CO, NO₂ and NOₓ urban pollution monitoring with on-field calibrated electronic nose by automatic bayesian regularization

Saverio De Vito [a,*], Marco Piga [b], Luca Martinotto [b], Girolamo Di Francia [a]

[a] ENEA, Centro Ricerche Portici, 80055 Portici (NA), Italy
[b] Pirelli Labs, Viale Sarca 222, 20126 Milano, Italy

ABSTRACT

Low cost gas multisensor devices can represent an efficient solution for densifying the sparse urban air pollution monitoring mesh. In a previous work, we proposed and evaluated the calibration of such a device using short term on-field recorded data for the benzene pollution quantification. In this work, we present and discuss the results obtained for CO, NO₂ and total NOx pollutants concentration estimation with the same set up. Conventional air pollution monitoring station is used to provide reference data. We show how a multivariate calibration can be achieved with the use of two weeks long on-field data recording and neural regression systems. Also for these pollutants, no significant performance boost was detectable when longer recordings were used. The influence of an appropriate feature selection for achieving optimal performances is also discussed comparing long term performance results of the obtained calibrations. Benefits and issues of multivariate correlation based calibration are evaluated during one year long measurement campaign.

© 2009 Elsevier B.V. All rights reserved.

1. Introduction

Urban pollution monitoring is currently carried out with networks of industrial spectrometers based stations. Their dimensions and costs make unrealistic the realization of a measurement mesh with appropriate density. The pollution diffusion dynamic inside cities is indeed very complex due to the intricate pattern of roads and buildings. Actually, the typical monitoring mesh granularity exceeds several kilometres and this can hamper the representativity of gathered data in terms of the actual pollutants concentration distribution in city areas. As an example, roads surrounded by high buildings can undergo very high pollution levels due to the so called canyon concentration effect even if a nearby monitoring station report extremely low pollution levels [1–3]. On the other hand, real time knowledge of the actual pollution concentration pattern in city areas is considered crucial for implementing savvy traffic management and air pollution management policies [4].

During the last years, some authors have proposed the use of gas multisensor devices as a tool for densing the urban pollution monitoring mesh due to the significantly low cost per unit [4,6–9].

Their reduced dimensional impact make them suitable even for cities historical centres where they can be easily hidden preserving architectural heritage fruition.

Unfortunately, the intrinsic long term stability and selectivity issues that affect the solid state sensors they rely on, can severely limit their reliability, in particular when compared with the well established performances of conventional analyzers. Some studies confirms how, when calibrated for single specie quantification in laboratory, they exhibit poor performances when operating in real conditions, revealing a low reliability of conventional in-lab calibration process when dealing with complex mixtures [5,6].

Recently, some on-field calibration strategies, have been proposed to overcome these issues [6–8]. In particular, our group proposed the use of a spectrometers based station as a reference for the on-field calibration of a small multisensor device obtaining very encouraging results for benzene quantification purposes. We also focused on the optimal duration of the calibration process finding that a neural network, trained using a 10 days measurements set, was capable to gain a relative estimation error of less than 0.02 (2%) over more than 6 months [8]. A long term performance degradation was observed and interpreted as caused by modification in different species absolute and relative concentration in the Winter time. A recalibration procedure, performed every 6 months, has been found to effectively restore the performance level. In a similar frame, Tsujita et al. proposed an interesting automatic on-field

* Corresponding author. Tel.: +39 0817723364; fax: +39 0817723344.
E-mail address: saverio.devito@portici.enea.it (S. De Vito).

recalibration scheme that could be useful for facing at least the sensor stability issues [9].

Pattern recognition algorithms are often used as automated multivariate analysis tools in gas sensing but their performances are rarely discussed especially for regression problems, e.g. quantitative concentration estimation problems [10–11]. In these problems, sensor fusion subsystems are devised for approximating Ψ in:

$$C_j = \Psi(RSens_1, RSens_2, \ldots RSens_n)$$

where C_j is the real jth pollutant concentrations as measured by conventional stations and $RSens_i$ is the ith sensor resistance measured by the multi-sensor device.

In a complex scenario such as urban pollution monitoring, an in-deep analysis of the results obtained by such a system can significantly help the feasibility assessment, guiding the development of appropriate monitoring strategies (device positioning, calibration methodology and long term performance assessment) as well as the instrumentation design and validation process.

In this paper, we study the performance assessment of on-field calibrated multisensor device, presenting the results obtained for CO, NO$_x$ and NO$_2$ concentration estimation. We analyze the results trying to reach an in deep knowledge on the main performance drivers in the selected scenario validating the hypothesis by correlation analysis, feature selection and long term performance assessment.

2. Experimental setup

A multisensor device, developed by Pirelli Labs (see Ref. [12]), has been co-located with a conventional air pollution analyzer, operated by the regional environmental protection agency (ARPA). The conventional analyzer response has been used to provide the true concentration values of the target pollutants at the measurement site. These values have been hence used as a reference for the tuning of a multivariate regression system designed for calibrating the multisensor device response. This multisensor device was designed to host five custom metal oxide chemoresistive sensors plus two commercial relative humidity and temperature solid state sensors. It is built on a very compact design (volume = 9.7×10^{-3} m^3) and was easily deployed in the operative scenario due to its limited weight (less than 2.5 kg). An on board microcontroller provided the computing platform, controlling analog to digital conversion, while an on board secondary memory device provided local storing capacity for up to 72 h of data measurement, with a 8 s sampling period. Hourly average of measurements was being transmitted via a GPRS modem toward PC class data sinks. A more detailed instrumentation and setup description can be found in [8].

Conventional fixed station provided reference concentration estimation for CO (mg/m^3), non-metanic hydrocarbons (NMHC) (μg/m^3), C$_6$H$_6$ (μg/m^3), NO$_x$ (ppb), NO$_2$ (μg/m^3). It was sampled recording hourly averages of the concentration values. Unfortunately, after only 8 days the NHMC analyzer went out of service. The multisensor device was sampled to provide the hourly average of the resistivity expressed by CO, NO$_x$, O$_3$ and NO$_2$ specific metal oxide (MOX) chemiresistors, a NMHC targeted MOX sensor whose characteristics are also detailed in [8], plus the commercial temperature and relative humidity sensors. Measurement campaign took place using as testing site one of the main street in the centre of an Italian city, characterized by heavy car traffic. Data acquisition campaign lasted from March 2004 until April 2005, building up a suitable one year long data set for the devised application.

3. Experimental scenario and methods discussion

According to the results of our previous investigation, the testing site was characterized by the presence of significant correlations among the recorded species concentrations time series [8]. In order to generate a non-linear multivariate calibration for the multisensor device, a back propagation neural network architecture was designed to solve a typical regression problem. The system was actually trained to estimate real pollutant concentrations hourly mean, as provided by the conventional analyzer, using the hourly mean of electrical resistance expressed by the multisensor device chemiresistors used as input features.

In the framework of statistical pattern recognition, full regression problems are characterized by the need of approximating the 'a posteriori' distribution of target random variables given the values of the measured variables (see [10]). Practically, the typical statistical regression system works by approximating the functional relationships (if existing) among measurements in the 'feature space' and expected output values starting from a predefined internal model and using a limited data set (training set) as examples to tune it. If the training set choice is representative of the real data statistical distribution we can expect the trained regressor to be able to correctly estimate the searched functional relationships retaining appropriate generalization properties. Of course, noise both on the measurement and target variables will affect the overall estimation performance at various extents.

In our pollutant concentration estimation scenario, we can expect reasonably good performances if the sensors pool, actually composing the feature space, includes a sensor unit showing good selectivity and stability properties for each target pollutant. In this case univariate calibration can represent a viable solution. Usually, this is very difficult to obtain for the intrinsic properties and limitations of solid state sensors.

Nice performance levels could be expected even if the target specific sensor is prone to significant selectivity issues but, in this case, a properly devised sensor pool could be capable to describe the influence of interferents on the target-specific sensor response. This can, for instance, occur if a sensor pool subset response is strongly correlated to interferents concentrations. In this case, a statistical regressor could extract sufficient knowledge about interferents concentrations and their effects on the other gas sensors responses, so to model such influences during the training phase and provide induced error cancellation. For instance, if CO concentration estimations are affected by the presence of a particular interferent specie, say H$_2$S, a sensor fusion subsystem can learn to estimate and subtract the influence on the CO targeted sensor given you can measure H$_2$S concentrations and you have a representative set of training samples to model it.

In our scenario, the usual presence of significant correlation among pollutant concentrations can also help a statistical regressor to extract knowledge about a particular gas concentration even if the related sensor performances are low or, ultimately, if no specifically targeted sensor is present in the measurement sensor pool. In this case the target gas concentration is estimated starting from the response of sensors towards gases whose concentrations are strongly correlated, not necessarily linearly, with the concentration of the pursued gas. This feature can have a strong drawback due to possible changes in the relative concentrations of gases on seasonal or spatial basis, as it is, for example, the differential increase of major pollutants concentration (see for example Ref. [3]) and the increase of NOx emissions by domestic heating both measurable in the Winter season. Actually, these changes could severely hamper the significance of the training set related knowledge.

Last paper was focused on the results of benzene concentration estimation. We showed how, in that case, it was possible to estimate the concentration of a pollutant for which no specific

sensor was present in the selected sensor pool, using the information brought in by the NMHC sensor. Actually, Benzene is the most significant part of NMHC related pollution in the selected scenario. We also found a measurable performance degradation using a Spring time obtained calibration during Winter time due to the above mentioned house heating pollution related effect.

In this work, we show results obtained calibrating feed forward neural networks in regression scheme for NOx, NO$_2$ and CO pollutant concentrations, in the same applicative scenario so to check the generalization of benzene related findings. Investigations on the performance relationships with training length have been performed in order to compare the results with those obtained for benzene. Influences of feature selection and correlation based learning have been investigated and evaluated.

Using conventional station output, neural networks have been trained and simulated in the Matlab environment; hyperbolic tangent has been selected as hidden neuron transfer function. Levenberg–Marquardt training algorithm was chosen for the experimentations, early stopping and automatic Bayesian regularization (ABR) were used as complexity control algorithms to avoid over training issues [13,14]. Training sets were built using consecutive samples, i.e. using fixed length specific interval of campaign data, while validation (when used for early stopping) was built by randomly choosing samples from the remaining data. Two approaches have been devised for uncertainty estimations on performance and for generalization purposes, one of which based on crossvalidation.

Automatic Bayesian regularization is a neural network capacity control technique for neural networks training; it was introduced in the late 90 s, but its usage is still uncommon in the electronic nose community. Training the regressor internal model usually involves the minimization of an error function. In conventional neural networks back propagation training, the error function basically take only into account the empirical error (i.e. the error on training set) while model capacity (i.e. the number of hidden neurons) is fixed and solution complexity should be carefully controlled. This training data centric procedure can easily lead to data overfitting. Approaches based in regularization theory (see Ref. [10] for a detailed description) try to balance the empirical error with a term related to model complexity, their global error function being described by:

$$\beta \sum_n^k \xi_n + \alpha \|w\|^2 \tag{2}$$

where n is the running training samples index, ξ_n the empirical error function computed on the nth of k total training samples and α, β are the regularization parameters. In neural networks, the $\|w\|$ term take into account the network weights values, trying to estimate network model complexity by their norm. In appropriate conditions, pursuing the minimization of (2) produces minimization of such superfluous connection weight values that could produce overtraining. This typically allows also for a smoother network response, and the effect is similar to what obtainable by superfluous network connections pruning. In fact, as regard as network architecture, by pursuing regularization, any modestly oversized network should be able to approximate regression function retaining good generalization properties.

In order to avoid both underfitting and overtraining issues, α/β value should be, of course, carefully tuned and this adds free parameters, to be optimized.

ABR is a regularization based training methodology that try to estimate α and β, in a Bayesian framework (see [10,15]), starting from initial hypothesis on weights statistical distribution, while simultaneously minimizing (2) by using Levenberg–Marquardt network training algorithm. In the Bayesian framework weights

are considered as random variables as their posterior probability density can be described as:

$$P(w|D, \alpha, \beta, M) = \frac{P(D|w, \beta, M)P(w|\alpha, M)}{P(D|\alpha, \beta, M)} \tag{3}$$

D being the gathered data set and M the particular neural network model.

Assuming Gaussian distribution for prior network weights distribution and training set noise, the optimal weights maximizing (3) also minimize the regularized error function given in (2). In ABR, each network training step adds knowledge on weights distribution practically obtaining an approximation, say w', of the weight values w^{MP} that maximize posterior probability in (3). This knowledge can be used to update suboptimal α and β values pursuing the maximization of

$$P(\alpha, \beta|D, M) = \frac{P(D|\alpha, \beta, M)P(\alpha, \beta|M)}{P(D|M)} \tag{4}$$

In fact, adding the likely assumption of $P(\alpha, \beta|M)$ being uniform then, after simple analytical steps, approximations of α^{MP} and β^{MP} can be found using (3):

$$\alpha' = \frac{\gamma}{2\|w'\|^2}, \quad \beta' = \frac{k - \gamma}{2\sum_n \xi_n(w')} \tag{5}$$

with γ being the so called effective parameters number, i.e. a measure of the number of neural network parameters that are effectively in use of error function minimization. The value of γ is computed as

$$\gamma = N - 2\alpha \cdot \mathrm{tr}(H)^{-1} \tag{6}$$

using the Hessian matrix H of the objective function whose Gauss–Newton approximation is readily available by using Levenberg–Marquardt training algorithm.

The obtained approximation of α^{MP} and β^{MP}, in turn, modifies Eq. (2) and hence the best w values obtainable by a further training iteration. The process is iterated until convergence. Note that, following network weights initialization, α and β are initialized at 0 and 1, respectively, before taking the first step of the training algorithm.

4. Results and discussion

Initially, in order to allow the reader to compare the error level for CO, NOx and NO$_2$ with benzene related findings reported in our previous work, a preliminary regression experiment has been set up. A two weeks long data segment, starting from the first day of measurements and using all the sensor responses, has been used as training set. Only in this preliminary experiment, a validation set consisting of 40% of the remaining samples was randomly extracted and reserved for the implementation of early stopping strategy. Remaining samples were used as test set for performance evaluation. Three different network were devised each one targeted to a specific analyte concentration estimation problem and Levenberg–Marquardt algorithm used for training purposes. The number of hidden layer was empirically fixed to 25 for all the three specie while results assessment was repeated 20 times and averaged in order to reduce local minima and initial weight choice influences. Mean relative error (MRE) computed as

$$MRE = \frac{1}{n} \sum_{i=1}^{n} \frac{|y_i - x_i|}{|y_i|} \tag{7}$$

where y_i is the true concentration value at the ith sample belonging to the test set, x_i the network concentration estimation and n the number of test samples involved, was set as primary performance

estimator both for this and for the following experimentations. We found $MRE_{(NO_2)} = 0.26$, $MRE_{(NO_x)} = 0.42$ and $MRE_{(CO)} = 0.32$. Although MRE figures are significantly worse than those computed for benzene with similar settings (MRE = 0.02), even in this case the ratio between mean absolute error (MAE), computed as

$$MAE = \frac{1}{n} \sum_{i=1}^{n} |y_i - x_i| \tag{8}$$

and the single specie yearly concentration range was encouraging, setting at 3% for CO, 8% for NO_2 and 7% for NOx. As a reference, their concentrations units of measure was, respectively, mg/m^3, $\mu g/m^3$, ppb. The most relevant negative impact on relative error was found at low concentrations levels typically reached at night.

It is also interesting to note that, for all these three species, the performance obtained on the normalized training set in terms of mean squared error (MSE) are typically by three order of magnitude than those obtained for benzene in the same setup. This result is explained with a significant "noisy" behaviour of feature space variables (sensor resistances) in response to these specie concentrations confirming the definitely worse descriptive power, in operative conditions, of the sensor array as regards as these species concentrations.

It should be underlined that the use of early stopping makes the above results depend on the merge of training and validation set related knowledge and not only on the training set data alone. However, similar but more rigorous results, as regards as training set length influence on performance, were obtained without the use of any validation set and are reported in the following paragraphs that represent the core of this paper, building up the basis for the drawn conclusions.

4.1. Optimal training length estimation

As a second investigation, we wanted to test the optimal training duration when using consecutive samples. In the benzene case, a 10 days long training set proved capable to recover the expected daily and weekly cycle of the pollutant concentration statistical distribution variation. Our hypothesis was that results obtained for benzene could be generalized as well for the other species.

For this reason, we set up a training-evaluation procedure using the whole sensor array response and different training set lengths for CO and NO_2. Networks training was conducted using Levenberg–Marquardt algorithm with automatic Bayesian regularization, for complexity control, so without the use of any validation set. Hidden layer number was initially set at 5 by using preliminary results.

As a first experiment we set up different length consecutive samples training set starting from the first sample of the available data set (March 1). For each selected training length, all the remaining samples were used as test set. In fact, each training duration produced a different regressor with different performances that we wanted to compare by using different length test sets. In a similar setup it is important to evaluate uncertainties on the performance estimation in order to understand significance of the reported differences possibly achieving generalization of findings. The primary source for uncertainties on performance estimation that should be tackled is the use of a limited and different duration test set. It is also well known that the outcome of the training process of a neural network itself, is prone to uncertainties due to the initial choice of weights and related local minima issues.

So, for each trained network characterized by its own (training, test) set couple, concentration estimation errors observations on test set were computed. For each training length the procedure has been repeated for 20 times in order to reduce the impact of initial NN weights choice, however variability on the performance

indicators was negligible except in the case of CO concentration estimation for a training length of 96 h for which test results were unstable.

As regards as uncertainty estimation, errors observation appeared not to be i.i.d.; actually they shown to be significantly autocorrelated. In fact, their autocorrelation function (ACF) was found definitely not being impulsive, instead absolute error ACF was only slowly fading under the white noise confidence intervals for ACF. For this reason, confidence intervals on samples mean could not be computed using CLT and t statistic just because observations failed to meet the i.i.d. (independent and identically distributed) requirement. In order to correctly compute the confidence intervals for the MRE and MAE, a novel algorithm, proposed by Zhang in [16] for the calculation of the uncertainty on the sample mean of autocorrelated measurements, was used. This methodology is based on a weak stationarity hypothesis that allow for an Auto regressive moving average modeling of the measurement process. It is interesting to note that the so computed confidence intervals are notably wider that the classic ones as one can expect given a lower number of independent variable observations in autocorrelated time series. It should be also noted that absolute errors observations show a notable trend in time, in this situation, author warn that the use of the MAE averaged value itself and the corresponding variance to characterize the regression systems may be misleading. For this reason, although they lead to the same conclusions, we based our comparisons on MRE computed values as a primary performance indicator. Results for CO and NO_2 have been reported in Tables 1 and 2, respectively. They show that for that year no particular benefit could be achieved by extending the duration of a training set starting form the first available sample for more than two weeks, when estimating pollutant concentrations over the remaining samples of that year.

A further experiment has been conducted for the sake of comparison, i.e. for comparing multiple training set durations in a slightly different setup. This time, we have used constant test set duration and fixed test set location in the data set time series. All the regressors were tested on the same test set i.e. the campaign final 6 months. Training sets start time, training procedures and regressors' architecture were the same as in the previous experiment. Although this setup may be correctly perceived as penalizing for short length training sets whose entire samples are located much before the start of test set, the results may represents an ideal comparison for the above presented ones. Results, presented in Table 3 shows a substantial confirmation of what already found.

Training and test localization in a particular time in the above setup may cast doubts on the possibility to generalize findings. Of course, the same apply to the training set choice process. For this reason, in order to compare and generalize performance figures resulted from the use of different training set length choices we decided to apply a particular crossvalidation approach that could take into account training (and test samples) located in different time locations. We set up the following crossvalidation procedure: in the entire campaign year, considering a fixed test set length h corresponding to 6 months, and for each selected training set duration i, then k mutually exclusive, consecutive training sets Tr_m^i, m being the training set index, have been considered, with

$$k = \frac{Total\ DB\ Lentgh - h}{i} \tag{9}$$

Actually, for each training set choice Tr_m^i, made by i consecutive samples, the h immediately following consecutive samples have been considered as its specific test set Te_m^k. For each Tr_m^i, Te_m^h couple, the neural network has been trained with ABR for 20 times. MRE, relative error standard deviation (STD_RE), MAE, absolute error standard deviation (STD_MAE) and squared correlation coefficient (SCC) computed over each Te_m^h, were used as perfor-

Table 1
CO concentration estimation performances of the neural regression scheme, computed over different training set lengths. All sensor responses have been used for feature vector composition, five neurons have been used in the hidden layer, ABR has been used for complexity control. For each training set choice, the remaining samples of the dataset have been used as test set. MAE figures are expressed in mg/m^3. Confidence intervals are reported for MRE and MAE. Only small performance enhancements are measurable when using more than two weeks data recording.

Hours	MRE	CI	STD_RE	MAE (mg/m^3)	CI (mg/m^3)	STD_MAE (mg/m^3)	SCC
24	0.49	±0.13	1.08	0.58	±0.10	0.43	0.77
96	0.78	±0.37	1.37	1.56	±0.31	1.19	0.18
240	0.34	±0.09	0.86	0.53	±0.16	0.63	0.79
360	0.32	±0.07	0.84	0.51	±0.10	0.51	0.82
600	0.31	±0.09	0.85	0.44	±0.13	0.46	0.86
1200	0.32	±0.13	0.98	0.41	±0.10	0.42	0.85
1800	0.31	±0.13	0.99	0.41	±0.10	0.42	0.84
2400	0.29	±0.11	0.74	0.46	±0.11	0.49	0.87

Table 2
NO$_2$ concentration estimation performances of the neural regression scheme computed over different training set lengths. All sensor responses have been used for feature vector composition, five neurons have been used in the hidden layer, ABR has been used for complexity control. MAE figures are expressed in µg/m^3. Confidence intervals are reported for MRE and MAE. For NO$_2$ a small performance worsening is observed when extending the ten days length training set.

Hours	MRE	CI	STD_RE	MAE (µg/m^3)	CI (µg/m^3)	STD_AE (µg/m^3)	SCC
24	0.60	±0.23	1.04	49.4	±9.53	35.0	0.09
96	0.31	±0.10	0.68	30.2	±10.1	25.4	0.36
240	0.22	±0.07	0.62	23.7	±7.10	21.0	0.66
360	0.21	±0.07	0.63	19.8	±5.32	18.3	0.71
600	0.23	±0.07	0.66	27.1	±8.70	25.9	0.62
1200	0.27	±0.07	0.65	26.5	±9.70	23.1	0.58
1800	0.26	±0.07	0.61	27.7	±10.0	24.6	0.59
2400	0.27	±0.10	0.68	28.2	±7.90	23.2	0.54

Table 3
CO concentration estimation performances of the neural regression scheme computed over different training set lengths. All sensor responses have been used for feature vector composition, five neurons have been used in the hidden layer, ABR have been used for complexity control. The last 6 months of the entire campaign dataset was used as a test set for all the different regressors. MAE figures are expressed in mg/m^3. Confidence intervals are reported for MRE and MAE. Again, only small performance enhancements are obtained by using more than two weeks data recording.

Hours	MRE	CI	STD_RE	MAE (mg/m^3)	CI (mg/m^3)	STD_MAE (mg/m^3)	SCC
24	0.49	±0.13	1.08	0.58	±0.10	0.43	0.77
96	0.87	±0.25	1.57	1.28	±0.17	1.11	0.11
240	0.36	±0.12	1.02	0.58	±0.07	0.67	0.77
360	0.38	±0.13	1.13	0.58	±0.07	0.68	0.74
600	0.35	±0.10	0.92	0.49	±0.10	0.50	0.82
1200	0.36	±0.16	1.15	0.43	±0.16	0.45	0.84
2400	0.35	±0.15	1.04	0.45	±0.10	0.47	0.84

mance indexes. Even in this case and for each couple, the results variability due to the NN training was found to be negligible, confirming ABR positive impact. Then, for each i, the k results coming from the different Tr$_m^i$, Te$_m^h$ couples have been considered as i.i.d. observations for performance figures and were finally averaged and their relative uncertainty computed so to correctly compare results coming from the different i dependant regressors.

In this case, we are evaluating the estimation capabilities of various regressors, each trained with a different length training set, tested on a fixed number of test samples immediately following the training set.

Obtained results for CO and NO$_2$ are shown in Tables 4 and 5, respectively. Fig. 1 depicts graphically MAE behaviour together with confidence intervals when changing training length for CO concentration estimation problem.

Single performance figures were significantly worse than that which has been found for the benzene case for which best MRE was found to be equal to 0.02. However, even for these two species, a relatively short length training set seems to be able to produce results that are very near to what obtainable with far longer data sets whether being significantly different from those obtainable at the shortest training set length. By using longer training set a slow positive trend is observable but uncertainty rapidly grow after 700 samples length for which we are not able anymore to reliably distinguish among average error performance. In fact, consider-
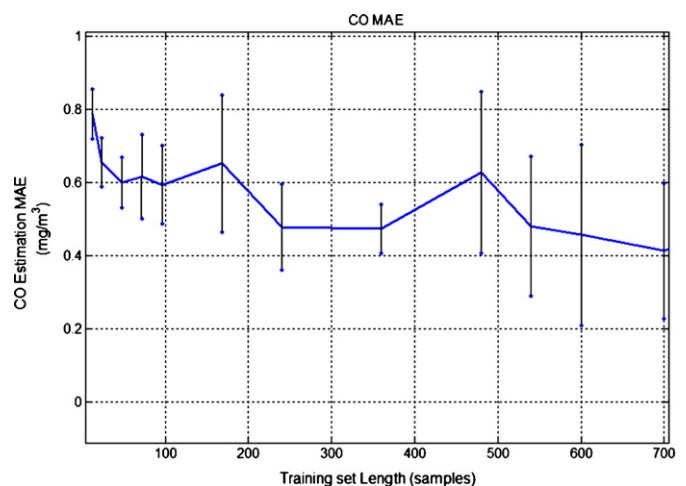

Fig. 1. CO concentration estimation MAE, expressed in mg/m^3, versus training set length measured in samples (hours) with related confidence intervals. All sensor responses have been used as feature vector (crossvalidation setting, see Table 4 for details).

Table 4
CO concentration estimation performances of the neural regression scheme computed over different training set lengths (fixed length test set, crossvalidation setting). All sensor responses have been used for feature vector composition, five neurons have been used in the hidden layer, ABR has been used for complexity control. Confidence intervals are reported for MRE, MAE and SCC. Welch test responses indicate that no significant performance enhancement could be obtained by extending training set length for more than two weeks.

Hours	MRE	CI	STD_RE	MAE (mg/m^3)	CI (mg/m^3)	STD_MAE (mg/m^3)	SCC	CI
12	0.55	±0.06	1.02	0.79	±0.07	0.67	0.76	±0.01
24	0.46	±0.05	0.94	0.66	±0.07	0.58	0.80	±0.01
48	0.42	±0.05	0.84	0.60	±0.07	0.54	0.82	±0.02
72	0.42	±0.07	0.82	0.61	±0.11	0.55	0.82	±0.02
96	0.39	±0.06	0.75	0.59	±0.11	0.53	0.82	±0.03
168	0.42	±0.11	0.80	0.65	±0.19	0.59	0.78	±0.08
240	0.34	±0.09	0.76	0.48	±0.11	0.49	0.84	±0.03
360	0.32	±0.08	0.76	0.45	±0.07	0.47	0.84	±0.03
480	0.38	±0.10	0.68	0.63	±0.22	0.62	0.78	±0.09
540	0.32	±0.14	0.71	0.48	±0.19	0.52	0.84	±0.02
600	0.30	±0.12	0.65	0.46	±0.24	0.45	0.85	±0.03
700	0.30	±0.18	0.76	0.41	±0.18	0.43	0.84	±0.13
800	0.34	±0.25	0.62	0.51	±0.44	0.48	0.84	±0.09
900	0.29	±0.18	0.80	0.40	±0.29	0.46	0.86	±0.12
1200	0.29	±0.24	0.75	0.40	±0.58	0.42	0.87	±0.03

Table 5
NO$_2$ concentration estimation performances of the neural regression scheme computed over different training set lengths (fixed length test set, crossvalidation setting). All sensor responses have been used for feature vector composition, five neurons have been implied in the hidden layer, ABR has been used for complexity control. Confidence intervals are reported for MRE, MAE and SCC. Welch test responses indicate that no significant performance enhancement could be obtained by extending training set length for more than two weeks.

| Hours	MRE	CI	STD_RE	MAE (μg/m^3)	CI (μg/m^3)	STD_AE (μg/m^3)	SCC	CI
12	0.46	±0.03	0.91	41.2	±3.22	30.8	0.48	±0.02
24	0.46	±0.04	0.91	42.7	±4.98	32.6	0.53	±0.02
48	0.44	±0.06	0.90	40.7	±6.81	32.8	0.54	±0.03
72	0.39	±0.04	0.78	36.9	±4.79	33.4	0.47	±0.05
96	0.40	±0.05	0.72	39.8	±7.89	34.35	0.51	±0.05
168	0.37	±0.06	0.75	36.1	±8.52	32.37	0.48	±0.09
240	0.34	±0.08	0.76	34.6	±11.1	33.8	0.48	±0.11
360	0.32	±0.13	0.73	32.6	±18.1	31.83	0.57	±0.12
480	0.31	±0.13	0.72	31.9	±16.7	34.9	0.51	±0.18
540	0.34	±0.32	0.72	36.5	±41.3	36.2	0.56	±0.17
600	0.28	±0.16	0.70	29.5	±26.0	30.0	0.48	±0.49
700	0.33	±0.35	0.70	33.9	±52.1	30.2	0.46	±0.66
800	0.28	±0.14	0.68	29.5	±20.7	27.4	0.49	±0.40
900	0.34	±0.61	0.74	37.6	±136	44.5	0.27	±3.28
1200	0.30	±0.81	0.68	30.4	±120	25.7	0.49	±2.44
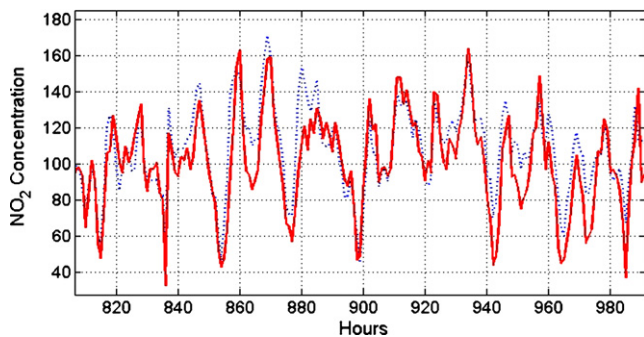
Fig. 2. Hourly concentration estimation of NO$_2$, expressed in μg/m^3, over one week period. Blue dashed line represent true concentration value as reported by the conventional analyzer. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

ing the performance figures observations as normally distributed (Kolmogorov–Smirnov test), a two samples right tailed Welch test[1], testing against an H_0 hypothesis of no significant difference in the mean, conducted for MAE observations in CO with test set lengths of 96 and 360 samples resulted in a $p = 0.02$, thus leading to reject

H_0. Conversely, null hypothesis cannot be rejected for any settings using 360 length test set together with longer training sets related results. Similar behaviour is found for MRE estimations, e.g. using a test length equal to 72 and 360 resulted in a $p = 0.04$. Similar tests have been conducted for NO$_2$ also, resulting in a substantial confirmation.

Overall results confirm the feasibility of the on-field approach for the calibration of multisensor devices in the city air pollution scenario, showing how a limited training campaign can be used for computing an optimal calibration for several of the following months. Absolute performance level however, depends on the descriptive power of the selected sensor array for the specific pollutant concentration estimation problem.

4.2. Feature selection

A further investigation involved testing the influence of a feature selection procedure on the obtainable performance in our scenario. Feature selection can broaden the experimentalist's knowledge about the information contribution brought in by the single sensor to the concentration estimation problem for each specie. This is particularly true if the selection is conducted using a brute force approach, i.e. exploring all available feature set combinations and their performance, though it can be very computationally expensive and hence it is not always a viable choice.

---

[1] Behrens–Fisher like settings, using Satterthwaite approximation for pooled variance at $\alpha = 0.05$.

Table 6
Performance evaluation for different feature vector compositions in the NO_2 concentration estimation problem. Ten days long training set has been used. Confidence intervals are reported for MRE and MAE.

CO	NMHC	NOx	NO2	O3	T	RH	MRE	MAE ($\mu g/m^3$)
X							0.31 ± 0.07	36 ± 13.0
	X						0.29 ± 0.09	29 ± 9.8
		X					0.43 ± 0.22	33 ± 6.7
			X				0.47 ± 0.13	50 ± 24.7
				X			0.29 ± 0.07	29 ± 8.25
		X		X			0.29 ± 0.11	26 ± 6.41
			X	X			0.36 ± 0.10	38 ± 18.3
		X	X				0.37 ± 0.11	34 ± 19.1
X	X						0.29 ± 0.09	29 ± 10.05
X		X	X				0.30 ± 0.08	32 ± 14.0
		X	X	X			0.35 ± 0.09	38 ± 16.9
			X		X	X	0.39 ± 0.14	45 ± 23.2
		X			X	X	0.28 ± 0.05	25 ± 4.53
		X	X	X	X	X	0.26 ± 0.06	28 ± 6.27
X	X			X			0.28 ± 0.08	27 ± 9.95
X	X		X				0.23 ± 0.10	19 ± 4.01
X	X	X	X	X			0.22 ± 0.07	19 ± 7.97
X	X	X	X	X	X	X	0.22 ± 0.08	20 ± 4.00
X	X	X		X			0.28 ± 0.08	29 ± 14.62
X	X		X	X			0.22 ± 0.09	20 ± 5.01
X	X		X	X	X	X	0.22 ± 0.09	20 ± 4.35

In our scenario, we have analysed the performances obtainable, in terms of MRE and MAE, by using different subset of the original sensor array for estimating NO_2, CO and NOx with separate regressors. Training set length was fixed to 10 days, starting from the first campaign day, according to the previous experimentations; the remaining samples has been used as test set. Confidence interval has been computed using Zhang method for uncertainty estimation for sample means of autocorrelated data series. Wilcoxon signed rank test (nonparametric paired samples with no normality assumption) has been used for evaluating the significance of observed performance differences at $\alpha = 0.05$. In order to estimate real operative system performance, feature selection validation and overall performance testing should be performed on different set, but our goal, as above mentioned, was different. i.e. investigating if there is significant effect of feature selection in this particular scenario.

Interestingly, for NO_2, the best performance values have been obtained using all the sensors responses; e.g. using NOx and NO_2 sensors responses, CO and NMHC sensors, or NOx, NO_2 and CO led to, respectively, 0.37, 0.29 and 0.30 MRE while using the entire gas sensor array led to a 0.22 MRE (see Table 6 and Fig. 2). Performance level that are not significantly different can be obtained by coupling NO_2 sensor response with both CO and NMHC sensors

response. This results confirms the possibility of achieving suitable performances, even when specific sensors are expressing low overall performance, by coupling their responses with information coming from sensors whose response is directed towards species showing high correlation factors with the specie under analysis (see [8]).

Table 6 confirms the influence of temperature and humidity on on-field specific sensor response calibration, in this case when NOx and NO_2 sensors are concerned. Their sensor response led to a significant performance increase when coupled with the response of single NOx sensor, single NO_2 sensor, NOx + NO_2 sensors, NOx + NO_2 + O_3 sensors. No improvement has been recorded when coupling RH and T sensor response to feature vectors containing CO, NMHC and NO_2.

Results obtained during CO directed sensors set selection, which are depicted in Table 7, revealed that a non-linear function of NMHC sensor response is able to follow the CO concentration significantly
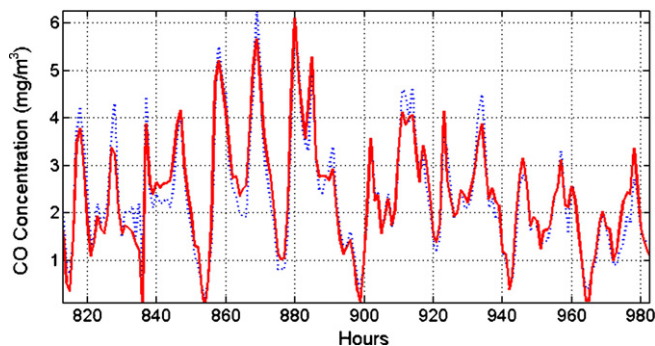

Fig. 3. Hourly concentration estimation, expressed in mg/m^3, of CO over one week period. Blue dashed line represent true concentration value as reported by the conventional analyzer. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)
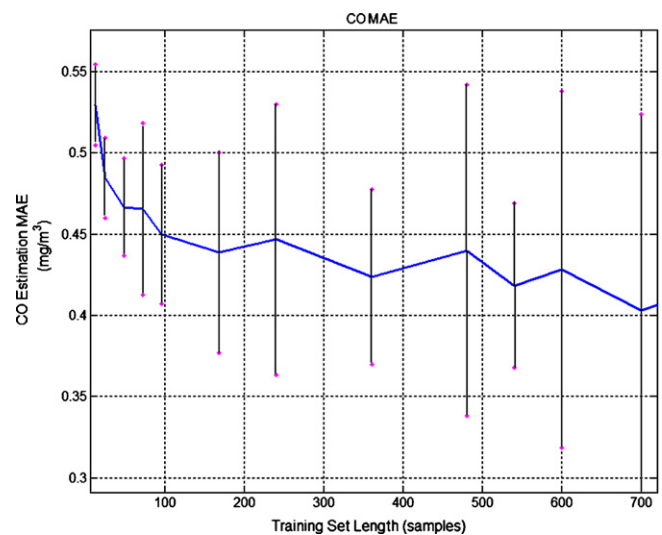

Fig. 4. CO concentration estimation MAE versus training set length measured in samples (hours) with related confidence intervals, only CO and NMHC sensor response have been used for feature vector composition (crossvalidation settings, see Table 9 for details).

Table 7

Performance evaluation for different feature vector compositions in the CO concentration estimation problem. Ten days long training set has been used. Confidence intervals are reported for MRE and MAE.

Feature set							MRE	MAE (mg/m^3)
CO	NMHC	NOx	NO$_2$	O$_3$	T	RH		
X							0.38 ± 0.05	0.72 ± 0.13
X	X						0.27 ± 0.06	0.35 ± 0.10
X	X				X	X	0.31 ± 0.10	0.43 ± 0.10
	X						0.34 ± 0.06	0.48 ± 0.06
X	X	X	X				0.36 ± 0.13	0.51 ± 0.16
X	X	X	X	X			0.33 ± 0.09	0.48 ± 0.14

Table 8

CO concentration estimation performances of the neural regression scheme computed over different training set lengths by using CO and NMHC solid state sensor response as feature vector. For each training length, starting from the start of campaign data time serie, the remaining data set samples has been used as test set. Five neurons have been used in the hidden layer, ABR has been used for complexity control. Confidence intervals are reported for MRE and MAE. No significant performance enhancement are obtained by using more than two weeks data recording.

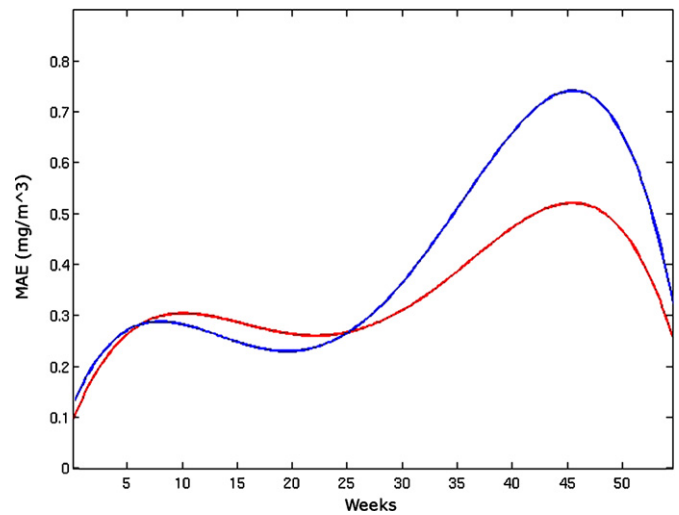| Hours	MRE	STD_RE	MAE (mg/m^3)	STD_MAE (mg/m^3)	SCC
12	0.27 ± 0.05	0.61	0.39 ± 0.07	0.39	0.87
24	0.27 ± 0.05	0.81	0.38 ± 0.05	0.39	0.85
48	0.39 ± 0.10	1.00	0.48 ± 0.10	0.37	0.83
72	0.37 ± 0.08	0.92	0.43 ± 0.06	0.37	0.88
96	0.37 ± 0.08	0.93	0.43 ± 0.07	0.38	0.87
168	0.27 + 0.08	0.74	0.37 ± 0.07	0.38	0.87
240	0.27 ± 0.06	0.72	0.34 ± 0.07	0.37	0.88
360	0.27 ± 0.05	0.74	0.37 ± 0.04	0.39	0.85
480	0.30 ± 0.07	0.84	0.41 ± 0.05	0.40	0.86
540	0.30 ± 0.06	0.83	0.39 ± 0.04	0.37	0.87
600	0.30 ± 0.05	0.80	0.37 ± 0.05	0.38	0.87
700	0.28 ± 0.05	0.81	0.38 ± 0.05	0.38	0.86
800	0.29 ± 0.05	0.82	0.36 ± 0.05	0.36	0.87
900	0.29 ± 0.06	0.81	0.36 ± 0.06	0.38	0.87
1200	0.29 ± 0.07	0.82	0.36 ± 0.07	0.38	0.87
1600	0.26 ± 0.09	0.76	0.36 ± 0.12	0.39	0.87
1800	0.26 ± 0.09	0.75	0.35 ± 0.11	0.38	0.87
2000	0.26 ± 0.09	0.74	0.38 ± 0.11	0.40	0.87
2200	0.27 ± 0.10	0.80	0.39 ± 0.12	0.44	0.86
2400	0.27 ± 0.09	0.80	0.40 ± 0.10	0.44	0.87

Fig. 5. Qualitative behaviour of the weekly mean absolute error (mg/m^3) in the CO concentration estimation problem. Blue line depicts MAE obtained by using all the sensors response for the feature vector composition while red line depicts MAE obtained by using only CO and NMHC sensor response. While there is a small advantage for the overall approach during the Summer time, at the start of the Winter time it accounts for a significantly higher error, this is very likely due to changes in the relative distribution of NO$_2$, NOx and CO concentrations. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

NO$_2$ and NOX sensor responses led to worse performance values. Fig. 3 depicts the differences between concentration estimation provided by the model and the true concentration as reported by the conventional analyzer, while Fig. 4 reports MAE results of this particular feature vector setup versus training set length choice.

Performance assessment was conducted by using ABR according to the procedure explained in the previous chapter with significant reduction of time needed by conventional procedures. As a control check, we report results of a conventional neural network hyperparameters optimization procedure for the CO, NMHC feature vector. Complexity control was achieved by controlling the target empirical error to be chosen in the [10e−3, 2 × 10e−3, 4 × 10e−3] set while the number of neurons varied in [5,10,15,20] set. Early stopping was not selected because it would have hampered results comparison by modifying the test set duration. For each empirical error and number of hidden neurons selection results had to be averaged through 20 different training runs. Best results were obtained for 10 neurons and 2 × 10e−3 target empirical error and were very similar (MRE = 0.27, MAE = 0.36) to the one obtained with

better than a CO targeted sensor response based one. This is probably due, again, to the high correlation factor existing between CO and Benzene concentrations in the selected scenario (r = 0.93, SCC = 0.86) and the very good performances of the NMHC targeted sensor. Best performance are achieved selecting a subset of the sensor array using both CO and NHMC sensor responses while inserting

Table 9

CO concentration estimation performances of the neural regression scheme computed over different training set lengths by using CO and NMHC solid state sensor response as feature vector (fixed test set length, crossvalidation setting). Five neurons have been used in the hidden layer, ABR has been used for complexity control. Confidence intervals are reported for MRE, MAE and SCC. No performance enhancement is obtained by using more than 10 days data recording.

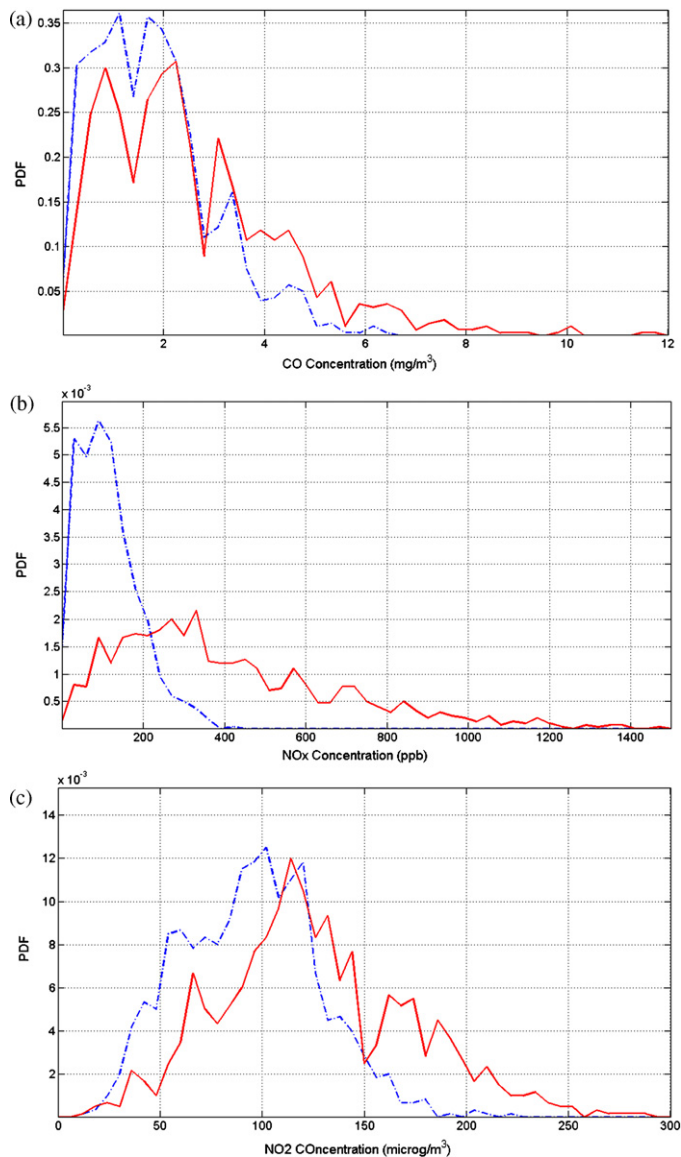| Hours	MRE	CI	STD_RE	MAE (mg/m^3)	CI (mg/m^3)	STD_MAE (mg/m^3)	SCC	CI
12	0.35	±0.01	0.67	0.53	±0.02	0.48	0.84	±0.01
24	0.31	±0.01	0.62	0.48	±0.02	0.46	0.85	±0.01
48	0.31	±0.01	0.62	0.47	±0.03	0.45	0.85	±0.01
72	0.30	±0.02	0.61	0.47	±0.05	0.46	0.85	±0.05
96	0.30	±0.02	0.62	0.45	±0.04	0.44	0.86	±0.02
168	0.29	±0.03	0.64	0.44	±0.06	0.42	0.87	±0.01
240	0.30	±0.04	0.64	0.45	±0.08	0.44	0.87	±0.01
360	0.28	±0.02	0.67	0.42	±0.05	0.43	0.87	±0.02
480	0.30	±0.04	0.66	0.44	±0.10	0.42	0.87	±0.02
540	0.29	±0.03	0.70	0.42	±0.05	0.41	0.87	±0.01
600	0.29	±0.04	0.67	0.43	±0.10	0.42	0.87	±0.02
700	0.27	±0.04	0.65	0.40	±0.12	0.41	0.87	±0.02
800	0.27	±0.06	0.63	0.42	±0.17	0.42	0.87	±0.03
900	0.27	±0.20	0.70	0.39	±0.21	0.41	0.87	±0.08
1200	0.30	±0.13	0.71	0.44	±0.81	0.43	0.87	±0.04

Fig. 6. Probability density function (PDF) non-parametric estimations in Summer time (blue, dash-dotted) and in winter time (red solid) for CO (a), NOx (b) and NO₂ (c) concentrations. Significant distribution changes are found for NOx and NO₂ leading to changes in concentration ratios with respect to CO. These modifications are identified as the main driver for performance degradation in winter time when NOx and NO₂ sensor responses are used for CO estimation. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

ABR using the same feature vector. The same procedure was applied to the entire array and led again to similar values (MRE = 0.32, MAE = 0.47) as referred to what obtained by ABR.

Summarizing, the major difference with respect to benzene concentration evaluation is due primarily to the worse performance of CO, NOₓ and NO₂ sensor in estimating their target gas. However, the presence of other sensors significantly contribute to mitigate this issue as a result of correlation effects.

These results warn against the unselective usage of all the sensor lines when training a sensor fusion subsystem in this scenario. Thus they also strongly suggest the use of different sensor fusion subsystem each one trained for the prediction of a single pollutant. In this case, feature selection algorithms could help to select the optimal composition of the array to be used for each pollutant concentration prediction problem. In Tables 8 and 9 we report the results of the performance evaluation versus training set duration for the CO by
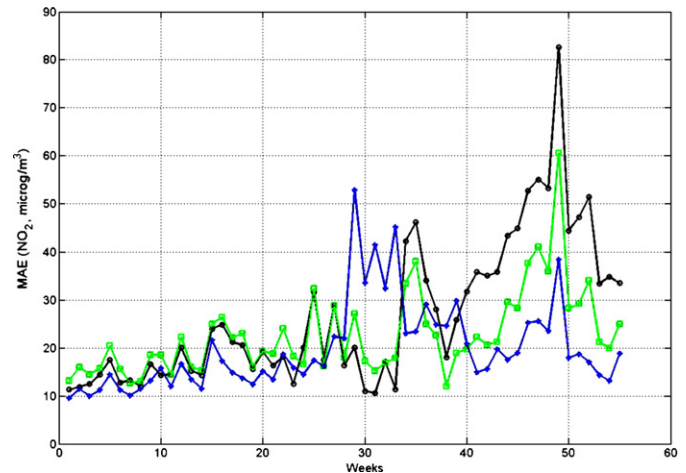

Fig. 7. Weekly mean absolute error in the NO₂ concentration estimation problem. Blue (dots) line depicts MAE obtained by using all the sensors response for the feature vector composition while green line (squares) depicts MAE obtained by using NO₂, NOx, O₃, T and RH sensors response. Black (circle) line depicts NO₂, T and RH sensors based estimation. All sensor approach retain a significant advantage over most of the test period and obtain best overall performance scores. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

using the best feature vector composition, using, respectively, the single test set and crossvalidated approach depicted in Section 4.1. Results confirm the optimality of a two weeks length choice regardless of feature vector composition while, of course, performance are generally better.

4.3. Performance estimation over time

Finally, starting from a 10 days training set length we have investigated the MAE performance index for CO in a week by week fashion over 55 weeks so to check their evolution over time. In addition, by using several feature vector composition, we investigated the influences of the use of different sensors responses on the concentration estimation during time.

As we found for benzene, short time performance indexes computed for this specie revealed to degrade in time, specifically, significant performance hits are detectable after 6 months from the calibration set end (see Fig. 5). Experimentally, we found that, with respect to the use of all available sensor responses, using only CO sensor response cause a small performance degradation in the Summer time and a significant performance boost in the Winter time. In particular, we believe that this is explainable by the definite changes in the concentration distribution of NOx and also NO₂ during Winter time. These changes disrupts relationships learnt by the regression system by looking at the Spring time training set. Instead, the combined use of CO and NMHC sensor responses have a positive effect during all test set. Summarizing, the descriptive power of NOx and NO₂ sensors and the correlation of their target gases with CO has less weight in the overall performance economy with respect to the misleading effects brought in by their targets distribution changes in the Winter time (see Fig. 6a–c). As regards as NO₂ week by week estimation error (see Figs. 7 and 8), we see, as above mentioned, that the all sensor based approach is able to obtain best performance by exploiting correlation positive effects, even here, however, performance worsening in Winter time is evident.

The combined findings reported in [8] and in the present paper lead us to consider as feasible the use of field data to obtain a suitable calibration for multisensor devices in city air pollution monitoring framework. These calibrations, however, suffer from
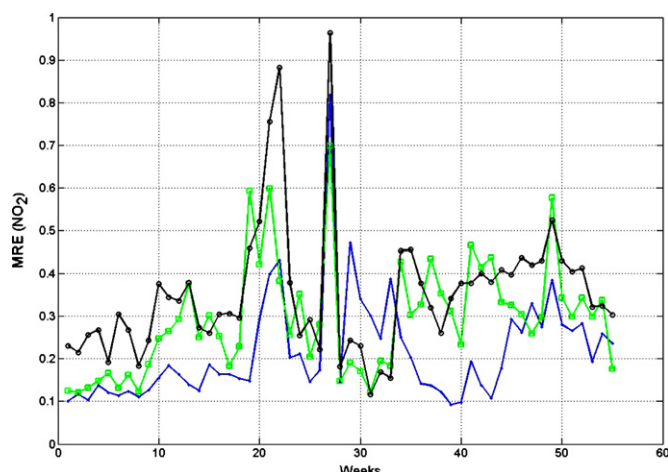
Fig. 8. Weekly mean relative error in the NO_2 concentration estimation problem. Blue (dots) line depicts MAE obtained by using all the sensors response for the feature vector composition while green line (squares) depicts MAE obtained by using NO_2, NOx, O_3, T and RH sensors response. Black (circle) line depicts NO_2, T and RH sensors based estimation. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

long term performance degradation and are sensitive to relative concentration distribution changes appearing on seasonal basis on as a results of particular events. This is far more evident when, due to poor specific sensor performances, the best statistical regressor has to rely on scenario-specific specie correlation factors. It remains still to prove if a site specific calibration can be used for concentration estimations at short distance sites. In this case, we believe that performance will be ultimately related to the site dependence of the specie correlation patterns that should be investigated.

5. Conclusions and further works

In this work we have further investigated the feasibility of the on-field calibration approach for a multisensor device in the air pollution monitoring scenario. The results obtained for typical city air pollutants, i.e. CO, NO_2, NOx and benzene have been reported and discussed overall results confirm that a relatively short training set duration, about two weeks long can cope with the cyclic behaviour of concentrations distribution leading to optimal performances and being sufficiently insensitive to outliers. This is very encouraging for the possibility of using multisensor devices, together with mobile conventional analyzers for the purpose of on-field calibration, for densening the sparse pollution monitoring network especially in the city historical centres.

Feature selection results definitely suggest the use of multiple regressors, each one being specifically designed and trained for a particular specie concentration estimation. In this way the designer is free to select the best feature set for the specific regression problem obtaining optimal results.

Feature selection analysis has furthermore confirmed the influences of scenario related species concentration correlations and specific sensors performances on the overall estimation scores. When specific sensors fails to correctly follow their target specie concentration, the existence of a strong correlation between this specie concentration and another one whose targeted sensor shows good performance may help the array to recover, ultimately leading to significantly better performance levels. However, in this case, changes in species relative statistical distribution, such as those that are due to Winter time house heating effect, may have a negative impact in the long term performances of a neural calibration. The same negative impact could probably be observed when using a calibration obtained for a specific site showing site specific mul-

tivariate concentration distribution and operating the multisensor device in a significantly distant site.

References

[1] 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.
[2] S. Vardoulakis, B.E.A. Fisher, K. Pericleous, N. Gonzalez-Flesca, Modelling air quality in street canyons: a review, Atmos. Environ. 37 (2003) 155–182.
[3] K.A. Kourtidis, I. Ziomas, C. Zerefos, E. Kosmidis, P. Symeonidis, E. Christophilopoulos, S. Karathanassis, A. Mploutsos, Benzene, toluene, ozone, NO_2 and SO_2 measurements in an urban street canyon in Thessaloniki, Greece, Atmos. Environ. 36 (2002) 5355–5364.
[4] B. Croxford, A. Penn, B. Hiller, Spatial distribution of urban pollution: civilizing urban traffic, in: 5th Symposium on Highway and Urban Pollution, Copenhagen, 1995.
[5] 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.
[6] 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.
[7] 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.
[8] S. De Vito, E. Massera, M. Piga, L. Martinotto, G. Di Francia, On field calibration of an electronic nose for benzene estimation in an urban pollution monitoring scenario, Sens. Actuators B Chem. 129 (2008) 750–757.
[9] W. Tsujita, A. Yoshino, H. Ishida, T. Moriizumi, Gas sensor network for air-pollution monitoring, Sens. Actuators B Chem. 110 (2005) 304–311.
[10] C.M. Bishop, Pattern Recognition and Machine Learning, Springer Science, 2006, ISBN 0-387-31073-8.
[11] 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.
[12] http://www.pirellilabs.com.
[13] M.T. Hagan, M.B. Menhaj, Training feedforward networks with the marquardt algorithm, IEEE Trans. Neural Networks 5 (1994) 989–993.
[14] F.D. Foresee, M.T. Hagan, Gauss–Newton approximation to Bayesian regularization, in: Proceedings of the 1997 International Joint Conference on Neural Networks, 1997, pp. 1930–1935.
[15] D.J.C. MacKay, Bayesian interpolation, Neural Comput. 4 (3) (1992) 415–447.
[16] N.F. Zhang, Calculation of the uncertainty of the mean of autocorrelated measurements, Metrologia 43 (2006) 276–281.

Biographies

Saverio De Vito received his MS in Informatics Engineering from University of Naples "Federico II" in 1998. During 1998 and 1999 he was a research fellow at Artificial Vision and intelligent systems laboratory of the above university working on breast cancer computer aided diagnosis. From 1999 to 2004 he was in ITC engineering as a R&D technical manager in the framework of satellite based telemedicine, earth observation and distance learning projects. In June 2004 he joined ENEA, as a researcher. His research interests include statistical pattern recognition, electronic noses, wireless sensor networks and computer aided diagnosis. Since 2005 he is contract professor of Applied Informatics at University of Cassino.

Marco Piga obtained his MSc in physics at Ferrara University (Italy) in 1999. Since 2004 researcher in Pirelli R&D department center in Milan. He is currently involved in Distributed Sensor Network project especially in environmental pollutants monitoring.

Luca Martinotto obtained his MSc in industrial chemistry at Milan University (Italy) in 1995. Since 1996 he was researcher in Pirelli R&D department center in Milan. He is responsible for Pirelli research activities in the field of DSN.

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