Concentration estimation of formaldehyde using metal oxide semiconductor gas sensor array-based e-noses

Lei Zhang, Fengchun Tian, Xiongwei Peng, Xin Yin, Guorui Li and Lijun Dang
College of Communication Engineering, Chongqing University, Chongqing, China

Abstract

Purpose – The purpose of this paper is to present a novel concentration estimation model for improving the accuracy and robustness of low-cost electronic noses (e-noses) with metal oxide semiconductor sensors in indoor air contaminant monitoring and overcome the potential sensor drift.

Design/methodology/approach – In the quantification model, a piecewise linearly weighted artificial neural network ensemble model (PLWE-ANN) with an embedded self-calibration module based on a threshold network is studied.

Findings – The nonlinear estimation problem of sensor array-based e-noses can be effectively transformed into a piecewise linear estimation through linear weighted neural networks ensemble activated by a threshold network.

Originality/value – In this paper, a number of experimental results have been presented, and it also demonstrates that the proposed model has very good accuracy and robustness in real-time indoor monitoring of formaldehyde.

Keywords Sensors, Gas, Neural networks, Arrays, Multi-sensor systems

Paper type Research paper

Introduction

Indoor air quality is attracting people's attention in recent years. As we know that new furniture, the wall with new paint and the new floors in a new house with fitment will release chemicals such as formaldehyde, benzene and toluene. Especially, as a kind of harmful gas to a human's health, formaldehyde occupies a larger proportion in the pollutant gas mixture. Therefore, accurate detection of the contents of formaldehyde indoor in real-time becomes the essential task in this work. Electronic nose (e-nose), as an artificial olfaction system, has a wide range of applications. For instance, e-nose has been used in food quality testing (Berna, 2010; Di Natale *et al.*, 1997; Gomez *et al.*, 2008), discrimination of tea (Yu and Wang, 2007) and milk (Ampuero and Bosset, 2003), environmental monitoring (Zhang *et al.*, 2011, 2012; De Vito *et al.*, 2008; Zhang *et al.*, 2013, 2011), medical treatments and diagnosis (Gardner *et al.*, 2000; D'Amico *et al.*, 2010), etc.

This paper concentrates on the study of indoor formaldehyde detection using the proposed e-noses developed in our lab. Our e-noses are designed based on metal oxide semiconductor gas sensors arrays consisting of four gas sensors (TGS2620, TGS2602, TGS2201A and TGS2201B). In addition, a temperature–humidity module is also embedded to sense the ambient temperature and relative humidity indoor. The cross-sensitivity and broad spectrum characteristic of gas sensors array make the gas detection possible. In recent years, new publications on qualitative analysis of various gases using e-noses have been proposed by many researchers (D'Amico *et al.*, 2010; Röck *et al.*, 2008; Brudzewski *et al.*, 2012; Güney and Atasoy, 2012; Cano *et al.*, 2011; Ehret *et al.*, 2011; Chen *et al.*, 2011). However, concentration estimation for quantification analysis is always a challengeable task compared with qualitative analysis in e-nose. To our knowledge, artificial neural network (ANN) has been widely used for concentration estimation (Yea *et al.*, 1997; De Vito *et al.*, 2007; Gao and Chen, 2007; Gao *et al.*, 2012; Huyberechts *et al.*, 1997; Pardo *et al.*, 2000) due to its nonlinear approximation ability. Generally, ANN was used for regression between responses of sensors array and the true concentrations, and then the learned weights of ANN would be transferred to the e-nose system for concentration estimation of unknown samples.

However, ANN is sensitive to environmental noise and cannot perform good detection in its usage; thus, it is trained using a few samples. Moreover, in practical application, the robustness is a very important characteristic of e-noses. Thus, this paper presents a novel quantification model in e-nose based on ANNs and aims to overcome the flaw of ANN's noise sensitivity and improve the stability of e-nose, although it is a new finding to apply ANN in an e-nose for prediction. In actual e-nose application for environmental monitoring, both accuracy and robustness should be equivalently important in technical assessment. In terms of the nonlinear e-nose system and the strong nonlinear regression ability of ANN, we propose a piecewise linearly weighted ANN ensemble model

The current issue and full text archive of this journal is available at **www.emeraldinsight.com/0260-2288.htm**

Sensor Review
34/3 (2014) 284–290
 [ISSN 0260-2288]
[DOI 10.1108/SR-05-2013-673]

The authors would like to express their sincere appreciation to the anonymous reviewers for their insightful comments, which have greatly improved the quality of the paper. This work was funded by the Hong Kong Scholar Program (Grant No. XJ2013044) and China Postdoctoral Science Foundation (Grant No. 2014M550457).

(PLWE-ANN) with characteristics of global linearity and local nonlinearity for formaldehyde concentration estimation by an e-nose. In the proposed ANN ensemble model, a threshold network is designed to make decisions on which prediction equation internally should be selected for usage. Accordingly, an embedded self-calibration module is also proposed for compensating the attenuation of the threshold network output and promising the reliability of the PLWE-ANN ensemble model.

Materials and methods

E-nose

The e-nose system has been introduced in our previous publication (Zhang *et al.*, 2012). The sensor array in an e-nose system consists of four metal oxide semiconductor gas sensors with TGS series including TGS2602, TGS2620, TGS2201A and TGS2201B. In addition, a module with two auxiliary sensors for sensing ambient temperature (T) and humidity (H) is also used. A 12-bit analog–digital converter is used as the interface between the field programmable gate array (FPGA) processor and the sensors. FPGA can be used for data collection, storage and processing. The e-nose system is then connected to a laptop via a joint test action group (JTAG) port which can be used for transferring data and debugging programs. The e-nose instruments developed in our laboratory are illustrated in Figure 1.

The proposed ensemble model

The proposed PLWE-ANN ensemble model in this paper is essentially designed with four ANNs: ANN_1, ANN_2, ANN_3 and $ANN_{threshold}$ (threshold network). Note that ANN_1, ANN_2 and ANN_3 are the predictive units in the PLWE-ANN model for concentration estimation through a weighted scheme. The structure of the PLWE-ANN model is described in Figure 2.

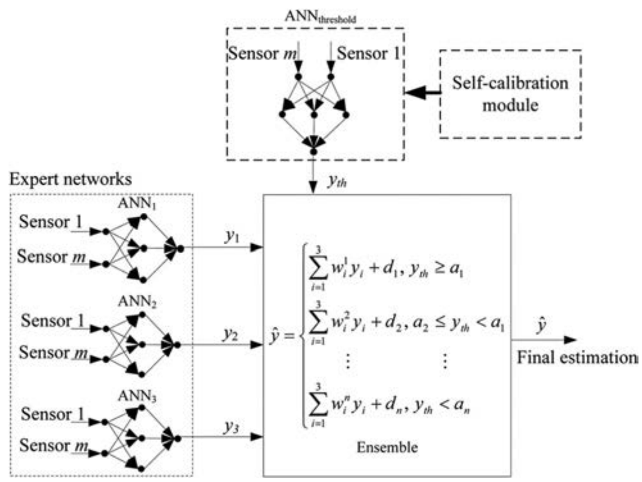
The mathematical representations of the PLWE-ANN model is shown by:

$$\hat{y} = \begin{cases} \sum_{i=1}^{3} w_1^i \cdot y_i + d_1, & y_{th} \ge a_1 \\ \sum_{i=1}^{3} w_2^i \cdot y_i + d_2, & a_2 \le y_{th} < a_1 \\ \vdots & \vdots \\ \sum_{i=1}^{3} w_{n-1}^i \cdot y_i + d_{n-1}, & a_n \le y_{th} < a^{n-1} \\ \sum_{i=1}^{3} w_n^i \cdot y_i + d_n, & y_{th} < a_n \end{cases} \quad (1)$$

where, $\hat{y}$ denotes the estimated concentration; $y_i (i=1,2,3)$ denotes the output of the ANN_i; w^i_j $(i=1,2,3;\ j=1 \ldots, n)$ denotes the weight coefficient of the ANN_i in the j^{th} section; d_j $(j=1 \ldots, n)$ denotes the constant value in the j^{th} section; y_{th} denotes the output of the $ANN_{threshold}$; a_j $(j=1 \ldots, n)$ $(a_1 > a_2 > \ldots > a_n)$ denotes the fixed threshold value in the j^{th} section; and n denotes the total number of sections of PLWE-ANN model. Each section corresponds to a given prediction equation.

Figure 1 E-nose instruments developed in our laboratory

Figure 2 Structure of the PLWE-ANN ensemble model



From equation (1), we can find from the PLWE-ANN model that the concentration estimation is used through a series of multivariate linear regression equations constructed between the outputs of the three ANNs and the true concentration. Note that each ANN is trained between a sensor response vector and the true concentration. Thus, a nonlinear estimation problem of an e-nose can be transformed into a simplified piecewise linearly weighted estimation problem, and the robustness and precision of the e-nose prediction can be effectively improved. Through the long-term observations of e-nose experiments, we set three fuzzy partitions of concentrations:

1. low concentration (0-0.3 ppm);
2. medial concentration (0.3-5 ppm); and
3. high concentration (5-20 ppm).

Then, the threshold values a_j $(j=1 \ldots ,n)$, the weights w^i_j $(i=1,2,3;\ j=1 \ldots ,n)$ and the constants d_j $(j=1 \ldots ,n)$ in the proposed PLWE-ANN model can be obtained through a multivariate linear regression method.

From the constructed PLWE-ANN model, we can find that the threshold network $ANN_{threshold}$ plays a very important role in e-nose prediction because it decides the specific section for robustness. Considering that the $ANN_{threshold}$'s attenuation resulting from the long-term sensor drift or sensor replacements in e-nose exists, and incorrect decision of the desired concentration section in the PLWE-ANN model would happen; thus, the final concentration estimation $\hat{y}$ of e-nose would become inaccurate due to the incorrectly used prediction equation. Therefore, a self-calibration module with adaptive correction of the $\mathrm{ANN}_{\mathrm{threshold}}$ output is also proposed to improve the concentration estimation in long-term use of the developed e-nose instrument and partially overcome the problem of sensor replacement.

Self-calibration module

For simplification of the self-calibration module, the correction method is proposed in a linear way, which is shown as:

$$\hat{y}_{th} = AC \cdot (y_{th} - AB) \tag{2}$$

where, AC and AB are the being-adjusted coefficients of the self-calibration model, $\hat{y}_{th}$ is the corrected value of the $\mathrm{ANN}_{\mathrm{threshold}}$ output y_{th}. Then, the deduced PLWE-ANN model with the self-calibration module is presented as follows:

$$\hat{y} = \begin{cases} \sum_{i=1}^{3} w_1^i \cdot y_i + d_1, & \hat{y}_{th} = AC \cdot (y_{th} - AB) \geq a_1 \\ \sum_{i=1}^{3} w_2^i \cdot y_i + d_2, & a_2 \leq \hat{y}_{th} = AC \cdot (y_{th} - AB) < a_1 \\ \vdots & \vdots \\ \sum_{i=1}^{3} w_{n-1}^i \cdot y_i + d_{n-1}, & a_n \leq \hat{y}_{th} = AC \cdot (y_{th} - AB) < a^{n-1} \\ \sum_{i=1}^{3} w_n^i \cdot y_i + d_n, & \hat{y}_{th} = AC \cdot (y_{th} - AB) < a_n \end{cases} \tag{3}$$

Due to the existence of a long-term sensor drift, calibration coefficients AC and AB should be adjusted adaptively. Therefore, the proposed experimental method for acquisition of AC and AB is illustrated in the following part.

Assume the actual output of the $\mathrm{ANN}_{\mathrm{threshold}}$ to be y_{th}, and the desired output to be $\hat{y}_{th}$. Due to that, $\hat{y}_{th}$ and y_{th} are easier to be obtained, and the two steps for determination of the calibration coefficients are shown as follows:

- *Step 1.* Calculate the coefficient AB of baseline calibration. The e-nose should be exposed to clean air, the initial AC is set as 1 and the desired output $\hat{y}_{th}$ of $\mathrm{ANN}_{\mathrm{threshold}}$ should be 0. That is, $\hat{y}_{th} = 0$. Then, according to equation (2), $y_{th}\text{-}AB = 0 \rightarrow AB = y_{th}$.
- *Step 2.* Calculate the coefficient AC of sensitivity calibration. The e-nose should be exposed to a chamber with a certain formaldehyde concentration C (C should be known in the experiment), then the desired value of $\hat{y}_{th}$ can be estimated by setting $\hat{y} = C$ in equation (3) and other known parameters of the PLWE-ANN model. According to equation (2), AC can be calculated by:

$$AC = \frac{\hat{y}_{th}}{y'_{th} - AB} \tag{4}$$

where, y'_{th} is the current estimation value of $\mathrm{ANN}_{\mathrm{threshold}}$.

From above analysis, we can find AB and AC that were used for correcting the baseline (exposed to clean air) and sensitivity (exposed to contaminants) of $\mathrm{ANN}_{\mathrm{threshold}}$, respectively. In this case, the self-calibration module can partially deal with the sensor drifts, including baseline drift and sensitivity drift, and sensor shifts caused by sensor replacement.

Learning of ANN

In the proposed PLWE-ANN model, the structure of the four neural networks (ANN_1, ANN_2, ANN_3 and $\mathrm{ANN}_{\mathrm{threshold}}$) is totally set as 24-10-10-1 with two hidden layers used. The numbers of input neurons, the 1^{th} hidden neurons, the 2^{th} hidden neurons and output neurons are set as 24, 10, 10 and 1, respectively. Note that the number 24 represents the number of input variables, and the number 1 represents one kind of formaldehyde concentration. Back-propagation algorithm is used for ANN training, the "log-sigmoid" function is used as the transfer function in the two hidden layers and the linear "purelin" function is used as the transfer function in the output layer.

To improve the diversity of ANN_1, ANN_2 and ANN_3, the training samples for each network are different according to the concentration value. ANN_1, ANN_2 and ANN_3 are learned by experimental samples with true low concentration (≤ 0.3 ppm), true median concentration (0.3-0.6 ppm) and true high concentration (≥ 0.6 ppm), respectively. In the proposed PLWE-ANN model, the threshold network $\mathrm{ANN}_{\mathrm{threshold}}$ plays a more important role than other three ANNs; thus, the $\mathrm{ANN}_{\mathrm{threshold}}$ is trained using the total experimental samples so that it can be used to make decisions on specific prediction equations of the PLWE-ANN model in the whole concentration level. In this case, three ANNs can effectively contribute to the final estimation by assigning them appropriate weights in each concentration section of the PLWE-ANN model.

Data analysis

In data analysis, we evaluate the proposed model using absolute error (AE, ppm), relative absolute error (RAE, per cent), mean relative absolute error ($MRAE$, per cent), variance of AE (VAE) and variance of RAE ($VRAE$).

Assume that the true concentration value of the j^{th} sample (j is the validation sample index, $j=1 \ldots, N$) is denoted as C_j (ppm) and the estimated concentration is denoted as $\hat{y}_j$ (ppm), then AE_j, RAE_j per cent, MAE, $MRAE$ (%), VAE and $VRAE$ can be calculated by the following equations:

$$AE_j = |\hat{y}_j - C_j|, \; j = 1, \ldots, N \tag{5}$$

$$RAE_j = \frac{|\hat{y}_j - C_j|}{C_j} \times 100, \; j = 1, \ldots, N \tag{6}$$

$$MAE = \frac{1}{N} \sum_{j=1}^{N} |\hat{y}_j - C_j| \times 100 \tag{7}$$

$$MRAE = \frac{1}{N} \sum_{j=1}^{N} \frac{|\hat{y}_j - C_j|}{C_j} \times 100 \tag{8}$$

$$VAE = \frac{1}{N} \sum_{j=1}^{N} AE_j^2 - \left(\frac{1}{N} \sum_{j=1}^{N} AE_j \right)^2 \tag{9}$$

$$VRAE = \frac{1}{N}\sum_{j=1}^{N} RAE_j^2 - \left(\frac{1}{N}\sum_{j=1}^{N} RAE_j\right)^2 \qquad (10)$$

Experimental data

The gaseous experiments of e-nose in this paper were conducted in the constant temperature and humidity chamber (LRH-150S), which can automatically adjust the temperature and humidity. The target gas is collected in a gas bag and then injected to the chamber through a flowmeter. A fan is fixed in the chamber for purging to make the gas diffuse evenly. In total, 10 minutes are consumed in each experiment for obtaining one sample. The specific experimental procedures can be illustrated as follows:

- *Stage 1.* Gas preparing and collection: collect each target gas in a bag, and dilute each target gas using pure nitrogen (N_2).
- *Stage 2.* Data collection (major part): in this stage, there are several steps shown as follows:
  - *Step 1.* Set the initial temperature and humidity of the chamber. To obtain the sample data by simulating the real environment, all samples are measured at the target temperatures of 15, 25, 30 and 35 degrees and target humidity of 40, 60 and 80 per cent relative humidity, through different combinations of these target temperatures and humidity.
  - *Step 2.* Turn on the e-nose system until the temperature and humidity in the chamber reach the initial setting, and then use the sensor baseline collection for 2 minutes.
  - *Step 3.* Inject target gas by using a flowmeter. Then, the sensors will have a quick response to target gas and until the sensors reach a steady-state response after about 8 minutes. Therefore, one experiment of sample collection would sustain 10 minutes totally. The steady-state response is also the extracted feature to represent the gas texture for pattern analysis.
- *Stage 3. Air exhaust and chamber cleaning:* after one experiment of sample collection, air exhaust by a pump is necessary for chamber cleaning to recover the sensor response as quick as possible.
- *Stage 4. Data transferring to PC:* sensor response data in one experiment is transferred to the PC conveniently through a JTAG connected between the e-nose and the PC for data analysis.

For ANN model learning, in e-nose experiments, we used 126 formaldehyde samples mentioned by Zhang *et al.*, 2011, and each experiment was conducted in the constant temperature and humidity chamber. For feature selection, three points at the steady-state response are extracted from each sensor; for each sample, its concentration (unit: ppm) is obtained using the spectrophotometer which can analyze the collected formaldehyde liquor through a gas sampler.

For testing of the approximated linearity of $ANN_{threshold}$ with the true concentration, two experiments were conducted by injecting formaldehyde gas from low to high concentrations gradually. In the first experiment, 24 samples (0-3 ppm) were obtained; for the second, 34 samples (0-2 ppm) were obtained. Besides, to validate the validity and robust performance of the proposed PLWE-ANN concentration estimation model, 60 samples (six samples were measured in each e-nose) using 10 e-nose instruments with completely the same type developed in our laboratory were designed.

Results and discussion

From the principle of the proposed model, the $ANN_{threshold}$ makes key decisions on the specific sections of PLWE-ANN model. Therefore, it should have a good linearity with the true concentrations and promise the correct decision. As can be seen in Figure 3(a and b), they prove the potential linearity of $ANN_{threshold}$ with the true concentrations from two experiments. Figure 3(a) presents the first experiment in which 24 samples between 0 and 3 ppm are collected, and the other experiment was shown in Figure 3(b), which includes 34 samples between 0 and 2 ppm. In Figure 3, the prediction of $ANN_{threshold}$ can effectively follow the true concentrations in a linear way, and it means that the $ANN_{threshold}$ obtained in this paper is a good choice to control the specific section in the proposed model.

The parameters of PLWE-ANN model in equation (1) are presented in Table I. In total, 18 sections are obtained including high concentration estimation from sections 1 to 4 ($5 \leq y_{th} \leq 20$ ppm), median concentration estimation from sections 4 to 15 ($0.3 \leq y_{th} < 5$ ppm) and low concentration estimation from sections 15 to 18 ($0.005 \leq y_{th} < 0.3$ ppm). Besides, we take the estimation level $y_{th} < 0.005$ ppm as the baseline prediction in clean air.

The prediction results of the 60 samples using the proposed PLWE-ANN model are shown in Figure 4. We can find from Figure 4 that the final predictions can approximate the true values very well. For quantitative evaluation of the prediction performance, Table II presents the specific validation results of all the experimental samples. Besides, the true values, the

Figure 3 Approximated linearity test between the output of $ANN_{threshold}$ and the true concentration

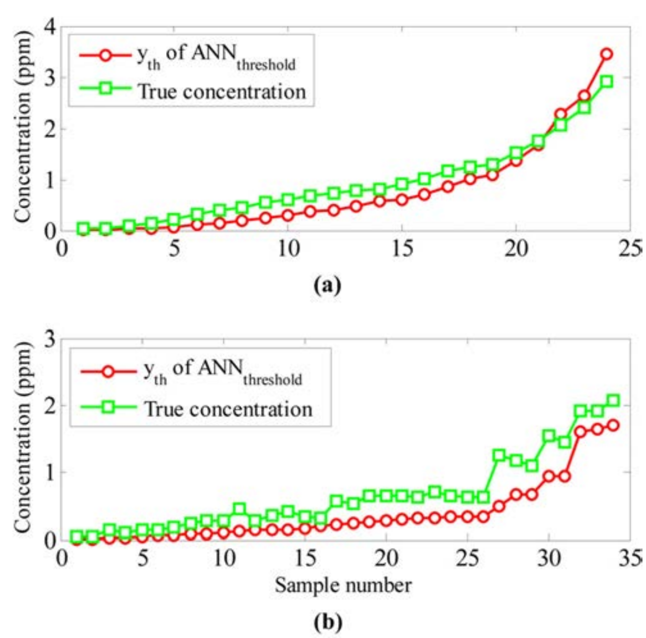
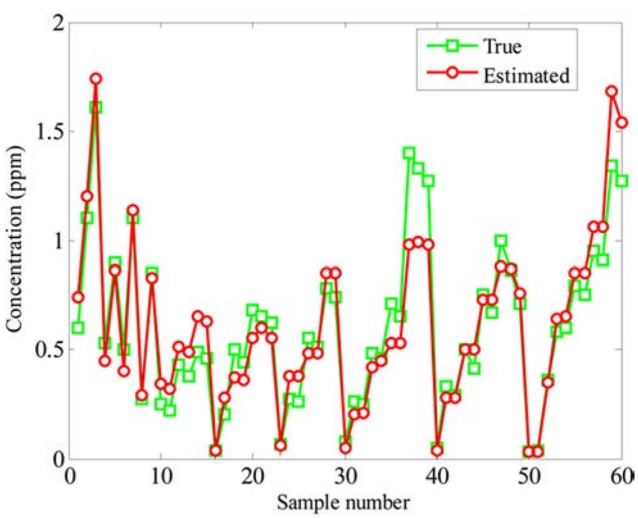


Table I Parameters of the PLWE-ANN ensemble model

Section i	w_1^i	w_2^i	w_3^i	d_i	a_i
1	0.1	0.4	1.0	3.6	20
2	0.1	0.4	1.0	3.2	15
3	0.1	0.4	1.0	2.8	10
4	0.1	0.4	1.0	2.4	5.0
5	0.1	0.4	1.0	2.2	4.5
6	0.1	0.4	1.0	2.0	4.0
7	0.1	0.4	1.0	1.8	3.5
8	0.1	0.4	1.0	1.6	3.0
9	0.5	0.5	1.0	1.4	2.5
10	0.5	0.5	1.0	1.2	2.0
11	0.5	0.5	1.0	1.0	1.5
12	0.5	0.5	1.0	0.8	1.0
13	0.5	0.5	1.0	0.7	0.8
14	0.5	0.5	1.0	0.6	0.5
15	0.5	0.5	1.0	0.4	0.3
16	1.0	0.4	1.0	0.2	0.1
17	0.8	0.5	1	0.1	0.03
18	0.5	0.1	0.5	0.03	0.005

Figure 4 Formaldehyde concentration prediction of the proposed PLWE-ANN ensemble model



estimated values, the RAE and absolute errors AE for all samples are included.

For visualization of the predictive performance, we have presented the bar plots of the RAE and AE in Figure 5(a and b), respectively. In Figure 5, the samples have been rearranged according to the concentrations from low to high so that we can see the prediction error in different concentration levels. For instances, the relative error is larger in low concentration, while the absolute error is larger in high concentration which can be easy to understand according to computations shown in equations (5) and (6). Through the equations (7–10), we can obtain that the $MRAE$ is 16.76 per cent, the $VRAE$ of RAE is 1.45 per cent, the MAE is 0.09 ppm and the VAE of AE is 0.0077 ppm. The $VRAE$ shows the uniform error distribution for the total samples from 10 e-nose instruments with completely the same type developed in our laboratory.

Besides, we present the proportions of samples in four different relative error grade 0-10, 10-20, 20-30 and 30-50 per cent in Table III, respectively. We can find that the proportion of 85 per cent was obtained with the relative error $<$ 30 per cent. The proportions of samples and mean relative absolute error in low (0-0.3 ppm), median (0.3-0.6 ppm) and high concentration ($\geq$ 0.6 ppm) levels are included in Table IV. We can see that the mean relative absolute errors are 23.9, 16.7 and 14.2 per cent in low, median and high concentration levels, respectively. We also use the statistical analysis with a confidence level of 95 per cent, as shown in Figure 6. The statistical results demonstrate the final concentration estimations that show a very good linear regression with the true values. These results demonstrate that the proposed PLWE-ANN model does not only have a good accuracy but also has good robustness in different concentration levels validated on multiple e-nose instruments.

In e-nose research, the robustness in actual application, especially in quantity of e-nose instruments, is also important. Nonlinear methods may show a seemingly good prediction, however, the robustness cannot be promised due to the broad-spectrum response characteristic of metal oxide semiconductor gas sensors which are sensitive to many kinds of gases and is easy to be affected by environmental elements such as temperature, humidity and pressure. Linear methods may have good robustness in actual prediction; however, the accuracy cannot be promised because of the strong nonlinearity of multiple sensors system in an e-nose. Therefore, a new prediction method is proposed based on an effective fusion of nonlinearity and linearity in this paper. The fusion means that the model for the whole (globally) is linear in robustness and nonlinear to its internal units (locally) in accuracy. The quantification of e-nose largely depends on the reliability of $ANN_{threshold}$ which decides the performance of the proposed PLWE-ANN concentration estimation model. First, it is very necessary to gain an obviously good linearity of $ANN_{threshold}$ with the true concentration in the model. Second, the $ANN_{threshold}$ should have a good stability after its long-term use. Note that we embed the self-calibration model for $ANN_{threshold}$ so as to improve its stability and avoid the possible attenuation results from the sensor drift after aging and sensor shift caused by sensor replacement. From the real validation results analyzed from 10 e-nose instruments, we can say that the proposed model is competent in e-nose prediction in accuracy and robustness.

Conclusions

In this paper, we propose a novel PLWE-ANN model with an embedded self-calibration module in e-noses for concentration estimation of formaldehyde. Unlike the single ANN, this proposed model is linear globally and nonlinear locally, and the ensemble model has better robustness than the single ANN. For improving the robustness and accuracy of e-nose prediction, we designed three expert ANNs as the prediction units in the model for accurate concentration estimation in each concentration level. Besides, the sensor drift after aging and the sensor shift caused by sensor replacement (i.e. discreteness) will also influence the long-term stability and

Sensor Review
Volume 34 · Number 3 · 2014 · 284–290

Table II Validation results of formaldehyde concentration estimation using the proposed PLWE-ANN ensemble model

j	C_j	$\hat{y}_j$	AE_j	RAE_j	j	C_j	$\hat{y}_j$	AE_j	RAE_j	j	C_j	$\hat{y}_j$	AE_j	RAE_j
1	0.60	0.74	0.14	23.3	21	0.65	0.60	0.05	7.69	41	0.33	0.28	0.05	15.1
2	1.10	1.20	0.09	9.09	22	0.62	0.55	0.07	11.214.2	42	0.29	0.28	0.01	3.44
3	1.61	1.74	0.13	8.07	23	0.07	0.06	0.01	40.7	43	0.50	0.50	0.00	0.00
4	0.53	0.45	0.08	15.0	24	0.27	0.38	0.11	46.1	44	0.41	0.50	0.09	21.9
5	0.90	0.86	0.04	4.44	25	0.26	0.38	0.12	12.7	45	0.75	0.73	0.02	2.66
6	0.50	0.40	0.10	20.0	26	0.55	0.48	0.07	5.88	46	0.67	0.73	0.06	8.95
7	1.10	1.14	0.04	3.63	27	0.51	0.48	0.03	8.97	47	1.00	0.88	0.12	12.0
8	0.27	0.29	0.02	7.40	28	0.78	0.85	0.07	14.8	48	0.86	0.87	0.01	1.16
9	0.85	0.83	0.02	2.35	29	0.74	0.85	0.11	37.5	49	0.71	0.76	0.05	7.04
10	0.25	0.34	0.09	36.0	30	0.08	0.05	0.03	23.0	50	0.03	0.03	0.00	0.00
11	0.22	0.32	0.10	45.4	31	0.26	0.20	0.06	16.0	51	0.04	0.03	0.01	25.0
12	0.43	0.51	0.08	18.6	32	0.25	0.21	0.04	12.5	52	0.36	0.35	0.01	2.77
13	0.38	0.49	0.11	28.9	33	0.48	0.42	0.06	2.17	53	0.58	0.64	0.06	10.3
14	0.49	0.65	0.16	32.6	34	0.46	0.45	0.01	25.3	54	0.60	0.65	0.05	8.33
15	0.46	0.63	0.17	36.9	35	0.71	0.53	0.18	18.4	55	0.79	0.85	0.06	7.59
16	0.04	0.04	0.00	0.00	36	0.65	0.53	0.12	30.0	56	0.75	0.85	0.10	13.3
17	0.20	0.28	0.08	40.0	37	1.40	0.98	0.42	25.5	57	0.95	1.06	0.11	11.5
18	0.50	0.37	0.13	26.0	38	1.33	0.99	0.34	22.8	58	0.91	1.06	0.15	16.4
19	0.44	0.36	0.08	18.1	39	1.27	0.98	0.29	20.0	59	1.34	1.68	0.34	25.3
20	0.68	0.55	0.13	19.1	40	0.05	0.04	0.01		60	1.27	1.54	0.27	21.2

Figure 5 Bar plots of relative estimation error and absolute estimation error

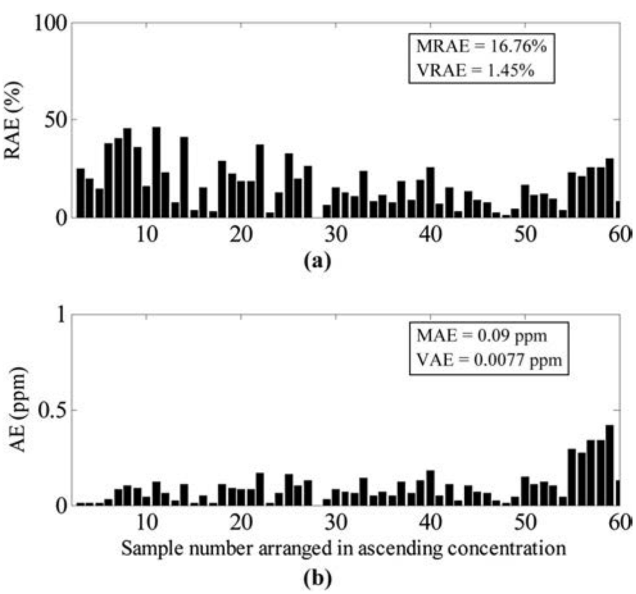


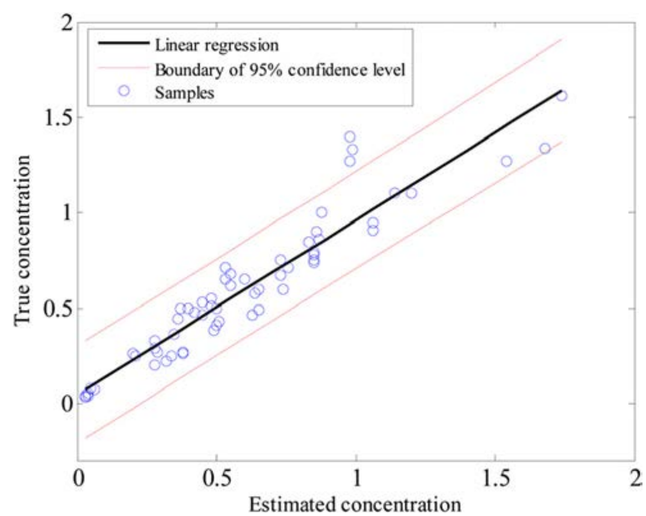
Table III Proportions of samples in total validation samples with different *RAE* grades

RAE grades (per cent)	Number of samples	Proportion (per cent)
0-10	20	33.33 (20/60)
10-20	19	31.67 (19/60)
20-30	12	20 (12/60)
30-50	9	15 (9/60)

Table IV *MRAE* of prediction in low, median and high concentration grades of formaldehyde

Concentration grades	Number of samples	Proportion (per cent)	*MRAE* (per cent)
Low (0, 0.3 ppm)	15	25	23.9
Median (0.3, 0.6 ppm)	19	31.67	16.7
High (0.6, 5 ppm)	26	43.33	14.2

Figure 6 Regression and statistical analysis of estimated and true concentrations



quantity application of e-noses. Therefore, an embedded self-calibration module is proposed for correcting the $ANN_{threshold}$'s attenuation in baseline and sensitivity of the proposed PLWE-ANN model. Experimental results demonstrate the effectiveness of the proposed estimation model in quantity of e-nose instruments.

References

Ampuero, S. and Bosset, J.O. (2003), "The electronic nose applied to dairy products: a review", *Sensors and Actuators B*, Vol. 94 No. 1, pp. 1-12.

Berna, A. (2010), "Metal oxide sensors for electronic noses and their application to food analysis", *Sensors*, Vol. 10 No. 1, pp. 3882-3910.

Brudzewski, K., Osowski, S. and Golembiecka, A. (2012), "Differential electronic nose and support vector machine for fast recognition of tobacco", *Expert Systems with Applications*, Vol. 39 No. 10, pp. 9886-9891.

Cano, M., Borrego, V., Roales, J., Idígoras, J., Lopes-Costa, T., Mendoza, P. and Pedrosa, J.M. (2011), "Rapid discrimination and counterfeit detection of perfumes by an electronic olfactory system", *Sensors and Actuators B*, Vol. 156 No. 1, pp. 319-324.

Chen, H.L., Liu, D.Y., Yang, B., Liu, J. and Wang, G. (2011), "A new hybrid method based on local fisher discriminant analysis and support vector machines for hepatitis disease diagnosis", *Expert Systems with Applications*, Vol. 38 No. 9, pp. 11796-11803.

D'Amico, A., Pennazza, G., Santonico, M., Martinelli, E., Roscioni, C., Galluccio, G., Paolesse, R. and Di Natale, C. (2010), "An investigation on electronic nose diagnosis of lung cancer", *Lung Cancer*, Vol. 68 No. 2, pp. 170-176.

De Vito, S., Castaldo, A., Loffredo, F., Massera, E., Polichetti, T., Nasti, I., Vacca, P., Quercia, L. and Di Francia, G. (2007), "Gas concentration estimation in ternary mixtures with room temperature operating sensor array using tapped delay architectures", *Sensors and Actuators B*, Vol. 124 No. 2, pp. 309-316.

De Vito, S., Massera, E., Piga, M., Martinotto, L. and Di Francia, G. (2008), "On field calibration of an electronic nose for benzene estimation in an urban pollution monitoring scenario", *Sensors and Actuators B*, Vol. 129 No. 2, pp. 750-757.

Di Natale, C., Macagnano, A., Davide, F., D'Amico, A., Paolesse, R., Boschi, T., Faccio, M. and Ferri, G. (1997), "An electronic nose for food analysis", *Sensors and Actuators B*, Vol. 44 Nos 1/3, pp. 521-528.

Ehret, B., Safenreiter, K., Lorenz, F. and Biermann, J. (2011), "A new feature extraction method for odour classification", *Sensors and Actuators B*, Vol. 158 No. 1, pp. 75-88.

Gao, D. and Chen, W. (2007), "Simultaneous estimation of odor classes and concentrations using an electronic nose with function approximation model ensembles", *Sensors and Actuators B*, Vol. 120 No. 2, pp. 584-594.

Gao, D., Yang, Z., Cai, C. and Liu, F. (2012), "Performance evaluation of multilayer perceptrons for discriminating and quantifying multiple kinds of odors with an electronic nose", *Neural Networks*, Vol. 33 No. 1, pp. 204-215.

Gardner, J.W., Shin, H.W. and Hines, E.L. (2000), "An electronic nose system to diagnose illness", *Sensors and Actuators B*, Vol. 70 Nos 1/3, pp. 19-24.

Gomez, A.H., Wang, J., Hu, G.X. and Pereira, A.G. (2008), "Monitoring storage shelf life of tomato using electronic nose technique", *Journal of Food Engineering*, Vol. 85 No. 4, pp. 625-631.

Güney, S. and Atasoy, A. (2012), "Multiclass classification of *n*-butanol concentrations with k-nearest neighbor algorithm and support vector machine in an electronic nose", *Sensors and Actuators B*, Vol. 166-167 No. 1, pp. 721-725.

Huyberechts, G., Szecowka, P., Roggen, J. and Licznerski, B.W. (1997), "Simultaneous quantification of carbon monoxide and methane in humid air using a sensor array and an artificial neural network", *Sensors and Actuators B*, Vol. 45 No. 2, pp. 123-130.

Pardo, M., Faglia, G., Sberveglieri, G., Corte, M., Masulli, F. and Riani, M. (2000), "A time delay neural network for estimation of gas concentrations in a mixture", *Sensors and Actuators B*, Vol. 65 Nos 1/3, pp. 267-269.

Röck, F., Barsan, N. and Weimar, U. (2008), "Electronic nose: current status and future trends", *Chemical Review*, Vol. 108 No. 1, pp. 705-725.

Yea, B., Osaki, T., Sugahara, K. and Konishi, R. (1997), "The concentration-estimation of inflammable gases with a semiconductor gas sensor utilizing neural networks and fuzzy inference", *Sensors and Actuators B*, Vol. 41 Nos 1/3, pp. 121-129.

Yu, H. and Wang, J. (2007), "Discrimination of Long Jing Green-tea grade by electronic nose", *Sensors and Actuators B*, Vol. 122 No. 1, pp. 134-140.

Zhang, L., Tian, F.C., Kadri, C., Pei, G., Li, H. and Pan, L. (2011), "Gases concentration estimation using heuristics and bio-inspired optimization models for experimental chemical electronic nose", *Sensors and Actuators B*, Vol. 160 No. 1, pp. 760-770.

Zhang, L., Tian, F.C., Nie, H., Dang, L., Li, G., Ye, Q. and Kadri, C. (2012), "Classification of multiple indoor air contaminants by an electronic nose and a hybrid support vector machine", *Sensors and Actuators B*, Vol. 174 No. 1, pp. 114-125.

Zhang, L., Tian, F.C., Liu, S., Guo, J., Hu, B., Ye, Q., Dang, L., Peng, X., Kadri, C. and Feng, J. (2013), "Chaos based neural network optimization for concentration estimation of indoor air contaminants", *Sensors and Actuators A: physical*, Vol. 189 No. 1, pp. 161-167.

Further reading

Zhang, L., Tian, F.C., Kadri, C., Xiao, B., Li, H., Pan, L. and Zhou, H. (2011), "On-line sensor calibration transfer among electronic nose instruments for monitoring volatile organic chemicals in indoor air quality", *Sensors and Actuators B*, Vol. 160 No. 1, pp. 899-909.

Corresponding author

Lei Zhang can be contacted at: leizhang@cqu.edu.cn