Towards the development of the Observability-in-Depth Safety Principle for the Nuclear Industry

Francesca M. Favarò[a], and Joseph H. Saleh[a]
[a]*Georgia Institute of Technology, Atlanta, GA 30332, USA*

Abstract: Defense-in-depth is a fundamental safety principle for the design and operation of nuclear power plants in the United States. Despite its general appeal, some authors have identified potential drawbacks in defense-in-depth in its potential for hazardous state concealment. To prevent this drawback from materializing, we propose in this work a novel safety strategy, namely "observability-in-depth". We characterize it as the set of provisions designed to enable real-time monitoring and identification of hazardous states and accident pathogens, and to conceive of a dynamic defense-in-depth safety strategy in which defensive resources, safety barriers and others, are prioritized and allocated dynamically in response to emerging of risks. To better illustrate the role of observability-in-depth in the nuclear industry, we examine in this work the exemplar case study of the Three Mile Island accident and several "event reports" from the U.S. Nuclear Regulatory Commission (NRC) database. The selected cases clarify some of the benefits of observability-in-depth, by contrasting outcomes in situations where this safety principle was violated with instances of proper implementation.

Keywords: Observability-in-depth, accident pathogen, latent failure, defense-in-depth, safety blind spot.

1. INTRODUCTION

Defense-in-depth is a fundamental safety principle for the design and operation of nuclear power plants in the United States. First conceptualized in the 1950s, defense-in-depth became the basis for risk-informed decisions by the U.S. Nuclear Regulatory Commission (NRC) [1,2]. It is also known and adopted under different names in other hazardous industries, for example layers of protection in the chemical industry [3]. In its bare essence, defense-in-depth consists in the design and implementation of multiple safety barriers, technical, procedural, and organizational, conceived to prevent accidents from initiating, to block accident sequences from escalating, and to mitigate adverse consequences should the previous barriers fail. Accidents typically result from the absence, inadequacy, or breach of such defenses [4,5].

Despite its general appeal, some authors have identified potential drawbacks in defense-in-depth [6,7]. For example, its successive lines of defense can (inadvertently) enhance mechanisms that conceal the transition of the system to an increasingly hazardous state, making "systems more […] opaque to the people who manage and operate them" [6]. As a result, system operators may be left blind to the possibility that hazard escalation is occurring, thus decreasing their situational awareness and shortening the time they have to intervene before an accident is released. In other words, defense-in-depth may create safety "blind spots". In extreme cases the efficiency of defense-in-depth may be degraded or worse may backfire, hampering the ability to safely operate the system. Several accident reports identified hidden failures and unobservable accidents pathogens as important contributing factors to the accidents, the Three Mile Island and the Texas City refinery accidents being representative cases [8,9]. The importance of these considerations cannot be underestimated. The NRC database for event reports, that we will analyse in detail later on in this work, reports more than 80 cases of unmonitored release paths for contaminated air and more than 1400 cases of loss of containment[1].

[1] The search, executed on the Licensee Event Reports (LERs) database, for "unmonitored AND release AND path" hits 89 results, while "loss AND containment" hits 1477 results if the search is executed only in the titles and abstracts of the reports. The count of loss of containment events exceeds 8000 cases when the search is

To avoid this potential negative effect, defense-in-depth, and more generally any safety strategy, ought to be augmented with additional guidelines for system design and operation. In previous works, we introduced the observability-in-depth principle as the requirement that all safety-degrading events or states that safety barriers are meant to protect against be observable/diagnosable [7,10]. This principle implies that various features be put in place to observe and monitor for the state and breaches of any safety barrier, and reliably provide this feedback to the operators.

In this work, we illustrate the essential role of observability-in-depth in the nuclear industry. To this aim, we briefly examine the well-known Three Mile Island accident and several "event reports" from the U.S. Nuclear Regulatory Commission (NRC) database. The reports and case studies selected clarify some of the benefits of observability-in-depth, by contrasting outcomes in situations where this safety principle was violated with instances of proper implementation (for example in detecting adverse conditions and guiding safety interventions before an incident unfolding).

The remainder of this work is organized as follows. In Section 2 we provide an overview of the observability-in-depth principle in the context of safety diagnosability[2]. Section 3 highlights the role of observability-in-depth in the nuclear industry through a detailed analysis of the scenarios selected from the NRC database and from well-known exemplar cases. Section 4 concludes this work.

2. SAFETY DIAGNOSABILITY AND OBSERVABILITY-IN-DEPTH

Observability is a Control Theoretic concept, which roughly indicates how well the internal states of a system can be inferred from the system's inputs and outputs[3]. More formally, a generic dynamical system given by Eq. (1)

$$\begin{cases} \dot{x}(t) = F(x(t), u(t)) \\ y(t) = G(x(t), u(t)) \end{cases} \qquad (1)$$

is said to be observable if the knowledge of the set of inputs **u**(t) and the set of outputs **y**(t) – measured from some initial time t_0 – are sufficient to obtain a unique estimation of the system's state vector **x**(t) for all future instants following t_0. Equation 1 indicates a functional relationship between the evolution of the internal states of the system and the system's inputs and current states. In Control Theory, the term *state vector* has a precise formal definition and it constitutes the foundation for most analytical techniques in this field. Roughly speaking, the state vector of a system is the minimum set of variables that contain all the necessary information about the internal condition of a system at some time t_0; and that knowledge, along with the input(s) to the system (e.g., operators' inputs) is sufficient to determine the system's outputs or behavior.

The ability to observe or diagnose the transition of a system to a hazardous state or the occurrence of a safety-degrading event is critical for the continued safety of operations. Roughly speaking, operators make decisions during system operation, which are both based on and affect the internal conditions/states of the system [11]. If process monitoring fails to provide information regarding the actual conditions/states of a system, there is a distinct possibility that operators will make flawed

---

executed in the entire document. The database is available at https://lersearch.inl.gov/LERSearchCriteria.aspx, query executed on 12/09/2013.

[2] This section builds upon and heavily relies on our previous publications on the subject: [7,9,10].

[3] The terms observability and diagnosability are used in a related manner. While there are some differences between them (in their domain of applicability and the nature of the underlying mathematical models they apply to, time-driven dynamical systems in the first case and discrete event systems in the other), these differences are not relevant for our purposes, and we will occasionally use these two terms interchangeably. Observability-in-depth remains the overarching category under which both observability and diagnosability will be subsumed.

decisions, which in turn can compromise the safe operation of the system or fail to check the escalation of an accident sequence (e.g., no decision when an intervention is warranted).

To better illustrate the importance of the notion of observability of hazardous states, consider the following illustrative example[4]. A system departs from nominal operating conditions and begins drifting toward an increasingly hazardous state as shown in Figure 1. Various safety barriers can be interposed between the nominal operating conditions (states) and the accident release (for some specifics about this point, in ref. [9] for example, the system is a splitting tower at an oil refinery, which is filling up with hydrocarbon. The barriers are safety pressure valves and specific design features designed to contain any overflow before the accident, namely loss of containment, occurs). We represent the accident trajectory by plotting the evolution over time of the hazard level of the system, here considered loosely speaking as the closeness of the accident to being released. Assume that safety barriers are implemented to prevent the system from reaching hazard level H_0 in Figure 1, and that additional barriers are in place to block further escalation past H_1 and H_2 should the previous barriers fail or prove inadequate.

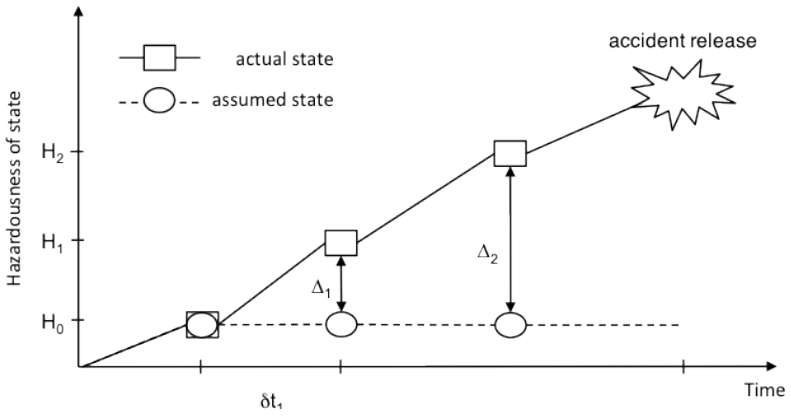

Figure 1. Schematic representation of system diagnosability/observability. Adapted from [9]

The solid line in Figure 1 represents the actual hazard level of the current state of the system, hereafter noted as *H(t)*, while the dashed line represents the operators' assumed hazard level, $\hat{H}(t)$, estimated from available information or through direct sensor observation.

The gap between these two quantities, the actual and the estimated hazard levels can be noted as:

$$\Delta H = \| H(t) - \hat{H}(t) \| \qquad (2)$$

This gap can result for example from the absence of observability into hazardous conditions in the system (e.g., missing sensors), or degraded observability (e.g., flawed or miscalibrated sensors), which can mislead the operators about the actual state of the system. Examples of these situations will be examined shortly.

In a previous work [9], we argued that all safety-degrading events or hazardous states that defense-in-depth is meant to protect against be diagnosable, that is, the failure or breach of any element in the implementation of defense-in-depth be observable—directly monitored or reliably estimated. This constitutes one aspect of the observability-in-depth safety principle. This principle implies among other things, and as a first step, that safety-critical elements in a system be properly instrumented to reflect their actual state, the extent of their degradation if any, and their breach if or when that occurs.

[4] This is based on ref. [9] by the authors and it is included here for convenience and illustrative purposes.

Many examples of accidents occurred, or were not prevented in a timely manner, because of a lack of implementation of this principle. We will examine such cases in Section 3.

In light of Figure 1, the purpose of observability-in-depth is (i) to minimize the gap between the actual and the estimated hazard levels (ΔH), and (ii) to ensure that at the hazards levels associated with various safety barriers, H_0, H_1, and H_2 in the figure, the two curves coincide if these hazard levels are reached (e.g., $\Delta H = 0$ if H_0 is reached—the safety barriers designed to prevent the system from reaching H_0 is breached). The end-objective is to provide sufficient time for the operators to understand an unfolding hazardous situation and intervene in a timely manner to abate it. By contrast, a persistent gap between the actual and the estimated hazard levels, as shown in Figure 1, leaves the operators blind to the developing hazardous situation, and it shrinks the time window, and options, available for the operators to intervene.

While these considerations may be viewed as general common sense, it is important not to underestimate their scope and the potential benefits arising from their formalization. Consider, for instance, the following incident occurred at the Davis-Besse Nuclear Power Station, where the *unobservability* of the degradation of the reactor pressure vessel (RPV) head barrier could have resulted in a massive loss of coolant with potential meltdown of the reactor [12]. In March 2002 a cavity of about 20-30 square inches was discovered by chance (and almost too late) in the reactor lid, and it "extended completely through the 6.63 inch thick carbon steel reactor pressure vessel (RPV) head down to a thin internal liner of stainless steel cladding" [12]. The degradation and breach of the reactor lid developed over an extended period of time unbeknown to the operators and plant managers. It was due to corrosion from a leak of boric acid. This lack of observability of the state or degradation of the reactor pressure vessel head barrier could have resulted in a massive loss of coolant with potential meltdown of the reactor [12]. This was a serious near miss, and the only element that prevented an accident from occurring was the internal cladding, which withstood the primary system pressure over the cavity during system operation and was neither designed for nor qualified to perform such function [12].

There are a number of lessons to be learned from this near miss at the Davis-Besse power plant, and many recommendations were provided in the NRC report [12], including for example heightened regulatory oversight of the plant. In addition to the specific recommendations provided, we propose that this and many other similar near misses support a more general recommendation, namely the adoption of the observability-in-depth safety principle, which was violated in this case, and whose implementation could have identified the degradation of this RPV safety barrier in a more timely manner.

Observability-in-depth can be implemented in many ways, and it requires creativity and technical ingenuity to design and implement in a variety of contexts and for monitoring different types of hazards and states of safety barriers. Regulations cannot be prescriptive in this regard, but a safety case can be required from the designers/operators to demonstrate compliance with this principle.

The *depth* qualifier in observability-in-depth has both a causal and a temporal dimension, and it characterizes the ability to identify adverse states and conditions far upstream (early) in an accident sequence. It also reflects the ability to observe emerging accident pathogens and latent failures before their effect becomes manifest on the system's output or behavior, or before an increasingly hazardous transition occurs in an accident sequence. To illustrate this point, consider Figure 2, which represents a set of safety barriers and various hazardous states. Each hazardous transition/escalation in an accident sequence has a set of underlying causes, and Figure 2 includes the underlying causes of a transition from S_i to S_j in the form of a Fault Tree.

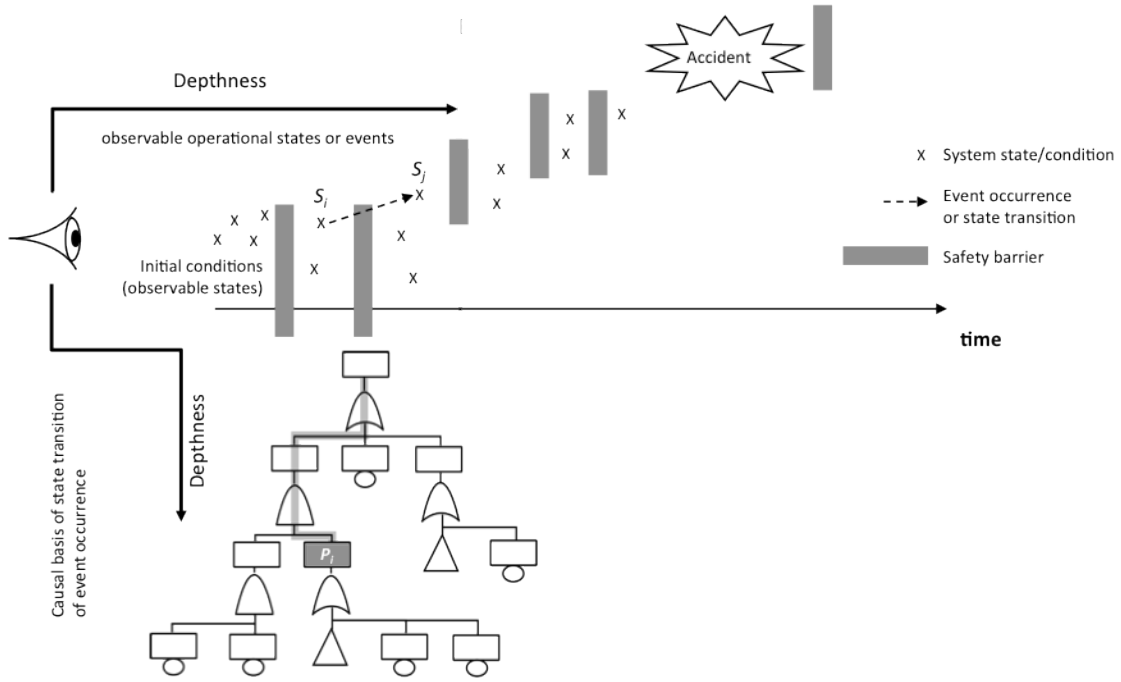

Figure 2. Schematic illustration of Observability-in-Depth

The condition P_i in the fault tree is a latent failure or accident pathogen [13]; it does not have a visible effect on the system behaviour or operation, until the second condition in its AND gate occurs. If the system reaches state S_i, the hazardous transition to S_j will occur, thus further advancing the accident sequence. The ability to observe such causal factors or accident pathogens in an accident sequence before they have a visible effect on the system operation is one measure of the *depthness* of observability. The other measure is that no safety barrier should conceal the fact that the system has breached any one safety barrier and has reached a hazardous state the engineers and system designers meant to protect against.

3. EXAMPLES OF ADVERSE CONSEQUENCES WHEN OBSERVABILITY-IN-DEPTH IS COMPROMISED OR NOT IMPLEMENTED

In this section we provide a few examples that illustrate some of the adverse consequences that can follow from the lack of, or degraded, observability into hazardous conditions. The purpose is to show both the importance of observability-in-depth by examining cases when it is not implemented, and by the same token to highlight the set of problems that it can help address or prevent. We begin with the well-known Three Mile Island accident and examine it from this perspective of observability-in-depth (or deficiencies in). Then we discuss several "event reports" from the U.S. Nuclear Regulatory Commission database, which reflect potential concealment of accident pathogens in the lines of defense.

3.1. The archetype case study: the Three Mile Island accident

The Three Mile Island (TMI) accident of March 1979 is perhaps the most famous incidents in the history of nuclear power plants in the United States. A complex sequence of events led to the loss of the water-coolant, which resulted in a partial core meltdown [8] and caused over $2 billion in damages [14].

The chain of causality leading to the accident has been widely discussed, see for example [8, 15, 16], and the accident became the subject of numerous debates for the complexity of the sequence of events

starting from a leaky valve and emergency pump shutdown and leading to the reactor partial meltdown. The accident resulted from a combination of factors, including four separate malfunctions in the internal and external cooling circuits, overall sloppy maintenance and organizational deficiencies[5] [8, 15], and operators' errors. Our purpose here is not to revisit the accident sequence, but to examine it from one particular perspective, namely that of observability-in-depth, and to highlight how deficiencies in the implementation of this principle contributed to the accident sequence or failed to prevent its escalation. Some brief technical knowledge is required for our discussion. A schematic representation of the reactor core with the cooling system circuits is shown in Figure 3.

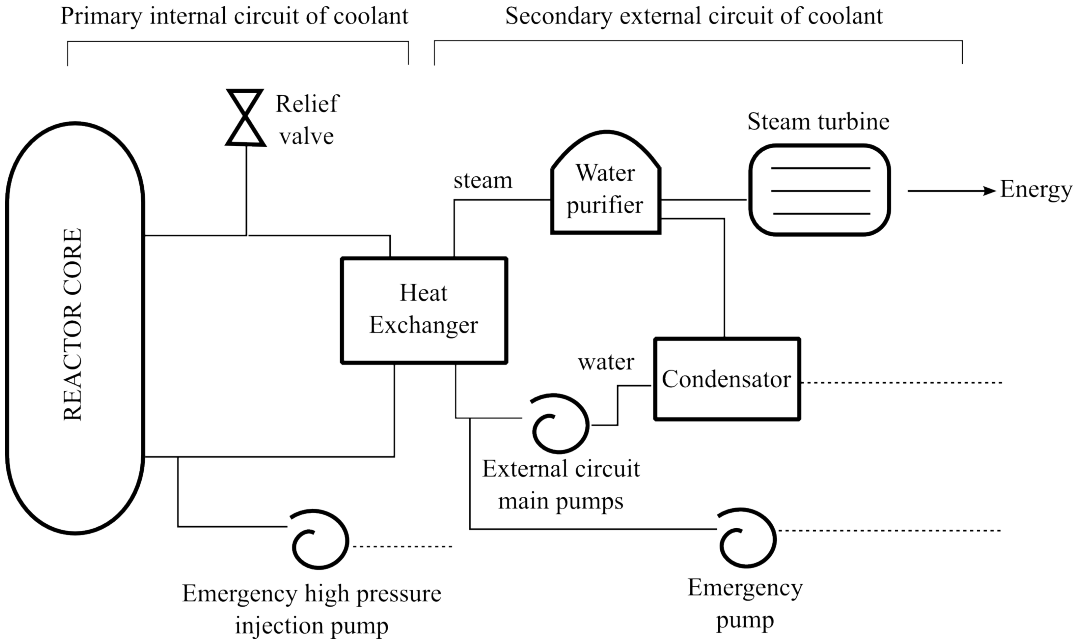

Figure 3. Simplified schematic of the reactor core and the cooling circuits, adapted from [20]

The heat generated by the reactor core at the TMI plant was removed by a heat exchanger at the intersection of two cooling circuits: a primary internal circuit directly connected to the reactor core, and a secondary external circuit connected to steam turbines (see Figure 3). Main pumps as well as emergency backups and pressure relief valves existed for both the internal and external circuits. Steam downstream of the heat exchanger drove the turbines (the power generation elements). This particular design, as well as the main pumps and emergency backups, and the pressure relief valves are specific elements in the implementation of defense-in-depth. And while they were particularly important for the safe operation of the plant, the fact that observability-in-depth was lacking or compromised in their design, as we will discuss shortly, rendered this a defense-blind strategy. Moreover, the inability to observe and assess the states of some of these safety barriers not only failed to prevent the escalation of the accident sequence, but also directly contributed to its advancement. In short, we argue that the Three Mile Island accident was to a large extent the result of a violation of the observability-in-depth safety principle, and while its proper implementation would not have prevented the initiating events from occurring—some of the factors noted previously were directly responsible for this, namely technical failures, sloppy maintenance, and organizational deficiencies—it would have ensured that the accident sequence was terminated in a timely manner before core meltdown.

[5] Gorinson et al. [17] and Hopkins [8] highlight how events similar to those indicated in Figure 4 had occurred in an incident 18 months earlier at the Davis-Besse reactor. Also previous failures of the relief valves had been witnessed in reactors manufactured by the same firm of the TMI plant. These and other near misses and warning signs apparently went unnoticed by the management of the TMI nuclear reactor.

The accident mainly concerned the primary and secondary cooling circuits of the reactor core. The adverse sequence of events was triggered by a leak in the external cooling circuit, which caused the main pumps to shut down [8]. Two days prior to this event, the emergency pumps of the external circuit had been shut down for maintenance work and were still inoperable. This condition, apparently unknown to the operators at the time of the accident (first unobservability), led to the impossibility to dissipate the heat from the internal circuit. As a result, the reactor core began to overheat, leading to its preventive shutdown. However, the pressure in the internal circuit kept increasing due to "decay heat" from the reactor core [8]. At this point, the emergency relief valve of the internal circuit opened, letting the coolant escape and lowering the pressure to the nominal value.

The relief valve should have closed when the pressure fell to proper levels, but it became stuck open. Instruments in the control room however indicated that the valve was closed [18]. The decrease in pressure activated the emergency high-pressure injection pumps in the internal circuit to prevent the core meltdown [8]. After noticing the pressure rise in the internal circuit, the operators were unaware of the loss of coolant from the internal circuit. This is the second major unobservability in this accident sequence, and it was due not only to the flawed sensor that was monitoring the status of the relief valve, but also to the absence, by design, of provisions to monitor and estimate the coolant flow in the primary internal circuit. We conceive of this situation as a gross violation of the observability-in-depth safety principle—two major elements in the implementation of defense-in-depth were not properly monitored and their status not observable.

The operators, still unaware of the loss of coolant from the internal circuit, manually throttled down the emergency pumps. This was considered in hindsight as a significant operator error and it led directly to the accident—the reactor's (partial) core meltdown. However, as shown here, this decision was the result of flawed or missing information that degraded the operators situational awareness and failed to convey the hazardous conditions of various safety barriers. It took them about 2 hours and 20 minutes to understand that a loss of coolant accident (LOCA) was ongoing. The total meltdown was then prevented by flooding the reactor core with cold water. While this extreme measure prevented the release of radioactive material, major irreversible damage had already been done [8].

Different authors have debated at length the controversial issue of "operator error" in the early termination of the high-pressure injection pumps [8, 15]. Hopkins points out that, "had the pumps been allowed to continue operating, the accident could have been avoided" [8]. The "design flaws of the relief valve and its monitoring system caused the control room to receive an incorrect signal of its position. The operators then acted on this incorrect understanding of the plant condition. To the best of our knowledge, there is no explanation in the literature as to why this condition unfolded.

The impossibility to monitor and diagnose an ongoing LOCA from the relief valve status is not the only source of unobservability. For instance, there was no instrument that allowed the operator to understand how much water covered the reactor core [18]. The time history of the water level could have improved the situational awareness of the operators and their understanding of the actual hazard level reached by the reactor.

Perrow choose this accident as the archetype of his "normal accident theory", where an accident "is termed *normal* because it is inherent in the characteristics of tightly coupled, complex systems and cannot be avoided" [15]. The normal accident argument, and specifically its applicability to the TMI accident, was criticized by Hopkins [8]. In his work, Hopkins provides a careful analysis of Perrow's point of view and notes that "Perrow claim[ed] that the information available to the operators [was] so flawed that there was no way they could have been expected to understand what was going on and react in an effective manner" [8]. Perrow's conclusion based on this observation is that the accident was indeed a "normal" occurrence, in the sense that its incomprehensible nature made prevention impossible. We agree that the flawed and missing information about the status of critical safety elements at TMI degraded the operators' situational awareness and hampered their ability to safely operate the plant. However instead of conceiving and accepting this and similar accidents as "normal", we trace back one important element in their causal chain, namely the lack of observability

of emerging hazardous states, and we conceive of a safety principle, observability-in-depth, whose implementation can help prevent similar occurrences.

3.2 Observability-in-depth and the NRC Database of Licensee Event Reports

In this subsection we discuss several event reports from the NRC's Licensee Event Reports (LER) database, which illustrate more situations that can results from the lack of, or degraded, observability into hazardous conditions. By the same token, theses examples highlight an additional set of problems that fall within the scope of observability-in-depth and which this safety principle can help address or prevent.

Caveat: It is important to note that this subsection does not constitute nor should it be considered a basis for statistical analysis of the problem of lack of observability of adverse conditions in the LER database—although this would be an interesting topic and a fruitful venue for future research. This subsection is merely for illustrative purposes and to better delineate the scope and extent of observability-in-depth.

The NRC has required nuclear power plants to submit LER since 1980, and more than 51,000[6] of these reports have since been submitted. Commercial nuclear reactor licensees are required to report certain event information when adverse conditions occur in a nuclear power plant, which are beyond its technical specifications [19, 20]. For example, the malfunction of a required safety barrier or the discovery of a potential design flaw would trigger the need for an LER. Once an LER is submitted, NRC staff review it to understand and confirm the licensee's assessment of the situation. NRC staff experts also determine whether the licensee's resolution of the issue continues to maintain adequate levels of safety and protection of the public [21]. The NRC provides public online access to the LER database. Each report consists of an abstract, a description of the events sequence, the event significance and implications, the identified causes, the implemented corrective actions, and additional information (e.g. information on similar previous occurrences).

Compared with the case study approach in Subsection 3.1, event reports (LER) allow the discussion of a broader set of situations, as the events reported are usually less serious in terms of their consequences and their occurrence relatively straightforward (or not as involved as in the TMI accident).

- *CASE I – Inoperable emergency diesel generator due to low fuel oil in storage tank [22]:* this case resulted from incorrect readings of the level of fuel oil contained in the storage tank of an emergency diesel generator. A low level of fuel oil (below the required minimum) was discovered during an inspection and investigation revealed that incorrect readings had been going on for more than a month. According to the report "the primary cause was a challenging method for determining tank level" [22]. The level indicator reading was also susceptible to exogenous disturbances, becoming "more unreliable under adverse conditions (e.g., poor weather, low light conditions)" [22]. Contributing factors included also a malfunctioning tank level indicator and the corresponding alarm. The investigation highlighted the "inadequate instrument design" of the fuel oil tank level alarm and the indicator.

  While this situation did not pose a considerable threat to the safety of the plant, it constituted a latent failure or adverse pre-existing condition, which when compounded with other factors, could have further advanced an accident sequence, for example if the emergency generators were called upon. As such, this condition constitutes a non-negligible accident pathogen [13]. The fact that it was not observable or its observability compromised is an instance of failed implementation of the observability-in-depth principle (specifically in this case a redundant safety barrier was inoperable and its breach was not monitored or reliably observable).

[6] Query executed on 12/10/2013.

- *CASE II – Unmonitored Flowpath in Safety Injection Cooling Pumps [23]*: In this case a review of the pump testing surveillances showed that unmonitored flowpaths existed for different pumps, including the safety injection pump of the cooling system at the Millstone Nuclear Power Station. According to the report, the regulations current at the time of the system design, did not "explicitly require" [23] to monitor the total pump flow. The unmonitored flowpaths diverted flow from the pump discharge prior to the point at which the flow measurement was taken, hence resulting in a condition of compromised observability of the pump flow.

  At the time of the instrumentation design, the potential impact of unmonitored flowpaths on the ability to test in order to detect pump degradation was not fully realized [23]. Indeed, unmonitored flowpaths have the potential to mask the detection of pump degradation by altering the expected measurements of flow and differential pressure. As with the previous case, this condition compromises the observability of the state of the pump, thus allowing an accident pathogen to emerge and go unnoticed.

- *CASE III – Design Deficiency - Potential for an Unmonitored Release Path [24]:* In this case an unmonitored release path of contaminated air was identified during an engineering evaluation of the station service water system circuit. The identified condition would allow "contaminated air to enter the service water piping [...] and to subsequently be released to the outside environment" in case of a Loss of Power/Loss of Coolant Accident event, thus resulting in a loss of secondary containment [24].

  The failure to identify this release path was considered in the report as non-compliance with General Design Criteria. Moreover, it shows the unobservability of a potential accident sequence. In other words, in case of loss of secondary containment through this particular path, the operators would not be able to identify the release of the contaminated air to the outside environment. While Hopkins claims that "it is not necessary to be able to *predict the precise trajectory* of an accident in order to prevent it" [8], we believe that it is necessary to be able to observe the trajectory in real-time should the accident sequence initiate to be able to properly manage and contain the incident/accident. Observability-in-depth was in this case violated by the unobservability of this particular release path.

- *CASE IV – Unmonitored Release Path Due to Radioactive Ash [25]:* In this case an unmonitored release path of contaminated ash was identified during the preparations to put a heating boiler into service for the winter season. The event report established that "if the ash on the fire side of the boiler contains radioactive constituents, some of the particulate matter could be discharged through the boiler exhaust" [25].

  This event may appear less severe and unremarkable compared with the previous ones. But the interesting point here is that the ashes in the boiler resulted from an original contamination and leak that occurred 25 years before the discovery of the unmonitored release path. This constitutes an interesting example not only of the unobservability of the accident pathogen, but also of the lack of a defense barrier against the release of the contaminated ashes.

Table 1: Selected LER – search keywords and scenarios summary

Case ID and Event Report #	Keyword	Highlighted Scenario
I - 3521996022	Malfunctioning Indicator	Compromised Observability
II - 4231998027	Unmonitored	Compromised Observability
III - 3541997025	Unmonitored	Unobservability
IV - 2451997037	Unmonitored	Unobservability

These cases only represent the tip of the iceberg of instances of adverse conditions that can be gleaned from the LER database, and which can be considered in some ways instances of violation of

observability-in-depth. Further examination of this database for events that include unobservability of adverse conditions and breaches of safety barriers would be a fruitful venue for future research.

4. CONCLUSION

To prevent the hazard-concealing potential of defense-in-depth from materializing, and more generally to introduce a real-time mind-set into risk analysis and management, we proposed in this work a new safety principle termed observability-in-depth. We characterized it as the set of provisions, technical (by design) and operational, designed to enable the monitoring and identification of emerging hazardous conditions and accident pathogens. Observability-in-depth requires among other things that all safety-critical elements in a system be properly instrumented to reflect their actual state, the extent of their degradation if any, and their breach if or when that occurs.

The objective of observability-in-depth is to minimize the gap between the actual and the estimated hazard levels during system operation, and in so doing to provide sufficient time for the operators to understand an unfolding hazardous situation and intervene in a timely manner to abate it. As such, we proposed that observability-in-depth is intimately related to situational awareness, and it supports one important subset of the latter, namely the awareness of the occurrence of hazardous states in the system in real time, and the understanding of the potential accident sequences that might follow. We explained that the *depth* qualifier in our principle has both a causal and a temporal dimension, and it is meant to characterize both the ability to identify adverse states and conditions far upstream in an accident sequence, and to observe emerging accident pathogens and latent failures before their effect becomes manifest on the system's output.

Changing mind-sets: Probability Risk Assessment (PRA), another important pillar in the regulatory regime of the U.S. Nuclear Regulatory Commission[7], has traditionally been performed offline and used as a static tool to help identify and prioritize various risks before system operation. Similarly, defense-in-depth has to some extent an implicit static connotation. Observability-in-depth introduces a real-time mind-set into risk analysis and management, and it supports the development of a "living" or online quantitative risk assessment, which in turn can help dynamically re-order risk priorities based on emerging hazards, and re-allocate some defensive resources accordingly. As such, observability-in-depth can help conceive of a **dynamic defense-in-depth safety strategy** in which some defensive resources, safety barriers and others, are prioritized and allocated dynamically in response to emerging risks.

This work constitutes a first step in the development of the observability-in-depth safety principle, and we hope this effort invites other researchers and safety professionals to further explore and develop this principle and its implementation.

References

[1] Sørensen, J. N., Apostolakis, G. E., Kress, T. S., and Powers, D. A. "*On the Role of Defense in Depth in Risk-Informed Regulation*". In: Proceedings of the PSA '99, 1999.

[2] NRC, US. "*Causes and Significance of Design Basis Issues at US Nuclear Power Plants*". Draft Report, Washington, DC: US Nuclear Regulatory Commission, Office of Nuclear Regulatory Research, 2000.

[3] Saleh, J. H., Marais, K. B., Bakolas, E. and Cowlagi, R. V. "*Highlights from the literature on accident causation and system safety: Review of major ideas, recent contributions, and challenges.*" Reliability Engineering & System Safety, Volume 95, Issue 11, pp. 1105-1116, 2010.

[4] Rasmussen, J. "*Risk management in a dynamic society: a modeling problem*". Safety Science, Volume 27, Issues 2-3, pp. 183–213, 1997.

[7] In addition to defense-in-depth.

[5] Svedung, I., and Rasmussen, J. "*Graphic representation of accident scenarios: mapping system structure and the causation of accidents*". Safety Science, Volume 40, Issue 5, pp. 397–417, 2002.
[6] Reason, J. T. "*Managing the risks of organizational accidents*". Aldershot, Hants, England; Brookfield, Vt., USA: Ashgate, 1997.
[7] Favarò, F. M., and Saleh, J. H. "*Observability in Depth: novel safety strategy to complement defense-in-depth for dynamic real-time allocation of defensive resources*". Presented at the ESREL Conference September 29 – October 2 2013, Amsterdam, 2013.
[8] Hopkins, A. "*Was Three Mile Island a 'Normal Accident'?*". Journal of Contingencies and Crisis Management, Volume *9,* Issue 2, pp. 65-72, 2001.
[9] Saleh, J. H., Haga, R. A., Favarò, F. M., Bakolas, E. "*Texas City Refinery Accident: Case Study in Breakdown of Defense-In-Depth and Violation of the Safety-Diagnosability Principle*". Engineering Failure Analysis, Volume 36, pp. 121-133, 2013.
[10] Bakolas, E., and Saleh, J. H. "*Augmenting defense-in-depth with the concepts of observability and diagnosability from Control Theory and Discrete Event Systems*". Reliability Engineering & System Safety, Volume *96*, Issue 1, pp. 184-193, 2011.
[11] Le Bot, P. "*Human reliability data, human error and accident models- illustration through the Three Mile Island accident analysis*." Reliability Engineering and System Safety. Volume 83, No. 2, pp. 153-167, 2004.
[12] NRC, US "*DAVIS-BESSE REACTOR VESSEL HEAD DEGRADATION LESSONS-LEARNED TASK FORCE REPORT*" available at www.-nrc.-gov/-reactors/-operating/-ops--experience/-vessel--head--degradation/-lessons--learned/-lessons--learned--files/-lltf--rpt--ml022760172.-pdf
[13] Saleh, J. H., Saltmarsh, E., Favarò, F. M., Brevault, L. "*Accident precursors, near misses, and warning signs: critical review and formal definition within the framework of Discrete Event Systems*". Reliability Engineering and System Safety, Volume 114, pp.148-154, 2013.
[14] Sovacool, B. K. "*The costs of failure: a preliminary assessment of major energy accidents, 1907–2007*." Energy Policy, Volume 36, No. 5, pp. 1802-1820, 2008.
[15] Perrow, C. "*The President's Commission and the normal accident*."Accident at Three Mile Island: The Human Dimensions pp. 173-84, 1982.
[16] Rogovin, M. "*Three Mile Island: A report to the Commissioners and to the public"*. No. NUREG/CR-1250 (Vol. 1). Nuclear Regulatory Commission, Washington, DC (USA), 1979.
[17] Gorinson, Stanley, M., and Kane, K. P. "*Report of the Office of Chief Counsel on the role of the managing utility and its suppliers"*. No. NP-25106. President's Commission on the Accident at Three Mile Island, Washington, DC (USA), 1979.
[18] NRC, US "*Backgrounder on the Three Mile Island Accident*", available at http://www.nrc.gov/reading-rm/doc-collections/fact-sheets/3mile-isle.html
[19] NRC, US – LERSearch database website, available at https://lersearch.inl.gov/Entry.aspx
[20] NRC, US "*10CFR50.73 Licensee Event Report System*" available at http://www.nrc.gov/reading-rm/doc-collections/cfr/part050/part050-0073.html
[21] NRC, US public blog "*Easy Searching for Licensee Event Reports*" available at http://public-blog.nrc-gateway.gov/2011/03/04/easy-searching-for-licensee-event-reports/
[22] NRC, US "*Licensee Event Report 96-022-00 Emergency Diesel Generator Inoperable Due to Low Fuel Oil in Storage Tank*" Limerick Geberating Station, Unit 1, December 31st 1996.
[23] NRC, US "*Licensee Event Report 98-027-00 Unmonitored Flowpath in Safety Injection Cooling Pumps May Prevent Detection of Pump Degradation*" Millstone Power Station Unit 3, April 16th 1998.
[24] NRC, US "*Licensee Event Report 97-025-00 Design Deficiency - Potential for an Unmonitored Release Path Through the Station Service Water System*" Hope Creek Generating Station, October 4th 1997.
[25] NRC, US "*Licensee Event Report 97-037-00 Unmonitored Release Path Due to Radioactive Ash in the House Heating Boiler*" Millstone Power Station Unit 1, September 10th 1997.