PERFORMANCE COMPARISON OF MACHINE LEARNING FAULT DETECTION AND DIAGNOSIS ALGORITHMS IN ORGANIC RANKINE CYCLE SYSTEMS

Andres Hernandez[1*], Aitor Cendoya[1], Basile Chaudoir[1], Vincent Lemort[1]

[1]Thermodynamics Laboratory / Université de Liège, Liège, Belgium

Corresponding Author: jahernandez@uliege.be

ABSTRACT

This study investigates fault detection and diagnosis (FDD) techniques aimed at enhancing the reliability and operational efficiency of organic Rankine cycle (ORC) systems used for waste heat recovery. By improving fault management, this approach can significantly reduce maintenance costs and unplanned downtime, contributing to the broader adoption of ORC technology in sustainable energy applications. The proposed methodology leverages advanced machine learning algorithms, specifically support vector machines (SVM) and Random Forest (RF), to identify and classify critical system faults. These faults include evaporator fouling, the presence of non-condensable gases, and mechanical issues in the expander and pump components. To evaluate the performance of the FDD framework, simulation-based experiments are conducted to address practical challenges such as data scarcity and noise, which are common in industrial applications. Results demonstrate the robustness of the SVM and RF models in accurately detecting and diagnosing faults, highlighting their potential to maintain high system performance in real-world scenarios. Additionally, the analysis provides insights into selecting appropriate strategies based on the quantity and quality of data available. Furthermore, the study explores the trade-offs between computational efficiency and diagnostic accuracy, offering insights into the applicability of these techniques for online monitoring systems. The findings underscore the critical role of machine learning in predictive maintenance strategies, paving the way for smarter and more resilient energy recovery solutions.

1 INTRODUCTION

The increasing global demand for energy efficiency and sustainability has intensified interest in Organic Rankine Cycle (ORC) systems for waste-heat recovery and renewable energy generation (Wieland et al., 2023). ORC technology offers a viable solution to improve energy utilization in industrial applications by converting low-grade heat into electricity. However, the fluctuating nature of waste heat sources often forces ORC systems to operate under off-design conditions (Dickes, 2019). This variability poses significant challenges in distinguishing between normal operational deviations and actual system faults based solely on available sensor data. Consequently, the development of effective Fault Detection and Diagnosis (FDD) techniques is crucial for maintaining optimal performance and ensuring the reliability of these complex energy systems (Wang et al., 2021).

FDD methodologies have evolved significantly, encompassing model-based, data-driven, and hybrid approaches. Model-based techniques utilize the physical laws governing system behavior to detect anomalies by comparing measured and predicted operational parameters (Dragan, 2011). Such methods have been successfully implemented in heat exchangers and other industrial equipment, enabling early fault identification and predictive maintenance (Choi and Krumdieck, 2016). However, they often require extensive calibration and may struggle with highly nonlinear or uncertain system dynamics.

Machine learning-based FDD approaches have gained traction due to their ability to learn complex patterns from historical data, offering adaptability and high detection accuracy. These methods have been extensively applied to Heating, Ventilation, and Air Conditioning (HVAC) and refrigeration systems (Elmouatamid et al., 2023; van de Sand, 2021). For instance, predictive frameworks leveraging unsupervised learning have been developed to assess system degradation states (van de Sand, 2021). Additionally, studies on air conditioning systems have shown that soft faults, which gradually deteriorate system performance, require sophisticated detection strategies due to their subtle nature (Elmouatamid et al., 2023).

Hybrid FDD approaches, which integrate model-based and data-driven techniques, have demonstrated promising results in complex applications. Digital twin frameworks and filtering algorithms such as the Extended Kalman Filter and One-Class Support Vector Machine (SVM) have been explored for real-time diagnostics in energy systems (Hosamo et al., 2022; Yan et al., 2017). These frameworks enhance fault detection accuracy and support decision-making by providing actionable insights into system health and performance.

Despite the progress in FDD research, literature specifically addressing ORC systems remains limited. Notable

contributions include the application of machine learning to detect multiple ORC faults using a small set of sensor data (Wang et al., 2021). This underscores the need for further research into robust FDD methodologies tailored to ORC applications, particularly in addressing challenges such as data scarcity and sensor noise.

This study investigates the effectiveness of Support Vector Machine (SVM) and Random Forest (RF) classifiers in diagnosing critical faults in ORC systems, including evaporator fouling, non-condensable gas presence, and mechanical issues in expanders and pumps. Through simulation-based experiments, we assess the impact of data scarcity and noise on model performance, providing insights into optimal feature selection for robust fault detection. The findings contribute to advancing predictive maintenance strategies, paving the way for smarter and more resilient ORC-based energy recovery solutions.

2 METHODOLOGY

This section provides an overview of the organic Rankine cycle architecture, along with a general analysis of the faults examined in this study and the adopted simulation approach.

2.1 ORC process architecture

The ORC power unit analyzed in this study features a 50 kWe expander and utilizes refrigerant R1233ZD(E) as the working fluid, with its architecture illustrated in Figure 1.

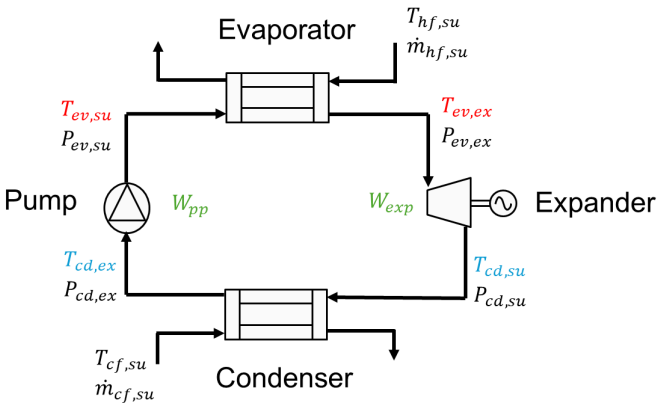


Figure 1: Process architecture

The refrigerant in liquid state is pumped to a higher pressure before reaching the evaporator inlet. In the evaporator, heat is exchanged between the refrigerant and the secondary fluid, which in this case is water at an average temperature of $T_{\mathrm{hf,su}} = 90°\mathrm{C}$ and a mass flow rate of $\dot{m}_{\mathrm{hf,su}} = 12.5$ kg/s. The refrigerant undergoes phase change, boiling and heating up until it reaches a superheated state at the expander inlet.

The vapor then passes through the expander, causing it to rotate and generate electrical power when connected to a generator. This expansion process results in a drop in temperature and pressure of the refrigerant at the condenser inlet. The condenser ensures that the refrigerant returns to a liquid state, thereby closing the thermodynamic cycle. Water is used as the secondary fluid in the condenser, with an average temperature of $T_{\mathrm{cf,su}} = 10°\mathrm{C}$ and a mass flow rate of $\dot{m}_{\mathrm{cf,su}} = 22.5$ kg/s. The overall overview of the ORC components modeling approach followed is encountered in Table 1. In this article our focus is on the development of the machine learning methodology for fault detection and diagnosis, the interested reader is referred to (Cendoya et al., 2024) and (Cendoya et al., 2025), for details regarding components and model validation.

2.2 Faults description

In this study, five conditions are analyzed: the healthy state, heat exchanger fouling, condenser fault due to non-condensable gases, expander mechanical fault, and pump mechanical fault. Initially, a physical interpretation of each fault is presented, followed by a description of our attempt for replicating those faults in simulation.

Evaporator Fouling: Over time, deposits such as crystallization, sedimentation, and corrosion accumulate on heat exchanger surfaces, reducing heat transfer efficiency and impairing system performance (Markowski et al., 2013).

Table 1: Overview of the ORC Component Modeling Approach

Component	Model	Correlations	Validation Error
Condenser/ Evaporator	(Cuevas et al., 2009): Brazed plate heat exchanger discretized according to different refrigerant phases (subcooling, two-phase and superheating).	**HTC:** Single phase (Martin, 2010) Condensation, (Shah, 2021) Boiling, (Amalfi et al., 2016a) **Pressure Drop:** Single phase, (Martin, 2010) Condensation, (Longo, 2010) Boiling, (Amalfi et al., 2016b) **Charge:** (Cioncolini and Thome, 2012)	**Condenser:** ±0.41% for the heat power, ±3.2% for the pressure drops, and ±1.12°C for the exhaust temperature. **Evaporator:** ±1.76% for the heat Power, ±0.05% for the pressure drops, and 0.2°C for the exhaust temperature
Pump	(Eck, 1973): Dimensionless factors (power, flow, and pressure) used to predict the performance and behaviour of pumps and fans	/	The R^2 for the factors curve fit based on the pressure factor are 99.5% and 97.6% for the flow factor and power factor, respectively
Expander	(Lemort et al., 2012) Semi-empirical model capable of accounting for over and under-expansion phenomena, as well as refrigerant leakages, electromechanical losses, and internal heat transfer effects	/	Number of manufacturer points: 18. Errors of ±5% for the electrical power, ±0.05% for the refrigerant mass flow rate, and ±6 K for the exhaust temperature.

In ORC systems, this leads to lower working fluid temperatures at the expander inlet, reducing enthalpy drop and overall power output.

Non-Condensable Gases (NCGs): In ORC condensers, NCGs enter through air leaks, fluid degradation, or material desorption, reducing heat transfer efficiency and increasing back pressure on the expander. This leads to a lower pressure ratio and reduced power output. Additionally, NCGs decrease mass flow rate and can cause cavitation and corrosion, further degrading performance.

Pump Mechanical Faults: The pump ensures the continuous circulation of the working fluid, but faults such as internal leakage, impeller wear, and reduced rotational speed hinder its performance. These issues lead to lower mass flow rates, affecting heat absorption in the evaporator and decreasing power output.

Expander Mechanical Faults: The expander converts thermal energy into mechanical work, directly affecting power generation. Common faults include excessive wear, irregular discharge volume, and structural degradation, which cause inefficiencies, abnormal vibrations, and overheating. These faults lead to reduced energy conversion efficiency and can accelerate mechanical wear.

2.3 Data generation

Using the simulation model described in Section 2.1, a series of simulations are performed to explore different temperatures at the secondary fluid supply of the evaporator ($T_{hf,su}$) and the condenser ($T_{cf,su}$), while superheating (ΔT_{sh}) and subcooling (ΔT_{sc}) are kept regulated. This regulation reflects real-world practice, where proportional-integral-derivative (PID) controllers are commonly employed for this task, adjusting the pump speed (N_{pp}) and the mass-flow rate in the condenser ($\dot{m}_{cf,su}$). Thus the nominal simulation parameters are presented in Table 2.

The numerical model, under nominal conditions, represents the healthy state, where simulations are performed at different heat source and heat sink temperatures ($T_{hf,su}$ and $T_{cf,su}$, respectively). The recorded dataset includes 150 steady-state samples of temperatures, pressures, power consumption and generation, as well as mass flow rates, which are stored for further processing. Next, it is essential to discuss how the faults are emulated to generate the complete set of simulation data.

In this study, a simplified approach is employed to emulate the effects of non-condensable gases (NCGs) or fouling

Table 2: Nominal simulation conditions

Simulation condition	symbol	value	unit
Heat source temperature	$T_{hf,su}$	80-130	°C
Heat source mass flow rate	$\dot{m}_{hf,su}$	18	kg/s
Heat sink temperature	$T_{cf,su}$	10-25	°C
Heat sink mass flow rate	$\dot{m}_{cf,su}$	22.5	kg/s
Superheating	ΔT_{sh}	5	K
Subcooling	ΔT_{sc}	3	K
Expander speed	N_{exp}	3000	rpm
Pump speed	N_{pp}	3000	rpm

in the condenser and evaporator by reducing the number of active plates in a Brazed Plate Heat Exchanger (BPHE). This method effectively represents a decrease in the effective heat transfer area, capturing the impact of thermal resistance buildup caused by NCG accumulation or fouling deposits. Although this approach does not explicitly model gas accumulation dynamics or partial pressure effects, it provides a practical means of assessing performance degradation in the condenser due to reduced heat exchange efficiency and increased back pressure. The number of plates is likewise reduced by 10%, 30%, or 50% to simulate different severity levels.

For the centrifugal pump, the fault is emulated by decreasing the isentropic efficiency, reflecting increased internal losses. The presence of such a fault can be detected through various sensors and operational variables, including the pressure at the pump inlet and outlet, as well as pressure drops and reduced mass flow rate. Additionally, lower heat absorption in the evaporator may be observed. Furthermore, the electrical power input to the pump can indicate inefficiencies, as a faulty pump may require more energy to maintain operation. Similar to the evaporator case, the efficiency is decreased by 10%, 30%, or 50% to simulate different severity levels.

Given the significant impact of mechanical faults on system performance originating from the expander, this study focuses on diagnosing a specific mechanical degradation fault, which is emulated by increasing the mechanical losses. The nominal value is increased by 30%, 50%, or 100%, resulting in a reduction of the isentropic efficiency, leading to higher energy dissipation and decreased mechanical power output. The presence of such faults can be observed through various operational variables, including pressure sensors at the expander inlet and outlet to detect irregular pressure changes and temperature sensors at the exhaust to identify potential overheating. Additionally, variations in power consumption under the same operating conditions may serve as indicators of performance degradation.

In Table 3, the sets of training data collected from the ORC thermodynamic model are presented, including a no-fault dataset and twelve datasets corresponding to different typical faults. These datasets are used to detect whether a fault occurs in the system. The fault detection algorithm is then trained to classify all the possible conditions, not only detecting the presence of a fault but also providing a diagnosis and determining the severity level of the detected fault. Finally, to make the simulation conditions more realistic, different noise levels are introduced into the recorded data at levels of 0.5%, 1%, 2%, and 5%, allowing for further evaluation of the robustness of the different methods against measurement noise.

3 MACHINE LEARNING FAULT DETECTION AND DIAGNOSIS ALGORITHMS

In this section, a brief description of the two machine learning algorithms employed—Support Vector Machine (SVM) and Random Forest (RF)—is provided, highlighting the techniques used during the tuning process.

3.1 Support Vector Machine (SVM)

Support Vector Machines (SVM) are supervised learning algorithms used for classification tasks (Cortes and Vapnik, 1995). The goal is to find an optimal hyperplane that separates data points from different classes while maximizing the margin. Given a training dataset $\{(x_i, y_i)\}_{i=1}^{L}$, where x_i represents feature vectors and y_i are class labels, the decision function is:

$$f(x) = \mathbf{w}^T x + b, \tag{1}$$

Table 3: Simulation fault conditions

Dataset	Simulation condition	Severity level	Class label
No fault	healthy state	-	0
Fault 1	evaporator fouling	10%	1
Fault 2	evaporator fouling	30%	2
Fault 3	evaporator fouling	50%	3
Fault 4	mechanical losses in expander	30%	4
Fault 5	mechanical losses in expander	50%	5
Fault 6	mechanical losses in expander	100%	6
Fault 7	mechanical fault in pump	10%	7
Fault 8	mechanical fault in pump	30%	8
Fault 9	mechanical fault in pump	50%	9
Fault 10	non-condensable gases in condenser	10%	10
Fault 11	non-condensable gases in condenser	30%	11
Fault 12	non-condensable gases in condenser	50%	12

where $\mathbf{w}$ is the weight vector and b is the bias term. The optimization problem is formulated as:

$$\min_{\mathbf{w},b} \frac{1}{2}\|\mathbf{w}\|^2 + C\sum_{i=1}^{L}\xi_i \quad \text{subject to} \quad y_i(\mathbf{w}^T x_i + b) \geq 1 - \xi_i, \quad \xi_i \geq 0, \tag{2}$$

where C is a regularization parameter that controls the trade-off between maximizing the margin and minimizing classification errors, and ξ_i are slack variables allowing misclassifications.

For non-linearly separable data, kernel functions map features into a higher-dimensional space where a linear separation is feasible. The Radial Basis Function (RBF) kernel (Vapnik, 2000)is commonly used:

$$K(x_i, x_j) = \exp(-\gamma\|x_i - x_j\|^2), \tag{3}$$

where γ controls sample influence.

Feature selection enhances classification by reducing dimensionality. The Iterative Relief (IR) method assigns importance scores to features based on their ability to separate classes (Kira and Rendell, 1992) Hyperparameter tuning is performed using k-fold cross-validation, where C and γ are optimized over a logarithmic grid (Bergstra and Bengio, 2012) Performance is assessed using accuracy and the macro-averaged F1-score for multi-class classification:

$$\text{F1-score} = \frac{1}{C_{F1}}\sum_{c}\frac{2P_c R_c}{P_c + R_c}, \tag{4}$$

where C_{F1} is the total number of classes, whilst P_c and R_c denote precision (True Positives divided by the sum of True Positives and False Positives) and recall (True Positives divided by the sum of True Positives and False Negatives) for class c. Since SVM is a binary classifier, multi-class classification is achieved using the One-Versus-One (OVO) strategy (Schölkopf and Smola, 2002) training an SVM for each class pair. The final prediction is based on majority voting, making SVM effective for multi-class problems with high accuracy and computational efficiency.

3.2 Random Forest (RF)

Random Forest (RF) is an ensemble learning method for classification and regression that constructs multiple decision trees and aggregates their predictions to improve accuracy and reduce overfitting (Breiman, 2001) Given a dataset $\{(x_i, y_i)\}_{i=1}^{N}$, where x_i are feature vectors and y_i are class labels, a single decision tree recursively partitions the feature space based on feature thresholds, minimizing impurity using measures such as the Gini index:

$$G = 1 - \sum_{k=1}^{K} p_k^2, \tag{5}$$

where p_k is the proportion of samples belonging to class k in a node. The final classification in RF is obtained through majority voting:

$$\hat{y} = \arg\max_k \sum_{t=1}^{T} I(h_t(x) = k), \tag{6}$$

where $h_t(x)$ is the prediction from tree t and T is the total number of trees.

Bootstrap sampling is a key technique in RF, introducing variability among trees by drawing samples with replacement from the original dataset. Approximately 63.2% of the data appears in each bootstrap sample, while 36.8% remains out-of-bag (OOB), enabling performance estimation.

Hyperparameter tuning is performed to optimize RF performance. Key parameters include the number of trees, maximum depth, minimum samples per split, minimum samples per leaf, and the number of features considered per split. Randomized Search Cross-Validation (Bergstra and Bengio, 2012)is used to efficiently explore the hyperparameter space. A 5-fold cross-validation strategy is applied, where the dataset is divided into five subsets, training occurs on four, and validation on the remaining one, rotating through all folds.

The best model is selected based on the macro-averaged F1-score as in equation (4), ensuring balanced evaluation in multi-class classification tasks. Once the optimal hyperparameters are identified, the best RF classifier is trained for final predictions.

4 PERFORMANCE OF THE FAULT DETECTION AND DIAGNOSIS CLASSIFIERS

4.1 Feature selection

Before analyzing the performance of the fault detection algorithms based on SVM and RF, it is essential to select the features that will be used to classify the different proposed faults. The complete list of candidate features is presented in Table 4, including sensor measurements as well as derived variables obtained from them.

Table 4: Variables used as features for training

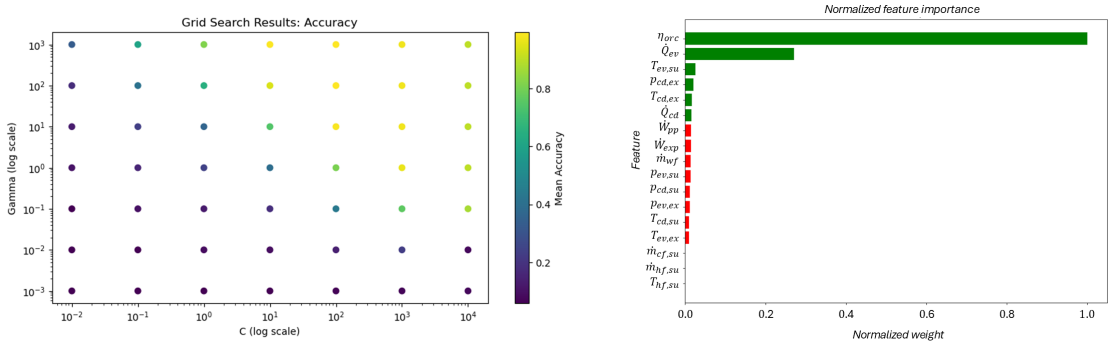
Feature	Name	Feature	Name
$T_{hf,su}$	Supply temperature heat source	$\dot{m}_{hf,su}$	Supply mas flow-rate heat source
$\dot{m}_{cf,su}$	Supply mas flow-rate heat sink	$T_{ev,su}$	Temperature evaporator supply
$p_{ev,su}$	Pressure evaporator supply	$T_{ev,ex}$	Temperature evaporator exhaust
$p_{ev,ex}$	Pressure evaporator exhaust	$T_{cd,su}$	Temperature evaporator supply
$p_{cd,su}$	Pressure condenser supply	$T_{cd,ex}$	Temperature condenser exhaust
$p_{cd,ex}$	Pressure condenser exhaust	$\dot{W}_{exp}$	Expander's electrical power
$\dot{m}_{wf}$	Working fluid mas flow-rate	$\dot{W}_{pp}$	Pump's electrical power
η_{orc}	Cycle efficiency	$\dot{Q}_{ev}$	Heat transfer rate in evaporator
$\dot{Q}_{cd}$	Heat transfer rate in condenser		

Then, from the list of candidates, the iterative Relief method described in Section 3.1 is applied, resulting in the following top six features ranked by importance: η_{orc}, $\dot{Q}_{ev}$, $T_{ev,su}$, $p_{cd,ex}$, $T_{cd,ex}$, and $\dot{Q}_{cd}$, as depicted in Figure 2b. Features with lower relevance or no significant impact were not selected. These selected features are used in all subsequent experiments to establish a consistent basis for comparison.

It is worth noting that the PID controller for superheating may mask pump faults by increasing the pump's rotational speed to maintain the desired superheating level. However, if the increase in speed is caused by a mechanical fault, this compensation leads to higher energy consumption, which ultimately affects the net energy output and, consequently, the ORC efficiency. This behavior aligns with the features identified by the algorithm—particularly the selection of η_{orc}—as it reflects variations in energy consumption due to potential pump faults.

4.2 Experiment 1: Baseline Classification Test (0.5% noise training, 0.5% noise test)

The first experiment represents the baseline for evaluating the classification performance. It is performed using 80% of the data for training and 20% for testing, the data is represented as 150 samples for each class, so the total data for this experiment is $L = 1950$ samples. The level noise is kept low at 0.5%.



(a) Heat map SVM tuning

(b) Feature selection

Figure 2: SVM hyperparameters tuning and feature importance.

First the SVM classifier is tuned using a k-fold cross validation of the hyperparameters C and $gamma$ as described in section 3.1. A heat map of the results are presented in Figure 2a, resulting in parameters $C = 100$ and $\gamma = 10$. For the case of the RF classifier after tuning a number of 500 estimators are employed with maximum depth of 30 and bootstrap.

The results of classification experiment 1, as later summarized in Table 5 indicate that both the Random Forest (RF) and Support Vector Machine (SVM) models exhibit high accuracy in fault detection and diagnosis. The results show that RF achieves perfect classification across all fault classes, while SVM has a misclassification. This minor error suggests that SVM's decision boundary may be slightly less robust in distinguishing certain classes, potentially due to the feature space distribution or kernel function limitations. Despite this, both models demonstrate strong generalization capabilities, given that the experiment was conducted under low noise conditions (0.5%) for both training and test data sets. Thus, this experiment, conducted with low noise and a 'large' training dataset, serves as the baseline and represents the highest performance achievable for each algorithm.

4.3 Experiment 2: Initial Robustness Test (0.5% noise training, 2% noise test)

A subsequent experiment is conducted to assess robustness to noise. A training dataset is first prepared with a noise level of 0.5% and a total length of $L = 1560$ samples, where the algorithms' hyperparameters are tuned to achieve the highest possible accuracy, resulting in parameters $C = 100$ and $\gamma = 100$. For the case of the RF classifier after tuning a number of 500 estimators are employed with maximum depth of 20 and bootstrap. Finally, a second, unseen dataset consisting of $L = 390$ samples with a noise level of 2% is used to evaluate the prediction performance.

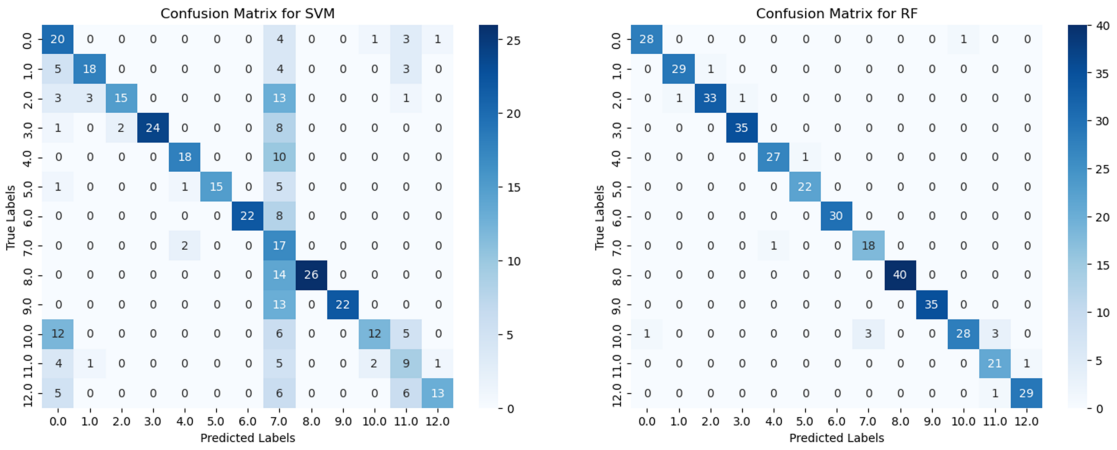
The robustness evaluation under increased noise conditions (2%), summarized in Table 5, shows a decline in accuracy for both classifiers compared to the baseline. RF maintains strong classification performance, with most predictions correctly assigned, though minor misclassifications occur. These small errors indicate that RF is slightly affected by noise but remains highly reliable. In contrast, SVM exhibits a higher misclassification rate, suggesting greater sensitivity to noise disturbances.

A comparative analysis confirms that RF is significantly more robust than SVM, retaining high accuracy with minimal classification errors. SVM struggles under noisy conditions due to its reliance on support vectors, which are more susceptible to distortion. Despite prior hyperparameter tuning, SVM remains less reliable than RF for fault detection and diagnosis in noisy environments. To further validate this finding, future work should evaluate performance under even higher noise levels (e.g., 5%).

4.4 Experiment 3: Extended Robustness Test (0.5% noise training, 5% noise test)

The extended robustness test consists of employing the same classifiers as in Experiment 2, trained with a noise level of 0.5%, but testing their prediction performance on an unseen dataset consisting of $L = 390$ samples with a noise level of 5%. The results are summarized in Table 5 but also depicted in Figure 3 through the confusion matrices for a deeper analysis.

The results of the extended robustness test, as shown in Figure 3, reveal a more pronounced degradation in classification performance due to the increased noise level of 5% in the unseen test dataset. The RF confusion



(a) SVM prediction

(b) RF prediction

Figure 3: Confusion matrix results for experiment with higher noisy data.

matrix demonstrates that the model maintains relatively high accuracy, with most predictions correctly classified. However, compared to Experiment 2, a higher number of misclassifications appear, particularly in classes 2, 7, 10, and 12. These errors indicate that while RF remains robust, its classification boundaries are increasingly affected by the noise, leading to minor prediction errors. Nevertheless, RF still outperforms SVM, suggesting that its ensemble-based nature provides better resilience to noisy data.

On the other hand, SVM experiences a significant drop in performance, as reflected by a substantial increase in off-diagonal elements in the confusion matrix. The model struggles particularly with classes 0, 2, 4, 10, and 12, where a large number of misclassifications occur. This suggests that the decision boundaries of SVM are more susceptible to noise disturbances, leading to greater confusion between similar fault classes. The results confirm that SVM's sensitivity to noisy data makes it less reliable under high-noise conditions, whereas RF retains a higher degree of accuracy. These findings reinforce the conclusion that RF is the preferred model for real-world fault detection and diagnosis applications, where robustness to noise is a critical requirement.

4.5 Experiment 4: Data Reduction Sensitivity Test (2% noise training, dataset reduced by 37%)

A critical aspect for any machine learning algorithm is the amount of available training data. This is particularly important for FDD using labeled data, as obtaining real-life process data for all fault scenarios can be challenging. Therefore, the objective of Experiment 4 is to investigate the sensitivity of the algorithms to data scarcity. In this experiment, the noise level is maintained at 2% for both the training and test datasets. However, the training dataset length is reduced by 37%, meaning that instead of using $L = 1560$ samples, only $L = 982$ samples are available for training.

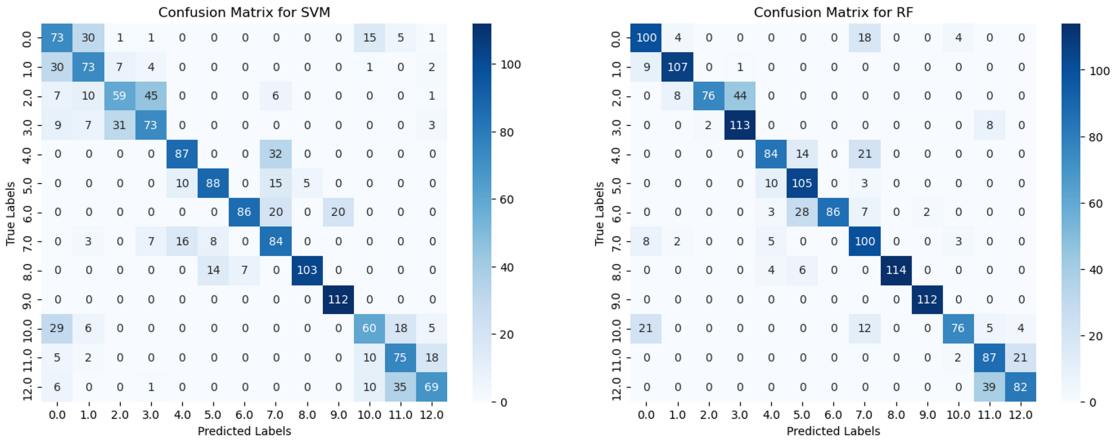
As in all previous cases optimal hyperparameter tuning is performed. For SVM, the best results are obtained with $C = 1000$ and $\gamma = 10$, while for the RF classifier, a total of 500 estimators is used, with bootstrap enabled and no limitation on the maximum tree depth.

The data reduction sensitivity test results, summarized in Table 5, show a noticeable impact on classification performance due to a 37% reduction in training data. Despite this, RF maintains relatively high accuracy, suggesting that while it is affected by the smaller dataset, its ensemble-based approach helps preserve robustness. Errors occur slightly more frequently than in previous experiments, but RF still demonstrates strong generalization capability even with limited data.

In contrast, SVM experiences a more significant performance decline, indicating that its reliance on well-defined decision boundaries makes it more sensitive to data scarcity. This reduced generalization ability causes SVM to struggle more than RF when trained on a smaller dataset. Overall, RF remains the more robust classifier under these conditions, reinforcing its ability to generalize effectively with fewer training samples, whereas SVM is more susceptible to data reduction effects.

4.6 Experiment 5: Extended Data Reduction Sensitivity Test (2% noise training, dataset reduced by 75%)

To further evaluate the impact of data reduction, the training set is reduced by 75% in Experiment 5, leaving only $L = 390$ samples while maintaining a 2% noise level. Optimal hyperparameter tuning for SVM results in $C = 10000$ and $\gamma = 1$, while RF uses 100 estimators with bootstrap enabled and a maximum tree depth of 20. As summarized in Table 5, this extreme reduction leads to a further decline in classification performance.



(a) SVM prediction

(b) RF prediction

Figure 4: Confusion matrix results for Experiment 5 - 75% reduced training data.

As shown in Figure 4, the RF classifier maintains relatively high accuracy but exhibits increased misclassifications, particularly in class 2, which is frequently misclassified as class 3. Higher confusion is also observed in classes 10 and 12, indicating that RF struggles to define clear decision boundaries with limited training data. Despite this, RF retains reasonable robustness, with most predictions along the diagonal, suggesting it can still generalize effectively with only $L = 390$ training samples.

In contrast, SVM suffers a more pronounced decline in classification performance, with a substantial increase in off-diagonal misclassifications. Classes 0, 1, 2, and 3 are particularly affected, with frequent misclassification into neighboring categories. Errors are more widespread compared to RF, highlighting that SVM's decision boundaries become unreliable when trained on a highly limited dataset. These results confirm that RF remains more robust under extreme data scarcity, while SVM's performance deteriorates significantly, making it less suitable for real-world applications with very limited labeled data.

5 RESOURCE UTILIZATION AND COMPUTATIONAL DEMAND

In fault detection and diagnosis (FDD), computation time and memory consumption are critical for deploying machine learning models in real-world applications. While high accuracy is desirable, models must also be computationally efficient to ensure fast inference and low resource utilization, particularly in industrial and IoT environments where hardware constraints exist (Lee et al., 2020) and (Zhang and Wang, 2021). Excessive CPU demand can cause delays in fault detection, impacting system safety, while high memory usage may hinder deployment on embedded devices (García et al., 2019). Thus, balancing accuracy and computational efficiency is essential for practical, real-time FDD applications (Chen et al., 2022). A table summarizing the main outcomes of this investigation relate to computational demand and prediction accuracy are presented in Table 5.

In order to create a more comprehensive visualization of the trade-offs between training time, accuracy, and memory consumption for both SVM and RF classifiers the results are also plotted in Figure 5.

As expected, RF consistently achieves higher accuracy across all experiments but at the cost of increased training time. In contrast, SVM generally requires less computation time but suffers significant accuracy degradation, particularly in experiments with increased noise or reduced training data. The results reveal a key distinction: SVM initially consumes more memory, especially in the baseline classification test (Experiment 1), whereas RF's memory usage increases as experiments become more computationally intensive. This suggests that while SVM has a higher initial memory footprint, RF becomes more demanding as training data decreases or noise levels rise.

A notable trend observed in the plot is that as dataset quality deteriorates (higher noise or fewer training samples),

Table 5: Summary of Classification Performance, Computation Time, and Memory Usage

Test Performed	ErrorPercentage (%)		Tuning +Training Time (s)		MemoryUsage (MB)	
	SVM	RF	SVM	RF	SVM	RF
Exp. 1: Baseline Classification	0.26	0	25.2021	85.6424	35.43	4.85
Exp. 2: Initial Robustness	11.28	1.54	27.5534	105.1941	16.22	16.04
Exp. 3: Extended Robustness	40.77	3.85	25.538	73.97	18.16	25.57
Exp. 4: Data Reduction	5.85	2.36	19.5152	59.3232	6.65	32.58
Exp. 5: Extended Data Reduction	33.21	20.38	11.8326	23.1468	2.82	28.47

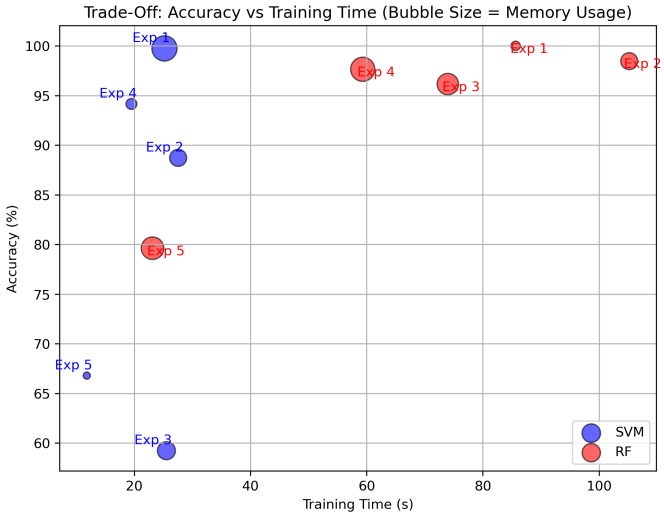


Figure 5: Trade-Off Between Accuracy and Computational Demand

SVM's performance drops dramatically, making it less reliable for fault detection under challenging conditions. Experiment 3 (Extended Robustness Test) highlights this issue, where SVM's accuracy decreases sharply, while RF remains stable with minimal accuracy loss. Similarly, in Experiment 5 (Extended Data Reduction Sensitivity Test), SVM reaches its lowest accuracy despite maintaining a small memory footprint and fast training time. This reflects its difficulty in generalizing from a highly reduced dataset, where reliance on a clear decision boundary becomes a disadvantage. Conversely, RF maintains a better trade-off between accuracy and computational demand, despite requiring more memory and training time.

Overall, the results confirm that RF is the preferable choice for fault detection and diagnosis for the given dataset, particularly in real-world scenarios where noise and data availability constraints are prevalent. While SVM may be suitable for applications requiring quick training with lower memory consumption, its rapid accuracy decline under noisy or limited data conditions makes it less reliable. Therefore, if computational resources are a limiting factor, the choice between RF and SVM should carefully balance accuracy, training efficiency, and memory usage based on the specific application requirements.

6 CONCLUSIONS

In this paper, two machine learning algorithms, namely Support Vector Machine (SVM) and Random Forest (RF), were employed to develop a Fault Detection and Diagnosis (FDD) methodology for Organic Rankine Cycle (ORC) power systems. The key findings are summarized as follows:

- Feature selection was optimized in the initial phase, identifying the six most relevant features for fault classification. This feature set remained constant across all experiments to ensure a fair comparison between SVM and RF while minimizing computational complexity.
- RF consistently outperformed SVM in classification accuracy, demonstrating superior robustness to increased noise levels and reduced training data availability.
- SVM exhibited a significant accuracy decline under challenging conditions, particularly in high-noise or data-scarce scenarios, highlighting its sensitivity to data scarcity and decision boundary distortions.

- While RF required more computational resources in terms of training time and memory usage, it remained the more reliable model for real-world fault detection, making it the preferred choice despite its slightly higher computational cost.
- The trade-off analysis between accuracy and computational efficiency emphasized that model selection should consider not only classification performance but also hardware constraints, particularly in real-time FDD applications where fast inference and optimized memory usage are critical.

Future work will focus on extending this study to scenarios involving unsupervised algorithms, eliminating the need for labeled data and enabling the detection of emerging faults over time, making the approach more adaptable to real-world applications.

NOMENCLATURE

ORC	Organic Rankine Cycle
FDD	Fault Detection and Diagnosis
SVM	Support Vector Machine
RF	Random Forest
NCG	Non-condensable Gases
p	Pressure (Pa)
T	Temperature ($^{\circ}C$)
N	Rotational speed (rpm)
$\dot{Q}$	Heat transfer rate (J/s)
$\dot{m}$	Mass flow rate (kg/s)
η	Efficiency (-)

Subscripts

su	supply
ex	exhaust
hf	hot fluid
cf	cold fluid
ev	evaporator
cd	condenser
exp	expander
pp	pump

REFERENCES

Amalfi, R. L., Vakili-Farahani, F., & Thome, J. R. (2016a). Flow boiling and frictional pressure gradients in plate heat exchangers. part 1: Review and experimental database. *International Journal of Refrigeration*, *61*, 166–184. https://doi.org/10.1016/j.ijrefrig.2015.07.010

Amalfi, R. L., Vakili-Farahani, F., & Thome, J. R. (2016b). Flow boiling and frictional pressure gradients in plate heat exchangers. part 2: Comparison of literature methods to database and new prediction methods. *International Journal of Refrigeration*, *61*, 185–203. https://doi.org/10.1016/j.ijrefrig.2015.07.009

Bergstra, J., & Bengio, Y. (2012). Random search for hyper-parameter optimization. *Journal of Machine Learning Research*, *13*, 281–305.

Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5–32.

Cendoya, A., Ransy, F., Dumont, O., Hernandez, A., Guo, B., & Lemort, V. (2025). Developing a control strategy for carnot batteries: Analysing temperature glide effects in organic rankine cycle and heat pump modes. *Proceedings of the 37th International Conference on Efficiency, cost, optimization, simulation and environmental impact of energy systems*.

Cendoya, A., Ransy, F., Lemort, V., Dumont, O., & Bentao, G. (2024). Numerical modelling of a 50 kwe reversible carnot battery coupled with waste heat. *In proceedings 1st Belgian Symposium of Thermodynamics*.

Chen, X., Li, H., Patel, S., & Kumar, A. (2022). Computational efficiency in real-time fault diagnosis systems. *IEEE Transactions on Industrial Informatics*, *18*(6), 4123–4135. https://doi.org/10.1109/TII.2022.3141234

Choi, S., & Krumdieck, S. (2016). Dynamic modeling of organic rankine cycle (orc) system for fault diagnosis and control system design. *Proceedings 38th New Zealand Geothermal Workshop*.

Cioncolini, A., & Thome, J. R. (2012). Void fraction prediction in annular two-phase flow. *International Journal of Multiphase Flow*, *43*, 72–84. https://doi.org/10.1016/j.ijmultiphaseflow.2012.03.003

Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, *20*(3), 273–297.

Cuevas, C., Lebrun, J., Lemort, V., & Ngendakumana, P. (2009). Development and validation of a condenser three zones model. *Applied Thermal Engineering*, *29*(17–18), 3542–3551. https://doi.org/10.1016/j.applthermaleng.2009.06.007

Dickes, R. (2019). *Charge-sensitive methods for the off-design performance characterization of organic rankine cycle (orc) power systems* [Doctoral dissertation, University of Liege].

Dragan, D. (2011). Fault detection of an industrial heat-exchanger: A model-based approach. *Journal of Mechanical Engineering*, *57*(6), 477–484.

Eck, B. (1973). *Fans: Design and operation of centrifugal, axial-flow and cross-flow fans* (1st English Edition). Pergamon Press.

Elmouatamid, A., Fricke, B., Sun, J., & Pong, P. (2023). Air conditioning systems fault detection and diagnosis-based sensing and data-driven approaches. *Energies*, *16*, 4721.

García, M., Torres, F., & Smith, R. (2019). Machine learning for industrial iot: Challenges and opportunities. *Journal of Process Control*, *85*, 24–36. https://doi.org/10.1016/j.jprocont.2019.10.002

Hosamo, H., Ragnar, P., Svidt, K., Han, D., & Nielsen, H. (2022). A digital twin predictive maintenance framework of air handling units based on automatic fault detection and diagnostics. *Energy and Buildings*, *261*, 111988.

Kira, K., & Rendell, L. A. (1992). The feature selection problem: Traditional methods and a new algorithm. *Proceedings of the Tenth National Conference on Artificial Intelligence (AAAI)*, 129–134.

Lee, J., Kim, B., & Choi, H. (2020). Edge ai for predictive maintenance: A survey. *Engineering Applications of AI*, *95*, 103861. https://doi.org/10.1016/j.engappai.2020.103861

Lemort, V., Declaye, S., & Quoilin, S. (2012). Experimental characterization of a hermetic scroll expander for use in a micro-scale rankine cycle. *Proceedings of the Institution of Mechanical Engineers, Part A: Journal of Power and Energy*, *226*(1), 126–136. https://doi.org/10.1177/0957650911413840

Longo, G. A. (2010). Heat transfer and pressure drop during hfc refrigerant saturated vapour condensation inside a brazed plate heat exchanger. *International Journal of Heat and Mass Transfer*, *53*(5–6), 1079–1087. https://doi.org/10.1016/j.ijheatmasstransfer.2009.11.003

Markowski, M., Trafczynski, M., & Urbaniec, K. (2013). Identification of the influence of fouling on the heat recovery in a network of shell and tube heat exchangers. *Applied Energy*, *102*, 755–764. https://doi.org/10.1016/j.apenergy.2012.08.020

Martin, H. (2010). N6: Pressure drop and heat transfer in plate heat exchangers. In H. Martin (Ed.), *Vdi heat atlas: Chapter*. Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-77877-6

Schölkopf, B., & Smola, A. J. (2002). *Learning with kernels: Support vector machines, regularization, optimization, and beyond*. MIT Press.

Shah, M. M. (2021). Heat transfer during condensation in corrugated plate heat exchangers. *International Journal of Refrigeration*, *127*, 180–193. https://doi.org/10.1016/j.ijrefrig.2021.02.011

van de Sand, R. (2021). *A predictive maintenance model for heterogeneous industrial refrigeration systems* [Doctoral dissertation, Tor Vergata University of Rome, Italy].

Vapnik, V. (2000). *The nature of statistical learning theory*. Springer.

Wang, J., Zuo, Q., Liao, G., Luo, F., Zhao, P., Wu, W., He, Z., & Dai, Y. (2021). Machine learning-based fault detection and diagnosis of organic rankine cycle system for waste-heat recovery. *Journal of Energy Engineering*, *147*(4), 04021016.

Wieland, C., Schifflechner, C., Braimakis, K., Kaufmann, F., Dawo, F., Karellas, S., Besagni, G., & Markides, C. (2023). Innovations for organic rankine cycle power systems: Current trends and future perspectives. *Applied Thermal Engineering*, *225*, 120201.

Yan, K., Ji, Z., & Shen, W. (2017). Online fault detection methods for chillers combining extended kalman filter and recursive one-class svm. *Neurocomputing*, *228*, 205–212.

Zhang, Y., & Wang, T. (2021). Efficient machine learning models for smart manufacturing. *International Journal of Automation Science*, *12*(4), 289–305. https://doi.org/10.1109/IJAS.2021.3156789

ACKNOWLEDGEMENT

This project DECAGONE has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101069740. The information presented on this document reflects only the authors' view. The Agency is not responsible for any use that may be made of the information it contains.