Data-Centric Engineering (2026), 7: e33
doi:10.1017/dce.2026.10072

CAMBRIDGE UNIVERSITY PRESS

RESEARCH ARTICLE

Adaptive digital twins for predictive decision-making: online Bayesian learning of transition dynamics

Eugenio Varetti[1], Matteo Torzoni[2], Marco Tezzele[3] and Andrea Manzoni[1]

[1]MOX—Dipartimento di Matematica, Politecnico di Milano, Italy
[2]Dipartimento di Ingegneria Civile e Ambientale, Politecnico di Milano, Italy
[3]Department of Mathematics, Emory University, USA
Corresponding author: Marco Tezzele; Email: marco.tezzele@emory.edu

Received: 16 December 2025 UTC; **Revised:** 26 June 2026 UTC; **Accepted:** 01 July 2026 UTC

Keywords: Bayesian networks; digital twins; parametric Markov decision processes; reinforcement learning; structural health monitoring

Abstract

This work shows how adaptivity can enhance the value realization of digital twins (DTs) in civil engineering. We focus on adapting the state transition models within DTs represented through probabilistic graphical models. The bi-directional interaction between the physical and virtual domains is modeled using dynamic Bayesian networks. By treating state transition probabilities as random variables endowed with conjugate priors, we enable hierarchical online learning of transition dynamics from one state to another through effortless Bayesian updates. We provide the mathematical framework to account for a larger class of distributions with respect to the current literature on DTs. To compute dynamic policies with precision updates, we solve parametric Markov decision processes through reinforcement learning. The proposed adaptive DT framework enjoys enhanced personalization, increased robustness, and improved cost-effectiveness. We assess our approach on a case study involving structural health monitoring and maintenance planning of a railway bridge.

Impact statement

We introduce an adaptive digital twinning framework that integrates dynamic Bayesian networks and hierarchical online learning to enable closed-loop, data-driven decision-making. Deep learning is leveraged for real-time structural health tracking, while reinforcement learning optimizes decision-making. By continuously refining state transition dynamics through conjugate Bayesian updates, the proposed approach enables digital twins (DTs) to learn and adapt over time. Application to the health monitoring, management, and maintenance of a railway bridge demonstrates that adaptive digital twinning achieves improved physical–digital synchronization, personalized predictive capabilities, and dynamically optimized decision-making under uncertainty.

1. Introduction

Monitoring structural integrity, maintaining planned functionality, and managing the operation of critical structural systems represent paramount challenges in modern engineering (Farrar and Worden, 2013; Avci

This research article was awarded Open Data and Open Materials badges for transparent practices. See the Data Availability Statement for details.

© The Author(s), 2026. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (http://creativecommons.org/licenses/by/4.0), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

https://doi.org/10.1017/dce.2026.10072 Published online by Cambridge University Press

et al., 2021; Torzoni et al., 2022). These systems are often threatened by material deterioration, cyclic loading, and extreme events. Such variability, combined with natural hazards increasingly driven by climate change (American Society of Civil Engineers, 2018), calls for a new level of "adaptivity" in the monitoring and maintenance of engineering systems (Torzoni et al., 2026). This is essential to respond effectively to changing conditions over time and ultimately to minimize the risks associated with compromised structures (Brownjohn, 2007; CR Farrar and Worden, 2007; ASCE, 2021). Additionally, asset managers must ensure the practical and effective use of services, avoiding unforeseen access limitations that could lead to higher costs for users (Thoft-Christensen, 2009; Jones et al., 2019). Decision-makers must also maximize profits (or minimize losses), which results in a multi-objective optimization problem under uncertainty. This challenge can be effectively addressed by combining structural health monitoring (SHM) with sequential decision-making within a DT framework, enabling effective damage identification strategies and optimal action-planning schedules.

A computational model that overlooks the uncertain and dynamic nature of engineering systems limits its personalization, which in turn reduces the model accuracy and the value of insights derived from it (Moya et al., 2020; Cherifi et al., 2022; Arcones et al., 2024). According to the 2024 report by the *National Academies of Engineering, Science, and Medicine* (National Academy of Engineering and National Academies of Sciences, Engineering, and Medicine, 2024), a DT can be distinguished from instances of digital models (forward models from parameters to outputs of interest) or digital shadows (data assimilation and model updating engines) (Kritzinger et al., 2018). A DT is defined as a personalized virtual representation of specific attributes of a natural or engineered system or process. It relies on computational models that dynamically mirror the physical counterpart by continuously assimilating sensor data and refining predictive capabilities accordingly. This continuous updating enables the simulation of what-if scenarios, feeding predictive decision-making that is tailored to realize value. Rooted in the "Dynamic Data Driven Applications Systems" paradigm (Darema, 2004), this bi-directional interaction between the virtual and the physical forms a feedback loop that propels the DT concept toward autonomy rather than simple automation. DT applications today span a variety of fields, including SHM and predictive maintenance (Li et al., 2017; Karve et al., 2020; Ye et al., 2020; Löhner et al., 2024; Torzoni et al., 2024), additive manufacturing (Phua et al., 2022), smart cities (Jans-Singh et al., 2020), urban sustainability (Tzachor et al., 2022), railway systems management (Arcieri et al., 2023, 2024), spacecraft operations (Henao-Garcia et al., 2025), and aerial vehicles monitoring and control (Kapteyn et al., 2020; Kapteyn et al., 2021; McClellan et al., 2022; Sisson et al., 2022; Tezzele et al., 2026), among others. Beyond engineering, the DT paradigm is also being deployed in healthcare for advancing medical diagnosis and personalized treatment (Corral-Acero et al., 2020; Chaudhuri et al., 2023), as well as in climate science (Bauer et al., 2021).

Adopting a DT perspective for SHM purposes is crucial to enable condition-based or predictive maintenance practices (Glaser and Tolman, 2008; Achenbach, 2009; Torzoni, 2024). To this aim, in situ inspections are inadequate for continuous and global monitoring. On the other hand, by assimilating sensor data from permanent data collecting systems (García-Macías and Ubertini, 2022; Torzoni et al., 2022), vibration-based techniques enable automated damage identification and evolution tracking. This paradigm holds the potential to unlock personalized monitoring and management programs (Thelen et al., 2022), yielding numerous benefits throughout the system life cycle, including more informed assessments of structural safety, better resource allocation, and increased system availability.

In Torzoni et al. (2024), the authors presented a DT framework for engineering structures based on probabilistic graphical models (PGMs) for predictive DTs introduced in MG Kapteyn et al., (2021). Specifically, the DT is built on a dynamic Bayesian network (DBN) with decision nodes (Koller and Friedman, 2009; Russell, 2020), which encodes the observation-to-decision interaction and its temporal evolution. The computational flow spans from physical to digital through data assimilation and inference, and then back to the asset via control actions, while also accounting for uncertainty quantification and propagation. As depicted in Panel 1 of Figure 1, vibration recordings are assimilated for structural health diagnostics (physical-to-digital information flow), leveraging the flexibility of deep learning (DL) models (Torzoni et al., 2023a, 2023b; Franco et al., 2024; Tong et al., 2024; Torzoni et al., 2025; Vlachas et al.,

https://doi.org/10.1017/dce.2026.10072 Published online by Cambridge University Press

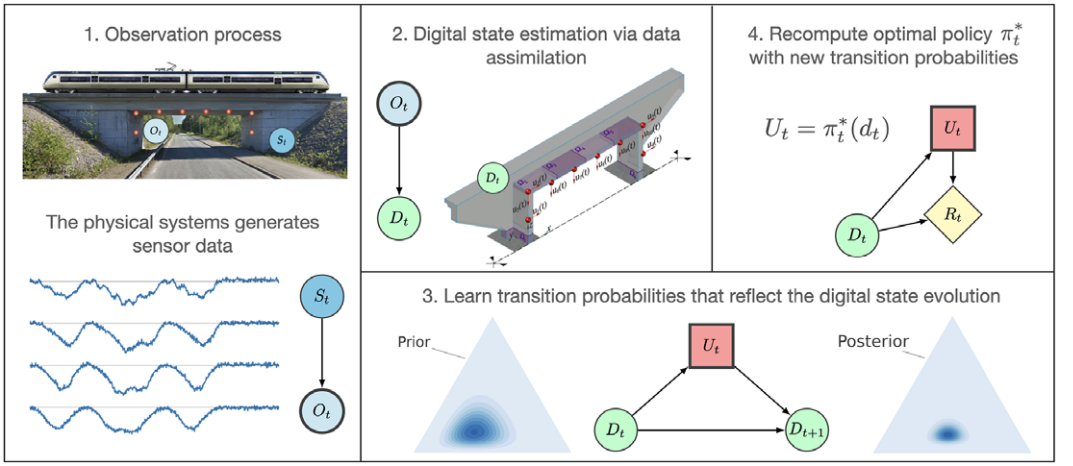


Figure 1. *Graphical abstraction of the end-to-end information flow enabled by the probabilistic graphical model. Capital letters denote time-dependent random variables, while the corresponding lowercase letters refer to their specific realizations. Circular nodes denote random variables, square nodes denote actions, and diamond nodes denote the objective function. Nodes with bold outlines denote observed quantities, while nodes with thin outlines denote unobserved variables to be inferred. Directed edges encode conditional dependencies between variables. (Panel 1) The hidden physical state $S_t \sim p(s_t)$ governs the emission of observational data $O_t \sim p(o_t)$. (Panel 2) The assimilation of observational data characterizes the physical-to-digital information flow through the inference of a digital state $D_t \sim p(d_t)$, which captures the essential features of the physical state. (Panel 3) The update of the digital state occurs within a sequential Bayesian inference formulation, where a stochastic parametrization of the state transition probabilities enables online learning through continuous updating of state transition beliefs. (Panel 4) The updated digital state informs the digital-to-physical information flow by guiding the selection of control actions $U_t \sim p(u_t)$ to maximize the performance of the asset–twin system, as measured by the reward $R_t \sim p(r_t)$. The optimal control policy π_t^* governing this process is continuously refined as the transition beliefs are updated.*

2025). These DL models are preliminarily trained offline on structural health scenarios generated from a numerical model that simulates the structural response to applied loads. Once trained, the DL predictors map vibration recordings to parameters describing the presence, location, and severity of damage. This initial estimation is integrated into a sequential Bayesian inference formulation to continually update a digital state encoding the health condition of the asset (see Panel 2 in Figure 1). This updated characterization is then used to forecast future damage evolution via DBN unrolling. The predicted trajectories guide a Markov decision process (MDP) for maintenance and management actions (digital-to-physical information flow). The health-aware control policy is also precomputed offline using dynamic programming (Bellman, 1957a) to solve the planning problem induced by the PGM. After an action is taken, the system state evolves, and the perception-action cycle begins again.

This work enhances the aforementioned framework by incorporating adaptivity to improve decision-making capabilities. Specifically, the reference PGM includes a transition model that serves as a control-dependent predictor, propagating digital state beliefs forward in time. While this internal model is typically predefined using historical data or physics-based models, as in our previous work (Torzoni et al., 2024), here we adopt a stochastic parametrization that enables online learning of state transition beliefs through Bayesian updates (see Panel 3 in Figure 1). This adaptivity ensures that the DT remains synchronized with its physical counterpart and can effectively respond to changing conditions over time. Moreover, moving beyond precomputed optimal policies—based on the unrealistic assumption of

https://doi.org/10.1017/dce.2026.10072 Published online by Cambridge University Press

precisely known transition dynamics—we formulate a parametric MDP and develop dynamic policies with continuous precision updates (see Panel 4 in Figure 1). Specifically, we leverage the refined transition dynamics and employ model-based reinforcement learning (MBRL) to optimize control policies over the expected remaining lifespan of the structure. A similar approach, not in a DT context, for partially observable MDP was proposed in (Memarzadeh et al., 2015), where the authors utilize a Bayes-adaptive framework to address real-time decision-making for the management of wind farms under uncertainty.

The paper is organized as follows. Section 2 reviews the components of the DT framework as presented in (Torzoni et al., 2024). Section 3 describes the proposed adaptive capabilities within the context of parametric MDPs, extending (Tezzele et al., 2026) to a broader class of probability distributions. This section also shows how optimal policies can be systematically updated using RL techniques. Section 4 assesses the computational procedure through the simulated monitoring of a railway bridge. Conclusions are drawn in Section 5.

2. Probabilistic graphical models for digital twins

DBNs are a specific type of PGM used to model probability distributions that evolve over time, making them suitable for inference and learning in dynamic environments (Koller and Friedman, 2009; Russell, 2020). This section introduces the asset–twin DBN model used in (Torzoni et al., 2024), which is based on the mathematical abstraction proposed in MG Kapteyn et al., (|2021). We begin by introducing DBNs and then detail the specific model used in this work.

2.1. Dynamic Bayesian networks

PGMs represent joint probability distributions through efficient factorization, leveraging the conditional independence among random variables induced by the graph structure. Bayesian networks are a specific type of PGMs represented by directed acyclic graphs. A directed graph $\mathcal{G} = \langle \mathcal{V}, \mathcal{E} \rangle$ is a graph composed of a set of vertices $\mathcal{V}$ connected by directed edges, in $\mathcal{E} \subseteq \{\mathcal{V} \times \mathcal{V}\}$ which is a set of ordered pairs of vertices.

In a Bayesian network, $\mathcal{V}$ represents a set of random variables, and the edges connecting the nodes represent causal relationships through conditional probability distributions (CPDs). Furthermore, given a set $\mathcal{X} = \{X_1, \ldots, X_N\}$ of N random variables (whether observed or hidden), we can factorize their joint probability distribution using the chain rule of probability, effectively leveraging the conditional independence structure induced by the graph topology. To this end, each node X_i, $i = 1, \ldots, N$, has an associated factor $\phi^i = p(X_i | \mathrm{Pa}_{X_i})$, which quantifies the influence of its parent nodes Pa_{X_i}, reflecting the CPDs implied by the joint distribution. Nodes with no parents represent independent variables, specified by a prior (unconditional) distribution.

To extend Bayesian networks to settings where variables evolve over time, we denote the template variable $X_i \in \mathcal{X}, i = 1, \ldots, N$, at a specific point in time t as $X_i^{(t)}$. The trajectories of the (dynamic) Bayesian network refer to the complete set of possible realizations over a time horizon $[0, T]$, with $t \in [0, T]$. The joint distribution for all possible trajectories reads as follows:

$$p\left(\mathbf{X}^{(t_1:t_2)}\right) = p\left(\mathbf{X}^{(t_1)}, \ldots, \mathbf{X}^{(t_2)}\right), \tag{2.1}$$

where $\mathbf{X} = (X_1, \ldots, X_N)^\top$, and $\mathbf{X}^{(t_1:t_2)}$ denotes the set of variables $\left\{\mathbf{X}^{(t)} : t \in [t_1, t_2]\right\}$, $t_1 < t_2$.

To alleviate the computational burden associated with representing the joint distribution in Equation (2.1), a DBN typically relies on a set of simplifying assumptions. First, we consider a time discretization, which involves partitioning the time horizon $[0, T]$ into time slices $t = 0, \ldots, T$, with an arbitrary step Δt. In this work, we assume regularly spaced time steps with granularity equal to the time between two consecutive observations of the system. Second, we assume that the conditional dependencies between consecutive random variables are represented by a Markov process. Accordingly, each

https://doi.org/10.1017/dce.2026.10072 Published online by Cambridge University Press

variable $X_i^{(t)}$ can have parents only in the current or immediately preceding time slice. This conditional independence structure allows us to compactly write $p\left(\mathbf{X}^{(0:T)}\right)$ as

$$p\left(\mathbf{X}^{(0:T)}\right) = p\left(\mathbf{X}^{(0)}\right) \prod_{t=0}^{T-1} p\left(\mathbf{X}^{(t+1)} | \mathbf{X}^{(t)}\right). \tag{2.2}$$

Hence, the probability distribution over different trajectories is fully characterized by the initial state distribution $p\left(\mathbf{X}^{(0)}\right)$ and the transition model $p\left(\mathbf{X}^{(t+1)} | \mathbf{X}^{(t)}\right)$. This implies that probabilities at time $t+1$ are conditionally independent of all states prior to time t, given the state at time t.

Typically, a third assumption concerns the stationarity of the underlying Markov process, which allows the conditional dependencies between consecutive random variables to be represented by an invariant transition model:

$$p\left(\mathbf{X}^{(t+1)} | \mathbf{X}^{(t)}\right) = p(\mathbf{X}' | \mathbf{X}) \quad \forall t, \quad \mathbf{X}, \mathbf{X}' \subseteq \mathcal{X}. \tag{2.3}$$

As detailed in the following, in this work, we relax this assumption, resulting in adaptive DTs capable of tracking systems that evolve according to a non-stationary process.

The transition model in Equation (2.3) can be represented by a two-time-slice Bayesian network (2-TBN) for a process over $\mathcal{X}$, which encodes the following CPD:

$$p(\mathbf{X}' | \mathbf{X}) = \prod_{i=1}^{N} p\left(X_i' | \mathrm{Pa}_{X_x'}\right), \tag{2.4}$$

where, in addition to $\mathbf{X}'$, the 2-TBN includes only those variables in $\mathbf{X}$ with a direct effect on $\mathbf{X}'$.

Given a Bayesian network $\mathcal{B}_0$ representing the initial distribution $p\left(\mathbf{X}^{(0)}\right)$, a DBN over the set of template variables $\mathcal{X}$ is defined by a pair $\langle \mathcal{B}_0, \mathcal{B}_{\rightarrow} \rangle$, where $\mathcal{B}_{\rightarrow}$ is a 2-TBN for the process over $\mathcal{X}$. The graph topology and conditional dependencies for $t = 0$ follow those in $\mathcal{B}_0$, while for $t > 0$ they follow those in $\mathcal{B}_{\rightarrow}$. The resulting DBN provides a single compact model that captures the system dynamics and yields distributions over different trajectories. It can be interpreted as a sequence of Bayesian networks, one for each time slice.

For any desired time horizon $T > 0$, the distribution over $\mathbf{X}^{(0:T)}$ is obtained by unrolling the DBN until T, repeating the core structure of the graph at each time step. Given a sequence of observations, a DBN can be unrolled until it fully encompasses the observed data. Additional time slices beyond the last observation do not affect inference within the observation period. Once unrolled, the DBN becomes a standard Bayesian network, and beliefs can be updated and propagated as for static Bayesian networks.

2.2. Probabilistic graphical model for predictive structural digital twins

The PGM used to represent the dynamic interaction between the physical and virtual domains is depicted in Figure 2 as a dynamic decision network, that is, a DBN with decision nodes. In the graphical representation, circle nodes denote random variables, square nodes indicate actionable inputs, and diamond-shaped nodes symbolize the objective function, all at discrete time steps. Each time the DT is updated by assimilating observational data, the PGM advances of one time step, with $t \in \{0, \ldots, T\}$. Here, $t = 0$ marks the time at which the DT enters operation, and $t = T$ represents its lifetime horizon. Nodes with bold outlines represent observed quantities, while nodes with thin outlines represent unobserved quantities to be inferred. The PGM is sparsely connected, with edges encoding known or assumed conditional dependencies between variables.

We use capital letters to denote random variables representing the quantities involved in our abstraction, with subscripts indicating the time index. Corresponding lowercase letters represent realizations of these random variables. Calligraphic script variables denote the space of possible values that these quantities can take. For instance, the unknown physical state, which reflects the ground truth variability of interest, is denoted as $S_t \sim p(s_t)$, where s_t represents the realization of the random variable S_t at time t, and $p(s_t)$ represents the probability that $S_t = s_t$.

https://doi.org/10.1017/dce.2026.10072 Published online by Cambridge University Press

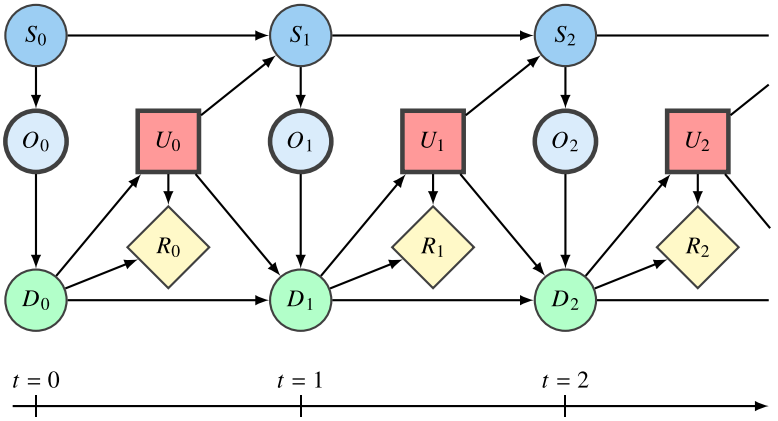


Figure 2. *Dynamic Bayesian network encoding the asset–twin dynamical system, unrolled over two time slices. Circle nodes denote random variables, square nodes denote actions, and diamond nodes denote the objective function. Nodes with bold outlines denote observed quantities, while nodes with thin outlines denote unobserved quantities to be inferred. Directed edges represent the variables' dependencies encoded via conditional probability distributions.*

The digital state $D_t \sim p(d_t)$ aims to capture the variability in the physical state. It may include various types of information, such as initial conditions, boundary conditions, material properties, and other relevant quantities describing the asset (or process) of interest. However, a DT does not need to be an exact virtual replica to provide maximum value; rather, it should capture the essential features and variations relevant for diagnosis, prediction, and decision-making (Ferrari and Willcox, 2024). Although there are no strict constraints in this regard, we define the digital state as a categorical random variable that can assume values among a set of predefined configurations in the digital state space $\mathcal{D}$.

The physical-to-digital information flow from S_t to D_t is governed by the continuous collection and assimilation of observational data $O_t \sim p(o_t)$. These may include sensor recordings, inspection data, or diagnostic information. Since the physical state S_t is only partially and indirectly observable, the digital state D_t cannot be uniquely determined. Instead, it provides evidence for digital state configurations that could be consistent with the given observations.

The updated digital state D_t informs the digital-to-physical information flow, that is, the control actions. In Figure 2, $U_t \sim p(u_t)$ is a decision variable representing the action to be taken. The space of actions $\mathcal{U}$ may include interventions affecting the physical state, adjustments to the operational conditions, or modifications to the observational process. Each action is characterized by its own transition dynamics across the system's state space.

Finally, the reward node $R_t \sim p(r_t)$ measures the performance of the asset–twin system. Specifically, the reward function provides values within a reward space $\mathcal{R}$, measuring the expected utility of DBN trajectories that we maximize when solving the planning problem induced by the PGM.

Since we assume discrete spaces for digital states and control actions, the related causal relationships are modeled using conditional probability tables (CPTs). In a CPT, each column contains the CPD of a node's value given the values of its parent nodes. Each column sums to one, as the entries represent an exhaustive set of possible outcomes for the variable. A node with no parents has only one column, indicating the prior probabilities of each possible value. CPTs can be viewed as a set of parameters representing prior knowledge about random variables. Thanks to the discrete DBN setting, we can propagate and update beliefs exactly with a single pass of the sum-product message-passing algorithm (Pearl, 1982).

The Bayesian network $\mathcal{B}_0$ and the 2-TBN $\mathcal{B}_{\rightarrow}$, jointly forming our PGM, are shown in Figure 3. $\mathcal{B}_0$ encodes the initial distribution of the random variables, which can be derived from expert knowledge to

https://doi.org/10.1017/dce.2026.10072 Published online by Cambridge University Press

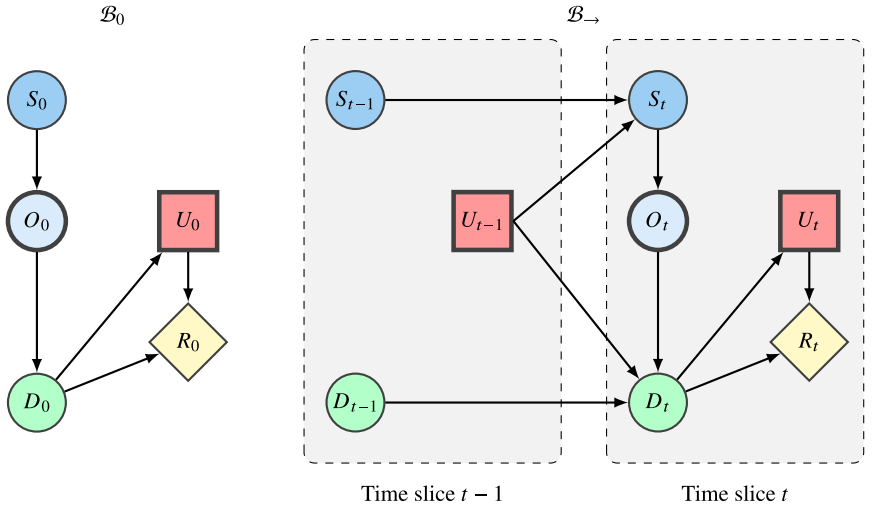


Figure 3. *Dynamic Bayesian network encoding the asset–twin dynamical system: Bayesian network representing the initial distribution $\mathcal{B}_0$; two-time-slice Bayesian network for the process $\mathcal{B}_\rightarrow$. Circle nodes denote random variables, square nodes denote actions, and diamond nodes denote the objective function. Nodes with bold outlines denote observed quantities, while nodes with thin outlines denote unobserved quantities to be inferred. Directed edges represent the variables' dependencies encoded via conditional probability distributions.*

assign a prior distribution to the initial state of the asset–twin system. In contrast, $\mathcal{B}_\rightarrow$ encodes the inter-slice conditional dependencies, modeling the temporal relationships between the random variables. The PGM enables the extraction of the joint distribution of the network's trajectories, conditioned on observed data $O_t = o_t$ and enacted actions $U_t = u_t$. Specifically, the joint belief state can be factorized from the initial time step $t = 0$ up to the current time t_c, using the following sequential Bayesian inference formulation:

$$p(D_{0:t_c}, U_{0:t_c}, R_{0:t_c} | o_{0:t_c}, u_{0:t_c}) \propto \prod_{t=0}^{t_c} \left[\phi_t^{\text{update}} \phi_t^{\text{control}} \phi_t^{\text{reward}} \right], \tag{2.5}$$

with factors:

$$\phi_t^{\text{update}} = p(D_t | D_{t-1}, U_{t-1} = u_{t-1}, O_t = o_t), \tag{2.6}$$

$$\phi_t^{\text{control}} = p(U_t | D_t), \tag{2.7}$$

$$\phi_t^{\text{reward}} = p(R_t | D_t, U_t = u_t). \tag{2.8}$$

The factor ϕ_t^{update} updates the digital state belief D_t to reflect the (hidden) physical state S_t. First, it accounts for a control-dependent transition model, describing the expected evolution of the digital state given the previous belief D_{t-1} and the most recent action taken $U_{t-1} = u_{t-1}$. Additionally, it integrates this prior with the data assimilation outcome by conditioning on $O_t = o_t$. We can explicitly split the contributions of transition dynamics and data assimilation by further factorizing ϕ_t^{update} into a predictor–corrector form update, as follows:

$$\phi_t^{\text{update}} \propto \phi_t^{\text{data}} \phi_t^{\text{history}}, \tag{2.9}$$

https://doi.org/10.1017/dce.2026.10072 Published online by Cambridge University Press

where:

$$\phi_t^{\text{history}} = p(D_t | D_{t-1}, U_{t-1} = u_{t-1}), \tag{2.10}$$

$$\phi_t^{\text{data}} = p(O_t = o_t | D_t). \tag{2.11}$$

The causal structure required for this factorization is not included in the set of conditional dependencies induced by the graph depicted in Figure 2. Graphically, this would require an arrow from D_t to O_t, but in our model, the arrow is depicted in the opposite direction to reflect the information flow within the DT, consistent with MG Kapteyn et al., (2021) and Torzoni et al., (2024). For further details, the reader may refer to Koller and Friedman, (2009). The transition model $\phi_t^{\text{history}} : \mathcal{D} \times \mathcal{D} \times \mathcal{U} \mapsto [0, 1]$ acts as a forward-time predictor. It can be visualized as a CPT describing the expected transition dynamics from D_{t-1} to D_t for each possible action $U_{t-1} = u_{t-1}$. This CPT should incorporate any prior knowledge the DT designer has regarding the asset and relevant operational conditions. It can be derived offline from historical data and accumulated experience (Thompson and Johnson, 2005; Morcous, 2006; Bianchi and Biondini, 2022), as well as from physics-based or empirical evolution models (Biondini and Frangopol, 2016, 2017; Garayalde et al., 2025). In this paper, we introduce a stochastic parametrization of ϕ_t^{history}, similar to Tezzele et al., (2026), enabling online learning of state transition beliefs via Bayesian updates, as described in Section 3.1. In contrast, the factor ϕ_t^{data} is associated with the assimilation of observational data $O_t = o_t$, as further described at the end of this section.

The factor $\phi_t^{\text{control}} : \mathcal{U} \times \mathcal{D} \mapsto [0, 1]$ evaluates the belief about the next optimal control action. This factor corresponds to a CPT encoding a control policy $\pi : \mathcal{D} \mapsto \mathcal{U}$. The actual action to be taken is selected as the best point estimate $u_t = \arg\max_{u \in \mathcal{U}} \pi(D_t)$, that is, the conversion from probabilistic to deterministic control. The computation of π is described in Section 3.2.

Finally, $\phi^{\text{reward}} : \mathcal{D} \times \mathcal{U} \to \mathcal{R}$ encapsulates the evaluation of the objective function. The two edges leading to R_t represent a reward function that quantifies the costs associated with the digital state and control actions, respectively.

Given the current digital state D_{t_c}, probabilistic estimates of future digital states and control actions are made by unrolling a simplified DBN up to a prediction time $t_p > t_c$. This prediction DBN, depicted in Figure 4, retains a structure similar to the one discussed above but is stripped of the nodes related to observations, data assimilation, and action execution. At the prediction stage, information from sensing and performed actions is no longer available. Accordingly, the actions U_t are depicted as random variables rather than decision nodes, as they represent predictions of what-if scenarios beyond data assimilation, made through uncertainty propagation via DBN unrolling. As a result, the factorization over the trajectories in Equation (2.5) can be extended over the prediction time steps $t = t_c + 1, \ldots, t_p$ by augmenting the belief state as follows:

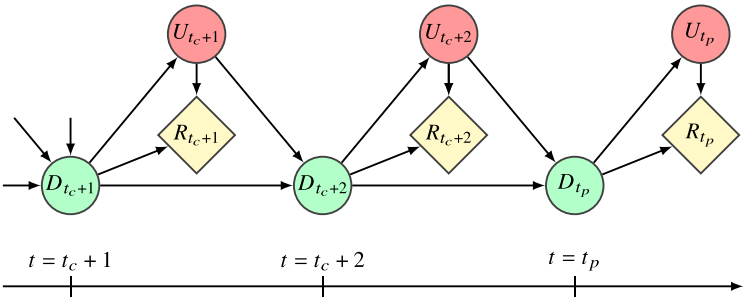


Figure 4. *Dynamic Bayesian network used to predict the future evolution of digital states and control actions. Circle nodes denote random variables, and diamond nodes denote the objective function. Directed edges represent the variables' dependencies encoded via conditional probability distributions.*

https://doi.org/10.1017/dce.2026.10072 Published online by Cambridge University Press

$$p\left(D_{t_c+1:t_p}, U_{t_c+1:t_p}, R_{t_c+1:t_p} | D_{t_c}, U_{t_c}\right) = \prod_{t=t_c+1}^{t_p} \left[\phi_t^{\text{history}} \phi_t^{\text{control}} \phi_t^{\text{reward}}\right]. \tag{2.12}$$

In this case, the factors $\phi_t^{\text{history}} = p(D_t | D_{t-1}, U_{t-1})$ and $\phi_t^{\text{reward}} = p(R_t | D_t, U_t)$ enable future state predictions and probabilistic reward estimates, respectively, conditioned on predicted control actions that have not yet been enacted. As clarified in Section 3.2, factors ϕ_t^{history} and ϕ_t^{control} actually encode dynamically updated transition models and control policies, respectively. Indeed, after updating the transition model ϕ_t^{history} with the evidence collected up to t_c, we also recompute an optimal policy π_{t_c}, such that $U_{t_c:t_p} \sim \pi_{t_c}\left(D_{t_c:t_p}\right)$.

In this work, the asset of interest is a deteriorating structural system, although the framework outlined above can be adapted to other types of systems or applications beyond SHM. The asset is dynamic, meaning its state evolves over time according to laws that depend on both the system itself and external factors. These external factors can either improve or degrade the condition of the asset and influence its deterioration rate. Such factors may include long-term degradation caused by chemical, physical, or mechanical processes, as well as sudden changes due to discrete damage events or interventions by external agents (Zakić et al., 1991).

We simulate the monitored asset using a physics-based computational model. Specifically, the structure is modeled as a linear-elastic continuum, and its dynamic response to applied loads is described through elasto-dynamics equations, assuming linearized kinematics. The model is spatially discretized using finite elements, and its solution is advanced over time to produce synthetic observational data. This computational model serves as a data generator, controlled by a vector of N_{par} parameters $\boldsymbol{\mu} \in \mathbb{R}^{N_{\text{par}}}$, which describe the operational, damage, and (possibly) environmental conditions. Further details on the modeling choices for the Hörnefors railway bridge case study considered in this work will be provided in Section 4.

The observational data consist of vibration recordings shaped as multivariate time series $\mathbf{U}(\boldsymbol{\mu}) = [\mathbf{u}_1(\boldsymbol{\mu}), \ldots, \mathbf{u}_{N_s}(\boldsymbol{\mu})] \in \mathbb{R}^{L \times N_s}$, made up of N_s time series, each containing L sensor measurements equally spaced in time, such as accelerations or displacements. Accordingly, the observation space $\mathcal{O}$ is described through the sensitivity of $\mathbf{U}$ with respect to $\boldsymbol{\mu}$. For the problem settings we consider, each observation process refers to a time interval that is short enough for the operational, environmental, and damage conditions to be assumed time-invariant, yet long enough to allow accurate identification of the structural behavior.

As the unknown physical state reflects the ground truth variability in the structural health of the asset, the categorical random variable modeling the digital state describes the presence, location, and magnitude of damage among a set of predefined configurations in $\mathcal{D}$. We describe changes in the structural dynamic response due to damage through a localized reduction of effective stiffness. If the time duration of the observation process is small compared to the time required for damage progression, this approach allows the structural behavior to be treated as linear within an observation interval, thus enabling a timescale separation between damage evolution and health assessment (Eftekhar Azam and Mariani, 2018). Examples of damage patterns that can be represented in this way include the failure of bolted connections and corrosion in steel structures (Pandey and Biswas, 1994; Mariniello et al., 2021), cracking and delamination in composite structures (Kapteyn et al., 2020), cracking in reinforced concrete structures (Teughels et al., 2002; Mariniello et al., 2021), cracking and crushing-induced stiffness losses in masonry brick structures (Hernández-González and García-Macías, 2024), cuts in steel beams (Seventekidis et al., 2020), and the erosion of piles and abrasion of deck surfaces (Fernandez-Navamuel et al., 2022). Here, the local stiffness reduction shaping $\mathcal{D}$ is achieved by means of two variables, $y \in \mathbb{N}$ and $\delta \in \mathbb{R}$, both included in $\boldsymbol{\mu}$. Specifically, $y \in \mathbb{N}$ labels the specific damage region among a set of $N_\Omega \in \mathbb{N}$ predefined damageable regions, where $y = 0$ denotes the undamaged baseline, while $\delta \in \mathbb{R}$ represents the magnitude of the stiffness reduction occurring in the region associated with y.

Data assimilation, represented by the ϕ_t^{data} factor, is performed using DL models to approximate the inverse mapping from measurements $\mathbf{U}$ to the target parameters y and δ. The deterministic output of the DL models is then converted into evidence for configurations in the digital state space that are consistent with the given observation, thereby serving as the likelihood mechanism informing the posterior.

https://doi.org/10.1017/dce.2026.10072 Published online by Cambridge University Press

This conversion is carried out through a confusion matrix that measures the offline (expected) performance of the DL models in correctly identifying the digital state among the predefined configurations in $\mathcal{D}$. The confusion matrix can be interpreted as a CPT, where rows represent the ground truth digital state, and columns represent the predicted digital state. For a more detailed description of the evaluation of the ϕ_t^{data}, the reader is referred to (Torzoni et al., 2024).

3. Online learning of transition dynamics

This describes the proposed adaptive capabilities within the context of parametric MDPs. Specifically, Section 3.1 reviews MDPs to formalize the problem of sequential decision-making under uncertainty and then discusses the online learning of transition dynamics. Section 3.2 explores the use of RL to compute control policies based on dynamically inferred transition models.

3.1. Parametric Markov decision processes

In a decision-making setting, an agent must choose from a set of possible actions, each of which may have uncertain effects on the system state. The goal of the decision-making process is to maximize the expected utilities associated with the outcomes of these actions. Sequential decision problems expand on this concept by requiring the agent to choose a sequence of actions that optimizes the long-term accumulation of rewards over the planning horizon.

MDPs provide a framework for modeling sequential decision problems in stochastic environments, using Markov transition models and additive rewards (Bellman, 1957b; Howard, 1960). A discrete MDP is defined as a 4-tuple $\langle \mathcal{S}, \mathcal{U}, \mathcal{P}, \mathcal{R} \rangle$, consisting of a set of states $\mathcal{S}$, a set of actions $\mathcal{U}$, a CPT $\mathcal{P} \in \mathbb{R}^{\mathcal{D}\times\mathcal{D}\times\mathcal{U}}$ describing the Markov transition model, and a set of rewards $\mathcal{R}$.

The control action is selected as the best point estimate, thereby mapping the probabilistic state representation to a deterministic decision. This is consistent with the standard MDP assumption of full observability. While alternative approaches account for partial observability (e.g., (Arcieri et al., 2023, 2024; Papakonstantinou et al., 2016), they require more complex and computationally demanding solution strategies. In the present work, we rely on the high accuracy of the deep learning models used for digital state classification, which provide a near-complete characterization of the system state. On this basis, we approximate the true state with its most likely estimate, effectively decoupling state estimation from control. We note, however, that this simplification neglects residual uncertainty and may lead to suboptimal decisions when classification uncertainty is non-negligible.

In our DT framework, the MDP associated with optimal control planning is encoded by the prediction DBN shown in Figure 4. While in (Torzoni et al., 2024) the transition model ϕ_t^{history} was kept fixed, we relax the assumption of a stationary Markov process for the conditional dependencies between random variables in consecutive time slices. Specifically, we treat the transition probabilities in $\mathcal{P}$ as random variables with conjugate distributions, similar to (Tezzele et al., 2026). Starting with an initial prior, the transition dynamics is learned online using Bayesian updates that incorporate new evidence about the system response to actions. This approach is scalable and allows (*i*) adapting DTs starting from (possibly, inaccurate) initial priors or (*ii*) tracking non-stationary, evolving contexts.

Actions with uncertain effects produce non-deterministic (and possibly unexpected) outcomes. Within an SHM context, inter-state transitions can be described using (stochastic) transition matrices $\mathcal{P}_u \in \mathbb{R}^{\mathcal{D}\times\mathcal{D}}$, one for each action $u \in \mathcal{U}$:

$$\mathcal{P}_u = \begin{pmatrix} p(d_1|d_1,u) & p(d_1|d_2,u) & \dots & p\big(d_1|d_{|\mathcal{D}|},u\big) \\ p(d_2|d_1,u) & p(d_2|d_2,u) & \dots & p\big(d_2|d_{|\mathcal{D}|},u\big) \\ \vdots & \vdots & \ddots & \vdots \\ p\big(d_{|\mathcal{D}|}|d_1,u\big) & p\big(d_{|\mathcal{D}|}|d_2,u\big) & \dots & p\big(d_{|\mathcal{D}|}|d_{|\mathcal{D}|},u\big) \end{pmatrix}, \tag{3.1}$$

https://doi.org/10.1017/dce.2026.10072 Published online by Cambridge University Press

where $|\mathcal{A}| \in \mathbb{N}$ denotes the cardinality of a generic set $\mathcal{A}$. Matrix $\mathcal{P}_u$ collects the probabilities $p(d' \mid d,u)$ of transitioning from state $d \in \mathcal{D}$ to state $d' \in \mathcal{D}$, under a given action $u \in \mathcal{U}$. Here, columns represent the starting state d, and rows represent the arrival state d'. Each column of $\mathcal{P}_u$ sums to one, reflecting the exhaustive set of transition probabilities from a given state d under the action u. By partitioning $\mathcal{D}$ in lexicographic order, with states ordered first by damage location and then by damage level, the diagonal entries in $\mathcal{P}_u$ denote the probability that the system has of remaining in the same state, while the lower-left and upper-right triangles represent the probabilities of system deterioration and improvement, respectively. The CPT describing the complete transition model, tensorial with respect to all available actions, is given by $\mathcal{P} = \left\{\mathcal{P}_1 \ldots, \mathcal{P}_{|\mathcal{U}|}\right\}$.

The specific content of the transition matrices should account for the structural system under consideration, its operational conditions, and the surrounding environmental variability (Torzoni et al., 2022). To address these sources of uncertainty, we adopt a hierarchical Dirichlet–Multinomial parametrization. For each state–action pair, the categorical transition distribution is endowed with a Dirichlet prior, whose hyperparameters encode prior beliefs over transition outcomes. This defines a hierarchical Bayesian structure in which transition probabilities are treated as latent random variables, and uncertainty is propagated through the higher-level Dirichlet prior, capturing both epistemic uncertainty and data-driven adaptation (Gelman et al., 2013). These parameters are then subject to temporal updating, that is, learned online as new data arrives, enabling adaptive Bayesian inference.

We model the transition probabilities $p(d'|d,u)$ with categorical distributions parameterized by $\boldsymbol{\theta}_{d,u} = \left(\theta^1_{d,u}, \ldots, \theta^{|\mathcal{D}|}_{d,u}\right)^\top \in \mathbb{R}^{|\mathcal{D}|}$, where $\sum_{k=1}^{|\mathcal{D}|} \theta^k_{d,u} = 1$ and $k = 1, \ldots, |\mathcal{D}|$ indexes the arrival states d'. The uncertainty over $\boldsymbol{\theta}_{d,u}$ is represented by a Dirichlet prior:

$$\boldsymbol{\theta}_{d,u} \sim \text{Dirichlet}(\boldsymbol{\alpha}_{d,u}), \tag{3.2}$$

where $\boldsymbol{\alpha}_{d,u} = \left(\alpha^1_{d,u}, \ldots, \alpha^{|\mathcal{D}|}_{d,u}\right)^\top \in \mathbb{R}^{|\mathcal{D}|}$ are (positive) concentration parameters encoding prior beliefs about the transition probabilities. We use the Dirichlet distribution because it is conjugate to the Multinomial likelihood, ensuring that the posterior distribution remains within the Dirichlet family after incorporating observations. This property allows efficient refinement of the transition probabilities and beliefs by updating the Dirichlet hyperpriors in closed form.

Managing $|\mathcal{D}| \times |\mathcal{U}|$ Dirichlet distributions, each with $|\mathcal{D}|$ parameters, can become computationally intractable for large state spaces. While we provide the formulation and update scheme for such a state-dependent transition model in A, a more computationally viable approach to handling multi-step transitions is to assign equal transition probabilities between consecutive elements of the partition, rather than a separate prior for each individual state transition (Ross et al., 2011; Papakonstantinou et al., 2016; Andriotis and Papakonstantinou, 2019; Tezzele et al., 2026). In other words, we assign probabilities of remaining in the same state, of transitioning by one step, two steps, and so on, independently of the starting state d. This state-independent transition model is thus parametrized through $\mathbf{p}_u = \left(p^0_u, p^1_u, \ldots, p^I_u\right)^\top \in \mathbb{R}^{I+1}$, representing transition probabilities of $0, 1, \ldots, I$ steps forward under action u. The corresponding transition matrix $\mathcal{P}_u$ is then constructed from $\mathbf{p}_u$, ensuring the same column structure for all starting states, except for absorbing states—that is, states from which no further transitions occur, meaning they retain probability 1 of remaining in the same state.

As an example, for a state space of cardinality $|\mathcal{D}| = 5$ with a single damage location, the transition matrix for an action under which the structural health state can deteriorate by up to $I = 2$ steps is:

$$\mathcal{P}_u(\mathbf{p}_u) = \begin{pmatrix} p^0_u & 0 & 0 & 0 & 0 \\ p^1_u & p^0_u & 0 & 0 & 0 \\ p^2_u & p^1_u & p^0_u & 0 & 0 \\ 0 & p^2_u & p^1_u & p^0_u & 0 \\ 0 & 0 & p^2_u & p^1_u + p^2_u & 1 \end{pmatrix}. \tag{3.3}$$

As per Equation (3.2), the probabilities $\mathbf{p}_u$ are assigned a Dirichlet prior:

https://doi.org/10.1017/dce.2026.10072 Published online by Cambridge University Press

$$\mathbf{p}_u \sim \text{Dirichlet}(\boldsymbol{\alpha}_u), \tag{3.4}$$

where $\boldsymbol{\alpha}_u = \left(\alpha_u^0, \alpha_u^1, \ldots, \alpha_u^I\right)^\top \in \mathbb{R}^{I+1}$ are the concentration parameters for the step transitions under action u.

Let $d_{0:t_c}$ and $u_{0:t_c-1}$ denote the history of system states and actions observed up to the current time t_c, respectively. Additionally, let $\mathbf{y}_u = \left(N_u^0, N_u^1, \ldots, N_u^I\right)^\top \in \mathbb{N}^{I+1}$ represent the counts of observed step transitions under $\{d_{0:t_c}, u_{0:t_c-1}\}$. The likelihood of $\mathbf{y}_u$, given the transition probabilities $\mathbf{p}_u$, is modeled as a Multinomial distribution:

$$\mathbf{y}_u \mid \mathbf{p}_u \sim \text{Multinomial}(N_u, \mathbf{p}_u), \tag{3.5}$$

where $N_u = \sum_{i=0}^{I} N_u^i$ is the total transition count observed under action u. Using Bayes' theorem, the posterior distribution of $\mathbf{p}_u$ after observing $\mathbf{y}_u$ remains in the Dirichlet family due to conjugacy:

$$\mathbf{p}_u \mid \mathbf{y}_u \sim \text{Dirichlet}(\overline{\boldsymbol{\alpha}}_u), \tag{3.6}$$

with updated concentration parameters $\overline{\boldsymbol{\alpha}}_u \in \mathbb{R}^{I+1}$, computed as:

$$\overline{\boldsymbol{\alpha}}_u = \left(\alpha_u^0 + N_u^0, \alpha_u^1 + N_u^1, \ldots, \alpha_u^I + N_u^I\right)^\top. \tag{3.7}$$

The parametrized Dirichlet–Multinomial transition model efficiently accounts for the possibility of multi-step transitions. However, it does not address the possibility of inconsistencies arising from evidence that contradicts the transition model. This limitation can be particularly critical, as assigning a zero probability to potential state transitions could compromise the inference process (Lindley, 2013). To resolve this potential inconsistency, a perturbation matrix $\Xi \in \mathbb{R}^{|\mathcal{D}| \times |\mathcal{D}|}$, with positive entries, arbitrarily small relative to the non-zero terms of $\mathcal{P}_u$, is introduced. This matrix is used to obtain a perturbed transition matrix as follows:

$$\mathcal{P}_u^{\Xi}(\mathbf{p}_u) = \mathcal{P}_u(\mathbf{p}_u) + \Xi, \tag{3.8}$$

which is then properly normalized.

A special case of the state-independent Dirichlet–Multinomial transition model is the Beta-Binomial model, which is parameterized by a single one-step transition probability. The formal derivation follows an approach analogous to the one presented in this section and is provided in B for completeness. The numerical experiments in Section 4 use Dirichlet–Multinomial state-independent transition models.

3.2. Dynamic control policies and precision updates

The ϕ_t^{control} factor in Equation (2.5) represents a CPT encoding a health-aware control policy $\pi_t(D_t)$. In (Torzoni et al., 2024), a stationary policy was computed offline for an infinite planning horizon, based on the simplifying assumption of known transition dynamics. Here, the sequence of actions $u_{t_c:T}^*$ defining the optimal control policy π^* is systematically recomputed using MBRL with a dynamically updated transition model. The policy is recomputed over the finite planning horizon $[t_c, T]$, representing the design lifetime of the asset. The corresponding optimization problem is expressed as:

$$\pi_t^*(D_t) = \arg\max_{\pi_t} \mathbb{E}\left[\sum_{t=t_c}^{T} R_t\right]. \tag{3.9}$$

The subscript t in π_t^* denotes a non-stationary control policy, as the optimal action depends on the remaining time before the end of the planning horizon, as well as the time t_c during which the transition dynamics is updated, and the control policy is recomputed. The additive reward function to be maximized is defined as follows:

$$R(d,u) = R^{\text{health}}(d) + \xi R^{\text{control}}(u), \quad \forall d \in \mathcal{D}, \forall u \in \mathcal{U}, \tag{3.10}$$

where $R^{\text{health}}(d)$ and $R^{\text{control}}(u)$ quantify the rewards associated with the health state of the system and the action taken, respectively. In general, these reward values may represent real-world costs or abstract values, tuned using the weighting factor $\xi \in \mathbb{R}$ to achieve the desired system performance.

One way to characterize an MDP is to consider the expected utility associated with a policy π when starting in a certain state and following π thereafter. For a finite horizon T, the state-value function

https://doi.org/10.1017/dce.2026.10072 Published online by Cambridge University Press

$V^\pi : \mathcal{D} \to \mathbb{R}$ quantifies, for every state and starting time t, the expected cumulative reward when starting at d_t and following policy π, as follows (Sutton and Barto, 2018):

$$V^\pi(d_t) = \mathbb{E}_\pi \left[\sum_{\tau=t}^{T} R_\tau \middle| d_t \right], \tag{3.11}$$

where $\mathbb{E}_\pi[\cdot]$ denotes the expected value, given that the agent follows policy π. For any control policy π and for each possible state d_t, with $t < T$, the state-value function satisfies the following recursive relationship, known as the Bellman equation for V^π:

$$V^\pi(d_t) = \sum_{u \in \mathcal{U}} \pi(u|d_t) \sum_{d_{t+1} \in \mathcal{D}} p(d_{t+1}|d_t, u)(r_t + V^\pi(d_{t+1})), \tag{3.12}$$

where r_t is the immediate reward at time t, and $p(d_{t+1}|d_t, u_t) \in \mathcal{P}$ are the transition probabilities. According to the principle of maximum expected utility, for each time $t < T$, there exists a time-dependent optimal policy π_t^* that maximizes expected utility. The time-dependent optimal action at state d_t is given as:

$$\pi_t^*(d_t) = \arg\max_{u \in \mathcal{U}} \left(\sum_{d_{t+1} \in \mathcal{D}} p(d_{t+1}|d_t, u)(r_t + V^*(d_{t+1})) \right), \tag{3.13}$$

where $V^*(d_{t+1})$ measures the utility associated with d_{t+1} when following π^*. The Bellman optimality equation for π^* can thus be written as:

$$V^{\pi*}(d_t) = \max_{u \in \mathcal{U}} \left(\sum_{d_{t+1} \in \mathcal{D}} p(d_{t+1}|d_t, u)(r_t + V^*(d_{t+1})) \right), \tag{3.14}$$

yielding a system of $|\mathcal{D}|$ recursions in $|\mathcal{D}|$ unknowns. The optimal state-value functions are the unique solutions to these equations, from which the corresponding optimal policy can be directly derived. However, the optimal action u_t^* is selected through the max operator, which introduces nonlinearity in the system of Bellman equations.

An approximation of the unique solution to Equation (3.14) is computed using finite-horizon dynamic programming. Dynamic programming is an algorithmic paradigm introduced by Bellman (1957a) that simplifies the optimization by decomposing it into smaller sub-problems that are solved recursively. Among these methods, we employ the value iteration algorithm, also known as backward induction, as it proceeds by solving the problem backward in time, from $t = T$ down to $t = t_c$.

Algorithm 1. Preliminary offline phase.

Input: parametrization of the physics-based models
training options for the reduced-order model
training options for the deep learning models
1: set up the physics-based numerical model of the structure
2: assemble snapshot matrix of full-order solutions
3: find reduced basis functions via singular value decomposition of the snapshots matrix
4: use the reduced-order model to populate the training dataset with recordings at sensor locations
5: train and validate the deep learning models for structural health estimation
6: compute the confusion matrix associated with ϕ_t^{data}
7: set the prior distributions over the transition probabilities defining ϕ_t^{history}
8: set the reward function encoding ϕ_t^{reward}
return trained deep learning models for state estimation
topology and factors for the online phase DBN
topology and factors for the prediction DBN

https://doi.org/10.1017/dce.2026.10072 Published online by Cambridge University Press

Algorithm 2. Online phase.

input: topology and factors for the online phase DBN
topology and factors for the prediction DBN
trained deep learning models for state estimation
observational data $O_{t_c} = o_{t_c}$
history of observed system states and actions $\{d_{0:t_c}, u_{0:t_c-1}\}$
1: $t_c \leftarrow t_c + 1$
2: assimilate o_{t_c} with deep learning models
3: infer D_{t_c}, given D_{t_c-1} and u_{t_c-1}
4: adapt transition matrices $\mathcal{P}_u(\mathbf{p}_u)$ by updating $p(\mathbf{p}_u|\{d_{0:t_c}, u_{0:t_c-1}\})$
5: **optional**: choose the statistic parametrizing the transition model to adjust risk aversion
6: recompute control policy $\pi = (\pi_{t_c}, \dots, \pi_T)$ under the updated transition model
7: forecast $D_{t_c+1:t_p}$ and $U_{t_c+1:t_p}$ using the updated transition model and control policy
8: select u_{t_c} as the best point estimate of U_{t_c}
return control action to be executed $U_{t_c} = u_{t_c}$
expected evolution $D_{t_c+1:t_p}$
expected evolution $U_{t_c+1:t_p}$
updated transition model CPT $\mathcal{P}$
updated control policy $\pi \approx \pi^*$

An algorithmic description of the offline and online phases of the proposed computational procedure is reported in Algorithm 1 and Algorithm 2, respectively.

Remark 3.1. For each action, the transition matrix $\mathcal{P}_u$ is parametrized by a statistic of the inferred transition posteriors leading to different control policies. The parametrizing statistic is directly connected to the risk aversion of the decision-maker, which involves balancing a trade-off between maximizing profits and ensuring structural safety. In the numerical experiments presented in Section 4, the transition matrices are parameterized using the maximum a posteriori estimate of the inferred Dirichlet posterior. Alternative choices could be adopted to induce different levels of risk aversion. For univariate probability distributions it is possible to incorporate different risk measures as proposed in (Tezzele et al., 2026). For the Dirichlet distribution, due to its multivariate nature and its support on simplexes, it is not possible to extend commonly used risk measures such as the value at risk (Morgan, 1996) or the conditional value at risk (Rockafellar and Uryasev, 2000) without using an intermediate scalar function.

4. Numerical experiments

This section assesses the proposed methodology on the simulated monitoring, management, and maintenance planning of the Hörnefors railway bridge.

4.1. Physical asset

The Hörnefors railway bridge, shown in Figure 5, is an integral concrete bridge located along the Swedish Bothnia line. The bridge features a span of 15.7m, a free height of 4.7m, and a width of 5.9m (edge beams excluded). The structural elements have a thickness of 0.5m for the deck, 0.7m for the frame walls, and 0.8m for the wing walls. The bridge is founded on two plates, connected by two stay beams and supported by pile groups. The concrete is of class C35/45, with mechanical properties: $E = 34\text{GPa}$, $\nu = 0.2$, $\rho = 2500\text{kg/m}^3$. The superstructure consists of a single track with sleepers spaced 0.65m apart, resting on a 0.6m deep, 4.3m wide ballast layer with a density of $\rho_B = 1800\text{kg/m}^3$. The bridge is subjected to the transit of *Gröna Tåget* trains traveling at speeds ranging from $v \in [160,215]\text{km/h}$. We specifically consider trains composed of two wagons, totaling 8 axles, with each axle carrying a mass of $\psi \in [16,22]\text{ton}$. The geometrical and mechanical data, as well as the traveling load model, are adapted from (Rosafalco et al., 2021; Torzoni et al., 2024).

https://doi.org/10.1017/dce.2026.10072 Published online by Cambridge University Press

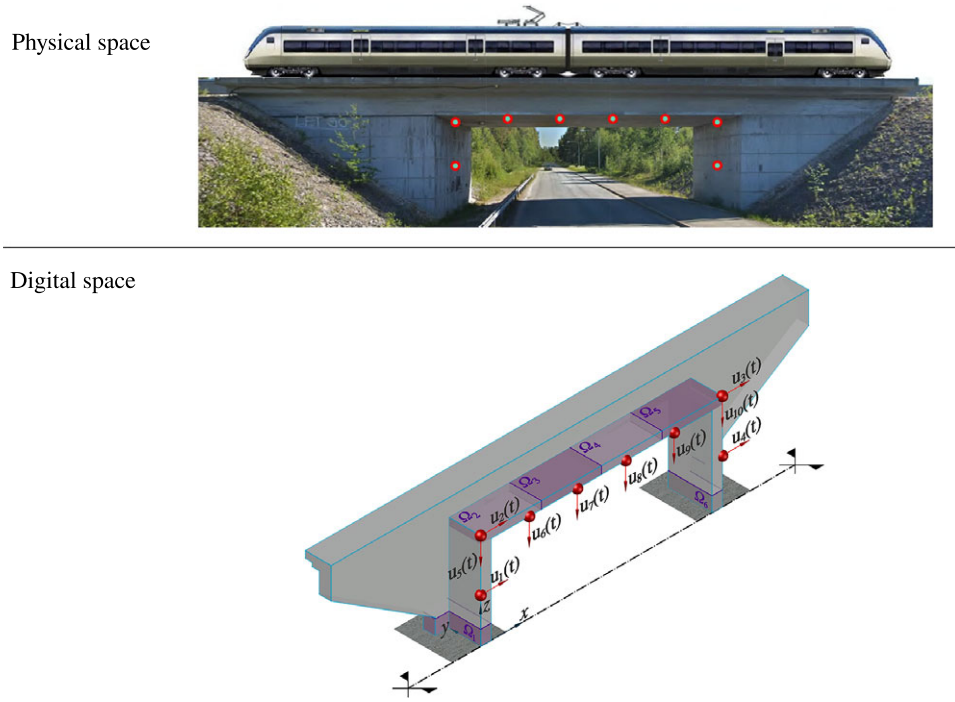


Figure 5. *Physical asset and its digital twin—The physical space corresponds to the Hörnefors railway bridge. The digital space is a schematization for structural health monitoring, including details of synthetic recordings related to displacements* $u_1(t), \ldots, u_{10}(t)$*, and predefined damage regions* $\Omega_1, \ldots, \Omega_6$.

The integral design of the Hörnefors bridge minimizes the ingress of moisture, debris, and corrosive agents, thereby enhancing its long-term durability. The types of damage that can affect such bridges depend on material and structural factors, environmental conditions, and traffic loads. Some damage patterns that can be described through a localized stiffness reduction include: concrete cracking caused by thermal gradients, freeze–thaw cycles, or overloading; gradual concrete degradation due to the chemical reaction between reactive silica aggregates and alkalis in the cement paste, potentially resulting in cracking and spalling; concrete cracking and structural damage induced by stress concentrations from differential foundational settlements; erosion of concrete surfaces due to vehicular collisions or prolonged exposure to environmental factors.

4.2. Dataset assembly and training options

The monitoring system is deployed to provide displacement recordings $\mathbf{U}$, corresponding to the $N_s = 10$ degrees of freedom highlighted in Figure 5. Each recording spans a time interval $[0, T = 1.5\text{s}]$, with a sampling frequency of 400 Hz. To simulate measurement errors caused by sensor self-noise, the recordings are corrupted with an additive Gaussian noise to yield a signal-to-noise ratio of 120. An exemplary set of displacement time histories is shown in Figure 6.

The parameter $y \in \{0, \ldots, 6\}$ labels the specific damage region. In addition to the damage-free baseline $y = 0$, the presence of damage in the structure is simulated by considering a set of $N_\Omega = 6$ predefined damageable subdomains Ω_m, with $m = 1, \ldots, 6$, as depicted in Figure 5 and corresponding to $y = 1, \ldots, 6$. Within these regions, a localized reduction of the material stiffness can occur with a magnitude $\delta \in [30\%, 80\%]$, and is kept fixed while a train travels across the bridge.

https://doi.org/10.1017/dce.2026.10072 Published online by Cambridge University Press

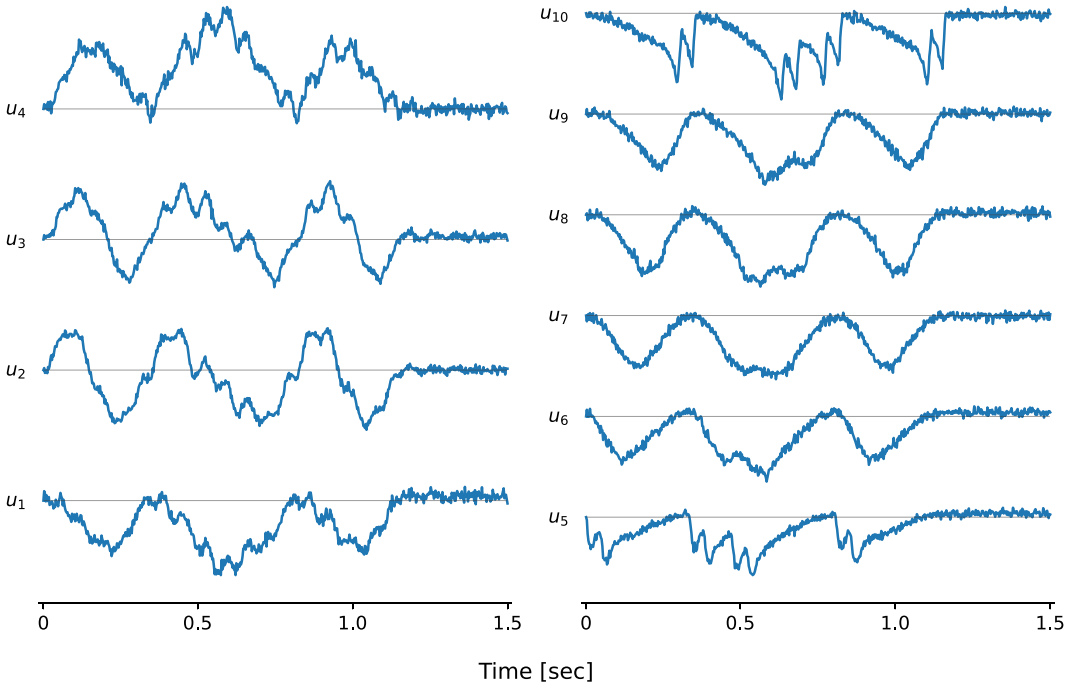


Figure 6. *Exemplary multivariate time series of structural vibrations (standardized displacements $u_1(t),\ldots,u_{10}(t)$) induced by the transit of a pilot train across the railway bridge.*

The full-order model is the same as that described in Torzoni et al., (2024). It features 17,292 degrees of freedom, resulting from a space discretization using linear tetrahedral finite elements with an element size of 0.80 m, reduced to 0.15 m for the deck. The ballast layer is modeled through an increased density for the deck and the edge beams. The embankments are modeled using distributed springs over the surfaces in contact with the ground, implemented as a Robin mixed boundary condition (with elastic coefficient 10^8 N/m^3). The structural dissipation is modeled using a Rayleigh damping matrix, calibrated to achieve a 5% damping ratio for the first two structural modes.

The solution time interval is uniformly partitioned into 600 time steps, and the time integration is performed using an implicit Newmark scheme (Hughes, 2000). We use the (unconditionally stable) constant average acceleration rule, so that the integration step can be arbitrarily selected based on the structural frequencies and the sensor sampling rate.

To reduce the computational cost of solving the full-order model for any sampled $\boldsymbol{\mu}$, a projection-based reduced-order model is used to speed up the dataset generation for training the DL models. We use a proper orthogonal decomposition Galerkin reduced basis method (Quarteroni et al., 2015; Chinesta et al., 2017). The reduced-order model is derived from a snapshot matrix assembled from 400 full-order solutions for varying input parameters $\boldsymbol{\mu} = (v, \psi, y, \delta)^\top$, taken as uniformly distributed and sampled via the Latin hypercube rule. The order of the reduced-order expansion is determined using an energy criterion. By prescribing a tolerance of 10^{-3} for the fraction of energy content to be disregarded, the number of degrees of freedom is reduced to 133. Using this reduced-order model, the training dataset is populated with 10,000 instances. These data are subsequently used to train DL models for detecting, locating, and quantifying structural damage, as described in (Torzoni et al., 2024).

The application of DL in SHM follows a pattern recognition paradigm (Bishop, 2006), where damage is assessed by comparing incoming measurements with data previously collected under known structural conditions. By capturing temporal correlations within and across time series data, DL optimizes the

https://doi.org/10.1017/dce.2026.10072 Published online by Cambridge University Press

extraction of damage-sensitive features for modeling the relationship between inputs and damage states. Moreover, as DL models are trained offline, they provide real-time predictions during operation.

Within our framework, each time new observational data are acquired—when a train crosses the bridge, marking a new DT time slice—they are first processed by a classification model to address damage detection and localization. The target class corresponds to one of the predefined damage scenarios described by the y parameter. The classifier maps the vibration recordings $\mathbf{U}$ onto confidence levels representing the likelihood of $\mathbf{U}$ belonging to each damage class. The most likely class is selected as the best point estimate for categorizing the measurements. Such a classifier is trained by minimizing the categorical cross-entropy loss function between the one-hot encoding of y and its predicted counterpart.

Whenever damage is detected and localized in the m-th region Ω_m, where $m = 1, \ldots, 6$, the observational data are further processed with a regression model—one for each damageable region—to quantify the extent of damage δ. Each regression model maps the vibration recordings associated with the m-th damage region onto the estimated magnitude of the stiffness reduction within that region. These regressors are trained by minimizing a mean-squared error loss function between the target stiffness reduction and its predicted value. For a detailed and formal description of the DL models, architectures, and their implementation, the reader is referred to (Torzoni et al., 2024).

4.3. Digital twin framework

Since the space of digital states is discrete, the outcomes of the DL models are accommodated into the PGM by discretizing the value range of δ into $N_\delta \in \mathbb{N}$ intervals. The number and width of these intervals are chosen arbitrarily, without specific restrictions. The δ range is discretized into $N_\delta = 6$ intervals $\{[30\%, 35\%], [35\%, 45\%], [45\%, 55\%], [55\%, 65\%], [65\%, 75\%], [75\%, 80\%]\}$. This results in $|\mathcal{D}| = 1 + N_\delta N_\Omega = 37$ digital states, ordered first by damage location and then by damage level.

In the absence of experimental data, the testing phase of the DL models is carried out using 4000 noisy full-order solutions. The ground truth digital state is classified with an accuracy of 91.39%.

We consider four different control actions, each associated with a CPT that models the transition probability $p(D_{t+1}|D_t, U_t = u_t)$. Together, these CPTs encode the ϕ_t^{history} factor. Note that these internal models predict the evolution of structural health but do not replicate the ground truth progression, which remains unknown to the DT. The considered control actions are as follows:

1. *Do nothing* (DN) action: no maintenance is scheduled, and the structural health state evolves according to a stochastic deterioration process, while regular revenue is maintained.
2. *Restrict operations* (RO) action: traffic flow on the bridge is restricted to lightweight trains carrying less than 18ton per axle. This reduces the structure degradation rate but also decreases the revenue generated by the infrastructure.
3. *Minor repair* (MR) action: a moderate-cost maintenance intervention is carried out to address specific damages, though the structure does not return to a damage-free state.
4. *Perfect repair* (PR) action: a high-cost maintenance intervention is performed, almost certainly restoring the bridge to a damage-free state.

The (unknown) ground truth stochastic model varies depending on the most recent control actions. We assume that damage can develop in any of the predefined regions, without propagating between different damageable subdomains. Based on this assumption, we define below the prescribed degradation (or improvement) evolution model associated with each action.

4.3.1. Do nothing action

In the absence of maintenance actions, a (simulated) stochastic deterioration process monotonically degrades the structural health. If the (unknown) state of the system is undamaged, the presence/location of damage y is sampled from a categorical distribution. This distribution assigns half of its probability mass to the undamaged configuration $y = 0$, indicating that the bridge remains intact. The remaining half is evenly distributed among the remaining damage classes $y \neq 0$. Damage may develop with magnitude δ

https://doi.org/10.1017/dce.2026.10072 Published online by Cambridge University Press

uniformly sampled within the range of the first damage interval $[30\%, 35\%]$. In contrast, if the system is already damaged, further deterioration propagates through δ increments sampled from a categorical distribution with Dirichlet hyperparameters $\boldsymbol{\alpha}_{\mathrm{DN}}^{\mathrm{true}} = \left(\alpha_{\mathrm{DN}}^{0,\mathrm{true}}, \alpha_{\mathrm{DN}}^{1,\mathrm{true}}, \alpha_{\mathrm{DN}}^{2,\mathrm{true}}\right)^{\top}$, scaled by the δ discretization step 10%, and finally perturbed by Gaussian noise with mean 2% and standard deviation 0.5% (negative increments are rounded to zero). The DN model can be formalized as follows:

$$y \sim \mathrm{Categorical}\left(\frac{1}{2}, \frac{1}{12}, \frac{1}{12}, \frac{1}{12}, \frac{1}{12}, \frac{1}{12}, \frac{1}{12}\right), \tag{4.1}$$

$$\delta_0 | y \neq 0 \sim \mathrm{Uniform}(0.3, 0.35), \tag{4.2}$$

$$\delta_t - \delta_{t-1} | y \neq 0 \sim 0.1 \cdot \mathrm{Categorical}\left(\mathbf{p}_{\mathrm{DN}}^{\mathrm{true}}\right) + \mathrm{Normal}_{\geq 0}(0.02, 0.005), \tag{4.3}$$

$$\mathbf{p}_{\mathrm{DN}}^{\mathrm{true}} \sim \mathrm{Dirichlet}\left(\boldsymbol{\alpha}_{\mathrm{DN}}^{\mathrm{true}}\right), \quad \boldsymbol{\alpha}_{\mathrm{DN}}^{\mathrm{true}} = (500, 400, 200)^{\top}. \tag{4.4}$$

The prescribed trajectory of the structural health parameters is arbitrarily chosen to fully showcase the capabilities of the DT. However, the DT would be equally capable of tracking the structural health evolution, whether considering more or less aggressive degradation processes.

4.3.2. Restrict operation action

If only lightweight trains are allowed to cross the bridge, the model resembles the one for the DN action but is characterized by a lower probability of damage initiation and a reduced deterioration rate. In this scenario, damage may occur with a probability of 0.25, evenly distributed among the damage classes. Once damage develops, it propagates with δ increments sampled from a categorical distribution with Dirichlet hyperparameters $\boldsymbol{\alpha}_{\mathrm{RO}}^{\mathrm{true}} = \left(\alpha_{\mathrm{RO}}^{0,\mathrm{true}}, \alpha_{\mathrm{RO}}^{1,\mathrm{true}}, \alpha_{\mathrm{RO}}^{2,\mathrm{true}}\right)^{\top}$, scaled by the δ discretization step 10%, and finally perturbed by Gaussian noise with mean 2% and standard deviation 0.5% (negative increments are rounded to zero). The RO model is formalized as follows:

$$y \sim \mathrm{Categorical}\left(\frac{3}{4}, \frac{1}{24}, \frac{1}{24}, \frac{1}{24}, \frac{1}{24}, \frac{1}{24}, \frac{1}{24}\right), \tag{4.5}$$

$$\delta_0 | y \neq 0 \sim \mathrm{Uniform}(0.3, 0.35), \tag{4.6}$$

$$\delta_t - \delta_{t-1} | y \neq 0 \sim 0.1 \cdot \mathrm{Categorical}\left(\mathbf{p}_{\mathrm{RO}}^{\mathrm{true}}\right) + \mathrm{Normal}_{\geq 0}(0.02, 0.005), \tag{4.7}$$

$$\mathbf{p}_{\mathrm{RO}}^{\mathrm{true}} \sim \mathrm{Dirichlet}\left(\boldsymbol{\alpha}_{\mathrm{RO}}^{\mathrm{true}}\right), \quad \boldsymbol{\alpha}_{\mathrm{RO}}^{\mathrm{true}} = (600, 300, 100)^{\top}. \tag{4.8}$$

4.3.3. Minor repair action

If the structural health state of the bridge undergoes partial improvement and $y = 0$, the system remains undamaged. If $y \neq 0$, the damage magnitude δ decreases. Specifically, the δ decrement is sampled from a categorical distribution with Dirichlet hyperparameters $\boldsymbol{\alpha}_{\mathrm{MR}}^{\mathrm{true}} = \left(\alpha_{\mathrm{MR}}^{0,\mathrm{true}}, \alpha_{\mathrm{MR}}^{0,\mathrm{true}}, \alpha_{\mathrm{MR}}^{0,\mathrm{true}}\right)^{\top}$, scaled by the δ discretization step 10%, and finally perturbed by Gaussian noise with mean -2% and standard deviation 0.5% (positive increments are rounded to zero). The state is updated to reflect a partial repair, ensuring that the damage magnitude does not fall below the undamaged state threshold of $\delta = 30\%$. The MR model is formalized as follows:

$$\delta_t - \delta_{t-1} \,|\, y \neq 0 \sim -0.1 \cdot \mathrm{Categorical}\left(\mathbf{p}_{\mathrm{MR}}^{\mathrm{true}}\right) + \mathrm{Normal}_{\leq 0}(-0.02, 0.005), \tag{4.9}$$

$$\mathbf{p}_{\mathrm{MR}}^{\mathrm{true}} \sim \mathrm{Dirichlet}\left(\boldsymbol{\alpha}_{\mathrm{MR}}^{\mathrm{true}}\right), \quad \boldsymbol{\alpha}_{\mathrm{MR}}^{\mathrm{true}} = (1, 100, 300)^{\top}. \tag{4.10}$$

https://doi.org/10.1017/dce.2026.10072 Published online by Cambridge University Press

4.3.4. Perfect repair action

Under PR, the system returns to the damage-free configuration, independently of its damage state.

4.3.5. Transition model

All numerical results are obtained using a Dirichlet–Multinomial state-independent transition model. Transition probabilities thus depend only on the selected action and not on the current state, and are learned online through action-specific Dirichlet priors updated via observed state transitions.

The assumed transition model $p(D_{t+1}|D_t, U_t = u_t)$ satisfies the assumption that only one damage scenario can be activated at a time. The transition dynamics associated with the different actions are formulated as described in Section 3.1. Starting from an initial belief over the transition probabilities, these are subsequently refined by integrating evidence about the system response to applied actions. Specifically, we instantiate two Dirichlet–Multinomial state-independent models to represent degradations of zero, one, or two δ intervals under the DN and RO actions, respectively, and one Dirichlet–Multinomial state-independent model to represent improvements of zero, one, or two δ intervals under the MR action. The initial Dirichlet hyperpriors are assigned as:

$$\boldsymbol{\alpha}_{\mathrm{DN}} = \left(\alpha^0_{\mathrm{DN}}, \alpha^1_{\mathrm{DN}}, \alpha^2_{\mathrm{DN}}\right)^\top = (10,7,4)^\top, \tag{4.11}$$

$$\boldsymbol{\alpha}_{\mathrm{RO}} = \left(\alpha^0_{\mathrm{RO}}, \alpha^1_{\mathrm{RO}}, \alpha^2_{\mathrm{RO}}\right)^\top = (10,5,2)^\top, \tag{4.12}$$

$$\boldsymbol{\alpha}_{\mathrm{MR}} = \left(\alpha^0_{\mathrm{MR}}, \alpha^1_{\mathrm{MR}}, \alpha^2_{\mathrm{MR}}\right)^\top = (3,10,9)^\top, \tag{4.13}$$

where $\boldsymbol{\alpha}_{\mathrm{DN}}$ and $\boldsymbol{\alpha}_{\mathrm{RO}}$ prioritize remaining in the same state over one-step and two-step degrading transitions, while $\boldsymbol{\alpha}_{\mathrm{MR}}$ prioritizes one-step and two-step recovery over remaining in the same state. Finally, for the PM action, we prescribe a transition matrix that assigns unit probability to transitioning to the undamaged state, independently of the current condition.

Transitions are defined in terms of discrete step variations, corresponding to changes of zero, one, or two δ intervals. The state space is bounded by construction, so that any transition exceeding the admissible range is mapped onto the nearest valid state, that is, the undamaged or most-damaged condition. Moreover, it is worth highlighting that our state-independent assumption is appropriate when the deterioration and repair dynamics are primarily governed by the applied action rather than the current damage level, or when available data are insufficient to reliably estimate fully state-dependent transitions. Under these conditions, pooling observations across states improves statistical efficiency and yields more stable learning, while still capturing the dominant degradation and recovery trends.

4.3.6. Reward model

At each time step, the DT selects a control action $u_t \in \{\mathrm{DN}, \mathrm{RO}, \mathrm{MR}, \mathrm{PR}\}$. Choosing a DN or RO action yields a positive reward but also entails the risk of structural deterioration. Conversely, the MR and PR actions mitigate structural health deterioration but incur a negative reward. The costs associated with the structural health state, encoded in D_t, and the costs related to the control actions U_t, appearing in Equation (3.10), are defined as follows:

$$R_t^{\mathrm{health}}(d) = \begin{cases} 0 & \text{if } y = 0, \delta = 0, \\ -\exp(5\delta) + 4 & \text{if } y \neq 0, 0 < \delta < 80\%, \\ -250 & \text{if } \delta \geq 80\%, \end{cases} \quad R_t^{\mathrm{control}}(u) = \begin{cases} +30 & \text{if } u = \mathrm{DN}, \\ +18 & \text{if } u = \mathrm{RO}, \\ -75 & \text{if } u = \mathrm{MR}, \\ -250 & \text{if } u = \mathrm{PM}. \end{cases} \tag{4.14}$$

These non-dimensional rewards represent indicative costs charged to the decision-maker. Actual values can be derived from service and cost lists provided by state agencies and companies (KG Papakonstantinou and Shinozuka, 2014). For instance, R^{health} is designed to penalize the progressive deterioration

https://doi.org/10.1017/dce.2026.10072 Published online by Cambridge University Press

of the structural condition as a function of δ. These costs may reflect various impacts, such as a reduction in service level, workplace accidents, or increased structural failure probability (KG Papakonstantinou and Shinozuka, 2014), with the last entry imposing a significantly negative reward for excessively compromised structural states.

4.4. Results

In this section, we present the results of DT simulations carried out over an expected lifespan of $T = 60$ time steps. At each time step, new observational data are simulated based on the (unknown) ground truth structural health and the most recently applied control action. The DT assimilates these data, estimates the digital state, and recommends the next control action.

By running a cluster of 100 simulations, each spanning 60 time steps and initialized with a different random seed for the ground-truth physical process, both the DT using a finite-horizon dynamic programming policy without precision updates—computed based on the initial prior transition model—and the adaptive DT using a finite-horizon MBRL policy with precision updates, where control policies are iteratively recomputed using the updated transition model, exhibit zero failures. Figure 7 compares the evolution of cumulative rewards, showing that the adaptive DT yields significantly higher cumulative rewards (+17%). This improvement stems from the online refinement of the transition model, which enables more informed planning and decision-making. Moreover, our experiments indicate that non-adaptive policies can lead to undesirable outcomes; depending on the interplay between the transition model, the reward model, and the ground-truth dynamics, such policies may cause the DT to fail, that is, the structural state reaches a critically compromised condition. For instance, digital failure (corresponding to $\delta \geq 80\%$ and symbolizing structural collapse) can occur if the initial prior over the transition dynamics is poorly aligned with the ground-truth process. In contrast, the robustness of MBRL ensures that the DT consistently maintains the structural health above critical thresholds.

In-simulation performance is further assessed through a tracking accuracy metric based on the maximum a posteriori estimate of the digital state relative to the ground-truth physical process. To account for the δ discretization, accuracy is computed using a 10% tolerance, corresponding to the discretization step. Under these conditions, the adaptive DT achieves a mean accuracy of 93.6% with a standard deviation of 4%, while the non-adaptive counterpart attains a mean accuracy of 93.15% with a

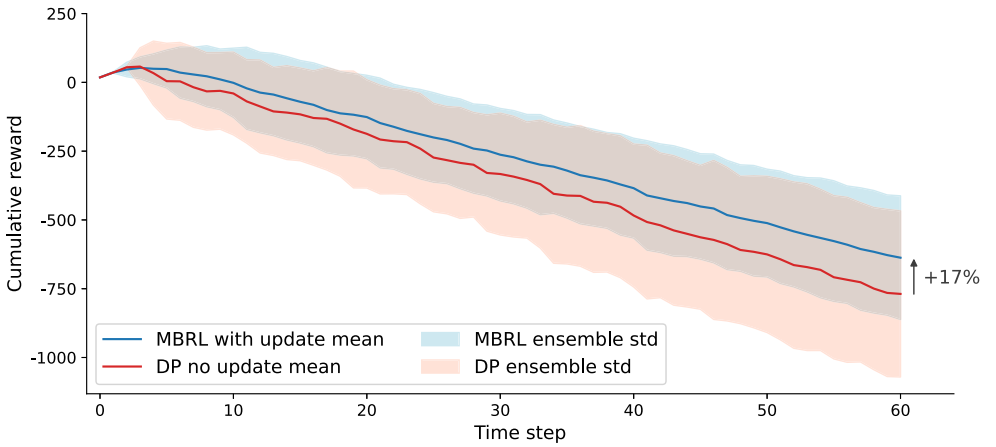


Figure 7. *Digital twin online phase—Comparison of cumulative rewards under control policies computed using finite-horizon model-based reinforcement learning (MBRL) with precision updates and finite-horizon dynamic programming (DP) without precision updates. Solid lines denote the mean cumulative rewards, while the shaded areas indicate the one-standard-deviation credibility intervals. Results are shown for two clusters of simulations, each comprising* 100 *runs initialized with different random seeds.*

https://doi.org/10.1017/dce.2026.10072 Published online by Cambridge University Press

standard deviation of 4.3%. To avoid biasing the metric, true undamaged cases are excluded from this computation, as they are easier to identify and would favor conservative policies that promote occupancy of the undamaged state. The similar tracking accuracies observed for the two configurations are expected, since the observation model itself is not modified by the adaptive updates.

Figure 8 illustrates a sample DT simulation using finite-horizon MBRL with precision updates. The DT accurately tracks the digital state evolution with relatively low uncertainty and a limited delay of at most two time steps with respect to the ground truth. This delay is attributed both to the need to update the prior belief based on data from previous time steps and to occasional misclassification by the DL model between digital states associated with the same damage location but adjacent δ intervals. The figure also demonstrates how accurate inference of transition dynamics and optimization of finite-horizon control

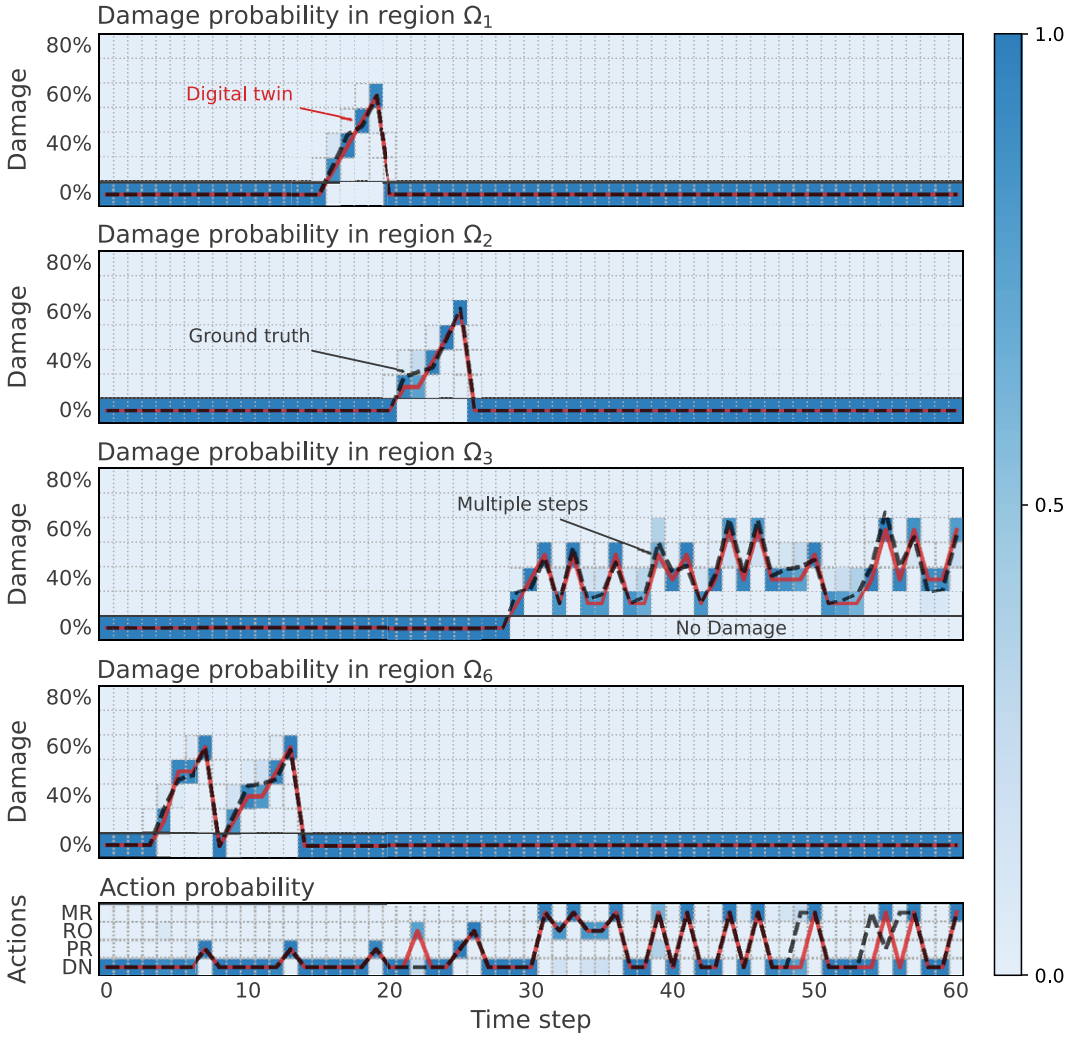


Figure 8. *Digital twin online phase—Simulation under finite-horizon model-based reinforcement learning with precision updates. Probabilistic and best point estimates of: (top four panels) digital state evolution against the ground truth physical state for regions Ω_1, Ω_2, Ω_3, and Ω_6; (bottom panel) control actions informed by the digital twin against the optimal action under the ground truth. The background colors represent the digital state beliefs in the top panels, and the control action beliefs in the bottom panel.*

https://doi.org/10.1017/dce.2026.10072 Published online by Cambridge University Press

policies allow the PR action to be avoided as the simulation approaches the end of the planning horizon (i.e., the decommissioning time). Specifically, the belief of selecting a PR action for the same damage magnitude progressively decreases as the simulation advances. Additionally, Figure 9 shows the predicted evolution of the digital state and control actions starting from $t_c = 60$ and spanning a prediction horizon of $t_p = t_c + 5$. The DT first forecasts the expected improvement in structural health after taking the MR action, based on the learned transition model. This is followed by a high-probability prediction of DN and MR actions, yielding a pseudo-stable belief propagation pattern that maintains most of the probability mass of the digital state within the first three δ intervals.

Figure 10 illustrates the initial Dirichlet priors on the transition probabilities for the DN, RO, and MR actions, together with the corresponding posteriors at the end of the simulation. In all cases, the posteriors exhibit clear convergence toward the (unknown) ground-truth transition probabilities while significantly reducing uncertainty. The corresponding posterior parameters read:

$$\overline{\boldsymbol{\alpha}}_{\mathrm{DN}} = \left(\alpha^0_{\mathrm{DN}} + N^0_{\mathrm{DN}}, \alpha^1_{\mathrm{DN}} + N^1_{\mathrm{DN}}, \alpha^2_{\mathrm{DN}} + N^2_{\mathrm{DN}}\right)^{\top}, \tag{4.15}$$

$$\overline{\boldsymbol{\alpha}}_{\mathrm{RO}} = \left(\alpha^0_{\mathrm{RO}} + N^0_{\mathrm{RO}}, \alpha^1_{\mathrm{RO}} + N^1_{\mathrm{RO}}, \alpha^2_{\mathrm{RO}} + N^2_{\mathrm{RO}}\right)^{\top}, \tag{4.16}$$

$$\overline{\boldsymbol{\alpha}}_{\mathrm{MR}} = \left(\alpha^0_{\mathrm{MR}} + N^0_{\mathrm{MR}}, \alpha^1_{\mathrm{MR}} + N^1_{\mathrm{MR}}, \alpha^2_{\mathrm{MR}} + N^2_{\mathrm{MR}}\right)^{\top}, \tag{4.17}$$

whose probability masses are centered around the true modes, indicating that the updated beliefs tend to reflect the actual transition probabilities while reducing uncertainty.

A dedicated sensitivity analysis to observation uncertainty is not repeated in the present work; instead, we refer the reader to (Torzoni et al., 2026), where this aspect has been systematically investigated. The analysis considers progressively degraded observation conditions, including incomplete data streams during train passages, varying observation failure rates, and uncertainty in damage extent. The results show a gradual performance degradation: reduced data availability increases state uncertainty and may delay optimal actions, while the transition model helps preserve acceptable performance even under severe data loss, although with increased uncertainty. On the other hand, uncertainty in damage extent mainly affects estimation accuracy, with negligible impact on failure rates.

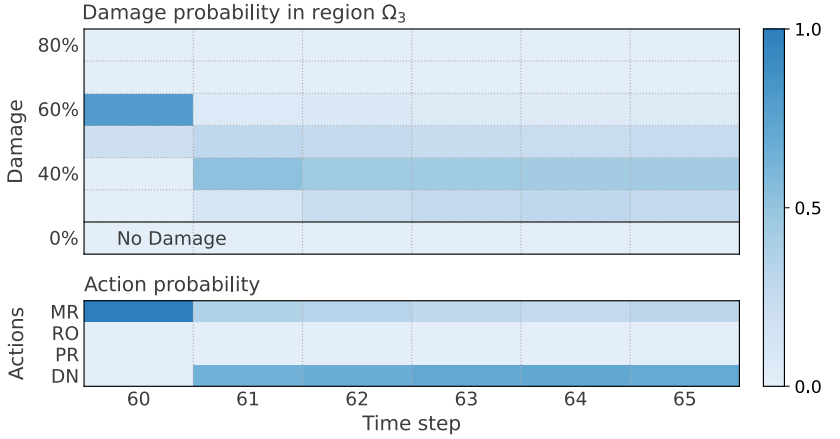


Figure 9. *Digital twin online phase—Future predictions starting at $t_c = 60$ under finite-horizon model-based reinforcement learning with precision updates. The background colors represent the digital state beliefs in the top panels, and the control action beliefs in the bottom panel.*

https://doi.org/10.1017/dce.2026.10072 Published online by Cambridge University Press

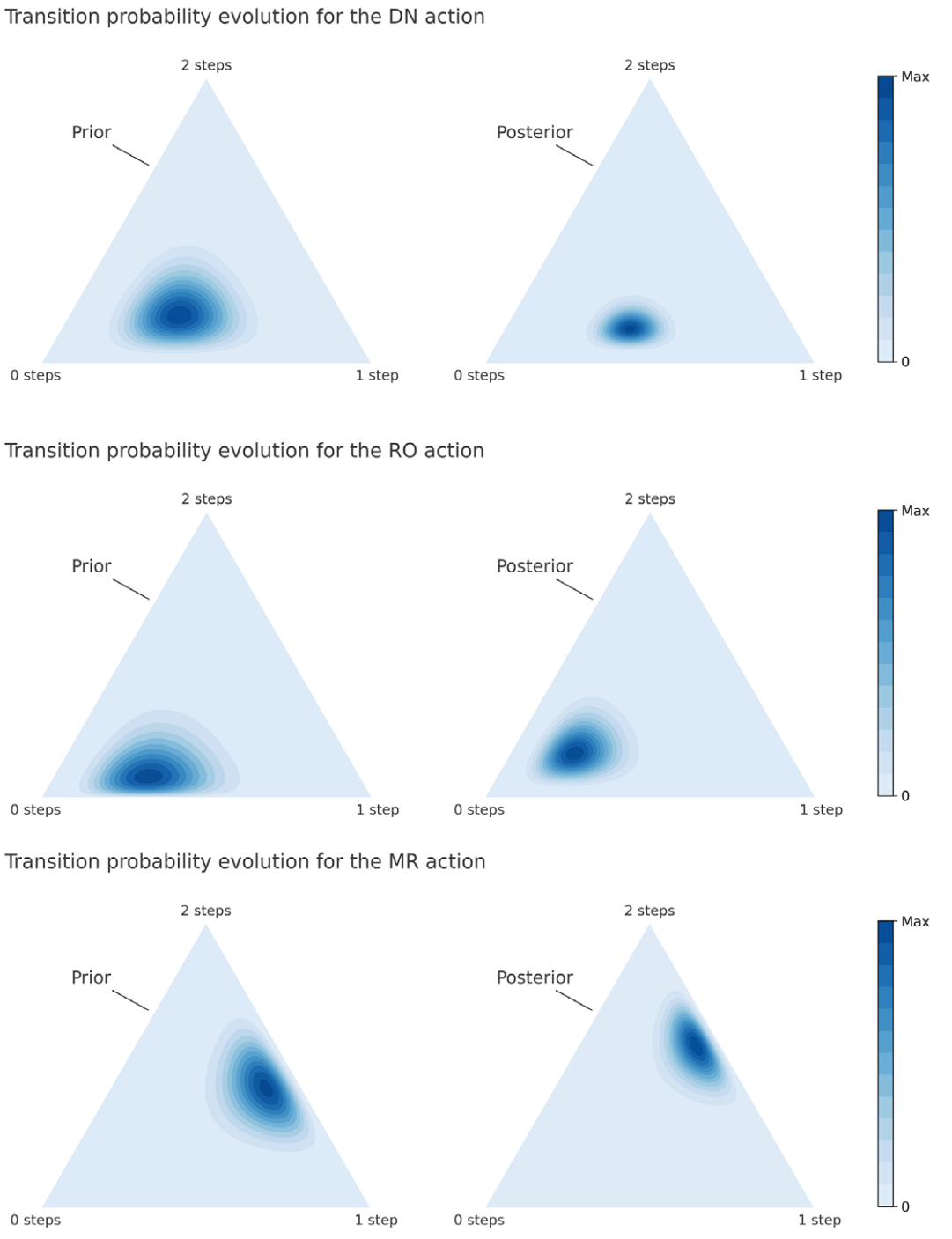


Figure 10. *Prior and posterior distributions of the transition probabilities associated with the do nothing action (first row), the restrict operations action (second row), and the minor repair action (third row). Initial priors correspond to* $t = 0$*, while the posterior distributions correspond to* $t = 60$*.*

https://doi.org/10.1017/dce.2026.10072 Published online by Cambridge University Press

5. Conclusion

In this paper, we have introduced adaptivity into the predictive DT framework for SHM proposed in (Torzoni et al., 2024) to enhance decision-making capabilities. The framework models the bi-directional information flow between the monitored asset and its virtual replica using DBNs. Adaptivity is achieved by treating the state transition probabilities as random variables endowed with conjugate prior distributions that are updated online through Bayesian inference. Adaptive planning is integrated via parametric MDPs, enabling dynamic control policies with precision updates. This work extends the results in (Tezzele et al., 2026) to a broader class of distributions, thereby allowing the proposed DT framework to be applied to real-life, large-scale systems.

The computational procedure has been assessed through simulations of monitoring, management, and maintenance planning for the Hörnefors railway bridge. The DT effectively tracks the structural health evolution, recommending timely and appropriate control actions with low uncertainty. By continuously refining the state transition models using online data, the DT dynamically adapts to evolving structural conditions, providing more reliable structural health assessments, improved forecasts accuracy, and enhanced decision-making. When discrepancies with the ground truth arise, the DT demonstrates the ability to self-correct after assimilating new observational data.

Future research will address additional challenges, such as environmental uncertainties that can cause sudden structural degradation (Sanchez-Silva et al., 2011; Guo et al., 2020), complementing the long-term deterioration effects considered in this study. Particular attention will be devoted to extreme events driven by environmental risks and climate change, which require robust control policies that proactively account for the probability of such events. Further efforts will focus on partial observability by incorporating information-gathering actions, such as installing new sensors and scheduling inspections. In addition, future work will explore explicitly non-stationary dynamics, such as time-dependent deterioration rates or abrupt changes in system behavior. Addressing these scenarios may require the integration of adaptive learning strategies, such as forgetting mechanisms or context-aware models (Gama et al., 2014; Padakandla, 2021), to effectively track evolving transition dynamics.

Data availability statement. The DT framework was implemented using the `pgmtwin` Python package (Fabris et al., 2025), available at https://github.com/pgmtwin/pgmtwin. The observational data used to run the experiments presented in Section 4 are available in the public repository https://github.com/mtezzele/digital-twin-SHM (Torzoni and Tezzele, 2026). The DL models trained according to the implementation details reported in Torzoni et al., (2024) are also made available in the same repository. The `Matlab` library for finite element simulation and reduced-order modeling of partial differential equations used to generate these data is available in the public repository http://redbkit.github.io/redbKIT (Negri, 2016).

Author contribution. Conceptualization: M.Te.; Data visualization: M.To., M.Te.; Formal analysis: E.V., M.To.; Funding acquisition: A.M.; Methodology: M.To., M.Te.; Software: E.V., M.To., M.Te.; Validation: M.To., M.Te.; Writing original draft: E.V., M.To.; Writing – review and editing: M.To., M.Te., A.M. All authors approved the final submitted draft.

Funding statement. M.To. acknowledges the financial support from the Politecnico di Milano through the interdisciplinary Ph.D. Grant "Physics-informed deep learning for structural health monitoring". M.To. and A.M. acknowledge the PRIN 2022 Project UQ-FLY (Numerical approximation of uncertainty quantification problems for PDEs by multi-fidelity methods), funded by the NextGenerationEU program within the PNRR scheme (M4C2, Investment 1.1). A.M. acknowledges the Project "Reduced Order Modeling and Deep Learning for the real-time approximation of PDEs (DREAM)" (Starting Grant No. FIS00003154), funded by the Italian Science Fund (FIS)—Ministero dell'Università e della Ricerca and the project FAIR (Future Artificial Intelligence Research), funded by the NextGenerationEU program within the PNRR-PE-AI scheme (M4C2, Investment 1.3, Line on Artificial Intelligence). M.To. and A.M. acknowledge the financial support from the ERC advanced grant IMMENSE (Grant Agreement 101140720), funded by the European Union. Views and opinions expressed are, however, those of the authors only and do not necessarily reflect those of the European Union or the European Research Council Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

Competing interests. The authors declare no competing interests have appeared to influence the work reported in this paper.

https://doi.org/10.1017/dce.2026.10072 Published online by Cambridge University Press

References

Achenbach JD (2009) Structural health monitoring – What is the prescription? *Mechanics Research Communications 36*(2), 137–142. https://doi.org/10.1016/j.mechrescom.2008.08.011.

American Society of Civil Engineers (2018) Adaptive design and risk management. In *Climate-Resilient Infrastructure*, Vol. *7*. Reston, Virginia: American Society of Civil Engineers, pp. 173–226.

Andriotis CP and Papakonstantinou KG (2019) Managing engineering systems with large state and action spaces through deep reinforcement learning. *Reliability Engineering & System Safety 191*, 106483. https://doi.org/10.1016/j.ress.2019.04.036.

Arcieri G, **Hoelzl C**, **Schwery O**, **Straub D**, **Papakonstantinou KG and Chatzi E** (2023) Bridging POMDPs and Bayesian decision making for robust maintenance planning under model uncertainty: An application to railway systems. *Reliability Engineering & System Safety*, 109496. https://doi.org/10.1016/j.ress.2023.109496.

Arcieri G, **Hoelzl C**, **Schwery O**, **Straub D**, **Papakonstantinou K and Chatzi E** (2024) POMDP inference and robust solution via deep reinforcement learning: An application to railway optimal maintenance. *Machine Learning*. https://doi.org/10.1007/s10994-024-06559-2.

Arcones DA, **Weiser M**, **Koutsourelakis PS and Unger JF** (2024) Model bias identification for Bayesian calibration of stochastic digital twins of bridges. *Applied Stochastic Models in Business and Industry.* https://doi.org/10.1002/asmb.2897.

ASCE (2021) tech. rep. A comprehensive assessment of America's infrastructure. Reston, Virginia: World Headquarters. https://2021.infrastructurereportcard.org/cat-item/bridges-infrastructure/

Avci O, **Abdeljaber O**, **Kiranyaz S**, **Hussein M**, **Gabbouj M and Inman D** (2021) A review of vibration-based damage detection in civil structures: From traditional methods to machine learning and deep learning applications. *Mechanical Systems and Signal Processing 147*, 107077. https://doi.org/10.1016/j.ymssp.2020.107077.

Bauer P, **Stevens B and Hazeleger W** (2021) A digital twin of earth for the green transition. *Nature Climate Change 11*(2), 80–83. https://doi.org/10.1038/s41558-021-00986-y.

Bellman R (1957a) *Dynamic Programming*. Princeton, NJ: Princeton University Press.

Bellman R (1957b) A Markovian decision process. *Journal of Mathematics and Mechanics 6*(5), 679–684.

Bianchi S and Biondini F (2022) Life-cycle assessment of deteriorating RC bridges using artificial neural networks. *Journal of Infrastructure Systems 28*(2). https://doi.org/10.1061/(ASCE)IS.1943-555X.0000672.

Biondini F and Frangopol DM (2016) Life-cycle performance of deteriorating structural systems under uncertainty review. *Journal of Structural Engineering 142*(9), F4016001. https://doi.org/10.1061/(ASCE)ST.1943-541X.0001544.

Biondini F and Frangopol DM (2017) Time-variant redundancy and failure times of deteriorating concrete structures considering multiple limit states. *Structure and Infrastructure Engineering 13*(1), 94–106. https://doi.org/10.1080/15732479.2016.1198403.

Bishop CM (2006) *Pattern Recognition and Machine Learning*. Information Science and Statistics. New York, NY: Springer.

Brownjohn JM (2007) Structural health monitoring of civil infrastructure. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences 365*(1851), 589–622. https://doi.org/10.1098/rsta.2006.1925.

Chaudhuri A, **Pash G**, **Hormuth DA**, **Lorenzo G**, **Kapteyn M**, **Wu C**, **Lima EA**, **Yankeelov TE**, **Willcox K** (2023) Predictive digital twin for optimizing patient-specific radiotherapy regimens under uncertainty in high-grade gliomas. *Frontiers in Artificial Intelligence 6*. https://doi.org/10.3389/frai.2023.1222612.

Cherifi K, **Mehrmann V and Schulze P** (2022) Simulations in a Digital Twin of an Electrical Machine. arsXiv preprint arXiv:2207.02171. https://doi.org/10.48550/arXiv.2207.02171.

Chinesta F, **Huerta A**, **Rozza G and Willcox K** (2017) Model reduction methods. In Stein E, deBR and Hughes TJR (eds.), *Encyclopedia of Computational Mechanics*, 2nd Edn. John Wiley & Sons, pp. 1–36.

Corral-Acero J, **Margara F**, **Marciniak M**, **Rodero C**, **Loncaric F**, **Feng Y**, **Gilbert A**, **Fernandes JF**, **Bukhari HA**, **Wajdan A**, **Martinez MV**, **Santos MS**, **Shamohammdi M**, **Luo H**, **Westphal P**, **Leeson P**, **Di Achille P**, **Gurev V**, **Mayr M**, **Geris L**, **Pathmanathan P**, **Morrison T**, **Cornelussen R**, **Prinzen F**, **Delhaas T**, **Doltra A**, **Sitges M**, **Vigmond EJ**, **Zacur E**, **Grau V**, **Rodriguez B**, **Remme EW**, **Niederer S**, **Mortier P**, **McLeod K**, **Potse M**, **Pueyo E**, **Bueno-Orovio A**, **Lamata P** (2020) The 'digital twin' to enable the vision of precision cardiology. *European Heart Journal 41*(48), 4556–4564. https://doi.org/10.1093/eurheartj/ehaa159.

Darema F (2004) Dynamic data driven applications systems: A new paradigm for application simulations and measurements. In Bubak M, vanAGD, Sloot PMA and Dongarra J (eds), *International Conference on Computational Science*. Berlin, Germany: Springer, pp. 662–669.

Eftekhar Azam S and Mariani S (2018) Online damage detection in structural systems via dynamic inverse analysis: A recursive Bayesian approach. *Engineering Structures 159*, 28–45. https://doi.org/10.1016/j.engstruct.2017.12.031.

Fabris L, **Torzoni M and Tezzele M** (2025) *pgmtwin library on GitHub.* Available at https://github.com/pgmtwin/pgmtwin (accessed 17 July 2026).

Farrar CR and Worden K (2007) An introduction to structural health monitoring. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences 365*(1851), 303–315. https://doi.org/10.1098/rsta.2006.1928.

Farrar C and Worden K (2013) *Structural Health Monitoring: A Machine Learning Perspective*. Chichester, United Kingdom: John Wiley & Sons.

Fernandez-Navamuel A, **Zamora-Sánchez D**, **Omella ÁJ**, **Pardo D**, **Garcia-Sanchez D and Magalhães F** (2022) Supervised deep learning with finite element simulations for damage identification in bridges. *Engineering Structures 257*, 114016. https://doi.org/10.1016/j.engstruct.2022.114016.

https://doi.org/10.1017/dce.2026.10072 Published online by Cambridge University Press

Ferrari A and Willcox K (2024) Digital twins in mechanical and aerospace engineering. *Nature Computational Science 4*, 178–183. https://doi.org/10.1038/s43588-024-00613-8.

Franco NR, **Fraulin D**, **Manzoni A and Zunino P** (2024) On the latent dimension of deep autoencoders for reduced order modeling of PDEs parametrized by random fields. *Advances in Computational Mathematics 50*, 96. https://doi.org/10.1007/s10444-024-10189-6.

Gama J, **Žliobaitė I**, **Bifet A**, **Pechenizkiy M and Bouchachia A** (2014) A survey on concept drift adaptation. *ACM Computing Surveys 46*(4). https://doi.org/10.1145/2523813.

Garayalde G, **Rosafalco L**, **Torzoni M and Corigliano A** (2025) Mastering truss structure optimization with tree search. *Journal of Mechanical Design 147*(10), 101702. https://doi.org/10.1115/1.4068300.

García-Macías E and Ubertini F (2022) Integrated SHM systems: Damage detection through unsupervised learning and data fusion. In Cury A, Ribeiro D, Ubertini F and Todd MD (eds), *Structural Health Monitoring Based on Data Science Techniques*. Cham, Switzerland: Springer International Publishing, pp. 247–268.

Gelman A, **Carlin JB**, **Stern HS**, **Dunson DB**, **Vehtari A and Rubin DB** (2013) *Bayesian Data Analysis*, 3rd Edn. Boca Raton, FL: CRC Press. https://doi.org/10.1201/b16018.

Glaser SD and Tolman A (2008) Sense of sensing: From data to informed decisions for the built environment. *Journal of Infrastructure Systems 14*(1), 4–14. https://doi.org/10.1061/(ASCE)1076-0342(2008)14:1(4).

Guo HY, **Dong Y and Gu XL** (2020) Two-step translation method for time-dependent reliability of structures subject to both continuous deterioration and sudden events. *Engineering Structures 225*, 111291. https://doi.org/10.1016/j.engstruct.2020.111291.

Henao-Garcia S, **Kapteyn M**, **Willcox KE**, **Tezzele M**, **Castroviejo-Fernandez M**, **Kim T**, **Ambrosino M**, **Kolmanovsky I**, **Basu H**, **Jirwankar P and Sanfelice R** (2025) *Digital-Twin-Enabled Multi-Spacecraft On-Orbit Operations*. In AIAA SCITECH 2025 Forum. American Institute of Aeronautics and Astronautics, Inc.

Hernández-González I and García-Macías E (2024) Towards a comprehensive damage identification of structures through populations of competing models. *Engineering with Computers*, 85–91. https://doi.org/10.1007/s00366-024-01972-6.

Howard RA (1960) *Dynamic Programming and Markov Processes*. Oxford, England: John Wiley.

Hughes TJ (2000) *The Finite Element Method: Linear Static and Dynamic Finite Element Analysis*. Mineola, NY: Dover Publications.

Jans-Singh M, **Leeming K**, **Choudhary R and Girolami M** (2020) Digital twin of an urban-integrated hydroponic farm. *Data-Centric Engineering 1*, 20. https://doi.org/10.1017/dce.2020.21.

Jones C, **Lethco T**, **Curley J and Zackin D** (2019) tech. rep. A preventable crisis: The economic and human costs of a Hudson River rail tunnel shutdown. New York: Regional Plan Association.

Kapteyn MG, **Knezevic DJ**, **Huynh D**, **Tran M and Willcox KE** (2020) Data-driven physics-based digital twins via a library of component-based reduced-order models. *International Journal for Numerical Methods in Engineering 123*(13), 2986–3003. https://doi.org/10.1002/nme.6423.

Kapteyn MG, **Pretorius JV and Willcox KE** (2021) A probabilistic graphical model foundation for enabling predictive digital twins at scale. *Nature Computational Science 1*(5), 337–347. https://doi.org/10.1038/s43588-021-00069-0.

Karve PM, **Guo Y**, **Kapusuzoglu B**, **Mahadevan S and Haile MA** (2020) Digital twin approach for damage-tolerant mission planning under uncertainty. *Engineering Fracture Mechanics 225*, 106766. https://doi.org/10.1016/j.engfracmech.2019.106766.

Koller D and Friedman N (2009) *Probabilistic Graphical Models: Principles and Techniques*. Cambridge, MA: MIT Press.

Kritzinger W, **Karner M**, **Traar G**, **Henjes J and Sihn W** (2018) Digital twin in manufacturing: A categorical literature review and classification. *IFAC-PapersOnLine 51*(11), 1016–1022. https://doi.org/10.1016/j.ifacol.2018.08.474.

Li C, **Mahadevan S**, **Ling Y**, **Choze S and Wang L** (2017) Dynamic Bayesian network for aircraft wing health monitoring digital twin. *AIAA Journal 55*(3), 930–941. https://doi.org/10.2514/1.J055201.

Lindley DV (2013) *Understanding Uncertainty*. Hoboken, NJ: John Wiley & Sons.

Löhner R, **Airaudo F**, **Antil H**, **Wüchner R**, **Meister F and Warnakulasuriya S** (2024) High-fidelity digital twins: Detecting and localizing weaknesses in structures. *International Journal for Numerical Methods in Engineering 125*(21), e7568. https://doi.org/10.1002/nme.7568.

Mariniello G, **Pastore T**, **Menna C**, **Festa P and Asprone D** (2021) Structural damage detection and localization using decision tree ensemble and vibration data. *Computer-Aided Civil and Infrastructure Engineering 36*(9), 1129–1149. https://doi.org/10.1111/mice.12633.

McClellan A, **Lorenzetti J**, **Pavone M and Farhat C** (2022) A physics-based digital twin for model predictive control of autonomous unmanned aerial vehicle landing. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences 380*(2229), 20210204. https://doi.org/10.1098/rsta.2021.020417.

Memarzadeh M, **Pozzi M and Kolter JZ** (2015) Optimal planning and learning in uncertain environments for the management of wind farms. *Journal of Computing in Civil Engineering 29*(5), 04014076. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000390.

Morcous G (2006) Performance prediction of bridge deck systems using Markov chains. *Journal of Performance of Constructed Facilities 20*(2), 146–155. https://doi.org/10.1061/(ASCE)0887-3828(2006)20:2(146).

Morgan J (1996) tech. rep. Risk Metrics – Technical Document. JP Morgan and Reuters, New York, 102–127.

Moya B, **Alfaro I**, **Gonzalez D**, **Chinesta F and Cueto E** (2020) Physically sound, self-learning digital twins for sloshing fluids. *PLoS One 15*(6), e0234569. https://doi.org/10.1371/journal.pone.0234569.

https://doi.org/10.1017/dce.2026.10072 Published online by Cambridge University Press

National Academy of Engineering and National Academies of Sciences, Engineering, and Medicine (2024) *Foundational Research Gaps and Future Directions for Digital Twins*. Washington, DC: The National Academies Press. https://doi.org/10.17226/26894.

Negri F (2016) *redbKIT, version 2.2.* Available at http://redbkit.github.io/redbKIT (accessed 17 July 2026).

Padakandla S (2021) A survey of reinforcement learning algorithms for dynamically varying environments. *ACM Computing Surveys 54*(6). https://doi.org/10.1145/3459991.

Pandey A and Biswas M (1994) Damage detection in structures using changes in flexibility. *Journal of Sound and Vibration 169*(1), 3–17. https://doi.org/10.1006/jsvi.1994.1002.

Papakonstantinou KG and Shinozuka M (2014) Planning structural inspection and maintenance policies via dynamic programming and Markov processes. Part II: POMDP implementation. *Reliability Engineering & System Safety 130*, 214–224. https://doi.org/10.1016/j.ress.2014.04.006.

Papakonstantinou K, **Andriotis C and Shinozuka M** (2016) Point-based POMDP solvers for life-cycle cost minimization of deteriorating structures. In Bakker J, Frangopol DM and Breugel K v (eds), *Life-Cycle of Engineering Systems: Emphasis on Sustainable Civil Infrastructure*. London: CRC Press, pp. 427–434.

Pearl J (1982) Reverend bayes on inference engines: A distributed hierarchical approach. In Geffner H, Dechter R and Halpern JY (eds), *Probabilistic and Causal Inference: The Works of Judea Pearl*. New York: ACM Books, pp. 129–138.

Phua A, **Davies C and Delaney G** (2022) A digital twin hierarchy for metal additive manufacturing. *Computers in Industry 140*, 103667. https://doi.org/10.1016/j.compind.2022.103667.

Quarteroni A, **Manzoni A and Negri F** (2015) *Reduced Basis Methods for Partial Differential Equations: An Introduction*, Vol. *92*. Cham, Switzerland: Springer.

Rockafellar RT and Uryasev S (2000) Optimization of conditional value-at-risk. *Journal of Risk 2*, 21–42. https://doi.org/10.21314/JOR.2000.038.

Rosafalco L, **Torzoni M**, **Manzoni A**, **Mariani S and Corigliano A** (2021) Online structural health monitoring by model order reduction and deep learning algorithms. *Computers & Structures 255*, 106604. https://doi.org/10.1016/j.compstruc.2021.106604.

Ross S, **Pineau J**, **Chaib-draa B and Kreitmann P** (2011) A Bayesian approach for learning and planning in partially observable Markov decision processes. *Journal of Machine Learning Research 12*(48), 1729–1770.

Russell SJ (2020) *Artificial Intelligence: A Modern Approach*. London, UK: Pearson Education.

Sanchez-Silva M, **Klutke GA and Rosowsky DV** (2011) Life-cycle performance of structures subject to multiple deterioration mechanisms. *Structural Safety 33*(3), 206–217. https://doi.org/10.1016/j.strusafe.2011.03.003.

Seventekidis P, **Giagopoulos D**, **Arailopoulos A and Markogiannaki O** (2020) Structural health monitoring using deep learning with optimal finite element model generated data. *Mechanical Systems and Signal Processing 145*, 106972. https://doi.org/10.1016/j.ymssp.2020.106972.

Sisson W, **Karve P and Mahadevan S** (2022) Digital twin approach for component health-InformedRotorcraft flight parameter optimization. *AIAA Journal 60*(3), 1923–1936. https://doi.org/10.2514/1.J060770.

Sutton RS and Barto AG (2018) *Reinforcement Learning: An Introduction*. Cambridge, MA: MIT Press.

Teughels A, **Maeck J and De Roeck G** (2002) Damage assessment by FE model updating using damage functions. *Computers & Structures 80*(25), 1869–1879. https://doi.org/10.1016/S0045-7949(02)00217-1.

Tezzele M, **Carr S**, **Topcu U and Willcox KE** (2026) Adaptive planning for risk-aware predictive digital twins. In *Physics-Based and Data-Driven Modeling for Digital Twins*. Cham, Switzerland: Springer.

Thelen A, **Zhang X**, **Fink O**, **Lu Y**, **Ghosh S**, **Youn BD**, **Todd MD**, **Mahadevan S**, **Hu C and Hu Z** (2022) A comprehensive review of digital twin – Part 1: Modeling and twinning enabling technologies. *Structural and Multidisciplinary Optimization 65*, 354. https://doi.org/10.1007/s00158-022-03425-4.

Thoft-Christensen P (2009) Life-cycle cost-benefit (LCCB) analysis of bridges from a user and social point of view. *Structures and Infrastructure Engineering 5*, 49–57. https://doi.org/10.1080/15732470701322818.

Thompson PD and Johnson MB (2005) Markovian bridge deterioration: Developing models from historical data. *Structure and Infrastructure Engineering 1*(1), 85–91. https://doi.org/10.1080/15732470412331289332.

Tong GG, **Long CAS and Schiavazzi DE** (2024) InVAErt networks: A data-driven framework for model synthesis and identifiability analysis. *Computer Methods in Applied Mechanics and Engineering 423*, 116846. https://doi.org/10.1016/j.cma.2024.116846.

Torzoni M (2024) Model-based and data-driven methodologies toward predictive digital twins of structures. PhD dissertation, Politecnico di Milano.

Torzoni M and Tezzele M (2026) mtezzele/digital-twin-SHM: v1.0. https://doi.org/10.5281/zenodo.21229439.

Torzoni M, **Manzoni A and Mariani S** (2022) Structural health monitoring of civil structures: A diagnostic framework powered by deep metric learning. *Computers & Structures 271*, 106858. https://doi.org/10.1016/j.compstruc.2022.106858.

Torzoni M, **Rosafalco L**, **Manzoni A**, **Mariani S and Corigliano A** (2022) SHM under varying environmental conditions: An approach based on model order reduction and deep learning. *Computers & Structures 266*, 106790. https://doi.org/10.1016/j.compstruc.2022.106790.

Torzoni M, **Manzoni A and Mariani S** (2023a) A deep neural network, multi-fidelity surrogate model approach for Bayesian model updating in SHM. In Rizzo P and Milazzo A (eds), *European Workshop on Structural Health Monitoring*. Cham, Switzerland: Springer International Publishing, pp. 1076–1086.

https://doi.org/10.1017/dce.2026.10072 Published online by Cambridge University Press

Torzoni M, **Manzoni A and Mariani S** (2023b) A multi-fidelity surrogate model for structural health monitoring exploiting model order reduction and artificial neural networks. *Mechanical Systems and Signal Processing 197*, 110376. https://doi.org/10.1016/j.ymssp.2023.110376.

Torzoni M, **Tezzele M**, **Mariani S**, **Manzoni A and Willcox KE** (2024) A digital twin framework for civil engineering structures. *Computer Methods in Applied Mechanics and Engineering 418*, 116584. https://doi.org/10.1016/j.cma.2023.116584.

Torzoni M, **Manzoni A and Mariani S** (2025) Enhancing Bayesian model updating in structural health monitoring via learnable mapping. *Engineering with Computers 41*, 4953–4975. https://doi.org/10.1007/s00366-025-02224-x.

Torzoni M, **Maisto D**, **Manzoni A**, **Donnarumma F**, **Pezzulo G and Corigliano A** (2026) Active digital twins via active inference. *Engineering Applications of Artificial Intelligence 174*, 114519. https://doi.org/10.1016/j.engappai.2026.114519.

Tzachor A, **Sabri S**, **Richards CE**, **Rajabifard A and Acuto M** (2022) Potential and limitations of digital twins to achieve the sustainable development goals. *Nature Sustainability 5*(10), 822–829. https://doi.org/10.1038/s41893-022-00923-7.

Vlachas K, **Simpson T**, **Garland A**, **Quinn DD**, **Farhat C and Chatzi E** (2025) A reduced order model conditioned on monitoring features for estimation and uncertainty quantification in engineered systems. *Mechanical Systems and Signal Processing 226*, 112261. https://doi.org/10.1016/j.ymssp.2024.112261.

Ye Y, **Yang Q**, **Yang F**, **Huo Y and Meng S** (2020) Digital twin for the structural health management of reusable spacecraft: A case study. *Engineering Fracture Mechanics 234*, 107076. https://doi.org/10.1016/j.engfracmech.2020.107076.

Zakić B, **Ryzynski A**, **Guo-Hong C and Jokela J** (1991) Classification of damage in concrete bridges. *Materials and Structures 24*(4), 268–275. https://doi.org/10.1007/BF02472082.

Appendix A. State-dependent Dirichlet–Multinomial parametrization

In this Appendix, we provide the formulation and update scheme for the general case of a state-dependent Dirichlet–Multinomial transition model. In such a case, $\boldsymbol{\theta}_{d,u} = \left(\theta^1_{d,u}, \dots, \theta^{|\mathcal{D}|-1}_{d,u}\right)$ is defined over the $|\mathcal{D}|-1$ probability simplex. The Dirichlet prior over it is given by:

$$p(\boldsymbol{\theta}_{d,u}) = \frac{1}{B(\boldsymbol{\alpha}_{d,u})} \prod_{k=1}^{|\mathcal{D}|} \left(\theta^k_{d,u}\right)^{\alpha^k_{d,u}-1}, \tag{A.1}$$

where $B(\boldsymbol{\alpha}_{d,u})$ is the multivariate Beta function that ensures normalization:

$$B(\boldsymbol{\alpha}_{d,u}) = \frac{\prod_{k=1}^{|\mathcal{D}|} \Gamma(\alpha^k_{d,u})}{\Gamma\left(\prod_{k=1}^{|\mathcal{D}|} \alpha^k_{d,u}\right)}, \tag{A.2}$$

with $\Gamma(z>0) = \int_0^\infty a^{z-1} e^{-a}\, da$ being the Gamma function.

Let $\mathbf{y}_{d,u} = \left(N^1_{d,u}, \dots, N^{|\mathcal{D}|}_{d,u}\right)^\top \in \mathbb{N}^{|\mathcal{D}|}$ denote the counts of transitions from state d to each state d' under action u, collected from $\{d_{0:t_c}, u_{0:t_c-1}\}$. Each entry is given as $N^{d'}_{d,u} = \sum_{t=0}^{t_c-1} \mathbb{1}_{\{d_t=d, u_t=u, d_{t+1}=d'\}}$, while $N_{d,u} = \sum_{k=1}^{|\mathcal{D}|} N^k_{d,u}$ is the total number of transitions from state d under action u.

The transition counts $\mathbf{y}_{d,u}$ conditioned on $\boldsymbol{\theta}_{d,u}$ follow a Multinomial distribution:

$$\mathbf{y}_{d,u} | \boldsymbol{\theta}_{d,u} \sim \text{Multinomial}(N_{d,u}, \boldsymbol{\theta}_{d,u}), \tag{A.3}$$

$$p\left(\mathbf{y}_{d,u} | \boldsymbol{\theta}_{d,u}\right) \propto \prod_{k=1}^{|\mathcal{D}|} \left(\theta^k_{d,u}\right)^{N^k_{d,u}}. \tag{A.4}$$

By applying Bayes' theorem and substituting the likelihood (A.4) and the prior (A.1), the unnormalized posterior distribution over $\boldsymbol{\theta}_{d,u}$ is given by:

$$p\left(\boldsymbol{\theta}_{d,u} | \mathbf{y}_{d,u}\right) \propto p\left(\mathbf{y}_{d,u} | \boldsymbol{\theta}_{d,u}\right) p(\boldsymbol{\theta}_{d,u}) = \prod_{k=1}^{|\mathcal{D}|} \left(\theta^k_{d,u}\right)^{N^k_{d,u}} \cdot \prod_{k=1}^{|\mathcal{D}|} \left(\theta^k_{d,u}\right)^{\alpha^k_{d,u}-1} = \prod_{k=1}^{|\mathcal{D}|} \left(\theta^k_{d,u}\right)^{\alpha^k_{d,u}-1+N^k_{d,u}}, \tag{A.5}$$

which remains Dirichlet distributed:

$$p\left(\boldsymbol{\theta}_{d,u} | \mathbf{y}_{d,u}\right) = \text{Dirichlet}(\overline{\boldsymbol{\alpha}}_{d,u}), \tag{A.6}$$

with updated concentration parameters $\overline{\boldsymbol{\alpha}}_{d,u} = \left(\overline{\alpha}^1_{d,u}, \dots, \overline{\alpha}^{|\mathcal{D}|}_{d,u}\right)^\top \in \mathbb{R}^{|\mathcal{D}|}$, computed as:

$$\overline{\alpha}^k_{d,u} = \alpha^k_{d,u} + N^k_{d,u}, \quad k = 1, \dots, |\mathcal{D}|. \tag{A.7}$$

https://doi.org/10.1017/dce.2026.10072 Published online by Cambridge University Press

Appendix B. State-independent Beta-Bernoulli parametrization

In this Appendix, we provide the formulation and update scheme for a state-independent Beta-Bernoulli transition model, parameterized by a one-step transition probability $p_u \in \mathbb{R}$. This model constitutes a special case of the Dirichlet–Multinomial framework for state-independent transitions. When applied to the illustrative exasmple introduced in Section 3.1, it yields the following transition matrix:

$$\mathcal{P}_u(p_u) = \begin{pmatrix} 1-p_u & 0 & 0 & 0 & 0 \\ p_u & 1-p_u & 0 & 0 & 0 \\ 0 & p_u & 1-p_u & 0 & 0 \\ 0 & 0 & p_u & 1-p_u & 0 \\ 0 & 0 & 0 & p_u & 1 \end{pmatrix}, \tag{B.1}$$

where the system transitions from state d to the subsequent state d' with probability p_u, and remains in the same state with probability $1-p_u$. Analogously to the Dirichlet–Multinomial case, the online learning procedure for the Beta-Binomial transition model can be formalized as follows:

$$p_u \sim \text{Beta}(a_u, b_u), \tag{B.2}$$

$$\mathbf{y}_u | p_u \sim \text{Binomial}(N_u, p_u). \tag{B.3}$$

Herein, p_u is assigned a Beta prior with shape hyperparameters $a_u, b_u \in \mathbb{R}$ for action u. The vector $\mathbf{y}_u = \left(N_u^0, N_u^1\right)^\top$ collects the observed transition counts $N_u^0 = \sum_{t=0}^{t_c-1} \mathbb{1}_{\{d_t = d_{t+1}, u_t = u\}}$ and $N_u^1 = \sum_{t=0}^{t_c-1} \mathbb{1}_{\{d_t \neq d_{t+1}, u_t = u\}}$; the likelihood term is modeled as a binomial distribution with $N_u = \sum_{t=0}^{t_c-1} \mathbb{1}_{\{u_t = u\}}$. The posterior distribution of p_u is then updated as:

$$p_u | \mathbf{y}_u \sim \text{Beta}\left(a_u + N_u^1, b_u + N_u^0\right). \tag{B.4}$$

While this transition model does not account for multi-step transitions—which can be restrictive for systems with multiple components deteriorating at varying rates—a state-independent one-step transition model can be highly efficient in scenarios characterized by slow deterioration rates.

Cite this article: Varetti E, Torzoni M, Tezzele M and Manzoni A (2026). Adaptive digital twins for predictive decision-making: online Bayesian learning of transition dynamics. *Data-Centric Engineering*, 7, e33. doi:10.1017/dce.2026.10072

https://doi.org/10.1017/dce.2026.10072 Published online by Cambridge University Press