Preprints of the 19th World Congress
The International Federation of Automatic Control
Cape Town, South Africa. August 24-29, 2014

A One-Class SVM Based Tool for Machine Learning Novelty Detection in HVAC Chiller Systems

A. Beghi *, **L. Cecchinato** **, **C. Corazzol** *, **M. Rampazzo** *, **F. Simmini** *, **G.A. Susto** *

* *Department of Information Engineering, University of Padova, via G. Gradenigo 6/B, I-35131 Padova, Italy.*
e-mail: {beghi, corazzol, rampazzom, simminif, sustogia}@dei.unipd.it.
** *Everest s.r.l., via Roveredo 20, I-33170 Pordenone, Italy.*
e-mail: lcecchinato@energyeverest.com

Abstract:
Faulty operations of Heating, Ventilation and Air Conditioning (HVAC) chiller systems can lead to discomfort for the occupants, energy wastage, unreliability and shorter equipment life. Such faults need to be detected early to prevent further escalation and energy losses. Commonly, data regarding unforeseen phenomena and abnormalities are rare or are not available at the moment for HVAC installations: for this reason in this paper an unsupervised One-Class SVM classifier employed as a novelty detection system to identify unknown status and possible faults is presented. The approach, that exploits Principal Component Analysis to accent novelties w.r.t. normal operations variability, has been tested on a HVAC literature dataset.

Keywords: Classification, Data reduction, Detection systems, Failure detection, Fault identification, Machine learning, Performance monitoring

1. INTRODUCTION

Operating problems associated with degraded equipment, failed sensors, improper installation, poor maintenance, and improperly implemented controls plague many Heating, Ventilation and Air Conditioning (HVAC) systems; these factors lead to inefficient operations (increased energy costs), discomfort, and increased wear of components (reduced reliability and shorter equipment life). *Fault Detection (FD)* systems are nowadays playing a fundamental role in monitoring complex HVAC systems, detecting anomalous behaviours in such way to keep the systems in their best operational condition with minimum costs: when unexpected anomalous events occur in HVAC systems, their consequences can be quite damaging and costly (in terms of money and time) to be treated.

Despite their use in other fields (i.e. automotive [Sievers and Mortonson, 1983], aerospace [Henry et al., 2010], manufacturing [Susto and Beghi, 2013], etc.), developing efficient FD systems for HVAC installations remains as a challenge due to the general unavailability of *labelled data*; labelled data contain qualitative information related to the functioning condition of the system (normal or anomalous for example) and they are usually:

- costly because labelling is done manually by a human expert;
- unfeasible because anomalous data instances are never seen in most of the modern HVAC plants, where none, or more generally, few of the potential anomalies already happened in the past.

For these reasons in this field we are dealing with *novelty detection* problems, that aim to monitor the behaviour of the system and identify if shifts from the nominal working conditions arise [Chandola et al., 2009].

In this paper we propose a method to detect anomalies in order to smartly monitor system operating conditions and predict in advance potential faults with application to vapour-compression chillers. In particular, a centrifugal chiller (which is identified by the type of employed compressor) is considered; centrifugal chillers are popular choices for facilities with medium and large cooling loads, [Stanford, 2012]: this kind of cooling machines are variable volume displacement units that use rotating impellers to compress the refrigerant vapour and the cooling capacity is regulated through the use of inlet vanes to restrict the flow of refrigerant to the impeller. It is worth noticing, that chiller component faults account for around 42% of the service resources and around 26% of repair costs [Comstock and Braun, 1999]: it goes without saying that maintaining HVAC chiller systems in healthy conditions and identifying anomalies in time is beneficial to both energy and operating costs savings. In this context, we consider anomalies as faults, a priori unknown, such as reduced evaporator or condenser water flows, refrigerant leakage or overcharge, condenser fouling, etc.

The main issue in HVAC anomaly detection is that most of the variability in the data is due to the usual functioning of the system (see the example depicted in Fig. 1); this unfortunately masks the appearing of anomalies and makes their detection complicated, since the variability is

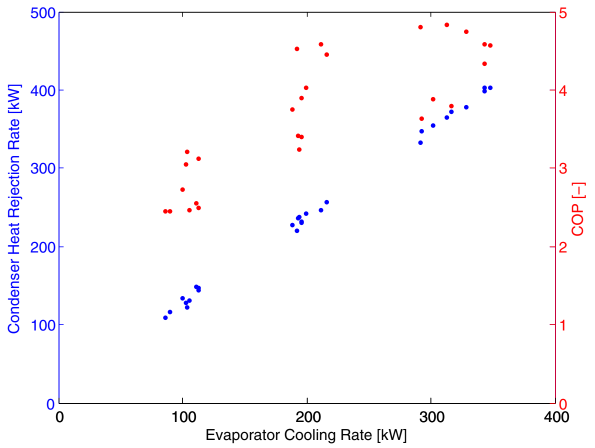

Fig. 1. Chiller operating points at normal (i.e. fault-free) conditions. The Coefficient of Performance (COP) is defined as the ratio of heat removal from the evaporator over energy input to the compressor. ASHRAE 1043-RP project.

dominated by the typical changes from different operating conditions. To deal with this issue, the proposed anomaly detection system, that is based on a kernelized One-Class Support Vector Machine (OCSVM) classifier, is assisted by Principal Components Analysis (PCA) to help the discrimination between normal operating and anomalies-related variability.

In order to test and validate the proposed anomaly detection technique, data-sets from the ASHRAE (American Society of Heating, Refrigerating and Air Conditioning Engineers) research project 1043-RP have been used [Comstock and Braun, 1999]; the project conducted experimental studies to produce both fault-free and faulty data sets on a centrifugal water-cooled chiller to develop and evaluate fault detection and diagnosis methods. As final result of our experiments we provide information about the severity of detected anomalies by exploiting the distance of new data from the OCSVM decision boundary as health factor indicator of potential faults.

The paper is organized as follows: Section 2 is dedicated to review Novelty Detection techniques and introduce One-Class SVMs, while Section 3 deals with PCA; Section 4 illustrates the centrifugal chiller data-sets employed in our experiments. The model corresponding to nominal behaviour is derived in Section 5, whereas the experimental part is detailed in Section 6. Some concluding remarks are given in Section 7.

2. NOVELTY DETECTION

Novelty detection is the identification of new or unknown data/situations that a machine learning system is not aware of during the training phase. In this domain, statistical classification tools [Hastie et al., 2001] that are constructed and developed for discriminate between different classes of labelled data, are employed as detectors, i.e. identifiers of "unfamiliar" cases. The statistical detection of novelties is fundamental in applications such as Fault Detection and Statistical Process Control and it is becoming a key issue as engineering systems are becoming more and more sophisticated (and more "data-rich") and downtime due to unexpected faults should generally be minimized [Susto et al., 2012]. Novelty detection belongs to the class of *unsupervised* or *semi-supervised* problems, an extremely challenging area: working with unlabelled data requires, not only appropriated statistical tools, but also experience and attention in describing the nature of the data without overcomplicating the model.

Several statistical approaches have been employed for novelty detection, amongst them: Gaussian Mixture Models, Hidden Markov Models, Hypothesis Testing; however, all the aforementioned methods assume that data distributions are Gaussian in nature: this, as in the modelling of HVAC systems, can be a strict assumption.

Other approaches that do not impose restrictions on data distribution are based on: k-Nearest Neighbour, Statistical Clustering, One-Class Support Vector Machines.

In the work detailed in this paper we have chosen One-Class Support Vector Machines (OCSVMs) to deal with novelty detection for certain of the positive aspects of this tool, mainly the capability of dealing with high-dimensional dataset, but also the possibility of providing *kernelized* solutions [1] [Schölkopf et al., 1999], that results into having:

- non-linear decision boundaries (complex classification rules available);
- convex optimality problems;
- flexibility in the solutions (hyperparameters available to adapt the solution to the data type).

In the next subsections we introduce OCSVMs and the related tuning issues.

2.1 One-Class Support Vector Machines

As stated above, One-class classification tries to identify objects of a specific class amongst all objects, by learning from a training set containing only the objects of that class. OCSVM may be viewed as a standard two-class SVM [Hastie et al., 2001], where all the training data lie in the first class, and only the origin is taken as member of the second class.

As introduced previously, SVMs have the good quality of being able to provide non-linear classification through the *kernel trick* [Schölkopf et al., 1999]: given a training dataset $X = \{\mathbf{x}_1, .., \mathbf{x}_n\}$, $\mathbf{x}_i \in \mathbb{R}^d$, the OCSVMs algorithm maps the data into a higher dimensional feature space and finds a hyperplane to separate all the data objects from the origin with maximum margin, by solving the following quadratic programming problem:

$$\arg\min_{\mathbf{w},\xi,\rho} \frac{1}{2}\|\mathbf{w}\|^2 + \frac{1}{\nu n}\sum_{i=1}^{n} \xi_i - \rho, \tag{1}$$

$$\text{subject to} \quad \begin{cases} \langle \mathbf{w}, \Phi(\mathbf{x}_i)\rangle \geq \rho - \xi_i \\ \xi_i \geq 0 \end{cases}, \tag{2}$$

where n is the number of training samples, $\xi = [\xi_1 \dots \xi_n]$ and $\Phi(\cdot)$ is a *kernel* function, a, generally non-linear, feature map. In eq. (1), $\mathbf{w}$ represents the normal vector and ρ is the offset of the desired hyperplane in the

[1] In the following we will implicitly consider the kernelized version of OCSVMs.

feature space. The slack variable ξ_i measures the degree of misclassification of the data. The trade-off parameter $\nu \in (0, 1]$ is an upper bound on the fraction of training samples outside the decision boundaries and a lower bound on the fraction of support vectors (i.e. the data points that cannot be discarded in simplifying the SVM solution).

In the SVM solutions (see Hastie et al. [2001] for details) the kernel functions appear through the inner product:

$$k(\mathbf{x}_i, \mathbf{x}_j) = \langle \Phi(\mathbf{x}_i), \Phi(\mathbf{x}_j) \rangle . \tag{3}$$

A smart choice of the kernel function allows to avoid the explicit mapping $\Phi(\cdot)$, computing just the inner product. Common choices for the inner product $k(\mathbf{x}_i, \mathbf{x}_j)$ are:

Radial Basis (RBF): $\exp[-\|\mathbf{x}_i - \mathbf{x}_j\|^2/(2\sigma^2)]$;
Polynomial: $(1 + \langle \mathbf{x}_i, \mathbf{x}_j \rangle)^p$;
Neural Network: $\tanh(a \langle \mathbf{x}_i, \mathbf{x}_j \rangle + b)$.

In this work we have employed RBF functions and in the following we will refer to RBF OCSVMs.

2.2 RBF OCSVM Tuning

It is worth noticing that the width parameter σ of Gaussian kernel plays a key role in classification problems, because it has a significant impact on accuracy and generalization performance. As σ increases, the number of support vectors decreases and the decision boundaries become looser. In addition, the parameter ν also affects the shape of the decision boundaries: as ν increases, the number of support vectors increases and the number of misclassified training samples grows (Hastie et al. [2001]).

Since ν is closely related to the fraction of training samples outside the decision boundaries, it is usually set to a small value to ensure a small misclassification rate on the training phase: therefore, choosing an appropriate value for σ is the main challenge of building a satisfactory OCSVM.

In the considered semi-supervised context, the heuristic approach proposed by Wang et al. [2012], which chooses σ via tightness detecting, is employed. Based on the assumption that training samples are representative, an ideal decision boundary of OCSVMs should be neither tight to ensure the generalization of classifiers, nor loose to ensure the sensitivity to outliers. Since the relationship between the tightness of boundaries and σ is monotonous, an iterative algorithm is used to choose the width of Gaussian kernel via tightness detecting so that an appropriate tightness of decision boundaries is guaranteed.

3. PRINCIPAL COMPONENT ANALYSIS

A classification problem becomes significantly harder as the dimensionality of the data increases. Sometimes data are sparse in the space they occupy leading to difficulties for the unsupervised learning; in the literature, this phenomenon is referred to the *curse of dimensionality.* A high data dimensionality is a problem for many classification algorithms given the consequent high computational cost and memory usage. Moreover, huge dimensionality in the data can lead to poor understanding of the describing model [Hastie et al., 2001].

Principal Component Analysis (PCA) can be employed for dimensionality reduction. PCA is a linear projection-based method that transforms a set of variables into a new set of uncorrelated variables, named *Principal Components (PCs)*. PCA is run for a dataset defined by an $n \times d$ design matrix X where the d columns are variables and the n rows are observations. X is written in terms of the $n \times l$ scores matrix T, where $l \le d$, and the $d \times l$ loadings matrix P, plus a residual matrix E, as follows:

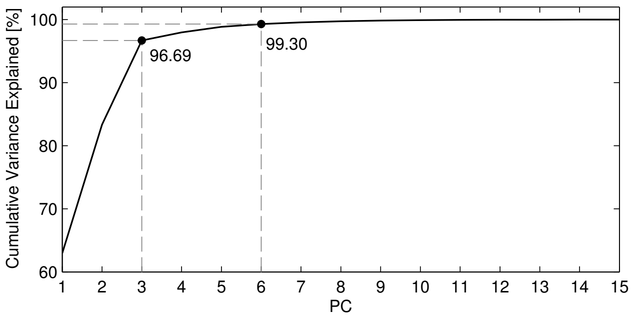

Fig. 2. Cumulative variance explained at the increase of the considered PCs.

$$X = TP^T + E = \sum_{i=1}^{l} \mathbf{t}_i \mathbf{p}_i^T + E, \tag{4}$$

where $\mathbf{t}_i = X\mathbf{p}_i$. The vectors $\{\mathbf{p}_i\}$ are the PCs and if $l = d$, then $E = 0$. PCs are arranged in order of magnitude variability of X explained: the first PC, $\mathbf{p}_1$, can be geometrically interpreted as the direction where most of the variability lies, then other PCs define orthogonal directions where less and less variability is contained. PCA is called also Eigenvalue Decomposition, as each of the PCs is related to an eigenvalue of the matrix in exam, ordered in terms of magnitude.

The transformation induced by PCA can therefore be employed for reducing the dimensionality of the problem at hand, as just $l < d$ variables can be employed to express a certain amount of variability in the input dataset (see Fig. 2). In this work PCA is employed also to group in the first PCs the dominant variability of the problem that is associated with the behaviour of system through the various operating conditions and that conceals the interesting changes associated to novelties.

4. CHILLER DATA-SETS

Centrifugal water-cooled chiller data from fault tests at different levels of severity are used as anomalous behaviours to test and validate the proposed semi-supervised anomaly detection strategy. The data were provided by the ASHRAE research project 1043-RP. Specifically, experimental data from a 316 kW centrifugal water-cooled chiller were collected. A wide variety of chiller faults was studied, and four of them are here considered (Tab. 1). Each fault was introduced at four levels of severity (10%-40%), in increments of about 10% denoted by SL1-SL4. In Tab. 1, fractional values indicate the level of fault severity; for example, the range 0.59-0.61 under SL4 for reduced condenser water flow indicates that the flow was reduced to about 60% of the normal value. Tests (lasted about 14.4 hours) consist of 27 experiments during which the following three control variables have been modified: chilled-water outlet temperature from chiller evaporator,

Table 1. Considered chiller faults.

Symbol	Fault Type	Normal Operation	SL1	SL2	SL3	SL4
fwc	Reduced condenser water flow	17 [L/s]	0.87-0.93	0.77-0.81	0.69-0.70	0.59-0.61
fwe	Reduced evaporator water flow	13.6 [L/s]	0.90-0.91	0.81-0.82	0.72-0.72	0.63-0.65
rl	Refrigerant leak	136 [kg]	0.1	0.2	0.3	0.4
cf	Condenser fouling	164 tubes	0.06	0.12	0.20	0.30

Table 2. Characteristic features.

Symbol	Characteristic features
$x_{i,1}$	Evaporator Water Temperature Difference
$x_{i,2}$	Condenser Water Temperature Difference
$x_{i,3}$	Calculated Condenser Heat Rejection Rate
$x_{i,4}$	Calculated Evaporator Cooling Rate
$x_{i,5}$	Refrigerant Suction Superheat Temperature
$x_{i,6}$	Refrigerant Discharge Superheat Temperature
$x_{i,7}$	Liquid-line Refrigerant Subcooling from Condenser
$x_{i,8}$	Compressor Power
$x_{i,9}$	Calculated Compressor Efficiency
$x_{i,10}$	Evaporator Approach Temperature
$x_{i,11}$	Condenser Approach Temperature
$x_{i,12}$	Oil Feed minus Oil Vent Pressure
$x_{i,13}$	Oil in Sump minus Oil Feed Temperature
$x_{i,14}$	Pressure of Refrigerant in Evaporator
$x_{i,15}$	Pressure of Refrigerant in Condenser

condenser water inlet temperature, and chiller thermal load.

Machine Learning algorithms generally deal with static data, therefore in order to remove the dynamical component the steady-state data filter developed by Rossi [1995] was employed: after the filtering procedure, about 60% of the data were retained for modeling porpuses.

Moreover, experience gained from past studies [Comstock and Braun, 1999], [McIntosh et al., 2000] indicates that anomaly/fault detection can be more sensitive if certain characteristic quantities or characteristic parameters are used instead of the basic sensor measurements. These characteristic features can be directly deduced from the sensor measurements using arithmetic operations and thermodynamic refrigerant property tables or correlations. Definitions of the $d = 15$ characteristic features subsequently used for anomaly detection are provided in Tab. 2.

5. REFERENCE MODEL

In the following we will describe how the "reference" model to characterize the baseline system behaviour has been derived: the analysis described has been developed working on the PCs derived from the 15 characteristics in Table 2. In deriving the reference (fault-free) model, the observations of the ASHRAE RP-1043 related to non-faulty runs [2] have been divided into training (67% of samples) and validation (the remaining 33%).

The tuning phase of the One-Class SVM has been tackled as follows: the parameter ν, which approximates the fraction of training errors and support vectors, is set to a small value (0.05) to ensure a small misclassification rate, as motivated in Section 2, whereas the parameter σ is chosen via *tightness detection*, a heuristic approach to evaluate the tightness of the decision boundaries; if there exists at least one large hole inside the boundaries, i.e. a region without samples, the tightness is considered "loose". On the other hand, if the boundaries nearby two neighbouring samples are concave, the boundaries are considered "tight". The algorithm described in Wang et al. [2012] implements the aforementioned idea and has been employed in this work to tune σ; in Fig. 3, three 2D OCSVMs reference models are depicted for visualization sake: the dimensions $\mathbf{p}_{14}$ and $\mathbf{p}_{15}$ are the principal components obtained by projecting the features $\{\mathbf{x}_i\}$ onto the eigenspace associated with the two smallest eigenvalues. In detail, the pink coloured boundary is considered loose, whereas the light blue coloured boundary is considered tight; the blue coloured one is considered sub-optimal, i.e. neither loose nor tight.

[2] In the following we will refer to non-faulty conditions as "normal" conditions.

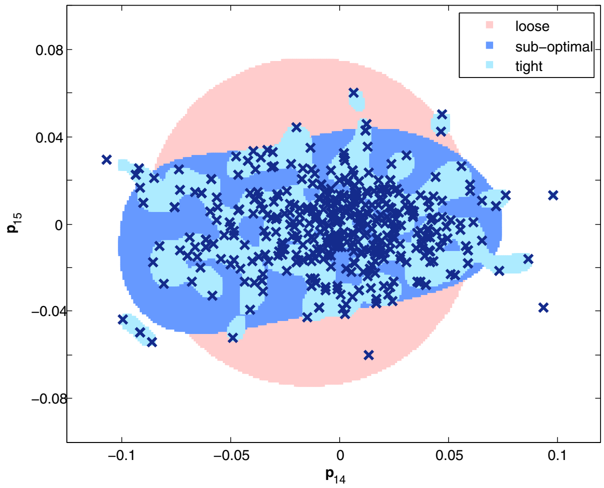

Fig. 3. Examples of tightness of the decision boundaries; the sub-optimal classifier can be seen as a trade-off between the loose (that barely adapts to the data at hand) and the tight solution (that clearly overfits the samples available).

Different classifiers have been computed depending on the input subset selection considered: in effect we will illustrate in the following how smart reduction of the PCs employed in the classification can strongly enhance the novelty detection performance. To verify this idea, we have tested classificators performance on new test data with faulty (see Tab. 1) and fault-free related runs.

Two experiment types have been developed based on different choice of the classificator input; in the first case, we have computed the classifiers $\left\{f^{(i)}\right\}_{i=0}^{d-1}$ by adopting different input spaces $\{U_i\}$ as follows:

$$f^{(i)}(U_i), \qquad U_i = [\mathbf{p}_{i+1} \dots \mathbf{p}_d], \tag{5}$$

for $i = 0, \dots, d-1$, with $d = l$. The performances, in terms of misclassification error rates, are provided in Fig. 4 for 3 different types of faults: the error rate is defined as the ratio of wrongly predicted data on total testing data. The abscissa of Fig. 4 represents the number of PCs discarded in the input space; more precisely, the leftmost point corresponds to the error rate without dimensionality reduction, while the next point corresponds to the error rate when data are projected onto the space spanned by all eigenvectors except that associated with the largest eigenvalue and so on. It is worth noticing that the curves generally have a U-shape: the minimum error rate is never

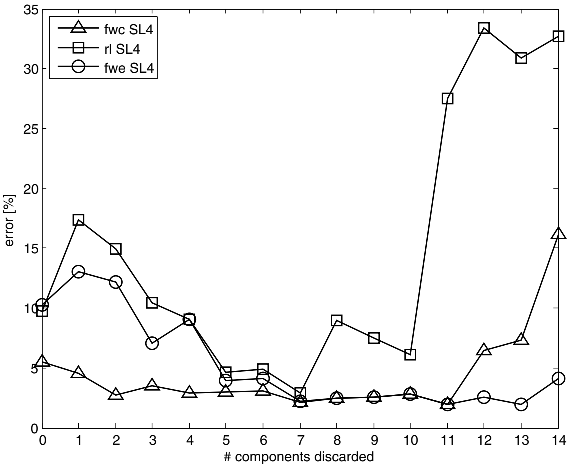

Fig. 4. Error rate as function of the number of discarded top eigenvectors.

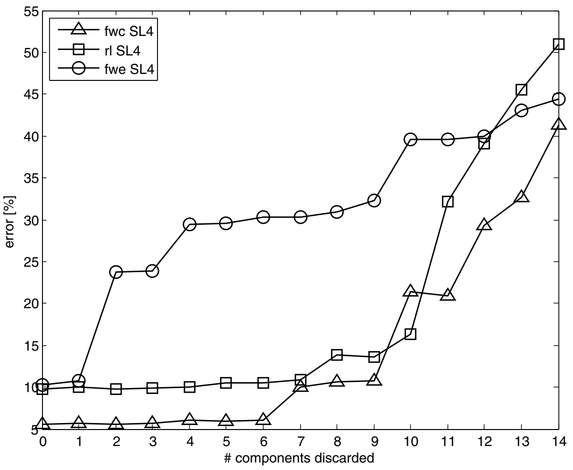

Fig. 5. Error rate as function of the number of discarded minor eigenvectors.

achieved when the PCs that contain most of the input variability are included into the input space; better results are obtained once those first PCs are discarded: this can be interpreted as an improvement of the classification performance once the most of the variability (mainly the one related to operating conditions changes) is left out from the input space. This result motivates the use of PCA analysis as procedure for ordering the variability of the system in exam and allowing the concealed anomalies related variability to be highlighted and more easily identified by the classificator. Projecting the data onto the directions of eigenvectors associated with smaller eigenvalues (e.g., from the 8th to the last one) before performing OCSVMs, the average error rate is at least halved, compared to the error without dimension reduction.

The previous idea that the first PCs describe most of the usual operating condition variability is also supported by a second type of experiment; the new classifiers $\left\{g^{(i)}\right\}_{i=0}^{d-1}$ differ from the working input space V_i employed:

$$g^{(i)}(V_i), \qquad V_i = [\mathbf{p}_1 \dots \mathbf{p}_{d-i}], \tag{6}$$

for $i = 0, \dots, d-1$. The results for this second type of experiments are reported in Fig. 5: it can be seen how the misclassication rate increases as the number of discarded PCs associated with the smallest eigenvalues is augmented. This proves how first PCs are not informative features for novelty detection.

Following the experiments outcome we have chosen to employ in Section 6 the classificator $f^{(7)}$, as it exhibits the lowest average error rate with the considered faults.

6. RESULTS

In order to assess the performances of the OCSVM we use data corresponding to different faults, and we evaluate the ability of the classification model to distinguish normal behaviour from anomalous one by Receiving Operating Characteristic (ROC) analysis, exploiting the Area Under the ROC Curve (AUC) as indicator of discriminatory power. ROC curve is created by plotting the fraction of true positives over the total actual positives vs. the fraction of false positives over the total actual negatives, for a range of different thresholds. A positive instance refers to a fault-free operational point, whereas a negative one corresponds to a faulty behaviour. The maximum value for the AUC is 1.0, thereby indicating a perfect test, while an AUC value of 0.5 indicates no discriminative value and it is represented by a straight, diagonal line extending from the lower left corner to the upper right.

In Fig. 6, we compute the ROC curve to analyze the classification model on a set composed by the nominal data (the fault-free portion of data employed for classifier validation) and data related to four different anomalies (i.e. reduced condenser water flow, reduced evaporator water flow, refrigerant leak and condenser fouling), at SL1 severity level. All the ROC curves are above the diagonal representing a good classification results (better than random classification): this fact is confirmed by the AUC values (reported in the Figure legend). The ROC analysis confirms that both the reduced condenser and evaporator water flow rate are easily detected [Comstock and Braun, 1999].

Furthermore, Fig. 7 depicts the performances of the classifier on test data considering just one type of anomaly, the reduced evaporator water flow, at different severity levels (SL1-SL4). The classification score increases as the severity levels raises: intense fault severities are easier to be classified and detected than low intensity faults (as intuitively expected).

For Predictive Maintenance purposes it is interesting to have not only an indication of the fault happening, but also an estimation of its intensity; we exploit the distance of the observations in exam from the OCSVM decision boundary as a mean for inferring information about anomaly severity: the underlying idea is that observations with small fault nature will be close to the normal data (and therefore to the decision boundary), while severe faults related samples will be distant from normally classified data and the OCSVM decision boundary.

Fig. 8 shows the boxplot representation of the signed-distances from the decision boundary of points related to the reduced evaporator water flow rate, at different severity levels. A positive distance value corresponds to a test point classified as normal, whereas a negative one refers to an anomalous behaviour; it can be seen that in

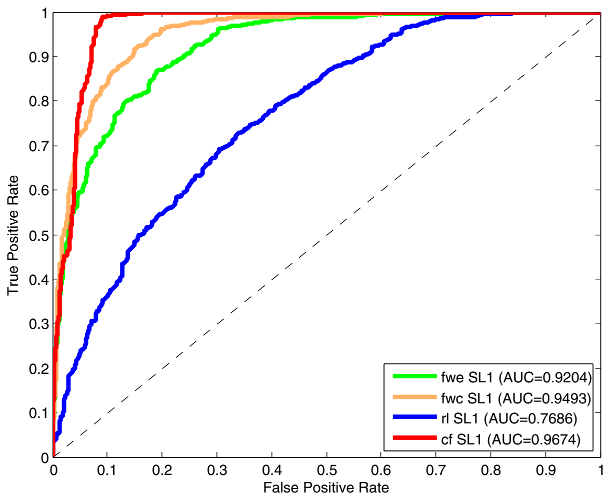

Fig. 6. ROC analysis related to different anomalies (faults) at SL1 severity level.

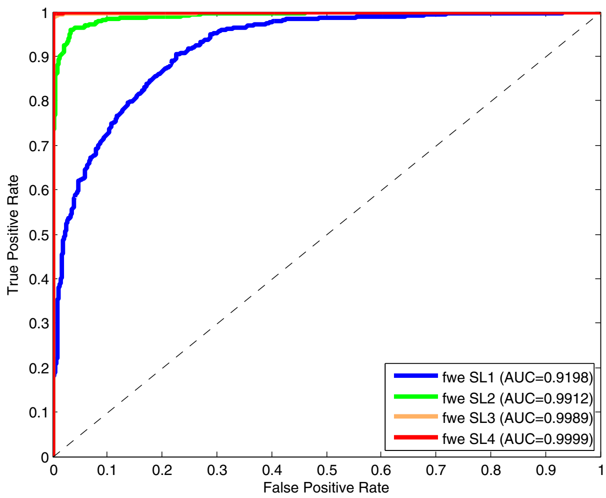

Fig. 7. ROC analysis related to the reduced evaporator water flow at different levels of severity.

correspondence with the increase of the anomaly severity level, the distance decreases in statistics; this shows how the distance from the decision boundary could be exploited for predictive maintenance purposes.

7. CONCLUSIONS

HVAC systems maintenance and energy efficiency can be increased by adopting Fault Detection systems. A major issue in developing efficient FD methods for HVAC installations is the unavailability of labelled data containing information about the operating conditions of the system. In this paper we have presented a novelty detection tool, being able to identify anomalous situations without using labeled data. We have shown how anomaly detection is not an easy task in HVAC world due to the fact that normal functioning conceals changes in the data related to anomalous conditions: the proposed combination of OCSVMs for classification and PCA for discarding the variability related to usual operating conditions changes has been shown to be effective in the detection performance.

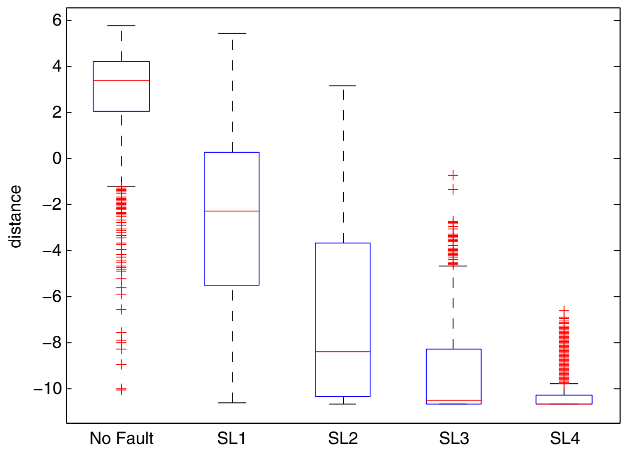

Fig. 8. Boxplot related to the signed-distance of points corresponding to the reduced evaporator water flow at different levels of severity.

REFERENCES

V. Chandola, A. Banerjee, and V. Kumar. Anomaly detection: A survey. *ACM Comput. Surv.*, 41(3):15:1–15:58, July 2009.

M.C. Comstock and J.E. Braun. Development of analysis tools for the evaluation of fault detection and diagnostics in chillers. Technical report, ASHRAE Research Project RP-1043, Report HL 99-20 4036-3, 1999.

T. Hastie, R. Tibshirani, and J. Friedman. *The Elements of Statistical Learning*. Springer New York, 2001.

D. Henry, S. Simani, and R.J. Patton. Fault detection and diagnosis for aeronautic and aerospace missions. In *Fault tolerant flight control*, pages 91–128. Springer, 2010.

I.B.D. McIntosh, J.W. Mitchell, and W.A. Beckman. Fault detection and diagnosis in chillers–part 1: Model development and application. *ASHRAE Transactions*, 106(2):268–282, 2000.

T.M. Rossi. *Detection, Diagnosis, and Evaluation of Faults in Vapor Compression Cycle Equipment*. PhD thesis, Purdue University, 1995.

B. Schölkopf, R.C. Williamson, A.J. Smola, J. Shawe-Taylor, and J.C. Platt. Support vector method for novelty detection. In *NIPS*, volume 12, pages 582–588, 1999.

K.A. Sievers and R.W. Mortonson. Fault detection and diagnostic system for automotive battery charging systems, April 12 1983. US Patent 4,379,990.

H.W. Stanford. *HVAC Water Chillers and Cooling Towers: Fundamentals, Application, and Operation*. CRC Press Taylor & Francis Gropu, 2012.

G.A. Susto and A. Beghi. A virtual metrology system based on least angle regression and statistical clustering. *Applied Stochastic Models in Business and Industry*, 29 (4):362–376, 2013.

G.A. Susto, A. Beghi, and C. De Luca. A predictive maintenance system for epitaxy processes based on filtering and prediction techniques. *IEEE Transactions on Semiconductor Manufacturing*, 25(4):638–649, 2012.

H. Wang, L. Zhang, Y. Xiao, and W. Xu. An approach to choosing gaussian kernel parameter for one-class svms via tightness detecting. In *2012 4th International Conference on Intelligent Human-Machine Systems and Cybernetics (IHMSC)*, volume 2, pages 318–323, 2012.