Journal of
Marine Science and Engineering

MDPI

Article

A Framework to Assure the Trustworthiness of Physical Model-Based Digital Twins for Marine Engines

Jaehan Jeon * and Gerasimos Theotokatos *

Maritime Safety Research Centre, Department of Naval Architecture, Ocean and Marine Engineering, University of Strathclyde, 100 Montrose Street, Glasgow G4 0LZ, UK
* Correspondence: jaehan.jeon@strath.ac.uk (J.J.); gerasimos.theotokatos@strath.ac.uk (G.T.)

Abstract: Digital twins (DTs) are gradually employed in the maritime industry to represent the physical systems and generate datasets, among others. However, the trustworthiness of both the digital twins and datasets must be assured. This study aims at developing a framework to assure the trustworthiness of marine engines DTs based on first-principle models. This framework considers the phases of the DT development, progressivity, and trustworthiness assurance, the latter being based on three steps, namely validation, verification, and robustness. Subsequently, a methodology is applied to develop the DT of a marine engine for healthy conditions, which is extended to represent a wider operating envelope considering systematically identified anomalies. The results demonstrate that the developed DT trustworthiness is assured, as the validation step provided errors within ±3%, the verification step provided sound trade-offs, whereas the robustness assessment step confirmed acceptable uncertainty ratios. Subsequently, the DT is employed to generate datasets required for developing a data-driven model for anomaly diagnosis, which exhibits an accuracy of 98.8% for anomaly detection, 97.6% for anomaly identification, and 90.1–91.8% for anomaly isolation. This is the first study addressing the trustworthiness of DTs for marine engines, and as such advances concepts of the fourth industrial revolution to the shipping industry.

Keywords: digital twin; first-principle models; trustworthiness assurance; simulation-based dataset generation; anomaly diagnosis; marine engines

check for updates

Citation: Jeon, J.; Theotokatos, G. A Framework to Assure the Trustworthiness of Physical Model-Based Digital Twins for Marine Engines. *J. Mar. Sci. Eng.* **2024**, *12*, 595. https://doi.org/10.3390/jmse12040595

Academic Editor: Leszek Chybowski

Received: 8 March 2024
Revised: 26 March 2024
Accepted: 27 March 2024
Published: 29 March 2024

1. Introduction

Intelligent technologies for managing machinery systems, which are of critical importance to assure safe, efficient, and reliable ship operations, are relatively less advanced compared to autonomous navigation and communication [1]. To retain the machinery systems in appropriate conditions, conventional ships rely on maintenance activities carried out by their crew. However, maintenance of crewless ship machinery systems requires different strategies [2]. Crewless ships require intelligent systems to assess their machinery health status during sailings and make decisions for effective maintenance at port [3,4]. The development of intelligent machinery systems addressing the functionalities of monitoring, diagnosis, prognosis, and health management is required for autonomous ships, whereas it is expected that such systems will also benefit conventional ships.

Prognostics and Health Management (PHM) methods have been gradually employed in the maritime industry [5]. The PHM systems monitor critical operating parameters, assess the machinery's health condition, and assist decision-making so that vessel machinery system maintenance is managed effectively [6]. However, implementing data-driven PHM methods and developing intelligent machinery systems require extensive datasets. The datasets must cover a wide envelope of operating conditions to effectively represent ship machinery systems. The quality of datasets is of paramount importance for the effectiveness of data-driven approaches [7].

Acquiring appropriate datasets, which represent machinery conditions comprehensibly, is an immense challenge attributed to the highly varying operating envelope of ship machinery [8]. Ship machinery operations are impacted by environmental conditions as well as faults and degradation of their sub-systems. Typically, shipboard-measured datasets represent only a limited operating envelope in healthy conditions [7,9] of well-maintained machinery systems [10].

To extend the operating envelope of the acquired datasets, artificial data augmentation methods are frequently employed for data-driven models [11,12]. However, these artificial datasets cannot represent the complex behaviour of ship machinery systems appropriately. The use of inappropriate datasets undermines the performance of data-driven models regardless of the training process's effectiveness.

A digital twin (DT), which is a virtual representation of a physical system [13], presents an alternative approach for extending the operating envelope of the datasets. To develop DTs, models (physical and/or data-driven) for the reference physical system components must be developed and integrated, whereas their validity can be appraised based on acquired datasets [14]. Moreover, DTs can be employed in health monitoring applications, when reference physical system data for current conditions are provided [15]. The DT is a useful tool for generating datasets that represent an extended envelope of operating conditions. Simulations with DTs can complement expensive experimental campaigns, facilitating the generation of synthetic datasets [16].

The following three types of modelling approaches are typically employed for marine machinery system simulations: physical models (PMs), data-driven models (DDMs), and hybrid models (HMs). PMs are based on first principles (e.g., laws of thermodynamics) and have the capability of adequately representing the machinery behaviours. However, they require thorough physical knowledge and high computational effort. DDMs do not require prior knowledge of the machinery systems; instead, they rely on an extensive amount of data. HMs combine PMs and DDMs to compensate for both methods' drawbacks; however, limited applications are reported in the pertinent literature [17].

Tsitsilonis et al. [15] employed a marine engine physical model to generate performance parameters under both healthy conditions and current conditions, facilitating engine health assessment. Altosole et al. [18] employed a marine engine physical model to generate performance parameters considering ten degradation types, subsequently developing engine anomaly diagnosis models. Djeziri et al. [19] applied data-driven augmentation methods to simulate various potential pathways of transistor degradation and developed an offline Remaining Useful Life prognosis model. Stoumpos and Theotokatos [20] employed a marine engine physical model for simulating sensor abnormality and demonstrated a methodology for sensor diagnostics and health management, including self-corrective actions.

Although the use of several simulation tools is reported in the pertinent literature, the trustworthiness of the DT is only partially addressed. Trustworthiness is defined as the ability to meet user's expectations despite operational disturbances [21]. Validation and verification are essential procedures to assure the trustworthiness of the DT [22]. Typically, the validation is addressed by comparing simulation results against acquired measurements [18]. However, due to the lack of measurements in specific anomaly conditions, the validation of such simulation results has not been addressed in the pertinent literature. Data-driven PHM models employing results derived by less trustworthy DTs fail to monitor system behaviour in actual operations [23].

The criteria for addressing the trustworthiness of DTs may vary depending on the specific industry sector. For city automation, Wang and Burdon [24] proposed the following three elements for assuring DTs' trustworthiness: ability (functional quality), integrity (conformity to industrial standards), and benevolence (information privacy). For manufacturing automation, Babiceanu and Seker [25] considered dependability and cybersecurity to assure cyber-physical system trustworthiness. For systems and software engineering, de la Vara et al. [26] employed system architecture, multi-concern dependability, interoperability, and cross/intra-domain reuse to assure the trustworthiness of cyber-physical systems in the development of the open-source platform.

From the preceding literature review, the following research gaps are identified: (a) the need for developing a systematic framework to develop trustworthy DTs; (b) the use of comprehensive methods to design simulation scenarios for generating required datasets; (c) a lack of methods/tools for comprehensively assessing the trustworthiness of the DT; (d) a lack of investigations of combined anomalies (faults and degradations) and their various severities replicating realistic operating conditions; and (e) a lack of simulation studies considering realistic operating profiles (e.g., voyage profiles and unmanned engine room operation). This study addresses the research gaps (a), (b), and (c).

This study aims at developing a framework to assure the trustworthiness of DTs based on physical (first principle) models for marine engines. The framework provides systematic phases considering DT development, progressivity, and trustworthiness assurance. Subsequently, an application methodology is introduced to develop the DT in an extended operating envelope as well as anomaly conditions. The developed DT is employed to generate datasets required for developing a data-driven anomaly diagnosis model.

To the best of the authors' knowledge, this is the first study in the pertinent literature to address the trustworthiness of marine machinery DTs. This study offers guidance to develop trustworthy DTs for emerging technologies as well as for the power plants of smart and autonomous ships.

The remainder of this study is organised as follows. The definition of trustworthiness is provided in Section 2. The reference engine system and its DT are described in Section 3. The systematic framework for developing trustworthy DTs and the application methodology for both generating datasets and developing a data-driven anomaly diagnosis model are presented in Section 4. Section 5 describes the employed methods and tools for developing the DTs, assuring their trustworthiness, generating datasets, and developing data-driven anomaly diagnosis models. Section 6 describes an overview of the case study. Section 7 presents and discusses the results demonstrating the developed methodology and tools. Section 8 summarises the main findings and concludes the study.

2. Trustworthiness Definition

Trustworthiness is a broad term, hence different users can interpret it differently [27]. This section defines "trustworthiness" in the context of this study. Connet and O'Halloran [28] reviewed the trustworthiness definition from four perspectives: design, behavioural, physical, and anomaly perspectives. The design perspective implies that the system fulfils the set critical requirements. The behavioural perspective is related to confidence in the similarity between expectations and the system behaviour. The physical perspective implies that the trust pertains to the intended system functionality. The anomaly perspective implies that the system maintains its functionality despite the presence of disturbances/anomalies. Trustworthiness cannot be addressed in a single step; continuous assessments are required during the system design, development, and operation phases.

Schneider et al. [21] argue that a trustworthy system should maintain its trustworthiness regardless of constraints throughout the system's lifetime. The trustworthiness assessment of the developed DT is not a one-time occurrence, but it is performed iteratively as the DT undergoes variations.

In this study, the trustworthiness of the DT for a marine engine is addressed through three steps, which pertain to validation, verification, and robustness. The trustworthiness decision criteria are defined based on the design perspective, and the assessment methods are developed considering the behavioural, physical, and anomaly perspectives. The progressivity (non-regressivity) of the DT is also taken into account in the DT application.

However, trustworthiness criteria pertaining to data-driven models, such as explainability and physical plausibility, are not considered in this study as the primary focus is on DTs based on physical (first principles) models. The method for addressing the DT trustworthiness is described in Section 5.

3. Reference Engine System Description

This study employs the reference system of a marine engine along with the sub-systems and components for its monitoring, control, and safety functions. The reference engine is part of a generator set used in electric propulsion systems for generating electric power. The reference system is the Wärtsilä 8L50DF engine, which is a four-stroke, turbocharged, and intercooled dual-fuel engine [29], offering flexibility, as it can operate in diesel or gas modes [30]. The main particulars of this engine are listed in Table 1, whereas the engine's sub-systems and components which may face anomalies during the engine's lifetime are listed in Table 2.

Table 1. Reference engine system technical specifications.

Engine Type	8L50DF
Maximum Continuous Rating (MCR) Power	7800 kW
Nominal Engine Speed	514 rev/m
Cylinder Bore	500 mm
Stroke	580 mm
Number of Cylinders	8
Turbocharger	1 TPL 76

Table 2. Reference dual-fuel generator engine along with sub-systems and components.

Combustion System	Air Supply System	Fuel Supply System	Lubricating Oil System	Others
Fuel Injector	Turbocharger	Fuel Pump	LO Pump	Cooling Water System
Gas Admission Valve	Waste Gate Valve	Fuel Filter	LO Cooler	Safety and Monitoring System
Intake Valve	Air Cooler	Fuel Rack and Governor	LO Filter	Compressed Air System
Exhaust Valve	Air Filter			(Control and Starting)

4. Framework and Methodology

4.1. Framework for Assuring Trustworthiness of Physical Model-Based Digital Twins

The framework for developing trustworthy physical model-based DTs is illustrated in Figure 1. It consists of three phases (listed in numerical order), whereas each phase includes several steps. The first phase focuses on the sub-model developments and their integration to develop the DT. The DT is subsequently calibrated by adjusting the sub-model coefficients/constants to achieve acceptable accuracy.

Phase 2 deals with the progressivity (otherwise called non-regressivity) of the DT. This phase includes three options: (a) extending the operating envelope, (b) modifying/replacing existing sub-models, and (c) adding sub-models. Option (a) enables simulation of engine anomaly conditions or/and varying ambient conditions. The ranges of input operating parameters are set considering the acquired reference operating data. Option (b) deals with the modification or replacement of the initially employed sub-models. Option (c) caters for new sub-model additions, such as component degradation sub-models or/and auxiliary system sub-models. Options (b) and (c) require re-calibration of the DT to achieve acceptable accuracy.

In Phase 3, the trustworthiness of the DT is assured through three steps: validation, verification, and robustness. The validation step employs error metrics to calculate the accuracy of the DT outputs against experimental data. If the estimated errors exceed predefined thresholds, the DT re-calibration and reassessment are required. The verification step estimates the performance and emissions parameters trade-offs, and assesses their soundness by qualitatively comparing to the reference trade-offs. When the estimated trade-off is not sound, the DT redesign is required, followed by repeating the verification step. The robustness step calculates the uncertainty ratio between the DT output and input parameters, using input parameter reference profiles. When the estimated uncertainty ratio exceeds predefined thresholds, the DT redesign and repetition of the robustness step are required. When the DT re-calibration and redesign activities do not result in improved trustworthiness, it is advisable to stop the trustworthiness assurance phase and return to the DT development phase.

Completing these three steps assures the DT's trustworthiness for the considered operating envelope. In cases where the validation, verification, and robustness criteria are partially fulfilled (e.g., only for specific performance parameters), partial assurance of the trustworthiness should be indicated. In cases where the DT version (Phase 1) fulfils all steps criteria, the operating envelope extension option (Phase 2-A) does not require repetition of the validation step (Phase 3-A).

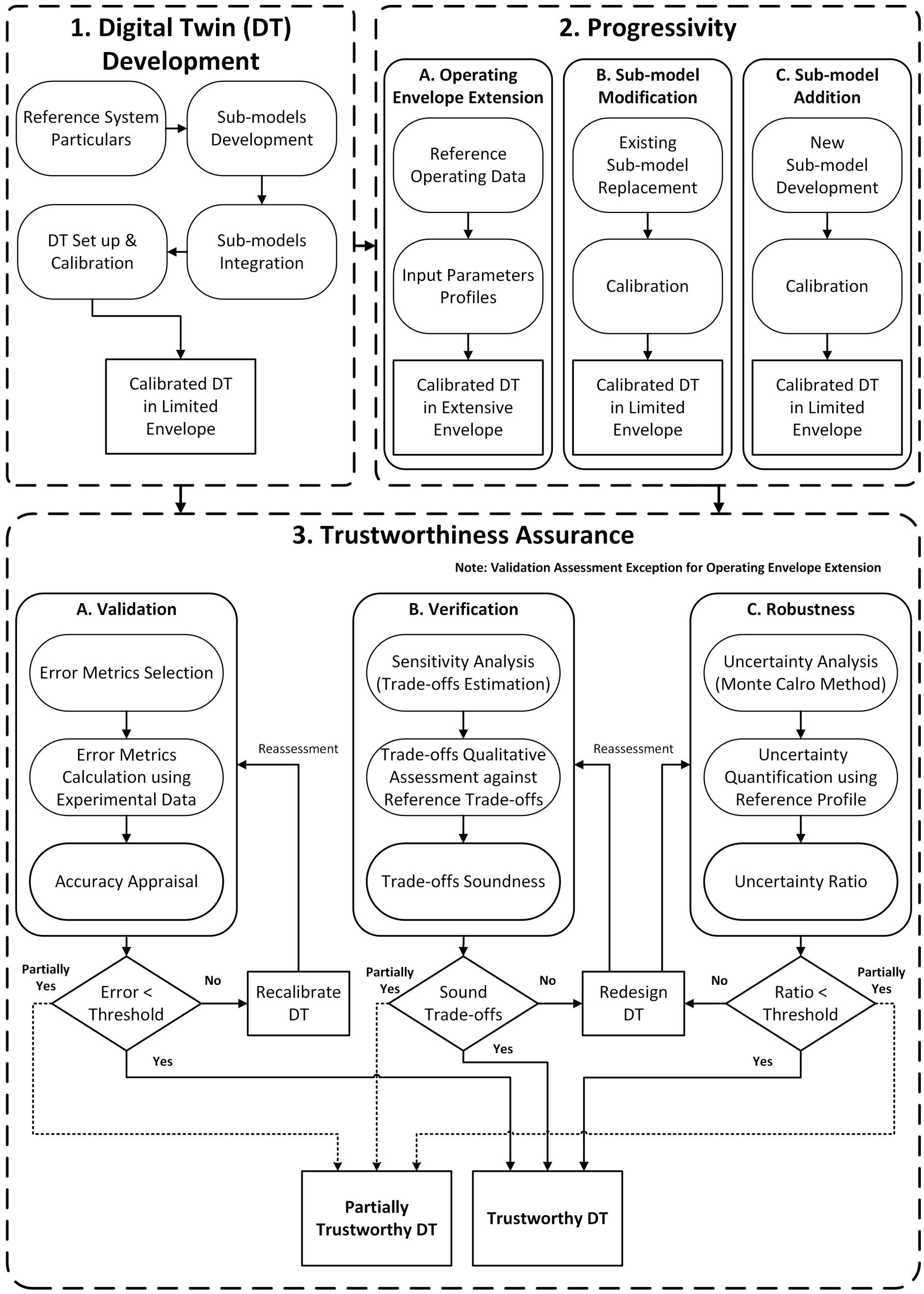


Figure 1. Systematic framework for assuring the trustworthiness of physical model-based DT.

4.2. Application Methodology in Marine Engine Anomaly Diagnosis

This section describes the application methodology for marine engine anomaly diagnosis, which employs the developed trustworthy DT framework (described in Section 4.1). This methodology consists of four stages (stages are used in this section to differentiate from the phases employed in the previous section), as illustrated in the flowchart of Figure 2.

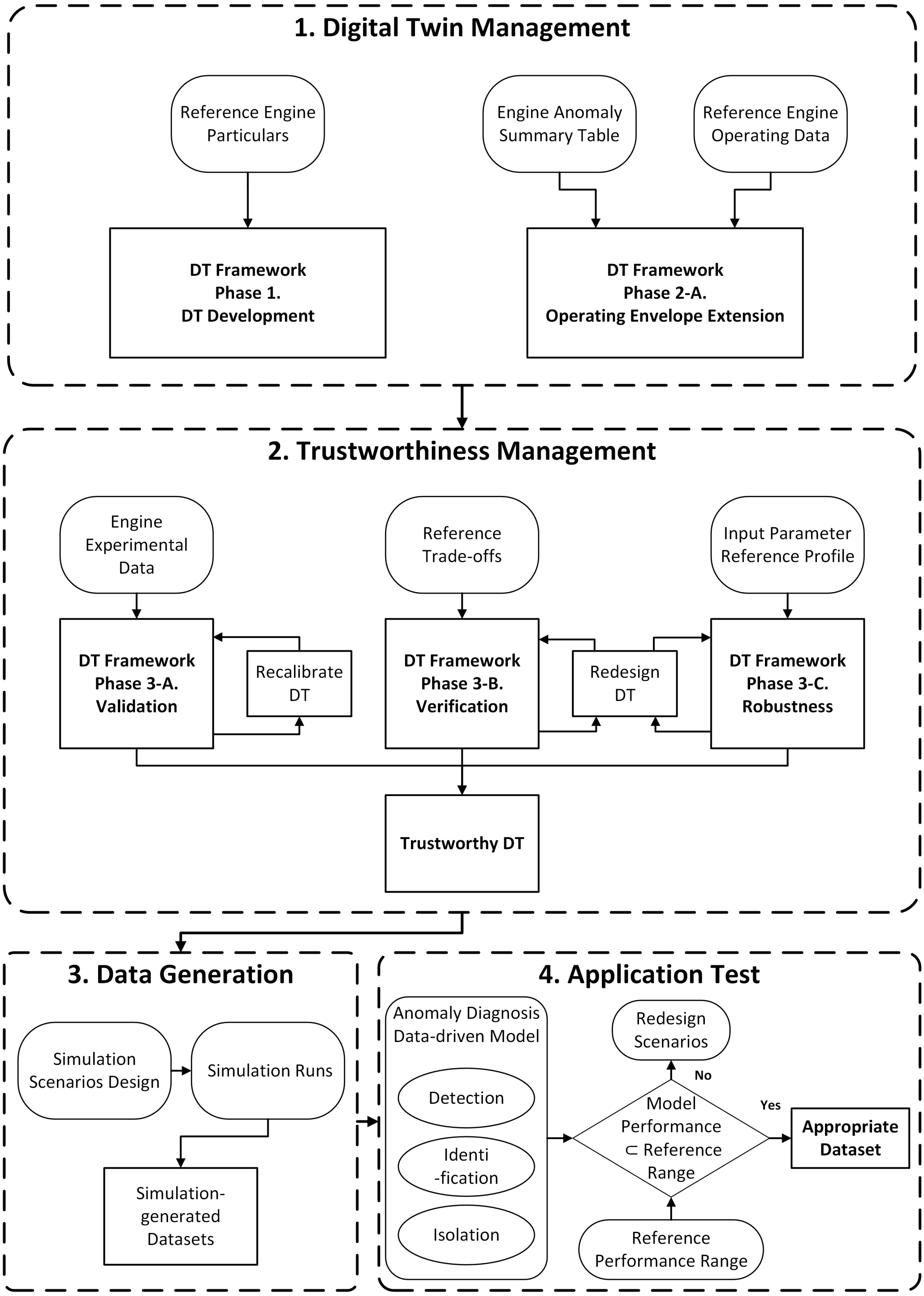


Figure 2. Application methodology in engine anomalies employing the trustworthy DT development framework.

Stage 1 deals with digital twin management. The reference engine particulars are employed, and the investigated engine DT is developed according to Phase 1 of the framework (physical model-based DT development). The engine particulars include geometrical information, turbocharger component performance maps, and measured engine performance parameters. To extend the engine operating envelope, which accommodates anomaly conditions and a wide range of ambient conditions, a reference engine anomaly summary table and reference engine operating data are employed. The anomaly summary table provides critical components for investigating anomalies and outlines their functions within the physical model. This information is provided as input to Phase 2-A of the framework (progressivity).

Stage 2 deals with DT trustworthiness management according to Phase 3 of the framework (DT trustworthiness assurance). Engine experimental data are employed for the validation step, reference trade-offs for the engine performance parameters are used for the verification step, whereas an input parameter reference profile is employed for the robustness step.

Stage 3 focuses on the dataset generation. Simulation scenarios are designed corresponding to the application of the generated datasets, considering the engine operating envelope, anomalies, and environmental conditions. Simulation runs are performed to estimate the engine performance parameters for these scenarios. The simulation results are integrated to generate datasets.

Stage 4 deals with the development of an anomaly diagnosis data-driven model using the simulation-generated datasets. Additionally, this application test assures the suitability of the generated datasets for anomaly detection, identification, and isolation tasks. The reported accuracy from pertinent diagnosis model studies is employed to compare with data-driven models developed herein. In cases where these models accuracy is not acceptable, the redesign of the data generation scenarios is required. Otherwise, it denotes the applicability of the simulation-generated datasets for their intended application.

5. Methods and Tools

5.1. Physical Model-based Engine Digital Twin

The physical model-based engine DT development was carried out in the GT-SUITE v2022 software, which provides a 0D/1D simulation interface with libraries for the system components and sub-models [31]. The DT is developed based on the author's previous studies [32,33]; however, in this study, the DT was extended by the inclusion of the modelling of sensors and their dynamics for the engine performance parameter measurements. The sensor sub-model monitors and collects virtual signals from the simulations representing the sensors deployed in engine operations. The layout of the modelled engine components, along with the sensors and monitoring system block, is illustrated in Figure 3.

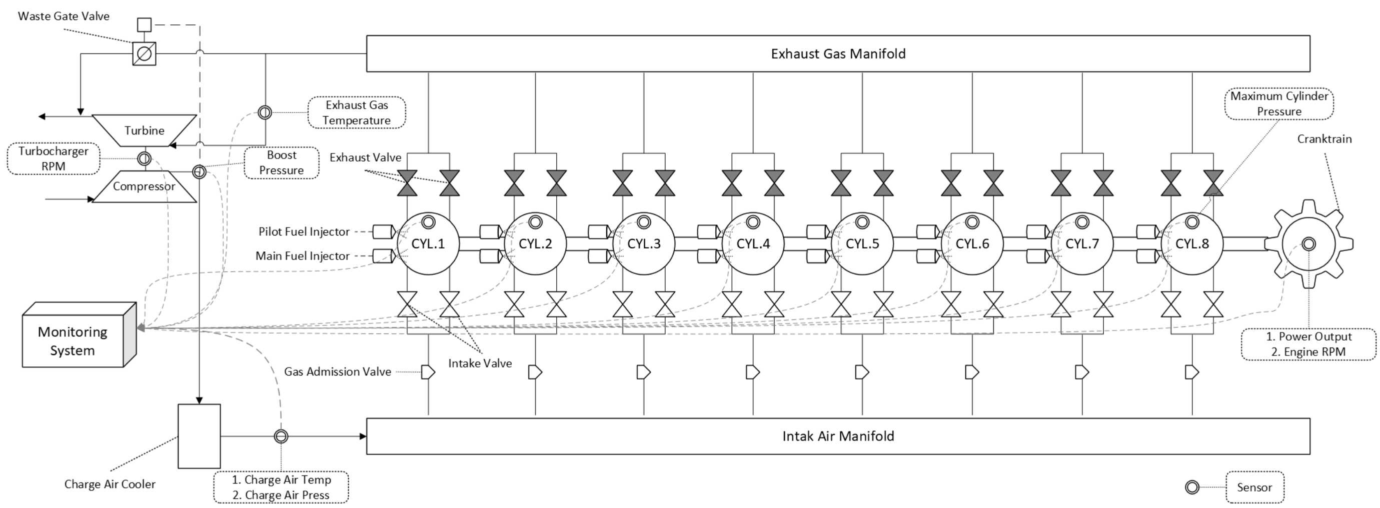


Figure 3. Marine engine components and sensors modelled in the developed DT.

The engine DT includes several sub-models for representing the engine components. The Woschni equation [34] is employed for representing the in-cylinder heat transfer process, whereas the single Wiebe function [35] is used for modelling the combustion process. The air and exhaust gas manifold sub-models employ one-dimensional approaches, which consider the fluid momentum, mass, and energy conservation equations. The engine control sub-models for the fuel supply systems and the wastegate valve employ Proportional-Integral-Derivative (PID) controllers. Detailed descriptions of the employed sub-models are provided in the author's previous studies [30,32,33].

In addition, the monitoring system was developed, including sensor sub-models and safety sub-models. The sensors were modelled using the additive noise model [36], which employs the Gaussian random process, along with appropriate time constants for capturing the sensor dynamics. The safety sub-model monitors if the input and output parameters remain within their specified ranges, whereas it terminates the simulation run when the parameters exceed their set thresholds.

5.2. Anomaly Summary Table

An anomaly summary table collects the necessary information to model the components required for extending the DT to represent anomaly conditions. The anomaly table is developed by employing identification and analysis processes as presented in Table 3. The identification process employs the Failure Mode, Effects, and Criticality Analysis (FMECA) method to identify the failure modes, failure causes, failure effects, as well as the criticality of these failures [37]. The anomaly criticality is evaluated by the Risk Priority Number (*RPN*), which is derived by the multiplication of occurrence level (*O*), severity level (*S*), and detectability level (*D*), according to the following equation [38]:

$$RPN = O\,S\,D \tag{1}$$

The following parameters are determined: anomaly control parameters (DT input), dependent parameters (DT output), manufacturer's limits, ranges of the pertinent input parameters, and steps; the latter is derived from the manufacturer's limits, ranging from healthy to severe anomaly conditions.

Table 3. Anomaly summary table format.

Anomaly Identification Process with FMECA										**Anomaly Analysis for Simulation Design**				
FMECA Output										DT Input parameters				DT Output parameters
Component	Function	Failure Mode	Failure Causes	Failure Effects	Detection Method	*O*	*S*	*D*	*RPN*	Name	Manufacturer limits	Input Range	Simulation Steps	Name

5.3. DT Calibration and Design of Experiments

The DT sub-models are calibrated by employing an optimisation method, which uses the reference performance parameters and calculates the optimal values for the sub-model constants to achieve the lowest value of an objective function [15,39]. GT-SUITE software provides the Advanced Direct Optimiser (ADO) that employs the Non-dominated Sorting Genetic Algorithm NSGA-III [40]. In this study, the ADO optimises turbocharger turbine mass flow and efficiency coefficients at 75% load using an objective function (weighted sum of errors) associated with three performance parameters (turbocharger shaft speed and turbine inlet and outlet temperature). An example for calibrating the turbocharger turbine model is shown in Table 4.

Table 4. Example of using the advanced direct optimiser to calibrate the turbocharger turbine model.

Factors			Responses			
Parameter	**Turbine Mass Flow Scale Factor [–]**	**Turbine Efficiency Scale Factor [–]**	**Parameter**	**TC Shaft Speed [rev/m]**	**Exhaust Gas Temperature Upstream Turbine [K]**	**Exhaust Gas Temperature Downstream Turbine [K]**
Search Range	0.8–0.99	1.0–1.12	Target Value	16,690	749	594
Optimal Value	0.9354	1.0753				

To explore the whole engine operating envelope and effectively identify the required simulation scenarios, a design of experiments (DOE) method [41,42] is employed. This study employs the GT-SUITE DOE tool [31] and the Latin hypercube method to define the required number of samples for sufficiently representing the considered operating envelope [43].

For selecting anomaly simulation scenarios, anomalies severity of the engine cylinder valve leakage range of 0–1.0 mm are investigated. The DOE results define that the valve leakage is noticeable between 0.2 mm and 0.3 mm, and thus the anomaly simulation steps are determined as acceptable level (0.1 mm), weak leakage (0.3 mm), and severe leakage (0.5 mm).

5.4. Trustworthiness Assurance

This study develops a comprehensive method for assuring the trustworthiness of the marine engine DT, considering the following three steps: validation, verification, and robustness. Each step partially ensures the trustworthiness of the DT due to limited data and information; however, the integration of these steps provides comprehensive assurance. Pre-set decision criteria for each step are decided based on the DT requirements, which are described in the following subsections.

5.4.1. Validation

The validation step depends on available measured data, which are expected to be limited. For marine engines, the engine shop tests (or factory acceptance tests) and ship trials provide the engine performance parameters under healthy conditions at limited operating points. Shipboard data acquisition systems are a recent trend in the maritime industry; however, their use is limited. The limited data restrict the validation, and thereby the validation ensures the trustworthiness of the developed DT only within the limited operating envelope where the measured dataset are available. The validation of the DT is based on the comparison between the predicted performance parameters and available measured datasets. The criteria employed for the validation step involve the percentage errors between the predicted and measured parameters, as calculated by the following equation:

$$E = \frac{OP - RD}{RD} 100 \tag{2}$$

where E denotes the percentage error, OP is the output parameter, and RD is measured value of the same parameter.

The acceptance limits for the error of performance parameters (pressure, temperature, rotational speed, and fuel flow) are based on the guidelines from the ISO standards applicable to engine testing measurements [44]. If the estimated percentage error is smaller than the guideline's allowance, the DT is considered validated. For cases where validation of only specific parameters is achieved, the DT is considered partially validated.

5.4.2. Verification

The verification step deals with identifying the DT behaviour considering several engine operating conditions. However, access to extensive measurements under specific operating conditions is limited. Instead, the evaluation is based on consistency with expected trade-offs from the pertinent literature. The verification step is based on sensitivity analysis to evaluate the influence of the input parameter changes on the output parameter variations [45]. The Spearman rank correlation coefficient [46,47] is employed to quantify the strength and the direction of the relationship (trade-off) between the input and output parameters. This coefficient is calculated by Equation (3), which results in a range from -1 (strong reciprocal correlation) to +1 (strong direct correlation) with 0 representing negligible correlation.

$$r_s = 1 - \frac{6\Sigma(R(X_i) - R(Y_i))^2}{N(N^2 - 1)} \tag{3}$$

where r_s denotes the Spearman's coefficient, X is the input parameter values, Y is the output parameter values, N is the number of samples, and R denotes the ranked variables.

This study takes into account varying operating conditions for ambient temperature as well as the anomalies of exhaust and intake valve leakages. The Latin hypercube method samples the combinations of these three parameters. The DOE tool of the GT SUITE is employed to perform the simulations and predict the trade-offs of performance parameters for each combination of input parameters. If the estimated trade-offs for the DT performance parameters match the trade-offs from the pertinent literature, the DT verification is assured.

5.4.3. Robustness

The last step of the trustworthiness assurance deals with the evaluation of the DT's robustness, which is based on the predicted performance parameter uncertainty considering anomaly conditions and varying environmental conditions (ambient temperature in this study).

Uncertainty analysis is employed to measure the probability of unexpected event occurrences [48]. This study employs the variance for quantifying the uncertainty [49] and uses the Monte Carlo (MC) method [50] to estimate the variance of the DT output performance parameters. The input (valve wear and ambient temperature) and the output (predicted performance parameters) are both normalised within the range of [0, 1] (to eliminate scale differences between parameters) by employing the min-max scaling method according to the following Equation (4) [51]:

$$\widetilde{X}_l = \frac{X_i - \min(X)}{\max(X) - \min(X)} \tag{4}$$

where $\widetilde{X}_l$ denotes the normalised value of the ith sample, X_i is the value of the ith sample, min(X) is the minimum value of the samples, and max(X) is the maximum value of the samples.

The variances of the input and output parameters are calculated according to the following equation:

$$S^2 = \frac{\sum_{i=1}^{n}(X_i - \bar{X})}{n - 1} \tag{5}$$

where S^2 denotes the variance, n is the sample number, X_i is the value of the ith sample, and $\bar{X}$ is the mean value of the samples.

The uncertainty ratio between the input and output parameters is calculated using the following Equation (6):

$$R_m = \frac{S_m^2}{S_{in}^2} \tag{6}$$

where R_m denotes the uncertainty ratio, S_m^2 is the variance of the mth predicted performance parameter, and S_{in}^2 is the variance of the input parameter.

The acceptance criteria for the DT robustness are established based on a requirement that the DT should not introduce additional uncertainty or propagate the uncertainty (from input to output parameters). This requirement implies that the uncertainty of the normalised output parameters should not exceed the uncertainty of the normalised input parameters. Hence, uncertainty ratio values below 1 indicate that DT robustness is assured.

5.4.4. Trustworthiness Assurance Table

From the results of these three steps, the calculated metrics are summarised in a tabular format, as presented in Table 5. The trustworthiness table provides an intuitive assessment, this supporting decision making.

Table 5. Trustworthiness assurance table.

Steps	Acceptance Criteria	Trustworthiness Checks	
		Environment Conditions	**Anomaly Conditions**
Validation	Acceptable Errors	Pass/Fail	
Verification	Trade-off Soundness	Pass/Fail	Pass/Fail
Robustness	Uncertainty Ratio	Pass/Fail	Pass/Fail

5.5. Data-Driven Anomaly Diagnosis Model

For testing the application suitability of the simulation-generated datasets, the generated datasets are employed to develop a data-driven anomaly diagnosis model. The diagnosis model includes sub-models for anomaly detection, identification, and isolation. The anomaly diagnosis is based on the following steps: (a) the anomaly detection model determines the engine condition as either healthy or abnormal; (b) the identification model categorises the anomaly types (intake valve, exhaust valve, or both); (c) the isolation model determines the anomaly location (engine cylinder) with the categorised anomaly type.

All these sub-models employ the same input parameters, specifically the turbocharger speed, exhaust gas temperature upstream turbine, maximum cylinder pressure, charge air pressure and temperature, and mass fuel consumption. The outputs of the sub-models are anomaly condition labels, which are allocated for each data sample before training the data-driven models corresponding to the diagnosis tasks.

The anomaly diagnosis model is developed using Support Vector Machines (SVM) with the Radial Basis Function (RBF) kernel, which demonstrated generalisation capability and advantageous performance with small datasets [52]. The SVM learns decision hyperplanes, which have minimum overall error [53] and estimates classes of new data with the trained hyperplanes [12]. However, the kernel function, which transforms nonlinear data into linearly separable data, is additionally required since the SVM was originally designed for linear problems [54]. The RBF kernel [55] is employed herein as it requires the calibration of only one parameter (σ), according to Equations (7) and (8). The grid search with five-fold cross-validation [56] is employed to tune the hyperparameters, and the one-versus-one method [57] is employed for multi-classification of anomaly identification and isolation.

$$\kappa(x_1, x_2) = \exp\left(-\frac{\|x_1 - x_2\|^2}{2\sigma^2}\right) \quad (7)$$

where κ denotes the kernel function of two-dimensional inputs (x_1, x_2) and σ is the width of the kernel.

$$F(x) = \sum_{i=1}^{n} w_i \kappa(x, x_\mathrm{i}) \quad (8)$$

where F denotes the multidimensional kernel function, n is the number of dimensions, and w is the weight of dimensions.

The performance of the anomaly diagnosis models is characterised by employing the confusion matrix, whereas the model accuracy is calculated according to Equation (9). The confusion matrix [58] illustrates the distribution of true and false predictions over all classes in the tabular format shown in Table 6.

$$A = \frac{TN + TP}{TN + TP + FN + FP} 100 \quad (9)$$

where A denotes the accuracy percentage, TN is the number of true negative samples, TP is the number of true positive samples, FN is the number of false negative samples, and FP is the number of false positive samples.

Table 6. Example of a multi-classification confusion matrix for class 2.

	Predicted Class			
True Class		Class 1	Class 2	Class 3
	Class 1	TN	FP	FN
	Class 2	FN	TP	FN
	Class 3	FN	FP	TN

6. Case Study Description

The considered marine engine DT must represent a wide envelope of engine conditions considering various engine loads, anomaly conditions, and environmental conditions. This study employs ambient temperature variations as representative of ship environmental conditions. The ambient temperature range was considered 15–45 °C according to the manufacturer's guidelines [29]. Table 7 provides the overview of the considered operating conditions, whereas Table 8 lists the simulated case studies, each accompanied by its corresponding application methodology step and scope.

For the anomaly simulations, the engine cylinder valve leakage faults based on the valve clearance error are investigated, which affects the engine output parameters. For large intake valve clearance, the engine performance is degraded due to the reduced air supply and delayed valve open timing. For small intake valve clearance, the engine performance is also degraded due to the early valve opening, air leakages, and a loss of compression pressure [59]. As the valve and seat wear is unavoidable over time, the valve clearance must be checked and adjusted regularly to avoid valve leakage. The engine manufacturer recommends checking the valve clearance every 2000 operating hours to prevent engine performance degradation.

The conventional method to monitor the valve clearance and leakage conditions relies on shipboard manual measurements, which cannot be carried out during the engine operation. However, intelligent health monitoring systems employing data-driven models developed herein, could monitor the valve clearance and leakage conditions by analysing the engine performance parameters without operational interruption. To overcome the lack of valve clearance data and develop anomaly diagnosis models for valve leakage anomalies, the developed data generation methodology is employed.

Table 7. Overview of the considered engine operating conditions.

Condition	Engine Load [%]	Anomaly Location [–]	Anomaly Severity [mm]	Amb_T [4] [°C]
Healthy	25; 50; 75; 100	–	–	15–45
EV_Leak [1]	25; 50; 75; 100	Cyl [3] 1–8	0.1; 0.3; 0.5	15–45
IV_Leak [2]	25; 50; 75; 100	Cyl 1–8	0.1; 0.3; 0.5	15–45

[1] IV_Leak: Intake Valve Leakage [mm]. [2] EV_Leak: Exhaust Valve Leakage [mm]. [3] Cyl: Cylinder Number [–]. [4] Amb_T: Ambient Temperature [°C].

Table 8. Overview of the simulated case studies.

Code	Simulated Case Study	Methodology Step	Scope
H1	Healthy (4 Loads, 1 Amb_T)	DT Management and Trustworthiness (Validation)	Model calibration/comparison to measured data for DT validation check
H2	Healthy Ambient Temperature Variation (4 Loads, Amb_T: 15–45 °C)	Trustworthiness (Verification)	Sensitivity analysis to confirm the predicted parameters trade-offs under ambient temperature variation
A1	Anomaly EVLeak or IVLeak (4 Loads, 1 Location, 3 Severities for each anomaly)	Trustworthiness (Verification)	Sensitivity analysis to confirm the predicted parameter trade-offs under various anomaly conditions
A2	Anomaly EVLeak or IVLeak (4 Loads, 1 Location, Severities of anomaly: 0.1–0.5 mm, Amb_T: 15–45 °C)	Trustworthiness (Robustness)	Uncertainty analysis based on Monte Carlo simulations to check predicted parameter variances
DG1	Healthy and Anomalies EVLeak or IVLeak (4 Loads, 1 Location, 3 Severities for each anomaly, Amb_T 25–35 °C)	Data generation	Multiple simulations to generate the datasets required for the data-driven anomaly detection models
DG2	Healthy and Anomalies EVLeak or/and IVLeak (4 Loads, 1 Location, 42 Severities with combined anomalies, Amb_T: 25–35 °C)	Data generation	Multiple simulations to generate the datasets required for the data-driven anomaly identification models
DG3	Healthy and Anomalies EVLeak or IVLeak (4 Loads, 8 Locations, 2 Severities for each anomaly, Amb_T: 25–35 °C)	Data generation	Multiple simulations to generate the datasets required for the data-driven anomaly isolation models
DD	Healthy and Anomalies EV or/and IV leakage	Application test	Developing data-driven models for anomaly diagnosis; check with the data-driven models' accuracy

7. Results and Discussion

7.1. DT Management and Trustworthiness Management

The summary table of engine cylinder valve anomalies is provided in Table 9. The engine cylinder valve seat wear causes valve stem clearance reduction. If this reduction exceeds the manufacturer's allowance, the valve does not fully close and causes air/gas leakages, leading to engine efficiency deterioration. To simulate these anomalies, the valve lift profile and the valve lash were determined based on the manufacturer's allowances for the valve stem clearance (0.147–0.199 mm). Hence, the following four severity levels were considered: No Leakage, Acceptable Clearance, Weak Leakage, and Severe Leakage. The engine performance parameters with the greater impact from these anomalies as well as their trade-offs were identified based on the pertinent literature [60,61].

The DT trustworthiness assurance was based on the outcomes of the validation, verification, and robustness checks, which are summarised in the trustworthiness decision table. In the validation step, the performance parameter predictions were compared with the respective reference engine shop trial data. The estimated percentage errors for six performance parameters are presented in Table 10. The maximum absolute percentage error was identified as 3% for the fuel oil consumption at 100% load. Nonetheless, all the predicted parameters (at all loads) satisfied the acceptance criteria (allowable errors), which appear in the right column of Table 10. Hence, it is deduced that the DT passed the validation check.

In the verification step, the trade-offs of the DT-generated performance parameters were compared with the trade-offs reported in pertinent experimental studies. The sensitivity analysis results (derived trade-offs) are presented in Table 11, whereas the comparisons of the trade-offs with pertinent studies are presented in Table 12 for the ambient temperature variations and Table 13 for the anomalies, respectively. The ambient temperature trade-offs were compared with the trade-offs reported by Whitehouse et al. [62], Serrano et al. [63], MAN [64], whereas the anomaly condition trade-offs were compared to those reported by Kowalski [60]. From these tables, it is deduced that the predicted trade-offs are sound corresponding to the reported ones. Therefore, it is deduced that the developed DT passed the verification test.

In the robustness step, 524 operating points were employed to perform an uncertainty analysis of the predicted engine performance parameters. The uncertainty ratios are presented in Table 14. The DT in the various ambient conditions and anomaly conditions satisfied the robustness acceptance criteria as the uncertainty ratios of all the predicted performance parameters are below 1. Hence, it is deduced that the developed DT satisfies the robustness acceptance criteria. The outcomes of the three steps are summarised in Table 15, from where it is confirmed that the developed DT is considered trustworthy.

Table 9. Anomaly summary table.

Component	Function	Failure Mode	Failure Causes	Failure Effects	Detection Method	*O*	*S*	*D*	*RPN*	Simulation Input				Simulation Output
										Name	Manufacturer Limit	Input Range	Simulation Steps	Name
Intake Valve	Supply intake air into a cylinder	The valve is not fully closed	Valve seat wear	Intake air leakage → Engine efficiency deterioration	Manual measurement of clearance	4	5	6	120	Valve remained lift and lash [mm]	Normal clearance: 1.0 mm (cold) Wear limit: 0.147–0.199 mm	Valve remained lift: 0–0.75 mm Valve lash: 0.25–1.0 mm	(4 steps of clearance) 0 mm—No leakage 0.1 mm—Acceptable clearance/No leakage 0.3 mm—Slight leakage 0.5 mm—Severe leakage	TC_RPM [1] Exh_T [2] Pmax [3] CA_P [4] CA_T [5] FOC [6]
Exhaust Valve	Releases burned gases from a cylinder	The valve is not fully closed	Valve seat wear	Exhaust gas leakage → Engine efficiency deterioration	Manual measurement of clearance	4	5	6	120	Valve remained lift and lash [mm]	Normal clearance: 1.5 mm (cold) Wear limit: 0.147–0.199 mm	Valve remained lift: 0–1.0 mm Valve lash: 0.5–1.5 mm	(4 steps of clearance) 0 mm—No leakage 0.1 mm—Acceptable clearance/No leakage 0.3 mm—Slight leakage 0.5 mm—Severe leakage	TC_RPM Exh_T Pmax CA_P CA_T FOC

[1] TC_RPM: Turbocharger Shaft Speed [rev/m]. [2] Exh_T: Exhaust Gas Temperature [K]. [3] Pmax: Maximum In-Cylinder Pressure [bar]. [4] CA_P: Charge Air Pressure [bar]. [5] CA_T: Charge Air Temperature [K]. [6] FOC: Fuel Oil Consumption [g/kWh].

Table 10. Validation results—percentage errors.

Load [%]	100	75	50	25	Acceptable Error
TC_RPM	0.2	0.0	−1.3	−1.2	± 2
Exh_T	−2.9	−2.0	0.0	−1.1	±5 (±25 K)
Pmax	1.0	0.9	0.4	−1.1	±5
CA_P	1.7	−1.4	−1.0	0.9	±2
CA_T	0.1	−0.5	−0.1	−1.1	±1.2 (±4 K)
FOC	−3.0	−1.6	−0.5	−0.3	±3

Table 11. Sensitivity analysis results.

Input Parameters	Spearman's Coefficient [–]					
	TC_RPM	Exh_T	Pmax	CA_P	CA_T	FOC
Ambient Temperature	0.19	0.84	−0.23	−0.58	0.58	0.35
Exhaust Valve Leakage	0.92	0.15	−0.31	0.61	0.76	0.43
Intake Valve Leakage	0.36	0.56	−0.95	0.40	0.37	0.85

Table 12. Verification results—ambient temperature variation trade-offs.

Output Parameters	Low Temperature		High Temperature		Reference
	DT	Reference	DT	Reference	
TC_RPM	↓	↓	↑	↑	[63,64]
Exh_T	↓	↓	↑	↑	[64]
Pmax	↑	↑	↓	↓	[64]
CA_P	↑	↑	↓	↓	[64]
CA_T	↓	↓	↑	↑	[62]
FOC	↓	↓	↑	↑	[64]

Table 13. Verification results—influence of anomaly trade-offs.

Output Parameters	Exhaust Valve Leakage		Intake Valve Leakage		Reference
	DT	Reference	DT	Reference	
TC_RPM	↑	↑	↑	↑	[60]
Exh_T	↑	↑	↑	↑	
Pmax	↓	↓	↓	↓	
CA_P	↑	↑	↑	↑	
CA_T	↑	↑	↑	↑	
FOC	↑	↑	↑	↑	

Table 14. Robustness results—uncertainty ratio.

Input Parameters	Load [%]	Uncertainty Ratio [–]					
		TC_RPM	Exh_T	Pmax	CA_P	CA_T	FOC
Amb_T	100	0.001	0.162	0.005	0.006	0.622	0.007
	75	0.001	0.151	0.004	0.006	0.605	0.008
	50	0.000	0.211	0.003	0.005	0.427	0.010
	25	0.000	0.125	0.000	0.001	0.729	0.009
EV_Leak	100	0.004	0.055	0.054	0.009	0.573	0.028
	75	0.005	0.025	0.030	0.010	0.750	0.024
	50	0.004	0.034	0.014	0.004	0.374	0.032
	25	0.002	0.060	0.005	0.000	0.082	0.045

Table 14. *Cont.*

Input Parameters	Load [%]	Uncertainty Ratio [–]					
		TC_RPM	Exh_T	Pmax	CA_P	CA_T	FOC
IV_Leak	100	0.001	0.207	0.009	0.127	0.071	0.028
	75	0.001	0.156	0.004	0.139	0.068	0.024
	50	0.000	0.084	0.001	0.179	0.105	0.032
	25	0.000	0.036	0.000	0.073	0.129	0.045

Table 15. Trustworthiness decision results.

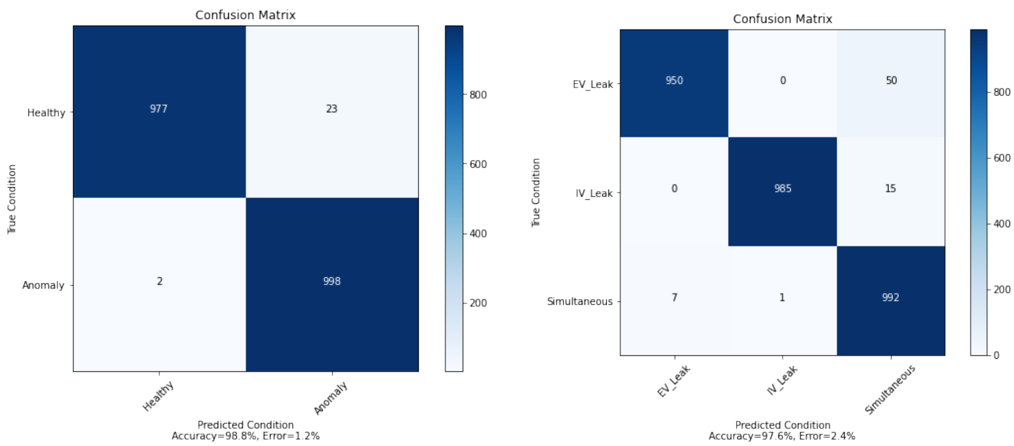
Steps	Acceptance Criteria	Trustworthiness Checks	
		Environment Conditions	Anomaly Conditions
Validation	Acceptable Errors	Pass	
Verification	Trade-off Soundness	Pass	Pass
Robustness	Uncertainty ratio	Pass	Pass

7.2. Data Generation and Application Test

Following the assurance of the developed DT trustworthiness, 79,980 data samples were generated within the considered operating conditions. Each sample contained several engine performance parameters (TC_RPM, Exh_T, CA_T, CA_P, FOC, Pmax for each cylinder). The generated datasets were randomly split for training and testing data-driven models.

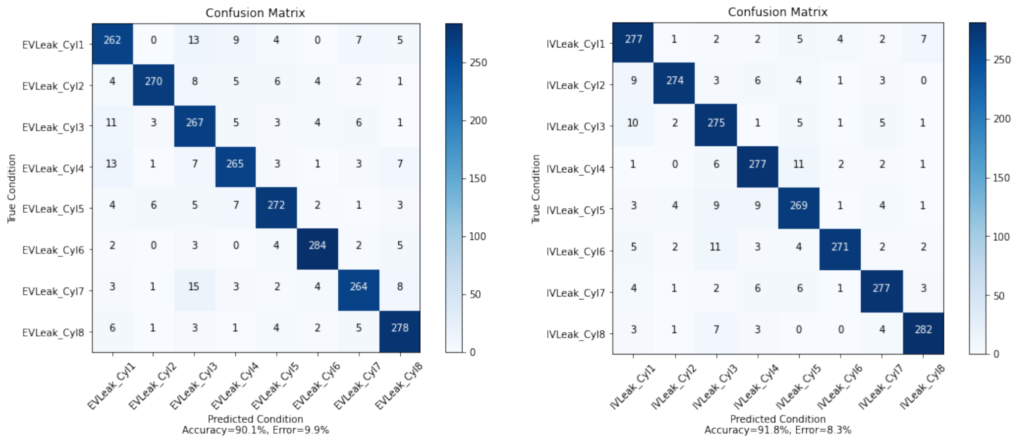
The results of the anomaly diagnosis model parts for detection, identification, and isolation in confusion matrices as well as the corresponding percentage errors as shown are presented in Figure 4. The anomaly detection model exhibits an accuracy of 98.8% with 2000 test datasets. The anomaly identification model exhibits an accuracy of 97.6% with 3000 test datasets. The anomaly isolation model exhibits an accuracy of 90.1% for determining the location (cylinder number) of the exhaust valve leakage and an accuracy of 91.8% for determining the location (cylinder number) of the intake valve leakage with 2400 test datasets.

The application test results indicate that the generated datasets are suitable for developing data-driven models. This is supported by the fact that the accuracy of the anomaly diagnosis models closely matched the values reported in the pertinent literature [65,66], which are in the range of 91.7–99.5% employing actual measurement data for binary and multi-classification with SVM.



(**a**) (**b**)

Figure 4. *Cont.*



(**c**) (**d**)

Figure 4. Anomaly diagnosis results. (**a**) Anomaly detection. (**b**) Anomaly identification. (**c**) Anomaly isolation (exhaust valve leakage). (**d**) Anomaly isolation (intake valve leakage).

7.3. Discussion

To address the challenge of limited datasets in the maritime industry, this study proposed a framework for assuring trustworthiness of physical model-based DT. This framework is expected to be a useful tool for developing trustworthy marine engine DTs considering lifetime updates (extension, modification, and addition). Although the framework was developed for marine engines, it can be extended for other marine systems, including mechanical, electrical, and hybrid components. In addition, the framework can be extended to accommodate data-driven or hybrid DTs by including the characteristics of explainability, robustness to adversarial noise, and physical plausibility.

The framework includes limitations in the selection of the thresholds used to characterise the verification and robustness of the DT. To enhance this aspect, further improvement can be achieved through extensive application and testing of the framework across diverse ship systems. Despite its limitations, the framework is valuable in facilitating the development of intelligent machinery monitoring and health management systems, which are necessary systems for smart and autonomous ships.

The framework was employed as part of a methodology aimed at generating appropriate datasets for a data-driven diagnosis model. The methodology was effective in extending the DT operating envelope considering the most critical anomalies, as identified by FMECA. This study focused on cylinder valve anomalies and ambient temperature variations; however, future studies can include anomalies of other components, environmental conditions, and their combinations. This study generated around 80,000 datasets, which required considerable computational effort for the simulation runs. A challenge that needs to be addressed in the future pertains to the computational efforts to generate datasets throughout the lifetime of the investigated system. The developed methodology can be further customised and extended for generating datasets of other machinery systems, including electric-battery propulsion systems, alternative fuel engine systems, and innovative technologies. This study employed only datasets from the engine shop trials. Future studies can also use acquired datasets during the lifetime of the system, along with their management, to improve both the quality of the DT and generated datasets.

8. Conclusions

This study developed a framework to assure the trustworthiness of a physical model-based DT for marine engines considering three phases, namely DT development, progressivity, and trustworthiness assurance. Subsequently, an application methodology was implemented, which employed the trustworthy DT framework to generate datasets for developing a data-driven anomaly diagnosis model. A marine four-stroke engine was used to demonstrate the developed framework and methodology. The main findings of this study are summarised as follows.

- The trustworthiness of the developed DT was assured as the validation step resulted in errors within $\pm 3\%$, the verification step resulted in sound trade-offs, and the robustness step confirmed the uncertainty was not propagated.
- The trustworthy DT generated 79,980 data samples representing a wide engine operating envelope. The generated dataset included ambient temperature variations and cylinder valve anomalies, making it suitable for training and testing anomaly diagnosis models.
- The accuracies of the anomaly diagnosis models were 98.8% for anomaly detection, 97.6% for anomaly identification, and 90.1–91.8% for anomaly isolation. These results demonstrated that the simulation-generated datasets can serve as viable alternatives to measurement datasets provided that the employed DT trustworthiness is appropriately assured.

This study provides insights towards DTs development and their trustworthiness management, Hence, it advances the use of tools from the fourth industrial revolution within the shipping industry, and addresses challenges pertaining to datasets generation required for developing data-driven models.

This study limitations pertain to the testing of the developed data-driven models using actual sensory data due to a lack of such information. The application of the generated datasets can be extended to prognostics and decision making cases. Future studies could consider to extend the proposed framework for data-driven and hybrid DTs as well as a wider envelope of engine operating conditions and anomaly combinations.

Author Contributions: J.J.: Conceptualisation, Methodology, Writing—original draft, Software, Visualisation, Validation. G.T.: Conceptualisation, Methodology, Validation, Writing—original draft, Writing—review and editing, Supervision, Funding acquisition. All authors have read and agreed to the published version of the manuscript.

Funding: The study was partially supported by the AUTOSHIP project funded by the European Union's Horizon 2020 research and innovation programme (agreement No. 815012) as well as the i-HEATS project funded by Innovate UK Smart Grants (project No. 99958).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data are contained within the article.

Acknowledgments: The authors affiliated with the MSRC greatly acknowledge the funding from DNV AS and RCCL for the MSRC establishment and operation. The opinions expressed herein are those of the authors and should not be construed to reflect the views of the EU, Innovate UK, DNV AS, and RCCL.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

0D/1D	0-Dimensional/1-Dimensional
ADO	Advanced Direct Optimiser
DD	Data Driven
DOE	Design of Experiments
DT	Digital Twin
FMECA	Failure Mode, Effects and Criticality Analysis
HM	Hybrid Model
ISO	International Organisation for Standardisation
LO	Lubricating Oil
MC	Monte Carlo

PHM	Prognostics and Health Management
PID	Proportional-Integral-Derivative
PM	Physical Model
RBF	Radial Basis Function
RPN	Risk Priority Number
SVM	Support Vector Machine

References

1. BahooToroody, A.; Abaei, M.M.; Banda, O.V.; Kujala, P.; De Carlo, F.; Abbassi, R. Prognostic health management of repairable ship systems through different autonomy degree; From current condition to fully autonomous ship. *Reliab. Eng. Syst. Saf.* **2022**, *221*, 108355. [CrossRef]
2. Karatuğ, Ç.; Arslanoğlu, Y.; Guedes Soares, C. Determination of a maintenance strategy for machinery systems of autonomous ships. *Ocean Eng.* **2022**, *266*, 113013. [CrossRef]
3. Ellefsen, A.L.; Æsøy, V.; Ushakov, S.; Zhang, H. A comprehensive survey of prognostics and health management based on deep learning for autonomous ships. *IEEE Trans. Reliab.* **2019**, *68*, 720–740. [CrossRef]
4. Abaei, M.M.; Hekkenberg, R.; BahooToroody, A. A multinomial process tree for reliability assessment of machinery in autonomous ships. *Reliab. Eng. Syst. Saf.* **2021**, *210*, 107484. [CrossRef]
5. Zhang, P.; Gao, Z.; Cao, L.; Dong, F.; Zou, Y.; Wang, K.; Zhang, Y.; Sun, P. Marine Systems and Equipment Prognostics and Health Management: A Systematic Review from Health Condition Monitoring to Maintenance Strategy. *Machines* **2022**, *10*, 72. [CrossRef]
6. Karatuğ, Ç.; Arslanoğlu, Y.; Soares, C.G. Review of maintenance strategies for ship machinery systems. *J. Mar. Eng. Technol.* **2023**, *22*, 233–247. [CrossRef]
7. Vanem, E.; Brandsæter, A. Unsupervised anomaly detection based on clustering methods and sensor data on a marine diesel engine. *J. Mar. Eng. Technol.* **2021**, *20*, 217–234. [CrossRef]
8. Saxena, A.; Goebel, K.; Simon, D.; Eklund, N. Damage propagation modeling for aircraft engine run-to-failure simulation. In Proceedings of the 2008 International Conference on Prognostics and Health Management, Denver, CO, USA, 6–9 October 2008; pp. 1–9.
9. Basurko, O.C.; Uriondo, Z. Condition-based maintenance for medium speed diesel engines used in vessels in operation. *Appl. Therm. Eng.* **2015**, *80*, 404–412. [CrossRef]
10. Coraddu, A.; Oneto, L.; Ilardi, D.; Stoumpos, S.; Theotokatos, G. Marine dual fuel engines monitoring in the wild through weakly supervised data analytics. *Eng. Appl. Artif. Intell.* **2021**, *100*, 104179. [CrossRef]
11. Raptodimos, Y.; Lazakis, I. Using artificial neural network-self-organising map for data clustering of marine engine condition monitoring applications. *Ships Offshore Struct.* **2018**, *13*, 649–656. [CrossRef]
12. Lazakis, I.; Gkerekos, C.; Theotokatos, G. Investigating an SVM-driven, one-class approach to estimating ship systems condition. *Ships Offshore Struct.* **2019**, *14*, 432–441. [CrossRef]
13. VanDerHorn, E.; Wang, Z.; Mahadevan, S. Towards a digital twin approach for vessel-specific fatigue damage monitoring and prognosis. *Reliab. Eng. Syst. Saf.* **2022**, *219*, 108222. [CrossRef]
14. Lebedevas, S.; Žaglinskis, J.; Drazdauskas, M. Development and Validation of Heat Release Characteristics Identification Method of Diesel Engine under Operating Conditions. *J. Mar. Sci. Eng.* **2023**, *11*, 182. [CrossRef]
15. Tsitsilonis, K.M.; Theotokatos, G.; Patil, C.; Coraddu, A. Health assessment framework of marine engines enabled by digital twins. *Int. J. Engine Res.* **2022**, *24*, 3264–3281. [CrossRef]
16. Altosole, M.; Donnarumma, S.; Spagnolo, V.; Vignolo, S. Performance Simulation of Marine Cycloidal Propellers: A Both Theoretical and Heuristic Approach. *J. Mar. Sci. Eng.* **2022**, *10*, 505. [CrossRef]
17. Coraddu, A.; Kalikatzarakis, M.; Theotokatos, G.; Geertsma, R.; Oneto, L. Physical and data-driven models hybridisation for modelling the dynamic state of a four-stroke marine diesel engine. In *Engine Modeling and Simulation*; Springer: Singapore, 2022.
18. Altosole, M.; Balsamo, F.; Acanfora, M.; Mocerino, L.; Campora, U.; Perra, F. A Digital Twin Approach to the Diagnostic Analysis of a Marine Diesel Engine. In *Technology and Science for the Ships of the Future*; Progress in Marine Science and Technology; IOS Press: Amsterdam, The Netherlands, 2022; Volume 6, pp. 198–206. [CrossRef]
19. Djeziri, M.A.; Benmoussa, S.; Benbouzid, M.E. Data-driven approach augmented in simulation for robust fault prognosis. *Eng. Appl. Artif. Intell.* **2019**, *86*, 154–164. [CrossRef]
20. Stoumpos, S.; Theotokatos, G. A novel methodology for marine dual fuel engines sensors diagnostics and health management. *Int. J. Engine Res.* **2021**, *23*, 974–994. [CrossRef]
21. Schneider, F.B.; Bellovin, S.M.; Inouye, A.S. Building trustworthy systems: Lessons from the PTN and Internet. *IEEE Internet Comput.* **1999**, *3*, 64–72. [CrossRef]
22. Sargent, R.G. Verification and validation of simulation models. In Proceedings of the 2010 Winter Simulation Conference, Baltimore, MD, USA, 5–8 December 2010; pp. 166–183.
23. Huang, X.; Kroening, D.; Ruan, W.; Sharp, J.; Sun, Y.; Thamo, E.; Wu, M.; Yi, X. A survey of safety and trustworthiness of deep neural networks: Verification, testing, adversarial attack and defence, and interpretability. *Comput. Sci. Rev.* **2020**, *37*, 100270. [CrossRef]

24. Wang, B.T.; Burdon, M. Automating trustworthiness in digital twins. In *Advances in 21st Century Human Settlements*; Springer: Singapore, 2021; pp. 345–365. [CrossRef]
25. Babiceanu, R.F.; Seker, R. Trustworthiness requirements for manufacturing cyber-physical systems. *Procedia Manuf.* **2017**, *11*, 973–981. [CrossRef]
26. de la Vara, J.L.; Ruiz, A.; Blondelle, G. Assurance and certification of cyber–physical systems: The AMASS open source ecosystem. *J. Syst. Softw.* **2021**, *171*, 110812. [CrossRef]
27. Cerbo, F.D.; Bisson, P.; Hartman, A.; Keller, S.; Meland, P.H.; Moffie, M.; Mohammadi, N.G.; Paulus, S.; Short, S. Towards trustworthiness assurance in the cloud. In *Cyber Security and Privacy Forum*; Springer: Berlin/Heidelberg, Germany, 2013; pp. 3–15.
28. Connett, B.; O'Halloran, B. Systems engineering design: Architecting trustworthiness in cyber physical systems using an extended aggregated modality. *Procedia Comput. Sci.* **2018**, *140*, 4–12. [CrossRef]
29. Wärtsilä. Wärtsilä 50DF Product Guide. 2012. Available online: https://cdn.wartsila.com/docs/default-source/product-files/engines/df-engine/product-guide-o-e-w50df.pdf?sfvrsn=9 (accessed on 7 March 2024).
30. Theotokatos, G.; Stoumpos, S.; Bolbot, V.; Boulougouris, E. Simulation-based investigation of a marine dual-fuel engine. *J. Mar. Eng. Technol.* **2020**, *19*, 5–16. [CrossRef]
31. Gamma Technologies. GT-SUITE Manual. 2022 .
32. Stoumpos, S.; Theotokatos, G.; Boulougouris, E.; Vassalos, D.; Lazakis, I.; Livanos, G. Marine dual fuel engine modelling and parametric investigation of engine settings effect on performance-emissions trade-offs. *Ocean Eng.* **2018**, *157*, 376–386. [CrossRef]
33. Stoumpos, S.; Theotokatos, G.; Mavrelos, C.; Boulougouris, E. Towards marine dual fuel engines digital twins–integrated modelling of thermodynamic processes and control system functions. *J. Mar. Sci. Eng.* **2020**, *8*, 200. [CrossRef]
34. Woschni, G. *A Universally Applicable Equation for the Instantaneous Heat Transfer Coefficient in the Internal Combustion Engine*; Technical Report, SAE Technical Paper; SAE International: Warrendale, PA, USA, 1967.
35. Merker, G.P.; Schwarz, C.; Stiesch, G.; Otto, F. *Simulating Combustion: Simulation of Combustion and Pollutant Formation for Engine-Development*; Springer: Berlin/Heidelberg, Germany, 2005.
36. Jerath, K.; Brennan, S.; Lagoa, C. Bridging the gap between sensor noise modeling and sensor characterization. *Measurement* **2018**, *116*, 350–366. [CrossRef]
37. Borgovini, R.; Pemberton, S.; Rossi, M. *Failure Mode, Effects, and Criticality Analysis (FMECA)*; Reliability Analysis Center: Rome, NY, USA, 1993.
38. Liu, H.C. *FMEA Using Uncertainty Theories and MCDM Methods*; Springer: Singapore, 2016; pp. 13–27. [CrossRef]
39. Andradóttir, S. Simulation optimization. In *Handbook of Simulation: Principles, Methodology, Advances, Applications, and Practice*; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 1998; pp. 307–333.
40. Deb, K.; Jain, H. An evolutionary many-objective optimization algorithm using reference-point-based nondominated sorting approach, part I: Solving problems with box constraints. *IEEE Trans. Evol. Comput.* **2013**, *18*, 577–601. [CrossRef]
41. Cox, D.R.; Reid, N. *The Theory of the Design of Experiments*; Chapman and Hall/CRC: Boca Raton, FL, USA, 2000.
42. Antony, J. *Design of Experiments for Engineers and Scientists*; Elsevier: Amsterdam, The Netherlands, 2014.
43. Huntington, D.E.; Lyrintzis, C.S. Improvements to and limitations of Latin hypercube sampling. *Probabilistic Eng. Mech.* **1998**, *13*, 245–253. [CrossRef]
44. *ISO 15550*; Internal Combustion Engines: Determination and Method for the Measurement of Engine Power-General Requirements. ISO: Geneva, Switzerland, 2016.
45. Kleijnen, J.P.C. Sensitivity Analysis of Simulation Models. *CentER Discuss. Pap. Ser.* **2009**. [CrossRef]
46. Hauke, J.; Kossowski, T. Comparison of values of Pearson's and Spearman's correlation coefficients on the same sets of data. *Quaest. Geogr.* **2011**, *30*, 87. [CrossRef]
47. Xiao, C.; Ye, J.; Esteves, R.M.; Rong, C. Using Spearman's correlation coefficients for exploratory data analysis on big dataset. *Concurr. Comput. Pract. Exp.* **2016**, *28*, 3866–3878. [CrossRef]
48. Vrijdag, A.; Schulten, P.; Stapersma, D.; Van Terwisga, T. Efficient uncertainty analysis of a complex multidisciplinary simulation model. *J. Mar. Eng. Technol.* **2007**, *6*, 79–88. [CrossRef]
49. Smith, E. Uncertainty analysis. *Encycl. Environmetrics* **2002**, *4*, 2283–2297.
50. Cox, D.C.; Baybutt, P. Methods for uncertainty analysis: A comparative survey. *Risk Anal.* **1981**, *1*, 251–258. [CrossRef]
51. Zheng, A.; Casari, A. *Feature Engineering for Machine Learning: Principles and Techniques for Data Scientists*; O'Reilly Media, Inc.: Sebastopol, CA, USA, 2018.
52. Yin, Z.; Hou, J. Recent advances on SVM based fault diagnosis and process monitoring in complicated industrial processes. *Neurocomputing* **2016**, *174*, 643–650. [CrossRef]
53. Cai, C.; Weng, X.; Zhang, C. A novel approach for marine diesel engine fault diagnosis. *Clust. Comput.* **2017**, *20*, 1691–1702. [CrossRef]
54. Tan, Y.; Zhang, J.; Tian, H.; Jiang, D.; Guo, L.; Wang, G.; Lin, Y. Multi-label classification for simultaneous fault diagnosis of marine machinery: A comparative study. *Ocean Eng.* **2021**, *239*, 109723. [CrossRef]
55. Mulgrew, B. Applying radial basis functions. *IEEE Signal Process. Mag.* **1996**, *13*, 50–65. [CrossRef]
56. Wainer, J.; Fonseca, P. How to tune the RBF SVM hyperparameters? An empirical evaluation of 18 search algorithms. *Artif. Intell. Rev.* **2021**, *54*, 4771–4797. [CrossRef]

57. Hsu, C.W.; Lin, C.J. A comparison of methods for multiclass support vector machines. *IEEE Trans. Neural Netw.* **2002**, *13*, 415–425. [CrossRef]
58. Heydarian, M.; Doyle, T.E.; Samavi, R. MLCM: Multi-label confusion matrix. *IEEE Access* **2022**, *10*, 19083–19095. [CrossRef]
59. Ftoutou, E.; Chouchane, M.; Besbès, N. Internal combustion engine valve clearance fault classification using multivariate analysis of variance and discriminant analysis. *Trans. Inst. Meas. Control* **2012**, *34*, 566–577. [CrossRef]
60. Kowalski, J. An experimental study of emission and combustion characteristics of marine diesel engine in case of cylinder valves leakage. *Pol. Marit. Res.* **2015**, *22*, 90–98. [CrossRef]
61. Rubio, J.A.P.; Vera-García, F.; Grau, J.H.; Cámara, J.M.; Hernandez, D.A. Marine diesel engine failure simulator based on thermodynamic model. *Appl. Therm. Eng.* **2018**, *144*, 982–995. [CrossRef]
62. Whitehouse, N.D.; Stotter, A.; Janota, M.S. Estimating the effects of altitude, ambient temperature and turbocharger match on engine performance. *Proc. Inst. Mech. Eng.* **1963**, *178*, 483–500. [CrossRef]
63. Serrano, J.R.; Guardiola, C.; Dolz, V.; Tiseira, A.; Cervelló, C. *Experimental Study of the Turbine Inlet Gas Temperature Influence on Turbocharger Performance*; Technical Report, SAE Technical Paper; SAE International: Warrendale, PA, USA, 2007.
64. MAN. *Influence of Ambient Temperature Conditions*; MAN Diesel and Turbo: Copenhagen, Denmark, 2014.
65. Zhan, Y.; Shi, Z.; Liu, M. The application of support vector machines (SVM) to fault diagnosis of marine main engine cylinder cover. In Proceedings of the IECON 2007—33rd Annual Conference of the IEEE Industrial Electronics Society, Taipei, Taiwan, 5–8 November 2007; pp. 3018–3022.
66. Tong, S.; Yanqiao, C.; Yuan, Z. Fault prediction of marine diesel engine based on time series and support vector machine. In Proceedings of the 2020 International Conference on Intelligent Design, ICID 2020, Xi'an, China, 11–13 December 2020. [CrossRef]