Proceedings of the 33rd European Safety and Reliability Conference (ESREL 2023)
Edited by Mário P. Brito, Terje Aven, Piero Baraldi, Marko Čepin and Enrico Zio
 Published by Research Publishing, Singapore.
doi: 10.3850/978-981-18-8071-1_P369-cd

Life-cycle Considerations for Trusting a Digital Twin for Safety Demonstrations

Ludvig Björklund

Department of Cybernetics, Norwegian University of Science and Technology, Norway.
E-mail: ludvig.bjorklund@ntnu.no

Mary Ann Lundteigen

Department of Cybernetics, Norwegian University of Science and Technology, Norway.
E-mail: mary.a.lundteigen@ntnu.no

Odd Ivar Haugen

DNV, Group Technology & Research, Trondheim, Norway E-mail: odd.ivar.haugen@dnv.com

Trust in evidence produced by the digital twin is crucial for critical decision-making. Establishing trust in the evidence produced by the digital twin require a robust qualification and assurance process of the digital twin. A use case of a digital twin for safety demonstrations is used to determine the need for qualification and assurance process throughout the complete lifecycle for the specific application. A battery system consisting of a battery management system and a battery pack is used as a theoretical case study for identifying specific considerations of the use case. Initially the digital twin require a qualification and assurance process that can extend to rare events and often unobserved states, Modeling the behavior in the presence of physical failures and degradation, corrupt data and other communication issues are crucial for demonstrating the capability of the logic to manage and control the battery system from reaching a failed state. During operation different triggers for updating the digital twin to manage changes in the physical space are identified and considerations for handling the updates are presented. Ultimately, the paper offers a recommendation for managing changes in the digital twin proposing a re-qualification and re-assurance process to be based on the change required in the digital twin.

Keywords: Safety demonstrations, Digital Twin, Trustworthiness

1. Introduction

The fundamental idea behind the digital twin is creating a virtual representation that mirrors or twins an object, process, or system during the entire life cycle. Three pillars are generally considered a part of the digital twin concept: the physical asset, a virtual representation of the asset, and a communication interface for exchanging data between the physical and digital space (Grieves and Vickers, 2017; Jones et al., 2020). Underlying computational models enable simulation-based capabilities for mirroring behavior for applications such as monitoring applications, fault detection, making predictions, optimizing production and testing processes, and more (Rasheed et al., 2020). For making decisions based on evidence produced by a digital twin, it is crucial to establish trust in the accuracy and fidelity of the digital twin. To address this challenge, DNV has developed a recommended practice that outlines the qualification and assurance process for digital twins (DNV GL, 2020). The authors propose a qualification and assurance process proportional to the consequence of the digital twin supporting the incorrect decision. For safety-critical systems the author highlights the importance of considering additional requirements from applicable safety standards (DNV GL, 2020). A significant part of using a digital twin for safety demonstrations is testing safety-critical functionality embedded in software-based control and diagnostics. The digital twin shall evaluate the ability of the controllers to satisfy the safety requirements, by simulating physical degradation, injecting failures, and communication issues (including interactions with other controllers and components). The life-cycle perspective of digital twins extends the application to encompass software updates,

hardware-based configuration managements and updating to account for degradation. Literature related to qualification and assurance of digital twins is limited, and therefore the main contribution of this article is to provide reasoning as it relates to qualification and assurance of digital twins for safety demonstrations.

1.1. *Objective*

This article aims to explore the qualification and assurance process for digital twins utilized to make simulation-based decisions regarding the safety of software-reliant systems. Reasoning of a requirement for a qualification and assurance process of a digital twin for safety demonstration during the complete lifecycle will be discussed, evaluating considerations from a theoretical perspective. Furthermore the aim is to provide an outline and suggestions for considerations regarding building trust in a digital twin for safety demonstrations, to enable the digital twin to be used for producing evidence when coupled to safety-critical functionality embedded in software.

1.2. *Article structure*

Section 2 provides an examination of the use case. In section 3 considerations for qualification and assurance are described in both the design and the operational phase of the digital twin. Section 4 discusses the considerations and recommends a proposed approach for qualification and assurance of the digital twin to cover occurring changes during the complete lifecycle of the physical asset. In section 5 the research is concluded and discussed in a larger aspect of digital twins.

2. A Digital Twin for Safety Demonstrations

Computational models are commonly used to capture the behavior of physical assets, and can be classified as data-driven, physics-driven, or a combination of both. Throughout this article, the term ”digital twin” is used to refer to computational models utilized in simulations for decision support. The digital twin concept, as simulation-based decision support, compared to standalone models, emphasizes a one-to-one connection with a instance of the physical asset twinned to a unique instance of the digital twin. Furthermore, the entire lifecycle of the asset is considered for the mirroring, and finally providing a complete virtual overview of the physical asset distinguishes the digital twin concept from traditional standalone models. The digital twin concept is illustrated in Fig. 1, covering the complete lifecycle of the asset.

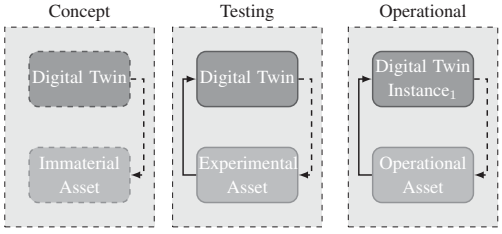

Fig. 1. The lifecycle and unique twinning of the digital twin concept.

The lifecycle starts with the immaterial conceptualization of the physical asset, at this stage computational models can be used to model behavior against the specification to aid decision making for the design of the physical asset. Efficient design is therefore dependent on the fidelity of the digital twin and the digital twin should be assured in order to trust the accuracy of the decision-aid. The testing stage is the final stage before taking the asset operational and requires trusting the digital twin to be used for complimentary evidence. The qualification and assurance process at this stage can integrate experimental data to reduce uncertainty in the digital twin (DNV GL, 2020; Bjorklund et al., 2022). In the operational stage each instance of a digital twin is connected to a unique instance of the physical asset. There are several triggers that can require updates or adjustments of the digital twin to adapt to changes in the physical space, some which may require a re-qualification and periodic assurance process to still make that claim that the digital twin can produce evidence regarding the physical asset.

2.1. *A safety-critical system*

A battery pack in a subsea all-electric actuation system for closing a safety valve, can be considered safety-critical if the overarching safety-critical system operates on an energize-to-close principle. The battery pack is a component of a subsea actuation system that functions to close a valve in the presence of either a process shutdown or emergency shutdown, ensuring overall safety. In the event of a power cut from the topside, the battery shall supply sufficient power to transition the valve and bring the system to a safe state. To maintain the battery and store an sufficient amount of energy at all times, a Battery Management System (BMS) is used. By providing advanced monitoring and control of critical battery parameters, such as voltage, current, and temperature, the BMS serves to maintain the safe and efficient operation of the battery. From this it follows that demonstrating compliance to the requirements on the BMS is important for demonstrating the overall safety integrity of the battery system, i.e the battery pack and the BMS. A simplified battery system is illustrated in Fig. 2. The dashed boxes highlights functionality embedded in logic and the other boxes illustrate the hardware required to realize the system.

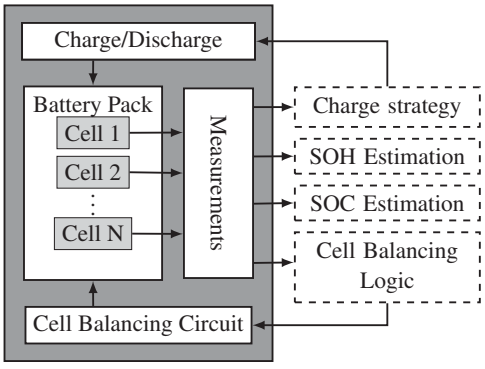

Fig. 2. The battery system separated into hardware and software, where the hardware is modeled in the digital twin.

In Fig. 2 the digital twin consists of the battery pack, the battery charger, sensors, hardware-functionality (protection circuits, cell balancing actuation components), and the load. In this paper the BMS refer to the programmed control logic, e.g. State-of-Charge (SOC) and State-of-Health (SOH) estimators, charging strategies, cell balancing algorithms, isolation control and protection measure.

2.2. *Demonstration of the safety using a digital twin*

The main idea of digital twins for safety demonstrations involves simulating physics-driven computational models interfaced with the control logic under test, as illustrated in Fig. 3. The software-based safety critical functionalities are assumed to be developed according to safety standards. For novel functional systems that are safety-critical, such a standard can for instance be IEC61508 (IEC 61508, 2010). The safety demonstrations covers testing related verification and validation activities using the digital twin as a software offline support tool. A database of test cases is created that are interesting from the point of violating the safety requirements of the BMS. The test manager creates a parameter set $P = p, f, c$ consisting of model parameters, failures, and communication issues.

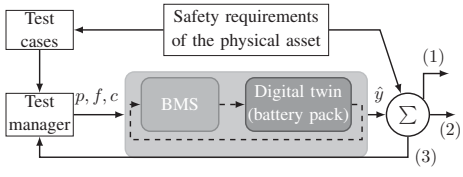

Fig. 3. A proposal for safety demonstrations of the BMS using a digital twin of the battery pack.

The test can also be configured to account for different operating conditions, such as varying charging strategies deployed by the BMS. The model parameters scales the behavior and are based in physical knowledge, e.g. the internal resistance, thermal resistance, cell impurities, etc. Fault injections encompasses issues such as e.g. sensor drifts and thermal issues, and communication issues captures issues such as data packet loss or corrupt information or interconnections with other physical components and interacting

controllers. The times series from the simulation $\hat{y}$ are compared to the safety requirements, with three possible outcomes.

(1) A safety violation occurs, the simulation iterations stops and returns the unsafe parameter set.
(2) If the time series are close to violating the safety requirement, a new iteration of simulation is set to increase the risk of violation by adjusting the parameter set.
(3) IEC 61508 requires sufficient testing in order to verify the functional safety, so after sufficient testing of the software under test, the safety is considered to be demonstrated (IEC 61508, 2010).

The strategy of utilizing a digital twin for these kind of demonstration, we would argue, would categorize the digital twin as a software offline support tool of class T2, which supports testing of the executable code, however it cannot directly create errors in the executable code. This proposed strategy is dependent on trusting the accuracy and fidelity of the digital twin, to produce $\hat{y}$ sufficiently close to what would occur in the battery pack. A qualified and assured digital twin can be considered able to produce expected results of the battery pack.

3. Considerations for a Qualification and Assurance Process.

Qualification and assurance processes, as per DNV's recommended practices, are crucial in establishing trust in a digital twin (DNV GL, 2020). Qualification of the computation model covers the process of providing evidence that the model conforms to the model specification. The model specification encompasses, but is not limited to, the use case, required accuracy or otherwise relevant performance metric, requirements on data quality to the model, and assumptions or simplifications. Making claims about the safety of the battery system during initial design using evidence produced by the digital twin, requires trust in the digital twin in the accuracy, i.e. the capacity of capturing the behavior of the battery pack. Assurance refers to the provision of justified confidence that a product or process conforms to existing safety, environmental, societal or regulatory requirements. Once operational, certain triggers will necessitate updating the digital twin, such as deviation between the physical and digital space, hardware configurations and updates to the software. A strategy for qualification and assurance of the digital twin to remain viable for providing evidence for safety demonstrations in the event of all these update triggers cases are to be considered.

3.1. *Initial qualification and assurance*

A useful digital twin of a battery pack shall sufficiently capture the expected behavior of the battery pack. For the evidence produced by the digital twin to be considered applicable to make claims about the battery pack, the difference in their responses to any stimuli must be sufficiently small, a concept illustrated in Fig. 4. In response to a stimuli u into the digital twin and the battery pack, y is the response of the battery pack, $\hat{y}$ is the response the digital twin, and $\tilde{y}$ is the difference between.

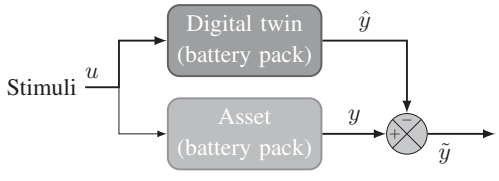

Fig. 4. In response to the same stimuli, a trustable digital twin of the battery pack ought to produce sufficiently close response to the battery pack it is set to mirror.

During the conceptual stage, assurance of the digital twin will be reduced to evaluating the captured behavior against the specifications of the battery pack, for example by achieving the behavior described in data sheets. In later developmental stages experimental data from the digital twin can be considered useful for calibrating the model parameters and increase the confidence level in the ability to accurately capture the behavior of the battery pack (Bjorklund et al., 2022).

3.2. *Update to manage deviations*

The digital twin can describe a specific battery pack with greater accuracy, by calibrating model parameters and attributing small deviations to incorrect assumptions on initial conditions or modeling uncertainties (Bjorklund et al., 2022). After this initial configuration period, deviations will be attributable to, degradation, faults and other issues in the battery pack. Deviations not explainable by aging models will eventually occur, resulting in a deviation between measurements and the expected values of the digital twin as illustrated in Fig. 5.

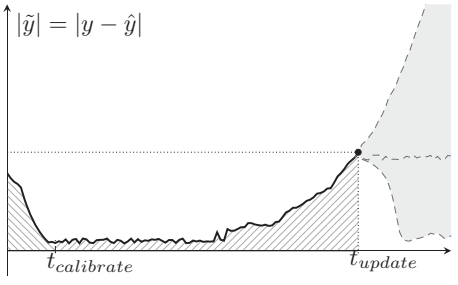

Fig. 5. During operation, the deviation between the digital twin and the physical asset is expected to increase.

Approaches for using the digital twin as a reference model and use the deviations to detect unwanted behavior are one natural application of such deviations (Milton et al., 2020; Xiong et al., 2022). However, the objective of the digital twin for safety demonstrations rely on an accurate digital twin. Therefore integrating the deviation to an explanatory variable inside the digital twin is of importance. Two versions of the digital twin are declared to be useful for enabling integrating the deviation, a offline and an online digital twin. The online digital twin is connected to the BMS and is used for detecting deviations, see Fig. 6. For testing and demonstrating the safety of the battery system, an offline digital twin can be used to directly interface with a replica of the BMS under test. The separation of activities in the digital twin allows for deviation detection, while providing evidence on the safety of software updates using the most up-to-date offline version of the digital twin.

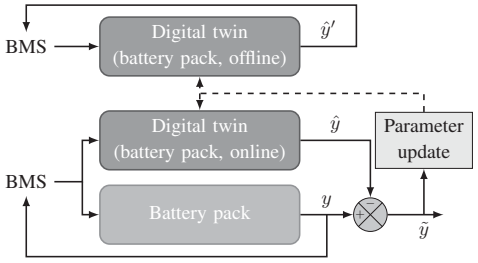

Fig. 6. Comparing the behavior of the digital twin with the actual behavior of the battery pack measured, will inevitably start to deviate.

3.3. *Update to manage hardware changes*

The battery pack is part of a subsea actuation system, providing a local energy storage for a safety valve installed on a X-mas tree. Due to the inaccessibility of the location, the entire development of the subsea actuation system is designed for a lifetime of about 25+ years. However, since the battery system is safety critical, intractable degradation or faults in the battery pack may warrant replacing components or the entire battery pack.

3.4. *Update to manage software updates*

The qualification and assurance of the digital twin is performed in order to enable the digital twin to produce evidence about the capability of the BMS to keep the system from entering a dangerous state. A lot of the functionalities, e.g. SOH are dependent on estimators, that may require software rejuvenation to remain reliable and remove round-off related errors. The lifespan of the battery pack can allow for updating or adding functionality to the BMS to optimize remaining useful life. An example of added functionality that may require offline safety demonstrations on the digital twin is an algorithm for detecting internal short circuits in battery cells. Although there are several ways of detecting internal short circuits, assume that the detection capability is not available at the time of deploying the battery system to illustrate the issue of added functionality. Demonstrating the safe execution of the algorithm for detecting internal

short circuits will be invalid if the digital twin is unable to capture this behavior.

4. Results and analysis of the use case

In this section, the use cases and considerations highlighted in the previous section are discussed and analyzed to define a step-wise process for qualification and assurance of a digital twin used for safety demonstrations. It is important to acknowledge that there is an assumption that the development of control and diagnostic logic adhere to the development flow specified in relevant standards, such as IEC 61508. Employing a digital twin for safety demonstrations is just one aspect of the overall process for developing safety-critical systems.

4.1. *Initial qualification and assurance*

For safety demonstrations, the digital twin must be capable of sufficiently capturing behaviors beyond normal operating conditions. The digital twin is required capture safety-related behavior during rare events, i.e. faults or other issues. While behavior of certain faults may require specific physical testing, it is unfeasible to cover all the behavior in experimental data from the physical space. Therefore, there will be a need to infer that the digital twin can function outside the usual operating conditions to some degree. To this end, the overall development process of a digital twin for safety demonstrations, will require transparency and traceability in the modeling decisions. Computational models grounded in first principles shall be developed, featuring clearly defined relationships between the physics-driven model parameters. Parameters that are excluded from the model must be documented with sufficient evidence to support the assumption that they do not contribute to safety-related behavior (Bjorklund et al., 2022). Proper documentation of parameter uncertainty, including a physically-supported range, should be included as part of the digital twin qualification and assessment process. If simplifications are made, such as a limited scope of the operational range for the computational models, this should be clearly specified (DNV GL, 2020). Since there is a need for assumptions related to rare and undesirable behavior, these assumptions should be clearly stated and the modeling process must be transparently documented. In the recommended practices by DNV the suggestion is to have a strategy of re-qualification and continuous assurance of the digital twin, however this implies a continuous update or continuous changes in the digital twin. In the context of safety demonstrations, continuous updates of the digital twin may not be required, as the objective of the use case does not mandate it. Rather, event-based or periodic updates are more appropriate.

4.2. *Process to manage deviations of measurements*

Violations of a certain threshold can be used as an indication for requiring an update of the model parameters. In Fig. 6 the parameter update step is intentionally left vague, since the issue of attributing the deviation to a (or multiple) model parameters or potential faults is an inherently uncertain process. A lower measured voltage can be attributable to parameters such as the internal resistance, accelerated capacity fade, the diffusion rate and migration of ions inside the battery. Configuring the model parameters of the digital twin to mirror the battery pack, can quickly become difficult even with diagnostic input from the BMS. The method for updating the model parameters can be considered as either a manual-, automated or semi-automated updating function.

Assuming that sensor inaccuracies can be ruled out, for example due to redundancy, the BMS will monitor and detect a capacity fade in the individual battery cells. The capacity fade can be captured in the model by parameters such as increased internal resistance, internal short circuit, or manufacturing defects (i.e. material impurities), a combination of the parameters, etc. However, in the design stage, a digital twin is used for simulation-based verification to assess the BMS's ability to comply with specifications, monitor, and detect faults and degradation in the battery pack's lifespan. This testing process includes adjusting the internal parameters of the battery pack, such as individual capacities, internal resistance, and

material impurities, to account for cell characteristics and model uncertainties.Therefore, an assumption is made that modifying the parameters or introducing faults into the digital twin does not necessitate any further qualification or assurance activities to be placed on the digital twin due to parameter updates or once it has passed the initial qualification and assurance process. It can additionally be argued that the use case of the digital twin is to demonstrate the safety of the BMS, and the computational model can be split into numerous simulations that could run offline. Therefore it will be possible to provide evidence about the safety using the digital twin, regardless of the parameter uncertainty.

4.3. *Process to manage hardware changes*

The long lifespan of the battery pack makes hardware adjustments or update an interesting consideration. Hardware updates can be split into two categories: those that require parameter updates and those that introduce new behavior. This distinction can be exemplified using the cell balancing function of the BMS. The objective of cell balancing is to manage the individual SOC of cells within a battery pack to prevent overcharging or over-discharging, thereby improving the overall performance and lifespan of the battery pack. Cell balancing can be achieved through either passive or active methods, with the former discharging cells through a dissipatative bypass route and the latter redistributing excess charge. Assuming that the original system is deployed with a passive cell balancing circuit that after some cycles requires a replacement. In the case of a similar circuit used to balance the cells, the update may be limited to updating the internal parameters. Introducing a active cell balancing strategy would in comparison require a significant update to capture the newly introduced behavior. Qualification and assuring of the digital twin for the first case would follow the same recommendations as for the parameter updates described in the deviation examples. For the latter case, the qualification and assurance process should be repeated in order to qualify the digital twin.

4.4. *Process to manage software updates*

A distinction can be made between software updates that necessitate capturing new safety-related behavior and those that do not. An example of software update that does not introduce any new safety-related behavior could be a novel algorithm for estimating the SOC. Assuming that the same basic measurements are already modeled in the digital twin. A change that does not warrant a significant update in the digital twin should not require a re-assurance of the digital twin. The same assumptions can be made for tuning the estimators, such as parameters of Kalman filters, updates that mainly affect the performance of the control logic does not require an update in the digital twin. An example of introducing novel functionality to the BMS is a internal short circuit detection algorithm. Assuming that the original BMS is deployed without the capability to detect short circuits, the digital twin may lack the capability to inject such a fault, requiring adjustments of the digital twin. Managing such a change will require a process of re-qualifying and re-assuring the digital twin to ensure trust.

4.5. *Recommendations for managing an updated digital twin*

The recommended practices for adapting the digital twin to changes in the physical space are summarized in Table 1. For each trigger two changes have been identified, examples are provided and the suggested action of re-assuring the digital twin is provided.

5. Conclusion

Qualifying and assessing a digital twin involves demonstrating that it can be utilized to generate evidence for a specific use case. The development of the computation models used for simulating safety related behavior, must be clearly documented and transparent. The computation models should be physics-driven to enable transparent testing of safety-critical software. Further work on this topic is of relevance, ultimately aiming to augment the digital twin with sufficient documentation to show the reasoning, assumptions and simplifications made during the initial design.

Table 1. Proposal for change management in a digital twin for safety demonstrations.

Trigger	Change	Example (s)	Action
Deviation	Parameter update	Capacity, internal resistance	No re-assurance required
	Fault Injection	Short circuit, sensor drift	No re-assurance required
Hardware	Direct replacement	Cell balancing board, battery cells	No additional qualification
	Augmented or new hardware	Active cell balancing battery board	Re-qualify the digital twin
Software	Software rejuvenation, tuning	SOC and SOH estimators	No additional qualification
	Novel functionality	Short circuit detection	Re-qualify the digital twin

During the operational stage of the physical asset, the ability of a digital twin to provide evidence regarding the safe and correct execution of software-based functionality can be considered independent of mirroring the current state of the physical asset. The digital twin is not required to correctly mirror the cause of a deviation, if it can provide evidence on the safety of the software-based functionality in the event of each of the potential causes. For added or adjusted software or hardware functionality that requires capturing novel behavior not introduced in the original computational models, the process of re-qualifying and assuring the digital twin must be repeated.

5.1. *Further work*

This article draws conclusions from argumentation and discussion about foreseeable considerations for qualification and assurance of a digital twin. In future research the proposed strategy will be evaluated using a case study for qualifying and assessing the capability of the digital twin to produce evidence of the safety of the physical asset. This includes the additional living documents that track the changes and simulations made in the digital twin and relates these to an initial parameter sheet that tracks data related to the physical parameters of the asset.

Acknowledgement

This work was carried out as a part of SUBPRO, a Research-based Innovation Centre within Subsea Production and Processing. The authors gratefully acknowledge the project support from SUBPRO (grant number 237893), which is financed by the Research Council of Norway, major industry partners and NTNU. Additional acknowledgment goes out to the team working on the ISSA project at the Institute for High Integrity Mechatronic Systems, Aalen University, for an interesting case study.

References

Bjorklund, L., M. Glaser, S. Imle, G. Skofteland, and M. A. Lundteigen (2022). Design of a digital twin of gate valves for partial stroke testing. In M. C. Leva, E. Patelli, L. Podofillini, and S. Wilson (Eds.), *Proceedings of the 32nd European Safety and Reliability Conference (ESREL 2022)*, Singapore, pp. S14–08–406–cd. ESREL2022 Organizers: Research Publishing.

DNV GL (2020, 10). Qualification and assurance of digital twins. Recommended Practice DNVGL-RP-A204, edition October 2020.

Grieves, M. and J. Vickers (2017). Digital twin: Mitigating unpredictable, undesirable emergent behavior in complex systems. In *Transdisciplinary perspectives on complex systems*, pp. 85–113. Springer.

IEC 61508 (2010). *Functional Safety of Electrical/Electronic/Programmable Electronic Safety-Related Systems.* Geneva: International Electrotechnical Commission.

Jones, D., C. Snider, A. Nassehi, J. Yon, and B. Hicks (2020). Characterising the digital twin: A systematic literature review. *CIRP Journal of Manufacturing Science and Technology 29*, 36–52.

Milton, M., C. D. L. O, H. L. Ginn, and A. Benigni (2020). Controller-embeddable probabilistic real-time digital twins for power electronic converter diagnostics. *IEEE Transactions on Power Electronics 35*(9), 9850–9864.

Rasheed, A., O. San, and T. Kvamsdal (2020). Digital twin: Values, challenges and enablers from a modeling perspective. *IEEE Access 8*, 21980–22012.

Xiong, J., H. Ye, W. Pei, L. Kong, Q. Huo, and Y. Han (2022). A monitoring and diagnostics method based on FPGA-digital twin for power electronic transformer. *Electric Power Systems Research 210*, 108111.