Bayesian Networks for data contextualization in Digital Twins of complex civil infrastructures

Leonardo Binni and Anna Brunetti and Fabrizio Gara and Berardo Naticchia

Polytechnic University of Marche, Italy
l.binni@pm.univpm.it, a.brunetti@pm.univpm.it, f.gara@univpm.it, b.naticchia@univpm.it

Abstract –

The management of complex civil infrastructures must adopt the principles of complex systems, focusing on identifying and modeling emergent behaviors for effective decision-making and predictive insights. Current methods face limitations in cross-scale and cross-domain analyses: process-based models are confined to single domains and fail under uncertainty, while data-driven approaches lack explainability, risking misaligned human-machine decisions. This work addresses this challenge by employing expert-elicited Bayesian Networks (BNs) to enhance the interpretative capabilities of Digital Twins (DTs) of civil infrastructures. BNs enable the modeling of expert knowledge regarding significant cross-domain and cross-scale relationships within complex systems, contextualizing data to develop knowledge and explicitly identify emergent system configurations. A case study focusing on the behavior of a reinforced concrete bridge pier under varying scenarios demonstrates the methodology. Qualitative testing validated the BN model's ability to reflect potential system behaviors and provide predictive insights. This study serves as a foundational step toward the development of civil infrastructure DTs with interpretative and decision-support capabilities. The presented knowledge development process facilitate the recognition of patterns that would otherwise remain undetected or be recognized only following time-consuming expert knowledge integration processes missing operational and emergency response needs.

Keywords –

Bayesian Networks; Data Contextualization; Digital Twin; Civil Infrastructure; Bridges;

1 Introduction

Large-scale civil infrastructures such as road infrastructures are complex systems [1] featuring nonlinear, interdependent interactions among numerous components across different domains and scales. Each component, though autonomous, influences the behavior of others, often resulting in dynamic equilibrium or instabilities that cannot be deduced by merely combining the rules governing individual components. This is not only due to the mathematical intricacy of the problem but also because of the extreme variability in outcomes that can arise from minimal changes or insufficient definition in input data.

Based on the principles governing complex systems, it is evident that the ability to identify and model their emergent behaviors must serve as the foundational basis for any decision-support tools aimed at forecasting within these systems. The contextualization of all the available heterogeneous information on the system (e.g., structural analyses and traffic flow) through cross-domain and cross-scales agent-based modeling becomes essential to develop knowledge [2] about the system and support meaningful high-level diagnosis [3]. It concerns the integration of observations (data) into a model in which both their consequences and potential causes are propagated.

Digital Twins (DTs) of complex civil infrastructure systems must include data contextualization procedures. A DT operates as a virtual shared representation of the physical system and its heterogeneous parts, aligned with it through both virtual and human agents, also exploiting real-time acquired data. This synchronization permits the perception of the system's true state, highlighting aspects that may be difficult to measure directly in the physical context. A DT of a complex system should allow for the observation of its behaviors and, crucially, the early detection of impending changes through the recognition of emergent behavior, potentially allowing for the anticipation of effects via forward simulations across time horizons that are commensurate with the system's inherent instability, typically of a short duration.

With that in mind, in this work we adopt Bayesian Networks (BNs) as a core modeling approach for data contextualization, also in accordance with the concept of human-centered digital systems. This paper therefore introduces the critical transition from information to knowledge through the use of expert-elicited BNs within a civil infrastructure DT framework. BNs are thereby intended as a superstructure of agent modeling that allows data contextualization to explicitly identify emergent system configurations for diagnostic and prognostic purposes.

2 Background

Probabilistic inference involves making conclusions about uncertain events using existing probabilities. This approach is based on probability theory, where results are not certain but rather defined by different levels of probability. Probabilistic inference enables the representation of uncertainty and the forecasting of outcomes or choices based on the likelihood of specific results. Bayesian inference is a method of probabilistic inference that involves using Bayes' theorem to adjust the probability estimate of a hypothesis as additional evidence or data is obtained. Bayes' theorem is rooted in Bayesian inference as it explains the connection between conditional probabilities. It is asserted that the probability $P(H|D)$ of the hypothesis H given the data D, is directly related to the prior probability of the hypothesis, $P(H)$, and the likelihood of the data given the hypothesis, $P(D|H)$:

$$P(H|D) = \frac{P(D|H)P(H)}{P(D)} \tag{1}$$

In this context, the prior indicates our assumptions about the hypothesis before observing the data, while the posterior shows a revised belief following consideration of the evidence.

BNs are a class of probabilistic graphical models representing a set of variables and their conditional dependencies through directed acyclic graphs [4]. BNs are well-suited for modeling expert knowledge in situations where uncertainty is prevalent. The concept of subjective probability, which underpins Bayesian theory [5], is rooted in modern probability theory. Unlike classical probabilistic approaches, which rely on frequentist assumptions, Bayesian probability incorporates beliefs and subjective interpretations (e.g., decision-making under uncertainty). BNs therefore excel at interpreting micro-level evidences at a macro scale by identifying patterns within system’s state transitions. BNs can also enable abductive inferences at a macro level, while offering the whole range of inference capabilities. Therefore, BNs present a versatile and robust approach to modeling the behaviors of complex systems [3] overcoming many limitations of traditional process-based algorithms, and purely data-driven methods as shown in Table 1. Expert knowledge-based BNs effectiveness lies in a set of key strengths that allow them to fully invert the traditional inferences logic and adapt to complex scenarios:

- explainable inferences: BNs are explicit models enabling human-computer shared awareness,
- expert knowledge-based: BNs are uniquely positioned to leverage expert knowledge, transforming experience and intuition (subjective probability – beliefs) into structured conditional probabilities,
- flexibility with input data: BNs can process partial or incomplete data,
- multi-domain and multi-scale variables,
- bidirectional inferences: BNs can perform causal (deductive) and diagnostic (abductive) inferences,
- data-driven compatibility, which helps prevent the introduction of biases caused by the subjectivity of expert knowledge

Building on the strengths of BNs this study integrates them into a DT framework for civil infrastructure diagnostics.

Table 1. Analysis of different modeling methodologies for diagnostics of civil infrastructures.

Models	**Features**
Knowledge-based	Multi-domain and scale variables Explainable inferences Handle uncertainty
Process-driven (physics-based)	Explainable inferences Single domain Cannot handle uncertainty
Data-driven	Multi-domain and scale feature Non-explainable inferences

3 Methodology

It is essential to conceptualize the DT not merely as a conventional simulation tool or as a simplistic virtualization of physical systems, although these aspects are certainly included. The DT possesses significant interpretative and decision-making implications, both from the service (operation and maintenance) and proactively resilience (disasters) management perspectives. The interpretive and decision-making features of a DT, as occurs in the human world, require high-level knowledge about the system under observation. This is necessary to produce meaningful and effective diagnoses. However, high-level knowledge is typically not readily accessible without substantial computational efforts or through human experts. Therefore, DTs must develop knowledge (Figure 1) starting from low-level data and information [2]. AI models can contribute to knowledge development on multiple levels. It must be said, however, that while at a low level data-driven and process-driven models prove useful, at a higher level of the pyramid there is a risk in the former of parallelized human-computer decision-making and in the latter a lack of multi-domain and multi-scale contextualization. Therefore, within the broader framework of a civil infrastructure DT, BN models are hereby intended to serve as high-level multi-scale and multi-domain data contextualization systems.

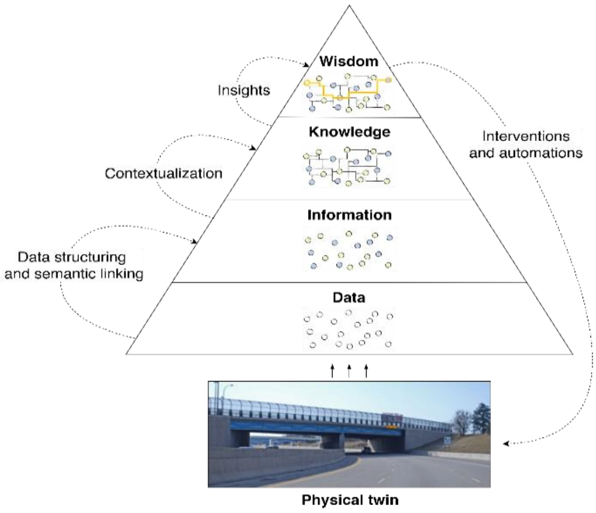


Figure 1. DT's knowledge development processes.

The proposed DT framework builds on a web-based Information Management Framework (IFM) integrating BN models that, by contextualizing a pool of significant evidences and observations through the elicitation of expert knowledge, is used as a bottom-up diagnostic framework for discovering patterns and supporting decisions. Data is collected and can describe events, observations, output from analytical models (e.g., system dynamics, discrete-events simulations, or further agent-based sub-models), real-time sensor data and design and survey data.The IFM taken as reference, and hereafter also generally referred to as DT platform, has been built to support the management processes of assets or systems of heterogeneous assets. It integrates knowledge graphs and unstructured data through a microservice-based architecture with backend and frontend services. The DT platform's backend service acts as an orchestrator of microservices, including a graph database, a document management system, data extraction, integration and query services, a REST API interface and a message broker. The BN models can be integrated with the DT platform as web-based microservices, enabling real-time inferences (Figure 2).

A Bayesian inference service was therefore developed using the Hugin Expert API REST web service [6]. It provides a Bayesian inference process server that can be queried via REST API calls to provide evidence to the network or retrieve the beliefs of nodes of interest after propagating the evidence. BN models can either be loaded or created directly within the web service by instantiating nodes and relationships through the Hugin REST API.

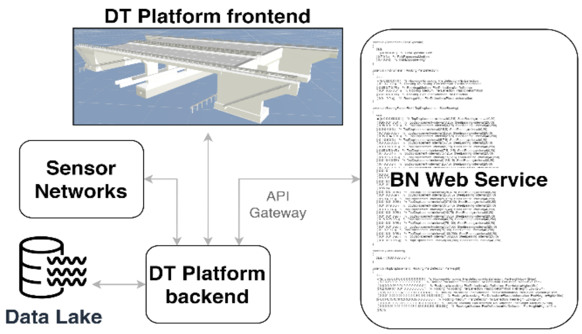


Figure 2. Schematization of the DT platform integrating the BN web service as a virtual sensor.

3.1 Case study implementation

To demonstrate the methodology, this section presents the development of a BN model representing the behaviors of a Reinforced Concrete (RC) bridge. The bridge is considered to have a simply-supported span scheme (Figure 3a). Specifically, the model focuses on the representation of high-level reasoning about the behavior of a single bridge pier, which under different actions is subject to different amplitudes of displacement at its top (Figure 3b), and may experience varying levels of bearing failure. The case study is not properly a large-scale complex infrastructure system by itself, but rather a significant part of it, given the number of multi-domain variables that influence its behavior. This case study is selected for its comprehensibility, facilitating the discussion and giving an understanding of the proposed approach along with its applicability and scalability to wider civil infrastructure systems. Key considerations include the following:

- expert knowledge was elicited to identify key variables, states, and interactions across scales and domains through five iterative 3-hour consultation session among the authors of this paper. .
- the model includes both static data (e.g., design information or data from surveys) and dynamic inputs (e.g., data from inspections, such as newly identified damages, analytical models like acceleration data analysis, or sensor data such as tilt values from inclinometers). This flexibility extends to real-time data integration,
- the BN can be updated with new evidence as it becomes available, ensuring its significance and accuracy over time,
- the BN model was integrated with the DT platform as explained in Section 3 and its output can directly

feed virtual agents as well as trigger actions from human experts,

- scalability and modularity: the network lies on a "plug-and-play" architecture, allowing for expansion and integration with other networks through loose connections, where significant. This modularity enables the development of multiple networks tailored to address further multi-domain and multi-scale interactions,
- for testing purposes, the BN was modeled using Hugin Expert [6] and exported as a HKB file.

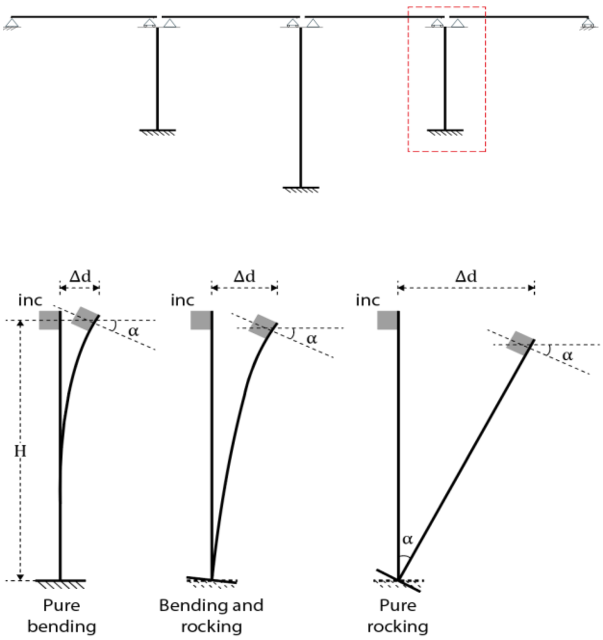


Figure 3. Structural scheme of the bridge pier taken as reference for the case study. The same inclination (α) causes different top displacements (Δd) depending on the pier's behavior.

3.1.1 Analysis of significant variables

An analysis of the significant variables was performed following the *Italian guidelines for existing bridges - classification and management of risk, safety assessment and monitoring* [7]. Given the quantity of the variables involved, a study has been performed to consider only those most appropriate for the purpose of the model presented: variables describing dynamic evidences (e.g., sensor readings or field inspection findings), context conditions (e.g., design data), and behaviors or phenomena (e.g., bending and rocking of piers, or bearing failure). The list of variables considered in the proposed model is shown in Table 2. It is not exhaustive of all possible variables, whose integration within the present model would go beyond the scope of this work. Therefore, the following key assumptions have been made:

- the model schematizes the behavior of a single circular pier of diameter **Ø** and height $\boldsymbol{H}$ in the longitudinal direction of the bridge. The pier is modelled as a fixed-base beam element free to longitudinally move at the top, involving its bending deformability (Figure 3b),
- the bearing can be either a steel or a neoprene support device (Figure 3a),
- the bridge is equipped with accelerometers, and a tilt sensor is positioned at the pier top to measure inclination α (Figure 3b),
- measurements of the relative displacement between the deck and the bearing (Δd) could be measured with displacement transducers. In some case, it might not be available,
- top displacement is caused by bending and/or rocking of the pier in the longitudinal direction.

Since the objective of this research is to validate the digital twinning methodology, these assumptions were necessary to define a specific use case. However, they do not constrain the model's updating with additional variables, limit its application scope and scalability, or compromise the network's effectiveness as a framework for contextualizing data and perform system diagnoses. Additional nodes (variables) and edges (relationships) can be modeled to capture currently omitted variables, interactions and system behaviors.

3.1.2 Model definition

The bridge pier system is described through the modeling of the variables presented in Table 2 as nodes of the network (Figure 4), and through the expert elicitation of conditional probabilities (arcs) in the Conditional Probability Tables (CPTs), which define the relationships between variables. For what concerns the variables, their states were discretized based on values deemed significant for the purposes of the model. For example, continuous variables like inclinometer readings were discretized into three ranges to simplify probabilistic modeling while capturing key behavioral thresholds that correlate with varying damage severities. Regarding the CPTs, an example is provided in Table 3. The table should be read as follows: in a scenario with "No rocking" and "No bending", the bridge pier is most likely (99%) experiencing a low level of inclination (0-0.005 rad). The probabilistic quantification of the network is achieved by specifying the CPTs for each node. For initial nodes (those without parent nodes), this quantification is assigned as prior probabilities. These priors are derived from statistical considerations when available, or by assigning equal probability to each interval into which the variable's domain is divided in cases of total uncertainty. The network's development reflected the paradigm that effects are conditioned by their causes. Therefore, the BN was designed with a hierarchical structure where variables that represent potential dynamic evidences $\{\alpha, B_d, g\}$ are conditioned by the variables representing phenomena and behaviors $\{R_v, S_f, D_p, \theta, \Delta d, B_{f1}, B_{f2}\}$ that act differently depending on the context condition variables $\{R_e, FT, Ag, K_y, \phi, H, AR, B_s, B_n\}$. The complete model is illustrated in Figure 4.

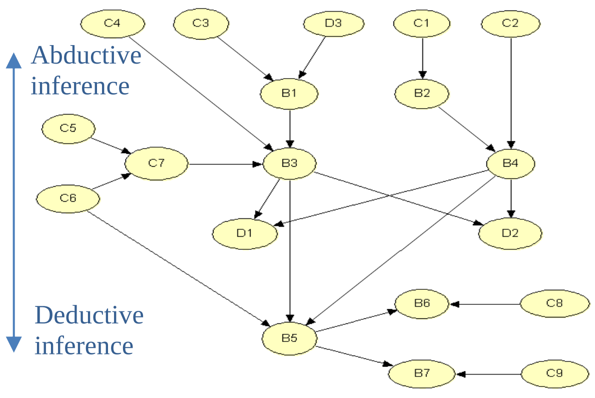


Figure 4. Visualization of nodes (variables) and relationships within the developed BN.

3.1.3 Model validation

Validating an expert-elicited BN involves determining whether the analysis accurately represents the specific concepts it is intended to model for its designated purpose. As a reflection of a mental construct designed to address possible consequences, which may or may not occur, a direct comparison between the BN's model outputs and observations from the real-world system is not feasible. In the context of this study, a comparison with historical data would be inconclusive given the little quantity of available data. However, validation can extend beyond comparisons with observed data by evaluating the internal consistency and logic of the model. Such approaches are widely applied in social science research, system dynamics modeling, and expert-driven BN modeling. A range of specific tests can be performed to assess whether the model meets predefined criteria [5]. Among those, predictive validity tests for BNs can be considered to encompass both the model behavior and the model output. Of particular relevance to establishing confidence in the predictive validity of BNs are Behavior Sensitivity Tests such as the Qualitative Features Tests (QFT). In this test the model response is evaluated in terms of a qualitative understanding on how the system is believed to respond under certain conditions and the sensitivity of the variables. This also helps unearth potential biases introduced by the modeler during the expert knowledge elicitation phase, including those caused by subjective judgment. The QFT is performed by triggering certain states of the BN-variables as inputs, and by inspecting the corresponding model response. The QFT test was performed in this study to qualitatively evaluate the BN and its consistency. Further testing will be performed in future developments, especially to assess the potential bias introduced by the heavy reliance on expert knowledge [5]. In this work, the test was conducted by two domain experts, distinct from those who contributed to the BN modeling phase. The experts have a grounded background in civil engineering with field experience related to the inspection and safety evaluation of existing bridges, obtained by conducting numerous back-office inspections and assessments following the requirements of state-of-the-art Condition Rating Systems [7]. The small validation sample size stems from the specificity of the developed model, which requires civil engineers with proven and established experience in the field of reinforced concrete bridge monitoring and inspection. This may partially limit the comprehensiveness of qualitative validation. However, this limitation does not impact the results of this research that aims at validating a digital twinning methodology. It is part of the future developments of this research to expand the validation sample by including a larger number of domain experts in the validation phase.

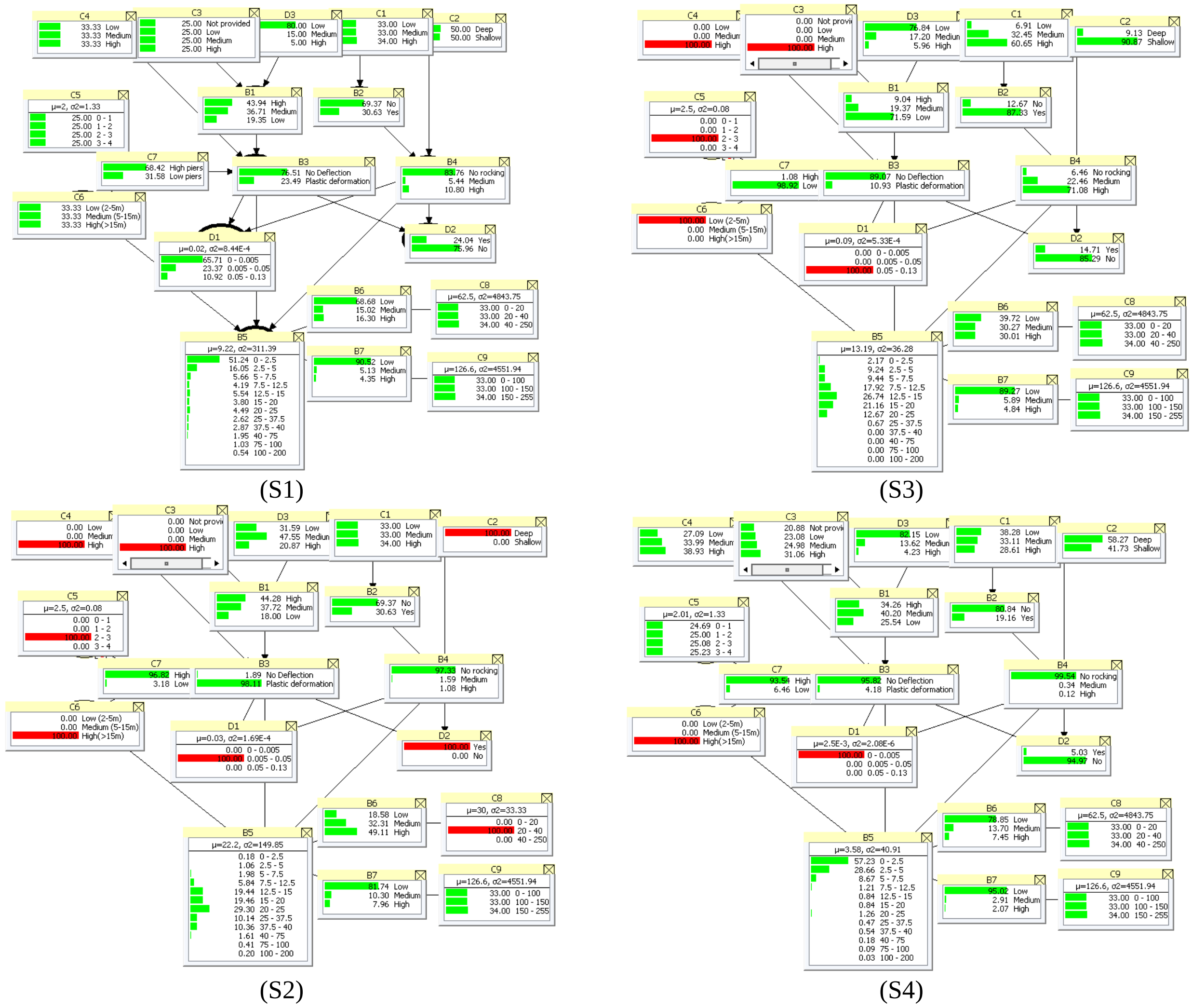
Table 2. Variables used to describe the bridge pier system under consideration.

Variable	Symbol	Node	Evidence from	Discretization
Inclinometer	α	D1	Inclinometer sensor data	{0-0.005, 0.005-0.05, >0.05} [rad]
PiersBaseDamage	B_d	D2	Inspection information	Yes/No
EarthquakeLevel	g	D3	accelerometer data	Low/medium/high
Risk exposure	R_e	C1	Expert assessment	Low/medium/high
FoundationType	FT	C2	Design/survey information	Deep/shallow
DesignGroundAcc	Ag	C3	Design information	Not provided/low/medium/high
RC_Yield_Curvature	K_y	C4	Design information	Low/medium/high
PierDiameter	ϕ	C5	Design/survey information	{0-1, 1-2, 2-3, 3-4} [m]
PierHeight	H	C6	Design/survey information	{2-5, 5-15, 15-40} [m]
AspectRatio	AR	C7	Design/survey information	Slender/Non-slender

SteelBearing	B_s	C8	Design/survey information	{0-20, 20-40, 40-250} [mm] [8]
NeopreneBearing	B_n	C9	Design/survey information	{0-100, 100-150, 150-255} [mm] [9]
BridgeSeismicHazard	R_v	B1	-	Low/medium/high
Foundation Settlement	S_f	B2	-	Yes/No
Bending	D_p	B3	-	No bending/plastic deformation
Rocking	θ	B4	-	No rocking/medium/high
TopDisplacement	Δd	B5	-	{0-2.5, 2.5-5, 5-7.5, 7.5-12.5, 12.5-15, 15-20, 20-25, 25-37.5, 37.5-40, 40-75, 75-100, 100-200} [mm]
BearingFailure	B_{f1}	B6	-	Minor/moderate/extensive [8]
BearingFailure2	B_{f2}	B7	-	Minor/moderate/extensive [9]

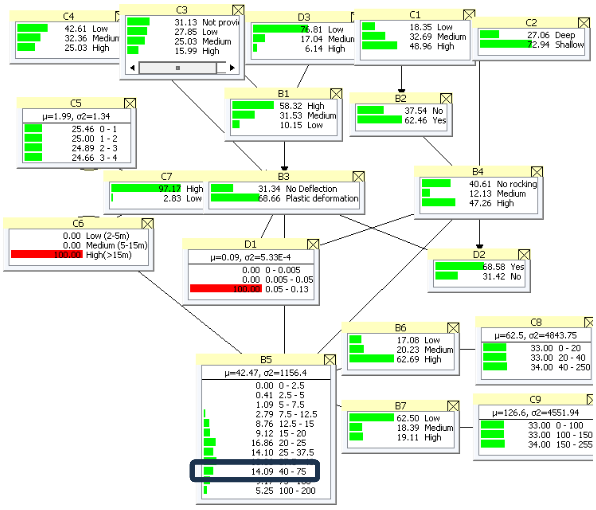
Table 3. Conditional Probability Table (CPT) of the node D1.

B4	No rocking		Medium		High	
B3	No bending	Plastic deformation	No bending	Plastic deformation	No bending	Plastic deformation
0 - 0.005	0.99	0.1	0.05	0.01	0.01	0.0
0.005 - 0.05	0.01	0.7	0.7	0.65	0.6	0.1
> 0.05	0.0	0.2	0.25	0.34	0.39	0.9

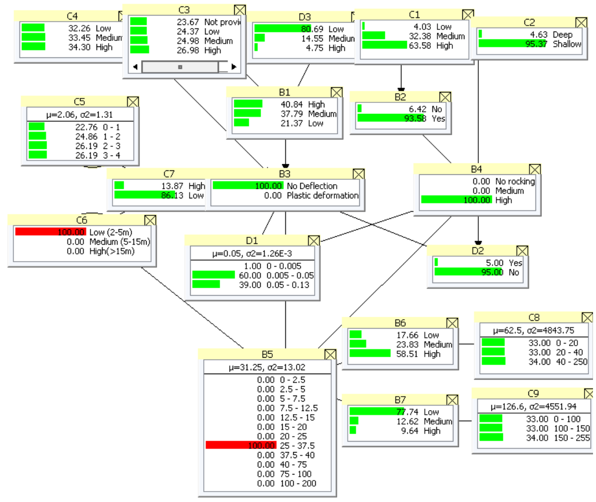


(S1) (S3)

(S2) (S4)



(S5)



(S6)

Figure 5. Tested scenarios (S1 to S6). Input variables for each scenario are highlighted in red.

Table 4. QFT results.

Test ID	Plausibility assessment
S1	The model schematizes a general conditions in which higher risks are not very likely, which demonstrates a good degree of generalization.
S2	A deep foundation implies a lower chance of rocking. Evidence of bending damage at the base of the pier implies a flexural plastic deformation even though the bridge was designed with appropriate standards. The inclination observed is therefore only given by bending, causing a medium level of displacement at the top of the pier (Δd). A steel bearing would experience extensive damages with high probability. Of interest is also the abductive reasoning process related to the potential causes of such damage: the beliefs of earthquake intensity grew in its *Medium* and *High* states compared to the baseline.
S3	The height of the pier is considered to be *Low (2-5m)*. The measured inclination angle therefore is most probably caused by rocking instead of bending. It implies that the belief of the state *No* of the *PiersBaseDamage* variables is higher than 85%. The significant inclination angle implies a *High* level of rocking. Nevertheless, the displacement on top seems to be not much influenced. This is plausible since the pier is low and displacement Δd grows along with the pier's height H. The believed displacement Δd would be critical in case of steel bearing causing at most moderate damages.
S4	Non-significant tilt angle is observed, bending and rocking effects are strongly believed to be absent. It implies that the displacement on top of the pier is low. Low or no risks for the bearings.
S5	Extremely significant tilt angle is observed, the model is predicting a combined bending-rocking scenario. Bending is more likely than rocking. Indeed the probability is higher, since there is a 50 % chance that rocking is not present. Displacement on top of the pier correctly increases and becomes significant. High risk of failure for the steel bearings, whereas for the neoprene ones additional evidences are required.
S6	Assuming that the displacement has been measured in a range between 20-25 mm, a low pier is 100% experiencing a significant rocking effect. The reason behind that could be the exposure of the pier and its foundation type, which are respectively believed to be most probably *High* and *Shallow*.

3.1.4 Querying the model

The model was integrated with the DT platform as explained in Section 3. The BN was then queried via the REST API service for testing the following processes:

- loading the developed BN model,
- assigning and propagating evidence after new observations,
- retrieving the beliefs of nodes of interest after evidence propagation.

Table 5 shows an example of a test query aimed at retrieving the a posteriori belief of one of the states of Node B5 after evidences in Nodes D1 and C6 are triggered in Scenario n. 5. The Response Body shows a belief of 14.09% for a top displacement in a range from

40 to 75 mm, which corresponds to the one outlined in Figure 5 (S5).

Table 5. Example of a query to the BN model.

Get the a posteriori belief of Node B5 {40-75 [mm]} in Scenario n. 5

GET
https://[redacted]/rest/domain/001210-0000-0000-0000-00000/node/00001210-0000-0000-0000-0000000008/getBelief?=9:
{…
"Response Headers": {
"content-type": "text/plain;charset=utf-8",
"content-length": "11",
"server": "Jetty(9.2.z-SNAPSHOT)"
},
"Response Body": "14.09374811"
}

4 Conclusions

This study aims at enhancing the management of complex civil infrastructures by integrating expert-elicited Bayesian Networks (BNs) with Digital Twins (DTs), offering a robust and scalable approach to modeling multi-scale and -domain emergent behaviors.

While the traditional process-based and data-driven models prove useful for data processing, this study offers a solution to the problem of developing high-level knowledge in DTs by endowing them with interpretive value. This is achieved through the development and integration of predictive and explainable BN models. BNs model expert knowledge, enabling the contextualization of heterogeneous multi-domain and multi-scale data (e.g., sensor data streams, observed damages, and design specifications) that are otherwise analytically non-integrable. This contextualization process, capable of assimilating significant new evidence as it arises—even in real time—facilitates the recognition of dynamic system configurations relevant to operational and emergency management needs. In this way, the proposed DT framework functions as a diagnostic and prognostic tool for civil infrastructures, including large-scale systems. It digitizes and accelerates the traditionally slow processes of expert knowledge integration.

The efficacy of this approach was demonstrated through a case study focusing on the dynamic behaviors of reinforced concrete bridge piers. A BN was modeled eliciting the authors' knowledge. The model was first integrated with a web-based DT platform through its REST API interfaces and then validated through Qualitative Features Tests conducted with two external domain experts.

The capacity to bridge human expertise with DT-based diagnostics ensures a cognitive human-machine collaboration that enhances a shared understanding of system behaviors, reflecting the concept of collective intelligence that characterizes complex systems.

Future developments will focus on the extension of the BN model, on its further validation through additional qualitative and quantitative testing and on the integration of the BN web service with the platform backend through a message broker. This would ensure seamless push communication between the service and the backend of the DT platform enabling real-time queries.

References

[1] S. Flumerfelt, K.G. Schwartz, D. Mavris, S. Briceno, Complex Systems Engineering: Theory and Practice, American Institute of Aeronautics and Astronautics, Inc., Reston, VA, 2019. https://doi.org/10.2514/4.105654.

[2] M. Grieves, DIKW as a General and Digital Twin Action Framework: Data, Information, Knowledge, and Wisdom, Knowledge 4 (2024) 120–140. https://doi.org/10.3390/knowledge4020007.

[3] M. Soleimani, F. Campean, D. Neagu, Diagnostics and prognostics for complex systems: A review of methods and challenges, Qual Reliab Eng Int 37 (2021) 3746–3778. https://doi.org/10.1002/qre.2947.

[4] J. Pearl, Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference, Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1988.

[5] J. Pitchforth, K. Mengersen, A proposed validation framework for expert elicited Bayesian Networks, Expert Syst Appl 40 (2013) 162–167. https://doi.org/10.1016/j.eswa.2012.07.026.

[6] A.L. Madsen, F. Jensen, U.B. Kjaerulff, M. Lang, The Hugin Tool for Probabilistic Graphical Models, International Journal on Artificial Intelligence Tools 14 (2005) 507–543. https://doi.org/10.1142/S0218213005002235.

[7] J.C. Matos, V. Nicoletti, J. Kralovanec, H.S. Sousa, F. Gara, M. Moravcik, M.J. Morais, Comparison of Condition Rating Systems for Bridges in Three European Countries, Applied Sciences 13 (2023) 12343. https://doi.org/10.3390/app132212343.

[8] J.B. Mander, D.-K. Kim, S.S. Chen, G.J. Premus, Response of Steel Bridge Bearings to Reversed Cyclic Loading, 1996.

[9] R.A. Cook, D.T., Allen, M.H. Ansley, Stiffness Evaluation of Neoprene Bearing Pads Under Long-Term Loads, 2009.