Abstract: The drift of metal oxide semiconductor (MOX) chemical sensors is one of the most important topics in this field. The work aims to test the performance of MOX gas sensors over the aging process. Firstly, sensors were tested with ethanol to understand their behavior and response changes. In parallel, beers with different alcoholic content were analyzed to assess what happened in a real application scenario. With ethanol analysis, it was possible to quantify drift of the baseline of the sensors and changes that could affect their responses over time (from day 1 to day 51). Conversely, the beer dataset has been exploited to evaluate how two different classifiers perform the classification task based on the alcohol content of the samples. A hybrid k-nearest neighbors artificial neural network (k-NN-ANN) approach and “standard” k-NN were used to evaluate to distinguish among the samples when the measures were affected by drift. To achieve this goal, data acquired from day one to day six were used as training to predict data collected up to day 51. Overall, performances of the two methods were similar, even if the best result in terms of accuracy is reached by k-NN-ANN (96.51%).

Keywords: MOX sensors; drift counteraction; aging; beer; k-NN; ANN

1. Introduction

Metal oxide semiconductor (MOX) gas sensors are characterized by high sensitivity, fast response, and low cost. Those characteristics make their application very promising in different fields, such as agri-food quality and safety [1–5], environmental monitoring [6–8], home security [9–11], and human health [12–14] among the most investigated. Although these features make MOX sensors one of the most promising technologies of recent years, their diffusion is limited due to some disadvantages. The most crucial is drift, which can change dramatically the information content of the features and cause a reduction of the performance over time.

In the last decade, various methods have been proposed to counteract sensors’ drift with good results. Among all those proposed, here some approaches that differ from one another are briefly presented to provide an overview that is as wide as possible. Orthogonal signal correction (OSC) was first applied to counteract gas sensors drift by Padilla, et al. [15]. They showed that sensors discrimination towards three gases at different concentrations was improved if compared to results obtained with non-corrected data. Vergara, et al. used an ensemble of classifiers based on support vector machine (SVM) where each SVM classification output was opportunely weighed [16]. Instead,

Chemosensors 2020, 8, 6; doi:10.3390/chemosensors8010006 www.mdpi.com/journal/chemosensors
Fornollosa, et al. used a master-slave approach with five twin sensors’ arrays for two reasons, but the most interesting was the mitigation of drift effect [17]. With the implementation of a multivariate technique called direct standardization, a multivariate technique that allows mapping an unknown space (in this case, the space from the slave unit) to a reference space (the one from the master unit), they were able to maintain the prediction error stable over time. Finally, Magna, et al. developed a system that improves significantly classification accuracy combining a static feature selection during the training phase and a dynamic feature selection during tests [18].

This study aimed to assess the possibility of using MOX sensors as soon as they are produced, skipping the standard aging step that usually lasts many hours or days, depending on the typology of the sensor [19]. Within the production process, aging is a phase that is induced and controlled to completely oxidize the material and to stabilize it. Hence, the instability for aged sensors is less impactful than for as-fabricated devices [20]. As examples, nanowires made of CuO may need up to a period of two weeks in ambient air at 300 °C to complete the aging [21], whereas it has been proved that gallium indium zinc oxide is more stable after being at 250 °C for 65 h [22].

This approach wanted to quantify the drift that affects the sensors and the changes in the response for 51 days after their production. During this period, sensors were kept at high temperatures for their correct working. In that way, a sort of aging has been induced in them, but not controlled, since in the same time interval measurements of ethanol and beers have been taken. Four types of sensors were considered and two or three replicas of the same typology of the sensor were used. These sensors were placed inside an S3 (small sensor systems), an e-nose-like device, to exploit its advantages: automatic execution of samples under analysis (once set) and transmission of data to a dedicated web app.

The datasets thus obtained were used to assess two machine learning algorithms regarding the ability to counteract drift in the “early life” of sensors and to assure reliable performances over time. On one hand, measures with ethanol have been elaborated to evaluate drift and changes in the sensors’ responses. On the other hand, a k-nearest neighbors (k-NN) approach has been used alongside a hybrid method that combines k-NN with artificial neural networks (ANN) for the “beer dataset”. The aim was to assess if the application of one of the most used computational systems, i.e., ANN, could enhance overall performances. To the best of our knowledge, such a machine learning approach has not been applied to chemical sensors, although it has been shown to yield promising results in other fields, such as classification of clouds through images [23] or disease like diabetes and cancer [24]; finally, it differs from the method proposed by [25], i.e., an ensemble of k-NN and ANN classifiers used to counteract sensors drift.

2. Materials and Methods

2.1. S3 Device

S3 is a device developed by the collaboration of Sensor Lab at the University of Brescia and Nasys S.r.l., a spin-off of the same University. S3 was previously used in numerous studies with considerable success; particularly, it has been applied to the field of food technology and quality control [26–28]. S3 consists of an electronic part to manage the signal, an element that allows its connection to the internet, a pump to bring the volatile compounds to the heart of the instrument from an autosampler HT2010H (HTA s.r.l., Brescia, Italy). It supports a 42-loading site carousel that can reduce the possible variables due to the preparation of the sample.

The sensors grown and developed for this work were created with two techniques, rheotaxial growth thermal oxidation (RTGO) and nanowire technology. The RGTO technique involves two deposition steps: the first stage of a metallic thin film by DC magnetron sputtering from a metallic target on a substrate at higher temperatures than the melting point of the metal, then the thermal oxidation cycle to get a metal oxide layer with stable stoichiometry [29]. In this work, tin oxide RGTO and tin oxide RGTO functionalized with gold particles were put in the chamber of S3. Nanowires (NW) exhibit exceptional crystalline quality and a very high length-to-width ratio, consequent in enhanced
sensitivity as well as long-term material stability for prolonged operation [30,31]. The experimental process consists of the evaporation of the powder (metal oxide) at high temperatures in a controlled atmosphere at pressures lower than hundreds of millibar and the following mass transport of the vapor towards the substrates kept at lower temperatures concerning the source evaporation region. The list of the 10 sensors produced at Sensor Lab and used for this study are presented in Table 1.

<table>
<thead>
<tr>
<th>Material (Type)</th>
<th>Number of Replicas</th>
<th>Morphology</th>
<th>Working Temperature (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SnO2Au (n)</td>
<td>3 sensors</td>
<td>RGTO</td>
<td>450</td>
</tr>
<tr>
<td>SnO2 (n)</td>
<td>3 sensors</td>
<td>RGTO</td>
<td>450</td>
</tr>
<tr>
<td>SnO2Au (n)</td>
<td>2 sensors</td>
<td>Nanowire</td>
<td>350</td>
</tr>
<tr>
<td>CuO (p)</td>
<td>2 sensors</td>
<td>Nanowire</td>
<td>350</td>
</tr>
</tbody>
</table>

2.2. Experimental Set-Up

Eight sessions of measures were realized within 51 days. The first session has been done before any aging of the sensors; other sessions were carried out at the following days: 2, 6, 8, 26, 30, 37, and 51. After the first measure, sensors were kept at their operating temperature throughout the measurement period. Two different tests were designed for each day of measure, waiting a period of 1 h among them. The first test was made with 300 ppm of ethanol (dry air), repeating the measures 10 times each day; the second one with four different beers. Beers differed for the alcohol content: two were nonalcoholic beers, having 0.0% and 0.5% of alcohol by volume (ABV); the other two had an ABV of 4.7% and 7.9%. The choice to analyze this type of product is mainly due to understanding the response of the sensors in a possible application scenario, bearing in mind that ethanol is one of the most volatile compounds that characterize it. Proceeding in this way, it was possible to make a comparison between the two tests. Every single sample was analyzed for 15 min: sensors were exposed to ethanol and beer VOCs for 2 min; 13 min were necessary to restore the baseline with air.

2.3. Data Analysis

The measures obtained in the 8 days have been divided to form two different datasets. The first one contained the analysis carried out using ethanol. From this dataset, two kins of information have been extrapolated: how sensors drifted and how much their response changes in the first 51 days after their production. As an indicator of the response, the difference between the first value of the acquired resistance ($R_0$) and its minimum value (for n-type sensors) or the maximum value (p-type sensors) during the analysis time was calculated; the difference was divided by $R_0$ to get the percentage variation respect to the baseline. Hence, $\Delta R/R_0$ has been extracted as a feature for the following analysis.

The second dataset (“beer dataset”) included an analysis of the four beer samples during the same period. In this case, the effect of drift in classification tasks has been evaluated. Principal component analysis (PCA) was used as an unsupervised technique to visualize data, reduce dimensionality and assess if the drift affected the four samples in the same way. From the PCA score plot, it has been decided to use the k-NN algorithm to obtain the accuracy of the system, using Euclidean distance as a metric. The possibility to use a hybrid algorithm based on k-NN was also investigated to try to enhance classification performances. In particular, a method similar to the one described in [23] has been used. The main difference stands in the use of “scaled conjugated gradient” as a learning algorithm instead of “extreme learning machine”; furthermore, different architectures of ANN were investigated to find the best one in terms of classification performances. This approach consists of 5 steps.

1. ANNs were trained using data from day 1 to day 6. Six different ANNs were tested: two of them had 1 hidden layer of 10 neurons and differed for the activation function, i.e., ReLu and
hyperbolic tangent sigmoid; the other four networks had 2 layers of 10 and 7 neurons and a combination of the abovementioned activation functions was used.

2. For the jth sample of the test set, the distance from the samples of the training set was calculated using Euclidean distance.

3. k-nearest neighbors were chosen; then the input for the ANN was calculated with the following formula:

\[
\text{Input} = \left[ \sum_{i=1}^{k} \frac{F_1^i}{k}, \sum_{i=1}^{k} \frac{F_2^i}{k}, \ldots, \sum_{i=1}^{k} \frac{F_n^i}{k} \right]
\]

where \(F_1, F_2, \ldots, F_n\) are the features that describe samples. Hence, the input of the network is a vector that has as features the mean value of the same parameters of the k-nearest neighbors.

4. ANNs were applied and the class of the sample was predicted.

5. Steps from 1 to 4 were repeated for all samples in the test set.

Data analysis has been performed with Matlab® R2015a software (MathWorks, Natick, MA, USA) and Python.

3. Results and Discussion

In Figure 1 sensors’ outputs at 300 ppm of ethanol for each typology of sensor are reported. One sensor for each type has been chosen; in particular, those that will be named as “sensor #1” from here on. For each day, the plotted signal represents the mean value of the ten measures done in that day. The two phases of the measure are clearly portrayed: during the first one, sensors are exposed to gas and hence their resistance decreases for n-type (SnO\(_2\)) and increases for p-type (CuO); in the second phase, pure air is fluxed into sensor and the baseline is restored.

Figure 1. Outputs from all sensors’ typologies to 300 ppm of ethanol, from day 1 to day 51. Each signal is the mean of the ten signals acquired the same day. (a) SnO\(_2\) Au RGTO sensor #1; (b) SnO\(_2\) RGTO sensor #1; (c) SnO\(_2\) NW sensor #1; (d) CuO sensor #1.
Figure 2 shows the trend of baseline from day 1 to day 51 for all the sensors, divided for typology. Results are presented in terms of mean value, with the error bars representing the standard deviation. Sensors of the same material and morphology are grouped to compare their behavior. Regarding SnO$_2$Au RGTO sensors, two of them had a stable baseline over time, whereas the third one started increasing after 8 days and reached a mean value at day 51 that is 53.67% higher respect to the initial value. On the contrary, SnO$_2$ RGTO sensors were all affected by a drift that caused an increase of the baseline at day 51 of 2.58 times for sensor #1 to 2.76 times for the other two respect to the value of the day 1.

Best results have been obtained with NW sensors. Indeed, the ones made of CuO had a resistance value in air that increased between 30.28% and 24.55% during the observation period, while the growth for tin oxide NWs between 105.37% and 76.44%.

In Figure 3, $\Delta R/R_0$ is reported. These values were taken as indicators of the responses of the sensors. From the data obtained, it results that both typologies of RTGO sensors have a response that is quite stable for 51 days, oscillating between 0.88 and 0.96 for SnO$_2$Au and between 0.81 and 0.89 for SnO$_2$. Regarding SnO$_2$Au RGTO, sensor #2 has not been included in Figure 3, but its response was lower than 0.08 for this type of measure. Conversely, NW sensors have different trends. Tin oxide NWs are characterized by an increase of their resistance variation in day 51 in respect to day 1, and fluctuation of the same parameter up to day 37. Finally, CuO sensor #2 has a behavior more similar to RGTO in terms of reproducibility, whilst the same feature for CuO sensor #1 has an oscillating trend from 0.18 to 0.78, that cannot be considered a linear drift if compared with the other sensors of the array.
Figure 3. $\Delta R/R_0$ for all the sensors from day 1 to day 51. (a) Response of 2 out of 3 SnO$_2$Au RGTO sensors; (b) response of the 3 SnO$_2$ RGTO sensors; (c) response of the 2 SnO$_2$ NW sensors; (d) response of the 2 CuO sensors.

For the “beer dataset”, PCA was performed as explorative analysis to visualize the drift that affected beer measures. In Figure 4, the PCA score plot of the first two principal components (PCs) is shown. The total explained variance (EV) is equal to 90.5% (87.08% for PC1 and 8.82% for PC2), meaning that most of the information is contained in two variables. A clear distinction between the four beers classes is achieved considering samples scores along PC1. However, the distance among the clusters tends to decrease as the ABV increases. This result could be explained by the fact that alcohol amount of 4.7% and 7.9% beers brings sensors near to saturation, making the distinction between these two classes less clear. On the contrary, on the PC2 axis drift can be seen. Additionally, in this case, different considerations can be done for the two lower and the two higher ABV beers. The entity of the dispersion for 0.0% and 0.5% samples is weak. Nevertheless, a separation between data acquired the first 6 days and those obtained from day 26 onwards can be observed. For the other two classes, the day-to-day difference is more pronounced within the same class. The same observation regarding measures before and after day six is valid in this case, too. The main cause of this result is due to the drastic change in CuO #1 sensor response that drops from the maximum value of 0.78 to 0.54. The subsequent discontinuous behavior is mitigated by the stable response recorded from the other sensors in the array, especially RGTO ones.

From PCA results, the k-NN algorithm seemed to be a good candidate to evaluate the classification performance of the sensors. It has been decided to use data from the first six days as training set (45 samples) and data from day 26 to day 51 as test set (86 samples). The aim was to assess the correct classification of the system when trained only with few initial patterns and forced to recognize patterns temporally distant belonging to the same classes. Hence, the intent was to verify whether the methods proposed succeeded in reducing the negative effect due to sensor drift. Side by side with k-NN, a hybrid algorithm that combines k-NN and ANNs has been investigated. The main target was to
exploit the non-linearity of the ANNs to improve the performance of k-NN. Both the approaches were tested varying the number of $k$ neighbors from 1 to 15.

From PCA, a subset of 3 features has been chosen to optimize the classification task. In detail, SnO$_2$Au RGTO #1 and the two SnO$_2$ NW sensors have been selected. Percentages of accuracy are shown in Figure 5. k-NN performances have been highlighted with a thicker and dashed line to better understand if the ones of the hybrid method could get over them.

**Figure 4.** PCA score plot for beer samples. Total explained variance (EV) equal to 95.9% from the first two principal components (PC1 EV = 87.08%, PC2 EV = 8.82%). Colors indicate the different ABV of the samples, markers denote different days (from 1 to 51).

**Figure 5.** Accuracies of k-NN and the different ANNs structures for the combined method. Among square brackets the number of neurons per layer is indicated; $tanh$ denotes the hyperbolic tangent sigmoid activation function, $relu$ is rectified linear unit function.

k-NN-ANN accuracies are consistent with those obtained with k-NN since there is no significant difference between the different percentages. However, with the increasing number of neighbors,
the combined method reduces performances in some architectures, especially those with ReLu in the last hidden layer. This reduction of performance is caused by the misclassification among the samples of classes of 4.7% and 7.9% ABV beers. It can be assumed that the class boundaries learned by the ANN were degraded when $k$ becomes higher: indeed, in this case, almost samples of the two classes used in training are considered as a reference for the calculation of the distance and ReLu function is not able to learn correctly. The only exception is represented by the ANN with ReLu and hyperbolic tangent sigmoid as activation functions of the first and the second layer respectively (cyan line). In this case, the best result is obtained with $k = 15$ reaching a value of 96.51%, which is also the highest accuracy achieved. Generally, it is possible to affirm that both methods are able to learn the classes of the sample using a small part of the dataset and to assign the measure used as test to the correct belonging class, thus limiting the negative effect of drift.

4. Conclusions

This study aimed to characterize how sensors, just produced and without induced aging, drift and how it is possible to counteract it. Ten different sensors divided into four groups were tested (RGTO and NW, tin oxide and copper oxide) with ethanol and beers with different alcohol content. Two parameters were considered: baseline drift and changes in resistance variation when exposed to gases. From ethanol measures, it resulted that NW sensors (both tin and copper oxide) had the most stable baseline during the period of analysis (days 1–51). Conversely, RGTO sensors showed more reproducibility over time.

One hybrid approach that combines k-NN and ANNs has been used to evaluate the possibility to counteract drift, using the dataset containing beers measures. Performances were compared with the k-NN algorithm. It resulted that the two approaches do not differ from each other in terms of accuracy, although the best classification result was achieved with the hybrid method.

**Author Contributions:** Conceptualization, V.S., E.C., and G.S.; methodology, E.N.-C., and V.S.; software, M.A.; validation, M.A., E.N.-C., V.S., E.C., and G.S.; formal analysis, M.A.; investigation, M.A., E.N.-C., and V.S.; resources, E.C. and G.S.; data curation, M.A.; writing—original draft preparation, M.A.; writing—review and editing, E.N.-C, V.S., E.C., and G.S.; supervision, E.C. and G.S. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Conflicts of Interest:** The authors declare no conflict of interest.

**References**

5. Galstyan, V.; Bhandari, M.P.; Sberveglieri, V.; Sberveglieri, G.; Comini, E. Metal Oxide Nanostructures in Food Applications: Quality Control and Packaging. *Chemosensors* 2018, 6, 16. [CrossRef]


27. Abbatangelo, M.; Nuñez-Carmona, E.; Duina, G.; Sberveglieri, V. Multidisciplinary Approach to Characterizing the Fingerprint of Italian EVOO. *Molecules* 2019, 24, 1457. [CrossRef] [PubMed]

