

Fault detection in the marine engine using a support vector data description method

Klaudia Wrzask¹, Jerzy Kowalski^{1,*}, Van Vang Le², Van Giao Nguyen³, Dao Nam Cao^{4,*}

¹*Gdansk University of Technology, Gdansk, Poland.*

²*Institute of Maritime, Ho Chi Minh City University of Transport, Ho Chi Minh City, Viet Nam*

³*Institute of Engineering, HUTECH University, Ho Chi Minh City, Viet Nam*

⁴*PATET Research Group, Ho Chi Minh City University of Transport, Ho Chi Minh City, Viet Nam*

***Corresponding authors**

***Jerzy Kowalski (jerzy.kowalski@pg.edu.pl)**

***Dao Nam Cao (nam.cao@ut.edu.vn)**

Abstract

Fast detection and correct diagnosis of any engine condition changes are essential elements of safety and environmental protection. Many diagnostic algorithms significantly improve the detection of malfunctions. Studies on diagnostic methods are rarely reported and even less implemented in the marine engine industry. To fill this gap, this paper presents the Support Vector Data Description (SVDD) method as applied to the fault detection of the fuel delivery system of two-stroke marine engine. The selected diagnostic data is the exhaust gas composition, with four components considered: oxygen, carbon oxide, nitric oxide, and carbon dioxide. With these diagnostics, the method distinguishes eight different engine faults from the efficient state. The manuscript presents in detail the methodology for applying the SVDD method in a marine engine. The method of obtaining diagnostic data and its scaling is described. The method of training and validating the algorithm is also presented, along with ready-made algorithms for use. The 100% accuracy of the proposed fault detection method. Based on the obtained results, the proposed fault detection method is promising for a simple application. Authors also present generalized algorithms that may be adapted to different technical solutions.

Keywords: Marine Engine; Engine fault detection; Diagnostic methods; Diagnostic algorithms; Support vector data description

Highlights

- SVDD was used for marine engine fault detection from exhaust gas composition
- Laboratory measurements were carried out on the two-stroke diesel engine
- The proposed algorithm detected the considered faults with 100% accuracy
- Generalized algorithm for adapting other complex technical objects was proposed

1. Introduction

Effective and reliable operations of complex technical facilities are closely dependent on their design, technical condition, and progression of degradation over time. One such complex technical object is the internal combustion engine, which is still one of the most popular mechanical energy sources, used both in the transport and energy industries (Hoang, 2019)(Hoang, 2020). In the case of internal combustion engines, the wear and corrosion of their structural elements leads to a significant power loss, an increase in fuel consumption, and an increase in the emission of harmful exhaust gas components into the atmosphere (Ayyagari et al., 2018)(Hoang et al., 2020). Therefore, an important element of both safety and environmental protection is the earliest possible detection and correct diagnosis of any changes in the engine's technical condition (Cai et al., 2020). In the case of engines used in shipbuilding, but also in other transport branches, the primary diagnostic tool is failure detection signaling based on exceeding the threshold value of the selected diagnostic symptom (Cao et al., 2015), as well as the use of low-carbon fuels for marine engine (Hoang et al., 2023). An example would be measuring the exhaust gas temperature behind the cylinder of a marine piston engine. Exceeding the temperature over the maximum value indicates a possible failure of the operator (Korczewski, 2015). On the other hand, the so-called slow steaming is a more and more commonly used method for reducing fuel consumption by ships. Slow steaming reduces the ship's speed when the main propulsion engine is running with partial load (Borkowski et al., 2019)(Hoang et al., 2022). As a result, the engine load is so low that even significant damage to the fueling apparatus does not cause the exhaust gas temperature to increase above the threshold value and therefore the alarm signaling is not triggered (Kalghatgi, 2013). An alternative is shown by Puzdrowska (Puzdrowska, 2022), where the quickly changing temperature of the exhaust gas was used as the diagnostic signal. Using a suitably "fast" thermocouple instead of a standard one during engine testing provided much more detailed diagnostic information.

In the field of science, many algorithms significantly improved the detection of malfunctions, yet they have not been widely used in engineering practice so far. The use of a mathematical model as a diagnostic tool is presented in studies of Lamaris et al. (Lamaris and Hountalas, 2010), Tamura et al. (Tamura et al., 2011), and Scappin et al. (Scappin et al., 2012). The aforementioned models use the measured signals as the input data, translating them into a virtual image of the engine's operation. Unfortunately, the use of online calculations requires significant simplification of the model. In order to increase the accuracy of the modeling results, approximating models based on Artificial Neural Networks (ANN) are applied increasingly (Basurko and Uriondo, 2015; Lazakis et al., 2018). Correctly trained ANN can be used in operational conditions, although only for specific learning solutions (Rudzki et al., 2022). Changing the configuration of a complex technical object usually leads to a long and costly process of ANN re-learning (Noyel et al., 2016). In recent years, a significant acceleration in the development of Machine Learning (ML) algorithms for diagnostic purposes has been noted. Wang et al. (Wang et al., 2013) apply adaptive wavelet packets and EEMD-fractal dimension to the vibration signal analysis. Kowalski (Kowalski, 2015) introduced the concept of the multidimensional diagnostic tool based on the analysis of exhaust gas composition. Further, execution of the process by an ensemble of Extreme Learning Machine (ELM) was presented by Kowalski et al. (Kowalski et al., 2017). Cai et al. (Cai et al., 2020) suggested applying a Bayesian network to vibration signal analysis for the purpose of technical diagnosis. Liang et al. (Liang et al., 2020) proposed a fault diagnostic method in power converters for wind power generation systems. Namigtle-Jimenez et al. (Namigtle-Jiménez et al., 2020) formulated and experimentally applied a fault detection and diagnosis scheme based on ANN for isolating the damaged injector of an internal combustion engine. Tan et al. (Tan et al., 2020) investigated data-driven approaches, based on one-class classifiers, among all SVDD, for condition monitoring of marine machinery systems. As for marine diesel engines, Liang et al. (Liang and



Chen, 2022) proposed a Big Data method based on a Gaussian Mixture Model and Density-Based Spatial Clustering to monitor the degree of degradation of the main engine performance. A comprehensive analysis of more ML algorithms applications for diagnostics and control is presented by Aliramezani et al. (Aliramezani et al., 2022).

Support Vector Data Description (SVDD) has also been successfully adopted for fault detection and diagnosis, including anomaly detection Banerjee et al. (Banerjee et al., 2007), analog circuit faults detection Luo et al. (Luo and Jiang Cui, 2011), chillers fault detection Zhao et al. (Zhao et al., 2013), fault diagnosis of rotating machinery such as rolling bearings and gears, widely used in wind turbines, helicopters, high-speed, railways and cranes Zhang et al. (Zhang et al., 2022), or fault diagnosis of gears Tang et al. (Tang et al., 2022). Regrettably, despite the development of and widespread access to more and more perfect diagnostic algorithms, the simplest and often unreliable diagnostic procedures are still widely used. The authors see the reason for this state of affairs in the insufficient popularization, among other things, of the already developed algorithms. The algorithms mentioned above are often presented on the basis of a complex mathematical apparatus, which may cause difficulties when adapted to specific technical solutions. For this reason, this work aims to develop a robust algorithm for recognizing selected faults of a Diesel engine similar in design to the two-stroke engines used in the main propulsion system of a ship. The contribution of this paper is the application of the SVDD method to new diagnostic data: the molar compositions of the exhaust gas components, to fault monitoring the overall fuel delivery system of a marine diesel engine. Both the application of the SVDD algorithm for the diagnosis of fuel delivery system of marine engines and the use of exhaust gas components for this purpose as carriers of diagnostic information for marine engines are not widely reflected in the literature. This work fills the gap for rarely reported robust fault detection algorithms in the marine engine industry. When compared to other diagnostic methods such as ANN, the choice of



this method is dictated by the multidimensionality of the input signal and the fast computation of the geometric description of the fault spaces. Authors also propose adapting this mathematical method to other complex technical objects or selected diagnostic signals by presenting generalized fault detection algorithms. This paper has been organized as follows. Section 2 presents the experimental setup. In Section 3, the SVDD method for fault detections is presented. As the success of the diagnostics is warranted by applying the right method to the data set characteristics, in Section 4.1 the data sets are described and analyzed. Section 4.2 focuses on the training procedure for calculating the boundary hyperspheres describing the faults. In Section 4.3, we describe a pretesting procedure to determine the main diagnostic parameter. Section 4.4 presents the validation of the method. The final conclusions are presented in Section 5.

2. Experiment setup

The data for the presented diagnostic model was collected during tests on a single-cylinder, two-stroke crosshead engine with longitudinal scavenging, loaded by a water brake. The engine parameters are presented in **Table 1**.

Table 1. Laboratory engine parameters

Nominal power [kW]	73.5
Maximum rotational speed [rpm]	320
Number of cylinders [-]	1
Cylinder bore [mm]	220
Piston stroke [mm]	350
Compression ratio [-]	18.5
Fuel delivery system	Bosh type pump, 1 injector
Specific fuel consumption at maximum load [g/kWh]	277.6

Two-stroke engines are commonly used in the main propulsion system of ships due to their high energy efficiency and low rotational speed required for marine propeller propulsion. An independently driven Roots blower charges the engine allowing for smooth regulation of its rotational speed. Single-cylinder engine design allows for undisturbed monitoring of the engine combustion process, the apparatus is shown in **Figure 1a**.

For the experiment, Authors used marine diesel oil with a density of 829.6 kg/m^3 at 15°C . The rationale behind choosing this type of fuel was to eliminate the impact of nitrogen content in the fuel at the level of nitric oxide emission. The engine was loaded to the maximum value at each observation during laboratory tests. The loads and rotational speeds were then reduced according to the marine propeller load characteristics, typical for two-stroke engines used in the ship propulsion system. The load pattern is presented in **Table 2** and conforms to the E2 characteristics of the ISO8178 standard for emission testing.

Table 2. The test cycle pattern

Load [%]	100	87.5	75.0	62.5	50.0	37.5	25.0
Speed [rpm]	320	305	290	270	250	225	200

Each measurement was performed after stabilization of the gas temperature at the exhaust manifold. We recorded 5 to 10 minutes of measurements with a 1-second sampling time. Each observation was repeated three times. The diagnostic signals chosen were fractions of CO , CO_2 , O_2 , and nitric oxides as a sum of NO and NO_2 . Exhaust gas components were measured for each considered load of the engine. The exhaust gas analysis process is shown in **Figure 1b**.

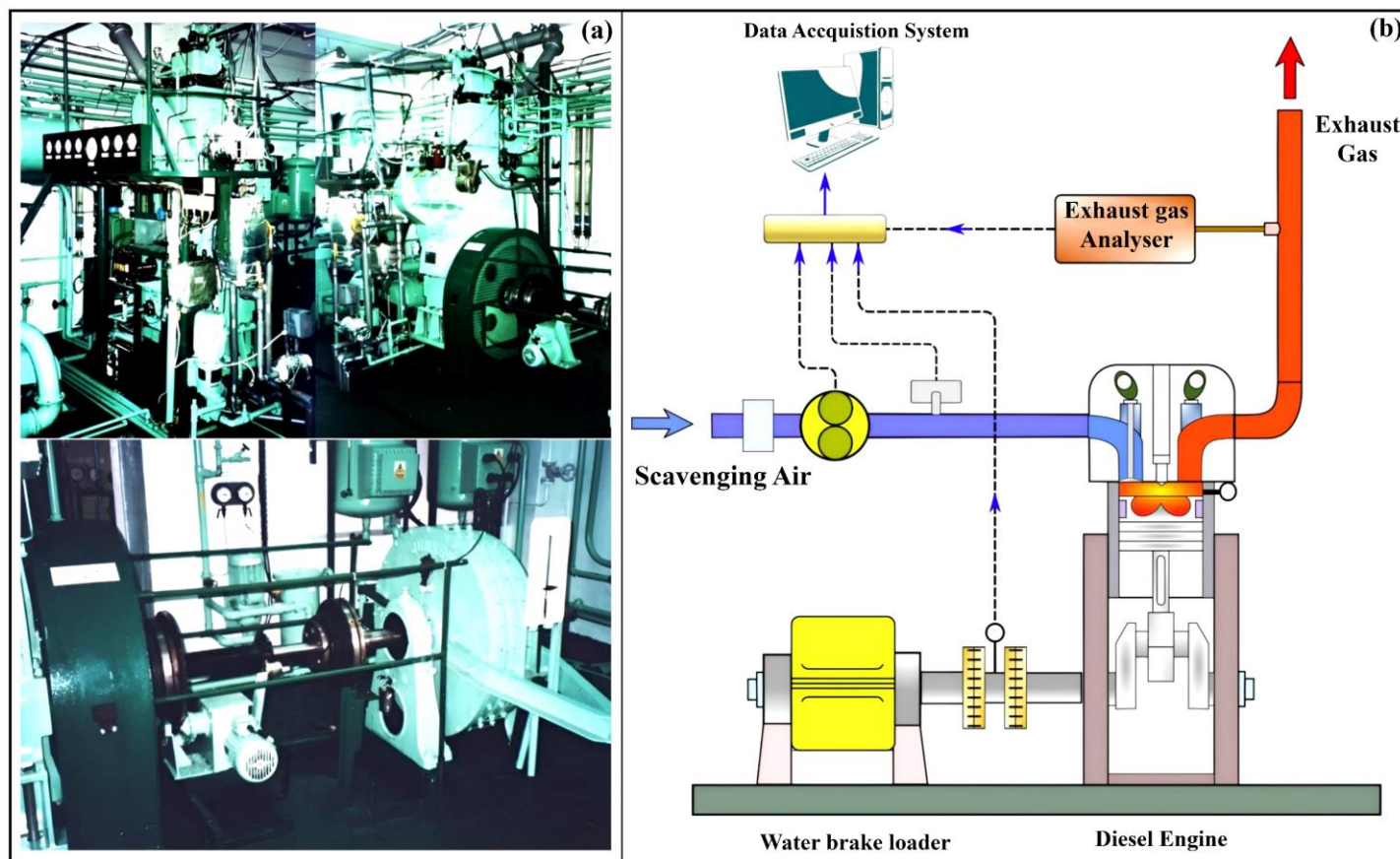


Figure 1. (a) Laboratory engine (top), water brake (bottom); (b) Exhaust gas analysis process

The content of the above chemical compounds strictly depends on the combustion process parameters. In general, the carbon compound content is a good indicator of combustion process efficiency and fuel consumption; the oxygen content depends on the combustion process parameters as well as the parameters of the engine scavenging process, while the nitric oxide content depends on the combustion process temperature. The rotational speed and engine load were measured to ensure the E2 engine test cycle operation. **Table 3** shows the range and accuracy of the measurement apparatus.

Table 3. Measurement apparatus parameters

Parameter	Range	Accuracy
Rotational speed of the engine	0 – 600 rpm	1 rpm
The engine load	0 – 80 kW	±1.0%
CO fraction in exhaust gas	0 – 10 000 ppm	±5.0%
NO fraction in exhaust gas	0 – 3000 ppm	±5.0%
NO ₂ fraction in exhaust gas	0 – 500 ppm	±5.0%
CO ₂ fraction in exhaust gas	0 – 50%	±0.3%
O ₂ fraction in exhaust gas	0 – 25%	±0.2%

During the measurements, the engine operating parameters were modified to simulate selected faults. The exact parameter modifications for each fault are presented in **Table 4**.

Table 4. The engine faults simulation description

Fault	Code description
F_1	The technical condition of the engine is assumed as proper
F_2	The start of the fuel injection is shifted from 7° to 4° BTDC
F_3	The start of the fuel injection is shifted from 7° to 10° BTDC
F_4	The opening pressure of the fuel injector is shifted from 22 MPa to 12 MPa
F_5	The opening pressure of the fuel injector is shifted from 22 MPa to 27 MPa
F_6	The opening pressure of the fuel injector is shifted from 22 MPa to 32 MPa
F_7	The throttling of the exhaust gas duct by 63% of the cross-section area
F_8	The throttling of the exhaust gas duct by 25% of the cross-section area
F_9	The reduction of the scavenging blower from a nominal 910 rpm to 450 rpm

3. Support vector data description

Data domain description comprises data set characterization and classification. To detect novel data or outliers, the boundary of a data set can be used. This procedure should cover most of

the target data while including the least redundant sets. The support vector data description (SVDD) method was proposed first by Tax and Duin (Tax and Duin, 1999) and elaborated in (Tax and Duin, 2004), this method was inspired by the Support Vector Classifier (Chapelle et al., 1999). SVDD achieves optimization for a boundary around the data set, which is spherically shaped. To introduce this method, let us first consider a set of N data objects:

$$\{\mathbf{x}_i \in R^n, \quad i = 1, \dots, N\}. \quad (1)$$

The goal is to find a minimum volume hypersphere containing most of the data objects. **Figure 2** shows an illustrative example. Here, some data objects are allowed to be outside the sphere. This may be an advantage in respect of classification problems because the smaller the volume of the sphere, the less it is going to intersect with other classifiers. For this purpose, slack variables ξ_i are introduced.

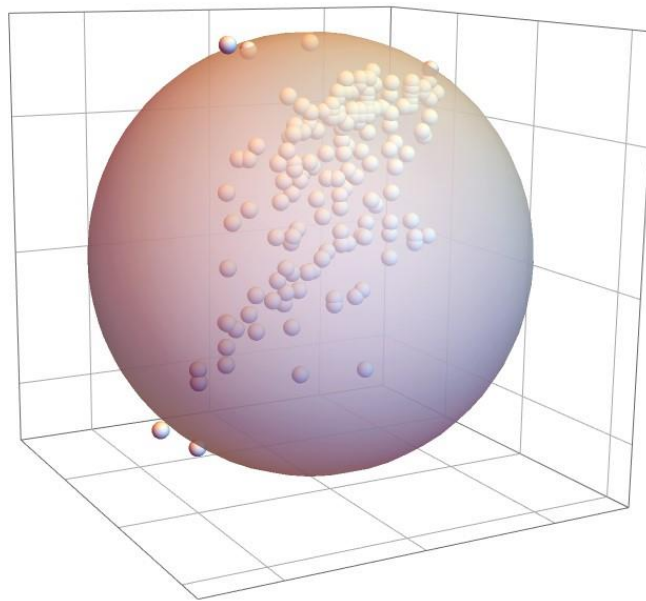


Figure 2. Boundary sphere obtained by the SVDD method: some data objects are allowed to be outside the boundary

Furthermore, the sphere is described by the center \mathbf{a} and radius R , such that:

$$H(\mathbf{a}, R, \xi_i) = R^2 + C \sum_i \xi_i, \quad (2)$$

where the variable C controls the trade-off between the volume of the hypersphere and the number of rejected objects (outside the sphere) and the slack variables are denoted by ξ_i . Hypersphere (2) should be minimized for each data object \mathbf{x}_i under the following constraints:

$$(\mathbf{x}_i - \mathbf{a})^T (\mathbf{x}_i - \mathbf{a}) \leq R^2 + \xi_i, \quad (3)$$

$$\xi_i \geq 0. \quad (4)$$

Constraints (3)-(4) refer to each of the data objects indexed $i = 1, \dots, N$, and give $2N$ inequalities. When including these constraints to the hypersphere (2), the following Lagrangian $L(\mathbf{a}, R, \xi_i, \alpha_i, \beta_i)$ is constructed:

$$\begin{aligned} L(\mathbf{a}, R, \xi_i, \alpha_i, \beta_i) &= R^2 + C \sum_i \xi_i - \sum_i \alpha_i [R^2 + \xi_i - (\mathbf{x}_i^2 - 2\mathbf{a} \cdot \mathbf{x}_i + \mathbf{a}^2)] \\ &\quad - \sum_i \beta_i \xi_i. \end{aligned} \quad (5)$$

Here, α_i and β_i are the Lagrange multipliers such that $\alpha_i \geq 0$ and $\beta_i \geq 0$. Setting the partial derivatives of the Lagrangian (5) to zero gives the following equations:

$$\frac{\partial L}{\partial R} = 0 \Rightarrow \sum_i \alpha_i = 1, \quad (6)$$

$$\frac{\partial L}{\partial \mathbf{a}} = 0 \Rightarrow \mathbf{a} = \frac{\sum_i \alpha_i \mathbf{x}_i}{\sum_i \alpha_i} = \sum_i \alpha_i \mathbf{x}_i, \quad (7)$$

$$\frac{\partial L}{\partial \xi_i} = 0 \Rightarrow C - \alpha_i - \beta_i = 0. \quad (8)$$

The next step is rewriting the Lagrangian (5) with the constraints (6)-(8), i.e.,

$$L = \sum_i \alpha_i (\mathbf{x}_i \cdot \mathbf{x}_i) - \sum_{i,j} \alpha_i \alpha_j (\mathbf{x}_i \cdot \mathbf{x}_j), \quad (9)$$

where $0 \leq \alpha_i \leq C$ and $\sum_i \alpha_i = 1$.



Maximizing (9) results in the Lagrangian multipliers, α_i that are the weight factors for the center of the hypersphere (7). For a small set of vectors, the equality in (3) will be satisfied; they are called the support vectors $\tilde{\mathbf{x}}_i$. These are the data objects which are on the boundary of the hypersphere. For those vectors, α_i multipliers will be non-zero and $\beta_i = 0$. These are the only vectors needed to describe the hypersphere's radius R , which can be calculated as the distance from the center of the hypersphere to the support vector with a weight smaller than C , i.e.,

$$\|\tilde{\mathbf{x}}_i - \mathbf{a}\|^2 = R^2 \Rightarrow 0 < \alpha_i < C, \quad \beta_i = 0. \quad (10)$$

Thus, all the vectors for which $\alpha_i = 0$ are inside the boundary, that is

$$\|\mathbf{x}_i - \mathbf{a}\|^2 < R^2 \Rightarrow \alpha_i = 0, \quad \beta_i = 0. \quad (11)$$

Finally vectors for which $\alpha_i = C$ are outside the hypersphere's boundary, i.e.

$$\|\mathbf{x}_i - \mathbf{a}\|^2 > R^2 \Rightarrow \alpha_i = C, \quad \beta_i > 0. \quad (12)$$

Having calculated the center and the radius of the hypersphere, Authors can determine whether a novel test point \mathbf{z} is within the radius; in such a case, the distance to the center of the hypersphere has to be less or equal to the radius R , i.e.,

$$(\mathbf{z} - \mathbf{a})^T(\mathbf{z} - \mathbf{a}) \leq R^2. \quad (13)$$

The presented mathematical description allows us to determine the location of the minimum size of the hypersphere in n-dimensional space. The mentioned hypersphere contains the values of engine diagnostic parameters, allowing for identification of the failure. Therefore, it is important to select the appropriate diagnostic signals that clearly identify the failure, otherwise the hyperspheres may penetrate each other. Such interference means the possibility of ambiguous classification.

4. Results and discussions

4.1. Experimental data

The experiments brought out 22 345 five-dimensional data points for nine engine faults.

Table 5 shows the number of 5D points for each fault space. Further, this data has been divided into three sets: the training set, the pretesting set, and the final testing set.

Table 5. Number of five-dimensional points for each fault $F_1 - F_9$

Fault	F_1	F_2	F_3	F_4	F_5	F_6	F_7	F_8	F_9
5D points	3270	3031	3002	2827	848	656	3011	2729	2971

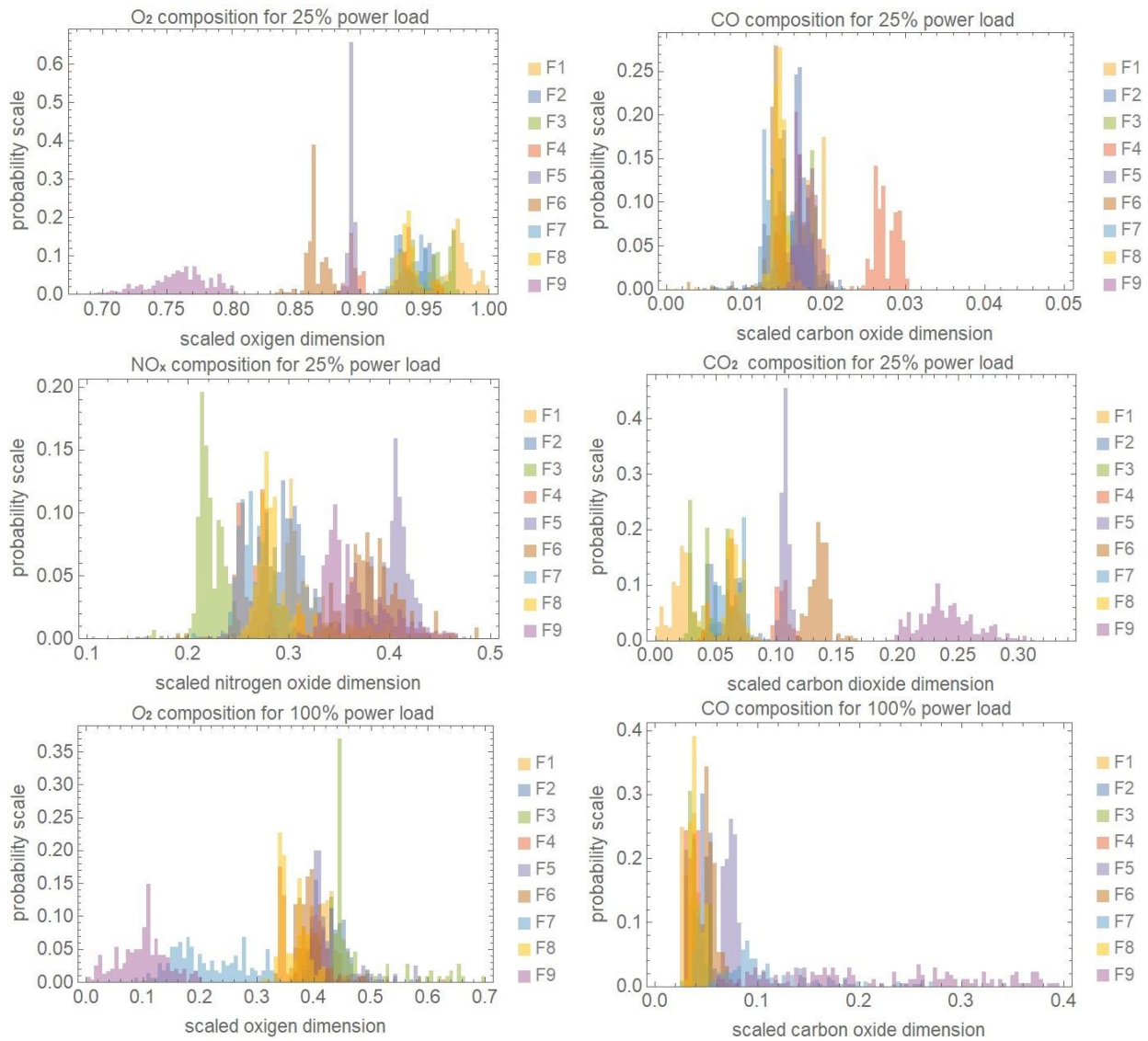
One of these dimensions is reserved for the parameter that can be set in the experiment, i.e., the engine power load of 25%, 37.5%, 50%, 62.5%, 75%, 87.5%, 100%. The other four dimensions are reserved for emission exhaust gases O₂, CO, NO_x, and CO₂. Space F_1 stands for the efficient engine, whereas spaces F_2 – F_9 indicate the faulty states. In the proposed model, the number of dimensions is adjusted to the research object. In the case of engine loading, such as with a fixed propeller, an additional dimension needs to be introduced. In this case, the engine may have variable load at a constant rotational speed. The presented research object is merely a simplified example of application. In real technical objects, multi-cylinder engine designs are used. In such cases, for fault detection purposes, additional signals measured on each cylinder are necessary, such as exhaust gas temperature or vibration signal. Consequently, the number of dimensions will increase proportionally with the number of cylinders. The most important aspect of selected diagnostic signals is their mutual independence. **Table 6** gives the overall data characteristics: the maximum and minimum values, the mean value, and mean deviation for each exhaust gas dimension.

Table 6. Exhaust gas data characteristics

Gas	O ₂ %	CO ppm	NO _x ppm	CO ₂ %
minimum value	12.88	23.	76.	0.82

maximum value	19.5	3557.	545.	4.45
mean	17.69	103.06	311.92	1.81
mean deviation	1.11	48.64	72.15	0.61

The maximum and minimum values for these four dimensions fall within a wide range. This is the main motivation for scaling each exhaust gas dimension for the same range $\langle 0,1 \rangle$. Moreover, **Figure 3** shows joined histograms of nine fault spaces for exhaust gas composition after the scaling procedure.



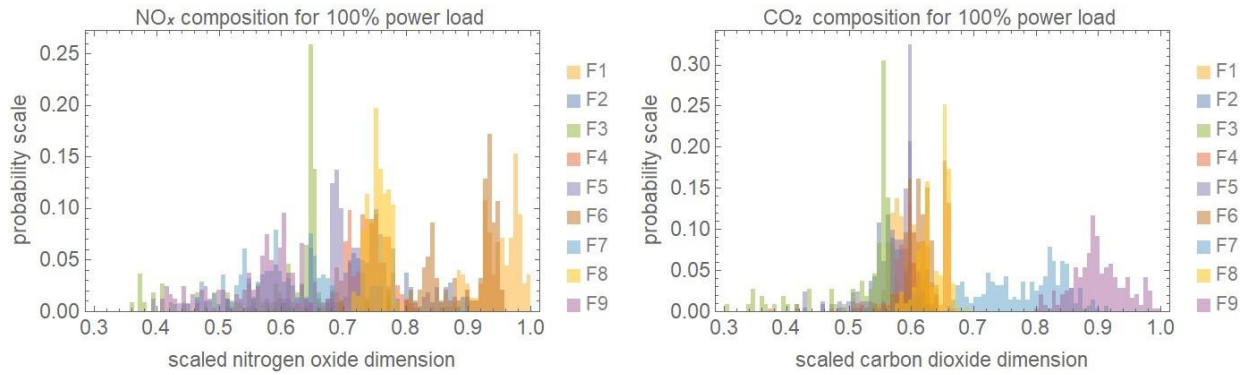


Figure 3. Histograms of four exhaust gas compositions after the scaling procedure for the 25% and 100% power load. Nine fault spaces are distinguished with different colors

Further, we show random points of all nine fault spaces for the selected 3D exhaust gas dimensions in **Figure 4**. This was done after the scaling procedure with minimum and maximum values shown in **Table 6**.

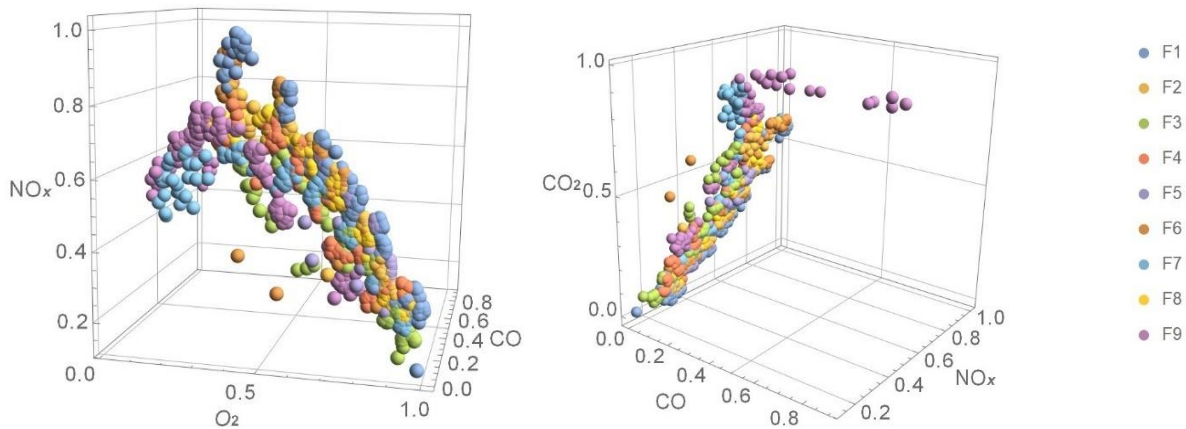


Figure 4. Random data points form fault spaces $F_1 - F_9$: the data cannot be easily clustered with one sphere per fault

The fault spaces are distinguished with different colors. As we can see, a single sphere for each fault could not easily cluster such data. Furthermore, in the case of faulty engines $F_2 - F_9$, the exhaust emission level of environmentally harmful gases such as CO, NO_x, and CO₂ is higher than

that for the efficient engine F_1 , this is better shown in **Table 7**. Based on the mean value and mean deviation for each fault-related exhaust gas dimension, we can see that the faulty engine $F_2 - F_9$ emits more environmentally harmful gases than the efficient engine.

Table 7. Data characteristics from each fault set: mean \bar{x} and mean absolute deviation (MAD)

Gas	O ₂ %		CO ppm		NO _x ppm		CO ₂ %	
	\bar{x}	MAD	\bar{x}	MAD	\bar{x}	MAD	\bar{x}	MAD
F_1	18.28	0.90	76.77	17.50	310.30	82.06	1.49	0.49
F_2	18.05	0.83	86.51	26.29	315.15	69.52	1.62	0.45
F_3	18.31	0.72	79.56	16.47	258.62	59.92	1.47	0.39
F_4	17.77	0.93	90.66	25.47	307.69	74.52	1.77	0.51
F_5	17.82	0.80	97.80	48.27	338.55	48.90	1.74	0.44
F_6	17.45	0.92	102.16	42.28	370.50	81.66	1.94	0.50
F_7	17.47	1.15	99.02	54.46	305.82	58.05	1.93	0.63
F_8	17.89	0.97	80.87	24.43	312.00	78.63	1.70	0.53
F_9	16.01	1.06	210.62	175.96	353.89	57.75	2.73	0.58

All previous conclusions lead us to investigate further the power load dimensions. The power load data are different from the exhaust emission ones. First of all, the power load is the input of the experiment, contrary to the exhaust gas emissions, which represent the output. Secondly, unlike the exhaust gas emission measurements, which yielded a stochastic distribution, the power load can be fixed in the experiment. **Figure 5** shows the selected three dimensions of the exhaust gas emission data for the efficient engine F_1 and fault F_9 . Here, the different subspaces of power load are colored. As we can see, the data clusters for different power load subspaces within the same fault. This is the main motivation for applying power load subspaces to the SVDD method generating hypersphere boundaries in our problem. It should be noted that adding the detection of

another fault in the engine involves determining the parameters of another hypersphere. Similarly, adding another symptom involves adding another dimension to the multi-dimensional space.

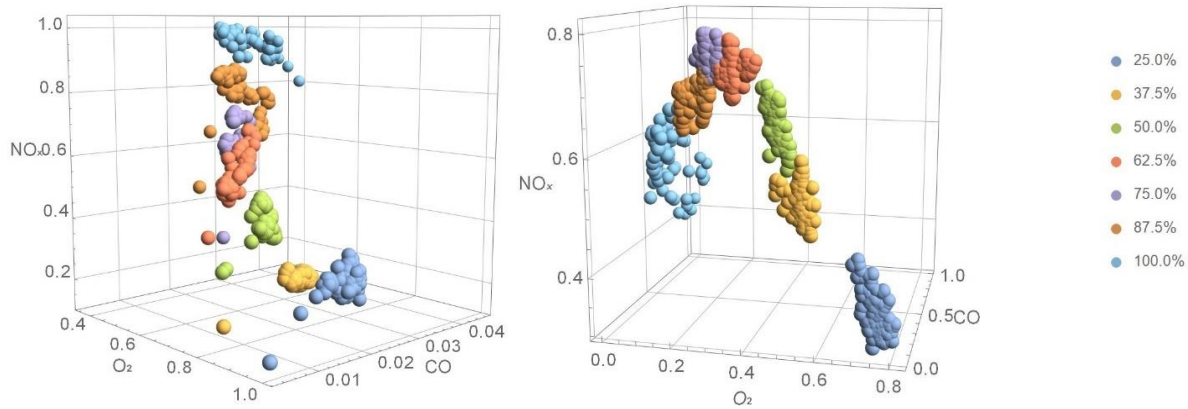


Figure 5. Random data points from power load subspaces 25% – 100% for fault spaces F_1 (right) and F_9 (left): data clusters for different power load subspaces

4.2. Engine fault detection SVDD-based method with power load subspaces

4.2.1. Training procedure

The training procedure will return the boundaries of the fault sets, described by the centers and radiuses of the hyperspheres. For each fault, there are 5D data points. One of these dimensions is reserved for the parameter set in the experiment, i.e., the engine power load of 25%, 37.5%, 50%, 62.5%, 75%, 87.5%, 100%. Some faults may be revealed differently under different power performances. This is why Authors subtract seven engine power load subspaces for each fault. The left four dimensions x represent the measured exhaust gases: O_2 , CO , NO_x , and CO_2 . The main challenge of the SVDD method in diagnostics is finding the fault boundary hyperspheres that gather the maximum number of data points at the minimum volume so that they are less likely to intersect with other faults (Deng and Zhang, 2020). For that reason, the four-dimensional data has been scaled to the $\langle 0,1 \rangle$ range with the minimum and maximum values shown in **Table 6**. Such data

preparation will result in better SVDD performance, i.e., less intersection of hyperspheres describing F_1 - F_9 .

Next, Authors randomly chose 200 training points from the testing set for each fault and power load subspace. With these data, Authors maximize the Lagrangian (9) and calculate the center coordinates and radius of the hyperspheres. The C variables are set to 0.25. The maximum number of outliers in the SVDD method is $(1 - C)/C$. This means that, for the chosen C variable and 200 data points, the boundaries will cover at least 98.5% of the data. For engine faults and subspaces where the engine failed to start, phantom hyperspheres are written with a zero center and relatively small radius compared with other results. This is done to unify the algorithm. The results of the SVDD training procedure, i.e., the center coordinates and radiuses of hyperspheres for each fault F_1 - F_9 and each power load subspace, are shown in **Table 8**. The selection of hypersphere radius is crucial for the accuracy of the proposed method. An increased radius means that the algorithm will identify a specific fault with relatively small changes in the values of diagnostic signals. In such a case, the fault will be indicated despite the correct operation of the system. Moreover, large hypersphere radius facilitates the possibility of overlapping hyperspheres in multi-dimensional space. According to the methodology, the value of the C coefficient is responsible for the radius of the hypersphere. In order to adapt the method to other complex technical objects, the value of this coefficient can be optimized for each failure. On the other hand, adopting smaller radius reduces the effectiveness of detecting faults, and only very large changes in the values of diagnostic parameters (in this case, the composition of exhaust gas) will indicate a fault. Therefore, it should be remembered that the effectiveness of the proposed diagnostic method depends not only on the selection of training parameters but also on the selection of appropriate diagnostic signals.

Table 8. The radius and coordinates of the center of hyperspheres for each fault and power subspace

Fault	Power	Centre of hypersphere coordinates a				Radius R
F1	25.0%	0.9777	0.0197	0.307	0.022	0.0539
	37.5%	0.8714	0.0075	0.2961	0.1288	0.1123
	50.0%	0.8409	0.0076	0.4447	0.1599	0.094
	62.5%	0.7512	0.012	0.6125	0.2479	0.096
	75.0%	0.6713	0.015	0.6368	0.3283	0.1119
	87.5%	0.5888	0.0197	0.7501	0.4115	0.1233
	100.0%	0.443	0.0372	0.8044	0.5572	0.1991
F2	25.0%	0.9283	0.0133	0.2626	0.0723	0.1139
	37.5%	0.8988	0.0123	0.3396	0.1012	0.0907
	50.0%	0.7728	0.0077	0.3924	0.2264	0.1675
	62.5%	0.728	0.0103	0.5819	0.2719	0.0756
	5.0%	0.6571	0.02	0.6642	0.343	0.0747
	87.5%	0.532	0.0289	0.6397	0.4669	0.1835
	100.0%	0.4651	0.0517	0.7197	0.5345	0.2175
F3	25.0%	0.9491	0.0155	0.2147	0.052	0.0427
	37.5%	0.898	0.0114	0.2839	0.102	0.0391
	50.0%	0.835	0.0125	0.3923	0.1657	0.0471
	62.5%	0.775	0.0106	0.4524	0.2258	0.0604
	75.0%	0.785	0.0131	0.4873	0.2152	0.2027
	87.5%	0.5654	0.0255	0.6015	0.4337	0.0873
	100.0%	0.5702	0.0309	0.5261	0.4304	0.2436
F4	25.0%	0.9282	0.0205	0.3176	0.0718	0.0792
	37.5%	0.8591	0.011	0.3401	0.1416	0.0583
	50.0%	0.7802	0.008	0.4435	0.2198	0.0901
	62.5%	0.6858	0.0112	0.588	0.3134	0.0642
	75.0%	0.611	0.0151	0.6407	0.3891	0.097
	87.5%	0.5083	0.0267	0.662	0.4917	0.1273
	100.0%	0.4826	0.0356	0.6231	0.5172	0.2497
F5	25.0%	0.	0.	0.	0.	0.001
	37.5%	0.8875	0.0137	0.3683	0.1134	0.0679
	50.0%	0.826	0.0078	0.451	0.1733	0.0821
	62.5%	0.7336	0.01	0.524	0.2665	0.0896
	75.0%	0.6644	0.0188	0.5533	0.3357	0.1444
	87.5%	0.5484	0.0427	0.676	0.452	0.09
	100.0%	0.4786	0.0653	0.5865	0.5216	0.2397
F6	25.0%	0.	0.	0.	0.	0.001
	37.5%	0.8763	0.011	0.3681	0.1225	0.1209
	50.0%	0.	0.	0.	0.	0.001
	62.5%	0.7455	0.0083	0.5299	0.2548	0.1227
	75.0%	0.6098	0.0158	0.5193	0.3899	0.2229
	87.5%	0.4872	0.0318	0.6391	0.5131	0.2522
	100.0%	0.4328	0.0498	0.8252	0.5661	0.1505
F7	25.0%	0.9332	0.0139	0.2697	0.0668	0.0378
	37.5%	0.8567	0.0105	0.3709	0.1439	0.0307
	50.0%	0.7665	0.0127	0.5011	0.2329	0.0406
	62.5%	0.6772	0.0108	0.59	0.3224	0.0475
	75.0%	0.5725	0.0214	0.6279	0.427	0.0523
	87.5%	0.4289	0.0332	0.6838	0.57	0.0774

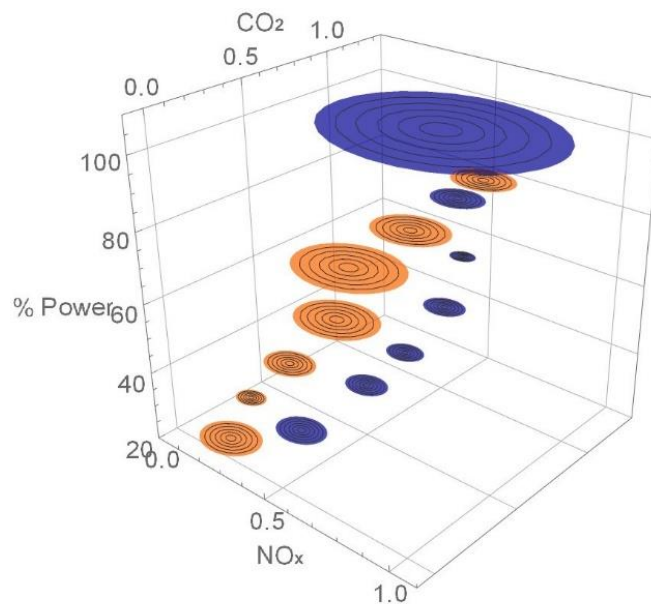
	100.0%	0.2187	0.1433	0.5489	0.7806	0.2235
F8	25.0%	0.9337	0.0141	0.3275	0.067	0.0755
	37.5%	0.8909	0.0106	0.3357	0.1095	0.0555
	50.0%	0.8035	0.0081	0.4753	0.1964	0.0444
	62.5%	0.7255	0.0092	0.5908	0.2745	0.0506
	75.0%	0.6334	0.0167	0.6829	0.3669	0.0577
	87.5%	0.5272	0.0332	0.7518	0.4724	0.0737
	100.0%	0.3694	0.0362	0.7296	0.63	0.0665
F9	25.0%	0.7477	0.0184	0.3966	0.2528	0.0955
	37.5%	0.539	0.021	0.5394	0.4605	0.0725
	50.0%	0.4914	0.0218	0.656	0.5088	0.0703
	62.5%	0.383	0.0287	0.7495	0.617	0.0707
	75.0%	0.3008	0.0403	0.7699	0.6985	0.0493
	87.5%	0.2432	0.0661	0.7052	0.7562	0.0929
	100.0%	0.1079	0.5001	0.5403	0.8909	0.4617

These calculations were performed on 12,000 four-dimensional data objects and took 52 seconds using Wolfram Mathematica 12. Below a general pseudo-code algorithm for the above training procedure is presented. It may be adapted to different technical solutions.

Algorithm 1: Training procedure	
Input:	training data from n fault spaces and m subspaces each $F_{ij}(\mathbf{x}_k)$, $i = 1, \dots, n, j = 1, \dots, m, k = 1, \dots, N$
Output:	hyperspheres $H_{ij}(\mathbf{a}_{ij}, R_{ij})$
1	take N input training data for each of m subspaces of all n faults F_{ij}
2	for each fault indexed with $i = 1, \dots, n$
3	for each fault subspace indexed with $j = 1, \dots, m$
4	for each training data indexed with $k = 1, \dots, N$
5	maximize Lagrangian (9)
6	return Lagrangian multipliers α_k
7	end for
8	from the Lagrangian multipliers calculate the center of hypersphere $\mathbf{a}_{ij} = \sum_k \alpha_k \mathbf{x}_k$
9	take any multiplier $0 < \alpha_k < C$ for the support vector $\tilde{\mathbf{x}}$
10	for this support vector calculate the radius $R_{ij} = \sqrt{(\tilde{\mathbf{x}} - \mathbf{a}_{ij})^T (\tilde{\mathbf{x}} - \mathbf{a}_{ij})}$
11	end for
12	end for
13	return hypersphere parameters $H_{ij}(\mathbf{a}_{ij}, R_{ij})$.

4.2.2. Cross-section pretesting procedure

The pretesting procedure will determine the main cutoff parameter and the testing sequence. In **Figure 6**, we show examples of the hypersphere' projections leaving two-dimensional spheres for all seven power load subspaces. Each graph presents two different faults. The spheres intersect differently for each power performance. In the left figure, we see faults that can be easily separated for lower power loads but get to intersect for the 100% load. In the right figure, for lower loads, the fault space F_2 contains F_7 , and both fault spaces get more separated for higher load performances. These two examples show that information from all power load subspaces for the final fault detection is needed.



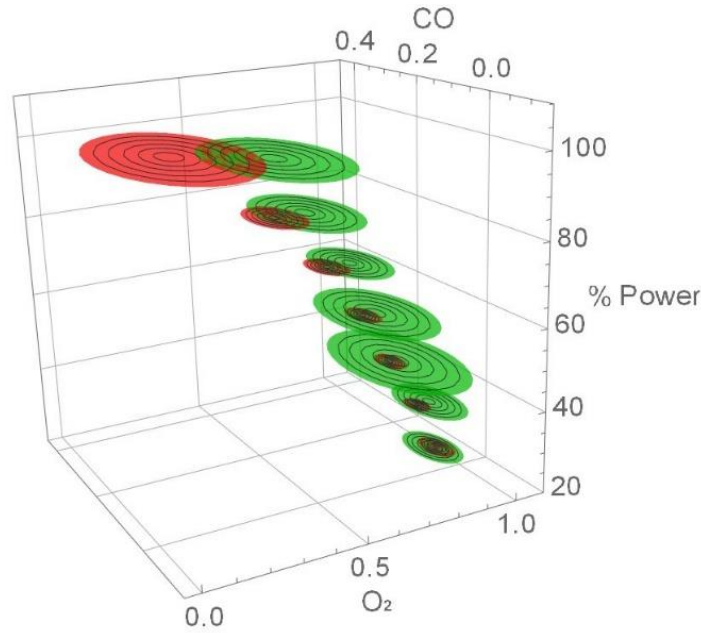


Figure 6. Projections to two dimensions are shown for seven power load subspaces. F1 - orange and F9 - blue, F2 - green and F7 - red: faults may reveal differently under different power performances

For the cross-section pretesting procedure, Authors randomly chose 50 data points from the pretesting set for each fault and power load subspace. These data points are scaled to the $(0,1)$ range. Next, cross-section tests are performed for new test data representing actual faults $F_1 - F_9$ with the parameters of the hyperspheres $H_1 - H_9$ calculated during the testing procedure. This Authors do separately for each power performance subspace. Using the formula (13), the output of this procedure is the percentage of points that are in the hyperspheres. Next, for each cross-section test, the mean value of the power load subspaces is calculated through the following formula:

$$P_{i,j} = \frac{\sum_{l=1}^m PR/N)_l}{m} \quad (14)$$

where PR denotes the number of positive results, N is the dimension of the test data set – in our case 50 and m is the dimension of subspaces in our case, seven power subspaces.

The result – a square confusion matrix P_{ij} – is shown in **Figure 7**. The true positive (TP)

diagnostic fractions are on the diagonal, whereas fractions of the false positives (FP) are outside the diagonal.

	F1	F2	F3	F4	F5	F6	F7	F8	F9
H1	0.98	0.9	0.74	0.89	0.7	0.59	0.4	0.97	0.
H2	0.64	0.98	0.79	0.96	0.83	0.47	0.61	0.98	0.
H3	0.04	0.03	0.97	0.08	0.05	0.01	0.11	0.1	0.
H4	0.17	0.5	0.29	0.97	0.2	0.13	0.69	0.86	0.
H5	0.31	0.54	0.66	0.49	0.97	0.32	0.2	0.55	0.
H6	0.57	0.7	0.57	0.7	0.83	0.98	0.48	0.71	0.
H7	0.	0.15	0.01	0.15	0.	0.09	0.98	0.14	0.07
H8	0.09	0.62	0.06	0.73	0.23	0.11	0.18	0.98	0.
H9	0.	0.	0.	0.	0.	0.	0.08	0.	0.98

Figure 7. Confusion matrix from cross-section pretests of hyperspheres $H_1 - H_9$ with novel data from actual faults $F_1 - F_9$: upper the diagonal, there are two high FP results; these can be eliminated by a proper testing sequence.

Usually, all the TP from the confusion matrix P_{ij} have to be higher than all the FP diagnostics. However, a weaker condition can also be applied; namely, all the positive diagnostics from the diagonal elements of the matrix P_{ij} have to be greater than the false-positive results below the diagonal, i.e.,

$$\forall_{ij} P_{i,i} > P_{i,j < i} \quad (15)$$

If this is not the case, we can try rearranging rows and columns to fulfill (15). The efficient engine F_1 should be tested as last; such a choice may signal the occurrence of novel and unknown

faults. Having fulfilled the weaker condition, we choose the main diagnostic parameter D to be

$$D = \frac{1}{2} (\min_i P_{i,i} + \max_{ij} P_{i,j < i}) \quad (16)$$

The final fault detection test can be then performed by eliminating faults by means of the proper testing sequence, i.e., from the last row's hypersurface to the first one. The proposed sequence testing based on the confusion triangular matrix is a trade-off for the robustness and simplicity of the algorithm. Below the pseudo-code algorithm for calculating the cross-section parameters is presented.

Algorithm 2: Cross-section pretesting procedure	
Input: test data $F_{il}(\mathbf{x}_k)$, $i = 1, \dots, n$, $l = 1, \dots, m$, $k = 1, \dots, N$, Hyperspheres $H_{jl}(\mathbf{a}_{jl}, R_{jl})$, $j = 1, \dots, n$, $l = 1, \dots, m$ from the Algorithm 1 output	
Output: P_{ij} cross-section matrix for positive results	
1	take N input test data for each of m subspaces for all n faults $F_{il}(\mathbf{x}_k)$,
2	for each fault indexed with $i = 1, \dots, n$
3	for each hyperspace indexed with $j = 1, \dots, n$
4	for each subspace indexed with $l = 1, \dots, m$
5	for each test data indexed with $k = 1, \dots, N$
6	test all data from $F_{il}(\mathbf{x}_k)$ with hyperspheres $H_{jl}(\mathbf{a}_{jl}, R_{jl})$ from formula (13)
7	count positive results PR
8	end for
9	calculate $(PR/N)_l$
10	end for
11	calculate the mean over all subspaces P_{ij} from formula (14)
12	end for
13	end for
14	return P_{ij}

4.2.3. Validation procedure

To validate the method, the final fault detection algorithm is performed. We choose the main diagnostic parameter $D = 90\%$. With the weaker condition (15), we can now perform the final testing by eliminating faults in the proper sequence. For our purpose, from the matrix presented in Table 7, the order of testing is $\{H_9, H_8, H_7, H_6, H_5, H_4, H_3, H_2, H_1\}$. For the algorithm's validation, Authors randomly pick new data points from the final testing data set for each nine fault spaces and each seven power load subspaces.

In our case, we first test novel data with hypersphere H_9 using the formula (13). If more than 90% of the data points fall to H_9 hyperspace, fault F_9 is detected, and the test is stopped. If less than 90% of the points fall into this category, fault F_9 is eliminated, and the tests are continued for the remaining faults only. To describe the function of sequence testing, let us take a closer look at the confusion matrix where the hypersphere H_2 could have falsely detected the actual fault F_8 as F_2 , were it not for the sequence testing. With the sequence testing, if the actual fault is F_8 , the test performed first with the hypersphere H_8 will return the result and end the procedure. **Figure 8** presents the final results of testing the novel data chosen randomly from each fault space.

	F1	F2	F3	F4	F5	F6	F7	F8	F9
H1	10	0	0	0	0	0	0	0	0
H2	0	10	0	0	0	0	0	0	0
H3	0	0	10	0	0	0	0	0	0
H4	0	0	0	10	0	0	0	0	0
H5	0	0	0	0	10	0	0	0	0
H6	0	0	0	0	0	10	0	0	0
H7	0	0	0	0	0	0	10	0	0
H8	0	0	0	0	0	0	0	10	0
H9	0	0	0	0	0	0	0	0	10

Figure 8. Final confusion matrix for predictions $H_1 - H_9$ with novel data from actual faults $F_1 - F_9$. Tests were performed with 10 data sets for each actual fault: to get 100% accuracy, five data points were taken for each power load subspace.

The test was performed ten times for each fault. In all cases, the actual faults were detected with 100% accuracy. The final diagnostics show that five exhaust composition measurements per seven power load subspaces are enough for 100% accuracy of this method. In practical use, for testing an actual engine fault, this would mean that the operator should take a minimum of five measures for each power load subspace. Below a generalized pseudo-code algorithm for the final testing is presented. It may be applied to different technical solutions.

Algorithm 3: Final testing procedure	
Input: test data for m subspaces $x_{kj}, k = 1, \dots, N, j = 1, \dots, m$	
hyperspheres $H_{ij}(\mathbf{a}_{ij}, R_{ij}), i = 1, \dots, n, j = 1, \dots, m$	
main diagnostic parameter - D	
order of fault testing $\{H_i\}$	
Output: the fault	
1	for the order $\{H_i\}$
2	for each subspaces indexed $j = 1, \dots, m$
3	for each test data indexed $k = 1, \dots, N$
4	test the data x_{kj} with hyperspheres $H_{ij}(\mathbf{a}_{ij}, R_{ij})$ with formula (13)
5	count positive results PR
6	end for
7	calculate the fraction of positive results $(PR/N)_j$ for each subspace

```

8      end for
9      calculate the mean over all subspaces using the formula  $M = \frac{\sum_j^m (PR/N)_j}{m}$ 
10     if  $M > D$  returns fault  $H_i$  and end testing
11     else exclude fault  $H_i$  and continue with the testing order for the remaining faults.
12     end for

```

Finally, for more clarification, a flowchart of all the introduced procedures is given in Figure 9.

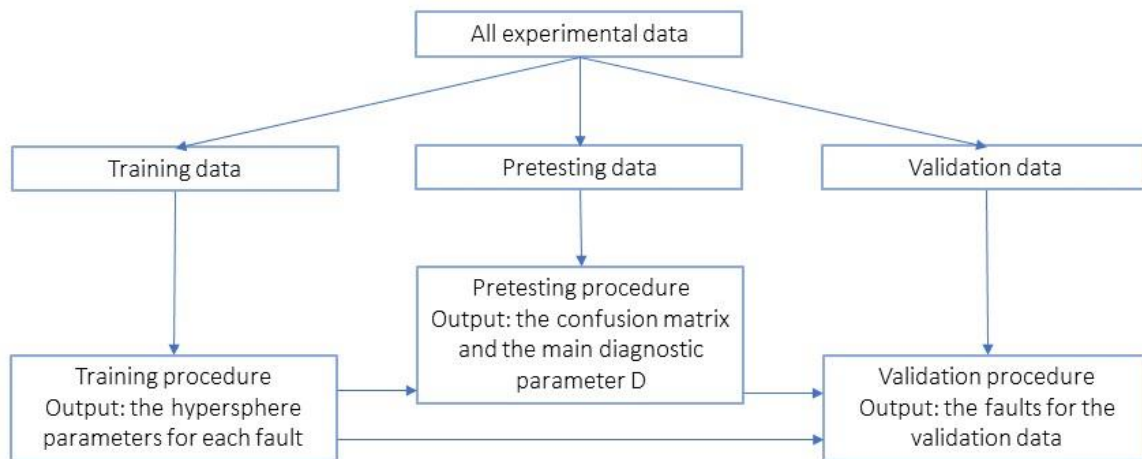


Figure 9. A flowchart of all the procedures

5. Conclusions

Our team has adeptly developed and rigorously tested an innovative approach for diagnosing faults in marine combustion engines. This approach employs a Support Vector Data Description (SVDD) methodology, integrated with power load subspaces. Our findings reveal that signals from exhaust gas composition are highly informative for diagnostic purposes. This newly proposed method demonstrates a high efficacy, accurately identifying eight different simulated faults as well

as the state of engine efficiency, achieving a perfect accuracy rate of 100%. This was accomplished using merely five data points from each power load subspace. The implementation of this method is straightforward, eschewing the need for complex parameter optimization and is not burdensome in terms of time. Consequently, we infer that this methodology holds significant promise for facile application across various engine types.

References

- Aliramezani, M., Koch, C.R., Shahbakhti, M., 2022. Modeling, diagnostics, optimization, and control of internal combustion engines via modern machine learning techniques: A review and future directions. *Prog. Energy Combust. Sci.* 88, 100967. <https://doi.org/10.1016/j.pecs.2021.100967>
- Ayyagari, A., Hasannaemi, V., Grewal, H.S., Arora, H., Mukherjee, S., 2018. Corrosion, erosion and wear behavior of complex concentrated alloys: A review. *Metals (Basel)*. 8, 603.
- Banerjee, A., Burlina, P., Meth, R., 2007. Fast Hyperspectral Anomaly Detection via SVDD, in: 2007 IEEE International Conference on Image Processing. IEEE, pp. IV-101-IV-104. <https://doi.org/10.1109/ICIP.2007.4379964>
- Basurko, O.C., Uriondo, Z., 2015. Condition-Based Maintenance for medium speed diesel engines used in vessels in operation. *Appl. Therm. Eng.* 80, 404–412. <https://doi.org/10.1016/j.applthermaleng.2015.01.075>
- Borkowski, T., Kowalak, P., Wiśnicki, B., Przywarty, M., 2019. Slow steaming of container carriers: A study. Maritime University of Szczecin Press.
- Cai, B., Sun, X., Wang, J., Yang, C., Wang, Z., Kong, X., Liu, Z., Liu, Y., 2020. Fault detection and diagnostic method of diesel engine by combining rule-based algorithm and BNs/BPNNs. *J. Manuf. Syst.* 57, 148–157. <https://doi.org/10.1016/j.jmsy.2020.09.001>

- Cao, W., Dong, G., Chen, W., Wu, J., Xie, Y.-B., 2015. Multisensor information integration for online wear condition monitoring of diesel engines. *Tribol. Int.* 82, 68–77.
<https://doi.org/10.1016/j.triboint.2014.09.020>
- Chapelle, O., Haffner, P., Vapnik, V.N., 1999. Support vector machines for histogram-based image classification. *IEEE Trans. Neural Networks* 10, 1055–1064.
<https://doi.org/10.1109/72.788646>
- Deng, X., Zhang, Z., 2020. Nonlinear chemical process fault diagnosis using ensemble deep support vector data description. *Sensors* 20, 4599.
- Hoang, A.T., 2020. Critical review on the characteristics of performance, combustion and emissions of PCCI engine controlled by early injection strategy based on narrow-angle direct injection (NADI). *Energy Sources, Part A Recover. Util. Environ. Eff.* 1–15.
<https://doi.org/10.1080/15567036.2020.1805048>
- Hoang, A.T., 2019. Experimental study on spray and emission characteristics of a diesel engine fueled with preheated bio-oils and diesel fuel. *Energy* 171, 795–808.
<https://doi.org/10.1016/j.energy.2019.01.076>
- Hoang, A.T., Foley, A.M., Nižetić, S., Huang, Z., Ong, H.C., Ölçer, A.I., Pham, V.V., Nguyen, X.P., 2022. Energy-related approach for reduction of CO₂ emissions: A critical strategy on the port-to-ship pathway. *J. Clean. Prod.* 355, 131772.
<https://doi.org/10.1016/j.jclepro.2022.131772>
- Hoang, A.T., Pandey, A., Martinez De Osés, F.J., Chen, W.-H., Said, Z., Ng, K.H., Ağbulut, Ü., Tarełko, W., Ölçer, A.I., Nguyen, X.P., 2023. Technological solutions for boosting hydrogen role in decarbonization strategies and net-zero goals of world shipping: Challenges and perspectives. *Renew. Sustain. Energy Rev.* 188, 113790.
<https://doi.org/10.1016/j.rser.2023.113790>



- Hoang, A.T., Tabatabaei, M., Aghbashlo, M., 2020. A review of the effect of biodiesel on the corrosion behavior of metals/alloys in diesel engines. *Energy Sources, Part A Recover. Util. Environ. Eff.* 42, 2923–2943. <https://doi.org/10.1080/15567036.2019.1623346>
- Kalghatgi, G., 2013. Fuel/engine interactions. SAE international.
- Korczewski, Z., 2015. Exhaust gas temperature measurements in diagnostics of turbocharged marine internal combustion engines part I standard measurements. *Polish Marit. Res.* 22, 47–54.
- Kowalski, J., 2015. Concept of the multidimensional diagnostic tool based on exhaust gas composition for marine engines. *Appl. Energy* 150, 1–8. <https://doi.org/10.1016/j.apenergy.2015.04.013>
- Kowalski, J., Krawczyk, B., Woźniak, M., 2017. Fault diagnosis of marine 4-stroke diesel engines using a one-vs-one extreme learning ensemble. *Eng. Appl. Artif. Intell.* 57, 134–141. <https://doi.org/10.1016/j.engappai.2016.10.015>
- Lamaris, V.T., Hountalas, D.T., 2010. A general purpose diagnostic technique for marine diesel engines – Application on the main propulsion and auxiliary diesel units of a marine vessel. *Energy Convers. Manag.* 51, 740–753. <https://doi.org/10.1016/j.enconman.2009.10.031>
- Lazakis, I., Raptodimos, Y., Varelas, T., 2018. Predicting ship machinery system condition through analytical reliability tools and artificial neural networks. *Ocean Eng.* 152, 404–415. <https://doi.org/10.1016/j.oceaneng.2017.11.017>
- Liang, J., Zhang, K., Al-Durra, A., Zhou, D., 2020. A novel fault diagnostic method in power converters for wind power generation system. *Appl. Energy* 266, 114851. <https://doi.org/10.1016/j.apenergy.2020.114851>
- Liang, M., Chen, M., 2022. Monitoring the Performance of a Ship's Main Engine Based on Big Data Technology. *Polish Marit. Res.* 29, 128–140.

- Luo, H., Jiang Cui, Y.W., 2011. A SVDD approach of fuzzy classification for analog circuit fault diagnosis with FWT as preprocessor. *Expert Syst. Appl.* 38, 10554–10561. <https://doi.org/10.1016/j.eswa.2011.02.087>
- Namigtle-Jiménez, A., Escobar-Jiménez, R.F., Gómez-Aguilar, J.F., García-Beltrán, C.D., Téllez-Anguiano, A.C., 2020. Online ANN-based fault diagnosis implementation using an FPGA: Application in the EFI system of a vehicle. *ISA Trans.* 100, 358–372. <https://doi.org/10.1016/j.isatra.2019.11.003>
- Noyel, M., Thomas, P., Thomas, A., Charpentier, P., 2016. Reconfiguration process for neuronal classification models: Application to a quality monitoring problem. *Comput. Ind.* 83, 78–91.
- Puzdrowska, P., 2022. Diagnostic Information Analysis of Quickly Changing Temperature of Exhaust Gas from Marine Diesel Engine Part I Single Factor Analysis. *Polish Marit. Res.* 28, 97–106. <https://doi.org/10.2478/pomr-2021-0052>
- Rudzki, K., Gomulka, P., Hoang, A.T., 2022. Optimization Model to Manage Ship Fuel Consumption and Navigation Time. *Polish Marit. Res.* 29, 141–153. <https://doi.org/10.2478/pomr-2022-0034>
- Scappin, F., Stefansson, S.H., Haglind, F., Andreasen, A., Larsen, U., 2012. Validation of a zero-dimensional model for prediction of NO_x and engine performance for electronically controlled marine two-stroke diesel engines. *Appl. Therm. Eng.* 37, 344–352. <https://doi.org/10.1016/j.applthermaleng.2011.11.047>
- Tamura, M., Saito, H., Murata, Y., Kokubu, K., Morimoto, S., 2011. Misfire detection on internal combustion engines using exhaust gas temperature with low sampling rate. *Appl. Therm. Eng.* 31, 4125–4131. <https://doi.org/10.1016/j.applthermaleng.2011.08.026>
- Tan, Y., Tian, H., Jiang, R., Lin, Y., Zhang, J., 2020. A comparative investigation of data-driven approaches based on one-class classifiers for condition monitoring of marine machinery

- system. *Ocean Eng.* 201, 107174. <https://doi.org/10.1016/j.oceaneng.2020.107174>
- Tang, Z., Liu, X., Wei, D., Luo, H., Jiang, P., Bo, L., 2022. Enhanced multiclass support vector data description model for fault diagnosis of gears. *Measurement* 194, 110974. <https://doi.org/10.1016/j.measurement.2022.110974>
- Tax, D.M., Duin, R.P., 1999. Support vector domain description. *Pattern Recognit. Lett.* 20, 1191–1199. [https://doi.org/10.1016/S0167-8655\(99\)00087-2](https://doi.org/10.1016/S0167-8655(99)00087-2)
- Tax, D.M.J., Duin, R.P.W., 2004. Support Vector Data Description. *Mach. Learn.* 54, 45–66. <https://doi.org/10.1023/B:MACH.0000008084.60811.49>
- Wang, X., Liu, C., Bi, F., Bi, X., Shao, K., 2013. Fault diagnosis of diesel engine based on adaptive wavelet packets and EEMD-fractal dimension. *Mech. Syst. Signal Process.* 41, 581–597. <https://doi.org/10.1016/j.ymssp.2013.07.009>
- Zhang, J., Zhang, Q., Qin, X., Sun, Y., 2022. A two-stage fault diagnosis methodology for rotating machinery combining optimized support vector data description and optimized support vector machine. *Measurement* 200, 111651. <https://doi.org/10.1016/j.measurement.2022.111651>
- Zhao, Y., Wang, S., Xiao, F., 2013. Pattern recognition-based chillers fault detection method using Support Vector Data Description (SVDD). *Appl. Energy* 112, 1041–1048. <https://doi.org/10.1016/j.apenergy.2012.12.043>