Sensors and Actuators B 177 (2013) 1017–1026

Contents lists available at SciVerse ScienceDirect

Sensors and Actuators B: Chemical

journal homepage: www.elsevier.com/locate/snb

# An adaptive classification model based on the Artificial Immune System for chemical sensor drift mitigation

Eugenio Martinelli[a,*], Gabriele Magna[a], Saverio De Vito[b,d], Raffaele Di Fuccio[a], Girolamo Di Francia[b], Alexander Vergara[c], Corrado Di Natale[a]

[a] Department of Electronic Engineering, University of Rome "Tor Vergata", Via del Politecnico 1, 00133 Roma, Italy
[b] Italian National Agency for New Technologies, Energy and Sustainable Development (ENEA), Portici Research Center, P.le E. Fermi, 1, 80055 Portici (NA), Italy
[c] BioCircuits Institute, University of California San Diego, 9500 Gilman Dr., La Jolla, CA 92093-0402, USA
[d] DAEIMI, University of Cassino, Italy

ARTICLE INFO

*Article history:*
Received 12 March 2012
Received in revised form
12 November 2012
Accepted 28 November 2012
Available online 6 December 2012

*Keywords:*
Gas sensors
Sensor drift
Artificial Immune System

ABSTRACT

Multivariate statistics is the standard instrument to analyze the data of chemical sensor arrays. However, the non-reproducibility of sensors and in some cases the drift of signals make the prediction of multivariate classifiers rather uncertain. In recent years, adaptive classifiers that can follow the evolution of sensor signals have been introduced for many pattern recognition applications. The performance of these methods is sometimes strongly dependent on the frequency of class occurrence of the problem. In particular, the adaptation fails for those classes characterized by a scarce rate of occurrence. But for chemical sensor arrays the detection of such classes is often the goal of the application.

A suggestion to overcome this problem is offered by a set of algorithm inspired by the natural immune system. In this paper a modified version of an Artificial Immune System algorithm is introduced. This algorithm can achieve a classification model that is substantially immune to the drift of the sensors. Noteworthy, the immunity is independent from the occurrence frequency of the classes of the problem. The algorithm properties are demonstrated with both synthetic and real data. In the case of real data a 5 classes problem with an array of metal-oxide gas sensors operated for 18 months has been considered.



## 1. Introduction

Adaptability is a persuasive feature enabling systems to maintain their properties even in the presence of unexpected disturbances.

More specifically, a system composed by a collection of interacting elements with at least one devoted to the adaptation process is called adaptive [1,2]. The scope of this process is to increase the correlation between the functionalities or structure of the system in order to reach certain specific goals. In the artificial system, this adaptive process is defined a priori in such a way that the system is able to increase the probability to achieve its own scopes [3–6]. Since adaptability is a peculiar quality of living beings, the observation of natural processes is a powerful source of inspiration to design artificial systems. To this regard, models inspired by the human immune system have received great attention from the scientific community [7–9]. Among the many fields of application, data clustering is one of the most promising scenarios in which the Artificial Immune System (AIS) has obtained interesting results for the intrinsic adaptive capability to follow the evolution of the clustered data [10–13].

The Artificial Immune System, by definition, incorporates some forms of memory, outperforming the great part of the learning system in a dynamic classification context. Actually, most of the existent learning systems have a limited memory and no mechanism to balance the need to maintain the information about the knowledge acquired in the past and the need to store new acquired knowledge for the future predictions [14,15].

The use of an array of non-selective chemical sensors in a real context for the identification of a specific odor is a typical example of dynamic classification problem. However, it is well known that the chemical sensors may be affected by unpredictable degradations mainly due to poisoning and aging of the sensor material [16]. These causes result in a progressive loss of reproducibility that makes impossible the correct identification of measured samples. This phenomenon, generally known as drift, has been deeply investigated in the past.

Straightforward counteraction considers frequent recalibrations of sensors. Examples of optimized measurement protocols including periodical calibrations with respect to reference samples

* Corresponding author. Tel.: +39 06 72597259; fax: +39 06 2020 519.
*E-mail address:* martinelli@ing.uniroma2.it (E. Martinelli).

0925-4005/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.snb.2012.11.107

have been suggested [17,18]. More recently, it has been shown that ad hoc gas exposure strategy may increase the sensors reproducibility reducing the poisoning of the sensor material [19].

Many authors focused their attention to data processing looking for alternative signal descriptors that could be less affected by drift [20–24].

Gas classification with sensor arrays is a special application where drift effects are considered. In this particular case, it is possible to develop classification models that can remove drift effects taking advantage from the periodical measurement of reference samples [25–28]. In spite of the good results achieved by these models, the on-line calibration with reference samples is not always possible in real applications. Adaptive models offer a valid approach for the solution of this problem and Adaptive Resonance Theory and Self Organizing Map have shown a certain degree of drift rejection [29–32]. However, in these models the counteraction of drift is strongly dependent on the statistics of class occurrence, and it fails in case of unbalanced classes presentation.

These neural networks are optimal examples of artificial systems based on biological computation paradigms. Adaptability can be further improved considering other biological paradigms. For instance, adaptive strategy based on evolutionary models achieves drift correction even with smaller training sets than those requested by standard adaptive neural network models [33]. However this model assumes a temporal linear drift.

The search of immunity toward external agents inspired a number of algorithms endowed with superior adaptability properties. In this paper, an Artificial Immune Network (AINET) algorithm has been considered [12]. The original version of the algorithm has been modified in order to remove the necessity of ordered samples sequences. Then, the modified algorithm can compensate the sensor drift also in the presence of non-uniform presentation of the samples of the different classes and without any a priori assumption on the kind of drift (for example, linear or non linear).

The algorithm has been evaluated with three datasets. The first two are synthetics datasets that simulate linear and non-linear drift in a multi classes classification problem. The third dataset describes a real gas recognition application. These data were collected from an array of metal oxide sensors exposed to different gases for eighteen months. The rate of correct classifications provided by the algorithm has been compared with those obtained by standard classifiers, such as K-Nearest Neighbor (KNN) and Partial Least Squares-Discriminant Analysis (PLS-DA) [34]. Of course, any algorithm aimed at removing drift is supposed to outperform standard classifiers, however, the classification rate of standard classifiers provides a quantitative estimate of the magnitude of drift and then it provides a sort of quality criterion for the drift removal algorithm.

## 2. Adaptive Artificial Immune System

The immune system is a multi-level defense architecture within a living being that protects the organism against external agents. The system identifies and neutralizes antigens and pathogens responsible for infecting the organism [35]. In humans, the immune system is composed of an innate and an adaptive part. The innate immune system, also referred to as the non-specific immune system, comprises the cells and mechanisms that defend the host from infection by other invading organisms in a non-specific manner. The adaptive immune system, on the other hand, reacts to previously unknown foreign cells to construct a response to them that can remain active over a long period of time. Two of the most important cells involved in these processes are the T cells and the B cells, a collection of special types of cells called leukocytes, also known as white blood cells. Both, B cells and T cells, circulate throughout the entire body via the blood and the lymphatic vessels, carrying receptor molecules that recognize specific targets, and are involved in the humoral and cell-mediated immune response, respectively. In particular, B cells are responsible for the production and secretion of antibodies, which are specific proteins that bind to the antigen. Each lineage of B cell expresses a different antibody, so the complete set of B cell antigen receptors represents all the antibodies that the body can manufacture. The antigen is found at the surface of the invading organism, and the binding of an antibody to the antigen generates the signal that provokes the destruction of the potentially harmful invading cell. T cells, in contrast, recognize a "non-self" target, such as a pathogen, only after antigens (small fragments of the pathogen) have been processed and presented in combination with the "self" receptor called major histocompatibility complex, a cell surface molecules that mediate interactions of leukocytes with other leukocytes or body cells.

The outstanding information processing capability of biological systems has captured the attention of computer scientists over the past recent years. As a result, novel computational techniques inspired by the main features of the biological immune system have been implemented in a wide variety of applications, including, but not limited to, system protection and fault tolerance, pattern recognition, noise reduction, function optimization, and biological modeling [10,11]. These systems, called Artificial Immune Systems, are the result of an effort to implement protection against external attacks and internal faults explicitly inspired by the biological immune system.

The theory behind the immune system involves an exceedingly complex formation of components that work in coordination with the system itself. Building an artificial network of interacting elements that represents all the principles and processes occurring at such connectivity structure stage would be a reliable indicator of the biological immune system's original logic operation, uncluttered by all the implementation details. Yet, it is still a distinct bio-inspired tool useful for artificial systems to solve problems. This work considers a modified version of the Artificial Immune Network (AINET) algorithm originally proposed by de Castro and Von Zuben in 2001 [12]. This algorithm is based on a clustering technique that produces a representation of the initial dataset with a reduced number of templates. In the following section, we describe the details of the low-level implementation of such an algorithm and proceed by showing how the concepts inspired by biological immune systems can be applied to the definition of Artificial Immune Systems to solve the problems considered in this paper.

### 2.1. The AINET algorithm

In AINET algorithm each measurement is an antigen. The algorithm evolves a population of templates (called antibodies) that represent the measurement dataset with a compact set of examples. At the end, a network based on the collection of the templates is produced. The learning algorithm is summarized in Table 1.

The procedure of evolving antibodies (*Abs*) to represent antigens (*Ags*) can be explained as follows. First, a set of *Abs* is generated and stored into a memory matrix called M. Then the affinity of each $Ag_i$ (data point of the training set) with respect to all the $Ab_j$ of the M set is calculated. Here the affinity is evaluated as the Euclidean distance between $Ag_i$ and $Ab_j$ pairs; the bigger the distance between the $Ab_j$ and $Ag_i$, the lower is the affinity.

As a consequence, the subset of *Abs* with the largest affinity with respect to the $Ag_i$ is selected and used to generate new antibodies by means of the cloning process. This procedure is applied to each of the selected *Abs* and it consists in the generation of several replicas of the initial antibody (mother *Ab*) in order to create a set of new *Abs*, called clones, close to it. The amount of cloned elements is

**Table 1**
Steps of the AINET algorithm training.

1. *Create* a random **antibodies'** matrix (*Abs* initial pool)
2. At each iteration, do:
   2.1 for antigen $Ag_i$:
      2.1.1 *Calculate* current antigen ($Ag_i$) – antibodies (*Abs*) **affinity**.
      2.1.2. **Clonal selection:** *clone* the greatest affinity antibodies. The clone size is proportional to the affinities of *Abs*. That is, the higher the affinity is, the more *Abs* are cloned.
      2.1.3. **Affinity maturation:** *mutate* each *Ab* inversely proportional to its affinity with $Ag_i$. As a result, the *Abs* with higher affinities experience has comparatively smaller mutation.
      2.1.4. **Reselection:** *calculate* the affinity between each *Ab* and $Ag_i$; reselect a subset of *Abs* with highest affinity and remove the *Abs* with low affinity and assign to them the $Ag_i$ label.
      2.1.5. **Network suppression**: *remove* redundant *Abs* and insert the resulting *Abs* into M.
      2.1.6. **Repeat** for each Ag.
   2.2. **Suppress Abs in M**: *Remove* similar *Abs* in M to avoid redundancy.
   2.3. **Add** a set of new randomly generated *Abs* into M.
   2.4. **Repeat** until a predefined number of iterations are reached or the average affinity between *Abs* and *Ags* is below an assigned threshold.

proportional to the affinity of the elements of the subset of *Abs* with $Ag_i$; that is, the higher the affinity the largest the number of cloned *Abs* and vice versa.

Eventually, the *Abs* with the largest affinity with the *Ags* will have comparatively the smallest distances (mutation) from the initial *Ab*. Finally the *Abs* with lowest affinity are removed, inserting thus the resulting set of pre-selected *Abs* into M. More details on process of clone generation and mutation are reported in Ref. [10].

The following parameters may be finely tuned to improve the mapping of the input data:

- The number of *Abs* selected for cloning.
- The suppression threshold that defines the spatial region to eliminate redundant *Abs* (step 2.1.4, 2.1.5).
- The percentage of reselected *Abs* (step 2.1.4).
- The suppression event, which defines the threshold to remove the low-affinity *Abs* after the reselection (step 2.1.4).

The described AINET model is an unsupervised data algorithm, namely it does not depend on individual data class membership.

**Table 2**
Steps of $A^2$INET algorithm test.

1. For each **antigen**:
   1.1 *Classification*
      1.1.1 *Calculate* antigen – antibodies **affinity**.
      1.1.2 *Assign* the best affinity *Ab* class to the antigen.
   1.2 *Assigned class mutation*
      1.2.1 **Clonal selection**: *Clone* greatest affinity antibodies of selected class.
      1.2.2 **Affinity maturation**: *Mutate* clones.
      1.2.3 **Clone suppression**: *Eliminate* too far and near clones.
      1.2.4 **Update M pool**: *Add* these templates to the assigned class pool.
      1.2.5 **Network suppression**: *Eliminate* too far and near *Abs* and *check* pool dimension.
   1.3 **Centroid normalization**
      1.3.1 *Calculate* new centroid.
      1.3.2 *Translate* every class proportionally to **shift vector** ($\Delta\mathbf{C}$).

For applications with chemical sensors, which are subject to a non-negligible drift, a supervised model might be preferable. Here a supervised version of the AINET classification model is introduced. This model, according to the process described in Table 1, considers a separate dataset for each class. The algorithm can be easily modified, as shown in Fig. 1, by a straightforward manipulation of the original unsupervised version. In this case, at the end of the process each class is represented by a set of templates that are used with distance criteria such as KNN to estimate the class-membership of new samples.

### 2.2. Adaptive artificial immune network ($A^2$INET)

*$A^2$INET* is a semi-supervised AINET classification model where the existent antibodies, obtained in a training phase by means of the supervised AINET model shown in Fig. 1, are modified in order to maintain the maximum affinity in the presence of mutated antigens. This is a well-known property of the immune system called antigenic drift [36,37]. In particular, an additional procedure is applied to each data presentation in order to modify the class templates according to the change of the sensors data. All the steps of this model are shown in Table 2 and summarized in Fig. 2. As Table 2 describes, when a new data is presented to the model, the classifier assigns this sample to the class for which the maximum affinity (i.e. the minimum distance) is found. Once the sample is assigned to a specific class, the antibodies whose affinity with the new data is maximum produce new random antibodies according to the affinity propagation rule

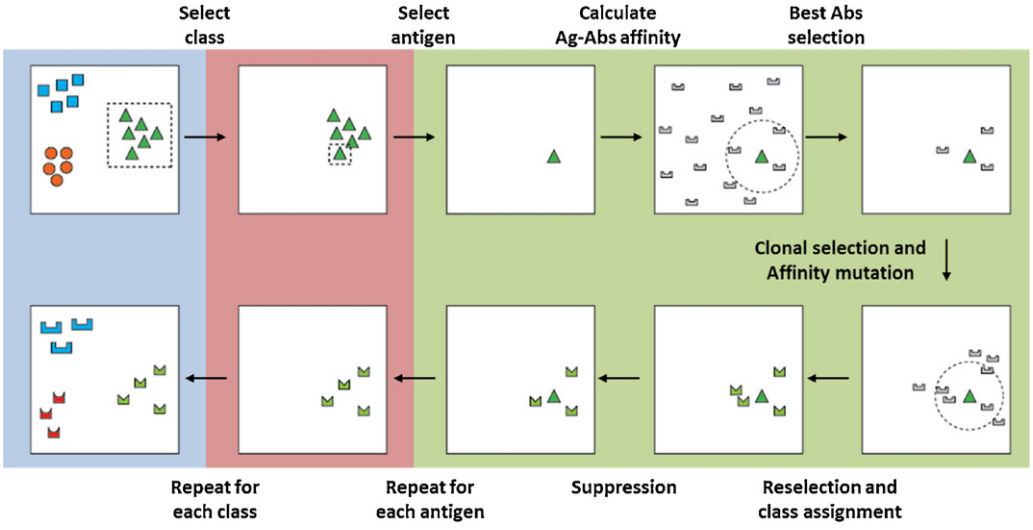


**Fig. 1.** Graphical illustration of the steps to train the supervised AINET algorithm.

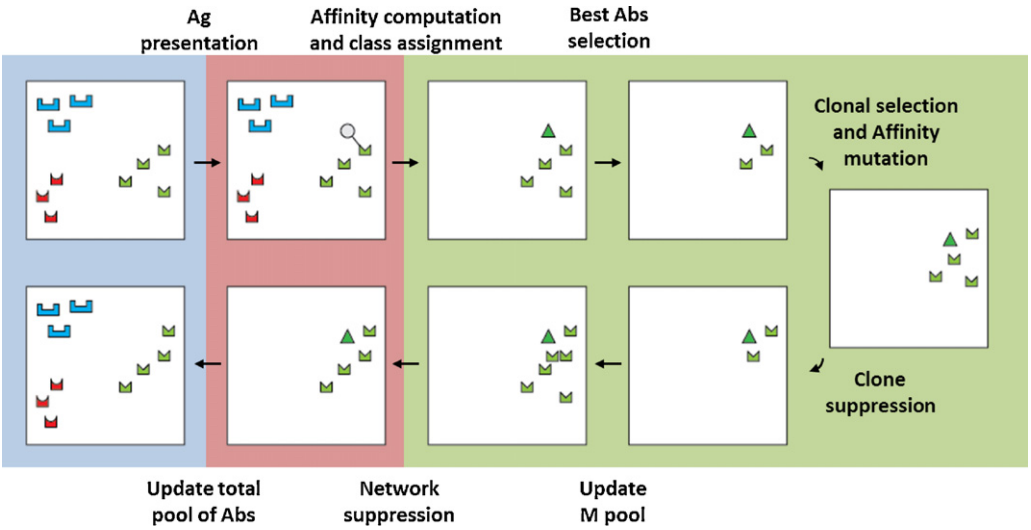


**Fig. 2.** Graphical illustration of the A²INET algorithm during test.

described in Table 2. As a result, only the antibodies with the largest sample affinity among those generated survive, adding thus a new set of class templates to the network. To compensate the addition of the new antibodies, the model eliminates, among the new pool antibodies, those with lowest affinity to the new data.

### 2.3. Centroid normalization

A²INET evolves according to the sensor drift but this capability requires that the frequency of occurrence of the classes is equal. This condition can be met in many practical cases; however, in many other cases this is not fulfilled. This is for instance the case of quality control of manufacture processes where the probability of "bad samples" is small.

In this case, the algorithm tends to follow the evolution of the most frequent classes failing in describing the evolution of the other classes.

For the Adaptive Artificial Immune System in an unbalanced dataset condition, the antibodies of the less frequent classes undergo a minor mutation reducing thus the adaptation capability.

To mitigate the dependency of the model adaptability to the class presentation, it is helpful to assume that the great part of the observed drift is common to all classes and global centroid normalization can be applied to all the antibodies to mitigate this effect.

In practice, after the presentation of a new antigen, all the antibodies shift according to the following expression:

$$\overrightarrow{\Delta C} = (\overrightarrow{C_{after}} - \overrightarrow{C_{before}}) * D, \quad (1)$$

where $C_{before}$ and $C_{after}$ are the average of all the templates before and after the template adaptation and $D$ is a parameter whose role is discussed below.

Templates averages are calculated according to the following:

$$\overrightarrow{C_{before}} = \left(\sum_{i=1}^{N} \overrightarrow{Ab_i}/N\right)$$
$$\overrightarrow{C_{after}} = \left(\sum_{i=1}^{N} \overrightarrow{Ab_i^*}/N\right) \quad (2)$$

where $N$ is the total number of antibodies (templates) and $Ab_i$, $Ab_i{}^*$ are the $i$th antibody before and after the templates adaptation. Eventually, all the antibodies are forced to follow the common variation of the centroid before and after the presentation of the new data.

The performance of the normalization depends on the parameter $D$ appearing in Eq. (1).

With $D = 0$ no common drift compensation is added to the class templates whereas with $D = 1$ the same drift is assumed for all the classes. Indeed through the setting of $D$, the evolution of the antibodies is tuned to the velocity of sensor properties variation. Thus, a large value of $D$ causes the antibodies to move too fast with respect to the sensor drift, while when $D$ is too small the shift of the antibodies could be not sufficient to follow the drift.

A careful evaluation of the parameter $D$ is then necessary to optimize the algorithm for the specific application.

## 3. Experimental

The algorithm has been tested with three datasets. Two datasets were synthetically generated to test the algorithm in the abstract situations of temporally linear and non-linear drifts. The third set of data was concerned with a laboratory experiment involving an array of metal-oxide semiconductor gas sensors applied to a five classes identification task.

### 3.1. Synthetic datasets

Each simulated dataset was composed of 900 measurements reproducing an array of five sensors measuring three equally probable classes. The training set formed by 150 measurements (50 each per class) was common for the two datasets. In both cases, the drift was assumed to be the same for all the classes.

The Principal Component Analysis (PCA) of the data provides a visual inspection of the datasets properties. PCA is a multivariate data analysis technique whose purpose is to linearly project the data in a subspace preserving the maximum amount of variance [34]. Figs. 3 and 4 show the scores plots of the first two principal components of the PCA of both the datasets. These plots show that the drift direction is a linear in the first case and curved in the second. However, in both cases all classes move in parallel.

In order to generate datasets that simulate a classification problem under linear or non-linear drift conditions, a two-step route has been followed. Firstly the train population, which consists of 150 samples equally divided in three classes, was generated. The intra-class dispersion was defined in order to avoid any overlapping among the three classes.

The second step consists in the generation of test set by an iterative procedure that simulates the effect of drift on train samples.

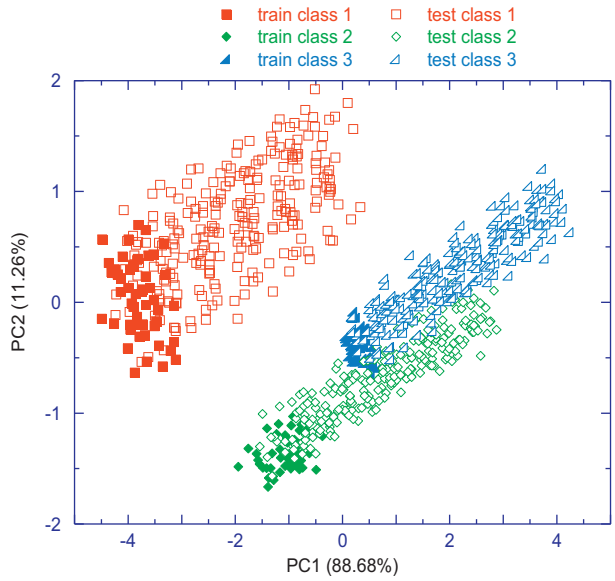


**Fig. 3.** Scores plot of the first two principal components of the artificial dataset characterized by a linear drift trajectory in the sensors space.

In particular, in each iteration a drift vector was added to the samples of the training set. The single elements of this vector represent the drift directions of each artificial sensors. Concerning the drift behavior, a linear trajectory occurs when the modulus of the vector is increased with a constant rate during the repetitions. On the other hand, a non-linear drift is obtained considering a quadratic relationship of the amplitude and a variable direction of the drift vector during the iterations. At the end of each iteration, a random component is added to each sample simulating the presence of noise.

The drifting parts of the datasets, formed by 750 data, were used to test A$^2$INET. The data were sorted according two different modalities: alternate and blocked. In the first mode the sequence of data presentation is alternated among the classes, while in the second mode data are alternated at blocks of 25 consecutive samples. In this way, although the samples are equally distributed for all the classes, their frequency along the experimental time is not uniform. These modes are schematically shown in Fig. 5.

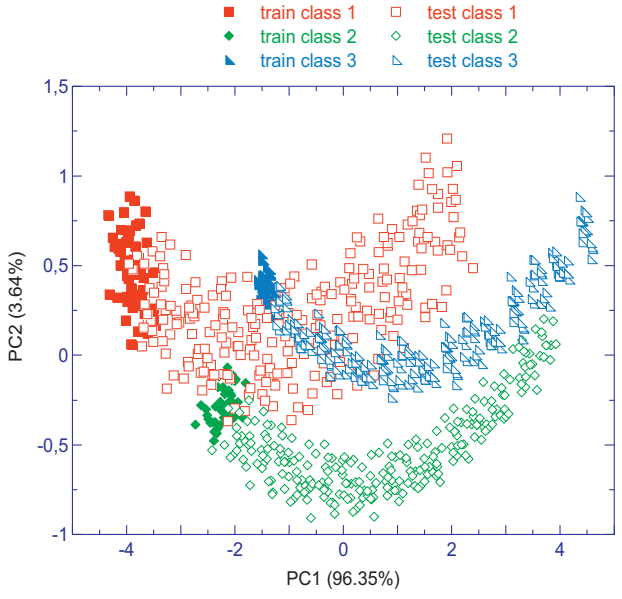


**Fig. 4.** Scores plot of the first two principal components of the artificial dataset featuring a non-linear drift trajectory in the sensors space.

### 3.2. Real dataset

The set of real sensors data was collected during an experiment performed at the Biocircuits Institute at the University of California San Diego [35]. The experiment involved an array of metal oxide semiconductor gas sensors formed by the following sensors provided by Figaro Inc.: TGS2600, TGS2602, TGS2610, and TGS2620.

Sensors were placed in a 60 ml volume Teflon/stainless steel cell into which gases were fluxed at a constant flow of 200 ml/min. Gas delivery conditions were maintained by mass flow controllers provided by Bronkhorst High-Tech B.V. Synthetic dry air was used as background and carrier gas. Humidity was kept constant at 10% R.H. (R.H. measured at $25 \pm 1\,^{\circ}$C) throughout the entire experiment.

The sensors were used in a five classes experiment aimed at identifying the following vapors and gases: acetaldehyde, acetone, ammonia, ethanol, and ethylene. A single concentration value for each compound is considered in the experiment. Gases were used from calibrated gas cylinders provided by Airgas Inc. The test compounds are quite different in terms of chemical structure, then they provide an interesting set to probe the selectivity of the array.

Gases were randomly sampled. Details about the gas-identity name, concentration values, and the sequence of measurements are listed in Table 3 where it is also indicated which part of the data was used to train the algorithm and which for test. A single measure was obtained exposing the sensors to gas mixed to carrier air for 100 s, afterward sensors were cleaned flowing only carrier air for 200 s.

The sensors resistances were measured at a rate of 100 Hz and each measurement produced a four-channel time series. Signals were recorded via a National Instruments data acquisition board controlled by *LabView* software on a computer platform fitted with the appropriate serial and/or digital-to-analog cards (DAC). Sensors temperature was fixed at 400 °C, this is a generic value at the midpoint of the temperature range suggested by the manufacturer.

Eventually, the dataset contained 509 measurements collected over a period of 18 months. The feature extracted from the sensors signals and used as input of the algorithm is the maximum shift of the resistance immediately before and at the end of the exposure to the gas.

### 3.3. Data analysis

To investigate the quality of the Artificial Immune System, two standard classifiers, such as the *k*-Nearest Neighbors (*k*-NN) and the Partial Least Square-Discriminant Analysis (PLS-DA), have been considered [34]. Since these classifiers do not implement any drift correction, A$^2$INET is expected to outperform. Nonetheless, the error of classifications of the two standard classifiers provides a simple method to evaluate both the magnitude of the drift and the complexity of the drift correction task.

For both standard classifiers and A$^2$INET, the datasets were divided into training and test subsets. One nearest neighbors ($k = 1$) was used with the *k*-NN classifier. For the PLS-DA classification model, the dataset was also auto-scaled and leave-one-out cross-validated to select the optimal number of latent variables [34]. For A$^2$INET the procedure was repeated 100 times, and the average and standard deviation of the whole analysis was then considered.

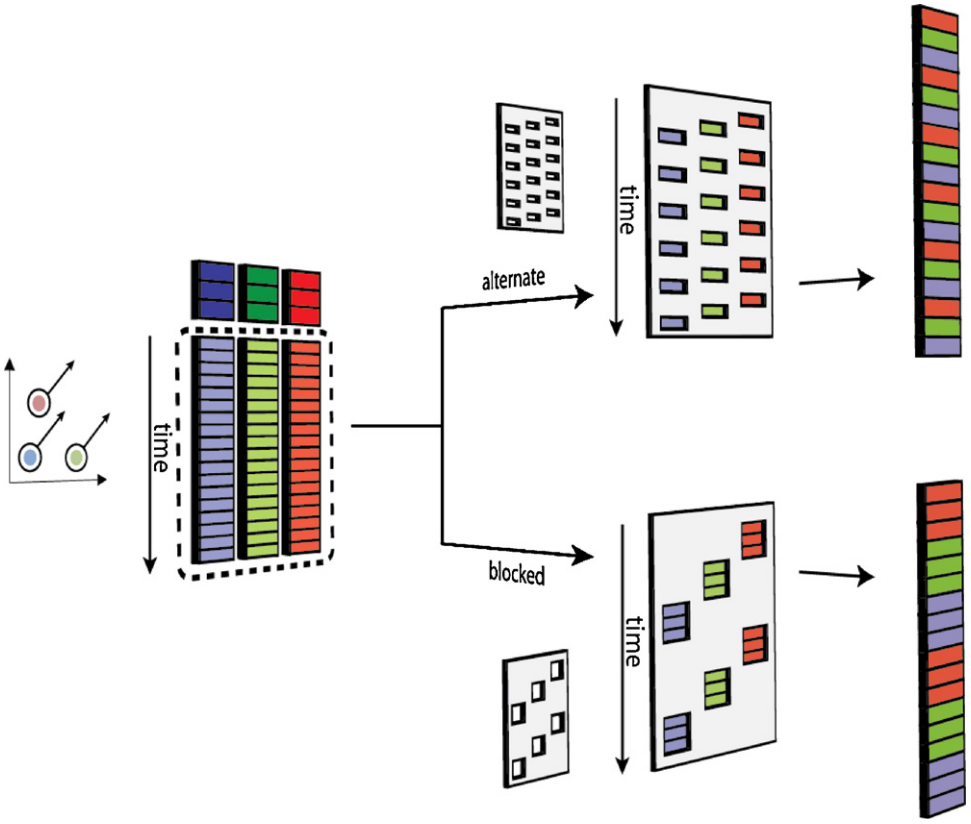


**Fig. 5.** Pictorial illustrations of alternate and block data sequence presentation.

**Table 3**
Number of samples and their distributions along the experiment. The experiment was performed in 18 sessions, one per month. In each session a different subset of classes were measured.

| | Month | Ammonia (@ 250 ppm) | Acetaldehyde (@ 100 ppm) | Acetone (@ 200 ppm) | Ethylene (@ 150 ppm) | Ethanol (@ 100 ppm) | |
|---|---|---|---|---|---|---|---|
| Training | 1 | – | – | – | 5 | 6 | 11 |
| Training | 2 | 23 | 12 | 11 | 7 | 11 | 64 |
| | | 23 | 12 | 11 | 12 | 17 | 75 |
| Test | 3 | – | 18 | 33 | 5 | 28 | 84 |
| Test | 4 | – | 5 | – | – | – | 5 |
| Test | 5 | – | – | – | – | – | 0 |
| Test | 6 | – | – | – | – | – | 0 |
| Test | 7 | – | – | – | – | – | 0 |
| Test | 8 | – | – | – | 25 | – | 25 |
| Test | 9 | – | – | – | – | – | 0 |
| Test | 10 | – | – | 30 | 20 | 32 | 82 |
| Test | 11 | – | – | – | – | 32 | 32 |
| Test | 12 | – | 20 | – | 26 | – | 46 |
| Test | 13 | 16 | – | – | – | 20 | 36 |
| Test | 14 | – | – | – | – | – | 0 |
| Test | 15 | 13 | 20 | – | – | – | 33 |
| Test | 16 | 20 | 9 | – | 30 | 12 | 71 |
| Test | 17 | – | – | – | 10 | – | 10 |
| Test | 18 | – | – | – | 10 | – | 10 |
| | | 49 | 72 | 63 | 126 | 124 | 434 |

Finally, the classification rate of A$^2$INET was compared with the performance of the supervised AINET.

## 4. Results and discussion

In analyzing the three datasets the influence of the parameter $D$ characterizing the centroid normalization (see Eq. (1)) was firstly studied. The classification performance of the optimized A$^2$INET was then compared with those obtained with the other classification models.

### 4.1. Synthetic datasets

As previously discussed the two synthetic dataset differ for the trajectory of the classes in the multidimensional sensors space. For sake of brevity, here the two cases are called linear and non-linear

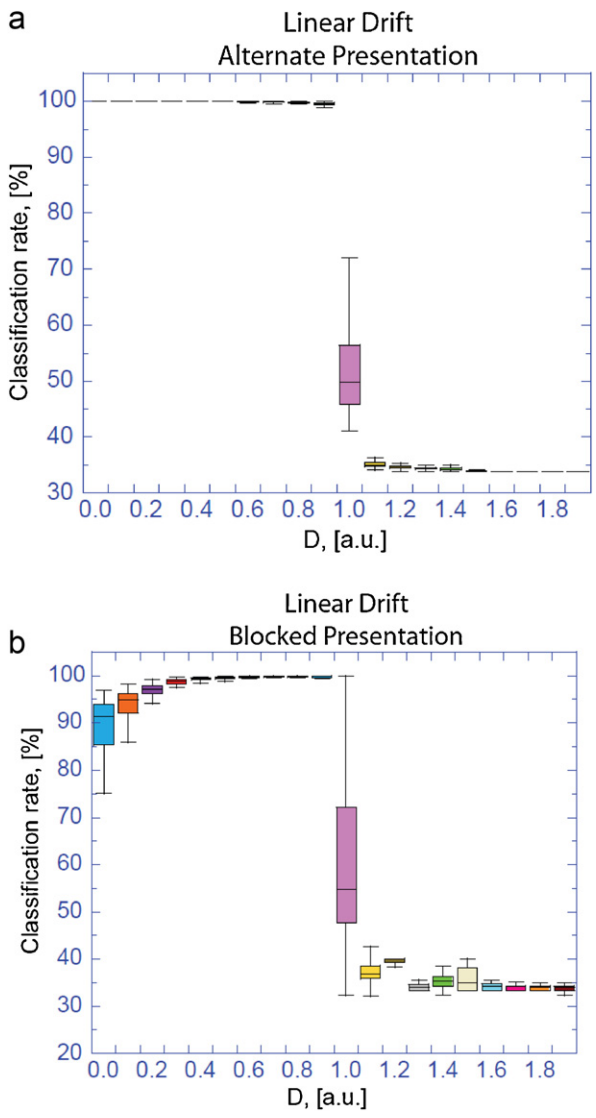


**Fig. 6.** Classification rate as a function of the parameter $D$ for the linearly drifted simulated data. (a) The results for the alternate data presentation and (b) for the blocked data presentation. For each value of the parameter $D$ a collection of models was calculated, here the statistics of the classification rates achieved by each individual model are plotted as box-plots.

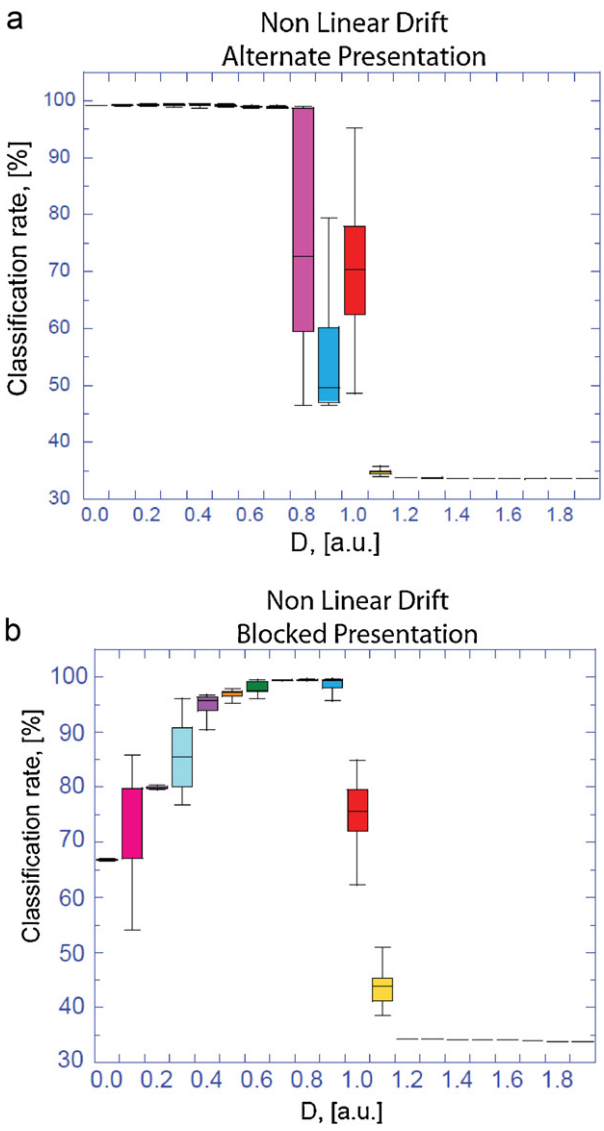


**Fig. 7.** Classification rate as a function of the parameter $D$ for the simulated data characterized by a non-linear drift. (a) The results for the alternate data presentation and (b) for the blocked data presentation. Performances are given as the box-plots of the distribution of the classification rates of 100 models replicas.

drift. It is clear that this definition is only related to the spatial direction of the sensor data in the sensors space. As mentioned in the previous sections, the performance of adaptive classifier generally depends on the sequence of class membership of the sorted data. In order to test the dependence of the algorithms here studied, the data were arranged in a purely alternate sequence and in a blocked sequence (Fig. 5).

Fig. 6 shows the correct classification rate of A$^2$INET as a function of the value of the parameter $D$ in the case of linear drift. Fig. 6a shows the results in the case of the alternate data sequence and Fig. 6b in the case of blocked data sequence. Similarly to all iterative algorithms, also in A$^2$INET the results depend from the initial values of the algorithm variables. For this reason, the algorithm was performed 100 times, the statistics of the results are reported in Fig. 6 as box-plots. A box-plot is a synthetic view of a statistical distribution where average (central line), standard deviation (box) and distribution extremities (external lines) are represented.

In both cases the best performance is obtained for $D < 1$, and the classification rate drops dramatically to less than 50% for $D \geq 1$. The sudden decrease of performance can be likely attributed to the fact that with $D \geq 1$ the normalization results in a movement of the templates is faster than the sensor drift.

In case of alternate data presentation, optimal classification rate is reached for any $D < 1$. Actually in this range of $D$ values, the alternation of all the classes drastically reduces the effect of the centroids normalization. Moreover the adaptive update of the class templates is sufficient to follow the drift evolution. On the other hand, in case of blocked data presentation the classification rate is rather sensitive to $D$, and the maximum value is obtained immediately before the condition $D = 1$.

Fig. 7 shows the results of A$^2$INET applied to the nonlinear drift. Here, instead of the sharp change between good and bad classification, around $D = 1$, the transition is rather smooth involving a range of $D$. In case of alternate data presentation the interval is extended from 0.8 to 1 and in the case of blocked data sequence the interval is between 1.0 and 1.2. For values of $D$ less than these transition intervals, the dependence of the classification rate from $D$ is similar the behavior observed for linear drift but the sensitivity of the classification rate with respect to $D$ is larger for the non linear drift.

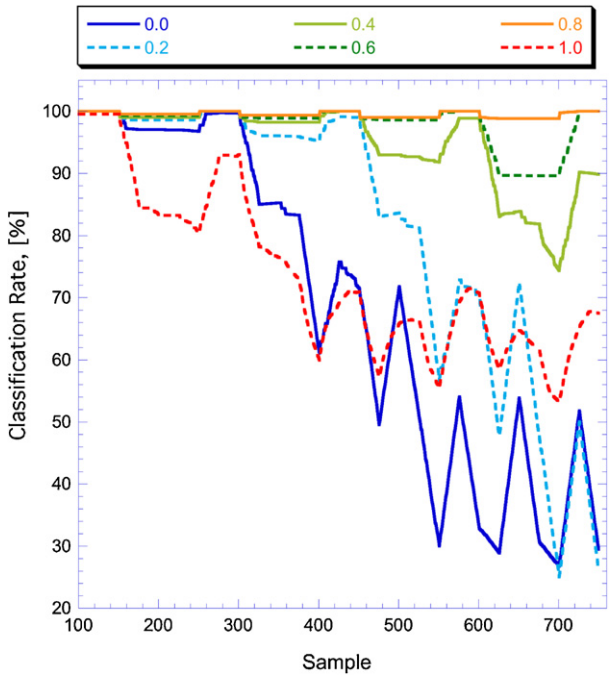


**Fig. 8.** Evolution of the classification rate for the nonlinearly drifted simulated data and for different values of the parameter $D$.

These results show that with an accurate choice of the parameter $D$, an almost error free classification rate is obtained even when the data sequence is not alternated among the classes.

A further insight into the adaptive capability of A$^2$INET model is obtained studying the evolution of the classification rate during the test and as a function of the number of presented samples. In a real situation this would correspond to a plot as a function of time. Fig. 8 shows the results for the non-linear drift and for different values of $D$. The classification rate is calculated using a moving window of 100 samples. Quite straightforwardly, the classification rate decays progressively as the number of samples increases and the drift progressively move the classes away from the training data. The decay changes for different values of $D$, and in particular for $D = 0.7$ the decay becomes negligible and about 100% of correct classification is achieved along the whole dataset.

The classification rate obtained by the A$^2$INET algorithm is compared in Table 4 for the linear drift and in Table 5 for the non-linear drift with the standard classifiers. Standard classifiers, where no drift compensation is considered, have a poor classification rate that drops below 50% in the case of non-linear drift.

As expected, a clear improvement is provided by the adaptive immune algorithms. No differences between the standard immune model and A$^2$INET are observed when the data are presented in the alternated mode. But the centroid normalization reveals its superior properties of drift rejection when samples are presented in the block mode. Indeed in this case it avoids that the correction is applied only to the class at which the sample belongs but the correction is extended to all the classes. The improvement of A$^2$INET is evident in the case of non-linear drift with a difference of more than 30% of the classification rates shown by the two immune models.

**Table 4**
Classification rates of adaptive immune models and standard classifiers applied to the synthetic data describing a linear drift trajectory in the sensors space. Both alternate and blocked data sequences are considered.

| Linear drift | Classification rates (%) | |
|---|---|---|
| | Alternate | Block-sample |
| PLS-DA | 57.07 | 59.6 |
| $k$-NN | 60.8 | 63.6 |
| Supervised AINET | 58 | 60.67 |
| A$^2$INET ($D = 0$) | 100 | 89.51 ($\sigma = 5.6$) |
| A$^2$INET best | 100 (with $D = 0.3, 0.4, 0.5$) | 99.40 ($\sigma = 0.4$) with $D = 0.6$ |

**Table 5**
Classification rates of adaptive immune models and standard classifiers applied to the synthetic data describing a nonlinear drift trajectory in the sensors space. Both alternate and blocked data sequences are considered.

| Non linear drift | Classification rates (%) | |
|---|---|---|
| | Alternate | Block-sample |
| PLS-DA | 58.3 | 57.87 |
| $k$-NN | 43.7 | 44.6 |
| Supervised AINET | 43.7 | 44.4 |
| A$^2$INET ($D = 0$) | 99.18 | 68.79 ($\sigma = 5.10$) |
| A$^2$INET best | 99.33 ($\sigma = 0.16$) with $D = 0.5$ | 99.55 ($\sigma = 0.22$) with $D = 0.8$ |

## 4.2. Real dataset

The properties of the experimental dataset are shown through the Principal Component Analysis. Fig. 9 shows the Bi-plot of the first two principal components. The drift trajectory in the real case is more complex with respect to the simulated data. The most evident difference is that the five classes do not undergo a common drift direction. However a limited class superimposition is observed. As a consequence the classification rates obtained by standard classifiers are considerably larger that those obtained with the simulated datasets (see Table 6). As shown in Table 3, the data presentation was rather irregular, with prolonged class absence during the experiment that lasted 18 months. The loadings show a similar behavior for two sensors (TGS2600 and TGS2620) while the other two sensors evidence a different contribution in the two principal components.

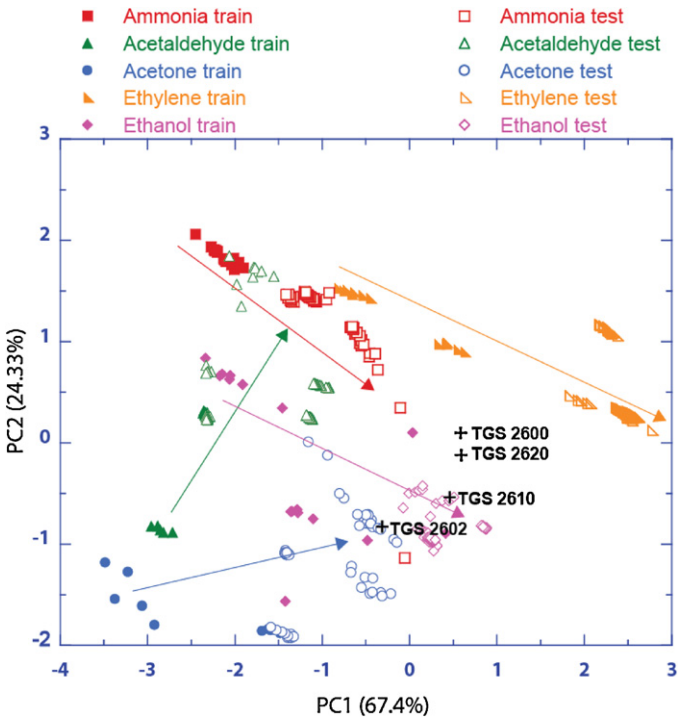


**Fig. 9.** Bi-plot of the first two principal components of the experimental dataset. Arrows indicate the drift direction of each class.

**Table 6**
Classification rates of adaptive immune models and standard classifiers applied to the experimental data.

| Experimental dataset | Classification rates (%) |
|---|---|
| PLS-DA | 73.73 |
| *k*-NN | 81.34 |
| Supervised AINET | 81.8 |
| $A^2$INET ($D=0$) | 84.19 ($\sigma=0.92$) |
| $A^2$INET best | 95.30 ($\sigma=0.38$) with $D=0.7$ |

The results achieved by $A^2$INET with the real dataset are qualitatively similar to those obtained with the simulated datasets. Fig. 10 shows the optimization of the parameter $D$. The relationship between classification rate and the parameter $D$ is similar to that observed with the synthetic data. The transition region in this case is more smooth, and it is extended from $D=0.9$ to $D=1.1$. The largest classification rate, and the smallest deviation among repeated trainings, is observed for $D=0.7$.

Fig. 11 shows the dynamic classification rate of the models calculated also in this case with a moving window of 100 samples. Optimal value of $D$, in the interval 0.7 and 0.8, achieves a constant classification rate except for the last part of the experiment. However, this is merely due to the fact that in the last period only the classes with the largest classification errors were measured.

A comparison between standard and modified immune models are shown in Table 6 together with the classification rates provided by standard, drift uncorrected, classifiers. Both the immune models reject part of the drift, but $A^2$INET better cope with the highly unbalanced data presentation that occurs in this experiment. About 10% of increases of correct classifications are provided by $A^2$INET with respect to the immune system without centroid normalization.

The results obtained with artificial and experimental datasets have put in evidence that the proposed algorithm is able to mitigate different kinds of drift using an unsupervised update of the classification model. This feature coupled with the capability of the model to add classes without losing the existent knowledge database represent the main differences and innovative aspects with respect to the drift counteraction techniques already published in literature. Nevertheless, a deeper investigation of the $A^2$INET performance has to be done with ad hoc experimental datasets comparing its results with those obtained with the standard drift mitigation algorithms.

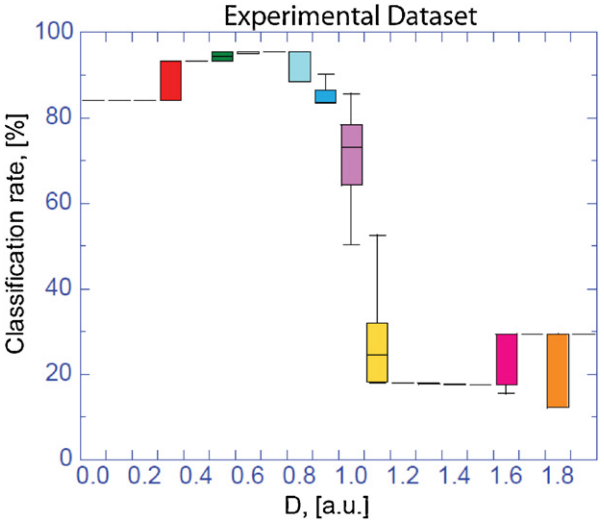


**Fig. 10.** Classification rate as a function of the parameter $D$ in the case of the experimental dataset. Performances are given as the box-plots of the distribution of the classification rates of 100 models replicas.

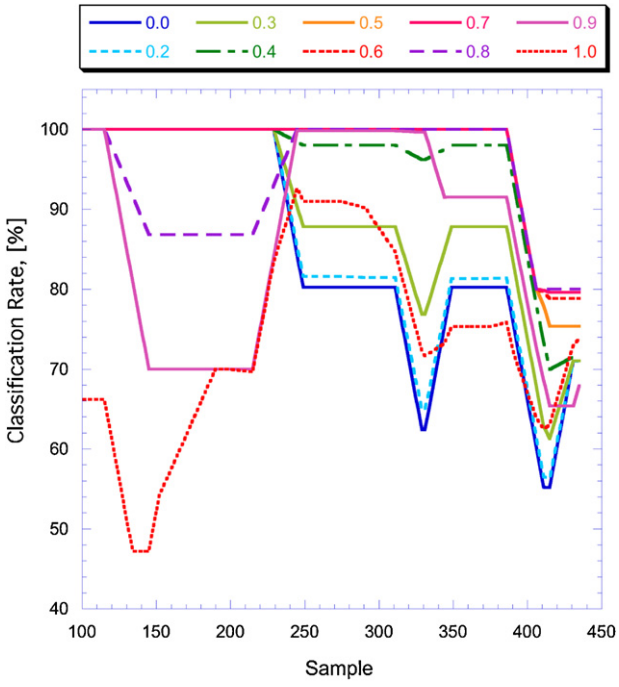


**Fig. 11.** Evolution of the classification rate along the experimental data and for different values of the parameter $D$.

## 5. Conclusions

In this paper an adaptive classification model called $A^2$INET, based on the biological immune system, is introduced and applied to artificial and real datasets. Such a classification model derives from a supervised AINET algorithm [12].

$A^2$INET has been conceived in order to provide a number of advantages particularly useful to process chemical sensor arrays data. The first advantage is the reduction of the model complexity that is achieved by an optimized templates selection. The reduced complexity does not jeopardize the properties of drift rejection. Indeed, the model can still actively follow the drift by changing or adding samples to the classes set (Fig. 2).

As common to other adaptive drift rejection methods, the results are strongly sensitive to the order at which the class membership of the samples occurs. Actually, these methods fail to correctly identify the drift when some of the classes do not occur, for a long time, in the samples sequence. It is worth to mention that this happens when the sample frequencies of the classes are not homogeneous.

It has been demonstrated in this paper, that a simple and potentially successful normalization can be used to adapt all the classes even those that are not actually presented to the sensor array. The effect of the normalization procedure is ruled by a single parameter that is required to be optimized in the presence of drift. This means that in real applications, a intermediate step between the usual training and test may be required to tune the normalization parameter.

The normalization parameter adapts the antibody updating to the speed of drift, namely to the time scale at which the sensor responses deviate from those recorded during the training. Then, once this parameter is optimized, drift rejection may be ensured until the speed of drift remains almost constant. To this regards, a strategy to continuously adapt the normalization parameter in the presence of a variable drift velocity is an important issue that will be subject of future investigations.

## References


[1] J.H. Holland, Adaptation in Natural and Artificial Systems: An Introductory Analysis With Applications to Biology, Control, and Artificial Intelligence, MIT Press, Cambridge, MA, 1992.
[2] A. Juarrero, Dynamics in Action: Intentional behaviour as a complex system, MIT Press, Cambridge, London, 2000.
[3] W. Cannon, The Wisdom of the Body, W.W. Norton & Company, Inc., New York, 1932.
[4] N. Wiener, Kybernetik, Econ-Verlag, Dusseldorf, 1963.
[5] T. Gross, B. Blasius, Adaptive coevolutionary networks: a review, Journal of the Royal Society Interface 5 (2008) 259–271.
[6] R. Ortega, M.W. Spong, Adaptive motion control of rigid robots: a tutorial, Automatica 25 (2006) 877–888.
[7] E. Munch, M.E. Launey, D.H. Alsem, E. Saiz, A.P. Tomsia, R.O. Ritchie, Tough, bio-inspired hybrid materials, Science 322 (2008) 1516–1520.
[8] F. Dressler, B.A. Ozgur, A survey on bio-inspired networking, Computer Networks 54 (2010) 881–900.
[9] B. Hannaford, K. Jaax, G. Klute, Bio-Inspired actuation and sensing, Autonomous Robots 11 (2001) 267–272.
[10] L.N. De Castro, J. Timmis, Artificial Immune Systems: A New Computational Intelligence Paradigm, Springer Verlag, London, 2002.
[11] E. Hart, J. Timmis, Application areas of AIS: the past, the present and the future, Applied Soft Computing 8 (2008) 191–202.
[12] L.N. de Castro, F.J. Von Zuben, Immune and neural network models: theoretical and empirical comparisons international, Journal of Computational Intelligence and Applications (IJCIA) 1 (2001) 239–257.
[13] S. De Vito, E. Martinelli, R. Di Fuccio, F. Tortorella, G. Di Francia, A. D'Amico, C. Di Natale, Artificial immune systems for Artificial Olfaction data analysis: comparison between AIRS and ANN models, in: Neural Networks (IJCNN), The 2010 International Joint Conference, 18–23 July 2010, 2010.
[14] C. Jacob, M. Pilat, P. Bentley, J. Timmis (Eds.), Artificial Immune Systems, Springer, Berlin, Heidelberg, 2005, number 3627 in LNCS.
[15] E. Hart, P. Ross, Exploiting the analogy between the immune system and sparse distributed memories, Genet. Prog. Evol. Mach. 4 (2003) 333–358.
[16] A. Hierleman, R. Gutierrez-Osuna, Higher-order chemical sensing, Chemical Reviews 108 (2008) 563–613.
[17] M.L. Salit, G.C. Turk, A drift correction procedure, Analytical Chemistry 70 (1998) 3184–3190.
[18] C. Sisk, N.S. Lewis, Comparison of analytical methods and calibration methods for correction of detector response drift in arrays of carbon black-polymer composite vapor detector, Sensors and Actuators, B: Chemical 104 (2) (2005) 249–268.
[19] E. Martinelli, M. Santonico, G. Pennazza, R. Paolesse, A. D'Amico, C. Di Natale, Short time gas delivery pattern improves long-term sensor reproducibility, Sensors and Actuators B 156 (2011) 753–759.
[20] A. Vergara, E. Llobet, E. Martinelli, C. Di Natale, A. D'Amico, X. Correig, Feature extraction of metal oxide gas sensors using dynamic moments, Sensors and Actuators B 122 (2007) 219–226.
[21] A. Vergara, E. Martinelli, E. Llobet, F. Giannini, A. D'Amico, C. Di Natale, An alternative global feature extraction of temperature modulated μ-hotplate gas sensors array using an energy vector approach, Sensors and Actuators B 124 (2007) 352–359.
[22] P. Boilot, E.L. Hines, M.A. Gongora, R.S. Folland, Electronic noses inter-comparison, data fusion and sensor selection in discrimination of standard fruit solutions, Sensors and Actuators, B: Chemical 88 (2003) 80–88.
[23] L. Carmel, S. Levy, D. Lancet, D. Harel, A feature extraction method for chemical sensors in electronic noses, Sensors and Actuators, B: Chemical 93 (2003) 67–76.
[24] B. Ehret, K. Safenreiter, F. Lorenz, J. Biermann, A new feature extraction method for odour classification, Sensors and Actuators B 158 (2011) 75–88.
[25] A. Ziyatdinov, S. Marco, A. Chaudry, K. Persaud, P. Caminal, A. Perera, Drift compensation of gas sensor array data by common principal component analysis, Sensors and Actuators B 146 (2010) 460–465.
[26] A.C. Romain, J. Nicola, Long term stability of metal oxide-based gas sensors for e-nose environmental applications: an overview, Sensors and Actuators B 146 (2010) 502–506.
[27] T. Artursson, T. Eklöv, I. Lundström, P. Mårtensson, M. Sjöström, M. Holmberg, Drift correction for gas sensors using multivariate methods, Journal of Chemometrics 14 (2000) 711–723.
[28] R. Gutierrez-Osuna, H.T. Nagle, A method for evaluating data-preprocessing techniques for odor classification with an array of gas sensors, IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics 29 (1999) 626–632.
[29] S. Marco, A. Ortega, A. Pardo, J. Samitier, Gas identification with tin oxide sensor array and self-organizing maps: adaptive correction of sensor drifts, IEEE Transactions on Instrumentation and Measurement 47 (1998) 316–321.
[30] M. Zuppa, C. Distante, P. Siciliano, K.C. Persaud, Drift counteraction with multiple self-organising maps for an electronic nose, Sensors and Actuators B 98 (2004) 305–317.
[31] E. Llobet, J. Brezmes, R. Ionescu, X. Vilanova, S. Al-Khalifa, J.W. Gardner, N. Bârsan, X. Correig, Wavelet transform and fuzzy artmap-based pattern recognition for fast gas identification using a micro-hotplate gas sensor, Sensors and Actuators B 83 (2002) 238–244.
[32] C. Di Natale, F.A.M. Davide, A. D'Amico, A self-organizing system for pattern classification: time varying statistics and sensor drift effects, Sensors and Actuators B 26–27 (1995) 237–241.
[33] S. Di Carlo, M. Falasconi, E. Sanchez, A. Scionti, G. Squillero, A. Tonda, Increasing pattern recognition accuracy for chemical sensing by evolutionary based drift compensation, Pattern Recognition Letters 32 (2011) 1594–1603.
[34] D.L. Massart, B.G. Vendeginste, S.N. Deming, Y. Michotte, L. Kaufman, Chemometrics: A Textbook, Data Handling in Science and Technology, Elsevier Science, The Netherlands, 1988.
[35] A. Vergara, S. Vembu, T. Ayhan, M.A. Ryan, M.L. Homer, R. Huerta, Chemical gas sensor drift compensation using classifier ensembles, Sensors and Actuators B 166–167 (2012) 320–329.
[36] P.S. Andrews, J. Timmis, On diversity and artificial immune systems: incorporating a diversity operator into aiNet, Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) 3931 LNCS, pp. 293–306.
[37] J.D. Farmer, N.H. Packard, A.S. Prelson, The immune system, adaptation and machine learning, Physica 22D (1986) 187–204.


## Biographies

**Eugenio Martinelli** is an assistant professor in electronics at the Faculty of engineering of the University of Rome Tor Vergata. His research activities are concerned with the development of chemical and biological sensors, artificial sensorial systems (olfaction and taste) and their applications, sensor interfaces and data processing. He authored more than 150 peer-reviewed papers on international journals and conference proceedings. He is member of the editorial board of Journal of Sensors and he is regular referee for a number of journals of the sector.

**Gabriele Magna** received an medical engineering degree from the University of Rome Tor Vergata in 2011. Currently he is a PhD student in Engineering of Sensorial and Learning Systems at the University of Rome Tor Vergata. His research interests are in chemical sensors and their data processing.

**Saverio De Vito** received his MS in Computer Engineering from University of Naples "Federico II", Italy in 1998. During 1998 He was with the image processing and understanding group at the Dipartimento di Informatica e Sistemistica of the above university working on breast cancer computer aided diagnosis. From 1999 to 2004 He was with a private ICT engineering company as an R&D team leader for telemedicine, earth observation and distance learning projects of ESA and ASI. In June 2004 he joined ENEA, as a full researcher. Since 2005 he is contract professor of Applied Informatics at University of Cassino. His research interests include artificial olfaction and electronic noses, statistical pattern recognition, and intelligent wireless sensor networks. He participated and was scientific responsible in several EU project in FP6 and FP7 programs as well as JTIs. He (co)authored more than 35 research papers and has served as referee for several international journals as well as for IJCNN, ISOEN and NaBIC. He is currently serving as MC member in EuNetAir Cost Action. Eng. De Vito is member of the IEEE and the International Association for Pattern Recognition (IAPR)-IC and is member of the XXVIII Italian research expedition in Antarctica.

**Raffaele Di Fuccio** received an medical engineering degree from the University of Rome Tor Vergata in 2011. Currently he is a research fellow in Institute of Cognitive Sciences and Technologies – National Research Council (ISTC-CNR). He works on funded European projects based on advanced learning technologies and data analysis.

**Girolamo Di Francia** received his degree in Physics from the University of Naples "Federico II". In 1985 he started his research activity in the field of fabrication and characterization of semiconductor solar cells (c-Si, GaAs), formerly in the Ansaldo comp. in Genova, and then in the ENEA research center of Rome, where he was appointed full time researcher in 1988. From 1991 he joined the ENEA research center of Naples where, starting from 1992, he investigated porous silicon based devices. In 1996 he established there the Gas Sensor Laboratory mainly devoted to the fabrication and characterization of devices based on nanomaterials and on polymers nanocomposites.

**Alexander Vergara** (Ph.D., 2006 – Universitat Rovira i Virgili) is a joint research associate working at the National Institutes of Health and National Institute of Standards and Technology (NIH/NIST) and a Visiting Researcher at the BioCircuits Institute, UC San Diego, where he was a Postdoctoral Researcher until summer 2011. His work mainly focuses on the use of dynamic methods for the optimization of micro gas-sensory systems and on the building of autonomous vehicles that can localize odor sources through a process resembling the biological olfactory processing. His areas of interest also include signal processing, pattern recognition, feature extraction, chemical sensor arrays, and machine olfaction.

**Corrado Di Natale** is an associate professor at the Faculty of Engineering of the University of Rome Tor Vergata where he teaches courses on Sensors and Detectors. His research activities are concerned with the development and application of chemical, bio-sensors, and artificial sensorial systems (olfaction and taste), and with the study of the optical and electronic properties of organic and molecular materials. He authored more than 420 papers on international journals and conference proceedings. He chaired the 9th International Symposium on olfaction and electronic nose (Rome, 2002) and Eurosensors XVIII Conference (Rome, 2004) and was member of the organizing committee of national and international conferences in sensors. He serves as component of the steering committee of the Eurosensors conferences series and as associate editor of IEEE Sensors Journal. He edited several proceedings volumes, and is a regular referee for a number of journals of the sector. In September 2006 he received the Eurosensors Fellow Award for his contribution in the fields of chemical sensors and artificial senses.