2021 International Conference on Electrical, Communication, and Computer Engineering (ICECCE) | 978-1-6654-3897-1/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICECCE52056.2021.9514086

Robustness analysis of PCA-SVM model used for fault detection in supermarket refrigeration systems*

1st Zahra Soltani
Dept. Electronic Systems
Aalborg University
Aalborg, Denmark
zns@es.aau.dk

2nd Kresten Kjaer Soerensen
Dept. Transport
Bitzer electronics
Soenderborg, Denmark
Kresten.soerensen@bitzerdk.com

3rd John Leth
Dept. Electronic Systems
Aalborg University
Aalborg, Denmark
jjl@es.aau.dk

4th Jan Dimon Bendtsen
Dept. Electronic Systems
Aalborg University
Aalborg, Denmark
dimon@es.aau.dk

***Abstract*—Supermarket refrigeration systems represent an important type of energy demanding appliances, which is in such widespread use that any development in the associated technology can have a huge impact on general health and global warming. Using automatic fault detection and diagnosis may for instance improve energy efficiency and reduce food waste as well as reduce expenses for the supermarket owners. In this paper, three model-free classification algorithms are tested on faulty/non-faulty data obtained from an actual refrigeration system. It is found that support vector machines (SVM) are able to classify fan faults in a real refrigeration system with near-100% classification accuracy, independent of the number of input variables. The classification performance and robustness against an unseen operation mode, low-resolution data, noisy data, and data of different operating points is tested for three different classifier configurations. The results show Principle Component Analysis (PCA)-SVM is highly robust to different operating points, disturbances, and gives the best computational efficiency, as it is able to reduce the feature space to only two dimensions. It is concluded that while all of the examined methods are insensitive to noise, and effective in terms of detecting faults from relatively small amounts of data, overall, PCA-SVM is slightly more computationally efficient.**

***Index Terms*—refrigeration, fault, robustness, classification, support vector, machine learning, dimensionality reduction.**

I. Introduction

Recently, data acquisition and data monitoring have become a part of business competitiveness in many industries, and the availability of data enables manufacturers to have more efficient and reliable systems. Automatic fault detection and diagnosis is one of the ways that ensure more efficient and reliable systems. In refrigeration systems (RS), for food storage, it is crucial to stay within a narrow temperature band; and therefore, it is important to detect faults before they turn into a system breakdown. If the airflow over the evaporator is reduced due to a faulty fan, it will normally not be noticed until the room temperature cannot be kept at its set-point. Traditional fault detection in supermarket refrigeration systems requires many expensive sensors and provide only limited identification of the root cause. Therefore, data-driven Automatic Fault Detection and Diagnosis (AFDD) of such systems is desired. One of the main challenges in designing automatic fault detection systems is that RS controllers, like the ones made by Bitzer Electronics, are used on many different refrigeration systems that exhibit different dynamical behaviour.

Different algorithms have been applied for fault detection and diagnosis (FDD) of refrigeration systems such as [1]–[5]. In [3], the rule-based fault classifier achieved higher effectiveness than data-driven models when the FDD performance index is a controlled variable. The components characteristics and operations anomalies for different types of SRS are studied in [2]. In this study, three sources of the industry including expert surveys, advisory messages such as alarms, and service calls are considered. This information can be used for expanded development of fault detection models. A Convolutional Neural Network was applied for fault detection of refrigeration systems in [4]. This algorithm achieved more than 99 % accuracy in fault classification. The results show that the model can be trained better using low-resolution data. However, as CNN is a deep learning model, it requires high amount of data and computation capacity. A Gaussian mixture model is used in [5], and data dimensions are reduced using Principal Component Analysis (PCA). This model classified four types of faults in Air-conditioning systems with about 99% accuracy, and the running time is reduced more than ten times. One of the binary classifiers that can classify the data based on a low number of samples is SVM. Compared to many types of ANN algorithms SVM has both fast computation and good accuracy. SVM is used in many fields for data classification see [6]–[8], condition forecasting [9], and fault detection [10]. SVM is also used in [11] for fault detection in

This work is partially funded by Innovations fund Denmark and supported by Bitzer electronics A/S, Denmark.

vapor compression refrigeration systems.

In this work, SVM is used to distinguish between a system with evaporator fan fault and a functional system in different operation points while data from evaporation side is not available. It is infeasible to design a bespoke fault detection algorithm for use in every supermarket or to have the technician set it up correctly for each new system and therefore, an automated adaptive fault detection method is required. The topology of the refrigeration systems controlled by the Bitzer condensing unit is generally the same, but they may vary in size and operational set-point. In other words, the challenge addressed in this paper is to design a fault detection algorithm that works effectively for ‘generic’ cooling systems, where the availability of particular combinations of signals cannot be guaranteed. In this paper, we present a method that is robust against the aforementioned types of variations. We show that, through careful selection of the inputs to the classifier, the amount of computation required can be reduced and that PCA can be used as a form of normalization of faulty data acquired at different set-points.

The remainder of this paper is structured as follows. Section II introduces SRS background and fault detection methodology used in this study. SVM classifier and PCA are explained in section III. The models structure and training sensitivity are studied in Section IV. Afterwards, robustness analysis, and comparison of the classifiers are introduced in section V. In Section VI, the results of the work are concluded.

II. Supermarket refrigeration systems

Supermarket refrigeration systems normally use the vapor-compression refrigeration cycle in which heat is moved from a low temperature space to higher temperature ambient air. The heat transferring in the cycle leads to phase change from liquid to vapor and vice versa. Fig. 1 represents an example of SRS which is later used in this paper. Nomenclature related to the figure is described in Table II. SRS might require several controllers for evaporation units and condensing units depending on the supermarket requirements and conditions. In Fig. 1, the system has two controllers that control condensing and evaporation side separately. In SRS, an evaporation unit controller ($Ctrl_{evap}$) controls superheat temperature (T_{sh}) or suction temperature (T_{suc}) to regulate the evaporator performance. A condensing unit controller ($Ctrl_{cond}$) controls cold room temperature by adjusting compressor work.

The evaporator fan plays a key role in transferring heat from the goods to the evaporator surface and consequently the refrigerant. Moreover, it circulates the air in the cold room to ensure an even temperature. An evaporator fan anomaly leads to uneven room temperature, wrong temperature readings by the sensors, higher power consumption, and finally food spoilage. Therefor, early fault detection algorithm is demanded to prevent those consequences. For an AFDD algorithm, data from a condensing unit is required. In this study, data of the normal condition is called *non-faulty*, and data when the fan is defective is called *faulty* data. Variations in the data are necessary to ensure that the developed fault detection model

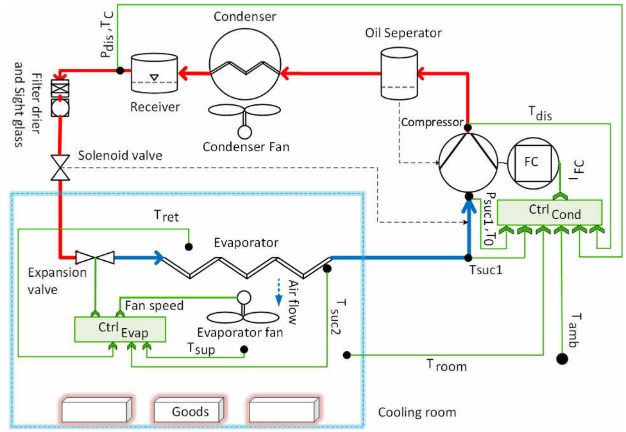

Fig. 1. Schematic of the laboratories’ SRS [4].

TABLE I
Symbols used in the Fig. 1 [4].

Symbols	description	SI unit
T_{room}	cooling room temperature (sensor)	$[°C]$
T_{amb}	ambient temperature (sensor)	$[°C]$
$T_{suc1,2}$	suction temperature (sensor)	$[°C]$
T_0	saturation evaporation (sensor)	$[°C]$
P_{suc}	suction pressure (sensor)	$[Pa]$
T_{dis}	discharge temperature (sensor)	$[°C]$
$Pdis$	discharge pressure (sensor)	$[Pa]$
T_c	saturated condensing temperature	$[°C]$
T_{ret}	returned air temperature (Sensor)	$[°C]$
T_{sup}	supplied air temperature (Sensor)	$[°C]$
I_{FC}	converter current	$[A]$
FC	frequency converter	[-]
$K_{\dot{m}}$	proportional mass flow rate	$[kg/m^3 s]$
$Ctrl_{Evap}$	evaporator controller	[-]
$Ctrl_{Cond}$	condenser controller	[-]

is robust against variations in refrigeration system dynamics. Such variations include evaporators size, air temperature set-points and suction super-heat. In this work, cooling load varied from 6 to 17 kW, the set-point is changed from 1 to 12 °C, and as a consequence, the compressor speed varied from 33 to 80 Hz. As seen in Fig 1, the laboratory set up has two evaporator fans. Fan fault is emulated in the laboratory set up so as one out of two evaporator fans is defective. In all data sets 14 measurements are logged from $Ctrl_{cond}$ which are relevant to show the system characteristics. The collected data is fed into the AFDD algorithm, which is described below.

A. Fault detection methodology

In this paper, three different fault classifiers are presented: SVM classifier using all available signals of relevance to the system characteristics, SVM classification using signals selected by experts based on system knowledge, and PCA-SVM classification in which PCA is used for feature extraction. Making a single fault detection algorithm that is capable of handling system variations requires that the features that are most important for detecting the fault are extracted from the data and normalized before being passed to the classifier. In this paper, the feature extraction or signal selection are tested both manually and automatically.

As for the second methodology, the most relevant signals are selected manually by experts. Reducing the number of inputs can be effective both for recording the data and also for classification computation. However, it is of course detrimental if any information-rich signals are removed. In this work, the most relevant signals that represent the system characteristics during fan fault detection are: P_{suc}, T_{sh}, compressor speed (V_{cpr}), and proportional mass flow rate ($K_{\dot{m}}$). Even though this methodology is computationally more efficient than the one that used all signals, it is vastly dependent on experts knowledge.

The last methodology is feeding all available signals to the PCA algorithm which extracts the most important information of the data and reduces its dimensions. Afterward, the reduced-order data is classified by SVM. This methodology should enable a simpler classifier to distinguish between faulty and non-faulty data from a range of systems with varying characteristics.

III. Methods

A. SVM classifier

Support Vector Machines is a type of supervised learning method used for classification purposes. Let a set of data be given as $\mathcal{D} = \{(x_i, y_i) | x_i \in \mathbb{R}^k, y_i \in \{-1, 1\}\}_{i=1}^n$ where k is the dimension of the samples x_i, y_i is the corresponding class and n is the number of the samples. If $\mathcal{D}$ is linearly separable, it is possible to select two parallel hyperplanes that separate the two classes of data, such that the distance between them is as large as possible. The region bounded by these two hyperplanes is called the *margin*, and the maximum-margin hyperplane is the hyperplane that lies halfway between them, as illustrated in Fig 2. A hyperplane h is the set of points $x \in \mathbb{R}^k$ satisfying an equation of the form

$$w \cdot x + b = 0 \tag{1}$$

where $\cdot$ is the standard vector dot product, $w \in \mathbb{R}^k$ (a.k.a. *weights*) is orthogonal to h, and $b \in \mathbb{R}^k$ is an offset (a.k.a. *bias*). Notice that the two classes are separated by two parallel hyperplanes h_1 and h_2 defined by

$$\begin{aligned} h_1 &: \quad w \cdot x + b = 1, \\ h_2 &: \quad w \cdot x + b = -1. \end{aligned}$$

Since h_1 and h_2 are parallel, they share the same w, and the distance between them is $2/\|w\|$. The distance between h_1 and h_2 is thus maximized by solving the following constrained optimization problem [12]:

$$\begin{aligned} &\min \; f(w) = \|w\| \\ &\text{s.t.} \\ &\begin{cases} w \cdot x_i + b \geq 1 & \text{if } y_i = 1 \\ w \cdot x_i + b \leq -1 & \text{if } y_i = -1 \end{cases} \end{aligned} \tag{2}$$

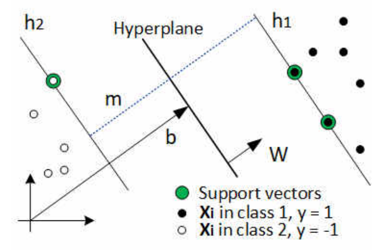

Fig. 2. Maximum margin (m) and optimal hyperplane.

In SVM, the hyperplanes are typically represented via *perceptrons* parameterized by the weight and bias vectors w as

$$H(x_i) = \text{sign}(w \cdot x_i + b). \tag{3}$$

The parameters are adjusted using an update rule in order to find the correct classification for each sample. Here, the technique of Lagrange multipliers is used, allowing the minimization problem (2) to be rewritten as:

$$\begin{aligned} &\min W(\alpha) = \Sigma_{i=1}^n \alpha_i - \frac{1}{2} \sum_i^n \sum_j^n \alpha_i \alpha_j y_i y_j K(x_i, x_j) \\ &\text{s.t.} \\ &0 \leq \alpha_i \leq C \\ &\sum \alpha_i y_i = 0 \end{aligned} \tag{4}$$

where $\alpha_i \geq 0, i = 1, \ldots, n$, are Lagrange multipliers, the constant C is a bound on the multipliers determining how SVM deals with classification errors, and $K : \mathbb{R}^k \times \mathbb{R}^k \to \mathbb{R}$ is a so-called *Kernel function*. The parameter C is a trade-off between narrow and wide margins. Selecting a too small value for C (corresponding to a wide margin) might result in hyperplanes that can not classify a validation data set, whereas hyperplanes resulting from too large C (narrow margin) might not handle noisy outliers well.

In some cases, the data set might not be linearly separable in its original representation. Then, it is often possible to transform the data by a kernel function such that the classes become separable in the transformed representation. Several types of kernel functions, such as polynomials, Radial Basis Functions (RBF), etc. may be used. A popular choice is the RBF kernel function, described by

$$K(x_i, x_j) = e^{-\gamma \|x_i - x_j\|^2} \tag{5}$$

where the bandwidth parameter γ is inverse of the variance of standardized samples which scales the distance between two samples. There are a number of heuristics to determine C and γ as hyperparameters of the optimization, which can be found in [12].

B. PCA

It is often the case in practice that some of the features are correlated with the others, thus providing less useful information for the classification. Principal Component Analysis (PCA) is a method that analyzes high dimensional data and identifies correlations among the data entries (features). PCA then projects the data down to a lower dimensional representation in which important relations between features, and other relevant information of the data set are preserved, but, unimportant information is discarded. The basis of this new representation, called *principal components*, is orthogonal by construction, as it is the span of eigenvectors of the covariance matrix of $\mathcal{D}$. The main advantage of PCA in this particular application is that it removes correlated features that do not make any contribution to the classification.

Correlation among the parameters can be identified by computing the covariance matrix $R_{xx}(\mathcal{D}) \in \mathbb{R}^k \times \mathbb{R}^k$. From $R_{xx}(\mathcal{D})$, we compute the k eigenvectors ν needed for projecting the k-dimensional samples onto the subspace spanned by the principal components. The eigenvectors are sorted by descending eigenvalues, and only the eigenvectors corresponding to the largest $m < k$ eigenvalues are used for the projection. Finally, a new data set $\mathcal{D}_{\text{pca}}$ can be obtained from the original data by computing

$$\tilde{x}_i = V(V^T V)^{-1} V^T x_{s_i}, \quad i = 1, \ldots, n \tag{6}$$

with $V = [\nu_1 \ \cdots \ \nu_m] \in \mathbb{R}^{k \times m}$, yielding $\tilde{x}_i \in \text{span}\{\nu_1, \ldots, \nu_m\}$ for all $i = 1, \ldots, n$.

IV. Model Training

As outlined above, first, an SVM classifier is designed for 14D and 4D input data. Then, the PCA-SVM algorithm is presented and compared with the two other proposed classifiers. The sensitivity of the classifiers during training phase is investigated against different sample rates and different data lengths (number of samples).

A. Training and Validation

In this work, data is classified into two categories of faulty and non-faulty. Fourteen measurements are logged from the condensing unit controller and fed into SVM when all information of the system is used. In this data set, neither human nor an algorithm selects the relevant signals. The SVM classifier used RBF kernel function with optimized hyperparameters of $C = 10$ and $\gamma = 1$. The result of SVM using 14D data represents 98% accuracy for training data and 100% accuracy for validation data classification. Afterwards, four of the aforementioned measurements are selected and supplied into a SVM algorithm. 98% classification accuracy in the training phase, and 100% classification accuracy in the validation phase is obtained. In this case, RBF kernel function is used with optimized hyperparameters, $C = 100$ and $\gamma = 1$.

As for the third algorithm, PCA is used to obtain the most correlated features. The scree plot in Fig. 3 illustrates the variation that each principal component accounts for in percentage. Therefore, the first two principal components, which has the most variance, is selected. Fig. 4, represents the training data classification using PCA-SVM classifier. The contour maps shows the choice of the decision boundary between the two classes of data using Radial basis kernel function (RBF), $\gamma = 1$, and $C = 100$. Fig. 4, represents the training data

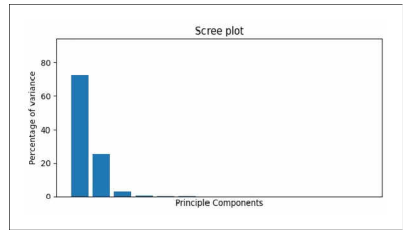

Fig. 3. The percentage of variation that each Principal component accounts for

classification using PCA-SVM classifier. The contour maps shows the choice of the decision boundary between the two classes of data using RBF kernel function, $\gamma = 1$, and $C = 100$. The faulty data is bounded by yellow surface and

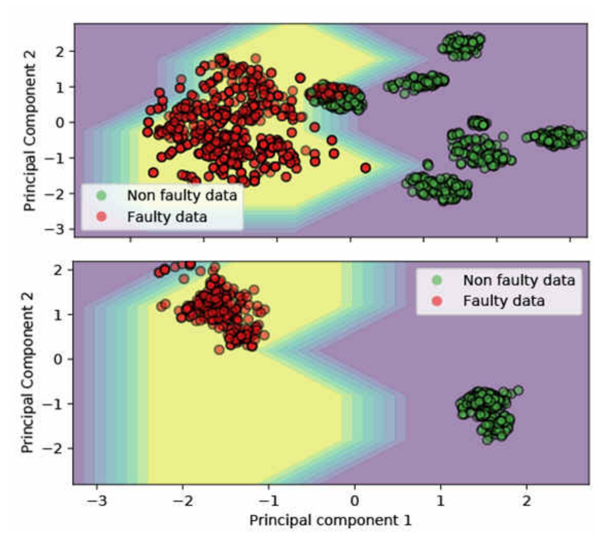

Fig. 4. The top plot, training result of PCA-SVM model. The bottom plot, the validation result of PCA-SVM.

the boundary become loser until a surface that belongs to the non-faulty data indicated by purple. Here, the data could have been classified with looser boundaries or more hyperplanes. However, the more restricted margin is selected due to two reasons. First, as seen in Fig. 4, in the top, the non-faulty data is more varied and distributed differently. Therefore, misclassification of unseen non-faulty data is avoided by more restriction for faulty data. Secondly, the smallest false positive rate is ensured, which is desired in the industry. In the bottom plot of Fig. 4, the validation data set is transformed into the principal components of the training data, which causes the different positions of the validation data compared to the training data. The classifier detects the training non-faulty data with 98% accuracy and faulty data with 97% accuracy. The validation data is classified with 100% accuracy for both faulty and non-faulty data. The validation result was more accurate than the training due to the distribution and overlap of the training data, which are less prominent in the validation data. The distribution of the training data is due to using data of different operation conditions while the validation set is taken only from one operating condition.

B. Training data sensitivity

In SVM, the number of samples to be used depends on the number of input measurements, meaning that if higher dimensional data is selected, the data set should be increased as well to achieve better performance. However, as the computation efficiency is important in this work, further tests with longer data set is ignored. Instead, 4D data with different lengths are tested for the SVM training.

In each test, training data with different sample rates are proposed. Training data is down-sampled from 1 Hz to 0.3, 0.1, 0.03, and 0.01 Hz. Here, it is not possible to analyze lower sample rates than 0.01 Hz, due to the limited data length. Table II represents the accuracy of the SVM classification to the various training data. In this analysis a specific validation data set is used which has different operating condition than training data. Remark, in this table, the length of the data is the number of samples of each class.

TABLE II
TRAINING DATA LENGTH AND RESOLUTION ANALYSIS

Length	Sample rate [Hz]	Training time (s)	Accuracy [%]
1800	1	0.57	93
	0.1	0.65	93
	0.01	0.63	93
900	1	0.09	99
	0.1	0.09	99
	0.01	0.1	99
300	1	0.07	94
	0.1	0.08	94
	0.01	0.07	94

From Table II, it can be recognized that different sample rates do not have effect neither on accuracy nor running time whereas data length has a considerable effect on both the accuracy and running time. It is found that the best training data length is about 900 samples for each class of data. The number of samples need to be sufficient enough to cover all information of the data. Therefore, an insufficient number of data leads to misclassification. Moreover, the classifier needs to handle too many outliers if it receives too large number of samples. In addition, by doubling the number of samples the training time increased about 60%.

V. ROBUSTNESS ANALYSIS

As mentioned in Section I, it is necessary to have an algorithm that is robust to different system configurations and operating conditions. Therefore, robustness tests for SVM and PCA-SVM classifiers are done. Note that in this section, data length is 900 samples and sample rate is 1 Hz for the training set.

A. Validation data resolution

In this test the training and validation data have different sample rates. Validation data with 1 Hz sample rate for all three classifiers obtained the same accuracy about 100%. Then, the validation data is down-sampled from 1 Hz to 0.1 and 0.01 Hz. However, the classification accuracy remains the same as using original data. The results shows that SVM is a robust classifier against data resolution as the same results are obtained after down-sampling of the original validation data. This test illustrates the validation data is accurately classifiable independent of the data resolution. A classifier trained in a specific sample rate can be used to classify the fault in a variety of RS with different sample rates.

B. System variations

To investigate robustness towards RS variations, the validation data was changed by adding noise, static perturbations (offset), and an operational disturbance as seen as On/Off operation of the compressor in RS. Every type of test is done 20 times to ensure the results. Table III, illustrates the classification results of system variations tests. The changes to the data was exacerbated compared to data from the field to ensure that the classifiers can handle a wide range of refrigeration systems. The noise is random with normal distribution $\mathcal{N}(0, 2)$ and values ranged [-4, 4] °C. As shown on Fig. 5, when adding noise to the data, non-faulty and faulty data overlap in some of the measurements and become harder to separate. Different

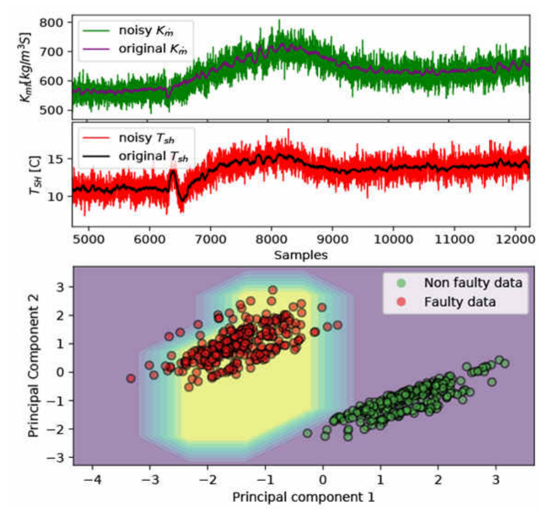

Fig. 5. The top plot , noisy data example (fault occurs at sample no. 6300); at the bottom, Noisy data classification using PCA-SVM.

system configurations and operating conditions in SRS can be considered as perturbations of the data assuming all or some of correlations between the measurements are preserved. On Fig. 6, the classifier's results for perturbed data is shown. Here, random offset of the superheat temperature in the range [-5 to -2]°C, and [2 to 12]°C is applied. In fact, perturbation might not make a huge impact on the result when using PCA as long as correlation of the data does not change. Comparing the classifiers results in Table III, the PCA-SVM classifier is more robust against perturbed data or different operation conditions. In SRS, when the temperature of the goods are on set-point, low cooling capacity is required to keep the goods at the same temperature. Thus, the SRS operation mode may alternate between stopped and running modes. A slow and periodic disturbance has been added to the data to simulate On/Off mode of operation for the compressor seen on Fig. 7. Table III represents better classification using PCA-SVM than two other classifiers in the on/ off mode.

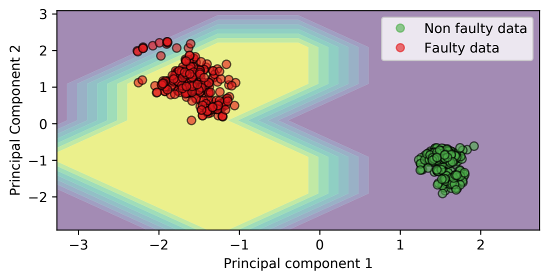

Fig. 6. Validation of perturbation test using PCA-SVM classifier.

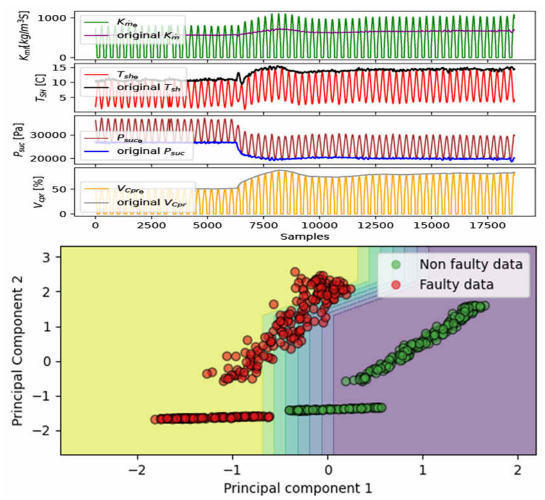

Fig. 7. The top plot, an example of data when disturbance is added; the bottom plot, validation of disturbance test using PCA-SVM.

TABLE III
THE CLASSIFIERS ROBUSTNESS TESTS.

	Algorithm	Non faulty[%]	Faulty[%]	Run time[s]
Noisy	14D SVM	98.5 -99.6	98 -99.4	0.31
	4D SVM	98 -100	98 -99.4	0.24
	PCA-SVM	98 -100	98 -99.6	0.25
Perturbed	14D SVM	89-100	97-100	0.32
	4D SVM	99.2-100	99-100	0.24
	PCA-SVM	100	100	0.23
On/Off	14D SVM	50-60	53-60.5	0.33
	4D SVM	55-60	54-61	0.25
	PCA-SVM	85-86	95.5-96.4	0.25

VI. Conclusion

In this study, it was shown that a SVM classifier can identify a fault in evaporation side using data from condensing unit with high accuracy, in both training and validation process, independent of the data resolution. It was shown that it is possible to do fault detection on refrigeration systems using Machine learning with lower amount of expert effort which is expensive and time consuming. Three models are proposed to classify the data using SVM classifiers. The difference among these classifiers are their inputs which were raw data from the controller for the first (14D) model, the most relevant measurements for the second (4D) model, and PCA transformed data for the third model. The classifiers are highly robust to different data sample rates as long as the dynamics of the system is preserved. PCA-SVM can overcome the significant difficulties that unseen data introduces for the classifiers such as noise, perturbation, disturbance and different running modes. PCA-SVM is more robust against system variations and about 25% more computationally efficient than SVM without dimension reduction.

Another advantage of the PCA-SVM algorithm is that it can be separated into two parts; a PCA algorithm, and an SVM algorithm. PCA can be processed in the controller hardware, and the transformed data with low dimensions can be sent to the third party for the fault classification. Therefore, PCA-SVM can be considered as the most accurate and cost-effective classifier among those three proposed classifiers.

References

[1] S. Srinivasan, A. Vasan, V. Sarangan, and A. Sivasubramaniam, "Bugs in the freezer: Detecting faults in supermarket refrigeration systems using energy signals," in *Proceedings of the 2015 ACM Sixth International Conference on future energy systems*, ser. e-Energy '15. ACM, 2015, pp. 101–110.

[2] A. Behfar, D. Yuill, and Y. Yu, "Supermarket system characteristics and operating faults (rp-1615)." *Science & Technology for the Built Environment*, vol. 24, no. 10, pp. 1104 – 1113, 2018.

[3] A. Behfar, Y. David, and Y. Yuebin, "Automated fault detection and diagnosis for supermarkets–method selection, replication, and applicability," *Energy and Buildings*, vol. 198, 2019.

[4] Z. Soltani, K. K. Soerensen, J. Leth, and J. D. Bendtsen, "Fault detection of supermarket refrigeration systems using convolutional neural network," in *IECON 2020 The 46th Annual Conference of the IEEE Industrial Electronics Society*, 2020.

[5] Y. Guo and H. Chen, "Fault diagnosis of vrf air-conditioning system based on improved gaussian mixture model with pca approach," *International Journal of Refrigeration*, 2020.

[6] C. Zhou, J. G. Chase, and G. W. Rodgers, "Support vector machines for automated modelling of nonlinear structures using health monitoring results," *Mechanical Systems and Signal Processing*, vol. 149, p. 107201, 2021.

[7] J. Wang, L. Shen, Y. Bi, and J. Lei, "Modeling and optimization of a light-duty diesel engine at high altitude with a support vector machine and a genetic algorithm," *Fuel*, vol. 285, 2021.

[8] A. Skariah, P. R, R. R, and B. C. R, "Health monitoring of rolling element bearings using improved wavelet cross spectrum technique and support vector machines," *Tribology International*, vol. 154, p. 106650, 2021.

[9] S. Rafiee-Taghanaki, M. Arabloo, A. Chamkalani, M. Amani, M. H. Zargari, and M. R. Adelzadeh, "Implementation of svm framework to estimate pvt properties of reservoir oil," *Fluid Phase Equilibria*, vol. 346, pp. 25 – 32, 2013.

[10] J. Liu, Y.-F. Li, and E. Zio, "A svm framework for fault detection of the braking system in a high speed train," *Mechanical Systems and Signal Processing*, vol. 87, pp. 401 – 409, 2017.

[11] H. Han, Z. Cao, B. Gu, and N. Ren, "Pca-svm-based automated fault detection and diagnosis (afdd) for vapor-compression refrigeration systems," *HVAC&R Research*, vol. 16, 2010.

[12] M. Jordan, J. Kleinberg, and B. Schölkopf, *Support Vector Machines*, 1st ed., ser. Information Science and Statistics. New York, NY: Springer-Verlag, 2008.