A Modelling Framework for Dynamic Safety Assessment

Silvia Tolo

Resilience Engineering Research Group, University of Nottingham, Nottingham, U.K.
E-mail: silvia.tolo@nottingham.ac.uk

Rundong Yan

Risk and Reliability Group, University of Loughborough, Loughborough, U.K.
E-mail: r.yan@lboro.ac.uk

Sarah Dunnett

Risk and Reliability Group, University of Loughborough, Loughborough, U.K.
E-mail: s.j.dunnett@lboro.ac.uk

John Andrews

Resilience Engineering Research Group, University of Nottingham, Nottingham, U.K.
E-mail: john.andrews@nottingham.ac.uk

The concept of resilience is progressively making its way into the design, operation and management practice of complex engineering systems. The core of such trend lies with the integration of failure mechanisms in the modelling of systems since the very design phase, focusing on the ability to efficiently absorb and rapidly respond to threats rather than merely avoid them. This is expected to overcome the limitations of traditional design-against-failure approaches, whose efficiency is often undermined by the strong uncertainty associated with rare or hardly predictable hazards. However, the potential advantages such a theoretical shift delivers have not yet been matched by the availability of adequate numerical tools and methodologies targeting the challenges associated with resilience analyses. The current literature and engineering practice lack of a widely agreed upon methodology for the assessment of systems resilience, or even for the definition of its metrics.
This study proposes a novel approach for the estimation of the dynamic response of complex systems to safety-threatening perturbations, aiming at providing a solid base for the evaluation of system resilience. The framework proposed relies on the use of Petri nets to capture both the physics of the processes entailed by the system operation and its interaction with the technological installation. The framework is applied to a case-study focusing on the response of a CANDU nuclear reactor to cyber incidents hindering the correct operation of the reactor control system and hence resulting in a loss of regulation threatening the structural integrity of the nuclear fuel.

Keywords: Resilience, Nuclear, Reactor, Safety, Cyber, Petri Nets

1. Background and Motivation

The extraordinary technological advances experienced in the last decades have on the one hand resulted in the continuously growing level of complexity of today's engineering systems, on the other, have provided the tools necessary to handle such complexity and maintain (when not enhance) operational safety. Hence, it is safe to claim that the research community is continuously advocating, along with the development of modern technology, the benefits of novel approaches able to expand the potential of existing techniques and complement traditional risk assessment methodologies.

In recent times, this has opened the way to a progressive shift in risk science, resulting in the rise of the concept of resilience in the design, operation and management of engineering systems. Rather than to design against failure, the current trend is to consider the existence of failures within the systems design, stressing the focus on their ability to efficiently absorb and rapidly respond to threats rather than to merely avoid them. This is far from being a radical alternative to risk analysis, but should be rather regarded as its natural evolution: the resilience perspective provides theoretical means to better understand threats that may be difficult to adequately predict and characterize, e.g. human error, cyber-attacks, extreme weather events etc.

In spite of the growing popularity of the concept

Proceedings of the 31st European Safety and Reliability Conference
Edited by Bruno Castanier, Marko Cepin, David Bigaud, and Christophe Berenguer
Published by Research Publishing, Singapore.
ISBN: 978-981-18-2016-8; doi:10.3850/978-981-18-2016-8_163-cd

of resilience in all engineering sectors, the potential advantages such theoretical framework delivers have not yet been matched by the availability of adequate numerical tools and methodologies targeting the challenges associated with such analyses. Indeed, while the agreement of the scientific community on the potential of resilience analysis is overwhelming, the same cannot be said of the process behind such analysis and at times even of its definition [Hosseini et al. (2016)]. Hence, the implementation of efficient numerical tools plays a crucial role in the feasibility of resilience analysis and subsequently in the enhancement of systems resilience. The current study is part of a wider research aiming at targeting these shortcomings and providing robust solutions for the quantification of complex systems resilience.

2. Methodology and Challenges

The first challenge encountered has been to identify an adequate metrics for the proposed analysis. As discussed in a previous work [Yan, Tolo, Dunnett, Andrews, and Patelli (Yan et al.)], several candidate parameters have been considered. Being the interest of current research mainly focused on safety-critical engineering systems, hence generally relying on the integration of technological and physical processes, physical parameters have been considered at first. The advantages associated with such solution lie mainly with the easiness of monitoring: physical parameters (e.g. pressure, temperature etc.) are generally easily measurable and can be then continuously monitored. Hence, the suitability of this approach would imply the possibility of tracking the evolution of systems' safety in real-time. Unfortunately, this resulted to not be a viable option, due to the strong limitations of physical parameters in depicting systems reliability. Indeed, the state of physical processes entailed by a system does not necessarily mirror the state of a technological installations. For instance, the failure of secondary safety systems would unquestionably affect the reliability of a power plant, hindering the speed and efficiency of the system response to possible threats, but would not directly affect the normal operation of the facility and hence the physical processes involved. In light of these and other considerations, the failure probability has been investigated in this study as a dynamic safety metrics. The choice was motivated by the capability of such parameter to fully capture the state of the system, and by the availability of well-established methodologies for its quantification, which can thus be adopted as a starting point for the extension of the analysis. It is worth to clarify that the selection of a probabilistic metrics does not lead this latter back to a mere reliability analysis. On the contrary, the aims is to extend traditional notions of reliability analysis to the dynamic context of systems response to threats. According to this interpretation, the focus of the analysis is not on the likelihood of the system failure but on the evolution of such likelihood over the time immediately following the occurrence of a safety-threatening hazard.

This work proposes a novel simulation framework for the numerical quantification of dynamic systems safety considering the failure probability as performance metrics. For this purpose, Petri nets [Petri and Reisig (2008)] have been selected as modelling language. This choice was motivated by the capabilities of such methodology in modelling asynchronous and concurrent processes as well as to integrate delays in timed systems.

In the analysis of safety-critical installations, which are the main focus of the current research, the interest lies mainly at the interface between technological and physical processes. Hence, the computational tools adopted in the current research provide the integration of traditional advanced Petri net language with mathematical equations capturing the dynamic evolution of physical processes. This allows to characterize the mutual influence and interaction between the technological system and physical processes, which are at the very core of the proposed analysis.

3. Dynamic Safety Analysis

The suggested approach has been applied to the case study of a nuclear reactor affected by a cyber attack (or incident) resulting in a loss of regulation accident. The system taken into account is based on the CANDU reactor design Hart (1997) and its Instrumentation and Control (I&C) architecture. According to these, during normal steady-state operation, the bulk reactor power is maintained at the set-point manipulating the reactivity within the reactor. This task is carried out by the Reactor Regulating System (RRS), which continuously monitors and controls the total reactor power to satisfy station load demands as well as the design neutron flux shape. Due to the large dimension of the CANDU reactor core, such tasks are carried out by the RRS on subsection of the reactor rather than its entirety. Indeed, the core is partitioned into 14 geographical zones for control purposes: each of these sections is provided with two platinum-clad Inconel flux detectors and one Liquid Zone Controller (LZC).

Although the in-core detectors are essential for power measurement purposes, they do not provide an absolute value for neutron power: the power estimate is quantified by the RRS logic, which calibrates the detectors reading against thermal power at regular intervals; the thermal power is in turn measured in specific reactor channels (namely Fully INstrumented CHannel, FINCH) provided with ad hoc instrumentation.

LZCs consist of a cylindrical compartment containing light water (which in heavy water reactors behaves as a neutron absorber) and characterized by a fixed outflow and a variable inflow regulated by a valve. Hence, manoeuvring the inflow valve, it is possible to control the light water level in the compartment and hence to inject positive or negative reactivity into the reactor according to the RRS logic demand. This latter is calculated combining the overall power error and any deviation of zone flux from the average over all 14 zones Garland (2016). Additional strategies are available to the RRS in case more negative reactivity is required to meet the power set-point: when the water level in the LZCs approaches the maximum capacity of the compartments, further reactivity reduction is accomplished driving banks of neutron-absorbing rods into the core. On the other hand, when a lack of reactivity is registered, adjuster rods, generally inserted into the core, can be withdrawn.

To ensure the prompt response of the system to abnormal behaviour (e.g. in the case of a loss of regulation) and guaranteeing the safety of the reactor, two independent shutdown systems are in place. The independence between the two is achieved through geographical separation (extended to the associated I&C components), equipment and design diversity and through the adoption of very different actuation technologies. The shutdown system generally referred as 1 (SDS1) relies on the use of 28 cadmium shutoff rods which drop into the core by gravity when the clutches are de-energized; conversely, the shutdown system 2 (SDS2) is expected to inject a neutron absorbing solution into the moderator if trip conditions are met. In order to prevent the erroneous activation of the shutdown systems, a two thirds voting logic is adopted for both SDS1 and SDS2: this implies triplicating the detection equipment (e.g. ion chambers) in order to reject inconsistent sensor readings.

3.1. *Case-Study*

The analysis carried out focuses on the immediate response (i.e. few minutes) of the reactor to a loss of regulation due to a cyber attack. Since the detectors and sensors considered are analogue, the attack is assumed to affect only the RRS logic, with no repercussion on the SDS logic which is independent from the RRS. The simulated attack is designed to undermine the calibration procedure necessary for the RRS power measurement, overwriting the estimated neutron flux and finally resulting in the underestimation of the reactor bulk power. In theory, this would result in the RRS reducing the light water level in the LZCs compartments, leading to a reactivity rise which would be not detected by the in-core detectors due to the fault of the RRS logic unit. However, an excessively fast increase of power, i.e. reactor log rate over 10% or the SDS logic power measurement approaching 122% of full power, would trigger the action of the shutdown systems, bringing the reactor back to safety. The failure of any subsystems involved in such process may hinder the correct implementation of the procedure described, potentially resulting in the further rise of bulk power and the subsequent structural damage of the fuel bundles.

Since the focus of this study is the design and testing of a simulation framework for the dynamic safety analysis rather than the accurate quantification of the latter, a simplified model of a CANDU6 reactor has been adopted. In more detail, the reactor core is assumed to be homogeneous, so that all the LZC are regulated simultaneously and on the basis of the only reactor bulk power, whereas any flux shape effect is neglected. This allows to reduce the analysis to only one of the 14 reactor zones and to extent the values estimated for this area to the remaining fuel bundles. Moreover, LZCs are assumed to be the only control strategy available for the normal regulation of the reactor, while the use of mechanical adjusters is not included in the model. Similarly, only the SDS1 is considered, while the use of neutron poison injection as alternative shutdown strategy (i.e. SDS2) is neglected.

3.2. *Petri Net Model*

The system discussed in the previous section has been simulated adopting traditional Petri nets, resulting in the not hierarchical model provided in Fig.1. The proposed model entails:

- Monitoring Equipment: such as two platinum-clad Inconel flux detectors (*Detector1* and *Detector2* in Fig.1) located inside the core, one Resistance Temperature Detector (*RTD* in Fig.1) monitoring the temperature of the fuel bundle and three ion chambers (*IonChamber1*,*IonChamber2* and *IonChamber3* in Fig.1) located outside the reactor; the two in-core detectors and the RTD are part of the RRS while the ion chambers communicate exclusively with the SDS1 logic. The failure of the in-core detectors (*Detector1Fail* and *Detector2Fail*) and RTD (*RTDfail*) is assumed to result in erratic readings, while the failure of the ion chambers (*IonChamber1Fail*, *IonChamber2Fail*, *IonChamber3Fail*) would cause the freezing of the signal.
- Reactor regulating system logic (*RRS-logic* in Fig.1): it receives the in-core detectors and RTD raw signals and calculates from these the corrected power measurement. This is calibrated against

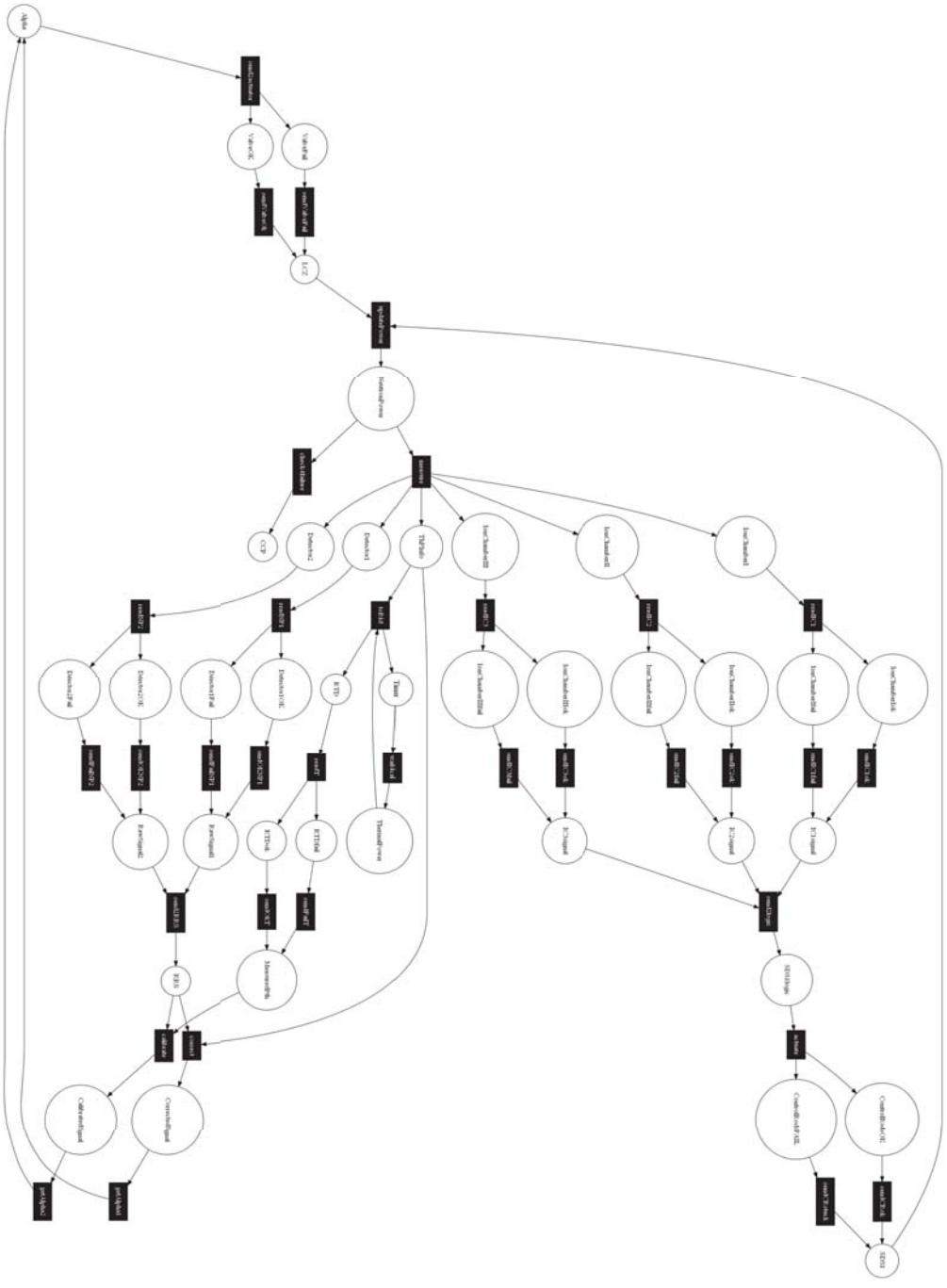

Fig. 1. Overview of the Petri net model adopted for the resilience analysis

thermal power measurements every 2 minutes or otherwise corrected on the basis of the last calibration. The measured reactor power so calculated is then compared against the set-point value, assumed equal to 2620 MW. The position of the LZC in-flow valve is then computed in order to eliminate the power error.

- Shutdown system logic (*SDS1logic* in Fig.1): it compares the neutron power value, measured by the three ion chambers, against two main trip parameters, i.e. the maximum acceptable neutron power (set to 3196.4 MW, namely 122% of full power) and the maximum power log rate (set to 10%), as suggested by Koclas (1996). When any of these threshold values is met, the logic triggers the release of the shutoff rods, initiating the shutdown of the reactor.
- Actuators: one light water compartment (*LZC* in Fig.1) regulated by one in-flow valve (*Valve* in Fig.1): the failure of this component (*ValveFail* in Fig.1) results in the valve being stuck and hence preventing the manipulation of the LZC water level. Similarly, the failure on demand of the release mechanism of the shutoff rods (*ShutoffRodsFail*) would result in the shutdown procedure not being activated.

The cyber malfunction is assumed to affect the RRS logic, tampering the computation of the reactor power and hence the calibration procedure, with subsequent effect on the control of the LZCs. In more detail, the power measurement ($P_{hackered}$) sent to the controller is assumed to drift from the legitimate signal ($P_{legitimate}$) along with time according to Equation 2:

$$P_{hackered} = P_{legitimate} \cdot \dot{d(t - t_0)} \qquad (1)$$

where t_0 refers to the time of the attack and d to a constant determining the drift speed. This latter is uniformly randomly generated between 40 and 80, so to explore the variability of the response against different attack strategies.

Measurement and actuation errors have been taken into account modelling the accuracy of sensors reading and of the LZC valve according to existent literature when available Basu and Bruggemn (1997). As previously mentioned, the failure of the neutron detectors and the RTD results in a significant increase of the measurement error, which causes the erratic behaviour of the signal. Differently, the failure of the ion chambers results in the associated reading to be frozen. Also, the accuracy of the LZC valve has been taken into account in the model Nasimi and Gabbar (2013); however, this is overwritten in the case of failure of the valve, which would result in it being stuck. It is worth to highlight that, in such case, the failure of the component may actually prevent the occurrence of worst case scenarios (i.e. the fuel structural damage in the case under study). Indeed, the failure of the inflow valve regulating the LZC water level would preclude the manipulation of this latter and hence the injection of positive reactivity requested by the compromised RRS logic.

Table 1. Failure rate values and associated references for components subject to failure

Component	Failure Rate (h^{-1})	Reference
Detector	$1.726 \cdot 10^{-8}$	Sion (2007)
Ion Chamber	$5.290 \cdot 10^{-5}$	IAEA (1997)
RTD	$5.707 \cdot 10^{-7}$	McAllindon et al. (1997)
Shutoff Rods	$4.000 \cdot 10^{-7}$	Agency (1988)
Valve	$1.000 \cdot 10^{-5}$	Agency (1988)

The failure probabilities for the mentioned components have been estimated from the failure rate of similar components available in the literature, which are presented in Table 1. No maintenance has been considered.

Beyond the modelling of individual technological components, physical processes such those associated with the change of neutron power (*NeutronPower* in Fig.1) due to variations in reactivity, have been included in the model through the use of a simplified solution to the point neutron kinetics equations AECB (1993), such that:

$$P(t) = P_0 \frac{\beta}{\beta - \rho} e^{\frac{\lambda \rho}{\beta - \rho} t} \qquad (2)$$

where P refers to the neutron power, β is the fraction of delayed neutrons (assumed equal to 0.005) ρ the reactivity and λ the precursors decay constant (assumed equal to 0.1 s^{-1}) CNSC (2003). Such solution is found assuming the reactivity to be lower than the delayed neutron fraction, a step variation in the reactivity to happen at $t = 0$ and the neutron density to be constant prior to the insertion of reactivity. Structural damage is assumed to occur in the fuel, i.e. resulting in core structural damage, when the thermal power overcomes 10 MW per fuel bundle.

3.3. *Results*

The described model has been simulated by means of a purpose-built software developed in C++ environment. The analysis carried out covers over 200,000 simulations, each of which depicts the response of the system along a maximum time

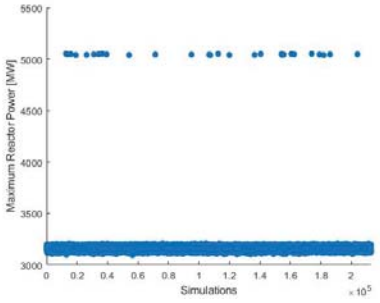

Fig. 2. Maximum reactor power registered per simulation

interval of 200 s following the attack (which is assumed to occur at $t = 0$ s). The range of the time interval has been selected in order to guarantee enough time for the reactor power to reach the fuel structural damage threshold in the case of failure of the safety systems. Indeed, according to the results obtained, fuel damage occurs between 133 and 141 s after the attack, while the shutdown system is triggered within 80 s from the attack. Hence, the maximum simulated time is reached only if the shutdown system is not activated and the power threshold value causing fuel structural damage is not reached. As mentioned previously, this may happen in the case of failure of the LZC inflow valve preventing the compromised RRS logic to inject positive reactivity and would result in the reactor power to assume a constant value along the time domain. However, this scenario has not been recorded in any of the simulations carried out due to the limited sample size compared to the value of the valve failure probability (i.e. equal to $8.3333 \cdot 10^{-08}$ per demand). Conversely, as shown in Fig.2, the vast majority of the simulations register a maximum reactor power compatible with shutdown conditions, with values oscillating around the threshold parameter of 3170 MW. The deviation from the exact value of the shutdown parameter is due to the accuracy of the flux measurements registered by the ion chambers, which explain the uncertainty affecting the measured power and subsequently the activation of the shutdown procedure. Only in 29 of the over 200,000 simulations the reactor power has reached values high enough to trigger structural damage in the fuel bundle (see 2). Figure 3 shows the trend of the reactor power along with time: as expected, the reactor power increases constantly along the time domain, resulting in structural damage of the fuel bundles in the case of failure of the safety systems. These may occur due to the malfunctioning of the electromagnetic clutches preventing the release of the shutoff rods into the core, or for the malfunctioning of at least two of the three ion chambers on which the SDS1 two-thirds voting system relies. However, all the 29 failure events recorded in the analysis fall in the first category and were hence caused by the failure of the shutoff rods insertion. Similarly to the valve failure mentioned above, also this result is due to the limited number of samples analysed and the small probability value associated with the simultaneous failure of two ion chambers.

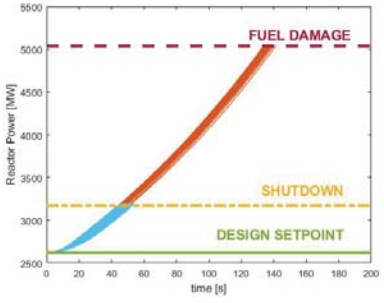

Fig. 3. System response after the attack

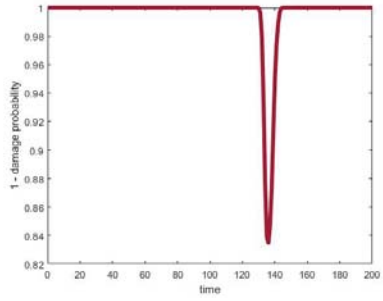

Fig. 4. Evolution of reactor reliability over the accident period

Finally, the discussed results have been elaborated in order to visualise the evolution of system reliability over the accident interval, adopting the probabilistic metric discussed in Section 2. A probability distribution, $p(t)$, function was build over the simulated failure data and the complementary values on the time domain $(1 - p(t))$ adopted for the construction of the dynamic safety profile shown in Fig. 4. This latter, suggests a good degree of safety of the system, since the drop of system performance results to be restricted to a very small region of the domain. As discussed, this area is associated with the failure of the shutdown system, and it is then expected to be an overestimation of real-systems response since these latter benefit from the availability of further

safety systems (e.g. SDS2) to be activated in the case of failure of the primary ones.

4. Conclusions

A simulation framework based on the integration of Petri nets and physical models has been proposed for the analysis of the safety of cyber-physical system subject to incident. The approach has been applied to a simplified model of a CANDU nuclear reactor where the reactor regulating system has been compromised. The framework allows to study the dynamic response of the system immediately after the cyber incident, evaluating the robustness of the system response. The methodology proposed focuses on the quantification of systems failure probability, or better its evolution along the time domain considered which represent the dynamic response of the system to the initial hazard. The implemented framework can be considered to be the first step towards the definition of a methodology for resilience quantification. Indeed, the integration of further aspects of systems' behaviour such as fault detection and restoration, would result in the shift from dynamic safety analysis to resilience analysis, finally providing the resilience profiles associated with the system. However, a simulation-based approach of this type implies the need for high computational power and may easily become not feasible for real-world engineering systems. Hence, future research will focus on implementing computational strategies to overcome the scalability limitations of the current approach.

Acknowledgement

This work has been supported by the UK Engineering and Physical Sciences Research Council (with the project "A Resilience Modelling Framework for Improved Nuclear Safety (NuRes)", Grant No EP/R021988/1), and by the Lloyd's Register Foundation, a charitable foundation in the U.K. helping to protect life and property by supporting engineering-related education, public engagement, and the application of research.

References

AECB, A. E. C. B. (1993). *Fundamentals of Power Reactors*, Chapter Science & Engineering Fundamentals. Canteach.

Agency, I. A. E. (1988). *Component Reliability Data for Use in Probabilistic Safety Assessment*. IAEA.

Basu, S. and D. Bruggemn (1997). Power raise through improved reactor inlet header temperature measurement at bruce a nuclear generation station.

CNSC, C. N. S. C. (2003). *Science and Reactor Fundamentals*, Chapter Vol.1: Reactor Physics. Canteach.

Garland, W. J. (2016). The Essential CANDU, A Textbook on the CANDU Nuclear Power Plant Technology. *Chap 17*, 12.

Hart, R. S. (1997). *CANDU: technical summary*. Atomic Energy of Canada Limited.

Hosseini, S., K. Barker, and J. E. Ramirez-Marquez (2016). A review of definitions and measures of system resilience. *Reliability Engineering & System Safety 145*, 47–61.

IAEA, I. A. E. A. (1997). Generic component reliability data for research reactor psa. iaea-tecdoc-930.

Koclas, J. (1996). Reactor Control and Simulation. *Chulalongkom University, Thailand (published and available to the public in 1996). Presently available online at* `https://canteach.candu.org`.

McAllindon, D., D. Sloan, and P. Mayer (1997). Rtd problems at darlington.

Nasimi, E. and H. A. Gabbar (2013). Analysis of liquid zone control valve oscillation problem in candu reactors. *Journal of Engineering 2013*.

Petri, C. A. and W. Reisig (2008). Petri net. *Scholarpedia 3*(4), 6477.

Sion, N. (2007). Neutron Flux Detector Diagnostic Methodology. *Intercan Technologies, Toronto, Canada. Presently available online at* `http://irpa11.irpa.net/pdfs/3h67.pdf`.

Yan, R., S. Tolo, S. Dunnett, J. Andrews, and E. Patelli. Resilience in the context of nuclear safety engineering. In N. Svartholm (Ed.), *Proceedings of the Reliability and Maintainability Symposium (RAMS) 2020*.