DIAGNOSTYKA, 2026, Vol. 27, No. 2

e-ISSN 2449-5220
DOI: 10.29354/diag/224917

STATISTICAL AND CONTEXT-BASED DIAGNOSTICS FOR DIGITAL TWIN ENVIRONMENTS

Artur KRÓL [1], Anna TIMOFIEJCZUK [2,*], Tomasz ŁUKASIK [1]

[1] Tenneco Inc. ul. Usługowa 20, 44-100 Gliwice, Poland
[2] Department of Fundamentals of Machinery Design, Silesian University of Technology, ul. Konarskiego 18A, 44-100 Gliwice, Poland
* Corresponding author, e-mail: anna.timofiejczuk@polsl.pl

Abstract

The Digital Twin concept is increasingly applied in engineering practice to support simulation-based design and validation processes. However, ensuring consistency between simulations and experimental data remains a critical challenge. In engineering systems identical statistical patterns can originate from different causes, so it may lead to incorrect diagnostic conclusions. This paper proposes an approach that integrates statistical validation methods with context-based reasoning to improve the interpretation of data in Digital Twin environments. The proposed framework combines data preprocessing techniques with statistical analysis and context-based reasoning. The applicability of the approach is shown through a representative case study. The results confirm that the integration of statistical methods with context-based reasoning enhances the robustness of Digital Twin systems and supports more effective use of simulation in modern engineering environments.

Keywords: Digital Twin, diagnostics, context-based reasoning, statistical validation, Industry 4.0

1. INTRODUCTION

The Digital Twin concept is increasingly applied in modern engineering to support simulation-based design and validation processes. However, its effectiveness strongly depends on the ability to ensure reliable agreement between simulation results and physical tests. In industrial practice, this correlation is difficult to achieve due to the complexity of systems, variability of operating conditions, and inconsistencies in data acquisition processes [13]. Identical statistical patterns may originate from fundamentally different causes, such as modelling assumptions or measurement conditions. Information describing design assumptions, operating conditions, and testing procedures provides contextual information that constitutes additional knowledge necessary for correct interpretation. Previous studies have shown that incorporating such information significantly improves the quality and interpretability of diagnostic processes [20], [21].

The main contribution of this paper is the integration of statistical validation methods with context-based reasoning, providing a structured approach to interpreting simulation and experimental data in Digital Twin environments. The proposed approach enables differentiation between model-related, measurement-related, and physical deviations, supporting more reliable engineering decision-making.

1.1. Transformation of automotive industry

The automotive industry is undergoing rapid transformation driven by electrification, digitalization, increasing system complexity, and the transition towards Industry 4.0 [7], [21]. Digital Twins support this transformation by enabling integration of simulation models with physical tests and improving the consistency of engineering analysis. Their growing importance in the automotive sector has been highlighted in recent studies [16]. However, despite the increasing availability of simulation tools, the interpretation of results remains a significant challenge.

In practice, engineering data are generated under varying operating conditions and influenced by multiple factors related to design, environment, and testing procedures. Without considering this broader context, the interpretation of this data may lead to incorrect conclusions. Therefore, the development of context-aware approaches becomes a necessary step.

1.2. Struggles of implementing Digital Twins

Digital Twin-based methods constitute an important element of Industry 4.0 systems 4.0 [3], [4], [6], [21]. Particularly relevant is improvement of simulation processes. It increases model fidelity and

Received 2026-04-07; Accepted 2026-06-22; Available online 2026-06-23

enhances the efficiency of product development. The implementation of Digital Twins is associated with both technical and organizational challenges.

In industrial practice, engineering processes involve multiple data sources, tools, and procedures, which makes it difficult to ensure data consistency and reliability across different stages of the product lifecycle [13]. These difficulties are further reinforced by the broader digital transformation of engineering organizations and the need to adapt business models and operational strategies [17].

A critical issue in this context is the lack of structured mechanisms for interpreting data in relation to their origin and conditions of acquisition. Simulation results and experimental data are often analyzed without sufficient consideration of contextual factors, which may lead to misinterpretation of discrepancies and incorrect diagnostic decisions [22], [23].

Effective implementation of the Digital Twin concept therefore requires not only the integration of data and tools, but also the incorporation of contextual knowledge into diagnostic processes. Such an approach enables more accurate identification of deviation sources and supports reliable validation of simulation models in complex engineering environments [8], [10].

1.3. Digital Twins supporting the design and development process

The ongoing industrial transformation associated with Industry 4.0 has introduced a broad range of intelligent solutions into manufacturing processes, enhancing automation, efficiency, and data integration [1], [4]. These developments are linked to the concept of cyber-physical systems, which form the technological foundation of modern Industry 4.0 architectures [12]. The emergence of such systems has introduced new challenges related to data integration, model fidelity, and system adaptability [14], [21]. Beyond manufacturing, these advancements increasingly affect product lifecycle activities, particularly in the areas of design and development, where Digital Twins have emerged as one of the most promising concepts [8], [11].

A Digital Twin enables the creation of a high-fidelity virtual representation of a physical asset, allowing engineers to replicate, simulate, and analyze real-world behavior in a digital environment. State-of-the-art studies indicate that these technologies are becoming a key industrial paradigm for integrating models, data, and lifecycle knowledge [20]. In this context, simulation plays a central role, enabling predictive analysis and system optimization in engineering applications [19].

Within the design and development process, a Digital Twin is typically realized through a simulation model that is calibrated against physical test results.

Achieving this requires iterative refinement of the model, including adjustment of material parameters, boundary conditions, and model assumptions [15], [19]. Model calibration against experimental data is therefore not a one-time validation step, but an ongoing process involving continuous model improvement and verification against real-world observations. This process ensures that the Digital Twin remains a reliable representation of the physical system under varying operating conditions, as illustrated in Fig. 1.

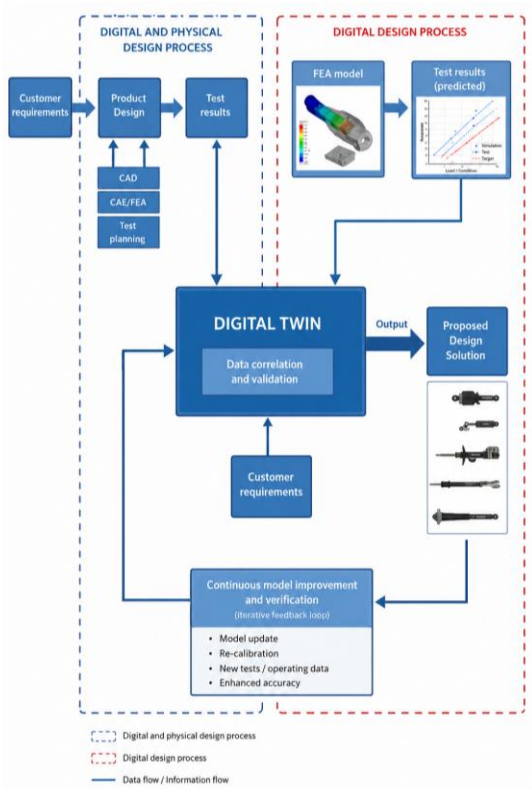

Fig. 1. Digital Twin implementation model with an iterative model calibration and continuous improvement loop

The feedback loop shown in Fig. 1 reflects the iterative nature of Digital Twin development, where simulation models are continuously updated and validated using experimental observations collected throughout the product lifecycle.

However, in practical engineering environments, discrepancies between simulation and experimental data are unavoidable. They may result not only from modelling inaccuracies, but also from variations in operating conditions, measurement procedures, and data acquisition processes. As a result, the interpretation of simulation results cannot rely solely on numerical agreement or statistical consistency.

Therefore, the correct interpretation of Digital Twin outputs requires the incorporation of contextual information describing the conditions under which both simulation and experimental data are generated. This includes design assumptions, operational conditions, and testing procedures, which together influence the observed system behavior. Without considering this broader context, the correlation between simulation and experimental results may be misinterpreted, leading to incorrect conclusions regarding model accuracy.

Only parameters relevant to the analysis should be considered. In cases of limited knowledge, Design of Experiments (DOE) may be applied to identify key input factors and define the design space. Correlation methods should be supported by statistical techniques such as the one-sample t-test and DOE-based ANOVA. However, as discussed in this paper, statistical methods alone are not sufficient to explain discrepancies, which further justifies the need for context-based diagnostic approaches.

2. ORGANIZATIONAL CONSIDERATIONS IN DIGITAL TWIN IMPLEMENTATION

The implementation of Digital Twin solutions requires not only technological readiness but also proper structuring of engineering processes and data. From a diagnostic perspective, the most critical aspect is the identification, consistency, and traceability of datasets exchange.

Process mapping techniques, such as swimlane diagrams, enable identification of inefficiencies related to data fragmentation and lack of integration. In many industrial environments, the same data are duplicated across multiple systems, leading to inconsistencies and reduced data quality. This directly affects the reliability of diagnostic analysis. Therefore, a structured approach to data mapping is required, ensuring that datasets are generated once and consistently reused across the entire engineering workflow. These organizational aspects directly influence the quality and interpretability of diagnostic processes in Digital Twin environments.

2.1. Process mapping for digital twin implementation

The implementation of Digital Twins, similarly to other complex IT-based systems, requires a deep understanding of existing business processes, which are critical to successful system integration [5], [6]. Without clear insight into operational workflows, data exchange points, and cross-departmental dependencies, even the most advanced technologies are at risk of underperformance or misalignment with actual organizational needs.

To support the analysis established methodologies from Lean Manufacturing, such as Value Stream Mapping (VSM) and Swimlane Diagrams, can be employed [2], [3]. These tools help visualize the sequence of activities, responsibilities across functional units, and data flow constraints. In the the paper, a Swimlane Diagram was used to illustrate the real-world product design and development workflow for a component of an automotive suspension system. The resulting process map is presented in Fig. 2 and reveals several critical inefficiencies.

Each department involved in the process operated within a semi-isolated environment, supported by independent ticketing systems for managing project requests and task assignments. As a result, data duplication and fragmentation occurred, increasing the risk of inconsistencies and reducing transparency. Lack of standardized nomenclature further increased the risk of misinterpretation. Ultimately, the lack of integration reduces data quality and process efficiency.

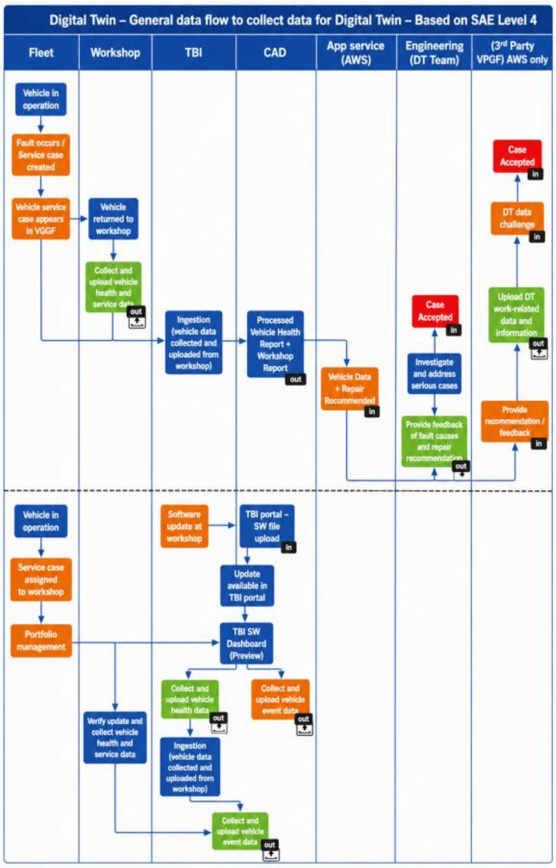

Fig. 2. Swimlane diagram of the design and development process based on a pilot project

2.2. Data mapping in Digital Twin implementation

The subsequent stage in the implementation of a Digital Twin involves the identification and tracking of datasets that are exchanged between individual process steps within the IT infrastructure. A widely recommended best practice is to develop data-process matrices, which align each recognized dataset with the corresponding process stages. An example of such a matrix is presented in Fig. 3.

This structured approach helps visualize how and where specific data points are generated, consumed, and transferred throughout the workflow. Data-driven approaches to Digital Twin implementation emphasize the importance of structured data integration and traceability across process stages [18]. The importance of real-time data acquisition and continuous updating of Digital Twin models has also been demonstrated in production system applications [24].

Referring to the process flow illustrated in the Swimlane Diagram, it becomes evident that datasets are often fragmented and siloed, with each

Generic XBOM (PV)	Specific XBOM	Material Spec 'database'	3D models	2D drawings	Proto request	Prototyping	Test Request
GPDM	GPDM		CATIA	CATIA	DA3	DA3	DA3
	Part Number				XBOM		
Material							
Type							
Material Specification #		Material Specification #					
Thickness [mm]							
	Part Number						
Material Specification #		Material Specification #					
RT thickness under Stabi Bracket [mm]							
RT OD under Stabi Bracker [mm]							
	Part Number						
Number of welds							
Weld Run Out							
Weld Configuration							
			DIR (3D)				
				DIR (2D, spec)			
					SB number		
					SB entity		
						XS	
							TR
							cycles
							loads
							criteria

Fig. 3. Datasets identification

department generating and storing data independently. Consequently, it is essential to map each dataset precisely, identifying its origin, format, labeling conventions, and subsequent usage across process stages. It also constitutes part of the operational context. The objective is to ensure that datasets are generated once and consistently reused in a given context across the workflow. In addition to improving traceability, data mapping supports the creation of a common information space shared by simulation engineers, test engineers, and product development teams. Such a unified view of data reduces the risk of inconsistent interpretations and facilitates the implementation of Digital Twin solutions across organizational boundaries. From a diagnostic perspective, data mapping also improves the transparency of relationships between data sources and engineering decisions, enabling more reliable identification of discrepancy origins. This minimizes redundancies, enhances traceability, and increases the overall efficiency and robustness of the Digital Twin environment, and ensure correct interpretation of the data.

3. DIAGNOSTICS BASED ON DIGITAL TWIN IMPLEMENTATION

Digital Twin environments generate large datasets from simulations and experiments, requiring interpretation. Diagnostic approaches based on data analysis and system modelling have been successfully applied in complex technical systems [25]. Preliminary studies indicated the importance of diagnostic and qualitative aspects in Digital Twin implementation processes [9].

One of the critical challenges is establishing a consistent and standardized nomenclature for variables, parameters, and test results. In organizations, especially these operating across multiple geographic locations, it is common to observe non-unified local standards and terminologies. This disparity often leads to miscommunication, redundancy, and errors. The root of the issue is often a lack of global alignment in data protocols and taxonomy.

To address this, it is essential to incorporate well-defined diagnostic frameworks into the software platforms supporting Digital Twins. These frameworks must include not only methods for data validation, filtering, and clustering, but also mechanisms for fault detection, anomaly recognition, and pattern analysis based on predefined quality thresholds.

Moreover, the selection and implementation of the appropriate diagnostics tools must be tailored to the specific application domain and the nature of the virtual-physical correlations required. When well-structured, such diagnostics provide not only verification of system behavior but also insights into failure modes, allowing for predictive and even prescriptive maintenance interventions.

3.1. Correlation methods

In Digital Twin environments, the evaluation of consistency between simulation outputs and experimental observations requires systematic diagnostic procedures. Rather than relying solely on direct comparison, it is necessary to assess the consistency, variability, and statistical significance of the observed differences. This involves the application of statistical methods that enable identification of deviations, outliers, and inconsistencies within datasets. Such analysis provides a quantitative basis for evaluating model accuracy, but it does not fully explain the origin of discrepancies, which may arise from modelling assumptions, measurement conditions, or actual system behavior.

An example of such a graphical correlation study is illustrated in Fig. 4. This figure presents the simulation output (FEA result) alongside a series of test samples. The visual comparison is crucial for identifying trends, deviations, or inconsistencies between physical and virtual results. It also highlights discrepancies that might not be immediately visible through statistical summaries alone.

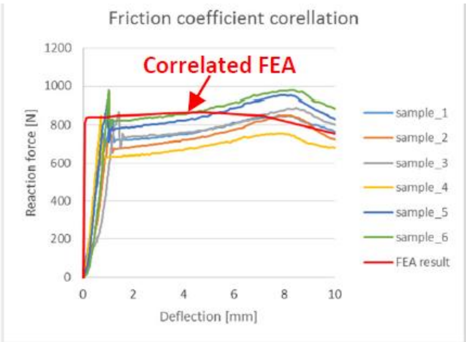

Fig. 4. Graphical correlation study

Moreover, this type of analysis provides insights not only into the validity of the simulation model, but also into the quality of the physical test procedure itself. For example, any oscillations or irregularities in the physical test curves may indicate systematic errors, such as improper mounting of components or variability in test bench setup. In this way, Digital Twins can indirectly support diagnostics of test execution quality.

Once the graphical correlation is established, it becomes necessary to apply statistical methods to quantify the degree of agreement between physical and digital data. An example of a widely used technique is the one-sample t-test, which compares the simulation result (treated as a target value) to a set of physical test results. This test provides information on whether the simulation lies within the expected statistical range of the physical outcomes. If the p-value of the test is above the predefined confidence threshold (e.g. 0.05), the simulation can be considered statistically equivalent to the physical test results.

Fig. 5 presents an example of a one-sample t-test, where the simulation output is positioned relative to the confidence interval of the physical test data.

This visualization provides an intuitive way to evaluate the quality of the correlation and support objective decision-making in simulation validation processes. However, these methods do not provide information about the underlying causes of observed discrepancies.

3.2. Identification of outliers

When analyzing test results related to product lifetime or durability, it is essential to apply appropriate statistical methods to interpret the collected data correctly. Furthermore, lifetime test analysis provides a means to verify the failure mode and the number of cycles to failure—two key parameters used to calibrate and update Digital Twin models for predictive maintenance applications. By integrating such statistical feedback, Digital Twins can continuously improve their accuracy and predictive capabilities.

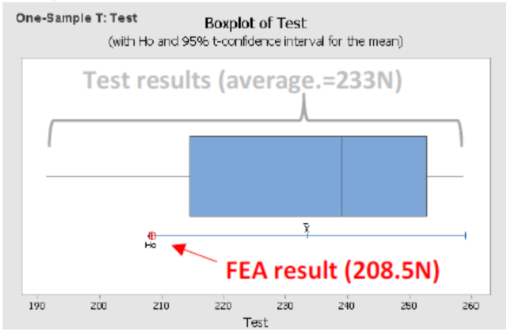

Fig. 5. 1-sample T test example

Among the most widely used techniques for lifetime analysis are the Weibull distribution and the Wöhler (S–N) method, both of which are foundational in fatigue and reliability studies. These methods provide insights into the variability and predictability of a component's performance under repeated loading or operational stress conditions. Both methods rely on confidence intervals, typically defined by industry standards (e.g. 90% or 95%), to assess statistical validity.

A crucial part of the analysis involves identifying and investigating data points that fall outside the defined confidence interval. Such points should not be disregarded automatically; instead, they should trigger a detailed root-cause analysis to determine whether they represent legitimate variations in material behavior or are the result of testing errors. Possible causes include testing errors, improper setup, or environmental conditions.

In the aspect of Digital Twins, these analyses are particularly valuable, as they allow the engineer to assess not only the quality of physical test data but also the fidelity of the virtual model. A mismatch between simulation and experimental data, especially in the tails of the distribution, may indicate that the simulation model lacks sufficient realism in modeling material fatigue or boundary conditions.

The curve illustrates the relationship between stress amplitude and the number of cycles to failure, showing the expected statistical scatter of experimental data and the boundaries within which most test points fall.

An example of a Wöhler curve (S–N diagram) with implemented confidence intervals is presented in Fig. 6.

3.3. Data clustering and filtering

Clustering methods play an important role in the analysis of data generated in Digital Twin environments, particularly in the context of diagnostic interpretation. In diagnostic applications, clustering can be used to group data with similar characteristics, facilitating the detection of anomalies, outliers, and inconsistencies. By organizing data into clusters, it becomes possible to distinguish between typical system behavior and irregular observations.

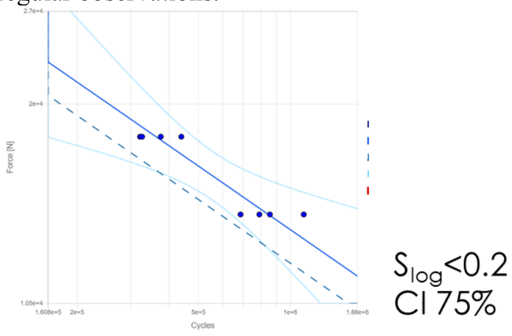

Fig. 6. Wöhler curve example

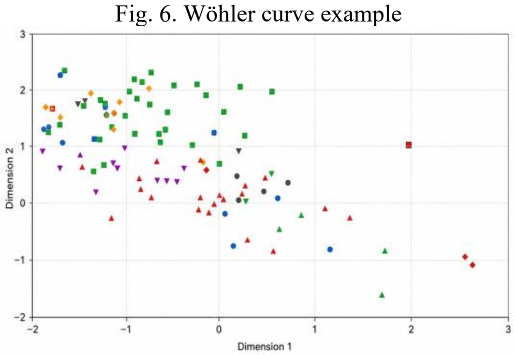

Fig. 7. Data clustering and filtering example

This supports more reliable identification of abnormal conditions and improves the interpretability of diagnostic results. The clustering-based grouping of data is illustrated in Fig. 7.

Clustering techniques may also assist in preprocessing data prior to statistical analysis, enabling more consistent datasets to be used in validation procedures. In practical Digital Twin applications, clustering methods may also support segmentation of operating conditions and identification of hidden data structures. As a result, datasets originating from different operational regimes can be analyzed separately, reducing the risk of drawing conclusions from heterogeneous populations. This capability is particularly valuable in systems operating under varying environmental and loading conditions. In this way, clustering contributes to improving the robustness of diagnostic conclusions in Digital Twin-based systems.

4. CONTEXT-BASED DIAGNOSTIC REASONING IN DIGITAL TWIN ENVIRONMENTS

While statistical methods provide a quantitative basis for comparing simulation and experimental data, they are often insufficient for reliable interpretation in complex engineering systems, where identical statistical patterns may arise from different physical causes depending on operating conditions, system configuration, or testing procedures. Therefore, the integration of context-based reasoning into Digital Twin environments constitutes a necessary extension of classical diagnostic approaches. Such context-aware methods have increasingly been recognized as essential for interpreting complex engineering data in uncertain and dynamic environments [22], [23].

In this study, context-based diagnostic reasoning is defined as a process of interpreting diagnostic signals and simulation outputs with respect to additional information describing the system state and its environment. Formally, the diagnostic process can be generally expressed as a mapping:

$$D{:}(X, C) \rightarrow S \tag{1}$$

where: X – set of observed data (measurements, simulation outputs), C – set of contextual variables, S – inferred system state or diagnostic decision.

In contrast to classical diagnostic approaches, where inference is based solely on X, the inclusion of C enables disambiguation of system behavior and improves robustness of diagnostic conclusions. The contextual space C may be structured into three primary layers:

- design context (C_d) – including material properties, geometric assumptions, and modelling simplifications used in simulation models,
- operational context (C_o) – describing real operating conditions such as load spectra, environmental parameters, and duty cycles; operational context includes load spectra, i.e., the distribution and sequence of loads experienced by a system during operation, as well as environmental conditions and operating regimes.
- process context (C_p) – related to testing and data acquisition conditions, including measurement setup, boundary conditions, and data preprocessing procedures.

Thus, the complete context vector may be represented as:

$$C = \{C_d, C_o, C_p\} \tag{2}$$

The integration of these context layers enables more precise differentiation between various sources of discrepancies observed in Digital Twin environments. In particular, deviations between simulation and experimental results may be classified into three categories:

- model-related deviations – resulting from simplifications or inaccuracies in simulation models,
- measurement-related deviations – caused by noise, sensor errors, or test setup inconsistencies,
- physical deviations – reflecting actual changes in system behavior or structural degradation.

The proposed context-based reasoning framework allows assigning observed deviations to one of the above categories by evaluating their dependence on contextual variables. This approach reduces the risk of incorrect diagnostic conclusions based solely on statistical analysis.

Furthermore, the incorporation of context supports decision-making processes by enabling prioritization of detected anomalies. For example, deviations occurring under nominal operating conditions (C_o) may indicate critical system faults, while similar deviations observed under extreme or non-standard conditions may be classified as acceptable variations.

Previous studies have indicated that the inclusion of contextual information significantly enhances diagnostic accuracy and interpretability, particularly in systems characterized by high variability and complex interactions between components [22], [23]. The importance of incorporating contextual information in diagnostic processes has been demonstrated in different systems. An example is an infrastructure system, such as water supply networks, where effective operation requires integration of technical, operational, and environmental data [13]. In Digital Twin applications, this capability is especially essential for maintaining consistency between virtual models and physical assets over time.

In summary, context-based diagnostic reasoning extends classical statistical diagnostics by introducing an additional layer of interpretative intelligence. Its integration with Digital Twin architectures enables more reliable fault identification, supports adaptive model updating, and enhances the overall robustness of engineering decision-making processes.

5. INTEGRATED STATISTICAL AND CONTEXT-BASED DIAGNOSTIC FRAMEWORK FOR DIGITAL TWIN ENVIRONMENTS

The diagnostic procedures discussed in the previous sections indicate that the effective operation of Digital Twin systems requires a structured approach to data analysis, interpretation, and decision support. Statistical methods provide a quantitative basis for assessing consistency between simulation and experimental data, while context-based reasoning enables interpretation of observed deviations with respect to design assumptions, operating conditions, and testing procedures.

Therefore, this section presents an integrated diagnostic framework for Digital Twin environments. The framework combines statistical diagnostics and context-based reasoning into a coherent structure supporting model validation, anomaly interpretation, and engineering decision-making.

5.1. General assumptions of the framework

The proposed framework assumes that discrepancies between simulation results and experimental observations cannot always be interpreted solely on the basis of numerical differences or statistical indicators. Similar statistical patterns may result from different causes, including model simplifications, measurement errors, inappropriate test setup, variable operating conditions, or real changes in the physical system.

5.2. Framework architecture

The proposed framework integrates statistical validation and context-based reasoning into a unified diagnostic process supporting Digital Twin development, validation, and operation.

The diagnostic procedures described in the previous sections indicate that effective operation of Digital Twin systems requires a structured approach to data analysis and interpretation. In complex engineering environments, large volumes of heterogeneous data originating from simulations, laboratory tests, and operational monitoring must be processed in a systematic manner. For this reason, it is useful to define a coherent diagnostic framework that integrates statistical analysis, data processing techniques, and contextual reasoning. The proposed diagnostic framework for Digital Twin environments was shown in Fig. 8. It consists of several complementary stages that together support reliable interpretation of engineering data.

The first stage involves data acquisition and integration, where datasets originating from simulation models, physical tests, and operational monitoring systems are collected and organized within a unified data structure. At this stage, particular attention must be paid to data consistency, standardized nomenclature, and traceability of measurement results.

Context identification, the second stage, is performed before or in parallel with statistical analysis. In practical applications, contextual information may influence data selection, preprocessing procedures, and the interpretation of anomalous observations. Therefore, context is treated not only as a reasoning mechanism applied after statistical validation but also as a factor affecting earlier stages of the diagnostic process.

The third stage focuses on data preprocessing and validation. In this phase, filtering, clustering, and statistical verification methods are applied to detect anomalies, remove inconsistent measurements, and prepare the dataset for further analysis. Techniques such as confidence interval analysis, Weibull distribution analysis, and outlier identification provide a foundation for ensuring the reliability of the available information.

The fourth stage involves correlation analysis between simulation and experimental data. Graphical comparison and statistical hypothesis testing allow engineers to evaluate whether simulation results remain within acceptable limits of experimental variability. This stage is essential for validating Digital Twin models and ensuring that simulation outputs accurately represent the behavior of physical systems.

In the next stage, context-based diagnostic reasoning is introduced. At this level, the interpretation of diagnostic results is extended by incorporating contextual information related to design assumptions, operating conditions, and testing procedures. The integration of contextual knowledge allows the diagnostic system to distinguish between deviations caused by modelling inaccuracies, measurement errors, or genuine structural phenomena.

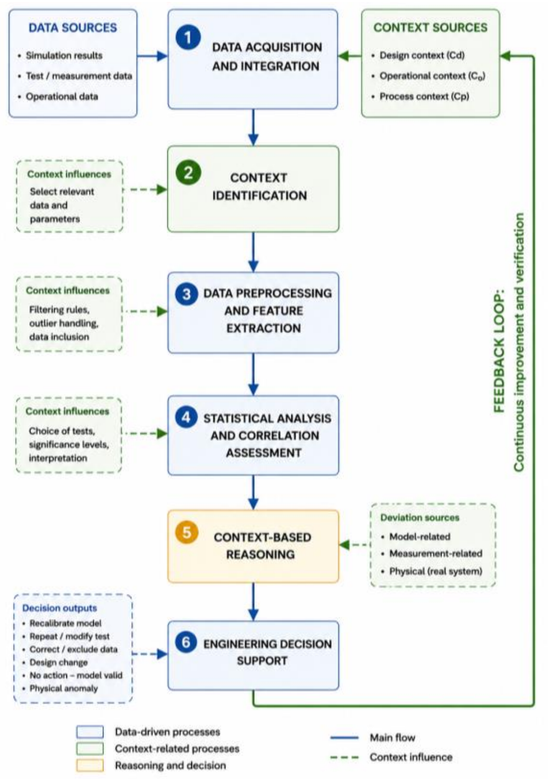

Fig. 8. Integrated statistical and context-based diagnostic framework for Digital Twin environments

The final stage of the framework involves engineering decision support. Based on the results of statistical analysis and contextual interpretation, engineers can determine whether simulation models require recalibration, whether additional physical testing is necessary, or whether the analyzed design solution meets the required performance criteria.

The proposed framework integrates statistical diagnostics and context-based reasoning into a unified structure. Such an approach supports the reliable operation of Digital Twin systems and enables more effective use of simulation-based engineering methods in modern product development processes. The case study presented in Section 6 demonstrates the application of the proposed framework.

5.3. Context Representation

In the proposed framework, diagnostic decisions are based not only on observed data but also on contextual information describing the conditions under which the data were generated, acquired, and processed. The contextual space is represented by the context model defined in Equation (2)

The design context includes assumptions adopted during model development, such as material properties, boundary conditions, and simulation settings. The operational context describes actual operating conditions, including load spectra and environmental influences. The process context contains information related to testing and measurement procedures, including sensor configuration, measurement uncertainty, and data acquisition conditions.

Within the proposed framework, contextual information supports data selection, interpretation of statistical results, and engineering decision-making. As a result, statistically detected deviations can be classified as model-related, measurement-related, or physical, improving the reliability of Digital Twin validation and diagnostic reasoning.

5.4. Diagnostic Reasoning Logic

In the proposed framework, diagnostic conclusions are derived from both observed data and contextual information. The diagnostic state may be expressed as:

$$S=f(X,C_d,C_o,C_p) \tag{3}$$

where (X) denotes the observed data, while (C_d), (C_o), and (C_p) represent the design, operational, and process contexts.

Statistical analysis identifies potential deviations between simulation results and experimental observations. However, the interpretation of these deviations requires contextual information. Depending on the available context, detected discrepancies may be classified as model-related, measurement-related, or physical deviations.

The combination of statistical validation and context-based reasoning improves the reliability of Digital Twin validation and supports more accurate engineering decision-making.

5.5. Role of Measurement Uncertainty

Measurement uncertainty influences the reliability of Digital Twin validation and should be considered during data interpretation. Experimental results may be affected by sensor accuracy, measurement resolution, environmental conditions, and test setup variability.

In the proposed framework, measurement uncertainty is treated as part of the process context (C_p), since it is directly related to the conditions under which data are acquired and processed. Information about uncertainty supports the interpretation of discrepancies between simulation and experimental results and helps distinguish between actual system deviations and measurement-related effects.

In engineering practice, uncertainty information may also support risk-based decision-making. Deviations that appear statistically significant may remain acceptable when uncertainty bounds are considered, while seemingly small differences may become important when measurement uncertainty is low. Consequently, uncertainty analysis provides an additional layer of confidence in Digital Twin validation procedures.

Considering measurement uncertainty improves the robustness of diagnostic conclusions and reduces the risk of incorrect decisions based on statistically significant but practically insignificant differences.

6. CASE STUDY: APPLICATION OF THE PROPOSED DIAGNOSTIC FRAMEWORK

To demonstrate the applicability of the proposed diagnostic framework integrating statistical analysis and context-based reasoning, a representative case study was developed. Due to industrial data confidentiality constraints, a synthetic dataset reflecting realistic engineering conditions was used. The dataset was constructed to represent typical variability observed in correlation studies between simulation and experimental testing in automotive components.

The presented case study is intended to demonstrate the application of the proposed diagnostic framework rather than to provide a comprehensive industrial validation. A simplified validation scenario was therefore considered, in which discrepancies between Digital Twin predictions and experimental observations were analyzed using both statistical methods and contextual information. The objective is to illustrate how the proposed framework supports engineering decision-making when interpreting deviations between the physical system and its digital representation.

The analyzed case concerns the comparison of simulation results with experimental measurements of a selected mechanical component subjected to cyclic loading. The dataset consisted of n = 7 samples, and the calculated mean and standard deviation were significantly affected by the outlier. The simulation output was assumed as a reference:

$$F_{sim}{=}1000\ N \qquad (4)$$

A set of experimental results was generated to reflect typical measurement variability:

$$F_{exp}{=}\{980, 995, 1005, 1010, 990, 1020, 1500\}\ N \ (4)$$

The last value (1500 N) represents an anomalous observation.

Although the dataset is limited, it reflects a common situation encountered during prototype development and Digital Twin calibration, where the availability of experimental measurements is restricted by testing costs, time constraints, or limited access to physical prototypes.

6.1. Scenario A: statistical analysis without contextual interpretation

Scenario A represents a conventional validation approach in which the available observations are analyzed using statistical methods only. The objective is to assess the level of agreement between simulation results and experimental measurements without considering additional contextual information. This scenario serves as a reference for evaluating the added value of context-based reasoning.

In the first scenario, diagnostic inference is based solely on statistical analysis. The sample mean and standard deviation are calculated, and a one-sample t-test is performed to evaluate the consistency between simulation and experiments. The presence of a significant outlier (1500 N) affects the statistical parameters, increasing variance and shifting the mean value. As a result, the statistical test indicates that the simulation result does not fall within the expected confidence interval.

From a purely statistical perspective, this leads to the conclusion that the simulation model is not sufficiently accurate and the correlation between simulation and experimental results is not acceptable.

Such a conclusion would typically trigger corrective actions, including model recalibration or additional simulation refinement. However, this interpretation does not account for possible external factors influencing the experimental data.

6.2. Scenario B: context-based diagnostic reasoning

In the second scenario, the same dataset is analyzed using the proposed context-based diagnostic framework. In addition to statistical analysis, contextual information related to the testing process is considered.

Scenario B extends the analysis by incorporating contextual information into the diagnostic process. Although statistical analysis identifies an anomalous observation, the final interpretation is supported by additional process context. This example demonstrates how contextual knowledge may influence diagnostic reasoning and improve the interpretation of deviations detected during Digital Twin validation.

The decision process in Scenario B follows the reasoning logic presented in Section 5.4. Although the statistical analysis identified an anomalous observation, the final decision was not based solely on statistical criteria. Additional process context (C_p) revealed that non-standard mounting conditions occurred during one of the tests. Consequently, the deviation was classified as a measurement-related deviation rather than a model-related or physical deviation. This contextual information justified the exclusion of the anomalous observation from the validation dataset. After excluding the contextually identified outlier, the remaining dataset becomes:

$$F_{exp}{^*}{=}\{980, 995, 1005, 1010, 990, 1020\}\ N. \qquad (5)$$

The recalculated statistical parameters show significantly improved agreement between simulation and experimental data. For the full dataset, the calculated p-value was below 0.05, while after removal of the context-identified outlier, the p-value exceeded 0.05. This demonstrates the sensitivity of statistical diagnostics to outliers and highlights the importance of contextual interpretation. The simulation result $F_{sim} = 1000\ N$ falls within the confidence interval of the corrected dataset, confirming the validity of the model.

The presented example demonstrates that statistically identified anomalies should not always be interpreted as evidence of model inadequacy. In engineering practice, additional contextual information may reveal external factors affecting the measurement process and provide an alternative explanation for the observed discrepancy.

6.3. Comparative analysis and discussion

The comparison between classical statistical diagnostics and context-based reasoning is summarized in Table 1.

The comparison presented in Table 1 illustrates that identical datasets may lead to different engineering decisions depending on the availability of contextual information. This observation is particularly important in Digital Twin environments, where simulation models are continuously updated based on operational and experimental data. Without contextual interpretation, organizations may unnecessarily invest resources in model recalibration or additional testing activities, even though the observed deviations originate from external factors rather than model deficiencies.

The presented case study demonstrates that statistical methods alone may lead to incorrect diagnostic conclusions when contextual information is not considered. In complex engineering systems, deviations observed in data may originate from multiple sources, and their correct classification is essential for reliable decision-making. The presented case study is simplified and intended to illustrate the methodological concept rather than provide a full-scale industrial validation.

Table 1. Comparison of diagnostic inference with and without context-based reasoning

Analysis approach	Input data interpretation	Diagnostic conclusion	Recom mended action
Statistical only	Outlier treated as valid data point	Simulation model inaccurate	Model recalibra tion required
Context-based reasoning	Outlier identified as measurement-related deviation	Simulation model valid	No model modifica tion required

The integration of context-based reasoning enables differentiation between deviations caused by model inaccuracies, measurement errors, and actual physical phenomena. This significantly reduces the risk of unnecessary model modifications and improves the efficiency of engineering processes. Furthermore, the proposed framework supports more informed decision-making by linking statistical analysis with engineering knowledge and process-specific information. This capability is particularly important in Digital Twin environments, where continuous alignment between simulation models and real-world data is required.

Although the presented example is intentionally simplified, it illustrates the main principles of the proposed framework and highlights the complementary roles of statistical validation and context-based reasoning in Digital Twin environments.

The results of the case study confirm that the integration of statistical diagnostics with context-based reasoning enhances the robustness of Digital Twin applications.

7. CONCLUSIONS

The results confirm that the Digital Twin concept supports engineering design and validation processes. However, maintaining consistency between simulation outputs and experimental observations remains a critical requirement for ensuring the credibility of Digital Twins. Statistical validation provides a quantitative basis for evaluation, while advanced diagnostic methods enable identification of inconsistencies and deviation sources.

The results also indicate that reliable Digital Twin validation requires consideration of measurement uncertainty and data acquisition conditions, as these factors may significantly influence the interpretation of discrepancies between simulation results and experimental observations.

The proposed framework integrates statistical analysis with context-based reasoning, allowing differentiation between model-related, measurement-related, and physical deviations, and improving the reliability of engineering decision-making. Future research should focus on integration of context-based reasoning with advanced data analytics and machine learning methods. The proposed framework may therefore serve as a practical foundation for the development of more reliable and context-aware Digital Twin validation systems in industrial applications.

7.1. Limitations of the proposed approach

Although the proposed framework provides a structured approach, several limitations should be considered. First, the effectiveness of the presented methodology strongly depends on the quality and availability of input data. In industrial environments, measurement data and simulation results are often obtained from heterogeneous sources and may be affected by inconsistencies, measurement errors, or incomplete datasets. Such issues may influence the accuracy of diagnostic conclusions.

Simulation models are often simplified, which may lead to discrepancies between simulation and experimental results. Moreover, in organizations operating across multiple locations, differences in local standards, naming conventions, or data management practices may limit the effectiveness of context-based reasoning mechanisms.

Finally, the implementation of the proposed diagnostic framework may require significant organizational and technological adjustments. These factors may limit the immediate applicability of the approach in organizations that are at an early stage of digital transformation.

7.2. Future research directions

The results presented in this study indicate several promising directions for further research. An important area involves the development of more advanced methods for integrating contextual information with statistical diagnostic techniques. This may enable more accurate interpretation of simulation and experimental data, particularly in complex engineering systems.

Another important research direction concerns the integration of Digital Twin technologies with machine learning and artificial intelligence methods. The use of advanced data-driven models may enhance the ability of Digital Twins to detect anomalies, predict system behavior, and support adaptive model updating based on operational data.

Future research should also focus on improving the automation of diagnostic procedures within Digital Twin systems. Automated data preprocessing, anomaly detection, and model validation procedures could significantly reduce the effort required for engineering analysis and increase the scalability of Digital Twin solutions in large industrial environments.

Finally, further studies should investigate the application of the proposed diagnostic framework in different industrial sectors. While the present study focuses primarily on the automotive industry, similar approaches may be applicable to other domains such as aerospace, energy systems, and advanced manufacturing. Evaluating the framework in diverse application contexts will help assess its generalizability and identify potential domain-specific adaptations.

Another promising direction concerns the development of context-aware Digital Twins capable of automatically acquiring and processing contextual information from engineering databases, manufacturing systems, and operational monitoring platforms. Such capabilities may significantly enhance the autonomy and adaptability of future Digital Twin systems operating in dynamic industrial environments.

Source of funding: *This research received no external funding.*

Author contributions: *Research concept and design, A.K., A.T., T.Ł.; Collection and/or assembly of data, A.K., A.T.; Data analysis and interpretation, A.K., A.T., T.Ł.; Writing the article, A.K., A.T., T.Ł.; Critical revision of the article, A.T., T.Ł.; final approval of the article, A.T.*

Declaration of competing interest: *The authors declares no conflict of interest.*

REFERENCES

1. Bartodziej CJ. The Concept Industry 4.0. Springer Berlin Heidelberg, New York. 2016. https://doi.org/10.1007/978-3-658-16502-4.
2. Elafri N, Tappert J, Rose B, Yassine M. Lean 4.0: Synergies between Lean Management tools and Industry 4.0 technologies. IFAC-PapersOnLine. 2022; 55(10):2060–2066. https://doi.org/10.1016/j.ifacol.2022.09.610.
3. Fuller A, Fan Z, Day C, Barlow C. Digital Twin: Enabling technologies, challenges and open research. IEEE Access. 2020;8:108952–108971. https://doi.org/10.1109/ACCESS.2020.2998358.
4. Garcia P, Radous J, Krol A, Bosek J, Baeten C. Lean Product Development. How to Create Flow? Reflection after a 4 Years Implementation in One Business Unit – Part 1. SAE 2016 World Congress and Exhibition. 2016. https://doi.org/10.4271/2016-01-0346.
5. Gilchrist A. Industry 4.0. Apress, Berkeley, 2016. https://doi.org/10.1007/978-1-4842-2047-4.
6. Grieves M. Vickers J. Digital Twin: Mitigating unpredictable, undesirable emergent behavior in complex systems. In: Transdisciplinary Perspectives on Complex Systems. Springer, 2017. https://doi.org/10.1007/978-3-319-38756-7_4.
7. Gudanowska A. Transformation towards Industry 4.0 – identification of research trends and aspects of necessary competences in the light of selected publications. Research in Logistics & Production. 2017;7(5):431–444.
8. Jones D, Snider C, Nassehi A, Yon J, Hicks B. Characterising the Digital Twin: A systematic literature review. CIRP Journal of Manufacturing Science and Technology. 2020; 29:36–52.
9. Król A, Timofiejczuk A, Łukasik T. Diagnostic and qualitative aspects in the digital twins implementation. In: Peruń G., Konieczny Ł. (eds.), Diagnostyka Maszyn. XLVIII Ogólnopolskie Sympozjum. Wisła, 2022.
10. Król A, Timofiejczuk A, Łukasik T. Cyfrowy bliźniak w procesie projektowania produktu. In: Projektowanie, budowa i eksploatacja maszyn. Studia i Monografie, Politechnika Opolska, 2022;574:121–130.
11. Kritzinger W, Karner M, Traar G, Henjes J, Sihn W. Digital Twin in manufacturing: A categorical literature review and classification. IFAC-PapersOnLine. 2018;51(11):1016–1022.
12. Lee J, Bagheri B, Kao HA. A cyber-physical systems architecture for Industry 4.0-based manufacturing systems. Manufacturing Letters. 2015;3:18–23.
13. Loska A, Moczulski W, Wyczółkowski R, Dąbrowski M. Integrated system of control and management of exploitation of water supply system. Diagnostyka 2016; 17(1)
14. Monostori L. Cyber-physical production systems: Roots, expectations and R&D challenges. CIRP Annals. 2014;63(2):621–641. https://doi.org/10.1016/j.cirp.2014.05.004.
15. Phanden RK, Sharma P, Dubey A. A review on simulation in digital twin for aerospace, manufacturing and robotics. Materials Today: Proceedings. 2021; 38:174–178. https://doi.org/10.1016/j.matpr.2020.06.332.
16. Piromalis D, Kantaros A. Digital twins in the automotive industry: The road toward physical-digital

convergence. Applied System Innovation. 2022;5(4):65.
17. Poniewierski A. SPEED: bez granic w cyfrowym świecie. Speednolimits.com Sp. z o.o., Wilcza Góra, 2020.
18. Resman M, Protner J, Simic M, Herakovic N. A five-step approach to planning data-driven digital twins for discrete manufacturing systems. Applied Sciences. 2021;11(8):3639.
19. Shao G, Janin S, Laroque C, Lee LH, Lendermann P, Rose O. Digital Twin for smart manufacturing: the simulation aspect. IEEE Conference Publications, 2019.
20. Tao F, Zhang H, Liu A, Nee AYC. Digital Twin in industry: State-of-the-art. IEEE Transactions on Industrial Informatics. 2019;15(4):2405–2415. https://doi.org/10.1109/TII.2018.2873186.
21. Temich S, Pollak A, Kucharczyk J, Ptasiński W, Mężyk A, Gąsiorek D. Prediction of energy consumption in the industry 4.0 platform - Solutions overview. Journal of Theoretical and Applied Mechanics. 2021;59(3):455–468. https://doi.org/10.15632/jtam-pl/140203.
22. Timofiejczuk A. Ujęcie kontekstowe w budowie i eksploatacji maszyn. Monografia, Politechnika Śląska, Gliwice, 2026.
23. Timofiejczuk A. An example of the use of contexts in diagnostics of manufacturing processes. Proceedings of the 12th International Conference on Modern Manufacturing Technologies in Industrial Engineering (ModTech), 2024.
24. Uhlemann THJ, Schock C, Lehmann C, Freiberger S, Steinhilper R. The Digital Twin: Demonstrating the potential of real-time data acquisition in production systems. Procedia Manufacturing. 2017; 9:113–120.
25. Wilczewski G, Siergiejczyk M. Testing of a data center environment with a generic model of a cloud-based telecommunication service. Diagnostyka. 2016;17(1).

Artur KRÓL, PhD. Eng. Six Sigma Master Black Belt.
Global lead of Lean & Six Sigma in Tenneco Monroe Ride Solutions. Responsible for managing several business processes within Global engineering structure, promoting implementation of Industry 4.0 into Engineering processes.
e-mail: akrol@driv.com

Anna TIMOFIEJCZUK, Prof. DSc hab.
Professor at the Silesian University of Technology. The area of scientific interest is technical diagnostics, signal and image analysis, applications of artificial intelligence and Industry 4.0
e-mail: anna.timofiejczuk@polsl.pl

Tomasz ŁUKASIK PhD. Eng.
Head of CAE department in Tenneco Monroe Ride Solutions. The area of professional activities, simulation processes and data management, advanced nonlinear simulation with usage of FEM, application of artificial intelligence and Industry 4.0.
e-mail: tlukasik@driv.com