A temporal-based SVM approach for the detection and identification of pollutant gases in a gas mixture

Mohand Djeziri, Oussama Djedidi, Nicolas Morati, Jean-Luc Seguin, Marc Bendahan, Thierry Contaret

▶ **To cite this version:**

Mohand Djeziri, Oussama Djedidi, Nicolas Morati, Jean-Luc Seguin, Marc Bendahan, et al.. A temporal-based SVM approach for the detection and identification of pollutant gases in a gas mixture. Applied Intelligence, 2021, 10.1007/s10489-021-02761-0 . hal-03332042

HAL Id: hal-03332042
https://hal.science/hal-03332042

Submitted on 15 Mar 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Applied Intelligence manuscript No.
(will be inserted by the editor)

A temporal-based SVM approach for the detection and identification of pollutant gases in a gas mixture

Mohand A. Djeziri · Oussama Djedidi · Nicolas Morati · Jean-Luc Seguin · Marc Bendahan · Thierry Contaret ·

Received: date / Accepted: date

Abstract Air toxicity and pollution phenomena are on the rise across the planet. Thus, the detection and control of gas pollution are nowadays major economic and environmental challenges. There exists a wide variety of sensors that can detect gas pollution events. However, they are either gas-specific or weak in the presence of gas mixtures. This paper handles this issue by presenting method based on a Temporal-based Support Vector Machine for for the detection and identification of several toxic gases in a gas mixture. The considered gases are carbon monoxide (CO), ozone (O_3) and nitrogen dioxide (NO_2). Furthermore, an incremental algorithm is proposed in this paper for the selection of the best performing kernel function in terms of accuracy and simplicity of implementation. Then, a decision-making algorithm based on the rate of appearance of a class on a moving window is proposed to improve decision making in presence of uncertainties. This algorithm allows the user to master the false-alarms and no-detection dilemma, and quantify the level of confidence attributed to the decision. Experimental results, obtained with different gas mixtures, show the effectiveness of the proposed approach with 100% of accuracy in the learning and testing stages.

M. A. Djeziri · Oussama Djedidi
Aix-Marseille University, Université de Toulon, CNRS, LIS, Marseille, France
E-mail: mohand.djeziri@lis-lab.fr
E-mail: oussama.djedidi@lis-lab.fr

Nicolas Morati · Jean-Luc Seguin · Marc Bendahan Thierry Contaret
Aix-Marseille University, Université de Toulon, CNRS, IM2NP, Marseille, France
E-mail: nicolas.morati@im2np.fr
E-mail: jean-luc.seguin@im2np.fr
E-mail: marc.bendahan@im2np.fr
E-mail: thierry.contaret@im2np.fr

Keywords Classification · Data analysis · Micro-sensors · pollutant gases · Support Vector Machines · Smart-sensors

List of abbreviation and acronyms

ACC	Accuracy
CO	Carbon monoxide
CO_2	Carbon dioxide
Fr_i	Rate of appearance of the i^{th} class in a moving window
HCD	High Confidence Decision
I	Electric current
LCD	Low Confidence Decision
LSSVM	Least Squares Support Vector Machine
LSTM	Long Short-Term Memory
NN	Neural Network
KNN	K Nearest Neighbors
EBR	Episode-Based Reasoning
MOX	Metal Oxide Gas
MC-SVM	Multi-classes Support Vector Machine
MR	Misclassification Rate
N	Number of samples in an observation window
NO_2	Nitrogen dioxide
P	Electric power
PCA	Principal Component Analysis
PRCN	Precision
O_3	Ozone
F1	F1 classification score
SO_2	sulfur dioxide
V	Electric voltage
VOC	Volatile Organic Compounds
y	Classifier Result (Output)
ϕ_R	Resistance characteristic
th_i	Decision threshold for the i^{th} class

1 Introduction

Recent advances in microelectronics and chemistry have allowed the development of a large number of micro-sensors whose low consumption and low cost have made it possible to generalise their use in a large number of industrial and public applications [39,40,41]. In the field of air quality monitoring, several research works are conducted for the the development of measuring stations able to detect pollutants such as CO_2, CO, NO_2, SO_2, VOC and Ozone [25,3,16,29]. One of the main issues of these sensors lies in their low selectivity and many research works have been carried out to improve it. The solutions proposed are initially material solutions which consist in exploring and improving the physico-chemical properties of the materials used [36]. Nowadays, existing approaches can be classified into three categories: material approaches, software approaches and hybrid approaches which combine the two. Among the material solutions proposed, F. Rasch reports in [36] the doping of the semiconducting oxides with metal ions [27], surface functionalization or decoration with metal (oxide) nanostructures and polymers [35], formation of heterostructures, core-shell structures or nanocomposites by combination of n-type and p-type metal oxides [28] and the use of permeable membranes [43]. Software approaches gather the multivariate analysis such as principal component analysis [31], machine learning tools such as neural networks [42] and deep learning methods like deep belief network [38] and deep echo state network [9]. Principal Component Analysis (PCA) is an attribute extraction and classification technique based on the projection of all the measured data in the principal space made up of the components that presents the greatest variability. Its use for improving the selectivity of gas sensors has given satisfactory results, especially in the presence of single gases [32,41]. However, this technique does not make it possible to reduce the number of variables measured and consequently the complexity of the device consisting of the sensor and its software.

In this paper, the aim is not only to select the relevant characteristics, but also to reduce the number of variables measured and consequently to improve the design and the ergonomics of the micro-sensor. This will facilitate its integration into reduced environments, where the available space is an additional constraint in the choice of sensors as in aeronautics. So, the proposed approach for gas discrimination is done in two parts in this work:

The relevant measurements are identified by analysing the physical relationships between the measurable variables. The study of these physical relations makes it possible to highlight the linear and non-linear dependencies between the variables and to select the relevant variables, non-redundant and with a minimum of dependencies.

The Multi-classes Support Vector Machine (MC-SVM) is then used for classification as it allows the separation of classes using a wide range of kernel functions [24]. The issue of the selection of the kernel function as well as the order or scale of some functions is dealt in this work by the proposal of an incremental selection algorithm. The proposed algorithm is iterative, so it is time consuming, but since the selection of the kernel is done offline, there will be no impact on the online execution of the classification algorithm.In addition, the selection of the kernel function is done incrementally, from the simplest function to deploy to the most complex function, so the selected function will be both the most suitable and the easiest to implement.

Although the classification rates obtained for each class are satisfactory in the case-study, cases of observations not correctly classified can arise because of the overlapping phenomenon, as well as the fact that the training data is often incomplete. To take these cases into account, a decision-making algorithm based on the rate of appearance of a class on a moving window is proposed [8]. This integration of the temporal notion and the notion of memory in classification algorithms based on machine learning has been the subject of several research works [2] [26]. Most of them introduce these notions into the network, so they become characteristics to be identified during the learning process [15] [13] [14]. This is the case of the Long Short-Term Memory (LSTM) [22], which are memory networks, for which the update of one of the outputs of a layer is governed by a switch whose activation parameter is identified during the process of learning. This allows information to be stored for a defined time interval. Episode-Based Reasoning (EBR) is another technique based on the abstraction of temporal sequences of cases, suitable when the aim is to detect similar temporal episodes of cases, rather than similar isolated cases [37]. A review of techniques for finding episodic events and repetitive events in a data stream is proposed in [20]. In this paper, we propose a post-learning decision-making algorithm, which relies on the accuracy obtained during the learning and testing phases. This temporal-based decision-making scheme takes into account the temporal relationship of the mixture gas current state with the previous mixture gas states, ensures temporal stability of the decision given to the user, and gives the user information on the relevance of this decision by displaying the level of confidence configurable by the user.

The paper is organized as follows. Section 3 is devoted to a description of the experimental setup, the used sensor, and the data acquisition process. The analysis and visualization of data for the identification of relevant measures are presented in section 3. Section 4 presents the proposed classification method where the MC-SVM method is presented and an incremental algorithm for kernel selection is given. A decision-making algorithm reducing the rate of non or false detection is presented in section 5, and experimental results are given ans discussed in section 6. A conclusion is given in section 7.

2 Related Works

The association the MOX micro-sensor measurements with artificial intelligence algorithms for the discrimination of gases in a mixture has been the subject of several research works over these past years. This is due to the significant development of data acquisition, storage and processing tools as well as the development of classification and analysis algorithms. In [17], the discrimination of three gases (CO, NO2 and O3) is obtained by an approach consisting of three main steps: the extraction of temporal attributes from raw measurements, the selection of the most relevant attributes by the ReliefF algorithm and the classification of gases by a multi-class SVM. The results presented in [17], show that the accuracy depends on the number of attributes selected and the sampling interval. In order to discriminate mixtures of three gases (methane, carbon monoxide and hydrogen), Principal Component Analysis(PCA) is used in [5] for feature extraction, and then compared to Linear Discriminant Analysis (LDA) and Neuroscale techniques. Compared to PCA, LDA method allows the reduction of the overlapping between the classes, and Neuroscale preserves the data structure, as well as the possibility of incorporating subjective information. In the classification stage, the authors compared three probabilistic methods: Nearest Neighbors (KNN), Gaussian mixture models (GMM) and Generative Topographic (GTM). The PCA-GMM combination gave the best result with an accuracy of 92.7%. Multiple mixtures of CO and NO2 are considered for discrimination in [11], where two approaches are proposed. In the first method the raw detest is augmented by calculating the response value, response time and recovery time, then, PCA is used for feature extraction. Gas discrimination is performed using Genetic Algorithms optimized Back Propagation Neural Network (GA-BPNN). The second method use a Convolutional Neural Network (CNN) for both feature extraction and selection after transforming response curves into gray image. Both approaches provide an accuracy of 100% in different operating conditions. The mapping of raw data into pixels associated with a CNN for feature extraction and selection is used in [21] for gas discrimination in a mixture of CO, Methane and Ethylene, with an accuracy of 97.67%. In [34] a Deep CNN is proposed and applied for gas discrimination. The network is built by stacking two consecutive convolutional layers in order to introduce more nonlinearity into the network.

Table 1 summarizes recent work combining MOX micro-sensor measurements with classification algorithms for the discrimination of polluting gases in a mixture, including C0, NO2 and O3. This table shows that some of the proposed approaches start with an extraction-selection of relevant features followed by a classification algorithm, and others directly use classification techniques that include a feature extraction step such as CNNs. The obtained accuracy for each method given in Table 1 shows the effectiveness of the proposed approaches, but also show that an uncertainty persists and can therefore lead to errors in decision-making. In this paper, we propose a decision-making algorithm that can be associated with the methods proposed in Table 1 to process the uncertainties, and give the user a final decision with a defined level of confidence, to better appreciate the relevance of the decision taken the discrimination algorithms.

3 Experimental Setup

The test bench used is developed by the laboratory *Institut Matériaux, Microélectronique et Nanosciences de Provence*[1] (IM2NP). An overview of this equipment is given in the synoptic diagram (Figure 1). This system allows the generation of one or more pollutants and their controlled dilution in a carrier flow of neutral gas (dry air).

The dilution of the polluting gases is controlled by the use of mass flow regulators controlling the proportion of each component in the total output flow. They thus make it possible to generate an output mixture at low concentration (ppb and ppm ranges). Each pollutant can be generated independently or in mixing mode. A mixing chamber is placed at the outlet of the dilution plate to ensure the homogenisation of the mixture generated. A water steam injection device makes it possible to humidify the flow generated, with a regulated relative humidity (adjustable from 0% to 80%). The final output flow is adjustable between $0.1\,\mathrm{L\,min^{-1}}$ and $0.5\,\mathrm{L\,min^{-1}}$. The entire installation is controlled by an

[1] Materials, Microelectronics and Nanosciences Institute of Provence

Table 1: Summary of used models and obtained performances

Reference	considered gas mixture	Features used	classification method	Accuracy
[17]	CO, NO_2, O_2	7 features selected by ReliefF	SVM(RBF kernel)	100%
[5]	CO, $CH4_4$, H_2	PCA for features extraction	GMM	92.7%
[11]	CO, NO_2	Extraction of response value, response time and recovery time + PCA	GA+BPNN	100%
[21]	CO, Methane and Ethylene	Maping Original data into pxels	CNN	96.67%
[33]	Air, CO, Ethylene and Methane	GRU	2L-ARNN	97.67%
[34]	C0, Methane, Hydrogen and Ethylene	-	DCNN	95.2%

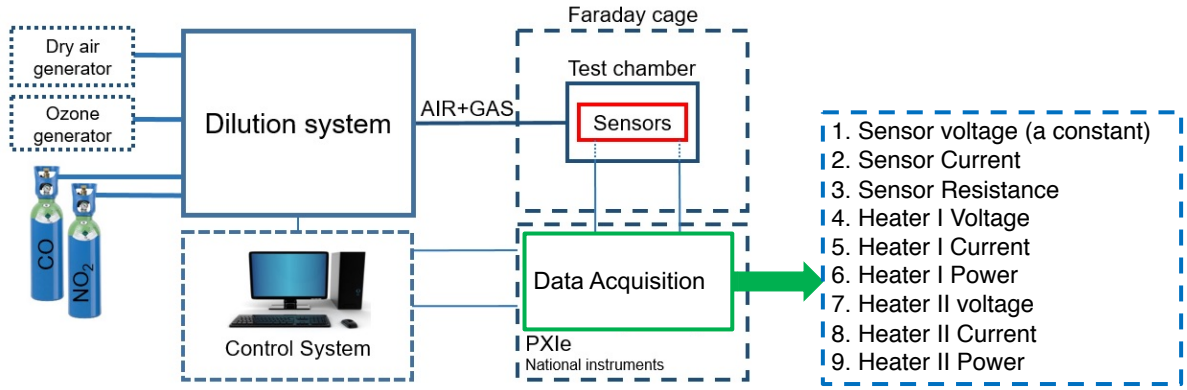


Fig. 1: Visual representation of the experimental Setup.

on-board automaton, which controls the dilution flow rates, the status of the valves as well as the alarms and safety devices. User control takes place through a user interface. Two operating modes are possible:

- user mode: the user controls the generation of a gas mixture (choice of dilution flow rates, flow check, valve sequencing)
- recipes mode: predefined recipes (or sequences) are edited and stored.

The used sensor in this study is developed by the IM2NP laboratory. It is a Metal oxide gas (MOX) sensor integrated with the heating elements in the same device in order to allow its heating online. The sensor operates in single sensor mode according to the diagram in Figure 2 . The sensor is biased at 0.9V using National Instrument PXIe 4140 source meter, and the heaters are powered by a triangular shape signal of amplitude 0.4V centered on 1.6V as illustrated in Figure 3. The measured variables are: Sensor voltage (a constant), Sensor Current, Sensor Resistance, Heater 1 Voltage, Heater 1 Current, Heater 1 Power, Heater 2 voltage, Heater 2 Current and Heater 2 Power.

The dynamic variation of the sensor's heating point improves its sensitivity to different gases, This behavior

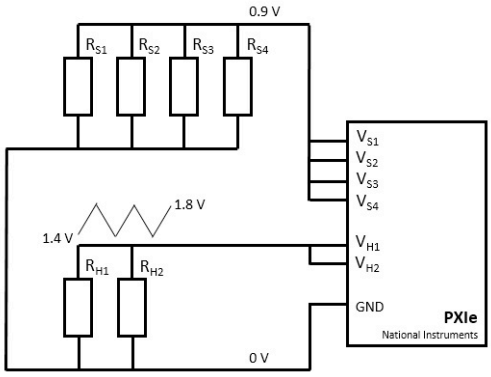


Fig. 2: Electrical configuration of the sensor: RS1, RS2, RS3 and RS4 are the resistances of the four detection zones. RH1 and RH2 are heating resistances of the two heaters [31].

can be explained by considering the temperature dependence of the surface coverage of chemisorbed species [6] [7]. It is demonstrated in [7] that at low temperature, the desorption phenomenon is weak and a total coverage of the sites can be obtained easily by ambient oxygen making any reversible detection impossible. Conversely, at high temperature the efficiency of the des-

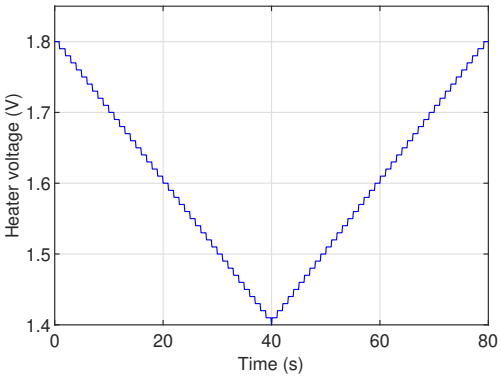


Fig. 3: A period of the voltage applied to the heaters.

orption mechanism increases and the system will be less sensitive. There is thus necessarily an optimum temperature at which the sensor has a maximum of sensitivity.

4 Identification of the relevant measurement

In this work, the aim is to propose an algorithm able to detect and identify the type of gas among three gases present in a mixture. The considered gases are: carbon monoxide (CO), ozone (O_3), nitrogen dioxide (NO_2), so, seven classes are possible. The plot of the individually measured signals, given in Figure 4, shows that it is not possible to distinguish the seven classes.

The combined plot of the signals measured in three dimensions, a sample of which is given in Figure 5, shows that the different classes are separable, and the profiles of the separating hyperplanes are relatively complex. The use of a Support Vector Machine (SVM) for the search of speration hyperplanes is therefore an adequate choice. Reducing the amount of variables input will reduce the complexity of the learning algorithm. In addition, the variables measured are interdependent, by looking at the physical meaning of the variables measured, we can easily see that the resulting physical variable (which capitalises the information contained in the other variables) is the power which is linked to the voltage, current and resistance by the following equations.

$$P(t) = f(u, i, \phi_R) \tag{1}$$

The function f can be linear or nonlinear, in the linear case the relations between the power and the measured variables, voltage u, current i and and resistance characteristic ϕ_R are:

$$P(t) = u(t).i(t) = \phi_R.i^2(t) \tag{2}$$

Equation.2 shows that the power is the physical variable which synthesizes the information carried by the other variables, the power of the sensor and the two powers of the two heaters are therefore selected as input variables of the classifier. The plot of the three powers is given in Figure 6, it shows that it is possible to separate the different classes by hyperplanes.

5 Gas detection and classification

5.1 Support Vector Classification

As illustrated in Figure 7, Support Vector Classification (SVC) is a supervised classification method, which aims to find a hyperplane with optimal margin, for the separation of two classes in a Hibertian space defined by a reproducing kernel associated with the scalar product of this space. In this work, the Multi-Classes Support Vector Machine (MC-SVM) in a "one against all" configuration is used [24]. In this configuration, n binary SVM classifiers are built, with n being the number of classes. Each of these classifiers is trained with a dataset composed in half of samples representing one of the classes (class $k \in [1, n]$) and a second half composed randomly of the rest of the classes. The classification results are obtained by a voting strategy: a pattern is classified to the class where the maximum number of votes is obtained.

This approach is known in the literature, it results from the work of Cortes and Vapnik in learning theory since 1995 [12], and is formally described as follows: Let consider that the learning data matrix (x) is composed of m attributes or variables representing monitoring indicators and a corresponding assigned label value $(y = Cl)$ $(l = 1, ..., n)$. The classifier builds a model which predicts the target class (y) from the testing input data (powers data in this case study), by searching an optimal hyperplane (Figure 7) optimizing a quadratic problem formalized in Equation 3:

$$\begin{aligned} &\min J(a) = \tfrac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} a_i a_j g_i(x) g_j(x) k(x, x) - \sum_{i=1}^{N} a_i \\ &s.t : \sum_{i=1}^{N} a_i g_i(x) = 0, 0 \leq a_i \leq D \; for \; i = 1, ..., N \end{aligned} \tag{3}$$

where $g(x) = 1$ if $x \in C1$ and $g(x) = -1$ if $x \in C2$, $a = [a_1, a_2, ..., a_N]^T$ are the Lagrange multipliers, D is the penalty parameter, and $k(x, x)$ is the Kernel function.

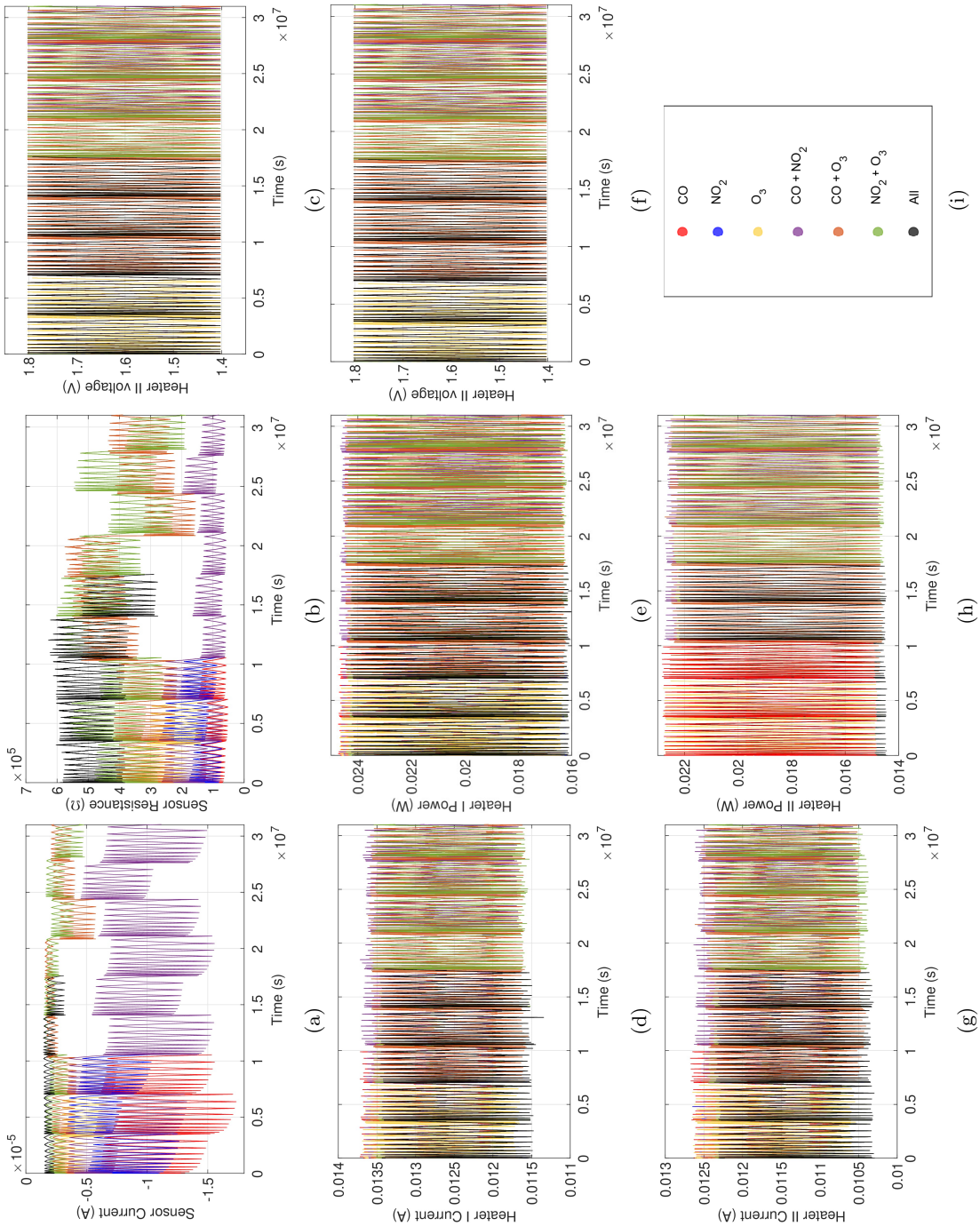


Fig. 4: Visual representation of the distribution of the classes according to the initial features.

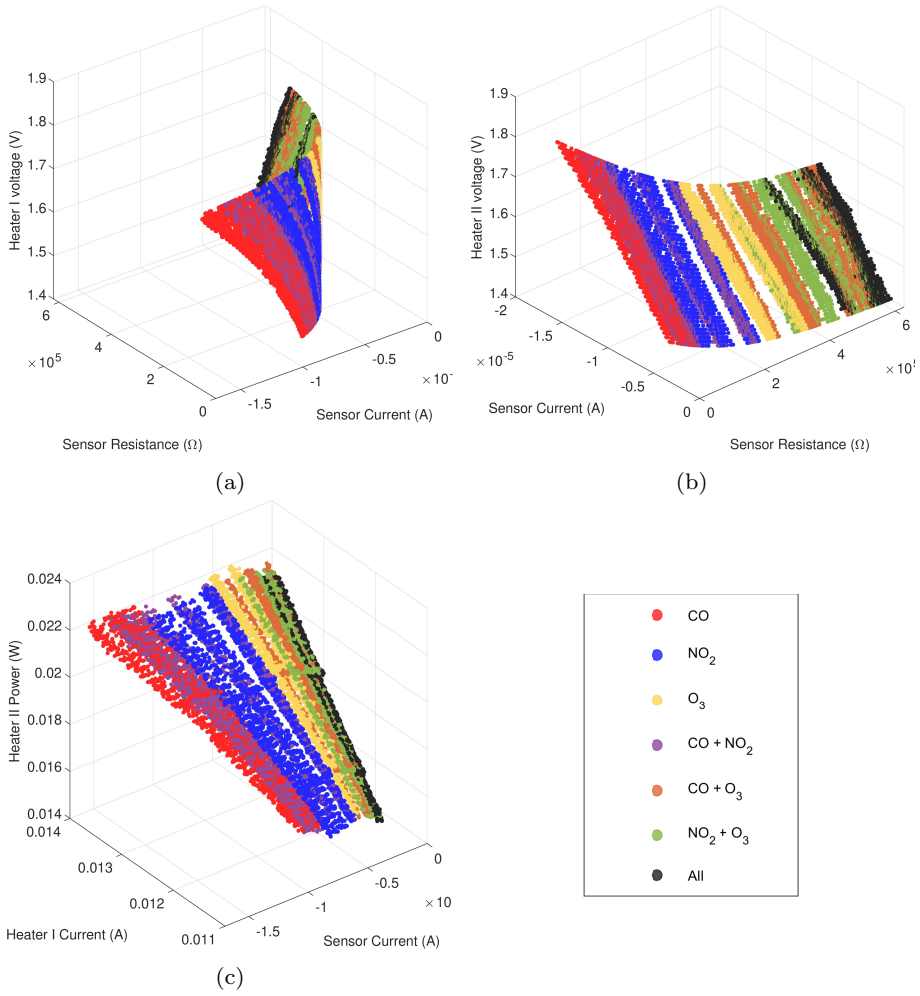


Fig. 5: Visual representation of the distribution of the classes according to the initial features. It is impossible to separate the classes linearly.

In the online stage, the classification of a new sample data is performed by using the following decision function:

$$y = \begin{cases} 1 & , if\, sign\left(\sum_{i=1}^{S} a_i^S g_i^S k(x_i^S, x) + b\right) = 1 \\ -1 & , elsewhere \end{cases} \quad (4)$$

where:

$$b = \frac{1}{S}\sum_{j=1}^{S}\left(g_j^S - \sum_{i=1}^{S} a_i^S g_i^S k(x_i^S, x_j^S)\right) \quad (5)$$

5.2 Algorithm for the selection of the Kernel function

It is commonly known that the innovative idea of the SVC is the reducing of the discrimination problem to a quadratic optimization one, which is based on the use of Lagrange multipliers for the optimization of the model (kernel function) parameters. However, one of the persistent problems concerns the selection of the kernel function for a given classification problem. Among the works focusing on the choice of kernel functions in the field of the detection of pollution events, [10] investigate the effects of different kernels embedded in Least Squares Support Vector Machine (LSSVM) algorithm and present a new kernel by using a logistic-based neural network for the detection of water pollution events. In [4], the Mercer condition is used to identify a composite kernel based on the most known kernel functions as bases. A Bayesian optimization is used to optimize the parameters of the defined composite kernel together with the hyperparameters of the SVM algorithm.

The issue of kernel identification is handled in this paper with an automatic search algorithm for the most suitable kernel function, considering a kernel-selection criterion as a predefined threshold for accuracy obtained

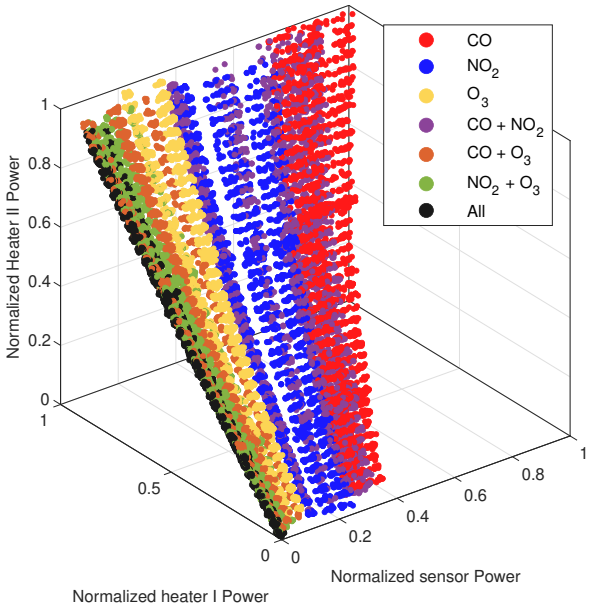


Fig. 6: 3D scatter plot of the classes according to the measured powers.

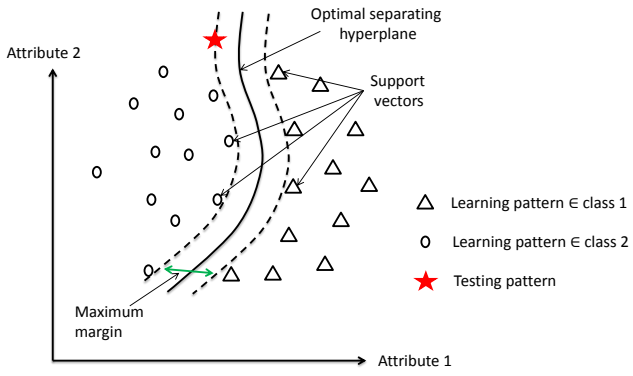


Fig. 7: Representation of the classification principle using Support Vector Classification

with the confusion matrices. This kernel-selection algorithm, given in Figure 8, is built incrementally, beginning by the linear kernel functions (linear function, polynomial,...) to the more complex one (Gaussian, Sigmoid,..) and then the combinations of kernel functions if the previous functions do not reach the required performance. The search for the kernel function stops when the preset performance threshold is reached.

To quantify the performance of the classifiers built with each of the kernels, several metrics can be used. In classification models, these metrics are the Accuracy (ACC), the Precision (PRCN), the Misclassification Rate (MR), the Recall, and the F1 score (F1). In a binary classifier as is the case of this work, Accu-

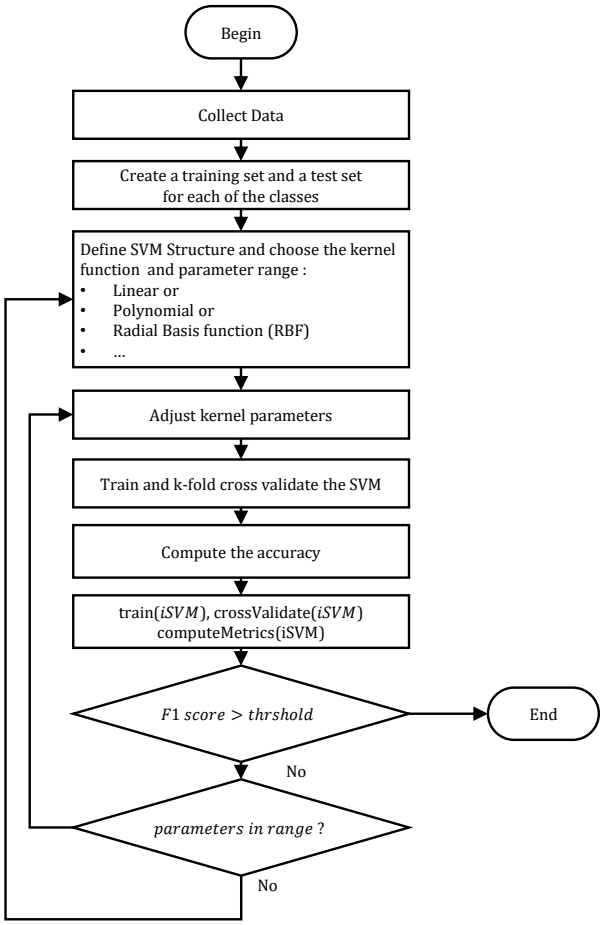


Fig. 8: Kernel-selection algorithm.

racy gives the overall performance at detecting the positive class and rejecting the negative class. Precision is the rate at which the model classifies the positive class amongst whole positively classified instances, whereas the Recall is the true positive (TP) rate of the model. Finally, the F1 score is a weighted average between precision and recall.

Since the aim of each of the binary classifiers is to accurately detect the positive class, the metrics that are sensitive to this class are precision and recall. Thus, the best quantifier for the performance of the model, in this case, is the F1 score as it is the harmonic mean of the precision and the recall.

6 Decision making

Decision-making is done more and more via software interfaces which provide a result without necessarily informing the user of its relevance. The current trend, in particular in the field of health and the environment, is to make available to the human operator the informa-

tion necessary for understanding the relevance of the decision taken by algorithms [23].

The presence of a single gas or a gas mixture is detected when the classifier result passes, during operation, from the normal operation to corresponding class $Ci/i = 1...n$. The SVM classifier gives a result (identifies the corresponding gas class) for each observation. In some cases, the observations cannot be perfectly classified due to the overlapping phenomenon, and the incomplete aspect of database used for training. So, an additional decision-making criterion is presented in this work.

In fault diagnosis theory, decision making always comes down to finding a compromise to the false alarms / no detection dilemma [18,1]. The word dilemma means that the improvement of one comes at the expense of the other. In the case of high-risk applications, it is preferable to have false alarms which will be checked and processed by the human operator than non-detection which can have serious human and material consequences [19]. In non-safety-critical applications, it is preferable to filter false alarms by using additional decision-making criteria based for example on the frequency of occurrence of false alarms in an observation window or the persistence of the alarm over a preset time .

In this work, the original idea is to replace instantaneous decision-making on each observation by a decision-making on a moving observation window, which will make it possible to master the false alarms and no detection dilemma. This criterion Fr_i is based on the rate of appearance of a class on a moving window, it is computed as follows:

$$Fr_i = \frac{\sum_{j=1}^{N} y == Ci}{N} \tag{6}$$

where y is the classifier result, and N is the number of samples in the observation window. A gas or a gas mixture is detected if the occurrence of a class exceeds a previously fixed threshold th (Eq.7). The rate of true positives (Recall) obtained in the confusion matrix of each class in the testing step can be used as an indicator to set the decision threshold th.

$$D_i = \begin{cases} 1 & Fr_i \geq th_i \\ 0 & Fr_i \leq th_i \end{cases} \tag{7}$$

To reduce cases of non-detection, the decision threshold should be chosen as follows:

$$th_i <= Recall_i \tag{8}$$

where $Recall_i$ is the Recall of the i^{th} class during the testing stages. The greater the difference between th_i and the true positive rate $Recall_i$, the more the rate of non-detection is reduced.

From the vector D, a high confidence decision (HCD) can be obtained by locating the non zero element of D:

$$HCD = \arg\max_i(D) \tag{9}$$

However, in practice, some sampling instances would result either in several elements of Fr surpassing the thresholds or none of them reaching it, thus, crippling the decision-making step. To avoid such shortfalls, the decision-making algorithm contains a second condition that would intervene in these cases. The retained class, in this cases, is the class that has the greatest Fr_i, resulting in low confidence decision (LCD):

$$LCD = \arg\max_i(Fr_i) \tag{10}$$

7 Experimental results

7.1 Data preparation

The dataset used in this work contains a different number of samples per class. For binary classifiers, Such distribution constitutes a problem called unbalanced learning. Hence, the first step is to balance the training and test sets. To do so, firstly, the number of samples of the set with the fewest samples is identified. Then, this number is used as a reference to build reduced datasets for the other 6 classes by randomly sampling rows from the original raw data.

Secondly, the newly reduced class sets are then divided into two parts, 70% for the training and 30% for the test. These subsets would constitute positive class information for each of the binary classifiers. Finally, to each training and test subsets, an identical number of samples is added randomly from the other classes to make up the information for the negative class.

7.2 Classification results

The data preparation step results in 7 training sets and 7 test sets. The training sets are then used to train 7 different support vector classifiers, each for a specific class (gas or a mixture). Figure 9 shows the confusion matrices of each class obtained from the test results using its RBF-based corresponding SVM.

To give an understandable interpretation of the results shown in the confusion matrices, let us consider the example of the CO gas confusion matrix (Figure 9a).

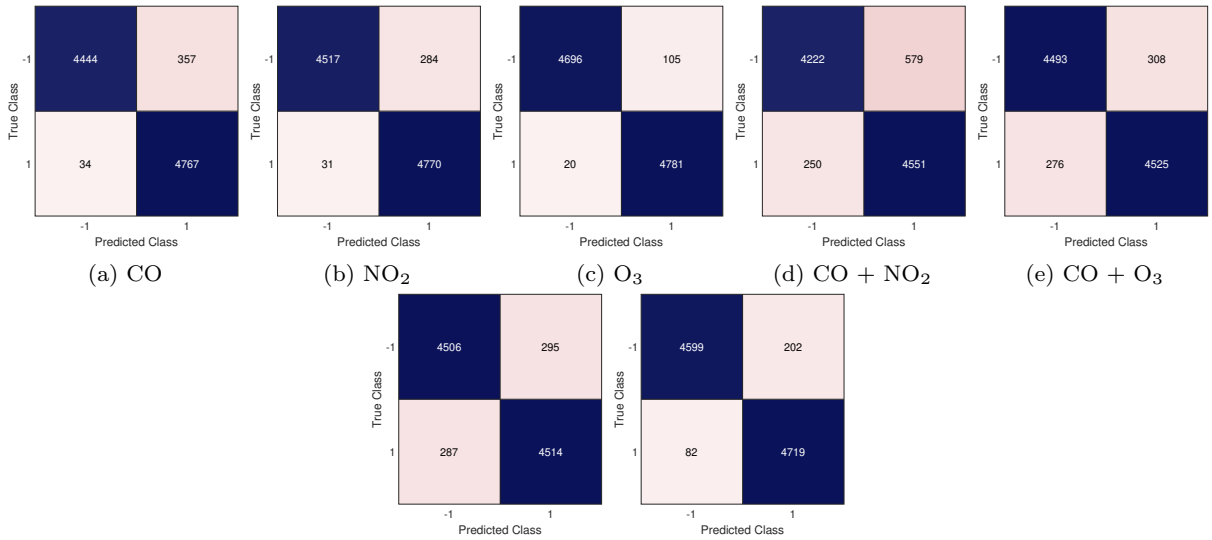
The said matrix shows that among 4801 (4767+34) observations belonging to class CO, 4767 (99.3%) are correctly classified as CO i.e. True Positive (TP), whereas 34 are been predicted as the negative class even though they belong to the positive one (CO), thus making them False Negatives (FN). On the other hand, among the 4801 (4426 + 375) observations belonging to the negative class, 4426 (92.3%) are predicted as True Negatives (TN), i.e. they truly do not belong to the CO class, and 375 ones are flagged as negatives while they are not, rendering them to be False Positives (FP).

The values deduced from each of the confusion matrices (TP, TN, FP, and FN) are used to compute the metrics evaluating the performance of the model, and comparing the results obtained from the use of each of the tested kernels, the linear one, the 2nd-degree polynomial, the 3rd-degree polynomial, and the radial basis function (RBF). The results for the training of each of the classifiers and their tests are summarized in Table 2, left to right from worst to best according to the F1 score.

The table shows the evolution of the performance from adequate in the case of a linear kernel to highly precise results in the case of an RBF kernel. It also clearly highlights where the classifiers fall short, and which classes are the most difficult to predict. These classes are the classes representing the mixtures CO + NO_2, CO + O_3, and NO_2 + O_3. As can be seen in Figure 6, these mixtures overlap with each other as well as with the classes making it harder to establish a linear or a polynomial hyperplane to separate them, and thus explaining poorer performance compared to the other classes.

Amongst the different kernels, the RBF function delivers the best results, overall 7 classes. While its precision suffers in the case of CO + NO_2 mixture, it still outperforms all the other kernels. Moreover, it delivers satisfactory results for all the other classes in terms of Precision and Recall, with an F1 score well over 90%. These high classification rates can be explained on the one hand by the efficiency of the proposed classification algorithm, and on the other hand by the nature of the application, since it can be seen in Figure 6 that the classes are separable, it is, therefore, possible to find a separating hyperplane which gives high accuracy. An example of such a hyperplane is given in Figure 10. The figure shows the decision boundary and the support vectors separating the positive class (CO). The profile of the hyperplane shows the efficiency of SVM for class separation thanks to a good choice of the kernel function.

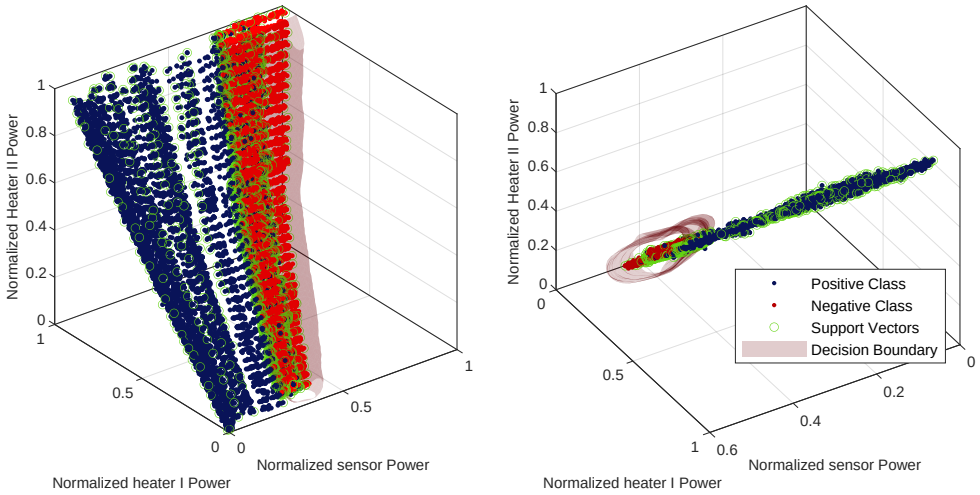
The Recall obtained for the CO class is 98.98%, so, according to the decision-making criterion proposed in equation 8, the decision threshold th is fixed in this case study at 70%, thus reducing the false alarm rate to zero. For each of the classes considered in this work, the decision threshold th is chosen while respecting the inequality of the equation 8.



(a) CO (b) NO_2 (c) O_3 (d) CO + NO_2 (e) CO + O_3

(f) NO_2 + O_3 (g) CO + NO_2 + O_3

Fig. 9: Confusion matrices of test results for each class.



(a) (b)

Fig. 10: The separation hyper-plane for CO class SVM viewed from two angles.

Table 2: Gas discrimination results using Multi-Support Vectors Machines and Decision Making. ACC: Accuracy. PRCN: Precision. MR: Misclassification Rate.

Kernel Function		Linear		Quadratic		Cubic		RBF		RBF + Decision Making		
Results		Training set	Test set	Training set	Test set	Training set	Test set	Training set	Test set	Test set		
										N = 5	N = 10	N = 20
CO	ACC (%)	93.96	93.81	94.31	93.95	94.31	94.31	94.63	94.25	98.61	99.79	**99.88**
	PRCN (%)	89.46	89.25	90.03	89.54	90.03	90.03	90.75	90.43	92.77	99.58	**100.00**
	MR (%)	6.04	6.19	5.69	6.05	5.69	5.69	5.37	5.75	1.39	0.21	**0.12**
	Recall (%)	99.65	99.63	99.65	99.52	99.65	99.65	99.38	98.98	97.91	98.96	**99.17**
	F1 (%)	94.28	94.15	94.60	94.27	94.60	94.60	94.87	94.51	95.27	99.27	**99.58**
NO2	ACC (%)	77.54	77.62	93.76	93.85	93.76	93.76	96.42	96.07	99.42	100.00	**100.00**
	PRCN (%)	72.98	72.86	89.82	89.91	89.82	89.82	93.36	92.80	96.18	100.00	**100.00**
	MR (%)	22.46	22.38	6.24	6.15	6.24	6.24	3.58	3.93	0.58	0.00	**0.00**
	Recall (%)	87.45	88.02	98.71	98.77	98.71	98.71	99.94	99.90	99.90	100.00	**100.00**
	F1 (%)	79.56	79.73	94.05	94.13	94.05	94.05	96.54	96.22	98.00	100.00	**100.00**
O3	ACC (%)	79.30	78.79	91.19	91.35	91.19	91.19	98.79	98.36	100.00	100.00	**100.00**
	PRCN (%)	74.32	73.38	85.45	85.87	85.45	85.45	97.70	97.23	100.00	100.00	**100.00**
	MR (%)	20.70	21.21	8.81	8.65	8.81	8.81	1.21	1.64	0.00	0.00	**0.00**
	Recall (%)	89.54	90.36	99.28	98.98	99.28	99.28	99.92	99.56	100.00	100.00	**100.00**
	F1 (%)	81.22	80.99	91.85	91.96	91.85	91.85	98.80	98.38	100.00	100.00	**100.00**
CO + NO2	ACC (%)	75.40	75.28	82.88	83.70	82.88	82.88	90.16	88.00	98.03	99.79	**99.88**
	PRCN (%)	74.35	74.30	75.59	76.40	75.59	75.59	83.90	81.54	97.58	98.96	**99.17**
	MR (%)	24.60	24.72	17.12	16.30	17.12	17.12	9.84	12.00	1.97	0.21	**0.12**
	Recall (%)	77.55	77.28	97.12	97.52	97.12	97.12	99.40	98.25	88.40	99.58	**100.00**
	F1 (%)	75.91	75.76	85.01	85.68	85.01	85.01	90.99	89.12	92.76	99.27	**99.59**
CO + O3	ACC (%)	74.78	74.68	75.73	74.84	75.73	75.73	95.91	93.85	99.91	100.00	**100.00**
	PRCN (%)	67.12	67.02	68.10	67.24	68.10	68.10	94.40	92.29	99.69	100.00	**100.00**
	MR (%)	25.22	25.32	24.27	25.16	24.27	24.27	4.09	6.15	0.09	0.00	**0.00**
	Recall (%)	97.16	97.19	96.83	96.90	96.83	96.83	97.62	95.69	99.69	100.00	**100.00**
	F1 (%)	79.39	79.33	79.96	79.39	79.96	79.96	95.98	93.96	99.69	100.00	**100.00**
NO2 + O3	ACC (%)	79.55	78.94	80.31	80.27	80.31	80.31	93.91	93.44	99.57	99.91	**100.00**
	PRCN (%)	73.20	72.42	71.91	71.91	71.91	71.91	90.77	90.02	99.57	100.00	**100.00**
	MR (%)	20.45	21.06	19.69	19.73	19.69	19.69	6.09	6.56	0.43	0.09	**0.00**
	Recall (%)	93.24	93.48	99.46	99.38	99.46	99.46	97.76	97.71	97.39	99.37	**100.00**
	F1 (%)	82.01	81.61	83.47	83.44	83.47	83.47	94.13	93.71	98.47	99.69	**100.00**
All	ACC (%)	90.96	90.60	91.40	91.12	91.40	91.40	93.26	92.33	99.60	99.91	**100.00**
	PRCN (%)	86.18	85.89	85.58	85.38	85.58	85.58	88.15	86.73	97.55	99.38	**100.00**
	MR (%)	9.04	9.40	8.60	8.88	8.60	8.60	6.74	7.67	0.40	0.09	**0.00**
	Recall (%)	97.57	97.15	99.58	99.23	99.58	99.58	99.95	99.96	99.69	100.00	**100.00**
	F1 (%)	91.52	91.17	92.05	91.78	92.05	92.05	93.68	92.88	98.60	99.69	**100.00**
Averaged	ACC (%)	81.64	81.39	87.08	87.01	87.08	87.08	94.72	93.76	99.30	99.91	**99.97**
	PRCN (%)	76.80	76.45	80.93	80.89	80.93	80.93	91.29	90.15	97.51	99.70	**99.88**
	Recall (%)	91.74	91.87	98.66	98.61	98.66	98.66	99.14	98.58	97.57	99.70	**99.88**
	microF1 (%)	83.61	83.45	88.92	88.88	88.92	88.92	95.05	94.18	97.54	99.70	**99.88**
	macroF1 (%)	83.42	83.25	88.71	88.66	88.71	88.71	95.00	94.11	97.54	99.70	**99.88**

7.3 Results with decision making

The results highlighted in the previous paragraph are obtained per class-specific SVM. In this section, on the other hand, are generalized by creating a new test set composed of each of the classes' test subsets and samples that were unused in previous sets. The new test set is then used on the Multi-SVM model N samples at a time. The model will generate a vector of 7 predictions per sample (One prediction per SVM-specific class), which will be processed by the decision-making algorithm described in Section 6. Naturally, the SVM used in this step are the highest performing for the previous step i.e. the RBF-based ones.

The results for this test are also indicated in Table 2, with multiple values for N. These results show how the decision-making layer drastically improves upon the existing results notably precision-wise as well as reducing the misclassification rate. They also show that these results can be further improved by increasing the sampling window N, reaching near-perfect results with N=20, at an averaged precision of 99.88%. The confusion matrix for the test with N=20 is shown in Figure 9. The latter shows that from 1680 sampling instances only 2 were misclassified, giving the model an overall accuracy of 99.97% (Figure 11).

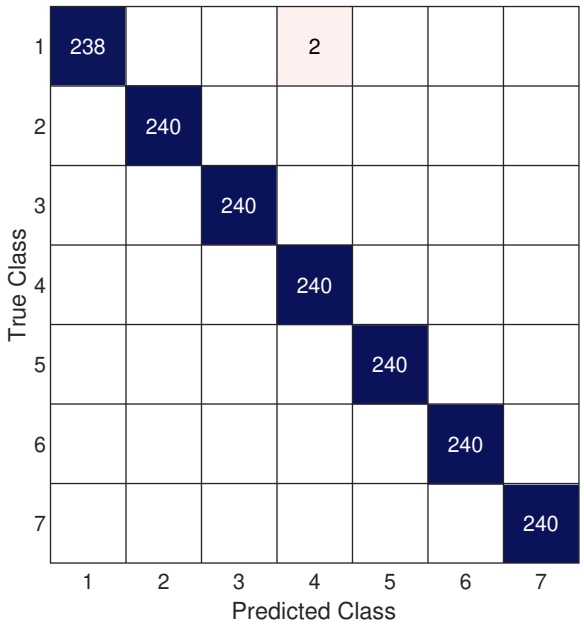


Fig. 11: Confusion matrix for the test of the Multi-SVM RBF-Based model with post-processing decision making (N=20).

Other techniques of extraction-selection of attributes associated with the most used classification techniques were applied to the database used in this work [30] [17]. To better situate the performance of the proposed approach compared to other techniques, a comparative analysis is given in the table below which shows that the proposed approach gives the best accuracy with a minimum number of measured inputs.

8 Conclusion

A smart gas micro-sensor capable of discriminating three types of gases, carbon monoxide (CO), ozone (O_3) and nitrogen dioxide (NO_2), present in a mixture of gases is presented in this paper. The proposed approach combines the dynamic online heating of the sensor with an automatic classification algorithm by Support Vector Machine. Dynamic heating makes it possible to permanently change the operating point of the sensor, an consequently its sensitivity to different gases. The analysis of the physical relationship between the measured variables allowed the identification of the power of the sensor and of the two powers of the two heaters as relevant variables, and the plotting of the observations in the spaces of the attributes motivates the choice of a support vector machine as an automatic classification method. Indeed, the SVM offers a large number of kernel functions from which it is possible to select the kernels able of building separating hyperplanes for each class of gas mixture. An incrementalselection algorithm of the kernel function is proposed in this work with an incremental search approach which ensures the selection of the most suitable kernel and also the simplest to implement. To deal with cases of uncertainty and master the false alarms / no detection dilemma, a decision-making algorithm based on the rate of appearance of a class on a moving window is proposed. The experimental results obtained using data from the IM2NP test bench show the efficiency of the proposed approach.

Acknowledgements The authors would like to thank Région Sud of France and Nanoz SAS for financial support. We also thank Tomas Fiorido for his technical support throughout this work.

Conflict of interest

The authors declare that they have no conflict of interest.

Table 3: Summary of the obtained results using other extraction-selection and classification methods

Refrence	Feature selection-extraction method	classification method	Accuracy
[17]	7 features selected by ReliefF	SVM(RBF kernel)	100%
[30]	Wavelet Decomposition of all the measured variables	KNN	54.4%
[30]	Wavelet Decomposition of all the measured variables	NN	59.8%
Proposed approach	Physical analysis	Temporal based SVM	100%

References

1. Adhikari, Y.R.: Inference and decision making methods in fault diagnosis system of industrial processes. IFAC Proceedings Volumes pp. 1–6 (2004)
2. Aggarwal, C.C.: Data streams. models and algorithms. In Series Advances in Database Systems. Springer **31**, 1–353 (2007)
3. Alrammouz, R., Podlecki, J., Abboud, P., Sorli, B., Habchi, R.: A review on flexible gas sensors: From materials to devices. Sensors and Actuators **2841**, 209–231 (2018)
4. Ay, M., Stenger, D., Schwenzer, M., Abel, D., Bergs, T.: Kernel Selection for Support Vector Machines for System Identification of a CNC Machining Center. In: IFAC-PapersOnLine, vol. 52, pp. 192–198. Elsevier B.V. (2019). DOI 10.1016/j.ifacol.2019.12.643
5. Belhouari, S.B., Bermak, A., Wei, C., Chan, P.C.: Gas identification algorithms for microelectronic gas sensor. MTC 2004 lnslrumentation and Measurement Technology Conference pp. 584–587 (2004)
6. Bendahan, M., Boulmani, R., Seguin, J.L., Aguir, K.: Characterization of ozone sensors based on wo3 reactively sputtered films: influence of o2 concentration in the sputtering gas, and working temperature. Sensors & Actuators **100**, 320–324 (2004)
7. Bendahan, M., Guerin, J., Boulmani, R., Aguir, K.: Wo3 sensor response according to operating temperature: Experiment and modeling. Sensors & Actuators **124**, 24–29 (2007)
8. Benmoussa, S., Djeziri, M.A., Sanchez, R.: Support vector machine classification of current data for fault diagnosis and similarity-based approach for failure prognosis in wind turbine systems. In: Artificial Intelligence Techniques for a Scalable Energy Transition: Advanced Methods, Digital Technologies, Decision Support Tools, and Applications, pp. 157–182. Springer International Publishing (2020). DOI 10.1007/978-3-030-42726-9_7
9. Bo, Y.C., Wang, P., Zhang, X., Liu, B.: Modeling data-driven sensor with a novel deep echo state network. Chemometrics and Intelligent Laboratory Systems **312**, 104062 (2020)
10. Chen, H., Xu, L., Ai, W., Lin, B., Feng, Q., Cai, K.: Kernel functions embedded in support vector machine learning models for rapid water pollution assessment via near-infrared spectroscopy. Science of the Total Environment **714**, 136765 (2020). DOI 10.1016/j.scitotenv.2020.136765
11. Chu, J., W.Li, Yang, X., Wu, Y., Wang, D., Yang, A., Yuan, H., Wang, X., Li, Y., Rong, M.: Identification of gas mixtures via sensor array combining with neural networks. Sensors & Actuators **329**, 129090 (2021)
12. Cortes, C., Vapnik, V.: Support-vector networks. Machine Learning **20**(3), 273–297 (1995). DOI 10.1007/bf00994018. URL https://link.springer.com/article/10.1007/BF00994018
13. Dagum, P., Galper, A., Horvitz, E.: Temporal probabilistic reasoning: Dynamic network models for forecasting. Knowledge Systems Laboratory. Section on Medical Informatics, Stanford University.
14. Dagum, P., Galper, A., Horvitz, E., Seiver, A.: Uncertain reasoning and forecasting. International Journal of Forecasting **11**, 73–87 (1996)
15. Dean, T., Kanazawa, K.: A model for reasoning about persistence and causation. Computational Intelligence **5**, 142–150 (1989)
16. Dey, N.: Semiconductor metal oxide gas sensors: A review. Materials Science and Engineering **229**, 206–217 (2018)
17. Djedidi, O., Djeziri, M.A., Morati, N., Seguin, J.L., Bendahan, M., Contaret, T.: Accurate detection and discrimination of pollutant gases using a temperature modulated mox sensor combined with feature extraction and support vector classification. Sensors & Actuators **339**, 129817 (2021)
18. Djeziri, M.A., Benmoussa, S.: Residual evaluation for fault diagnosis: Comparison of three approaches. Energy Procedia **ISSN:1876-6102** (2016)
19. Djeziri, M.A., Benmoussa, S., Zio, E.: Review of health indices extraction and trend modeling methods for remaining useful life estimation. Book Chapter Springer Nature Switzerland AG 2020 (2020)
20. Gaber, M.M., Zaslavsky, A., Krishnaswamy, S.: Mining data streams: A review. SIGMOD Record **34**, 18–26 (2005)
21. Han, L., Yu, C., Xiao, K., Zhao, X.: A new method of mixed gas identification based on a convolutional neural network for time series classification. Sensors **19**, 1960 (2019)
22. Hochreiter, S., Schmidhuber, J.: Long short-term memory. Neural Computation **9**, 1735–1780 (1997)
23. Holzinger, A.: Explainable ai and multi-modal causability in medicine. Wiley i-com Journal of Interactive Media **19**, 171–179 (2020)
24. Hsu, C., Lin, C.: A comparison of methods for multiclass support vector machines. IEEE Transactions on Neural Networks **13**(2), 415–425 (2002)
25. James, F., Fiorido, T., Bendahan, M., Aguir, K.: Development of mox sensors for low vocs concentrations detection: responses comparison for wo3, sno2, and zno sensitive layers with interfering gases as co and co2. International Journal on Advances in Systems and Measurements **10**(3-4), 158–162 (2017)
26. Joao, G.: Knowledge discovery from data streams. data mining and knowledge discovery. Chapman and Hall p. p255 (2010)

27. Kim, H.J., Lee, J.H.: Highly sensitive and selective gas sensors using p-type oxide semiconductors: overview. Sensors and Actuators **192**, 607–627 (2014)
28. Miller, D., Akbar, S., Morris, P.: Nanoscale metal oxide-based heterojunctions for gas sensing: a review. Sens. Actuators **204**, 250–272 (2014)
29. Mokoena, T.P., Swart, H.C., Motaung, D.E.: A review on recent progress of p-type nickel oxide based gas sensors: Future perspectives. Journal of Alloys and Compounds **80515**, 267–294 (2019)
30. Morati, N.: Système de detection ultra-sensible et séléctif pour le suivi de la qualité de l'air interieur et exterieur. PhD Thesis. Aix-Marseille University p. 234 (2021)
31. Morati, N., Contaret, T., Seguin, J., Bendahan, M., Djedidi, O., Djeziri, M.: Data analysis-based gas identification with a single mox sensor operating in dynamic temperature regime. AllSensors pp. 1–5 (2020)
32. Nallon, E., Schnee, V., Bright, C., Polcha, M., Li, Q.: Chemical discrimination with an unmodified graphene chemical sensor. ACS Sen **1**, 26–31 (2016)
33. Pan, X., Zhang, Z., Zhang, H., Wen, Z., Ye, W., Yang, Y., Ma, J., Zhao, X.: A fast and robust mixture gases identification and concentration detection algorithm based on attention mechanism equipped recurrent neural network with double loss function. Sensors & Actuators **342**, 129982 (2021)
34. Peng, P., Zhao, X., Pan, X., Ye, W.: Gas classification using deep convolutional neural networks. Sensors **18**, 157 (2018)
35. Postica, V., Vahl, A., Strobel, J., Carballal, D., Lupan, O., A. Essadek, N.d.L., Schütt, F., Polonskyi, O., Strunskus, T., Baum, M., Kienle, L., Adelung, R., Faupel, F.: Tuning doping and surface functionalization of columnar oxide films for volatile organic compound sensing: experiments and theory. J. Mater. Chem **6**, 23669–23682 (2018)
36. Rasch, F., Postica, V., Schütt, F., Mishra, Y.K., Lupan, O.: Highly selective and ultra-low power consumption metal oxide based hydrogen gas sensor employing graphene oxide as molecular sieve. Sensors and Actuators **3201**, 128363 (2020)
37. Sanchez-Marre, M., Cortes, U., Martinez, M., Comas, J., Rodriguez-Roda, I.: An approach for temporal case-based reasoning: Episode-based reasoning. ICCBR pp. 465–476 (2005)
38. Tang, S., Chen, W., Jin, L., Zhang, H., Li, Y., Zhou, Q., Zen, W.: SWCNTs-based MEMS gas sensor array and its pattern recognition based on deep belief networks of gases detection in oil-immersed transformers. Sensors and Actuators, B: Chemical **312**, 127998 (2020). DOI 10.1016/j.snb.2020.127998
39. Thai, N.X., Van Duy, N., Hung, C.M., Nguyen, H., Tonezzer, M., Van Hieu, N., Hoa, N.D.: Prototype edge-grown nanowire sensor array for the real-time monitoring and classification of multiple gases. Journal of Science: Advanced Materials and Devices (2020). DOI 10.1016/j.jsamd.2020.05.005
40. Tonezzer, M., Kim, J.H., Lee, J.H., Iannotta, S., Kim, S.S.: Predictive gas sensor based on thermal fingerprints from pt-sno2 nanowires. Sensors and Actuators B: Chemical **281**, 670 – 678 (2019). DOI https://doi.org/10.1016/j.snb.2018.10.102. URL http://www.sciencedirect.com/science/article/pii/S092540051831880X
41. Tonezzer, M., Le, D.T.T., Iannotta, S., Van Hieu, N.: Selective discrimination of hazardous gases using one single metal oxide resistive sensor. Sensors and Actuators B: Chemical **277**, 121 – 128 (2018). DOI https://doi.org/10.1016/j.snb.2018.08.103. URL http://www.sciencedirect.com/science/article/pii/S0925400518315417
42. Topalović, D.B., Davidović, M.D., Jovanović, M., Bartonova, A., Jovašević-Stojanović, M.: In search of an optimal in-field calibration method of low-cost gas sensors for ambient air pollutants: Comparison of linear, multilinear and artificial neural network approaches. Atmospheric Environment **21315**, 640–658 (2019)
43. Zhang, X., Liu, Y., Li, S., Kong, L., Liu, H., Li, Y., Han, W., Yeung, K., Zhu, W., Yang, W., Qiu, J.: New membrane architecture with high performance: Zif-8 membrane sup- ported on vertically aligned zno nanorods for gas permeation and separation. Chem. Mater **26**, 1975–1981 (2014)