On field calibration of an electronic nose for benzene estimation in an urban pollution monitoring scenario

5 authors, including:

Saverio De Vito
ENEA
142 PUBLICATIONS   **1,849** CITATIONS

SEE PROFILE

Ettore Massera
ENEA
191 PUBLICATIONS   **1,838** CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Project CONVERGENCE View project

Project CONVERGENCE View project

On field calibration of an electronic nose for benzene estimation in an urban pollution monitoring scenario

S. De Vito , E. Massera , G. Di Francia
ENEA, Centro Ricerche Portici, 80055 Portici (NA), Italy

M. Piga, L. Martinotto
Pirelli Labs, Viale Sarca 222, 20126 Milano, Italy

Abstract

Low-cost gas multi-sensor devices could be efficiently used for densifying the sparse urban pollution monitoring mesh if equipped with a reliable calibration able to counter specificity and stability issues of solid-state sensors they rely on. In this work, we present a neural calibration for the prediction of benzene concentrations using a gas multi-sensor device (solid-state) designed to monitor urban environment pollution. The feasibility of a sensor fusion algorithm as a calibrating tool for the multi-sensor device is discussed. A Conventional air pollution monitoring station is used to provide reference data. Results are assessed by means of prediction error characterization throughout a 13 months long interval and discussed. Relationship between training length and performances are also investigated. A neural calibration obtained using a small number of measurement days revealed to be capable to limit the absolute prediction error for more than 6th month, after which seasonal influences on prediction capabilities at low-concentrations suggested the need for a further calibration.

Introduction

Urban atmospheric pollutants are considered responsible for the increased incidence of respiratory illness in citizens, and some of them (e.g. benzene) are known to induce cancers in case of prolonged exposure [1,2]. Precise estimation of pollutants distribution is hence relevant for traffic management in the municipalities and more generally for the definition of integrated mobility plans designed to face these problems. Nowadays, urban air pollution monitoring is primarily carried out by means of networks of spatially distributed

fixed stations. These equipments, mostly based on industrial spectrometers, can selectively and precisely estimate the concentrations of many atmospheric pollutants, but their costs and sizes seriously hamper the deployment of adequately dense measurement networks. Unfortunately, pollutants diffusion is heavily affected by atmosphere dynamics and the availability of a limited number of measurement nodes may lead to the misevaluation of the real distribution of gases and particles concentrations in a complex and turbulent environment such as a city [3–5]. In this framework, gas multi-sensor devices, often referred as electronic noses, could play an important role helping to raise the density of the monitoring network. However, their concentration estimation capabilities are seriously limited by the known stability and selectivity issues of solid-state sensors they often rely on, especially when using univariate calibrations (i.e. when relying, for regression purposes, on the single sensor response). In harsh traffic environments, such effects are revealed even more evident [6–8]. When coupled with ad hoc sensor fusion algorithms to obtain a multi-variate calibration, i.e. a calibration relying on multiple sensor responses, gas multi-sensor devices have shown interesting capabilities for dealing with concentration estimation problems in complex mixtures. Often, the selected calibration model has been artificial neural networks [9–11], partial least squares (PLS) [12,13] and more recently Kernel methods like support vector machines [14,15]. Their capability to exploit partial selectivity of sensors as an advantage can give them a chance of well performing in modelling this complex framework. It is also reasonable to expect that, at any measurement point, the concentrations of different gases building the urban pollution may show significant correlations that these architectures could also exploit to perform concentration estimation. This can ultimately lead to systems able to estimate analytes for which no specifically targeted sensor is included in the sensor array, with sufficient accuracy. Sensor fusion algorithms, when used in regression schemes, need to be properly tuned via supervised training, i.e. using reference data. Unfortunately, for this scenario, the attempt to extend to field a neural calibration obtained via training with synthetic mixtures, was revealed to be unsuccessful, strongly suggesting the need to use on field recorded data [16]. At present, only few attempts to verify the possibility to obtain a calibration with on field campaigns have been made. Kamionka et al. [17] used an array of three solid-state sensors (screen printed tin dioxide thick layers) and obtained a neural calibration with on field data for estimation of NO 2 and O 3 . The obtained calibration showed good performances at ppb levels for 1-week intervals, after which sensors drifts imposed a new calibration. Carotta et al. [18] achieved an on field univariate calibration with least square methodology for real time NO x and CO concentration estimation, presenting results for a 3 days length test. Tsujita et al. performed a calibration for their multi-sensor device capable to estimate NO_2 with a 10 ppb mean absolute error increasing up to 40 ppb within 2 months [19]. They showed how a simple automatic re-calibration procedure based on the base

resistance, if performed every 2–4 weeks, could fix the error under 10 ppb levels. In this work, we present the results of the evaluation of an on field, multi-variate calibration of a multi-sensor device based on seven solid-state sensors, developed in cooperation with Pirelli Labs [20]. The overall work was targeted at the estimation of CO, NOx , NO_2 and benzene concentrations. Here, we focus on reporting benzene-related findings, delaying to a further paper the presentation of the remaining species. In order to characterize properly the performance of the proposed system, coping with long-term meteorological influences, seasonal car traffics, house heating variability and sensor instabilities, a long-term measurement campaign has been set up. The campaign was performed at a main road located in an Italian city and lasted for 13 months. Calibration was obtained by means of a statistical sensor fusion algorithm; in particular a neural network architecture was used [21]. This architecture, when using integer weights, could be, in fact, easily implemented on-board by means of a simple microcontroller. Performances are discussed by characterizing the estimation error with respect to conventional station output over the measurement campaign investigating the influence of training length. Long-term accuracy of a neural calibration obtained with data recorded on March 2004 was also investigated to understand the period of performance optimality. Highlighted seasonal influences and sensor drifts imposed a further calibration to keep adequate levels of accuracy in the winter time.

Experimental and methods

A novel multi-sensor device was developed by Pirelli Labs for the use in urban pollution monitoring scenarios. In particular, the device was designed as a low-cost compact, versatile and reliable tool to support the city air quality monitoring network managed, in Italy, by local or government agencies for environmental protection. The device was characterized by a three-stage design (sensor array, data processing unit, communication unit). The device was built up by a 31 cm × 26 cm × 12 cm metallic case hosting the power management unit, signal conditioning and acquisition electronics, a microcontroller board hosting a microprocessor eventually capable to run simple sensor fusion algorithms, a GSM (Global System for Mobile Communications) data trasmission unit and of course a sensor array subsytem. Total weight was 2.5 kg. The microcontroller board took care of first-stage data processing operations, storing up to 72 h measurements at an 8 s sample rate and controlled the communication unit in order to transfer processed data (8, 15 or 60 min mean sampled values) to data sinks. In this work, the proposed multi-sensor device was equipped with five independent sensor slots for easy management and replacement, hosting metal oxide chemoresistive sensors whose brief characteristics are depicted in Table 1. Further two slots were devoted to host commercial temperature and humidity sensors. A middle layer provided resistance to voltage conversion and

adaptation to the on-board ADC specifications. In order to validate its response, the device has been co-located with a fixed conventional monitoring station equipped with spectrometer analysers owned by Lombardy Regional Agency for Environment Prevention (ARPA). Measurement campaign took place, in fact, at a main street located in the centre of an Italian city characterized by heavy car traffics. The campaign lasted from March 2004 until April 2005, building up a suitable data set for the devised application. During this quite long-time interval the effect of several different issues characterizing this scenario, such as meteorological influences or sensor drifts, was investigated. The conventional fixed station provided reference concentration estimation for five different atmospheric pollutants, i.e. CO (mg/m^3), non-metanic hydrocarbons (NMHC) $(microg/m^3)$, C_6H_6 $(microg/m^3)$, NO x (ppb), NO 2 $(microg/m3)$ and was sampled recording hourly means of the concentration values. At the same time, the multi-sensor device was sampled to provide the hourly mean of the resistivity expressed by the seven solid-state sensor units specified above. As expected, significant correlations have been found among the recorded species concentrations; in particular, in Table 2 we report the linear correlation coefficients computed for a notable subset of the analysed species. In order to correctly estimate the values of single pollutant concentrations, resistivities were used as input of back-propagation neural networks [22], devised for approximating Ψ in the regression problem:

$$C_j = \Psi (RSens_i)$$

validation (for early stopping) and test set (for performance assessment) were built by randomly choosing samples from the remaining data. Reference data was provided using conventional station output. (1) where C j is the real pollutant concentrations as measured by conventional stations and RSens i are the sensor resistances measured by the multi-sensor device. Due to its architecture, this sensor fusion algorithm is expected to be capable to use any meaningful correlation existing between the different pollutants in the gas mixtures related to the problem under investigation. We focused our attention on concentration estimation of benzene, for which no specific sensor was devised in the multi-sensor array, assessing the array capability to cooperatively perform estimations exploiting both non-specificity and scenario-related correlations. We tested several architectures differing only by the number of neurons embedded in their single hidden layer, the basic architecture being a 7-X-1 Feed-Forward network with X in [3,5,15,20]. The hidden neuron transfer function was the MATLAB tansig function. In the preliminary runs, networks were trained using the resilient back-propagation algorithm [21] and early stopping as a measure to prevent over-training problems [22]. Training sets were built using consecutive samples, i.e. using a fixed length specific interval of the campaign data, while validation (for early

stopping) and test set (for performance assessment) were built by randomly choosing samples from the remaining data. Reference data was provided using conventional
station output.

Table 1

Parameters	Sensor types				
	CO	NO$_x$	O$_3$	NMHC (non-metanic hydrocarbons)	NO$_2$
Technology	Tin oxide	Tungsten oxide	Indium oxide	Titania	Tungsten oxide
Concentration nominal range	0.1–100 mg/m^3	20–1000 µg/m^3	20–600 µg/m^3	–	0–2 ppm
Detection limit	0.1 mg/m^3	20 µg/m^3	20 µg/m^3	–	0.020 ppm
Expected operating life (months)	12	12	12	12	12
Sensor response time, T_{90} (s)	60	90	90	90	<240
Operating temperature range (°C)	−20 to 50	−20 to 50	−20 to 50	−20 to 50	−20 to 50
Operative relative humidity range (%)	10–90	10–90	10–90	10–90	10–90
Output drift (%) (6 months)	<2	<2	<2	<2	<2

Selected Sensor Array principal parameters as obtained in single specie lab based characterization

Table 2

$r_{\text{NMHC-C}_6\text{H}_6}$	0.98
$r_{\text{CO-NO}_x}$	0.78
$r_{\text{CO-NO}_2}$	0.67
$r_{\text{C}_6\text{H}_6\text{-NO}_x}$	0.72
$r_{\text{C}_6\text{H}_6\text{-NO}_2}$	0.60
$r_{\text{NO}_x\text{-NO}_2}$	0.76
$r_{\text{CO-C}_6\text{H}_6}$	0.90

Linear correlation coefficients computed among analysed species using on filed recorded data

Table 3

MRE	0.0069
STD_RE	0.0891
MAE (µg/m^3)	0.0337
MSE	9.10E−04
SCC	0.9998

Performance indexes of a 7-X-1 back-propagation network computed for a training length of 8 months (benzene concentration prediction). Results were threefold cross-validated. MRE and MAE are extremely low while relative error standard deviation value depicts a significant dispersion in the obtained values.

Results and Discussion

The first investigation was set up trying to estimate the benzene concentrations using a huge data set for training. The   training set was actually built using the initial 8 months segment while after a randomization, one third of the remaining samples were chosen to act as the validation set in the early   stopping training procedure. All other samples were used as   the test set. Results were threefold cross-validated by running   the performance assessment using different validation/test sets.   The process was repeated for the above-mentioned architectures   and the best results were obtained for benzene concentration estimation by the five hidden neurons architecture. In Table 3,   we present these results by means of multiple performance   indexes: mean relative error (MRE) defined as the sample mean   of the ratio between absolute prediction error and true concentration value, relative error standard deviation (STD RE), mean absolute error (MAE) defined as the sample mean of absolute   prediction error, mean squared error (MSE) and squared correlation coefficient (SCC). As a reference, the empirical distribution   of benzene concentration values, as recorded by the conventional   station, is reported in Fig. 1. Values obtained for square correlation coefficients between expected (conventional station output) and predicted values were indeed very good with the mean relative error standing well under 1%. In Fig. 2, the cumulative   distribution function (CDF) of the relative and absolute error   is reported, showing that in the 99% of test samples, the absolute error was less than 0.3 microg/m^3 . In Fig. 3 where we show   the system performances over a test subset covering a week,   the ground truth estimation performed by the conventional analyser is almost perfectly overlapped by estimations carried out   by the multi-sensor device when coupled with the sensor fusion   subsystem and the error remains very limited.   Of course, in order to be feasible for this scenario, neural calibration should use a very limited segment of data for   training purposes so as to limit the overall period needed for   device calibration to a minimum number of days. Hence, it is   of great relevance to try to explore the relationship between performances and the training set length in terms of consecutive days. For this reason, we set up different training runs with different training set lengths. Training was conducted using   Levenberg–Marquardt [23,24] algorithm pursuing a 5×10^{-7} mean square error (MSE) on the training set with a maximum number of training epochs equal to 1000 to control the   learning phase, hence without using early stopping. Results   were obtained repeating the training and performance estimation cycle for 20 times in order to reduce the uncertainty, due to   the training methodology instabilities, about the estimated performance indexes. Of course, performance was assessed running the trained sensor fusion subsystem on the remaining samples   that were not used, in each turn, for the training stage. We   found that for the benzene concentration estimation problem,   in order to keep the MRE under a 2% threshold with respect to   the conventional fixed station output, a 10-day training length   was revealed to be sufficient (see Table 4). Furthermore,

theabsolute error was kept well below the 1 microg/m3 threshold for almost all the test samples. This results showed the feasibility of obtaining a well performing calibration with on field data, even when using a very limited training length. For car traffic generated pollution, the NMHC pollutant mixture has a notable benzene-related component. Furthermore, as previously mentioned, significant correlations exist among the other species, so that the remaining sensors can also give a contribution to the knowledge about benzene concentrations. Here, the sensor fusion subsystem has learned to exploit the capabilities of the NMHC sensor (primarily) and the scenario-related correlations, including the contributions of the remaining sensors, to perform the estimation of benzene concentrations. By using this strategy,it was possible for the neural network to exploit partial specificities and scenario-related correlations to perform adequate estimations of the benzene concentrations. However, the MRE when computed over the entire 12 months period has a limited descriptive power with respect to the variations occurring on this several months interval. As such, we analysed the weekly scores in order to report information useful for understanding performance degradations and eventually, evaluated the period after which a further calibration was required. Note that weekly scores were also computed with a mean over 20 training/performance estimation cycles. In Fig. 4 we report the weekly computed mean relative error and the mean absolute error as a function of time for benzene concentration estimation problem. We found a significant oscillation superimposed to a constant degenerative trend. This oscillation is due to the seasonal differences among the gas mixtures presented to the device, while the low-frequency trend is more likely to be due to the sensor drifts. During winter time, many pollutant concentration ranges shift toward higher values with respect to spring time; for example, the pollution component due to building heating devices become more significant modifying the ratios among pollutants concentration in the environment. During the measurement campaign we actually measured these effects, and changes were particularly evident in the case of nitrogen oxides where the empirical mean and standard deviation change from $\mu = 134$ ppb $\sigma = 91.8$ ppb (April to September) to $\mu = 382$ ppb $\sigma = 243.8$ ppb (October to March). The combined effects of empirical distribution changes and sensor drifts could make the knowledge obtained by training using spring time measurements partially obsolete, thus requiring a system re-calibration, in order to pursue the best performances. From March to April, pollutant statistical parameters seem to recover toward their original values. Accordingly, we observe a decreasing trend in the MRE. Anyway, in our case, results seem to show that a single calibration operation with 10 days length can be capable to keep results in a very small error range for more than 6 months. Furthermore, peaks in weekly MRE mostly occur at the lowest benzene concentrations, keeping the absolute error well below 0.2 microg/m3 . It is also worth to note that these peaks seem to occur at low temperatures after a rapid decrease in relative

humidity.In Fig. 5 we show a scatterplot of estimation error and relative error versus real benzene concentration in order to describe their distribution along different concentration values. It is possible to note how the estimation error range is almost constant and shows no bias between 8 and 30 microg/m3 , while a slight positive bias is apparent at the lowest concentrations (between 0 and 2 microg/m^3). These effects are far more evident for the highest concentrations where the model seems not adequate, probably because of the very low number of training samples having this characteristics. However the relative error became significant only below a 3 microg/m3 concentration level. As already mentioned, from Fig. 4 it is possible to spot a definite rise of weekly mean absolute error starting appreciatively at the 31th week with a rapid decrease starting at the 50th week. In order to better characterize the concentration levels at which these performance variations mostly occur, in Fig. 6 we depict the different distribution of MRE using two scatterplots, computed for (a) the first 30 weeks and for (b) the remaining weeks. Fig. 6 shows how significant changes occur at concentrations below 10 microg/m3 , so that from the 31st to the 50th week, the system showed difficulties in the modelling of sensor responses at these concentrations. After that we evaluated the effects of further calibration to occur at the 34th week (mid-November) using the 34th week samples as a training set and the remaining 19 weeks as a test bed. Results reported in Fig. 7 depicts that a significant decrease in both relative and absolute error is obtained but, surprisingly, peaks of MRE continue to occur almost at the same weeks, definitely revealing a slight difficulty for the multi-sensor device to cope with low levels of benzene in the winter time.
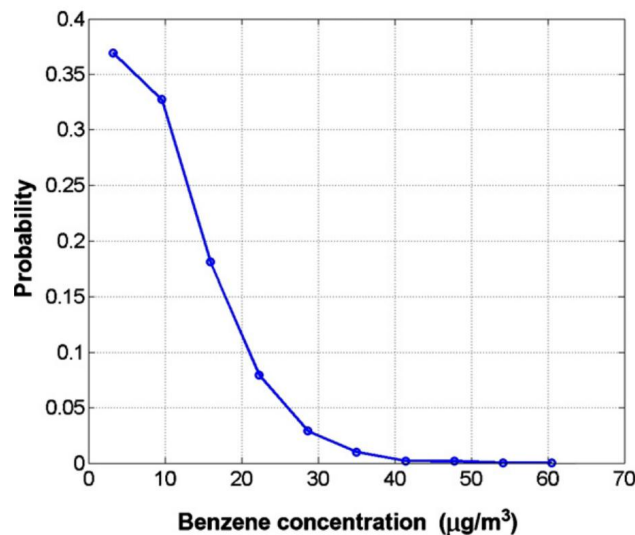
Fig. 1. Empirical probability distribution function of benzene concentration measured during the on field measurement campaign.
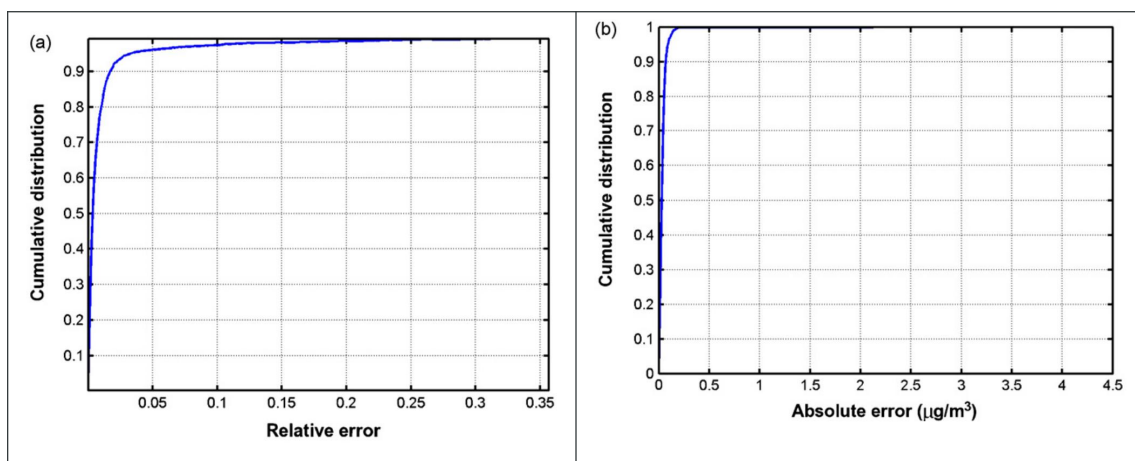
Fig. 2. Empirical cumulative distribution function of (a) relative and (b) absolute estimation error for benzene concentration (8-month training length). About 95% of test samples show a relative error smaller than 5% while about 99% of test samples show an absolute error smaller than 0.25 microg/m3 .
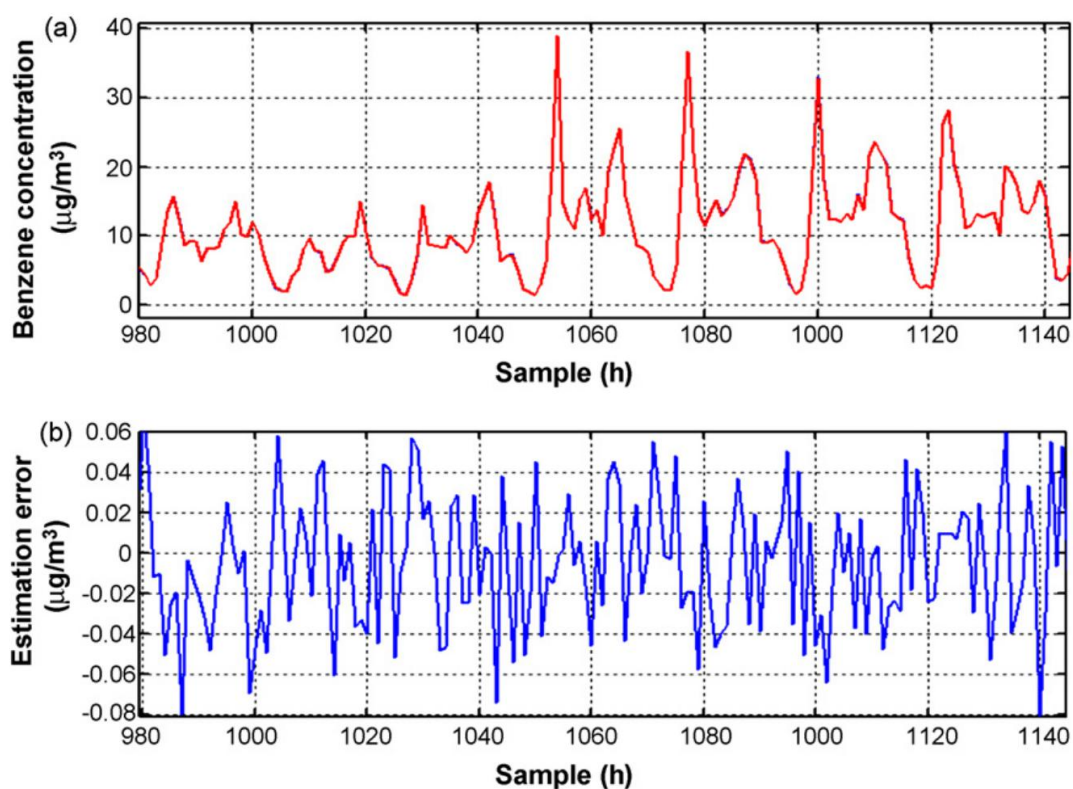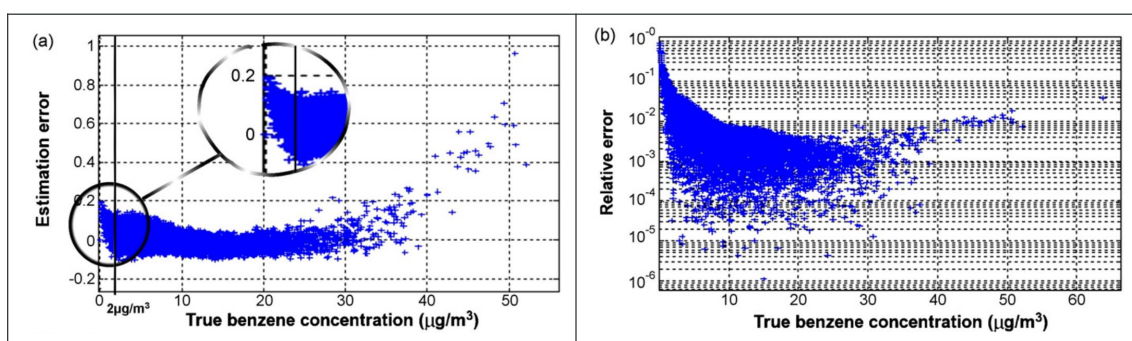

Fig. 3. (a) Benzene concentration estimation and (b) absolute error during 1 week (samples vs. hourly concentration and estimation absolute error in microg/m3). Results obtained with the 8-month training length preliminary run: fixed station output (red) almost totally overlaps network estimations (blue) over a 1-week test subset length.

Table 4

Training length	MRE	STD_RE	MAE ($\mu g/m^3$)	STD_AE ($\mu g/m^3$)	SCC
24 h	0.50	1.07	5.09	5.63	0.4383
96 h	0.16	0.24	1.31	1.65	0.8849
10 days	0.020	0.11	0.13	0.25	0.9938
25 days	0.017	0.071	0.081	0.19	0.9986
50 days	0.012	0.057	0.050	0.19	0.9984
100 days	0.009	0.076	0.050	0.16	0.9995

Relationship between training set length and performance indexes for benzene concentration prediction.Results show, as expected, a definitely positive trend extending the duration of training set. Surprisingly, a 10-day training length is sufficient to achieve a 2% relative error; still, error standard deviation remains significant.

Fig. 5. (a) Test set estimation error expressed in microg/m3 : between 3 and 30 microg/m3 of benzene concentration, we report an almost unbiased estimation with error limited in the −0.1 to 0.1 microg/m3 range. At higher values (i.e. >30 microg/m3) significant over-estimation errors occur. Relative error (logarithmic plot) remains under the 10 −2 order of magnitude unless below the 2 microg/m3 concentration level where a definite relative over-estimation trend is shown.
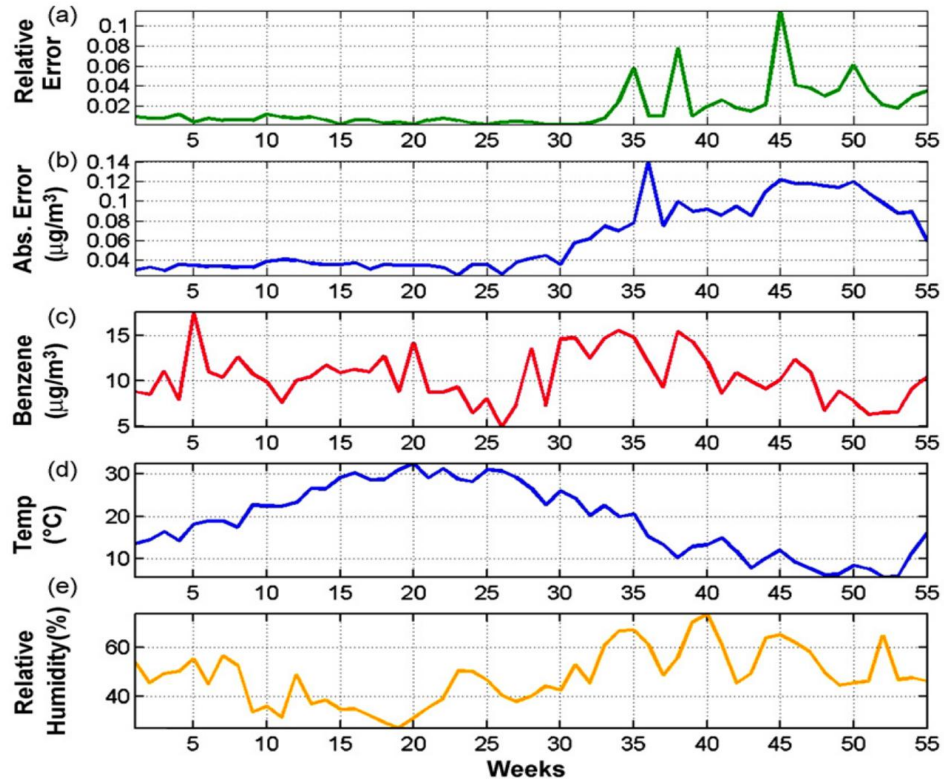
Fig. 4. Benzene concentration estimation results in the 10-day training length run (weekly averages): (a) MRE; (b) MAE; (c) benzene concentrations as measured by a conventional station; (d) temperature measured by multi-sensor device; (e) relative humidity measured by the multi-sensor device. Results are affected by what we expect to be seasonal meteorological effects, evident after the 30th week (starting of November) superimposed to a slow degradation due to sensor aging effects (drifts). After the 50th week the absolute error shows a definite recovery trend.
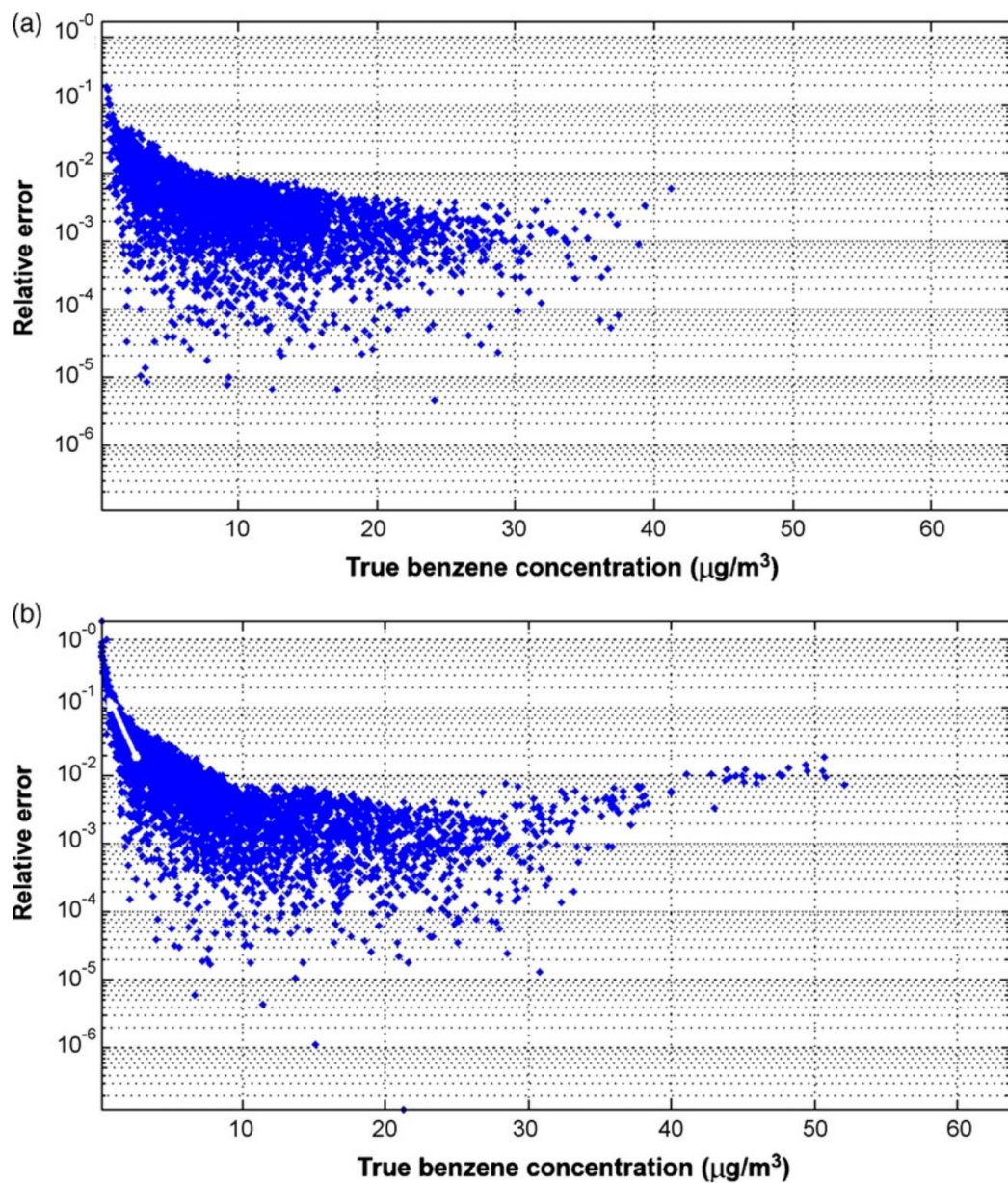
Fig. 6. Scatter plot of test set relative estimation error vs. the true benzene concentration in (a) the first 30 weeks and (b) the following 25 weeks. Distribution variations can be observed at lowest values where relative errors became more significant in the second period. Note that, in this period, a limited number of measurement points with very high values of benzene concentration are also present. The absence of such measurements in the training set is probably the cause of model misbehaviour (significant absolute errors).
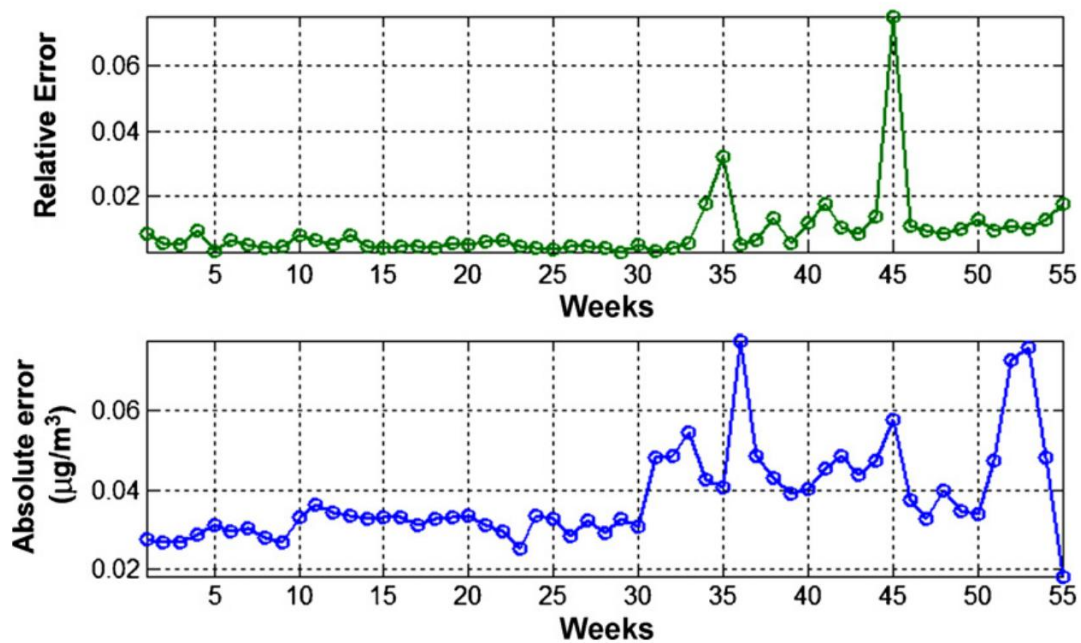
Fig. 7. Performance with the use of a second calibration performed at 34th week (weekly averages). With respect to Fig. 4 we can observe a definitely better behaviour with lower values of both absolute and relative errors. However, peaks in relative error seem to persist in the same weeks showing a slight difficulty for the obtained model to deal with relatively low-benzene concentrations during winter time.

Conclusions and further works

We have evaluated the use of neural networks together with on field data recordings for calibrating a multi-sensor device for benzene estimation. The scenario is characterized by significative correlations among several pollutant species. The proposed sensor fusion subsystem has been selected for exploiting both single sensor specificity and scenario-related correlations. Performance was assessed over a 1-year period, coping with both sensor aging and scenario-related seasonal influences. The latter have shown to have a stronger influence in the short-term causing definite trends in performance expressed by the proposed methodology. Results show that it is possible to obtain, using a limited number of days as the training period, a neural calibration capable to let the featured multi-sensor device to operate in the field, successfully carrying out benzene estimation with optimal results for more than 6 months. This is very encouraging as regarding the possibility to obtain a similar calibration for different species and hence to use the

multisensor device for densifying the conventional urban pollution monitoring network. Performance estimators were characterized and an almost steady, no bias absolute error distribution has been reported in the concentration range of interest. However, the neural model seems to over-estimate when fed with very low or very high benzene concentrations. This is very likely due to the limited number of samples with the characteristics recorded in the training phase. The problem is intrinsic when using only on field data that do not allow to control the training set gas concentration distribution. Generalization properties of the selected sensor fusion algorithm, hence, are decisive if using this methodology. Overall performances show also a slight degradation over time that affects primarily low-concentration values; significant degradation attenuation can be obtained performing a further calibration near the beginning of winter season. We are now focusing on the other pollutant species and on the portability level of the obtained calibration in order to fully exploit the capabilities of the proposed architecture to complement an urban atmospheric overall pollution monitoring network. The obtained calibration could be strongly dependent on the particular instance of the sensor array or even on the particular site eventually featuring a site-related characteristic pattern for the concentrations of several pollutants. Further works should clarify, in fact, the possibility to use a single calibration withdifferent multi-sensor devices or to use a calibration for applications in a mobile pollution monitoring scenario.

References

[1] D. Mage, G. Ozolins, P. Peterson, A. Webster, R. Orthoferj, V. Vanderweerds, M. Gwynnet, Urban air pollution in megacities of the world, Atmos. Environ. 30 (1996) 681–686.

[2] D. Dockery, C.A. Pope, X. Xu, F. Speizer, J. Schwartz, An association between air pollution and mortality in six US cities, N. Engl. J. Med. 329 (1993) 1753–1759.

[3] N.A. Mazzeo, L.E. Venegas, Evaluation of turbulence from traffic using experimental data obtained in a street canyon, Int. J. Environ. Pollut. 25 (2005) 164–176.

[4] R. Sozzi, La micrometeorologia e la dispersione degli inquinanti in aria, APAT CTN-ACE 16-17, 2003, pp. 387–395.

[5] B. Croxford, A. Penn, B. Hiller, Spatial distribution of urban pollution: civilizing urban traffic, in: Proceeding of the 5th Symposium on Highway and Urban Pollution, Copenhagen, 1995.

[6] H. Nakagawa, S. Okazaki, S. Asakura, K. Fukuda, H. Akimoto, S. Takahashi, S. Shigemori, An automated car ventilation system, in: Proceeding of the 7th International Meeting on Chemical Sensors, Beijing, 1988, pp. 187–189.

[7] M.C. Carotta, G. Martinelli, L. Crema, M. Gallana, Merli F., G. Ghiotti, E. Traversa, Array of thick film sensors for atmospheric pollutant monitoring, Sens. Actuators B: Chem. 68 (2000) 1–8.

[8] C. Pijolat, C. Pupier, M. Sauvan, G. Tournier, R. Lalauze, Gas detection for automotive pollution control, Sens. Actuators B: Chem. 59 (1999) 195–202.

[9] G. Huyberechts, P. Szekowka, J.R.N. Licznerski, Simultaneous quantification of carbon monoxide and methane in humid air using a sensor array and an artificial neural network, Sens. Actuators B: Chem. 45 (1997) 123–130.

[10] M. Pardo, G. Faglia, G. Sberveglieri, M. Corte, F. Masulli, M. Riani, A time delay neural network for estimation of gas concentrations in a mixture, Sens. Actuators B: Chem. 65 (2000) 267–269.

[11] S. De Vito, A. Castaldo, F. Loffredo, E. Massera, T. Polichetti, I. Nasti, P. Vacca, L. Quercia, G. Di Francia, Gas concentration estimation in ternary mixtures with room temperature operating sensor array using tapped delay architectures, Sens. Actuators B: Chem. 124 (2007) 309–316.

[12] D. Then, A. Vidic, C. Ziegler, A highly sensitive self-oscillating cantilever array for the quantitative and qualitative analysis of organic vapor mixtures, Sens. Actuators B: Chem. 115 (2006) 322–329.

[13] B.H. Kim, Multicomponent analysis and prediction with a cantilever array based gas sensor, Sens. Actuators B: Chem. 78 (2001) 12–18.

[14] J. Brezmes, E. Llobet, S. Al Khalifa, S. Maldonado, J.W. Gardner, Support Vector Machines: Theory and Applications, Berlin, Springer, 2005, pp. 365–386.

[15] S.A. Khalifa, S. Maldonado-Bascón, J.W. Gardner, Identification of CO and NO 2 using a thermally resistive microsensor and support vector machine, IEEE Proc. Sci. Meas. Technol. 150 (2003) 11–14.   757

[16] G. Martinelli, M.C. Carotta, G. Ghiotti, E. Traversa, Thick film gas sensors based on nano-sized semiconducting oxide powders, MRS Bull. 24 (1999) 30–36.

[17] M. Kamionka, P. Breuil, C. Pijolat, Calibration of a multivariate gas sensing device for atmospheric pollution measurement, Sens. Actuators B: Chem. 18 (2006) 323–327.

[18] M.C. Carotta, G. Martinelli, L. Crema, C. Malagu, M. Merli, G. Ghiotti, E. Traversa, Nanostructured thick-film gas sensors for atmospheric pollutant monitoring: quantitative analysis on field tests, Sens. Actuators B: Chem. 76 (2001) 336–342.

[19] W. Tsujita, A. Yoshino, H. Ishida, T. Moriizumi, Gas sensor network for air-pollution monitoring, Sens. Actuators B: Chem. 110 (2005) 304–311.

[20] Pirelli Laboratories. Web Site: http://www.pirelli.com/web/group/labs/default.page.

[21] C.M. Bishop, Neural Networks for Pattern Recognition, Oxford University Press, 1995, ISBN 0-19-853849-9.

[22] M. Pardo, G. Sberveglieri, Remarks on the use of multilayer perceptrons for the analysis of chemical sensor array data, IEEE Sens. J. 4 (2004) 355– 363.

[23] D.W. Marquardt, An Algorithm for least squares estimation of nonlinear parameters, SIAM J. 11 (1963) 431–441.

[24] M.T. Hagan, M.B. Menhaj, Training feedforward networks with the marquardt algorithm, IEEE Trans. Neural Networks 5 (1994) 989–993.