and emission matrix. The model determines the most likely sequence of states by inference (MC sampling with rules) on the observations;
- The second model has the same observations and transition probabilities as the first one. The emission probability and the most probable sequence of states are determined by inference;

- In the third model, only the observations are given. There is no information on either transition or emission probabilities. The model can infer all the information and determine the most likely sequence of states.

The heart of an HMM is the inferential process, made by the means of the *Viterbi* algorithm. Given a HMM with state space S , initial probabilities π_i of being in state i , transition probabilities $a_{i,j}$ of transitioning from state i to state j , and a set of observations y_1, \dots, y_T , according to the *Viterbi* algorithm the most likely state sequence x_1, \dots, x_T , that produces the observations is given by the recurrence relations:

$$V_{1,k} = P(y_1 | k) \pi_i \quad (9)$$

$$V_{t,k} = \max_x (P(y_t | k) a_{x,k} V_{t-1,x}) \text{ with } x \in S. \quad (10)$$

Here $V_{t,k}$ is the probability of the most probable state sequence $P(x_1, \dots, x_t, y_1, \dots, y_t)$ responsible for the first t observations that have k as its final state. The *Viterbi* path V can be retrieved by saving back pointers that remember which state x was used in the second equation.

The HMMs are developed in python using the packages *PyMC3* and *Theano* [66].

PyMC3 has been used for the implementation of the Metropolis–Hastings (MCMC-MH) algorithm to perform forward and backward inference by computing the distribution space of the model parameters and determine the most likely outcome. This technique requires a simple distribution called *the proposal* $Q(\theta'|\theta)$ to help draw samples from an intractable posterior distribution $P(\theta=\theta|D)$. These two distributions are called *conjugate distributions*.

MH uses Q to randomly walk in the distribution space, accepting or rejecting jumps to new positions based on how likely the sample is.

To decide if θ' is to be accepted or rejected, the following ratio must be computed for each new proposed θ' :

$$\frac{\prod_i^n f(d_i | \theta = \theta') P(\theta')}{\prod_i^n f(d_i | \theta = \theta) P(\theta)} \quad (11)$$

where f is the above-mentioned proportional function. The acceptance rule is set as follows:

- If (Equation (11)) < 1: $P(\text{accept}) = (\text{Equation (11)})$
- If (Equation (11)) ≥ 1 : $P(\text{accept}) = 1$

This means that if a θ' is more likely than the current θ , then θ' will always be accepted. If it is less likely than the current θ , then it might be accepted or rejected randomly with decreasing probability, the less likely it is.

By varying one input at a time (process variables values) and correspondingly analyzing the output sequences, it is possible to perform a sensitivity analysis to determine the most critical variable.

To evaluate the application of the proposed resilience assessment framework, it has been tested on a LNG bunkering operation.

2.3. AN APPLICATION OF THE RA FRAMEWORK

The LNG-based fuel system technologies show a better sustainability performance than the conventional marine fuel technologies [67]. LNG is natural gas cooled to approximately $-260\text{ }^{\circ}\text{F}$ ($-162.7\text{ }^{\circ}\text{C}$) and it is reasonably easy to store and transport it. In its liquefied state, LNG is odorless, colorless, non-toxic, and non-corrosive.

A main issue connected to the adoption of LNG as fuel is the lack of bunkering (fueling) facilities available yet, so getting an LNG-powered ship re-fueled may be problematic. Though there are plans for more fueling depots to be established to serve LNG-fueled ships, some worries are related to the possible threat that this plant and its operational activity might create, several administration and port authorities are addressing the safety issues the refueling operations might create for citizens and coastal environments.

Basically, the possible technical solutions under the current development in Europe and worldwide are [68]:

- Truck-to-Ship—TTS;
- Ship-to-Ship—STS; and
- Terminal (Port)-to-Ship—PTS

The resilience assessment will be carried out for shore-to-ship refueling and a schematic layout is depicted in figure 24.

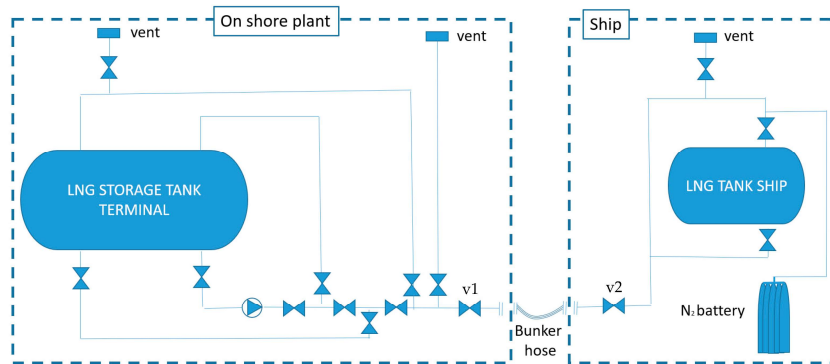


Figure 24: Representative layout of a shore-to-ship LNG refueling plant.

The basic steps of the bunkering process are summarized as follows, derived with some simplification from EMSA [68]:

- Precooling of the line (landside), cargo pump included;
- Actions to avoid ground fault arcing;
- Loading arms are usually used for bunker hose connection;
- The hose is put in place;
- Inert gas is used to remove oxygen and moisture from the piping of the receiving ship;

- Then, the receiving system is purged from the residual nitrogen using the natural gas remained in the LNG tank on board the ship;
- Closure of the onshore side valve (v1);
- Closure of the ship side valve (v2);
- Liquid line stripping;
- Bunker line inerting; and
- Disconnection of the bunkering hose.

For the purpose of this applicative study, specific attention will be given to the liquid transfer phase, with focus on the following critical steps:

- Analysis during the actual bunkering phase; and
- Analysis during the immediate post-bunkering phase with the pressure increment.

As already discussed, resilience analysis requires in the first instance to make a distinction between the condition at a given moment of time, i.e., static (performed adopting conventional QRA approaches), and trends during plant operation, i.e., dynamic (relying on innovative tools). The whole resilience assessment can be very large; to understand the validity of the proposed methodology, the investigation will be limited to the leakage hazard originating in the part of the system between the two flanges of the connecting hose, technically indicated as “LNG transfer system” [69].

2.3.1. Fault Tree Analysis

The FTA for the bunkering operations is carried out starting from technical reports [70]. The following Table 5 and Table 6 summarize the FTA elements.

Table 5: Equipment Count for Leak Frequency Estimate.

Root Component	Quantity (Diameter)
Manual valves	3 (3 in.)
Activated valves (ESDs)	2 (3 in.)
Flanges	12 (3 in.)
Small bore fittings	2 (1 in.)
Flexible hose	1 (3 in.)
Manifold piping	100 m (3 in.)

Table 6: Risk assessment assumptions.

Client Type	Source (m ³)	Client (m ³)	Rate (m ³ /h)	Op. time (h)	F (occ/y)
Ferry	500	200	50	4	365

OSVs	400	200	2	183
Container	2400	600	4	52

The developed FTA related to leakage hazard originated in the LNG transfer system is shown in Figure 25.

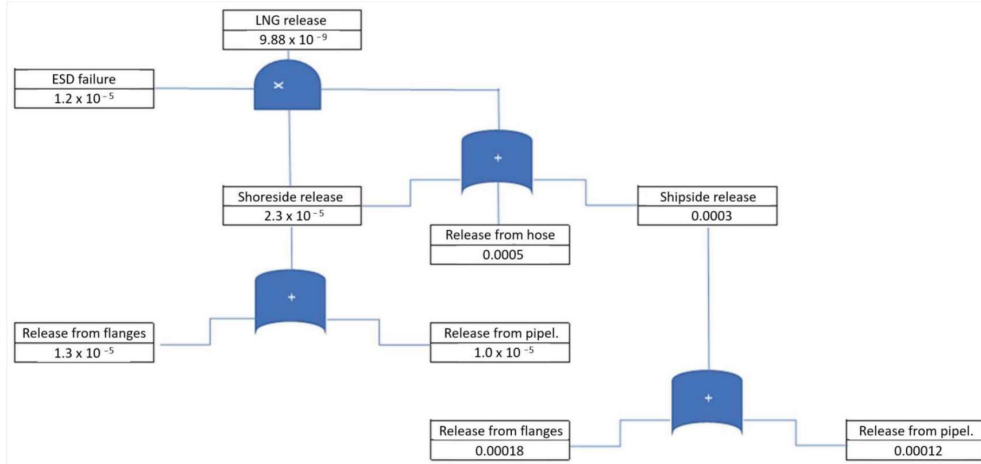


Figure 25: Fault tree developed for LNG loss of containment (LOC).

The failure probability of the single component is taken from the literature.

The relevant process variables in the LNG bunkering operations are pressure, temperature, and bunkering rate. The conventional pressure in a bunker hose is around 5–6 bar (g) and the corresponding LNG temperature during bunkering is around $-145\text{ }^{\circ}\text{C}$, based on the assumption that bunkering operations can maintain a constant temperature by managing the boil-off vapors.

2.3.2. Development of the predictive model

The FTAs can be dynamically updated by considering the root failures frequency as prior probabilities, applying the above mentioned MCMC-MH algorithm to update the basic event on real time information, thus reducing uncertainty and capturing the dynamic behavior of the system. The posterior probabilities are estimated with reference to the solution of the FTAs, as depicted in Figure 26.

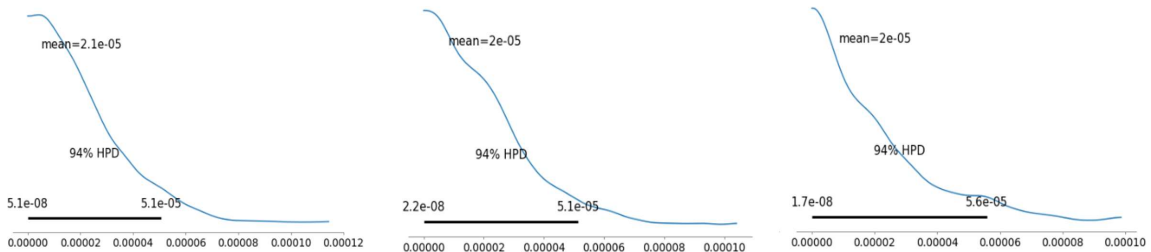


Figure 26: Posterior probabilities distribution for the leakage on the shipside (left), shoreside (center) and connection (right).

At last, by combining the contributions, it is possible to obtain the posterior predictive PDF as exemplified in Figure 27, where the red line indicates the expected probability of system failure in that moment.

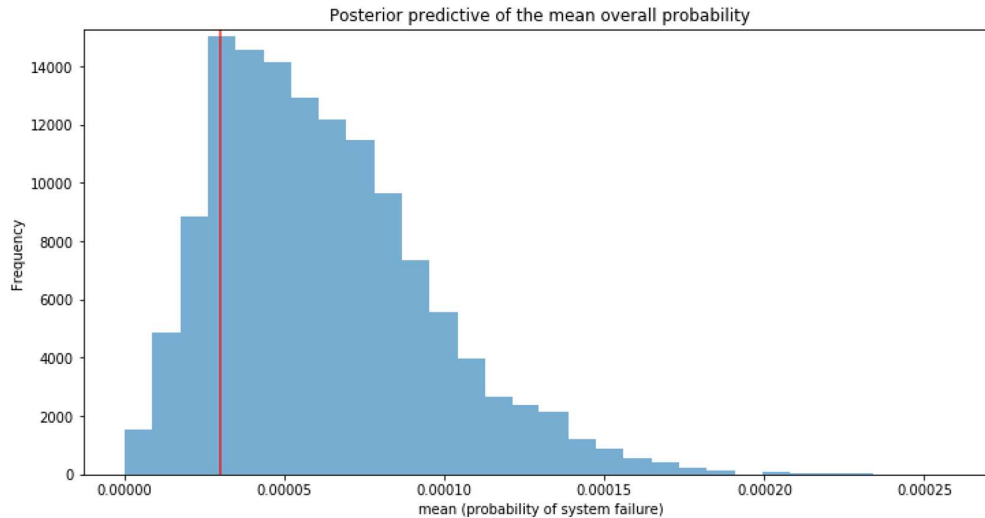


Figure 27: posterior predictive distribution, and expected probability of system failure.

All the detected signals and indicators will be properly treated by the HMM approach, as described above. The input parameters are the process variables, pressure, temperature and density, each one with appropriate set points as detailed in the following.

LNG line process equipment and hose

- Operating pressure is set to 10 bar(g). This is the maximum operating pressure for LNG process equipment according to European design standard EN1472-2;
- Operating temperature is set to $-162\text{ }^{\circ}\text{C}$ to keep the inventory in liquefied state. The bunker vessel (discharging unit) is assumed to be able to maintain this constant temperature during the transportation to site;
- Density depends on temperature and pressure. Based on the defined process parameters the density is 425 kg/m^3 .

Vapor return line (NG)—process equipment and hose

- Pressure is set to 2 bar(g) as it will be reduced compared to LNG line;
- Temperature is set to $-100\text{ }^{\circ}\text{C}$. The liquid has been warmed and is now in a vapor state;
- Density 4.3 kg/m^3 .

Tank

- The pressure in the tanks is set at 2 bar(g).

The operational parameters are acquired by the model using json (Javascript object notation) [65]. An extract from the input dataset for each sub-section is provided in Table 3.

Table 7: Head of an input dataset.

Timestamp	Pressure (Barg)	Temperature (°C)
2020-10-21 10.10.00.000	9.88	-162.02
2020-10-21 10.15.00.000	9.91	-162.03
2020-10-21 10.20.00.000	9.99	-162.05

The HMM process is depicted in Figure 28, where s_0 - s_n represent the simulated future time-steps, while the bar-charts on the left hand side represent the most probable expectations deriving from the inferential sampling shown on the right hand side (nr. of MCMC samples vs. state). By combining all the most probable outputs, the most probable state sequence of the system is at last obtained.

According to the outlined approach, the failure state is anticipated by the intermediate state, allowing an early warning. By analyzing the HMM traces, it is possible to determine the probability for each state of the given system. Additionally, based on the model outputs making backward inference, one can immediately refer the hazardous deviation to some sub-systems failure. This can strengthen the understanding and application of early process fault detection in effective process risk management.

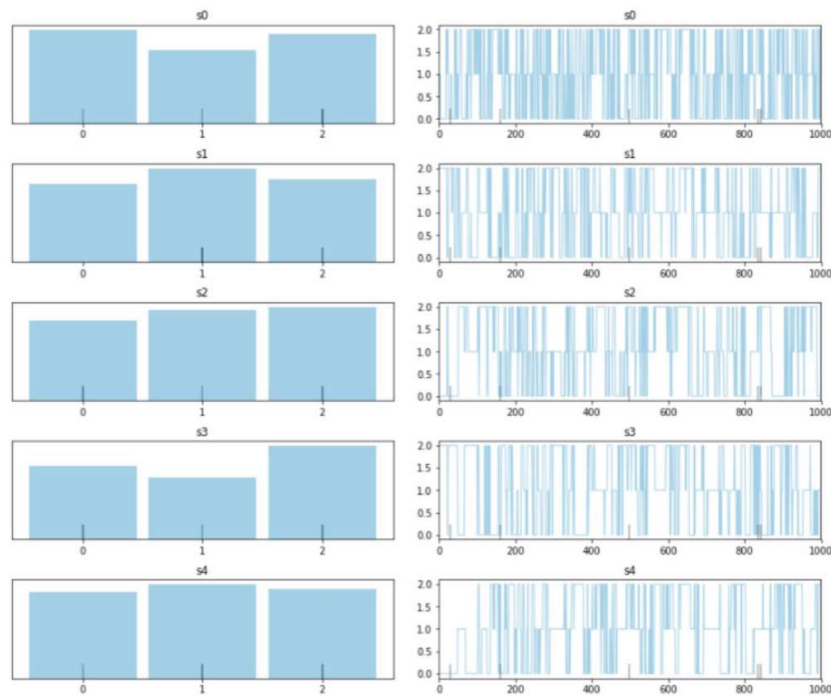


Figure 28: Excerpt obtained by HMM inferential sampling and prediction.

Table 8 summarizes the expected probabilities for the states of the root components (expressed in terms of MAP—Maximum A Posteriori), obtained by the Resilience model and compared with the results of the traditional FTA.

Table 8: Expected probabilities of occurrences of possible states in the root components.

Root Component	Traditional FTA	Resilience Model
SHORESIDE VALVES		
Safe	0.999	0.228 (MAP)
Intermediate	NA	0.761 (MAP)
Fail	1.2×10^{-6}	$1.6 \times 10^{-8} - 1.6 \times 10^{-5}$ (94%HPD)
PUMP		
Safe	0.999	0.166 (MAP)
Intermediate	NA	0.833 (MAP)
Fail	1.3×10^{-6}	$1.8 \times 10^{-8} - 1.8 \times 10^{-5}$ (94%HPD)
SHORESIDE PIPELINE		
Safe	0.999	0.387 (MAP)
Intermediate	NA	0.612 (MAP)
Fail	1×10^{-6}	$1.7 \times 10^{-8} - 1.7 \times 10^{-5}$ (94%HPD)
HOSE		
Safe	0.999	0.055 (MAP)
Intermediate	NA	0.854 (MAP)
Fail	7.5×10^{-6}	$5.7 \times 10^{-9} - 1.8 \times 10^{-5}$ (94%HPD)
SHIPSIDE VALVES		
Safe	0.999	0.297 (MAP)
Intermediate	NA	0.702 (MAP)
Fail	1.8×10^{-6}	$8.7 \times 10^{-9} - 1.7 \times 10^{-5}$ (94%HPD)
SHIPSIDE PIPELINE		
Safe	0.999	0.307 (MAP)
Intermediate	NA	0.692 (MAP)
Fail	1.2×10^{-6}	$1.2 \times 10^{-8} - 1.7 \times 10^{-5}$ (94%HPD)

2.3.3. Definition of a resilience score

A resilience indicator can be estimated as the dynamic response of a system (sub-system, system variables) to disturbances; the resilience of the whole system can thus be expressed as an indication of how likely the system is changing its state.

Combining the most probable sequence of system states, with a Monte Carlo asynchronous sampling from the posterior predictive probability of failures (Figure 26), into the same HMM model, it is possible

to represent the *resilience score* with a single parameter R varying between 1 (corresponding to safe system mode) and 0 (corresponding to failure system mode). The value of R can be continuously updated with the state evidence obtained by the previously mentioned random walks. Each step of the overall resilience score considers the CDF (Cumulative Distribution Function) of the different state probability.

Figure 29 represents how the resilience score R is changing over time during the operation, when different perturbative situations appear. In this way, the approach would open the way for continuous monitoring of the resilience level and provide anticipated indications for when and where to adopt corrective actions.

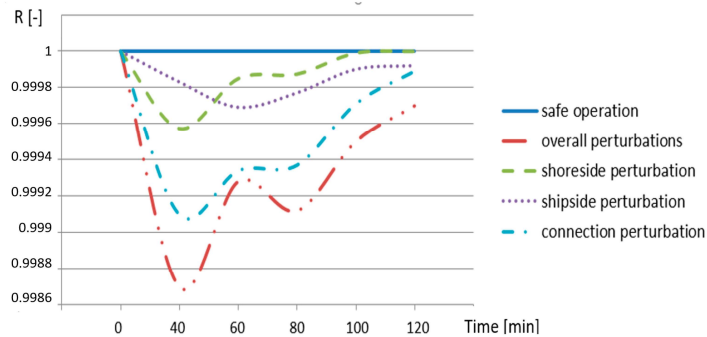


Figure 29: Value of the resilience score over time, corresponding to different perturbative situations.

The approach used for the presented case study, based on Bayesian statistical modelling and probabilistic machine learning, which focuses on advanced Markov Models and variational fitting algorithms, has proven to be a useful and flexible tool to study, analyze and verify the achievement of the four basic needs of the resilience paradigm.

2.4. CONCLUSION OF SECTION 2

The idea behind the resilience analysis is that safety is an emerging property of the system that can be framed with the four guidewords, i.e., monitor, learn, anticipate and respond.

RE aims at providing tools to proactively manage risk, acknowledging the inherent complexity of system functioning and the correspondent need for performance variability. In complex systems, safety is not a constant property, it's a continuous function of the interactive properties and activities of its constituent components [74]. Safety is related to how system performs, generating the need to focus on whole system and the connection between agents, rather than individual agents.

So resilience is a feature of some systems that allows them to prosper even during unanticipated disturbances that can lead to failure.

To catch the intrinsic dynamicity of a complex system, it would be possible to sketch resilience according to the different aspects to threat, i.e. the *predictability* of the threat, its *potential to disrupt* the

system, and its *origin*. Resilience is situational, based on specific capabilities to manage different typologies of threats [1], generating a strong consensus about identifying different patterns in adaptive capacity.

It's clear that resilience engineering is a more comprehensive and advanced concept of risk management. This notion is based on a systemic view of accidents, i.e., accidents are caused by a nonlinear combination of performance variability of system functions rather than a linear combination of component failures [75].

Recalling the four needs for resilient performance, the following achievements of the proposed model can be underlined:

- The model allows identifying how the state of the plant is changing over time, thus detecting the occurrences of perturbations during the operations and responding to the perturbation. The intermediate state defines the precursor of a perturbative event;
- The approach is able to monitor by analyzing in real time the data derived from the plant, finding the corresponding actual state;
- Through the learning Bayes-based algorithm, the model can produce increasingly reliable forecasts on the progress of the operation, as the training dataset is constantly updated by the actual operative evidence;
- By identifying the precursor events, the model anticipates the states transitions, providing an early warning to take appropriate countermeasures.

The main features of the model are the following:

- A dynamic representation of the loss of containment risk, related to the values of the process variables, is obtained by combining a Bayesian network for inferential sampling, with an HHM in a resilience model for the determination of hidden states probabilities;
- The sequences of the most probable system states represent relevant information for taking the most appropriate actions on time, in order to increase the *antifragility* of the system.

The proposed resilience score R can represent a valid metric to define how much the perturbations of systems and subsystems can be absorbed without leading to failure.

The application case show that, upon proper refinement, the approach can be effectively used to capture the dynamic evolution of internal and external risk conditions and effectively support critical decisions to improve the overall safety of a system.

As a general conclusion of the Section 2, the following statement should be emphasized: It is possible to evaluate the system's resilience through dynamic analyses, connected with what is happening in the plant at that precise moment in the operation in progress, and that a powerful tool for performing resilience assessment is represented by proper, precisely tailored, Machine Learning algorithms.

The next step will be identifying a set of top-level principles enabling the resilience against disruptions caused by threats of various sources, and defining them as system properties.

SECTION 3

3. ENGINEERING SYSTEMS RESILIENCE PROPERTIES

According to Woods [76], *RE uses the insights from research on failures in complex systems, including organizational contributors to risk, and the factors that affect human performance to provide systems engineering tools to manage risks proactively.*

The assumption is that resilience can be engineered into a complex system, in order to support the use of adaptive capacity. RE recognizes that a portion of variability is unavoidable and beneficial (§ 2.2.1.), and due to this fact it should be managed rather than dampened.

3.1. SYSTEMIC RESILIENCE

The systemic notion of resilience will not distinguish between normal and abnormal system conditions. While conventional risk management aims at maintaining risks below the acceptable limit, risk management in resilience engineering aims at enhancing the ability of a system to absorb performance variability under uncertainties. Resilience, therefore, deals with every system condition: stable operations in normal conditions, prevention of accidents in abnormal conditions, minimization of losses after accidents, and fast recovery.

The International Council on Systems Engineering (IN-COSE) Handbook [77] defines a system as *an integrated set of elements, subsystems, or assemblies that accomplish a defined objective. These elements include products (hardware, software, and firmware), processes, people, information, techniques, facilities, services, and other support elements.*

Jackson and Ferris [78] reported the following global principles:

- *Adapt* means to restructure before or during an encounter with a threat. Or it can mean to sense the approach of a threat and execute means to avoid it or minimize damage from it.
- *Avoid* means to eliminate contact between the system and the threat and to suffer no damage or disruption of functionality from the threat.
- *Prepare for* means to engineer the system in advance to enable recovery following an encounter with a threat.
- *Withstand* means to retain full functionality following an encounter with the threat.
- *Recover* means to retain or restore full functionality following an encounter with the threat.

In the present section, the aim is find ways of addressing system resilience within Systems Engineering (SE) methodologies. It is thus useful to introduce some key concepts for defining the SE context in which resilience requirements can be properly specified.

3.1.1. Defining the SE context for resilience

Systems by definition deliver desired capability. It is the quality of the delivery of such capability – in the face of adversity – that resilience addresses [79]. Systems interact with their environments. Nominal environmental conditions often dominate the focus of SE activities. The concept of resilience explicitly adds the consideration of adversity and requires a shift in requirements analysis, architecture, and design methods to establish an approach that addresses nominal and adverse conditions under which the system should operate. An influence diagram representing this meaning is shown in Figure 30.

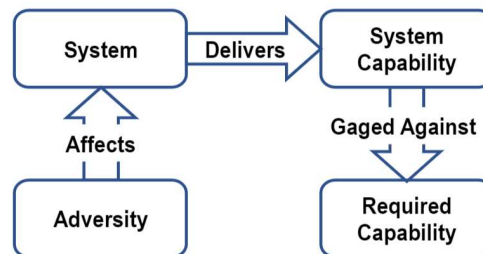


Figure 30: Influences in system resilience

The sources of adversity may be natural, technological, or human, and may include sources external to or within the system. A high-level view of the steps to assess systemic resilience should include:

- Knowledge on the system architectures and/or designs.
- Knowledge on the system functional behavior, data and control flows to deliver the required capability.
- Knowledge on the capabilities of interest, how are they measured, and the required levels of delivery.
- Knowledge on the adversities that may affect the system.
- Knowledge on the system behavior in response to the adversities.

The term *capability* represents the system's ability to achieve desired effects. This provides an umbrella term for considering the many objectives and outcomes achieved by SE activities that are relevant to resilience, such as: mission objectives, user needs, user requirements, system requirements, derived requirements, ... A fundamental understanding is that SE is not just a set of processes, but mainly a mindset, a point of view, a method for dealing with complexity, with a strong focus on the fundamental objectives to be achieved.

The first macro-phase of the SE approach is focused on the definition of needs and functional requirements, design, and validation of the system; the second macro-phase regards costs, performance, delivery and disposal. SE combines technical and design requirements with the purpose of obtaining the required level of delivery. The seven pillars of SE approach [80] are:

- *Life Cycle*: it's the time period between concept phase and retirement.
- *Gates*: are intended to ensure a safe progression along the project life cycle by providing intermediate checks during the development.
- *Requirements*: connections between the problem space and the solution space.
- *Perspectives*: integration of the different stakeholder's viewpoints in the early design stage.
- *Trade-offs*: the project is a matter of decisions. The goal is to find the optimal solution satisfying all requirements.
- *Modelling and simulations*: to forecast efficiency and performance of the system.
- *Operational effectiveness*: must be ensured in a long-term vision of the project.

The seven pillars allow to develop a broader view of the system. The approach can be represented in the so-called V-model (represented in Figure 31), which represents a useful guideline to manage the project, starting from the definition (*the system is right*), to satisfying the defined requirements (*it is the right system*).

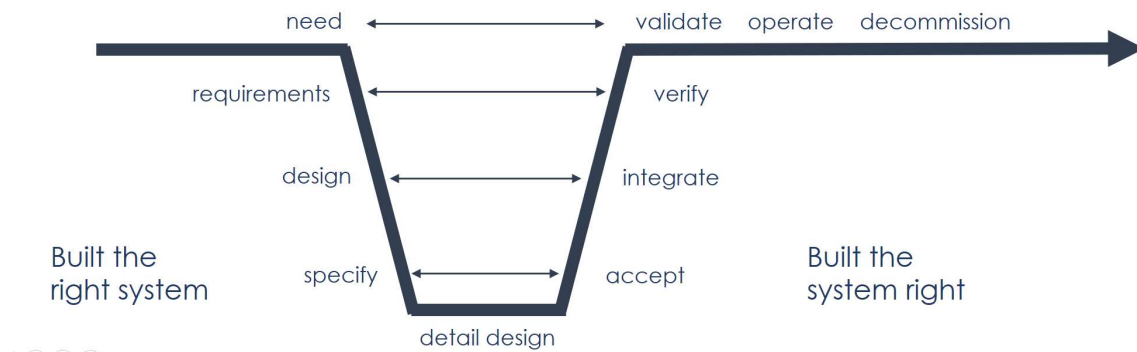


Figure 31: The V-model

In the V-model, two sides can be identified. The left one represents the analytical approach to the problem. Here a complex system is divided into sub-systems, that can be easily managed. In the right side, the sub-systems are combined back, so the whole system can be validated for ensure that the requirements are met. The process is a sequence of steps:

- design;
- detail design;
- implementation;
- verification;
- validation.

Verification ensures that all the steps to reach the goal have been well developed, validation aims to test whether the sub-systems put together meet the requirements and that the whole system fulfills the requested performance.

Business and stakeholder needs establish the foundation for resilience analysis. Stakeholder objectives are the reasons that capabilities are needed, and will be the basis for identifying the needed capabilities, and the expectation of performance in the face of adversity. Capabilities may or may not have direct value to the stakeholder, but they are needed – and justified – by their ability to support the achievement of the objectives, which have intrinsic value.

Scenarios are a useful way to represent business and stakeholder needs. A scenario should describe the *effect* to be achieved and the *environment* in which this will be performed. This establishes the baseline for the measures, targets and conditions (including adversities) by which acceptable capabilities will be judged. A range of scenarios should be developed that properly represents the scope of operations that the system is expected to support. An example of resilience scenario components is depicted in Figure 32.

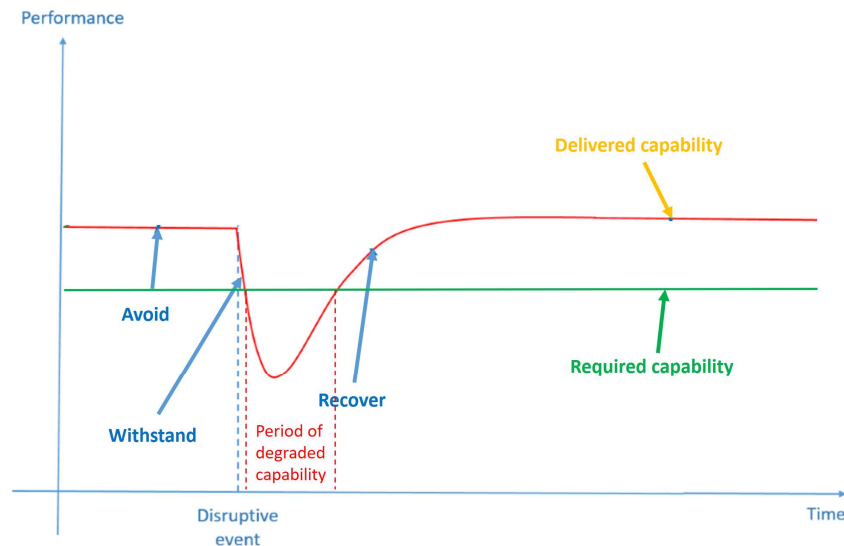


Figure 32: Notional resilience scenario life cycle components

The graph shows a highly simplified view of a resilience scenario. At the start, the system is shown to deliver more than the required capability, while the adversity may be in the environment of the system. When an adversity begins to affect the system a period of delay is shown during which the system withstands the affect, after which the system's delivery of capability degrades, eventually dropping below the required capability. After the period of affect ends the capability delivery can get back to the required level and even exceed it.

In a real situation several complications should be considered:

1. There may be multiple discrete adverse event scenarios that need to be considered.
2. There may be multiple adverse events affecting the system over the period of interest.

3. The required capability may vary in a much more complex manner.
4. As with risk, the ability to avoid, withstand and recovery from the adversity may or may not have a probabilistic component.

One approach to addressing items 1, 2 and the probabilistic aspect of 4 is to sum the probability weighted resilience of all of the relevant resilience scenarios.

While resilience should be considered throughout the systems engineering life cycle, it is critical that resilience be considered in the early life cycle activities that lead to the development of resilience requirements. Once resilience requirements are established, they can and should be managed along with all of the other requirements in the trade space throughout the system life cycle. Britis and McEvilley [78] recommend specific considerations, listed below, to be included in early life cycle activities:

- *Business or Mission Analysis Process*
 - Defining the problem space should include identification of adversities and expectations for performance under those adversities.
 - Concept of Operations (ConOps), Operations Concepts (OpsCon), and solution classes should consider the ability to avoid, withstand, and recover from the adversities
 - Evaluation of alternative solution classes must consider ability to deliver required capabilities under adversity
- *Stakeholder Needs and Requirements Definition Process*
 - The stakeholder set should include persons who understand potential adversities and stakeholder resilience needs.
 - When identifying stakeholder needs, identify expectations for capability under adverse conditions and degraded/alternate, but useful, modes of operation.
 - Operational concept scenarios should include resilience scenarios.
 - Transforming stakeholder needs into stakeholder requirements includes stakeholder resilience requirements.
 - Analysis of stakeholder requirements includes resilience scenarios in the adverse operational environment.
- *System Requirements Definition Process*
 - Resilience should be considered in the identification of requirements.
 - Achieving resilience and other adversity-driven considerations should be addressed holistically.
- *Architecture Definition Process*
 - Selected viewpoints should support the representation of resilience.

- Resilience requirements can significantly limit and guide the range of acceptable architectures. It is critical that resilience requirements are mature when used for architecture selection.
- Individuals developing candidate architectures should be familiar with architectural techniques for achieving resilience.
- Achieving resilience and other adversity-driven considerations should be addressed holistically.
- *Design Definition Process*
 - Individuals developing candidate designs should be familiar with design techniques for achieving resilience.
 - Achieving resilience and the other adversity-driven considerations should be addressed holistically.
- *Risk Management Process*
 - Risk management should be planned to handle risks, issues, and opportunities identified by resilience activities.

Requirements are a core consideration in SE. To be achieved through SE, resilience must be effectively represented as system requirements. The challenge is that resilience aggregates the considerations of functional, performance and environmental requirements. This compound requirement must be captured, so standard system engineering practices can trade system resilience against other system properties expressed in the system requirements. Thus, specifying resilience requires that several parameters, which aggregation is the resilience scenario, be identified. The following must be known in order to specify resilience:

- The capability of interest (note: a system may deliver several capabilities each of which may have different levels of resilience.).
- The measure(s) (and units) of the capability.
- The target(s) (required amount) of the capability
- System modes of operation (e.g., operational, contingency, training, exercise, maintenance, update).
- The adversity(s) being considered for this resilience scenario and the level of affect that the adversities can impose on the system.
- Understanding of the affects that the adversity imposes on the system and how the system reacts to those affects in terms of its ability to deliver capability.
- The timeframe of interest.

- The required resilience (performance) of the capability in the face of each identified resilience scenario (e.g., expected availability, maximum allowed degradation, maximum length of degradation, etc.).

Any of these factors and parameters may vary over the timeframe of the scenario, and this fact must be addressed by the systems engineer. The capability is likely to be a functional requirement of the system. Resilience then extends such requirements into a resilience scenario by adding environmental requirements (adversities) and performance requirements. This leads to a specific structure among the crucial parameters. The entity-relationship diagram for this information [78] is shown in Figure 33. This structure must be addressed in SE considerations for requirements traceability and management, architecting, design, verification, and validation.

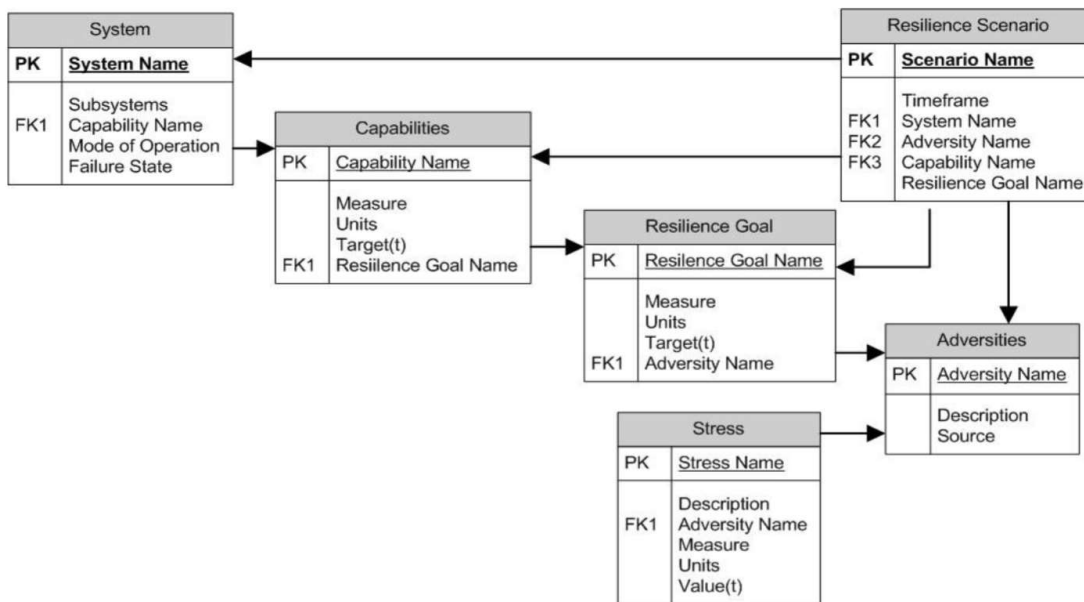


Figure 33: Data structure for specifying resilience requirements [78].

Table 9 identifies the modeling information that needs to be captured during the various lifecycle stages to support the effective development and documentation of resilience scenarios and resilience requirements.

Table 9: Modeling information during lifecycle phases

Life cycle stage	Information
Mission and Stakeholder Needs Analysis	Insert adversities in the context diagram as actors. Insert resilience scenarios as use cases.
Stakeholder Requirements	Develop use case interaction diagrams to document the interaction of actors and architectural modules during the scenarios. Develop sequence diagrams to represent the activity flow during scenarios.

System Requirements	Develop activity diagrams to show the states of the system (and adversities) during scenarios.
Architecture and System Design	Develop state models of the scenarios. Model events and signals among the architectural nodes.

The two major considerations addressed for improving SE's ability to deliver resilient systems are:

- generate quality requirements for resilience;
- extend the SE lifecycle to address resilience.

3.2. INTEGRATING THE RA FRAMEWORK IN THE SE LIFECYCLE

The next step is to check whether the developed RA framework (§ 2.2) can be integrated in the SE lifecycle design, and how the steps of RA, as stated in the proposed framework, are suitable to meet the SE requirements.

Implementing resilience from the design stage, and assessing it during the mission, allows a huge improvement in the system performance. With reference to Fig. 32, a performance curve in face of adversities, as improved by applying the resilience concepts, is shown in the following Fig. 34.

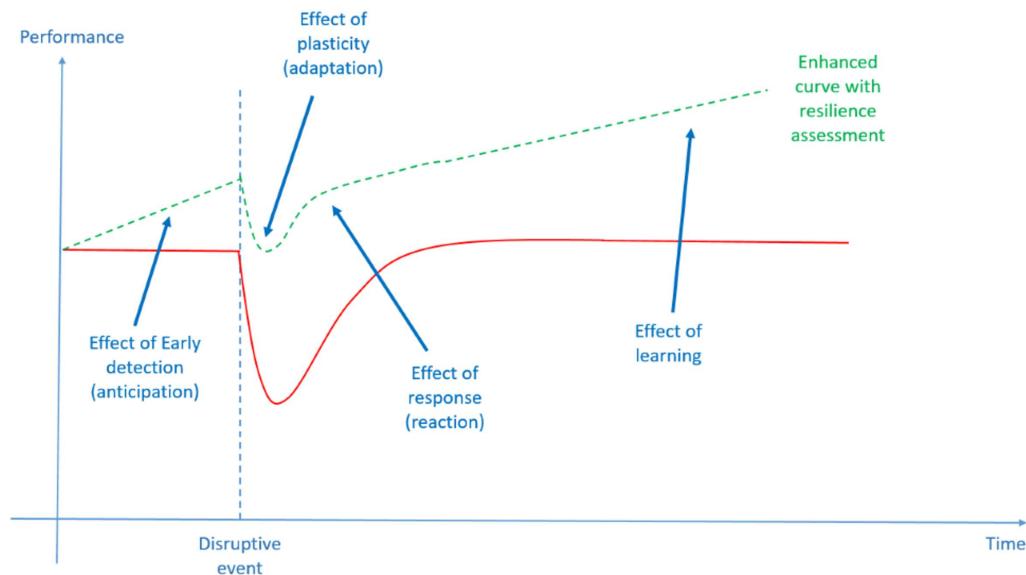


Figure 34: Standard and enhanced system performance

In the following table, a comparison between the SE lifecycle stages, as declined for the resilience requirements, and the main points of the resilience assessment framework.

Table 10: SE life cycle and Resilience Assessment stages

Life cycle stage	Information	Resilience Assessment	RA Output
Mission and Stakeholder Needs Analysis	Insert adversities in the context diagram as actors. Insert resilience scenarios as use cases.	Operative control: ability to monitor	The adversities are inserted by identifying the perturbations. The resilience scenarios are inserted by evaluating the weak signals in the safety assessment context
Stakeholder Requirements	Develop use case interaction diagrams to document the interaction of actors and architectural modules during the scenarios. Develop sequence diagrams to represent the activity flow during scenarios.	Decision: ability to respond	The use case interaction diagrams are rooted in the anticipation of expected values. The activity flow describe how the anticipation capability influences the safety assessment
System Requirements	Develop activity diagrams to show the states of the system (and adversities) during scenarios.	Early detection: ability to anticipate	The system states are predicted by the ML system.
Architecture and System Design	Develop state models of the scenarios. Model events and signals among the architectural nodes.	Forecast: ability to learn	States, event and signals represent the training datasets for design and develop the dynamic model

3.2.1. Defining a Resilience metric

Brtis [78] surveyed and evaluated a number of potential resilience metrics and identified the following:

- Maximum outage period
- Maximum brownout period
- Maximum outage depth
- Expected value of capability: the probability-weighted average of capability delivered

- Threat resiliency; i.e. the time integrated ratio of the capability provided divided by the minimum needed capability
- Expected availability of required capability; i.e. the likelihood that for a given adverse environment the required capability level will be available
- Resilience levels; i.e. the ability to provide required capability in a hierarchy of increasingly difficult adversity
- Cost to the opponent
- Cost-benefit to the opponent
- Resource resiliency; i.e. the degradation of capability that occurs as successive contributing assets are lost

Brtis found that multiple metrics may be required, depending on the situation. However, if one had to select a single most effective metric for reflecting the meaning of resilience, Brtis proposed that it would be "the expected availability of the required capability".

In the context of major accident plants, the required capability is undoubtedly the *System Safety*. Safety is the required emerging property the system must have, in order to ensure the desired performances (accident prevention, environmental protection, occupational health & safety, production quality, regulatory compliance, ...) in a sustainable way. Thus, an appropriate resilience metric, accordingly to Brits, is the *Expected availability of required capability*.

3.2.2. Dynamic Risk Index

The abovementioned metric can be represented through a Dynamic Risk Index (DRI). The DRI is a dynamic indicator of the accidental risk of a plant (DRI) which is updated in real time during operations, through the dynamic model described in the previous sections, and, in this way, allows to identify whether or not the critical unit of the plant is approaching failure, i.e. the probability of the related Top Event.

The DRI, since it captures the precursors of any risks, will allow operators to implement the appropriate counter measures when the index approaches a risk condition.

3.3. CONCLUSION OF SECTION 3

A system's resilience typically requires that a system be open to new information. As a process, resilience is a measure of how well an engineered system integrates shocks and initiates new operational regimes. Whereas stressors may originate from within the system, or from a co-occurring system, the resources to accommodate a stressor are typically the result of complex, reciprocal relationships between systems that involve many different factors at once [81].

Engineering resilience is about *nonlinearity* and *dynamism*; for example, understanding how an infinitesimally small change in initial conditions like an assumption in a software line of code can lead to huge consequences for a system as a whole [82].

Table 10 shows how the expected output of the Resilience Assessment framework (§2.2) can fit the Resilience Life Cycle requirements.

In the following section, an application of the whole Dynamic Asset-integrity and Risk Management System (DARMS) to a LNG bunkering facility is shown.

SECTION 4

4. AN APPLICATION OF THE DYNAMIC ASSET-INTEGRITY AND RISK MANAGEMENT SYSTEM (DARMS) TO A LNG BUNKERING FACILITY

The present section concerns the design, development and application of the DARMS to a LNG bunkering facility. The system is intended to be a predictive decision-making tool, which provides a constantly updated risk, based on the ongoing operations, for the whole plant and all the components. In accordance with the resilience engineering practice, the proposed system enhances the operational control, focusing on the overall performance of the complex system constituted by plant and environment.

As previously described, the system relies on a robust Machine Learning architecture, that provides the system the ability to automatically learn and improve from experience without being explicitly programmed.

The overall model includes Deep Neural Networks (DNN), for predicting the critical process variables values, Hidden Markov Models (HMM), for inferring the most probable sequence of states for the system, and Hierarchical Bayesian Networks (HBN), for combining Boolean events, such malfunctions and failures of critical equipment, and system states predictions. The outcome is a dynamic risk index, which is also the chosen resilience metric.

4.1. DARMS ARCHITECTURE

The conceptual architecture of the DARMS is depicted in the following figure.

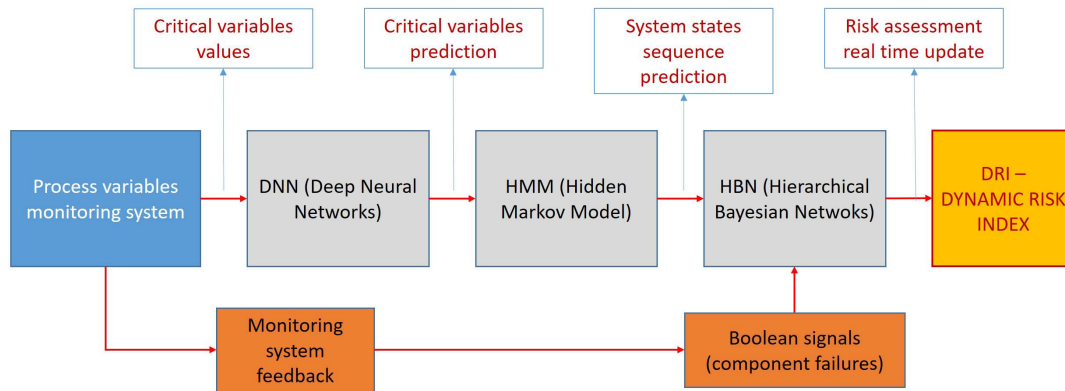


Figure 35: conceptual architecture of the DARMS

The logical steps are:

- prediction of critical variables values;
- prediction of system states sequence;
- real time update of the risk parameters, which define a dynamic risk index.

Prediction of critical variables values

The model component in duty of predicting the critical variables values is a series of Deep Neural Networks (DNN) (§1.4.1.), with the following characteristics:

- hidden layers: 10;
- neurons in each hidden layer: 24,
- learning rate: 1E-7,
- step max: 1E8,
- activation function: Tanh,
- error function: SSE,
- prediction time interval: 5 min.

Prediction of system states sequence

As already mentioned, the Bayesian approach has been proven to be a robust probability reasoning method under uncertainty, providing a tool for incorporating evidence during operations. Hidden Markov Models (HMMs) seems to be one of the most promising and reliable approaches (§2.2.1.).

In the DARMS design, with reference to the model description in §2.2.1., the hidden states are the states between a regular performance and a failure of a sub-system. The only known states are the first (the component is performing well) and the last (the component fails), and the hidden states in between may represent the precursors of accidental events. The emissions of the system are the process variable values.

Dynamic Risk Index

The Boolean elements of the risk analysis and the predictions of system states sequence are integrated in a Hierarchical Bayesian Network (HBN). The HBN is obtained by remapping the Fault Trees (FT) [83].

The quantitative analysis of a BN may proceed along two lines. A forward (or predictive) analysis, in which the probability of occurrence of any node of the network is calculated on the basis of the prior probabilities of the root nodes and the conditional dependence of each node. A more standard backward (diagnostic) analysis that concerns the computation of the posterior probability of any given set of variables given some observation (the evidence), represented as instantiation of some of the variables to one of their admissible values.

In the DARMS, the prior probabilities of the nodes are the FT failure rates. The parameters are updated, from one side, by the process variables values predictions, and, from another side, by the prediction of the state sequences, which defines the parameters of the probability distributions for each network node.

Hierarchical modelling is used when information is available on several different levels of observational units. In the present paper the different sources of information are:

- prior probabilities from Fault and Event Trees;
- boolean failures;
- predictions on critical variables values.

The sub-models combine to form the hierarchical model, and Bayes' theorem is used to integrate them with the observed data and account for all the uncertainty that is present. The result of this integration is the posterior distribution, also known as the updated probability estimate, as additional evidence on the prior distribution is acquired (§1.3.).

4.2. CASE STUDY

As case study, a LNG Shore-to-Ship bunkering operation was analysed.

The transfer unit is equipped with:

- quick release hooks;
- fenders;
- dock monitoring system to check the ship's position and speed of approach, weather and sea conditions;
- pier control room.

The quick release hooks will be installed on the dock. All hooks are capable of moving both vertically and horizontally and each is designed to be released independently of the other.

The pier control room is equipped with controls for the emergency stop of the LNG transfer, for the release of the LNG transfer connection and equipment for the remote control of the fire extinguishing system.

The Ship-to-Shore connection is used to reciprocally exchange Emergency Shut-Down (ESD) between the ship and the ground system.

The connection between the ship and the plant takes place via a loading arm, with two independent lines: one for the liquid phase (LNG) discharged from the ship to the plant and a flexible line for the gas phase (steam return from the plant to the ship); vice versa, the steam to the plant and the LNG to the ship, during the bunkering of a barge.

The loading arm is equipped with safety devices, such as:

- a quick release system (PERC);
- a PLC, dedicated to the loading arm and connected to the plant control system, integrated into the Hydraulic Processing Unit;
- the arm will be connected to the ship by means of 2 flanged connections, one for the liquid and one for the vapor.

The operations of connection / disconnection of the loading arm is monitored through the control system (pressure gauges and thermometers).

The lines are:

- LIQUID line (10");
- VAPOR line (8").

4.2.1. QRA of the bunkering operation

The main causes of loss of containment during bunkering reside in the coupling operation of the bunkering manifold to the receiving vessel and are due to damage to the connection pipe during normal operations and SIMOPS (simultaneous operations).

During bunkering operations, loss of containment can occur in different parts of the process. In particular, the situations that can lead to a loss of containment concern failures of critical equipment and failures of the receiving vessel.

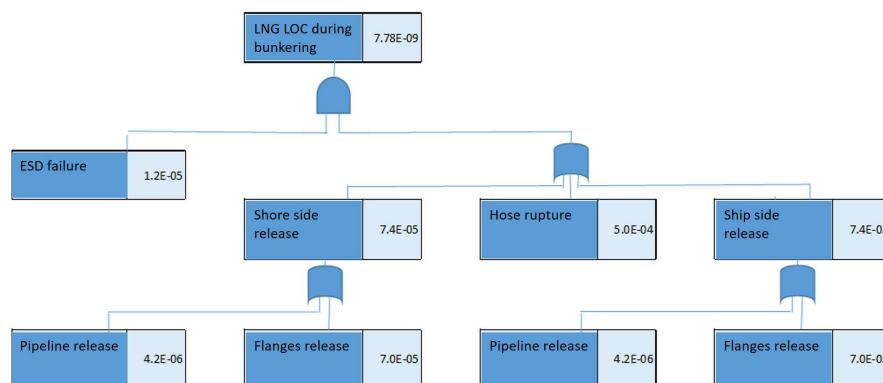


Figure 36: FTA for the bunkering operation

Top Event	Immediate ignition	Delayed ignition	Consequences	Frequency
Yes	0.03		Jet fire	2.33E-10
TES		0.003	Flash fire	2.26E-11
No	0.97		Dispersion	7.52E-09
		0.997		

Figure 37: ETA for the bunkering operation

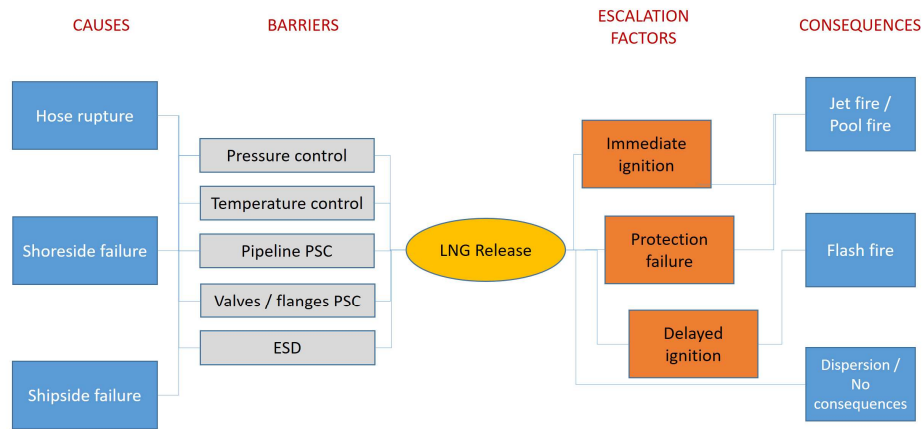


Figure 38: Bow-Tie for the bunkering operation

The bow-tie is transposed into a HBN through a specific .xml files (Zurheide et al., 2021) which contains the dependencies, as shown in the following figure.

4.2.2. Model development

The Bow-Tie structure represents the high-level model structure, i.e. the HBN. The transposition is performed according to the following flowchart.

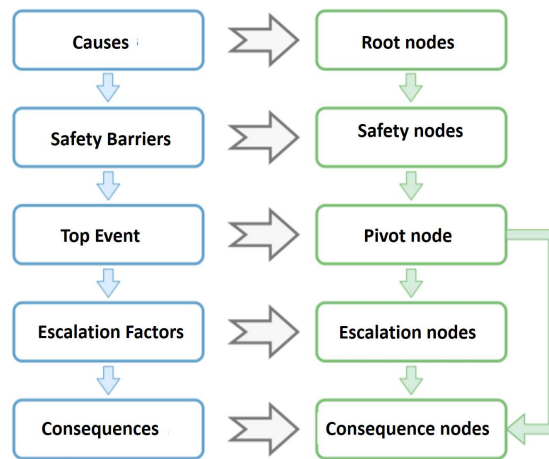


Figure 39: Bow-Tie to HBN transposition steps

The resulting structure is represented in the following figure.

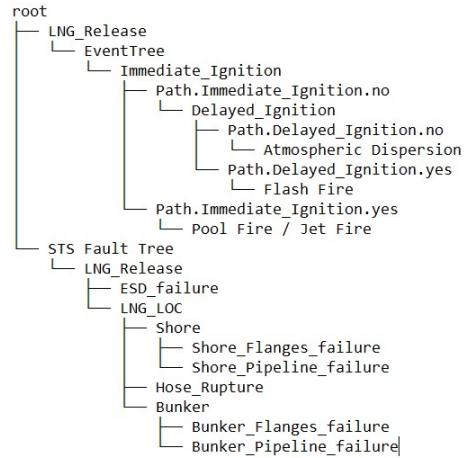


Figure 40: HBN structure with logical dependencies

According to the logic described above, root failures (leaks from flanges, valves, breakage or detachment of the hose, ...) are related to the trend of critical process variables (temperature, pressure, as monitored by the control system) through a hidden state Markov model (HMM).

The sequences predicted by the HMM, are integrated, as evidences, in the HBN, which consequently updates the risk parameters, as shown in the following tables.

Table 11: Root nodes Probability distributions



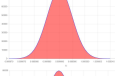
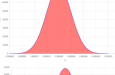

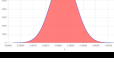


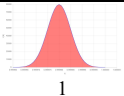
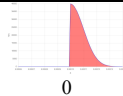
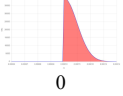
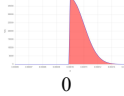
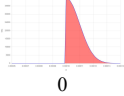

Stage	Expected probability	State	PDF
ESD (works)	0.99998	Safe	
Hose (works)	1	Safe	
S_Pipeline (works)	0.99999	Safe	
S_Flanges (works)	0.99997	Safe	
B_Pipeline (works)	0.99989	Safe	
B_Flanges (works)	0.99982	Safe	

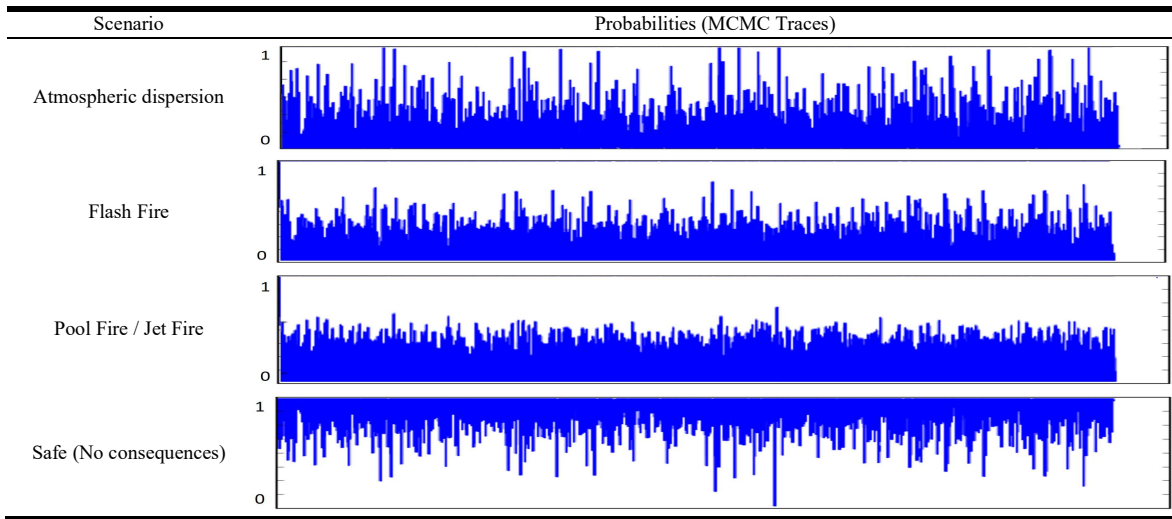
Table 12: Intermediate nodes dependencies

ESD LOC (or gate)	ESD (works – PD Tab.1) LOC (works - inference)	ESD (works – PD Tab. 1) LOC (fails - inference)	ESD (fails – marginal Tab.1) LOC (works - inference)	ESD (fails – marginal Tab.1) LOC (fails - inference)
LNG (works) Expected prob.	 1	 1	 1	 0
LNG (fails) Expected prob.	 0	 0	 0	 1

And so on for all the trees gates.

In accordance with the real time update of the risk parameters, the probability distributions of the accident scenarios are updated as well. The MCMC traces of the dynamic probability distributions of the consequences are shown in the following table.

Table 13: Scenarios probabilities (MCMC traces)



The transitions prediction between the hidden states of the system are thus expressed by the dynamic probability of system failure, as shown in the following figure.

The confidence interval reflects the accuracy of the prediction based on the collected evidences.

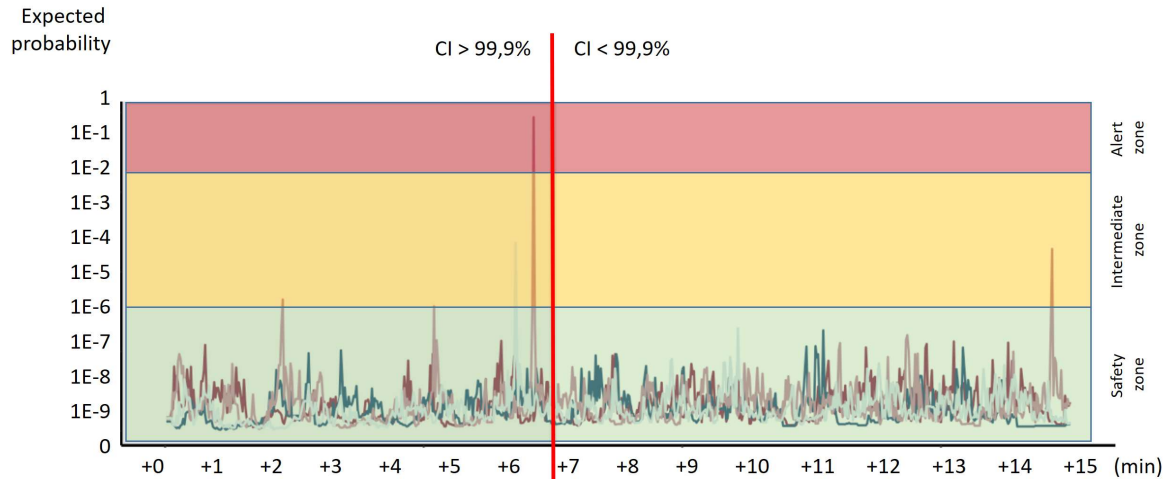


Figure 41: Expected probability prediction

The three ranges (safety, intermediate and alert zones) are reported as a global dynamic indicator, where the values are normalized in the interval 0-10:

Dynamic Safety Indicator

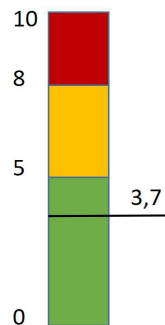


Figure 42: Representation of the actual value of the dynamic safety indicator

According to §3.2.1., the required capability is the system safety, considered as an emergent property of the system. The actual value of the dynamic indicator represents a measure of the *expected availability of required capability*, therefore **a measure of how the system is resilient in that very precise moment**.

By analyzing the dynamic safety indicator, which is related to the ongoing activity, the operator can modify the operating parameters, intercepting the cause of the deviation, and check, at the next update, how the indicator varies.

4.3. CONCLUSION OF SECTION 4

The DARMS is a hybrid model incorporating different data driven models in a complete logical and interconnected model (DNN – HMM – HBN). The former exhibits a robust predictive capability on the process variables, the second, which is the real predictive model, performs abductive inference on the states sequence of the system, and the latter explores the interdependencies among the system components and

their modification alongside process variables fluctuation. The combined model outcomes are then used to generate a dynamic risk indicator connected with the process variables prediction. The outcomes analysis demonstrated that the framework could be used for the reliable predictions in terms of accuracy and test error performance.

CONCLUSION

CONCLUSION

The need for a more dynamic and affordable risk assessment process is underlined by several authors [3], [23], [38], and the majority of the proposed ideas are about a deep change in the inferential process. Such a need has also emerged vigorously both from the analysis of recent industrial accidents and from the evolution of industrial safety regulations. As represented in the introduction, the focal point of industrial safety has passed from the *technical aspects*, where the issue was the *failure of components*, to the *human factor*, where the issue was the *human error*, then to *socio-technical interactions*, where the issue was the *improper interaction*, finally to *resilience*, where the issue is *vulnerability against unanticipated situations*.

These changes in the safety focal point have reflected both on the evolution of legislation and on the risk analysis techniques.

The first Seveso Directive (82/501/EEC) introduces the mandatory requirement for risk assessment and employees training; in fact, it is temporally located, between the *era of technology* and the *era of human factor*. In those years, the classical risk analysis techniques have been developed, such as Domino model, HazOp and Failure Mode and Effects Analysis (FMEA).

The second Seveso Directive (96/82/EC) introduces the Safety Management System, and the External Contingency planning; in fact, it is temporally located in the *era of socio-technical interaction*. In those years more detailed risk assessment techniques have been developed, such as the Swiss-cheese model, and the Cognitive Reliability and Error Analysis Method (CREAM).

The third Seveso Directive (2012/18/EU) introduces the integrated risk assessment for areas with a high concentration of major accident plants, and the plant ageing management. We are at the dawn of the *era of resilience*! the first attempts to define adequate tools for risk analysis led to the Functional Resonance Analysis Method (FRAM) and to the Systems-Theoretic Accident Model and Processes (STAMP), which was discussed in the introduction.

It is crystal-clear that there is a pressing need for a risk assessment tool that can determine how much a complex system, or system of systems, is capable of a resilient performance.

The research and development phases for defining a tool capable of responding to the requirements of systemicity and dynamism, and that is able to evaluate the response of the system in relation to the four resilience guide-words - monitor, anticipate, learn, adapt -, allow to answer the open research questions posed at the beginning of the present work.

Research Question I: "How can real time measurements of process variables and Boolean failures be integrated into a dynamic risk model, for predicting the safety state transitions of a complex system"?

For designing the DARMS, a detailed information on the current risk during operations, or better, on the ongoing safety of the whole system moment by moment, is required. In fact, one of the critical issues

to be addressed in the risk assessment is the coexistence of Boolean elements (e.g. failure of instruments) and analogical elements (deviation of process variables). The results of the risk analysis are obtained from logical concatenations of failures (the fault trees) and events (the event trees), and are characterized by a Boolean (true / false) approach. What is not yet evaluated is the transition between a safe state and a risk state.

This issue was addressed with the *Research Question I*.

In the classic techniques of risk analysis, static variables are used, which, as defined, do not lend themselves to representing a dynamic system, in which the risk, or rather, the safety, varies at any time, in relation to ongoing activities. The traditional QRA is based on a frequentist approach, that means it calibrates the plausibility of propositions by considering repeated sampling of a population distribution. The frequentist properties of a statistical proposition can be quantified as fixed values, that's why the traditional QRA is static.

For designing a dynamic risk assessment model, the need is to explore a recent branch of inferential statistics, the abductive inference. Abduction is defined as "a syllogism in which the major premise is evident but the minor premise, and therefore the conclusion, only probable." Basically, it involves forming a conclusion from the information that is known. Abductive reasoning typically begins with an incomplete set of observations and proceeds to the likeliest possible explanation for the set. The dynamism is evident: the conclusion is adjusting itself with the observations, and the only observations available, are the process variables fluctuations.

The system described in §1.4.1. allows to integrate the process variables predictions and the component failures in a robust, integrated, hybrid model. The prediction of critical variables and the consequent automatic updating of the risk assessment parameters, enable a dynamic representation of the safety level of the system.

Research Question II: "Can Data-Driven models be considered an appropriate approach to Resilience Assessment, complementing the Knowledge-Driven models describing systems physical behavior"?

Over the past decades, few concepts have gained such attention as resilience, but however, a set of tools to verify and measure the resilience of complex systems has not yet been defined. The most relevant attempts are those of Hollnagel, with the FRAM [4] and of Leveson, with the STAMP [3]. The first, however, cannot be applied to industrial safety analysis, involving elements that are too distant from each other, and often not inherent to the industrial context, and the second one does not fit into a dynamic representation of system safety.

A data-driven model is based on the analysis of the data about a specific system. The main concept of data-driven model is to find relationships between the system state variables (input and output) without explicit knowledge of the physical behavior of the system.

But does it fit into the resilience paradigm? (*Research Question II*)

To analyze and evaluate the resilience of a system, it is necessary to start from the four resilience guide-words – monitor, anticipate, learn, adapt. Each of these concepts has to do with the data collected in the system during the various operative phases. There is an immense source of information in these data, which, in most cases, is unknown, and making that information available is the only way to grasp the suggestions of the 4 pillars of resilience.

Monitoring means having knowledge of the data at all times.

Anticipating is possible only by observing the behavior of the system over time, identifying the weak signals.

Learning is feasible by knowing the results of what is being done, to recognize the precursors of undesired events.

Adapting can only be done if you have reliable forecasts and knowledge on the consequences.

The application of the resilience assessment framework presented in §2.2., which relies on a robust Machine Learning architecture, allows to catch all the four needs for a resilient performance.

Research Question III: “How to integrate the Resilience Assessment into the Systems Engineering framework”?

Once it is understood how to design a dynamic model of safety analysis, and how, through appropriate machine learning techniques, to assess and measure the resilience of a complex system, it is possible to state that ***safety is an emergent property of the system***, and ***resilience is the enabling property***.

Starting from these considerations, how can systemic requirements be defined to achieve, maintain and measure the resilience of systems? How can we measure “how much” safety emerges from the interactions between all system components?

These concepts are embodied in the *Research Question III*.

Systems, by definition, deliver desired capability. It is the quality of the delivery of such capability – in the face of adversity – that resilience addresses. While resilience should be considered throughout the Systems Engineering life cycle, it is critical that the development of resilience requirements is led by resilience considered in the early life cycle activities. Once resilience requirements are established, they can and should be managed along with all of the other requirements in the trade space throughout the system life cycle.

By defining the life cycle stages as described in §3.2., the resilience assessment phases can be included in the system requirements definition.

PUBLICATIONS LIST

- Gualeni, P., Perrera, F., Raimondo, M., **Vairo, T.** *Accessibility for maintenance in engine room: development and application of a prediction tool for operational costs estimation.* Ship Technology Research, Taylor and Francis, 2022.
- Markowski, A.S., Krasławski, A., **Vairo, T.**, Fabiano, B. *Process safety management quality in industrial corporation for sustainable development.* Sustainability (Switzerland), 2021, 13(16), 9001.
- Vairo, T.**, Gualeni, P., Reverberi, A.P., Fabiano, B. *Resilience dynamic assessment based on precursor events: Application to ship lng bunkering operations.* Sustainability (Switzerland), 2021, 13(12), 6836.
- Bragatto, P., **Vairo, T.**, Milazzo, M.F., Fabiano, B. *The impact of the COVID-19 pandemic on the safety management in Italian Seveso industries.* Journal of Loss Prevention in the Process Industries, 2021, 70, 104393.
- Vairo, T.**, Pontiggia, M., Fabiano, B. *Critical aspects of natural gas pipelines risk assessments. A case-study application on buried layout.* Process Safety and Environmental Protection, 2021, 149, pp. 258–268.
- Vairo, T.**, Reverberi, A.P., Bragatto, P.A., Milazzo, M.F., Fabiano, B. *Predictive model and soft sensors application to dynamic process operative control.* Chemical Engineering Transactions, 2021, 86, pp. 535–540.
- Magri, S., **Vairo, T.**, Reverberi, A.P., Fabiano, B. *Oil Spill Identification and Monitoring from Sentinel-1 SAR satellite earth observations: A machine learning approach.* Chemical Engineering Transactions, 2021, 86, pp. 379–384.
- Vairo, T.**, Gualeni, P., Fabiano, B., Benvenuto, A.C. *Resilience assessment of bunkering operations for A LNG fuelled ship.* Proceedings of the 30th European Safety and Reliability Conference and the 15th Probabilistic Safety Assessment and Management Conference, 2020, pp. 3693–3700.
- Vairo, T.**, Rapuzzi, A., Lecca, M., Fabiano, B. *A data driven model for ozone concentration prediction in a coastal urban area.* Chemical Engineering Transactions, 2020, 82, pp. 379–384.
- Vairo, T.**, Reverberi, A.P., Fabiano, B. *From risk assessment to resilience assessment. an application to a hazmat storage plant.* Chemical Engineering Transactions, 2020, 82, pp. 151–156.
- Vairo, T.**, Milazzo, M.F., Bragatto, P., Fabiano, B. *A Dynamic Approach to Fault Tree Analysis based on Bayesian Beliefs Networks.* Chemical Engineering Transactions, 2019, 77, pp. 829–834.
- Pontiggia, M., **Vairo, T.**, Fabiano, B. *Risk assessment of buried natural gas pipelines. Critical aspects of event tree analysis.* Chemical Engineering Transactions, 2019, 77, pp. 613–618.
- Magri, S., Quagliati, M., De Gaetano, P., **Vairo, T.**, Fabiano, B. *Fuel spill after ship collision: Accident scenario modelling for emergency response.* Chemical Engineering Transactions, 2019, 74, pp. 1363–1368.
- Vairo, T.**, Lecca, M., Trovatore, E., Reverberi, A.P., Fabiano, B. *A Bayesian belief network for local air quality forecasting.* Chemical Engineering Transactions, 2019, 74, pp. 271–276.
- Cermelli, D., Currò, F., **Vairo, T.**, Fabiano, B. *Hydrogen jet-fire: Accident investigation and implementation of safety measures for the design of a downstream oil plant.* Chemical Engineering

Transactions, 2018, 67, pp. 415–420.

Vairo, T., Reverberi, A.P., Milazzo, M.F., Fabiano, B. *Ageing and creeping management in major accident plants according to Seveso III directive*. Chemical Engineering Transactions, 2018, 67, pp. 403–408.

Vairo, T., Magrì, S., De Gaetano, P., Quagliati, M., Fabiano, B. *Multicomponent dispersion of hydrocarbons at sea: Source term evaluation and hydrodynamic simulation of the spill*. Chemical Engineering Transactions, 2018, 67, pp. 61–66.

Vairo, T., Quagliati, M., Del Giudice, T., Barbucci, A., Fabiano, B. *From land- to water-use-planning: A consequence based case-study related to cruise ship risk*. Safety Science, 2017, 97, pp. 120–133.

Vairo, T., Magrì, S., Quagliati, M., Reverberi, A.P., Fabiano, B. *An oil pipeline catastrophic failure: Accident scenario modelling and emergency response development*. Chemical Engineering Transactions, 2017, 57, pp. 373–378.

Vairo, T., Pastorino, R., Rehman, A., Fabiano, B. *An approach to risk evaluation in connection with fire scenarios from a cruise ship*. Chemical Engineering Transactions, 2015, 43, pp. 1939–1944.

Pastorino, R., **Vairo, T.**, Benvenuto, A., Fabiano, B. *Area risk analysis in an urban port: Personnel and major accident risk issues*. Chemical Engineering Transactions, 2014, 36, pp. 343–348.

Vairo, T., Currò, F., Scarselli, S., Fabiano, B. *Atmospheric emissions from a fossil fuel power station: Dispersion modelling and experimental comparison*. Chemical Engineering Transactions, 2014, 36, pp. 295–300.

Vairo, T., Pastorino, R., Reverberi, A.P., Fabiano, B. *HazMat liquid release following a tank truck accident: Cross-check modelling and field data validation*. Chemical Engineering Transactions, 2013, 32, pp. 97–102.

LIST OF FIGURES

Introduction

Figure 1: Timeline of Risk assessment methods

Figure 2: Heinrich's Domino model

Figure 3: Bird and Loftus' Domino model

Figure 4: Reason's Swiss Cheese model

Figure 5: Bias - Variance tradeoff

Figure 6: Overfitting and Underfitting

Open Research Questions

Figure 7: Analytical models

Figure 8: changing of systems safety focal point

Section 1 – Inference and Machine Learning

Figure 9: Known Unknowns matrix

Figure 10: Dynamic Risk Assessment framework

Figure 11: FTA for leakage on the VRU

Figure 12: BN for leakage on the VRU

Figure 13: Posterior probability density function for a leakage on the VRU

Figure 14: Map of analytical modelling

Figure 15: BBN with evidence of the interdependencies between the nodes

Figure 16: scatterplot observed vs. predicted

Figure 17: scatterplot observed vs. predicted – LightGBM

Figure 18: Hybrid model concept

Figure 19: Schematic diagram of the Vapor Recovery Unit (VRU)

Figure 20: conceptual elements of a Backpropagation Neural Network

Figure 21: Scatterplots predicted values vs. ground truth for temperature sensors

Figure 22: An example of ML pipeline

Section 2 – The Resilience of Complex Systems

Figure 23: The resilience assessment framework

Figure 24: Representative layout of a shore-to-ship LNG refueling plant

Figure 25: Fault tree developed for LNG loss of containment (LOC)

Figure 26: Posterior probabilities distribution for the leakage on the shipside (left), shoreside (center) and connection (right)

Figure 27: posterior predictive distribution, and expected probability of system failure

Figure 28: Excerpt obtained by HMM inferential sampling and prediction

Figure 29: Value of the resilience score over time, corresponding to different perturbative situations

Section 3 – Engineering Systems Resilience properties

Figure 30: Influences in system resilience

Figure 31: The V-model

Figure 32: Notional resilience scenario life cycle components

Figure 33: Data structure for specifying resilience requirements

Figure 34: Standard and enhanced system performance

Section 4 – An application of the Dynamic Asset-integrity and Risk Management System (DARMS) to a LNG bunkering facility

Figure 35: conceptual architecture of the DARMS

Figure 36: FTA for the bunkering operation

Figure 37: ETA for the bunkering operation

Figure 38: Bow-Tie for the bunkering operation

Figure 39: Bow-Tie to HBN transposition steps

Figure 40: HBN structure with logical dependencies

Figure 41: Expected probability prediction

Figure 42: Representation of the actual value of the dynamic safety indicator

LIST OF TABLES

Introduction

Table 1: Systematic and Systemic thinking and action

Section 1 – Inference and Machine Learning

Table 2: Sensitivity analysis for leakage on the VRU

Table 3: Hardware sensors and relevant process parameters monitoring VRU process section

Table 4: Errors and accuracy

Section 2 – The Resilience of Complex Systems

Table 5: Equipment Count for Leak Frequency Estimate

Table 6: Risk assessment assumptions

Table 7: Head of an input dataset

Table 8: Expected probabilities of occurrences of possible states in the root components

Section 3 – Engineering Systems Resilience properties

Table 9: Modeling information during lifecycle phases

Table 10: SE life cycle and Resilience Assessment stages

Section 4 – An application of the Dynamic Asset-integrity and Risk Management System (DARMS) to a LNG bunkering facility

Table 11: Root nodes Probability distributions

Table 12: Intermediate nodes dependencies

Table 13: Scenarios probabilities (MCMC traces)

REFERENCES

- [1] Woods, D., Leveson, N., Hollnagel, E. *Resilience engineering concepts and precepts*. CRC Press, 2017.
- [2] Hollnagel, E. *Safety I and Safety II – The Past and Future of Safety Management*. *Cognition Technology & Work* 17(3):461-464, 2017.
- [3] Leveson, N. *Safety III: A Systems Approach to Safety and Resilience*. MIT Engineering Systems Lab, 2020.
- [4] Hollnagel, E. *FRAM: The Functional Resonance Analysis Method: Modelling Complex Socio-Technical Systems*. Ashgate, 2012.
- [5] Heinrich, H.W., Peterson, D., Roos, N. *Industrial Accident Prevention, 5th Edition*, McGraw-Hill, 1980.
- [6] Bird, F.E., Lotus, R.G. *Loss control management*. Loganville, Ga. Institute Press, 1976.
- [7] Lawley, H.G., *Operability Studies and Hazard Analyses*. *Chemical Engineering Progress*, 70(4):45. 1974.
- [8] Reason, J. *The Contribution of Latent Human Failures to the Breakdown of Complex Systems*. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*. 327 (1241): 475–84. 1990.
- [9] Mahmood, Y. A., Ahmadi, A., Verma, A. K., Srividya, A., Kumar, U. Fuzzy fault tree analysis: a review of concept and application. *International Journal of System Assurance Engineering and Management* volume 4, 19–32. 2013.
- [10] Perrow, C. *Normal Accidents: Living with High-Risk Technologies*. New York: Basic Books, 1984.
- [11] Leveson, N., Dulac, N., Zipkin, D., Cutcher-Gershenfeld, J., Carroll, J., Barrett, B. *Engineering resilience into safety-critical systems*. MIT Engineering Systems Lab, 2012.
- [12] Merriam Webster Dictionary (<https://www.merriam-webster.com/dictionary/artificial%20intelligence>) accessed 12.12.2020.
- [13] Domingos, P., *A Unified Bias-Variance Decomposition*. Department of Computer Science and Engineering University of Washington. AAAI-00 Proceedings. 2020.
- [14] Cloutier, R. *Model-Based Systems Engineering*. INCOSE, 2019.

- [15] Checkland, P. *Systems thinking, systems practice*. Chichester: John Wiley, 1999.
- [16] Capra, F. *The Web of Life: A New Scientific Understanding of Living Systems*. Anchor Books, 1996.
- [17] Denyer, D. *Organizational Resilience: A summary of academic evidence, business insights and new thinking*. BSI and Cranfield School of Management. 2017.
- [18] Pasma, H.J., Genserik, R. *Past, present and future of Quantitative Risk Assessment (QRA) and the incentive it obtained from Land-Use Planning (LUP)*. Journal of Loss Prevention in the Process Industries 28:2–9. 2014.
- [19] Vairo, T., Milazzo, M.F., Bragatto, P., Fabiano, B. A Dynamic Approach to Fault Tree Analysis based on Bayesian Beliefs Networks. Chemical Engineering Transactions 75, AIDIC, 2019.
- [20] Yang, M., Kahn, F., Lye, L. *Precursor-based hierarchical Bayesian approach for rare event estimation: a case of oil spill accident*. Process Safety and Environmental Protection, 91, 333-342, Elsevier, 2013.
- [21] Kalantarnia, M., Khan, F., Hawboldt, K. *Dynamic risk assessment using failure assessment and Bayesian theory*, Journal of Loss Prevention in the Process Industries, 22, 600-606, Elsevier, 2009.
- [22] Hollnagel, E. *Safety-II in practice*. Abingdon, Oxon, UK, Routledge, 2017.
- [23] Jain, P., Rogers, W.J., Pasma, H.J., Mannan, M.S. *A resilience-based integrated process systems analysis. Part II management system layer*, Process Safety and Environmental Protection, 118, 115-124, Elsevier, 2018.
- [24] Swaminathan, S., Smidts, C. *The Event Sequence Diagram framework for dynamic Probabilistic Risk Assessment*. Reliability Engineering & System Safety, Elsevier, 1999.
- [25] People, H.E. *On the mechanization of abductive logic*. Proceedings of the 3rd International Joint Conference on Artificial Intelligence, 147–152, 1973.
- [26] Thagard, P., Cameron, S. *Abductive reasoning: Logic, visual thinking, and coherence*. Waterloo, Ontario: Philosophy Department, University of Waterloo, 2005.
- [27] Vairo, T., Lecca, M., Trovatore, E., Reverberi, A.P., Fabiano, B. A Bayesian Belief Network for Local Air Quality Forecasting, Chemical Engineering Transactions, 74, 271-276, AIDIC, 2019.
- [28] NATO Press Conference by US Secretary of Defense, Donald Rumsfeld, 2002 June 6th. (<https://www.nato.int/docu/speech/2002/s020606g.htm>). Accessed 24.02.2021.

-
- [29] Luft, J.; Ingham, H. *The Johari window, a graphic model of interpersonal awareness*. Proceedings of the Western Training Laboratory in Group Development. Los Angeles, 1955.
- [30] Sato, T., Kameya, Y. *A Viterbi-like algorithm and EM learning for statistical abduction*. Dept. of Computer Science, Graduate School of Information Science and Engineering Tokyo Institute of Technology, 2001.
- [31] Selman, B., Levesque, H. *Abductive and default reasoning: A computational core*. Proceedings of the Eighth National Conference on Artificial Intelligence, Boston, 1989.
- [32] Geotor, L., Taskar, B., *Introduction to Statistical Relational Learning*. MIT Press, 2007.
- [33] Richardson, M., Domingos, P. *Markov Logic Networks*. Department of Computer Science and Engineering University of Washington, 2006.
- [34] Kakas, A.C., Kowalski, R.A., Toni, F. *The role of abduction in logic programming*, Handbook of Logic in Artificial Intelligence and Logic Programming 5. Oxford University Press, 1993.
- [35] Ng, H.T., Mooney, R.J. *Abductive plan recognition and diagnosis: A comprehensive empirical evaluation*. Proceedings of the Third International Conference on Principles of Knowledge Representation and Reasoning, 499-508, Cambridge, 1992.
- [36] Ishihata, M., Sato, T. *Bayesian inference for statistical abduction using Markov chain Monte Carlo*. JMLR: Workshop and Conference Proceedings 20, 81–96, 2011.
- [37] Vairo, T., Pontiggia, M., Fabiano, B. *Critical aspects of natural gas pipelines risk assessments. A case-study application on buried layout*. Process Safety and Environmental Protection, 149, 258-268, Elsevier, 2021.
- [38] Vairo, T., Reverberi, A.P., Fabiano, B. *From Risk Assessment to Resilience Assessment. an Application to a Hazmat Storage Plant*, Chemical Engineering Transactions, 82, 151-156, AIDIC, 2020.
- [39] Amin, T., Khan, F., Ahmed, S., Imtiaz, S. A novel data-driven methodology for fault detection and dynamic risk assessment. Canadian Journal of Chemical Engineering, 1-20, Wiley, 2020.
- [40] Allenby, G.M., Rossi, P.E., McCulloch, R.E. *Hierarchical Bayes Models: A Practitioners Guide*. SSRN, 2005.
- [41] Fang, W., Wu, J., Bai, Y., Zhang, L., Reniers, G. *Quantitative risk assessment of a natural gas pipeline in an underground utility tunnel*. Process. Safety Progress, 38, 12051. Wiley, 2019.
- [42] Vairo, T., Gualeni, P., Reverberi, A.P., Fabiano, B. *Resilience Dynamic Assessment Based on*

- Precursor Events: Application to Ship LNG Bunkering Operations*. Sustainability 13(12):6836, MDPI, 2021.
- [43] Vairo, T., Quagliati, M., Del Giudice, T., Barbucci, A., Fabiano, B. *From land- to water-use-planning: A consequence-based case-study related to cruise ship risk*, Safety Science, 97, 120-133, Elsevier, 2017.
- [44] Baklanov, A., Korsholm, U., Mahura, A., Petersen, C., Gross, A. *ENVIRO-HIRLAM: on-line coupled modelling of urban meteorology and air pollution*, Advances in Science and Research, 2, 41-46, Wiley, 2008.
- [45] Aguilera, P.A., Fernández, A., Fernández, R., Rumí, R., Salmerón, A. *Bayesian networks in environmental modelling*. Environmental Modelling & Software 26, 1375-1774. Elsevier, 2011.
- [46] Murphy, B.D., Ohr, S.Y. *The limitations of atmospheric dispersion data and their contribution to uncertainties in dose assessment*. Health Physics, 48(3):315-24, Wolters Kluwer, 1985.
- [47] Vairo, T., Rapuzzi, A., Lecca, M., Fabiano, B. *A Data Driven Model for Ozone Concentration Prediction in a Coastal Urban Area*. Chemical Engineering Transactions, 82, 379-384, AIDIC, 2020.
- [48] Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., Liu, T. *LightGBM: A highly efficient gradient boosting decision tree*. Advances in Neural Information Processing Systems 30, 3146-3154, MIT Press, 2017.
- [49] Taleb, N.N. *The Black Swan: The Impact of the Highly Improbable*. Random House Publishing, 2007.
- [50] Kadlec, P., Gabrys, B., Strandt, S. *Data-driven Soft Sensors in the process industry*. Computers & Chemical Engineering, 33 (4), 795-814, 2009.
- [51] Vairo, T., Reverberi, A.P., Bragatto, A.P., Milazzo, M.F., Fabiano, B. *Predictive model and Soft Sensors Application to Dynamic Process Operative Control*. Chemical Engineering Transactions, 86, AIDIC, 2021.
- [52] Reverberi, A.P., Fabiano, B., Dovì, V.G. *Use of inverse modelling techniques for the estimation of heat transfer coefficients to fluids in cylindrical conduits*. Int. Commun. Heat Mass Transf., 42, 25-31, 2013.
- [53] Paltrinieri, N., Comfort, L., Reniers, G. *Learning about risk: Machine learning for risk assessment*, Safety Science, 118, 475-486, 2019.

-
- [54] Berkes, F., Colding, J., Folke, C. *Navigating Social- Ecological Systems: Building Resilience for Complexity and Change*. Cambridge University Press, 2003.
- [55] Pillay, M. *Resilience Engineering: An Integrative Review of Fundamental Concepts and Directions for Future Research in Safety Management*. Open Journal of Safety Science and Technology, 7, 129-160, 2017.
- [56] Woods, D.D. *Creating Foresight: Lessons for Enhancing Resilience from Columbia*. In: Farjoun, M. and Starbuck, W.H. *Organization at the Limit: Lessons from the Columbia Disaster*, Blackwell, 289-308. 2003.
- [57] ACSNI Study Group on Human Factors. *Third report: Organizing for safety*. Advisory Committee on the Safety of Nuclear Installations. HSE Books, 1993 (reprinted 1998). ISBN 0717608654.
- [58] Provan, D.J., Woods, D.D., Dekker, S.W.A., Rae, A.J. *Safety II professionals: How resilience engineering can transform safety practice*. Reliability Engineering & System Safety, 195, 106740, 2020.
- [59] Mendonca, D. *Measures of Resilient Performance*. In: Hollnagel, E., Nemeth, C.P., Dekker, S. *Resilience Engineering Perspectives: Remaining Sensitive to the Possibility of Failure*, Vol. 1, Ashgate Publishing Ltd, Aldershot, 29-48. 2008.
- [60] Sheridan, T.B. *Risk, Human Error, and System Resilience: Fundamental Ideas*. Human Factors: The Journal of the Human Factors and Ergonomics Society, 50, 418-426. 2008.
- [61] Jain, P., Paskan, H.J., Waldram, S., Pistikopoulos, E.N., Mannan, M.S. *A resilience-based integrated process system hazard analysis (RIPSHA) approach: Part I plant system layer*. Process Safety and Environmental Protection, 116, 92-105, Elsevier, 2018.
- [62] Woods, D.D. *Resilience as Graceful Extensibility to Overcome Brittleness*. In: IRGC. Resource Guide on Resilience. Lausanne: EPFL International Risk Governance Center. 2016.
- [63] Linkov, I., Trump, B.D., Fox-Lent, C. *Resilience: Approaches to Risk Analysis and Governance*. In: IRGC. Resource Guide on Resilience. Lausanne: EPFL International Risk Governance Center. 2016
- [64] Sarkara, S., Vinayb, S., Rajc, R., Maitia, J., Mitra, P. *Application of optimized machine learning techniques for prediction of occupational accidents*. Computers and Operations Research, 106, 210–224, 2019.
- [65] Galagedarage, D.M.; Khan, F. *Process Fault Prognosis Using Hidden Markov Model–Bayesian Networks Hybrid Model*. Industrial & Engineering Chemistry Research, 58, 12041–12053, 2019.

- [66] Salvatier, J., Wiecki, T.V., Fonnesbeck, C. *Probabilistic programming in Python using PyMC3*. Peer J. Computer Science, 2, 55–65. 2016.
- [67] Iannaccone, T., Landucci, G., Tugnoli, A., Salzano, E., Cozzani, V. *Sustainability of cruise ship fuel systems: Comparison among LNG and diesel technologies*. Journal of Cleaner Production, 260, 121069, 2020.
- [68] EMSA, *Guidance on LNG Bunkering*; European Maritime Safety Agency: Lisboa, Portugal, 2018.
- [69] ISO. *ISO 20519—Specification for Bunkering of Liquefied Natural Gas Fueled Vessels*; ISO: Geneva, Switzerland, 2018.
- [70] DNV-GL *Liquefied Natural Gas (LNG) Bunkering Study*. PP087423-4, Rev 3. 2014.
- [71] Taleb, N.N. *Anti-Fragile*. London: Penguin, 2012.
- [72] Aven, T., *The Concept of Antifragility and its Implications for the Practice of Risk Analysis*. Risk Analysis, 34. 2014.
- [73] Cook, R., *How complex systems fail*. Cognitive technologies Laboratory University of Chicago. 2002.
- [74] Patriarca, R., Bergström, J., Di Gravio, G., Costantino, F. *Resilience engineering: Current status of the research and future challenges*, Safety Science, 102, 79-100, 2018.
- [75] Furuta, K. *Resilience Engineering*. In: Ahn, J., Carson, C., Jensen, M., Juraku, K., Nagasaki, S., Tanaka, S. *Reflections on the Fukushima Daiichi Nuclear Accident*. Springer, 2015.
- [76] Woods, D., *Creating foresight: how resilience engineering can transform NASA's approach to risky decision making*. Testimony on the future of NASA for committee on commerce, science and transportation; 2003.
- [77] INCOSE. *Systems Engineering Handbook: A Guide for System Life Cycle Processes and Activities*, version 4.0. Hoboken, NJ, USA: John Wiley and Sons. 2015.
- [78] Jackson, S., Ferris, T.L.J. *Resilience Principles for Engineered Systems*. Wiley online library. 2012.
- [79] Britis, J.S., McEvelley, M.A. *Systems Engineering for Resilience*. Report nr. MP1909495. The MITRE Corporation, 2019.
- [80] Boardman, J., Sauser, B. *System Thinking: coping with 21st century problems*. Taylor & Francis, 2008.
- [81] Prilleltensky, I., Prilleltensky, O. *Promoting well-being: linking personal, organizational, and*

community change. Wiley, 2007.

- [82] Ungar, M. *Multisystemic Resilience: Adaptation and Transformation in Contexts of Change*. Oxford Scholarship, 2021.
- [83] Zurheide, F.T., Hermann, E., Lampesberger, H. *pyBNBowTie: Python library for Bow-Tie Analysis based on Bayesian Networks*. *Procedia Computer Science* 180, 344–351, 2021.

ACKNOWLEDGEMENTS

A doctorate over 20 years after graduation... Although I have never got too far from the university, this has been a peculiar experience.

In some ways it was simpler, because as a "grown-up", and above all working in the industrial safety sector, many of the research ideas developed in these three years have their roots in my work experience; on the other hand, finding myself taking courses with young people initially puzzled me.

Surely I am grateful for having had the opportunity to deepen and study ideas and concepts on which I had been reflecting for quite a while, and for the opportunity to study, for me the most beautiful occupation!

The list of thanks is long ... I start with the curriculum coordinator, Professor Attilio Converti, who has always followed me with interest and attention, also giving me useful suggestions.

I thank all the teachers of the courses I attended, each of whom provided me with very important ideas for my research work.

Heartfelt thanks to my supervisor, and friend, Professor Bruno Fabiano. It is because of him that I have embarked on the path of these three years of doctorate!

And then I thank the extraordinary people I have had the opportunity to meet and get to know.

Eng. Vincenzo Arrichiello, who opened the doors to the universe of Systems Engineering for me.

Professor Paola Gualeni, with whom I explored the resilience of complex systems, as well as the endless philosophical discussions on society, risk, knowledge, all ... she is a real "lighthouse in the night".

I thank the guys who I followed as supervisor in their graduation theses, for their important contribution in my research work.

And finally, I sincerely thank my life partner, Stefania, who listened to my ideas with love and patience, and every time she helped me to give them a broader meaning, ... and for correcting my English!

It was a really nice trip, thanks to everyone who accompanied me!

Tomaso Vairo