Abstract

The performance of any electronic nose is ultimately determined by the properties of its constituent parts (e.g., the sensors, signal processing and pattern-recognition engine). Electronic noses currently exploit the technologies of several classes of sensor material (e.g. semiconducting oxide, conducting polymer, phthalocyanines and lipid coatings) as well as a variety of pattern-recognition paradigms (e.g. back-propagation, self-organizing map and discriminant function analysis). Consequently, there is a need to compare objectively the performance of the increasing number of both research and commercial electronic noses. This paper addresses this problem and suggests the need for odour standards to quantify both the ability of an electronic nose to discriminate between similar odours (i.e. its 'resolving power') and a number of dissimilar odours (i.e. its 'range'). We present a generic model from which we can define these two fundamental parameters, and hence develop a benchmark for the performance of these different electronic noses against two proposed odour standards. This model can be employed not only as a design tool to predict the performance of an electronic nose against an odour standard, but also as a diagnostic tool that can determine, for example, the effect of random errors in the sampling method, sensor characteristics or the effect of systematic errors associated with sensor drift or changes in ambient temperature. We believe that our definition of odour standards and performance parameters for electronic noses could be used to create a European standard, which is now required in this rapidly expanding field and marketplace.

Keywords: Electronic noses, Odours, Sensor arrays

1. Introduction

A considerable amount of research has been directed towards the development of electronic-nose instrumentation over the past decade. Numerous research groups now exist in countries such as Australia, Denmark, France, Germany, Japan, Sweden, UK and USA. The general approach has been to develop an array of sensors based upon a single class of material and type of transducer in a monotype electronic nose. This approach has led to the arrival of three commercial instruments that currently dominate the world market, namely, the Fox 2000 (Alpha MOS SA, France), Digital Aroma Analysis System (Aromascan Plc, UK) and the NOSE (Neotronics Scientific Ltd, UK) [1]. A number of other instruments are also under development and we may see 10 or more commercial instruments by the end of this decade.

The main aim of this paper is to consider how the ability of an electronic nose to discriminate between odours can be standardized across the industry. Fig. 1 shows a general representation of multidimensional odour space $\Omega_2$. The dotted lines demarcate the seven 'primary' aromatic odours, i.e. camphoraceous, floral, etc. (A simple classification system proposed by Amoore [2] is used to illustrate our model. The concept of a primary odour is perhaps oversimplistic and is still under debate.) An electronic nose may be designed to work in all of these regions or, more realistically, a subset of them relevant to the needs of a particular end user. A set of primary odours could be chosen (labelled A) to test the ability of an electronic nose across odour space and thus determine its range. Secondly, and more importantly, a second set of odours could be chosen, which in each case lie in close proximity in odour space to odour A (labelled B). The ability of an electronic nose to classify these known odours will determine both its resolving power and range. The resolving power is the more important parameter...
because users need to discriminate between similar smells arising from the same product, rather than, say, discriminating between a floral and musky perfume.

2. Electronic nose model

Fig. 2 shows the basic stages of signal processing in an electronic nose. Sample space $\Omega_1$ is defined by either the chemical composition of the sample, which may be a single compound (e.g. 100 $\mu$g/l of (+)-limonene in water for a simple citrus (lemon) odour) or, and more likely, a number of compounds (e.g. a combination of citrus compounds such as undecyl alcohol (lime) with limonene). Although some odours can be engineered chemically in this manner, many contain tens if not hundreds of compounds and so the sample space must be defined in terms of the headspace composition (e.g. 100% Brazilian coffee bean, or 50% Brazilian and 25% Colombian coffee beans for a blend) diluted in a carrier gas. The zero point in sample space represents the baseline value of the sensor array in odour-free air or a calibration standard (e.g. the headspace from a solution of 5% ethanol in pure deionized water held at 300°C). The sensor array responds to the sample A and maps one point in sample space onto sensor space $\Omega_2$. The choice of sensor parameter (e.g. the fractional change in conductance) and data preprocessing is crucial and broadly determines the resolving power of the electronic nose. In the human olfactory system, the front-end processing is elaborate and manages to boost the sensor sensitivity to the sub-ppb level and yet, at the same time, suppresses the signal noise. The condition of the signal from the sensor array thus largely determines the performance of an electronic nose. Key parameters are the number and type of sensors, the specificity of the sensors, their noise level, stability, thermal sensitivity and so on. In other words, if the samples do not create distinct vectors $S_j$ in sensor space, then the pattern-recognition (PARC) engine cannot be expected to discriminate them. Any attempt to do so will be unproductive as the noise in sensor space will be mapped as uncertainty in the classification space $\Omega_3$. For example, it is possible to boost the ppm sensitivity of commercial tin oxide sensors to the ppb level through electronic circuitry [3] but the concurrent increase in noise and drift provides no benefit in resolving power. Therefore, we assume in this model that the performance of the sensor array dictates the overall performance of the nose and that any limitations due to the accuracy of the analogue-to-digital converter (ADC) [4] are not fundamental to the problem. So provided replicate samples form distinct clusters in sensor space, then the PARC method should be able to classify them by employing a method (such as metric transformations, back-propagation or self-organizing map) appropriate to the nature of sensor space, e.g. linear in concentration, high degree of sensor colinearity, etc. Previous work has shown that the function that maps sensor space onto classification space is reasonably well behaved for sensor arrays based upon oxide, polymer and lipid materials (see references in [1] for details).

2.1. Model of sensor array

The exact relationship between sample space and sensor space clearly depends upon the type of sensor used. In order to develop an analytical model we shall assume an empirical power-law dependency, so the output $S_i$ from a sensor $i$ in an array of $n$ sensors can be modelled as

$$S_i = a_{i0} + a_{i1} x_1^k + a_{i2} x_2^k + \cdots + a_{im} x_m^k$$ (1)

where there is a regression coefficient $a_{ij}$ (sensitivity factor when $j \neq 0$) for each of the $m$ compounds present in the sample, and $a_{i0}$ is the baseline output of the sensor in the reference air. $x_j$ represents either the concentration of compound $j$ (ppm or $\mu$g l$^{-1}$) or percentage of odour present in the sample. The exponent $k_j$ takes a value of 1 for a linear sensor but is typically in the range 0.5–0.9 for semiconducting oxide conductance sensors. (In fact this is...
inadequate to describe the full working range of polymer sensors and so in the computational model discussed later we have extended the choice to include a Langmuir adsorption isotherm that models a sorption process. The model assumes that the compounds do not react with each other, which is normally the case for polar organic molecules (and chosen to be the case when defining our standard odours).

The general response $S$ of the entire array can be written in matrix form as $AX + B$ or explicitly

$$
\begin{pmatrix}
S_1 \\
S_2 \\
\vdots \\
S_n
\end{pmatrix}
= 
\begin{pmatrix}
a_{11} & a_{12} & \cdots & a_{1m} \\
0 & a_{22} & \cdots & a_{2m} \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & a_{nm}
\end{pmatrix}
\begin{pmatrix}
x_1^k \\
x_2^k \\
\vdots \\
x_m^k
\end{pmatrix}
+ 
\begin{pmatrix}
a_{10} \\
a_{20} \\
\vdots \\
a_{n0}
\end{pmatrix}
$$

(2)

For an array of perfectly specific (orthogonal) sensors, the regression matrix $A$ reduces to the diagonal matrix as shown in Eq. (3). This sensor system can perform multi-component analysis ($m \leq n$) but clearly lacks the range of a human nose (several thousand smells) as it could only identify a limited number of odours (albeit of varying odour quality scale).

$$
A = 
\begin{pmatrix}
a_{11} & 0 & 0 & 0 \\
0 & a_{22} & 0 & \vdots \\
0 & 0 & \cdots & a_{nm}
\end{pmatrix}
$$

(3)

In practice, gas sensors are not completely orthogonal, which means that some of the off-diagonal terms are non-zero and this increases the range of the system while decreasing the resolving power through overlapping vectors. Thus, in principle, hundreds of different odours could be discriminated with a small set of non-specific sensors.

2.2. The role of errors

In an ideal sensor system of, for example, a set of six orthogonal linear sensors, there are no errors and so in effect there exists an infinite number of distinct response vectors (in practice restricted only by the full-scale deflection of the sensor and the ADC resolution). Vectors that are very close together can be resolved by simply changing the distance metric (e.g. to Mahalanobis) and thus enhancing the signal gain. In practice, there are a number of sources of error that will determine the number of samples that can be discriminated in sensor space and hence the performance of the electronic nose. Errors can arise from the variation of the concentrations of the components in the sample and the relative concentration of components due to differential evaporation, etc. These may be due to variation in the sampling method or due to natural variation in the sample itself. For example, Mackay-Sim et al. have identified sources of experimental variability for piezoelectric odour sensors [5]. Clearly, an odour standard must be defined with a concentration variation below a specified limit of, say, 1% and be used with a standard sampling procedure (e.g. standard temperature, pressure, a flow rate of 100 ml min$^{-1}$ and a humidity of 50%) to compare reliably different electronic noses. Variations can also occur in the sensor coefficients $a_{ij}$ ($j \neq 0$), which may be random in nature due to materials variability or systematic due to changes in, say, temperature or humidity. Indeed, in some cases deliberately changing the sensor's operating temperature can aid the resolving power of the array [6]. In addition, a variation in the baseline parameter $a_{i0}$ could represent a systematic drift due to the ageing or poisoning of the sensor. In this case the sensor output will follow a drift line and there is likely to be a correlation between the regression coefficients. All of these cases can be described within a sensor error model (error on $k_i$ is added for completeness below) such as that given in Eq. (4).

$$
S_i(1+e_{S_i}) = a_{i0}(1+e_{a_{i0}}) + \sum_{j=1}^{m} a_{ij}(1+e_{a_{ij}})(x_j(1+e_{x_j}))^{k_i}
$$

(4)

where $e_{a_i}$ and $e_{x_i}$ represent the fractional errors on the sensor coefficients and concentration, respectively.

2.3. Resolving power of 'stationary' noses

Fig. 3 illustrates the model of an electronic nose with two samples labelled A and B with error in sample space

Fig. 3. Schematic diagram showing the propagation of sample and sensor errors through a computational electronic nose. Vectors are either classified as odours (in this case by plotting the first two canonical variables) or used in multivariate analysis (MVA) to predict the original component concentrations $x_j$. 

producing two response vectors with error in sensor space \( \Omega_b \) with separation \( S_{AB} \). For randomly distributed errors, each response vector falls into a hyperellipsoid with the axes given by the standard deviations of the sensor outputs. The variance on the sensor signal \( \sigma^2_x \) can now be related to the variance on the sensor coefficients and component concentrations from Eq. (4).

\[
\sigma^2_{A_i} = \sigma^2_{A_{in}} + k_i \sum_{j=1}^{m} \sigma^2_{a_{ij}} + k_i \sum_{j=1}^{m} \sigma^2_{e_{ij}} x_j^2
\]

(5)

where the first term contains the error in the baseline sensor signal, the second term contains the component concentration errors and the last term contains the errors in the sensitivity of the sensor to \( m \) different components.

A statistical parameter \( RP \) may now be defined to quantify the power of a nose to resolve two similar odour samples (see A and B in Fig. 1).

\[
RP = \frac{|S_{AB}|}{\sqrt{\sigma^2_A + \sigma^2_B}}
\]

(6)

where \( |S_{AB}| \) is the separation of response vectors (see Fig. 3) and \( \sigma_A \) and \( \sigma_B \) are the errors calculated along the direction of the distance metric. This parameter is essentially that referred to by Müller as the ‘electronic selectivity’ [7].

The volume that the error on the response vector \( \sigma_x \) occupies in sensor space determines both the local resolving power \( RP \) of the electronic nose and the maximum number of odours that can be classified as distinct (see below). The volume \( V_n \) of a hyperellipsoid comprising a set of \( n \) sensors with semiaxes defined by the standard error is

\[
2\pi^{n/2} \prod_{i=1}^{n} \sigma_x S_i
\]

(7)

The use of a hyperellipsoid model rather than a hyper-cuboid model is significant when considering a large sensor array as illustrated in Table 1. The ratio indicates the reduced volume that a hyperellipsoid occupies compared to a hypercuboid calculated from the prefactor in Eq. (7) and its hypercuboid equivalent.

The maximum number of response vectors \( N_n \) (i.e. a measure of the range) that can be discriminated (this assumes that the FARC engine is error-free) in sensor space is simply given by

\[
N_n = \frac{\text{Volume of sensor space}}{\text{Volume of sensor error}} = \frac{\prod_{i=1}^{n} \text{FSD}(S_i)}{\sum_{i=1}^{n} \text{FSD}(S_i)}
\]

(8)

where the FSD(\( S_i \)) is the full-scale deflection of the output signal \( S_i \) from sensor \( i \), and \( V_n \) is defined in terms of the model parameters and their errors by Eqs. (5) and (7). For sensors that operate over a finite FSD, the error volume determines the number of samples that can be discriminated. The volume of the sensor error, \( V_n \), is a function of the sensor output and so it is not in general uniform over sensor space.

An error on the regression coefficients or material sensitivity causes a concentration-dependent error volume, i.e. the error volume monotonically increases with increasing component concentration, see the third term in Eq. (5). To illustrate this, Fig. 4 shows the concentration dependence of the number density (number of distinct response vectors per unit sensor output) due to a high 10% standard error on the regression coefficient \( a_{ij} \) of linear (\( k = 1 \)) and non-linear (\( k = 0.5, 2.0 \)) sensors. The number

![Graph showing concentration dependence of number density](image-url)
density \( N' \) (using pre-factor for \( n = 1 \)) of a specific sensor (this is similar to the one-dimensional parameter proposed by Mülller [7]) was calculated using,

\[
N' = \frac{1}{V_1} = \frac{2}{\sqrt{a_0^2 + k_1a_0^2 + k_2a_1^2 + k_3a_2^2 + k_4a_3^2}} \quad (9)
\]

Here the number density of the sensor system falls with increasing concentration. Note that making the sensor non-specific generates additional error terms (see Eq. (5)) from the other component concentrations, which further increase the fall of the number density with concentration. The number density is also modified by the exponent \( k_i \) and the relative size of the three error terms in the denominator of Eq. (9). The number density of an array of any size can be determined from Eq. (7) and when integrated over the hypervolume gives the maximum number of distinct vectors \( N_\mu \).

It is thus possible to compute the error volume, maximum number of distinct vectors and number density of any sensor array from the values of sensor model parameters, component concentrations and their associated errors.

### 3. Definition of odour standards

It should be noted that the number of odours \( M_\mu \) that can be discriminated as distinct from the number of response vectors \( N_\mu \) depends upon the definition of classification space, \( Q_3 \). The organoleptic response of an odour can be classified on a six-point odour-quality scale (such as that formatted by the Institute of Olfactory Sciences, 211 Tampa St., Park Forest, IL 60466, USA, in 1977) as shown in Fig. 5. For simple odours this means that all concentrations below the olfactory threshold (e.g. \( 1 \mu g/1 \) for (+)-limonene) are classified as ‘odourless’ while all concentrations above a certain level are classified as ‘extremely odorous’, etc. The values of the threshold and scales vary from odour to odour. Consequently, the number of classes of odour that can be discriminated will be related to \( N_\mu \) but usually a much lower value (the exception being when the electronic nose is very bad!).

The range of an electronic nose can be tested by the use of a number of very different odour standards (labelled A in Fig. 1) consisting of commercial products with known organoleptic properties. For example, Table 2 lists 29 distinct odour descriptors (this is a superior set of odour descriptors to the seven proposed by Amoore [2]) which are associated with chemical compounds commercially available from Aldrich [8]. Thus, a standard odour of type A can be defined as a specific concentration of a compound in pure water (e.g. 1 ppm of diacetyl in water for butterscotch). A number of different type-A odour standards can then be selected (e.g. 29) and used to quantify the range of an electronic nose by their correct classification. Each standard will also have a specific concentration error (e.g. 1%) so that any error in the sampling system is apparent to the users.

More importantly, there is a need to discriminate between similar odours that lie close to A in odour space (labelled B in Fig. 1). Again these can be selected as single chemical compounds. For example, citrus odours can be classified into type-A lemon ((+)-limonene) and type-B lime (undecyl alcohol). More sophisticated noses could classify the samples into the six-point odour-quality scale rather than the binary scale of odourless/odorous. In this particular example, more difficult tests could be used whereby odour-standard B could be another compound that produces a lemon citrus note, such as fencyl alcohol. The resolving power of an electronic nose can now be defined as the value of the test parameter RP (defined by Eq. (6)) for standard odours A and B associated with a specific olfactory descriptor.

### 4. Computational nose

#### 4.1. Sensor model

Although it is possible to obtain analytical expressions with which to define the performance of an electronic nose, the situation is rather complicated when modelling sensor arrays with non-zero off-diagonal regression coefficients, random sampling and sensor coefficient errors.
and systematic errors due to such as temperature changes, ageing, etc. Consequently, we have designed a computational model of an electronic nose with which to evaluate the performance of a generalized sensor array. The sensor model now offers the choice of a Langmuir isotherm in which adsorption sites can be saturated as well as conventional linear and power-law responses. In addition, the exponents $k_i$ can also have errors $M_k$. Thus the sensor model of Eq. (4) is extended to become

$$S_i(1+\sigma_{s_i}) = a_{i0}(1+\epsilon_{a_{i0}}) + \sum_{j=m}^{j=m} a_{ij}(1+\epsilon_{a_{ij}})x_j(1+\epsilon_{x_j})^{k_i(1+\epsilon_{k_i})}$$

$$+ \sum_{j=m}^{j=m} b_{ij}(1+\epsilon_{b_{ij}})x_j(1+\epsilon_{c_{ij}})$$

Now the sensor parameters can be set for the appropriate choice of sensor (e.g. semiconducting oxides, conducting polymers). The effect of errors on the performance of the electronic nose can be simulated by introducing errors such as a 10% increase in the baseline value $a_{i0}$, or a 5% decrease in the main sensitivity term $a_{ij}$ ($j\neq 0$, $i=j$), or the effect of interference through non-zero off-diagonal regression coefficients. In addition, when it is not possible to distinguish between standard odour samples (e.g. types A and B) with an existing electronic nose, it is possible to use this model to diagnose problems and improve the design (e.g. by simulating the effect of, for example, increasing the number of sensors or changing the type of sensor used).

4.2. Computational results

Fig. 6 shows results obtained from our computational model for a simple nose consisting of two non-linear odour sensors detecting a mixture of three odours (two variable). A uniformly spaced set of concentration points with values of $[0, 20, 40, 60, 80, 100]$ has been used with a 5% random error on all regression coefficients to model materials variation. Fig. 6a shows a plot of sensor space when the sensors are specific to one component, the exponent $k$ is equivalent to a typical semiconducting oxide gas sensor ($k = 0.8$) and there are non-zero baseline values ($a_{i0} = 1$, $a_{00} = 2$). Clearly, the range of this nose is very limited. Notice that the spacing between clusters of 10 replicate samples reduces with increasing concentration because of the initial choice of uniformly spaced concentration points with non-linear sensors. Notice also that the clusters become indistinct at higher component concentrations. Fig. 6b shows the additional effect of making the sensors non-specific. In this case the sensors are sensitive to all three components but with a reduced (positive) sensitivity typically observed for doped tin oxide gas sensors. There is a 20% (30%) sensitivity of sensor 1 (2) to the other variable component whilst having a 30% sensitivity to the third component, which has a fixed concentration of 100 units. As can be seen, the sensor space is now both distorted and shifted. So for non-specific sensors, sample space maps onto a transformed sensor space in which certain regions are now less commonly accessed and hence have a different information content. Regions that are commonly accessed due to cross-sensitivity lie close to the lines intersecting the sensor axes and here clusters are hard to resolve.

Fig. 7 shows the computational results on a more sophisticated electronic nose. In this case there are six non-linear ($k=0.8$) non-specific ($SnO_2$-like) sensors detecting three complex odours. Each complex odour comprises four components of differing amounts (70%, 15%, 10%, 5%).
Fig. 7. Effect of random sampling errors upon a complex computational nose with six SnO$_2$-like ($k = 0.8$) sensors responding to three odours (labelled A, A' and B) each containing a set of four out of seven different components. Each sensor has a specificity of 0.5 ($a_{ij}$ with $i = j$) and four non-zero regression coefficients ($a_{ij}$ with $i = j$) set to 0.25, 0.15, 0.07 and 0.03. Principal-component plots are used to show the effect of (a) 5% random sampling error and (b) 20% random sampling error (10 replicate samples).

5%) out of a total of seven possible components. Principal-component analysis is used to display sensor space and generally indicates the success of using a predictive classifier such as a back-propagation technique. Fig. 7 shows the first and second principal components for a sample set consisting of 10 replicates with random concentration errors $\sigma_x$ of (a) 5% and (b) 20% to simulate sampling variation. Although the odours are separable there is considerable variation in the second principal component associated with higher sampling errors: a feature which has been observed experimentally with commercial tin oxide sensors. Also note that the spread of the clusters varies between sample types. This is because odours A and B were given similar components, whereas odour A' was quite different. Plotting the first and third principal components helps ameliorate this effect, as shown by Fig. 8 where there is still a 20% sampling error.

Fig. 9 shows the effect of drift on this complex nose. In this case the sensitivity of the sensors (regression terms) has drifted by 20%. The direction of the drift in principal-component space is shown and is similar to that followed for sampling errors. However, a 20% drift in the sensor sensitivity is less detrimental than a 20% sampling error. Clearly the resolving power of the nose has been reduced by these errors.

5. Conclusions

At this time there is a need to design new electronic noses for an ever-increasing range of uses. We have constructed a computational model of a generic electronic nose. The model assumes that the sensors obey the principle of linear superposition. This is often a good approximation for low concentrations of analytes in metal oxide and polymer films, when it is not Eqs. (1) and (4) must be modified. This can be used as a design tool to evaluate the effect of, for example, the number of sensors,
the type of sensors (linear, power law, Langmuir, specific, non-specific) and the choice of PARC on its response to known odours. The model can also be used as a diagnostic tool to study the effect of both random and systematic errors (e.g. sampling and sensor characteristics) upon its resolving power as well as providing large data sets for optimizing PARC methods.

There also exists a need to quantify the performance of an electronic nose, so as to compare the different instruments and to show that they can meet the specification laid down by the various customers. Such a specification has to be in the form of a European standard. In this paper we have proposed odour standards and analytical expressions which can be used to define the working range and resolving power of electronic noses. First, a standard of type A, which determines the range of an electronic nose through the use of a set of different organoleptic chemical products. Secondly, and more importantly, further samples of similar odours (type B) can be used to determine the resolving power of the electronic nose.

The commercial future of electronic noses clearly depends upon them meeting well-defined standards (European or International). We believe that we have shown that such standards can be developed to provide a reliable measure of the performance of these new instruments.

References


Biographies

*Dr. Julian Gardner* is a Reader in the Department of Engineering at the University of Warwick. He is also Director of the Sensors Research Laboratory which is part of the Centre for Nanotechnology and Microengineering. His research interests include chemical silicon sensors, electronic noses, intelligent sensor systems and pattern recognition techniques.

*Phil Bartlett* is Professor of Electrochemistry at the University of Southampton. His research interests include bioelectrochemistry and bioelectronics, surface modification of electrodes, conducting polymers and the applications of electropolymerizes films in sensors and nanotechnology.