

Investigating an SVM-driven, one-class approach to estimating ship systems condition

Iraklis Lazakis^a, Christos Gkerekos^a, Gerasimos Theotokatos^a

^aDepartment of Naval Architecture, Ocean and Marine Engineering, University of Strathclyde, 100 Montrose Street, Glasgow, UK.

ARTICLE HISTORY

Compiled June 26, 2018

ABSTRACT

Maintenance is a major point that can affect vessel operation sustainability and profitability. Recent literature shows that condition monitoring of ship systems has a great potential at the cost of significant data requirements. In this respect, this paper presents a novel methodology for intelligent, system-level modelling for the monitoring of main engine performance utilising data acquired through noon-reports, with a minimal amount of data assumptions. The proposed methodology is utilised to train a one-class Support Vector Machine (SVM) that represents a diesel generator's normal behaviour. Acquired data representing the engine's new conditions are input to the model and model output reflects a gauge of their normality, when compared to the original dataset, for analogous operating conditions. This aids the dynamic detection of ship machinery incipient faults, contributing to the minimisation of ship downtime and increase of its operability. A case study presenting applications of this modelling technique on actual ship machinery raw data is included, complemented by a sensitivity analysis. The case study demonstrates the applicability of the developed methodology in the identification of deviant, abnormal ship machinery conditions.

KEYWORDS

ship machinery; condition monitoring; SVM; machine learning; Diesel-generator set; behavioural prediction

1. Introduction & Research Background

Ships are a significant asset of the global goods transportation system as over four-fifths of merchandise are carried by sea (UNCTAD 2015). Unplanned maintenance constitutes almost one-fifth of the overall operating costs of a vessel and overall maintenance combined with stores and lubricants contributes approximately one-tenth of the overall running cost of a vessel (Stopford 2009). Additionally, maintenance costs of a vessel increase with its age and Moore Stephens (2017) report an expected increase by approximately 2% year-by-year for the next two years. Combining the above facts with an average global merchant fleet vessel age of almost twenty years (UNCTAD 2015), a high level of maintenance optimisation is required for ships to remain sustainable and profitable.

Meanwhile, current maintenance state-of-practice in shipping offers ample room for

improvement, as time-based (preventive) maintenance implementations are currently preferred (Lazakis and Ölçer 2015). Thus, there exists a justifiable need for novel methods of monitoring the condition of machinery equipment, combined with the suggestion and optimised scheduling of suitable maintenance actions.

1.1. *Maintenance strategies*

Maintenance strategies can be split into four broad categories: reactive, preventive, predictive and proactive. Reactive maintenance concerns maintenance where components get replaced only following a complete failure (Mohanty 2014). This method of maintenance provides the longest time between shutdowns but failures are catastrophic and can possibly affect multiple components and/or machines (Randall 2011). At the same time, a large spares inventory is required, there is high downtime and a large cost attached to any failure. Therefore, reactive maintenance is mainly applied to relatively not expensive and non-critical machines or where redundancies have been implemented so that production is not interrupted.

Preventive maintenance refers to maintenance that takes place at a fixed frequency (also known as Planned Maintenance System (PMS)), usually following Original Equipment Manufacturers (OEMs), International Association of Classification Societies (IACS), or International Safety Management (ISM) code recommendations. Compared to reactive maintenance, preventive maintenance offers significant increase in machine lifespan as the probability of catastrophic failures diminishes (Randall 2011). Additionally, preventive maintenance is, especially in the case of critical equipment, more cost-effective as the number of components or machines requiring complete replacement is reduced. Additionally, the overall cost attached to retaining spare parts is reduced.

Predictive maintenance utilises an estimation of the current and future condition of a component in order to provide optimised maintenance scheduling that prevents failures without however resorting to over-maintenance (Verbert et al. 2017). This optimised maintenance scheduling offers extended machine lifespan, coupled with reduced maintenance costs and downtime. There, present and past condition of each component is taken into consideration to offer bespoke maintenance scheduling for each component and each machine. Predictive maintenance requires a higher Capital Expenditure (CAPEX) but over an extended period, becomes more economical than preventive or reactive maintenance, due to reduced Operating Expenses (OPEX). Especially in industries where machines are expected to run for long periods without any shutdowns, it has been shown that predictive maintenance can reduce relevant costs by up to 65% (Neale and Associates 1979). Furthermore, in terms of downtime, planned downtime is minimised to the bare necessary minimum and unplanned is almost diminished. This optimised maintenance scheduling permits the maximisation of machine lifespan. Still, while predictive maintenance proves to be more economical during a machines lifespan, results may not be immediately evident.

Proactive maintenance enhances the insights obtained from predictive maintenance by both offering optimised maintenance scheduling and identifying the underlying causes of a fault following its detection in order to offer a more holistic maintenance approach (Manzini et al. 2009). This reduces maintenance constant downtime as faults are being avoided instead of just being predicted and mitigated. At the same time, proactive maintenance dictates a high preliminary cost as for each detected fault, its root cases have to be identified and analysed. Accordingly skilled personnel are

required to perform the above tasks.

1.2. *Maintenance in the maritime sector*

In sectors such as defence, aviation, manufacturing, automobile, and nuclear power generation, maintenance focus has shifted from reactive and preventive towards predictive. However, in the maritime sector, ship maintenance has been considered an area of needless expenditure and advanced monitoring methods have not yet been widely applied (Lazakis et al. 2010; Lazakis and Ölçer 2015). Nevertheless, attempts towards predictive maintenance in shipping have been made in the past years and are rapidly progressing. For example, a methodology where vibration data are combined with performance data (cylinder pressures) for the condition monitoring of a main engine has been suggested (Chandroth 2004). Accordingly, thermodynamic model of main and auxiliary engines have been developed and used to perform condition monitoring (Watzenig et al. 2009; Lamarinis and Hountalas 2010). Accordingly, Hountalas (2000) developed a diesel engine performance model that can account for both normal and faulty conditions. Besides, a self-learning algorithm for fault diagnosis in the combustion system of a marine diesel engine has been developed (Li et al. 2010). Dikis and Lazakis (2016) present the framework of Inspection Capabilities for Enhanced Ship Safety (INCASS) project that developed tools to enhance machinery monitoring by combining real time information with machinery risk analysis tools. In this scope, a machinery risk analysis tool that performs condition monitoring and maintenance decision support was developed. Dikis et al. (2017) further elaborated on this tool by considering and assessing components' risk of failure and reliability degradation by utilising raw input data. Furthermore, Gkerekos et al. (2016, 2017b) developed a self-learning model for the condition monitoring of ship machinery based on vibration measurements and raw physical data. Along these lines, Gkerekos et al. (2017a) further extended this work by implementing a more thorough optimisation step in order to derive a more robust model. Coraddu et al. (2016) and Cipollini et al. (2018b) suggest a regression method for the estimation of component degradation in a marine COmbined Diesel eLectric And Gas (CODLAG) propulsion plant type in the existence of degradation information in the model training dataset. In these studies simulated data were used, therefore decay state information was also explicitly available. Cipollini et al. (2018a) then utilised the same dataset to compare supervised and unsupervised algorithms for fault detection. Raptodimos and Lazakis (2018) investigated the application of self-organising maps, an unsupervised neural network, for monitoring the condition of a two stroke marine diesel engine by identifying clusters containing data representing abnormal engine operating conditions.

1.3. *Performance monitoring of machinery*

Performance monitoring of machinery is a discipline that requires the development of a suitable model. This model can either use a first-principles analysis (i.e. white-box model), or use a more "brute-force" approach by developing a model employing self-learning algorithms coupled with an acquired dataset (i.e. black-box model). Often, a combination of both techniques is applied, leading to grey-box models. Yang et al. (2013) developed a framework for the analysis of data acquired through wind turbine Supervisory Control and Data Acquisition (SCADA) bus measurements to perform condition monitoring based on correlations between measurements. Kowalski et al.

(2017) proposed a data-driven methodology that learns from a dataset containing 14 faulty conditions additional to normal behaviour and is then able to discern between them in new data points. Widodo and Yang (2007) presented an overview of Support Vector Machine (SVM) techniques for fault diagnosis and monitoring in engineering applications. Yin and Hou (2016) built on that review, providing an updated overview of SVM applications for fault diagnosis and process monitoring.

1.4. *Aim of this paper*

While current maintenance practices incorporate proactive maintenance aspects in many other industries, in shipping proactive maintenance remains way beyond current practice (Dikis et al. 2017; Raptodimos and Lazakis 2018). Before the mass deployment of proactive maintenance strategies in a turnkey fashion, a robust predictive framework is required to be developed. In this respect, and contrary to most other condition monitoring approaches examined, where either specific dataset requirements have to be satisfied or an first-principles engine model is used, in the proposed methodology, model training will only assume the existence of a dataset containing normal data points. Additionally, the data requirements of the methodology proposed within this study can be satisfied using ubiquitous, noon-report data, without requiring the additional CAPEX cost of a specialised Data Acquisition (DAQ) system – something that has so far affected the implementation rate of advanced maintenance strategies. This majorly simplifies the data gathering that is required prior to model training and increases model deployment efficacy.

To this end, a multi-component, multi-criteria, SVM-based one-class condition monitoring framework is proposed that can be used as a basis for a robust predictive framework. Multi-component and multi-criteria refers to the fact that several performance measurements are combined as input and analysed suitably so that an estimation of system condition is returned. One-class refers to the fact that only measurements referring to the “normal” condition are required for the suggested methodology to discern between “normal” and abnormal” states. This aspect significantly simplifies model training as no “abnormal” recordings are required for model training.

Based on the above, this paper aims to present an SVM-driven methodology for the condition monitoring of ship machinery that will offer a simplified model training process, using only easily obtainable data.

Section 1 introduces this study’s scope and motivation of research along with an overview of the research background. Section 2 elaborates on the proposed methodology concerning dimensionality reduction, data processing, and one-class model training. Sections 3 and 4 detail the setup of the case study developed to validate the proposed methodology focusing on vessel’s Diesel Generator set (D/Gen) components along with its results and relevant discussion. Finally, in section 5, overall conclusions are provided.

2. Proposed methodology

The methodology elaborated in this paper includes a) the description of a suitable pre-processing technique for the acquired dataset, and b) the development of a self-learning model that can estimate whether a given data point corresponds to a reference (nominal) condition considered during model training. As such, a self-learning model can be trained without the need of obtaining data corresponding to “faulty” conditions.

A flowchart of the proposed methodology is presented in Figure 1.

Compared with other relevant methodologies available in pertinent literature, the novel characteristic of this framework concerns is modelling of a condition-estimation system using only “normal” data and the transformation of its output to a time-dependent (dynamic) normality metric without the explicit need for any additional information. Additionally, the developed methodology can be applied for the establishment of relevant measurement thresholds when such values are not readily available.

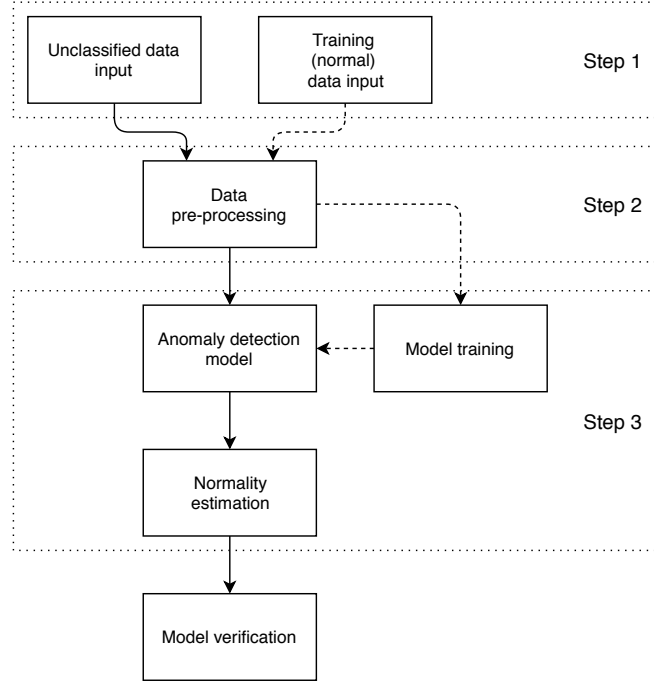


Figure 1. Visual representation of proposed methodology.

2.1. *Data pre-processing*

The data required for model training can either be acquired through an on-line DAQ system or through the processing of noon reports depending on relevant availability. While the data collection characteristics do not affect the proposed methodology, either process may include inconsistent and/or faulty data (e.g. due to recording inconsistencies, human error, or sensor faults) entries which need to be discarded before model training. Therefore, while loading the dataset, these elements are detected and eventually replaced by “Not a Number” (NaN), a notation used to denote non-numerical values in cases where a numerical value is expected. The pre-processing methodology consists of the following steps:

- (1) Observations containing NaN values are filtered.
- (2) Measured performance parameters are corrected to ISO ambient conditions (MAN B&W Diesel A/S 2004).
- (3) Feature vectors x_T are scanned for elements with values beyond $\mu \pm 3\sigma$ and corresponding observations are dropped from all feature vectors. μ corresponds to the mean value of each vector and σ to its standard deviation. Assuming a normal distribution, 99.7% of normal data should be within $\mu \pm 3\sigma$. Therefore,

this formulation filters out most abnormal data points from the training dataset, without affecting the vast majority of normal points.

- (4) Remaining dataset is split into training (70%), and testing (30%) sub-datasets. Validation is performed using K -folding (Kohavi 1995) on the training set, so no additional splitting is required.
- (5) Elements μ and σ of each feature x_T contained in the training sub-dataset are extracted and used to produce the normalised sub-dataset x'_T , derived by using Equation (1).
- (6) Testing sub-dataset is also normalised using μ and σ values extracted from the training sub-dataset.

$$x'_T = \frac{x_T - \mu}{\sigma} \quad (1)$$

Training dataset through K -folds cross-validation is used both for model training (i.e. deriving optimal model parameters) and hyperparameter optimisation (i.e. deriving optimal hyperparameter values). Model parameters are learnt during model training and depend on the training dataset peculiarities. Contrarily, model hyperparameters cannot be learnt during training but instead are selected by the user in order to optimise a selected metric (e.g. accuracy). Hence, while model parameters vary for distinct training data sets of the same application, hyperparameters tend to be common for similar models.

K -fold cross validation works by splitting the training dataset X into K roughly equal parts X_1, X_2, \dots, X_K and using $K - 1$ of them for training and 1 for validation, going through a for-loop to ensure that all K possible combinations are evaluated. Through K -fold cross validation, the generalisation capabilities of a model can be established as the model is trained in K different scenarios and the average performance is evaluated. Additionally, through the introduction of a second for-loop, different hyperparameter tuples can be evaluated in order to identify which one optimises model performance. This is presented in pseudo-code in Algorithm 1.

Algorithm 1 Model training and hyperparameter optimisation using K -folding

Require: X, X_{abn}, K , a set of n hyperparameter tuples $h_i, i \in (1, 2, \dots, n)$

- 1: Divide data X into K roughly equal parts
 - 2: **for** $i = 1, 2, \dots, n$ **do**
 - 3: $Hyps \leftarrow h_i$
 - 4: **for** $k=1, 2, \dots, K$ **do**
 - 5: $TrainSet \leftarrow X \setminus X_k$ (set subtraction)
 - 6: $ValidSet \leftarrow X_k \cup X_{abn}$
 - 7: Train model $M_{(k,i)}$ using $TrainSet$ and $Hyps$
 - 8: Evaluate model $M_{(k,i)}$ quality $MCC_{(k,i)}$ on $ValidSet$
 - 9: **end for**
 - 10: Calculate mean model quality $MCC_{(\mu,i)}$ (Eq. (9)) by averaging $MCC_{(k,i)} \forall k$
 - 11: **end for**
 - 12: Obtain $h^* = \underset{h_i}{\operatorname{argmax}}(MCC_{\mu,i})$ and use it to train final model M^* on whole X
 - 13: **return** M^*
-

Finally, testing sub-dataset is used to evaluate the generalisation capabilities of the final selected model (Hastie et al. 2001) using a previously unseen dataset, in order to

yield a more accurate, real-world usage scenario.

2.2. Anomaly detection model

A one-class SVM classifier (Cortes and Vapnik 1995) was trained, one the most commonly chosen algorithms for One Class Classification (OCC) (Khan and Madden 2010). Compared to other machine learning tools, SVMs offer superior generalisation capabilities (Widodo and Yang 2007). This is a semi-supervised algorithm that learns a decision boundary using only “normal” data points and then testing the likelihood of a test instance being within the boundary of the learnt model. Schölkopf et al. (1999, 2000, 2001) present a method of creating an OCC by implementing a suitable separating hyperplane of the form:

$$w^T x + b = 0 \quad (2)$$

where w is the weight vector, always normal to the hyperplane, x is the input vector and b is the bias term. This hyperplane aims to separate the surface region containing data from the region containing no data. This is achieved by “constructing a hyperplane which is maximally distant from origin, with all data points lying on the opposite side from the origin and such that the margin is positive” (Khan and Madden 2010).

In their ν -SVM implementation (Schölkopf et al. 2000), SVMs utilise the parameter ν as a degree of freedom in the trade-off between a large margin and a small training error. Accordingly, the parameter ν is “an upper bound on the fraction of training margin errors and lower bound on the fraction of support vectors” (Wu and Srihari 2003). In more practical terms, $\nu \in (0, 1]$ represents both an upper limit on the number of miss-classifications in the training dataset (at the cost of a possibly smaller margin) and a lower limit in the number of training samples used as support vectors (Chang and Lin 2001).

Following the formulation suggested by Schölkopf et al. (1999), the objective function of a SVM OCC can be described as:

$$\min_{w, \xi_i, \rho} \left[\frac{1}{2} \|w\|^2 + \frac{1}{\nu \ell} \sum_{i=1}^{\ell} \xi_i - \rho \right] \quad (3)$$

subject to: $(w \cdot \Phi(x_i)) \geq \rho - \xi_i, \xi_i \geq 0$

where ℓ is the number of observations, ρ represents the offset, and ξ_i are slack variables introduced to allow some points to lie within the margin in order to avoid overfitting.

Solving this minimisation problem using Lagrange multipliers, the decision function rule for a data point x then becomes:

$$f(x) = \text{sgn} \left(\sum_i \alpha_i \kappa(\mathbf{x}_i, \mathbf{x}) - \rho \right) \quad (4)$$

Coefficients a_i can be found as solutions to the dual problem:

$$\begin{aligned} \min_{\alpha} \quad & \frac{1}{2} \sum_{i,j} \alpha_i \alpha_j \kappa(\mathbf{x}_i, \mathbf{x}_j) \\ \text{subject to:} \quad & 0 \leq \alpha_i \leq \frac{1}{\nu \ell}, \quad \sum_i \alpha_i = 1 \end{aligned} \tag{5}$$

where κ refers to the kernel function that will be elaborated in subsection 2.2.1.

In order to calculate the offset ρ , we can exploit the fact that for any α_i that is not at the upper or lower bound, the corresponding pattern \mathbf{x}_i satisfies:

$$\rho = (w \cdot \Phi(x_i)) = \sum_j \alpha_j \kappa(\mathbf{x}_i, \mathbf{x}_j) \tag{6}$$

2.2.1. Kernel function

A key selection point in a SVM implementation concerns the kernel function κ , that for two vectors \mathbf{u}, \mathbf{v} takes the form:

$$\kappa(\mathbf{u}, \mathbf{v}) = \Phi(\mathbf{u})^T \Phi(\mathbf{v}) \tag{7}$$

Kernels operate as a similarity function, offering a gauge of similarity between two inputs and, especially in the case of SVMs, as a transformation that helps linearly separate linearly inseparable data. In that case, the kernel function offers a map of the originally inseparable data to a higher-dimensional space where they can be linearly separable. While multiple kernel functions can be applied for model training, Radial Basis Function (RBF) is, in practice, considered to work well as a SVM kernel and is usually a reasonable first choice (Hsu et al. 2010). As indicated in Equation (8), a significant advantage of RBF compared to other kernel functions is that it can be easily calibrated as it only depends on one parameter, $\gamma \in (0, \infty)$. A reasonable range and search spacing for γ when used as a SVM kernel, is $\gamma = 2^{-15}, 2^{-13}, \dots, 2^3$ (Hsu et al. 2010). Parameter γ controls the region that a single training sample can affect, with a small value of γ increasing the size of this region; and conversely. In other words, a large value of γ leads towards over-fitting in the Bias-Variance trade-off (Geman et al. 1992) whereas selecting a small value risks creating a model too constrained to efficiently capture the complexity of the training dataset. For two vectors \mathbf{u} and \mathbf{v} , RBF function $\kappa(\mathbf{u}, \mathbf{v})$ is determined by Equation (8).

$$\kappa(\mathbf{u}, \mathbf{v}) = \exp(-\gamma \|\mathbf{u} - \mathbf{v}\|^2) \tag{8}$$

2.2.2. Hyperparameter optimisation

Recapping the above, the performance of a one-class SVM classifier utilising RBF as kernel function and given a specific training dataset depends on the two hyperparameters: γ and ν . The optimal values of these hyperparameters are selected through random search (Bergstra et al. 2011); a predefined number of ν and γ values are randomly selected from their relevant search spaced and fed as hyperparameter tuples to the model optimiser.

Models are trained for each hyperparameter tuple in the two-dimensional grid using the training sub-dataset through cross-validation. The hyperparameter tuple h^* that

achieve the best results (lower error) in average, are finally selected and a model M^* is trained based on the whole dataset and h^* .

As previously mentioned, K -fold cross validation is implemented in order to safeguard model’s generalisation. However, in the case of OCC, only having validation data points belonging to one class does not suffice to ensure adequate generalisation. At the same time, access to abnormal data points at the validation stage cannot be merely assumed while developing a methodology, as in that case, developing a multi-class classifier with imbalanced data sets (e.g. Chawla et al. (2004); Van Hulse et al. (2007)) would be more appropriate. For this reason, a number of abnormal data points are algorithmically simulated without any requirements for a priori knowledge of the physical system. These points are generated by adding 4 or 5 standard deviations to the actual values, therefore, the probability that they are normal is in the range of 6×10^{-5} to 6×10^{-7} . The algorithm implemented to produce this simulated abnormal dataset is shown in Algorithm 2. Having derived the abnormal validation dataset, this is appended to the “normal” one, i.e. the one derived through K -folding and the two are concatenated.

Algorithm 2 Algorithmic derivation of abnormal points in array X_{abn}

Require: X

```

1:  $n_{attr} \leftarrow$  number of attributes in  $X$ 
2: Initialise  $X_{abn}$  with  $4 \times n_{attr}$  randomly selected observations from  $X$ 
3: for  $i = 1, 2, \dots, n_{attr}$  do
4:    $attr_{max} \leftarrow \max(X[i])$ 
5:    $attr_{min} \leftarrow \min(X[i])$ 
6:    $attr_{std} \leftarrow \text{std}(X[i])$ 
7:    $X_{abn}[4i - 3, i] \leftarrow attr_{max} + 4 \times attr_{std}$ 
8:    $X_{abn}[4i - 2, i] \leftarrow attr_{min} - 4 \times attr_{std}$ 
9:    $X_{abn}[4i - 1, i] \leftarrow attr_{max} + 5 \times attr_{std}$ 
10:   $X_{abn}[4i, i] \leftarrow attr_{min} - 5 \times attr_{std}$ 
11: end for
12: return  $X_{abn}$ 

```

In order to evaluate the measure of the quality of the derived model, the Matthews Correlation Coefficient (MCC) of the validation dataset was calculated for each hyperparameter tuple. MCC returns the correlation coefficient between actual and predicted binary classifications and returns values in the range $[-1, +1]$. MCC can be calculated by using the following equation (Matthews 1975):

$$\begin{aligned}
\text{MCC} &= \frac{TP/N - S \times P}{\sqrt{P \times S(1 - S)(1 - P)}} \\
N &= TN + TP + FN + FP \\
S &= \frac{TP + FN}{N} \\
P &= \frac{TP + FP}{N}
\end{aligned} \tag{9}$$

where TP corresponds to true positives, TN to true negatives, FP to false positives and FN to false negatives.

2.3. Normality estimation

In the “anomaly detection model” step (described in subsection 2.2), a hyperplane that encompasses the original normal points is derived in order to label accordingly new points as either normal or abnormal. At the same time, the distance between the hyperplane and each new point can be used as a normality metric (Gkerekos et al. 2018). The furthest a point is from the hyperplane towards the normal side, the normality of the system examined increases, and vice versa. This also presents an inherent limitation of one-class models, i.e. their ability to only identify abnormalities at system level, without the ability to examine individual components for maintenance optimisation.

Additionally, as every data point is evaluated independently, an Exponential Weighted Moving Average (EWMA) filter is implemented to smooth-out unnecessary kinks and make the evaluation of the overall degradation easier (Gkerekos et al. 2018). EWMA for a series Y can be calculated recursively by using the following equation (Croarkin et al. 2018):

$$S_t = \begin{cases} Y_1, & t = 1 \\ a \cdot Y_t + (1 - \alpha) \cdot S_{t-1}, & t > 1 \end{cases} \quad (10)$$

where $t = 1$ refers to the first point, with t increasing for each subsequent point and parameter α can be calculated from the Centre Of Mass (COM) property through the following equation:

$$\alpha = 1 / (1 + COM) \quad (11)$$

2.4. Methodology verification

The model verification is a challenging task, especially in cases where modelling is performed at system-level, as it is generally extremely difficult or outright impossible to model the entire possible input domain and therefore, a simulation model can only be an approximation to the actual system regardless of effort spent on the model development (Law 2006). Even when a model is considered to be validated against an observable system, in cases where the model is used for alternative configurations, the underlying assumptions may no longer be valid (Dinwoodie 2014).

To verify the proposed methodology, sensitivity analysis was employed, considering multiple sensitivity case scenarios. The sensitivity analysis process examines the flexibility of the developed methodology in introducing partial deviation of the utilised input data (Saltelli et al. 2004; Dikis 2017).

The case of this study, while retaining the original values of all features of the testing dataset, one feature is manipulated in order to verify that the methodology performs as expected. More specifically, an abnormal variation of a model feature is expected to reduce similarity between nominal and unclassified observations and consequently increase the probability of abnormal condition.

3. Methodology application

In this section, a case study based on actual ship measurements is presented. More specifically, D/Gen measurements from a diverse set of load conditions are used in or-

der to train a “normal” model capable of classifying a number of D/Gen performance parameters. A sensitivity analysis is then performed on the model, evaluating model behaviour under a number of simulated scenarios. A diagram of the system considered is shown in Figure 2, highlighting in bold the location where measurements are obtained.

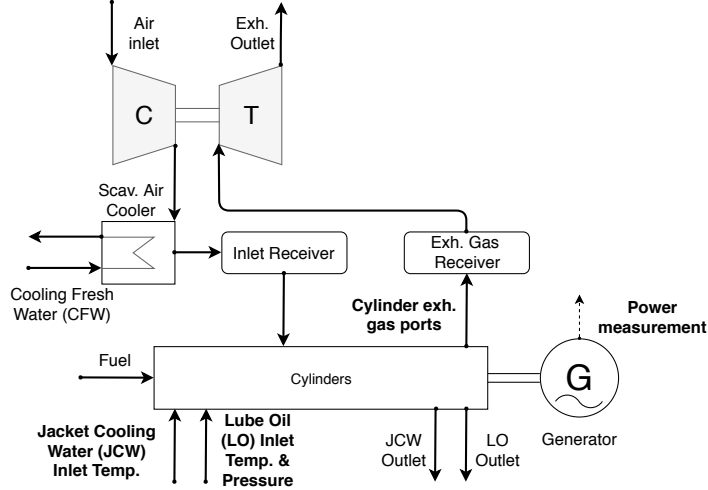


Figure 2. Diagram of a D/Gen system, showing parameter measuring locations in bold.

The features considered for D/Gen model training are presented in Table 1. Specifically, minimum and maximum exhaust gas temperature measurements were provided, inlet and outlet temperature of the scavenge air receiver, fresh water cooler inlet temperature, lubricating oil inlet temperature and pressure as well as power output. 1095 recordings were available, corresponding 317.5 days of recording, at 6 points per day. As the D/Gen examined in this case study was not always operational, points with “NaN” values were excluded from the analysis carried out, following the first step of the pre-processing methodology (described in sub-section 2.1). Following that, 804 points remained available. Next, the second step of the pre-processing methodology was applied, removing an additional 35 data points.

Subsequently, the remaining dataset is split into 70% – 30% respectively for training and testing.

Table 1. Noon-report measurements considered as input for D/Gen model training

Component	variable name	description	units
Misc.	output	Power output	kW
Lub. Oil	lo_inl_t	Lube Oil (LO) inlet temperature	°C
	lo_inl_p	LO inlet pressure (manometric)	bar
Fresh Water Cool.	fw_inl_t	Cooling Fresh Water (CFW) inlet temperature	°C
Cylinder Exh. Gas port	max_exh_t	Max temperature	°C
	min_exh_t	Min temperature	°C

Following the removal of extreme outliers, the results of a cursory univariate analysis on the remaining training dataset is shown in Table 2, where the mean, minimum and maximum values of each feature are presented, along with the standard deviation and the values at different (25, 50, and 75%) quantiles.

Table 2. Descriptive statistics of training dataset

	output (kW)	lo_inl_t (°C)	lo_inl_p (bar)	fw_inl_t (°C)	max_exh_t (°C)	min_exh_t (°C)
count	538	538	538	538	538	538
mean	652.23	64.05	4.50	72.34	305.45	285.69
std	51.94	1.48	0.14	2.92	13.20	12.82
min	500.00	60.00	4.10	64.00	270.00	260.00
25%	600.00	63.00	4.40	69.00	300.00	280.00
50%	650.00	64.00	4.50	73.00	300.00	280.00
75%	700.00	65.00	4.60	75.00	320.00	300.00
max	800.00	68.00	4.90	77.00	330.00	310.00

In order to evaluate model diagnostic performance, the testing dataset was augmented with simulated faults in the form of a sensitivity analysis. To achieve that, for each selected attribute in the testing dataset, the testing dataset was replicated and then the values of that attribute were linearly altered in order to reach the alarm or shutdown thresholds set by the OEM. The two testing datasets are then concatenated and fed to the trained condition monitoring model as input with its output being observed. Expected model behaviour is to return a positive distance from the hyperplane (i.e. normal behaviour) for points belonging to the “normal” dataset and then that distance decrease, cross zero and eventually take negative values (i.e. abnormal behaviour) as the manipulated attribute exceeds the normal range.

4. Results

This section presents and discusses the results obtained through the model training and relevant case studies introduced in Section 3.

In Figure 3, the results of the random search employed to determine the optimal values of hyperparameters ν and γ in the model trained on the D/Gen dataset is presented, with a heatmap denoting the MCC values obtained in the validation phase and dots depicting the hyperparameter combinations that were tested. Darker colours correspond to better MCC performance on the validation dataset; and conversely. Lower values of ν , especially in the $(0, 0.2)$ range and γ values in the $(10^{-1}, 10^{-4})$ range yielded higher MCC values on validation. Higher MCC values indicate better model performance at accurately classifying data points as normal or abnormal. In order to obtain this heatmap, the OCC SVM model was trained for 3000 randomly selected (ν, γ) tuples from the relevant ranged discussed in Section 2.2. Through that, it was found that the optimal (ν, γ) tuple was $(0.0049, 0.0016)$ for this case study.

In Figure 4, model performance in increasing CFW inlet temperature is presented. Normal behaviour (approximately first 230 points) of the CFW inlet temperature are associated within the $68 - 76$ °C range. Concurrently, the EWMA of the model output is steadily positive, therefore correctly identifying these points as normal.

As the CFW inlet temperature starts to increase, the distance of the relevant points from the hyperplane defining “normality” starts decreasing, eventually turning negative (i.e. points being classified as abnormal) slightly before the temperature crosses the alarm threshold. By the time the CFW inlet temperature reached the OEM shutdown threshold, the distance from hyperplane has taken a large negative value, with those points being labelled as abnormal with a high confidence. Based on the above

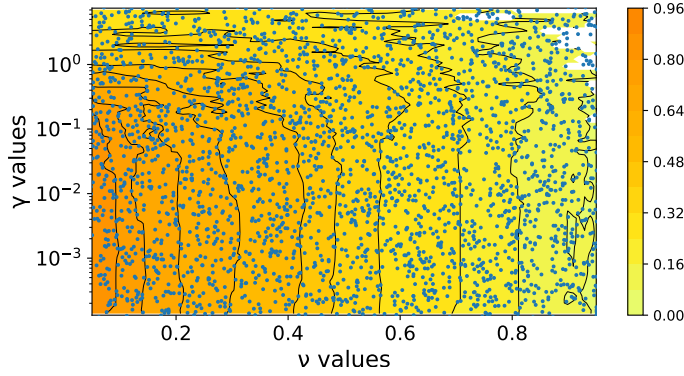


Figure 3. Contour plot depicting hyperparameter optimisation results for different combinations of (ν, γ) . Tested combinations are shown in dots and the MCC score achieved in K -folding is depicted by the colour of the contours.

results, it is inferred that the OCC SVM model implemented can accurately identify deviating patterns in a time-series dataset, and reflect that in its normality prediction.

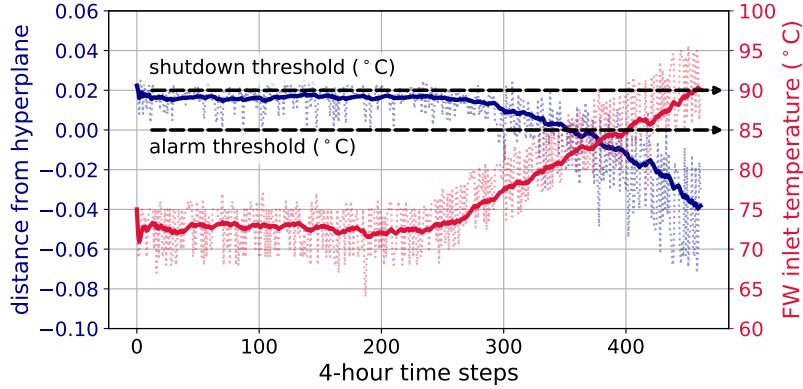


Figure 4. CFW inlet temperature sensitivity analysis. Actual distance from hyperplane (dark) and temperature measurements (light) are shown in thin, dotted lines and the EWMA-filtered values in a thick, solid line. Positive distance from hyperplane signifies points identified as “normal”; and conversely. A larger distance conveys a higher prediction confidence.

In Figure 5, the same process is repeated for the LO inlet pressure measurement. Analysing the testing dataset, normal measurement range is within the 4.2 – 4.7 bar range, with the OEM alarm threshold at 4.0 bar and the shutdown threshold at 3.5 bar.

As the LO pressure starts dropping, it is observed that the distance of the relevant points from the hyperplane decreases, and eventually becomes negative slightly before the pressure drops beyond the alarm threshold. This manifests the model’s ability to identify incipient deviations from the “normal” pattern that was learnt during training. Similarly to the preceding CFW inlet temperature case study, as the pressure reaches the OEM shutdown threshold, the corresponding points are assigned a large negative distance value.

Complimentary to the LO inlet pressure sensitivity analysis presented above, the same methodology is repeated for the LO inlet temperature in Figure 6. For this measurement, the OEM only provides the shutdown threshold, at 75°C and normal

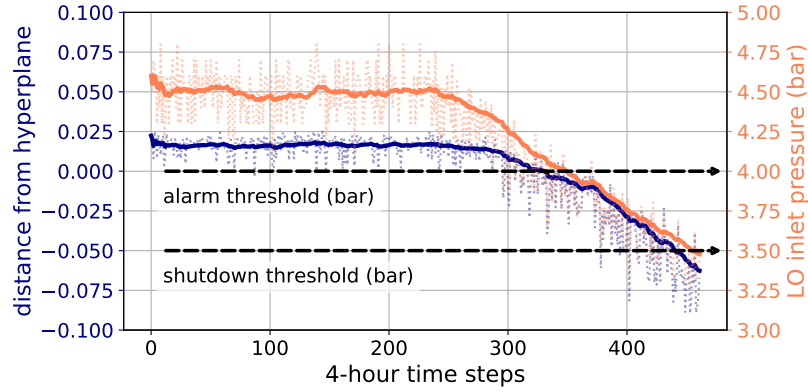


Figure 5. LO inlet pressure sensitivity analysis. Actual distance from hyperplane (dark) and pressure measurements (light) are shown in thin, dotted lines and the EWMA-filtered values in thick, solid line. Positive distance from hyperplane signifies points identified as “normal”; and conversely. A larger distance conveys a higher prediction confidence.

measurement range at the 62 – 66°C range.

In this case, it can be observed that as the LO inlet temperature increases towards the shutdown threshold, the distance from the separating hyperplane becomes increasingly negative. Following the results of the previous two case studies where it was observed that the alarm threshold is crossed at values where the hyperplane distance is approximately zero, an reversed process could be employed to identify an alarm threshold at approximately 70°C for LO inlet temperature measurements.

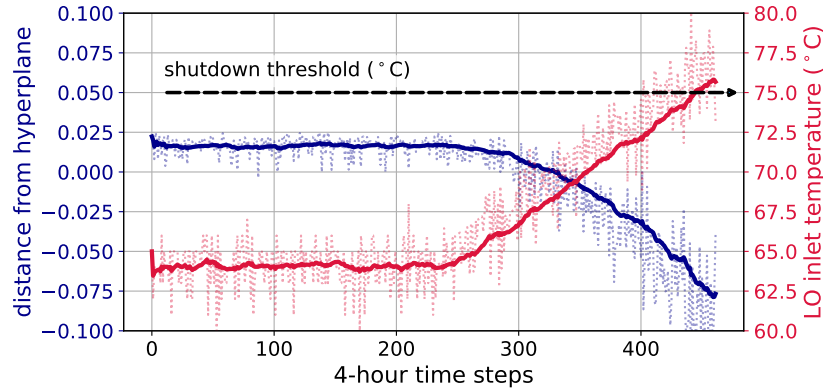


Figure 6. LO inlet temperature sensitivity analysis. Actual distance from hyperplane (dark) and temperature measurements (light) are shown in thin, dotted lines and the EWMA-filtered values in a thick, solid line. Positive distance from hyperplane signifies points identified as “normal”; and conversely. A larger distance conveys a higher prediction confidence.

Finally, in Figure 7, the sensitivity analysis for the maximum cylinder exhaust port gas temperature measurement is demonstrated. This measurement is utilised as the measurements for all cylinders are aggregated on board the vessel and only their max/min values are transported on shore. OEM provides an overall exhaust gas temperature alarm threshold at 480°C while measurement range observed in the testing dataset was at 300 – 320°C.

In this case, it was observed that as the maximum exhaust gas temperature increases, the distance from hyperplane decreases and becomes negative. However, at this case study the deviation that triggers the distance-from-hyperplane drop is de-

tected at approximately two-thirds of the actual alarm threshold. This presents the main shortcoming of the proposed methodology, i.e. the fact that the normality predicted by the model is affected by the range of inputs at the training stage. This effect is mostly present in the case of the D/Gen exhaust gas temperature, as this measurement correlates majorly with engine operating conditions whereas all other measurements examined fluctuate in a OEM preset range through the operation of coolers and pumps. If only a subset of normal operating conditions are provided, the model will focus on that range, identifying anything beyond that as abnormal. However, this can be easily resolved by re-training the model when the range of normal operating conditions is modified.

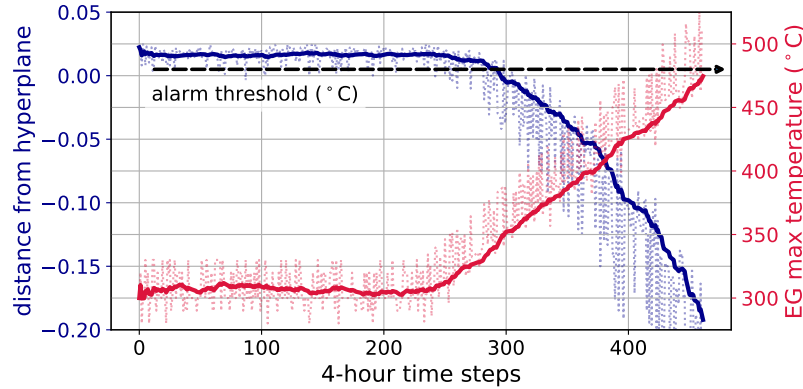


Figure 7. Maximum exhaust gas temperature sensitivity analysis. Actual distance from hyperplane (dark) and temperature measurements (light) are shown in thin, dotted lines and the EWMA-filtered values in a thick, solid line. Positive distance from hyperplane signifies points identified as “normal”; and conversely. A larger distance conveys a higher prediction confidence.

5. Conclusions

This study presents a methodology for the monitoring and detection of operating anomalies in ship machinery using measured performance data. For this, an OCC SVM model is trained using a dataset corresponding to normal behaviour of a D/Gen under varying operating conditions, with modelling selections being extensively discussed. The main findings of the research conducted are as follows:

- The developed model can accurately discern between normal and abnormal machinery condition, providing a suitable basis for an incipient-fault detection system.
- Due to limited amount of assumptions, this methodology can be applied to a diverse set of machinery.
- Contrary to most other condition monitoring approaches, model training only requires normal data, majorly simplifying the process.
- Proposed methodology data requirements can be satisfied from noon-report data, significantly decreasing, or even avoiding, CAPEX costs traditionally attached to condition monitoring applications.
- Using the EWMA filter, the proposed methodology becomes dynamic as previous data points are taken into consideration when predicting the normality of new data points.

- Through the performed sensitivity analysis, model alarm thresholds were obtained. Comparing these values with the ones provided by OEMs, a significant degree of similarity was observed.
- Even in cases where only a limited amount of OEM thresholds are available, relevant threshold can be suggested by analysing model output.
- The major shortcoming of the suggested modelling approach is that data points that correspond to a normal machinery operating condition still be flagged as abnormal if no similar points are part of the training set. This can also be seen as a beneficial feature, as it can highlight incipiently anomalous operation if dissimilar to regular.

The proposed methodology was elaborated and showcased through case studies utilising actual D/Gen raw data and applications of sensitivity analysis on each relevant measurement. It was demonstrated as no other assumptions apart from a dataset corresponding to normal behaviour of the system was required for model training. Accordingly, this methodology can be applied to a diverse set of machinery, without a requirement for OEM thresholds or a dataset with both “normal” and “abnormal” observations. It is expected that this methodology will form a useful tool for the development of dynamic, self-learning condition monitoring systems without requirements for extensive data gathering or input from system experts. Additionally, this methodology combined with the sensitivity analysis implemented within can be a useful aid in the derivation of measurement thresholds in a formalised manner.

Acknowledgements

The work presented in this paper is partially funded by INCASS project. INCASS Project has received research funding from the European Unions Seventh Framework Program under grant agreement no 605200. This publication reflects only the authors views and European Union is not liable for any use that may be made of the information contained within.

Abbreviations

CAPEX Capital Expenditure.
CFW Cooling Fresh Water.
CODLAG COmbined Diesel eLectric And Gas.
COM Centre Of Mass.
D/Gen Diesel Generator set.
DAQ Data Acquisition.
EWMA Exponential Weighted Moving Average.
IACS International Association of Classification Societies.
INCASS Inspection Capabilities for Enhanced Ship Safety.
ISM International Safety Management.
LO Lube Oil.
MCC Matthews Correlation Coefficient.
OCC One Class Classification.
OEM Original Equipment Manufacturer.
OPEX Operating Expenses.

PMS Planned Maintenance System.
RBF Radial Basis Function.
SCADA Supervisory Control and Data Acquisition.
SVM Support Vector Machine.

References

- Bergstra JS, Bardenet R, Bengio Y, Kégl B. 2011. Algorithms for hyper-parameter optimization. In: *Advances in neural information processing systems*. p. 2546–2554.
- Chandroth G. 2004. Condition monitoring: the case for integrating data from independent sources. *Journal of Marine Engineering & Technology*. 3(1):9–16.
- Chang CC, Lin CJ. 2001. Training ν -support vector classifiers: Theory and algorithms. *Neural Comput*. 13(9):2119–2147.
- Chawla NV, Japkowicz N, Kotcz A. 2004. Editorial: Special issue on learning from imbalanced data sets. *SIGKDD Explor Newsl*. 6(1):1–6.
- Cipollini F, Oneto L, Coraddu A, Murphy AJ, Anguita D. 2018a. Condition-based maintenance of naval propulsion systems: Data analysis with minimal feedback. *Reliability Engineering & System Safety*.
- Cipollini F, Oneto L, Coraddu A, Murphy AJ, Anguita D. 2018b. Condition-based maintenance of naval propulsion systems with supervised data analysis. *Ocean Engineering*. 149:268 – 278.
- Coraddu A, Oneto L, Ghio A, Savio S, Anguita D, Figari M. 2016. Machine learning approaches for improving condition-based maintenance of naval propulsion plants. *Proceedings of the Institution of Mechanical Engineers, Part M: Journal of Engineering for the Maritime Environment*. 230(1):136–153.
- Cortes C, Vapnik V. 1995. Support-vector networks. *Machine Learning*. 20(3):273–297.
- Croarkin C, Tobias P, Filliben JJ, Hembree B, Guthrie W, Trutna L, Prins J, editors. 2018. *NIST/SEMATECH e-Handbook of Statistical Methods*. NIST/SEMATECH. Available from: <http://www.itl.nist.gov/div898/handbook/>.
- Dikis K. 2017. Establishment of a novel predictive reliability assessment strategy for ship machinery [dissertation]. University of Strathclyde, Department of Naval Architecture, Ocean and Marine Engineering.
- Dikis K, Lazakis I. 2016. Dynamic risk and reliability assessment of ship machinery and equipment. In: *The 26th International Ocean and Polar Engineering Conference*. International Society of Offshore and Polar Engineers. p. 969–976.
- Dikis K, Lazakis I, Michala A, Raptodimos Y, Theotokatos G. 2017. Dynamic risk and reliability assessment for ship machinery decision making. In: *Risk, Reliability and Safety: Innovating Theory and Practice - Proceedings of the 26th European Safety and Reliability Conference, ESREL 2016*. p. 685–692.
- Dinwoodie I. 2014. Modelling the operation and maintenance of offshore wind farms [dissertation]. University of Strathclyde, Department of Electronic and Electrical Engineering.
- Geman S, Bienenstock E, Doursat R. 1992. Neural networks and the bias/variance dilemma. *Neural computation*. 4(1):1–58.
- Gkerekos C, Lazakis I, Theotokatos G. 2016. Ship machinery condition monitoring using vibration data through supervised learning. In: Lazakis I, Theotokatos G, editors. *Proceedings of MSO2016, International Conference on Maritime Safety and Operations*; Oct. University of Strathclyde Print Services. p. 103–110.
- Gkerekos C, Lazakis I, Theotokatos G. 2017a. Implementation of a self-learning algorithm for main engine condition monitoring. In: Soares C, Teixeira A, editors. *Maritime Transportation and Harvesting of Sea Resources*; vol. 2; 7. p. 981–989.
- Gkerekos C, Lazakis I, Theotokatos G. 2017b. Ship machinery condition monitoring using performance data through supervised learning. In: *Proceedings of Smart Ships Technology*

- 2017 Conference; 1. Royal Institution of Naval Architects (RINA). p. 105–111.
- Gkerekos C, Lazakis I, Theotokatos G. 2018. Exploiting machine learning for ship systems anomaly detection and healthiness forecasting. In: Proceedings of Smart Ships Technology 2018 Conference; 1. Royal Institution of Naval Architects (RINA). p. 1–6.
- Hastie T, Tibshirani R, Friedman J. 2001. The elements of statistical learning. New York, NY, USA: Springer New York Inc. Springer Series in Statistics.
- Hountalas DT. 2000. Prediction of marine diesel engine performance under fault conditions. *Applied Thermal Engineering*. 20(18):1753 – 1783.
- Hsu CW, Chang CC, Lin CJ. 2010. A practical guide to support vector classification.
- Khan SS, Madden MG. 2010. A survey of recent trends in one class classification. In: Coyle L, Freyne J, editors. *Artificial Intelligence and Cognitive Science*; Berlin, Heidelberg. Springer Berlin Heidelberg. p. 188–197.
- Kohavi R. 1995. A study of cross-validation and bootstrap for accuracy estimation and model selection. Morgan Kaufmann. p. 1137–1143.
- Kowalski J, Krawczyk B, Woniak M. 2017. Fault diagnosis of marine 4-stroke diesel engines using a one-vs-one extreme learning ensemble. *Engineering Applications of Artificial Intelligence*. 57:134 – 141.
- Lamaris V, Hountalas D. 2010. A general purpose diagnostic technique for marine diesel engines application on the main propulsion and auxiliary diesel units of a marine vessel. *Energy Conversion and Management*. 51(4):740 – 753.
- Law AM. 2006. How to build valid and credible simulation models. In: Proceedings of the 38th Conference on Winter Simulation. Winter Simulation Conference. p. 58–66. WSC '06.
- Lazakis I, Ölçer A. 2015. Selection of the best maintenance approach in the maritime industry under fuzzy multiple attributive group decision-making environment. *Proceedings of the Institution of Mechanical Engineers, Part M*. 230(2):297–309.
- Lazakis I, Turan O, Aksu S. 2010. Increasing ship operational reliability through the implementation of a holistic maintenance management strategy. *Ships and Offshore Structures*. 5(4):337–357.
- Li P, Liu L, Gong H. 2010. The research of the intelligent fault diagnosis optimized by aca for marine diesel engine. In: Luo Q, editor. *Advancing computing, communication, control and management*. Springer; p. 174–181.
- MAN B&W Diesel A/S. 2004. Instruction Book ‘Operation’ for 50-108MC/MC-C Engines. 2nd ed. Copenhagen, DK.
- Manzini R, Regattieri A, Pham H, Ferrari E. 2009. *Maintenance for industrial systems*. Springer Science & Business Media.
- Matthews B. 1975. Comparison of the predicted and observed secondary structure of t4 phage lysozyme. *Biochimica et Biophysica Acta (BBA) - Protein Structure*. 405(2):442 – 451.
- Mohanty AR. 2014. *Machinery condition monitoring*. CRC Press.
- Moore Stephens. 2017. Future operating costs report. Report No: DPS38615.
- Neale M, Associates. 1979. A guide to the condition monitoring of machinery. London, Great Britain: Committee for Terotechnology. Report No:.
- Randall RB. 2011. *Vibration-based condition monitoring*. John Wiley & Sons.
- Raptodimos Y, Lazakis I. 2018. Using artificial neural network-self-organising map for data clustering of marine engine condition monitoring applications. *Ships and Offshore Structures*. 13(6):649–656.
- Saltelli A, Tarantola S, Campolongo F, Ratto M. 2004. *Sensitivity analysis in practice: a guide to assessing scientific models*. John Wiley & Sons.
- Schölkopf B, Platt JC, Shawe-Taylor JC, Smola AJ, Williamson RC. 2001. Estimating the support of a high-dimensional distribution. *Neural Comput*. 13(7):1443–1471.
- Schölkopf B, Smola AJ, Williamson RC, Bartlett PL. 2000. New support vector algorithms. *Neural Computation*. 12(5):1207–1245.
- Schölkopf B, Williamson R, Smola A, Shawe-Taylor J, Platt J. 1999. Support vector method for novelty detection. In: Proceedings of the 12th International Conference on Neural Information Processing Systems; Cambridge, MA, USA. MIT Press. p. 582–588. NIPS'99.

- Stopford M. 2009. Maritime economics. Third edition ed. Routledge.
- UNCTAD. 2015. Review of maritime transport. New York and Geneva: United Nations Conference on Trade and Development. Report No: UNCTAD/RMT/2015.
- Van Hulse J, Khoshgoftaar TM, Napolitano A. 2007. Experimental perspectives on learning from imbalanced data. In: Proceedings of the 24th International Conference on Machine Learning; New York, NY, USA. ACM. p. 935–942. ICML '07.
- Verbert K, Schutter BD, Babuka R. 2017. Timely condition-based maintenance planning for multi-component systems. *Reliability Engineering & System Safety*. 159:310 – 321.
- Wätzenig D, Sommer MS, Steiner G. 2009. Engine state monitoring and fault diagnosis of large marine diesel engines. *Elektrotechnik und Informationstechnik*. 126(5):173–179.
- Widodo A, Yang BS. 2007. Support vector machine in machine condition monitoring and fault diagnosis. *Mechanical Systems and Signal Processing*. 21(6):2560 – 2574.
- Wu X, Srihari RK. 2003. New ν -support vector machines and their sequential minimal optimization. In: Proceedings of the 20th International Conference on Machine Learning (ICML-03). p. 824–831.
- Yang W, Court R, Jiang J. 2013. Wind turbine condition monitoring by the approach of SCADA data analysis. *Renewable Energy*. 53:365 – 376.
- Yin Z, Hou J. 2016. Recent advances on SVM based fault diagnosis and process monitoring in complicated industrial processes. *Neurocomputing*. 174, Part B:643 – 650.