



OPEN

Statistical data pre-processing and time series incorporation for high-efficacy calibration of low-cost NO₂ sensor using machine learning

Slawomir Koziel^{1,2✉}, Anna Pietrenko-Dabrowska², Marek Wojcikowski² & Bogdan Pankiewicz²

Air pollution stands as a significant modern-day challenge impacting life quality, the environment, and the economy. It comprises various pollutants like gases, particulate matter, biological molecules, and more, stemming from sources such as vehicle emissions, industrial operations, agriculture, and natural events. Nitrogen dioxide (NO₂), among these harmful gases, is notably prevalent in densely populated urban regions. Given its adverse effects on health and the environment, accurate monitoring of NO₂ levels becomes imperative for devising effective risk mitigation strategies. However, the precise measurement of NO₂ poses challenges as it traditionally relies on costly and bulky equipment. This has prompted the development of more affordable alternatives, although their reliability is often questionable. The aim of this article is to introduce a groundbreaking method for precisely calibrating cost-effective NO₂ sensors. This technique involves statistical preprocessing of low-cost sensor readings, aligning their distribution with reference data. Central to this calibration is an artificial neural network (ANN) surrogate designed to predict sensor correction coefficients. It utilizes environmental variables (temperature, humidity, atmospheric pressure), cross-references auxiliary NO₂ sensors, and incorporates short time series of previous readings from the primary sensor. These methods are complemented by global data scaling. Demonstrated using a custom-designed cost-effective monitoring platform and high-precision public reference station data collected over 5 months, every component of our calibration framework proves crucial, contributing to its exceptional accuracy (with a correlation coefficient near 0.95 concerning the reference data and an RMSE below 2.4 µg/m³). This level of performance positions the calibrated sensor as a viable, cost-effective alternative to traditional monitoring approaches.

Keywords Air pollution, Low-cost sensor calibration, Machine learning, Data pre-processing, Neural networks

Nitrogen dioxide (NO₂) pollution is a significant environmental concern stemming from various sources such as vehicle emissions, industrial processes, and combustion. This gas is a part of nitrogen oxides (NO_x) and contributes to poor air quality, leading to respiratory issues and environmental damage. NO₂ reacts in the atmosphere to form harmful particles and ozone, impacting human health, ecosystems, and even contributing to climate change¹⁻⁶. Specifically, emissions of NO_x play a key role in creating photochemical smog, triggering acid rain, and causing ecological harm in water reservoirs⁷. Furthermore, high NO_x levels also elevate O₃, adversely affecting agriculture. Needless to say, monitoring and reducing NO₂ levels are critical for mitigating its adverse effects on both human health and the environment. Strict regulations have been implemented to control NO₂ levels, such as the CAFE Directive, setting an annual average below 40 µg/m³ and hourly concentrations not exceeding 200 µg/m³ for over 18 h per year⁸. The World Health Organization (WHO) has proposed even stricter limits⁹. However,

¹Engineering Optimization and Modeling Center, Reykjavik University, 102 Reykjavik, Iceland. ²Faculty of Electronics, Telecommunications and Informatics, Gdansk University of Technology, 80-233 Gdansk, Poland. ✉email: koziel@ru.is

about one-sixth of European monitoring stations indicate NO₂ levels that surpass these boundaries, especially in urban zones, particularly along transportation corridors. The economic toll of air pollution, including NO₂, amounts to significant costs^{2,10}.

Traditional methods for NO₂ monitoring rely on stationary and bulky equipment, demanding controlled environments and regular maintenance. Commonly used measurement approaches encompass photofragment chemiluminescence¹¹, long-range differential optical absorption spectroscopy¹², laser-induced fluorescence¹³, and cavity ring down spectroscopy¹⁴. Although these methods exhibit high sensitivity, some present limitations (e.g. unsuitability for localized monitoring¹²) or require intricate hardware (e.g., a vacuum system and a pulsed laser¹³). These deficiencies in traditional monitoring systems have driven the development of alternative methods that are cost-effective, easily deployable, and straightforward to maintain. In recent years, considerable research efforts have been directed towards development of portable platforms, which may be useful to enhance the spatial resolution of air quality monitoring. The latter is essential for urban areas with diverse pollutant distributions^{15–17}. Notwithstanding, low-cost sensors encounter reliability limitations^{18–20} due to instability²¹, fabrication inaccuracies^{22,23}, and cross-sensitivity to multiple gases^{24–26}. They are also sensitive to environmental conditions, especially temperature and humidity^{27,28}. In spite of these constraints, affordable sensors may complement sparsely positioned reference stations and serve as cost-efficient air quality monitoring solutions²⁹. They may also become foundations of integrated sensor networks^{30,31}, including those deployed on cars or aerial vehicles^{32,33}.

Enhancing the reliability of low-cost sensors has been a focal point in research, primarily focusing on refining calibration methods. These techniques are typically categorized into two types: laboratory-based and field-based³⁴. While laboratory procedures are more precise in theory, they often fall short in practice as the actual operating conditions of sensors seldom align with controlled laboratory settings^{18,19}. Consequently, field-based techniques are more prevalent, relying on reference data collected from public air monitoring stations. Numerical modelling for calibration typically involves either rudimentary regression techniques or more advanced machine learning approaches. In Ref.³⁵, methods such as multivariate linear regression (MLR), support vector regression (SVR), and random forest regression (RFR) were employed to calibrate electrochemical NO and NO₂ sensors based on temperature and humidity data. A study presented in Ref.³⁶ utilized ridge regression, random forest regression (RFR), Gaussian process regression (GPR), and MLR to correct low-cost NO₂ and PM₁₀ sensors based on temperature and humidity. In Ref.³⁷, calibration of a chemiluminescence NO-NO₂-NO_x analyser using MLR was showcased, also integrating temperature and humidity data. Further investigations into diverse regression models have been reported in Refs.^{38–40}.

In recent times, there has been a surge in interest in employing artificial intelligence methods, specifically neural networks (NNs) and diverse machine learning techniques, to achieve more dependable correction of low-cost sensors. For instance, Ref.²⁹ employed single linear regression (SLR), multivariate linear regression (MLR), random forest regression (RFR), and long short-term memory networks (LSTM) for calibrating CO, NO₂, O₃, and SO₂ sensors, noting LSTM's superior performance compared to regression procedures. Meanwhile, in Ref.¹⁵, convolutional neural networks (CNNs) and recurrent neural networks (RNNs) were used to calibrate CO and O₃ sensors using temperature and humidity data, showcasing advantages over linear regression (LR), SVR, or LSTM combined with CNN. Extensive literature, as observed in Refs.^{41–44}, showcases the application of various ANN surrogates, e.g. Bayesian NNs, shallow NNs, or dynamic NNs for low-cost sensor calibration.

In this research, we introduce an innovative method for precise calibration of affordable NO₂ sensors. The technique revolves around statistical preprocessing of low-cost sensor data to align its distribution with reference data before further refinement. Central to this approach is an artificial neural network (ANN) surrogate, tailored to predict sensor correction coefficients that encompass additive adjustment and multiplicative scaling. The surrogate model is trained using environmental variables (temperature, humidity, atmospheric pressure), data cross-referenced from auxiliary NO₂ sensors, and short time series of previous readings from the primary sensor. Global data scaling is also integrated as an additional calibration mechanism. To validate our calibration methodology, we applied it to a custom-designed autonomous monitoring platform equipped with NO₂ and environmental detectors, supported by electronic circuitry for monitoring implementation and data transfer protocols. Reference data was collected over five months from high-precision public stations in Gdansk, Poland. The results demonstrate exceptional calibration efficacy, achieving a correlation coefficient close to 0.95 with reference data and an extremely low RMSE below 2.4 µg/m³, even within a broad NO₂ measurement range (from zero to sixty µg/m³). Additional experiments conducted with different sets of surrogate model inputs and by excluding certain algorithmic tools highlight the vital role of each mechanism within the calibration framework, reaffirming their significance in enhancing correction quality.

Autonomous NO₂ monitoring platform

The article will showcase the sensor calibration methodology implemented on a custom-designed autonomous monitoring platform developed at Gdansk University of Technology, Poland. Section "[Hardware description](#)" details the hardware specifications, while Section "[Monitoring platform: output data](#)" delves into the data output from the platform's sensors.

Hardware description

The system is a comprehensive setup comprising multiple sensors for monitoring environmental factors such as temperature, humidity, and atmospheric pressure. It integrates a primary nitrogen dioxide sensor and two redundant sensors for cross-validation purposes. Furthermore, it includes a GSM modem for wirelessly transmitting measurement data to the cloud. Managing the air quality monitoring protocols are off-the-shelf components

coordinated by the BeagleBone® Blue microprocessor system⁴⁵, which houses a 1 GHz ARM® Cortex-A8 processor, 512 MB DDR3 RAM, and 4 GB eMMC memory, operating on the Linux OS.

The system relies on a rechargeable 7.4 V/4400 mA battery capable of sustaining operations for at least twenty hours without external power sources. The block diagram of the platform, featuring sensor details, is illustrated in Fig. 1. Data transmission occurs via the GSM modem, making the measurement data available online. The system is mounted on a polyethylene terephthalate base plate, as depicted in Fig. 2. The gas sensors (ST, SGX, MICS) are closely positioned (see Fig. 2a) along with environmental detectors monitoring their operational conditions. An auxiliary environmental sensor is placed at the device's edge.

The employment of auxiliary sensors serves to address variations between external and internal temperatures and humidity, primarily influenced by heat generated by the electronic circuitry. An Intel USB Stick module is also installed for potential on-board execution of calibration procedures. The platform is accommodated in a weatherproof enclosure, cf. Fig. 2c.

Monitoring platform: output data

The monitoring platform, detailed in Section "Hardware description", gathers NO₂ measurements from the primary sensor and two redundant sensors, along with environmental sensor data (internal and external temperature, humidity, and atmospheric pressure). Figure 3a visually represents these outputs, while Fig. 3b introduces the notation used in this study. It is crucial to note that this platform captures environmental parameters both within the system (close to the NO₂ sensors) and externally (at the edge of the platform). The variations in internal and external temperature and humidity stem from the heat produced by the electronic circuitry. Given the influence of these parameters on sensor performance, incorporating both sets of temperature and humidity data can significantly enhance the reliability of the calibration process. Additionally, although the accuracy of

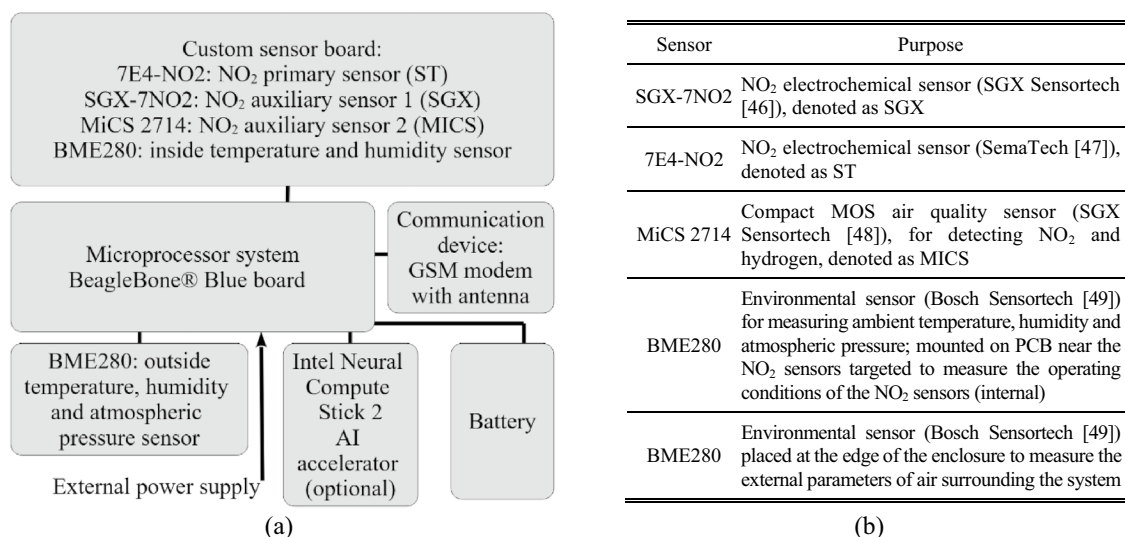


Figure 1. Autonomous air monitoring platform designed at Gdansk University of Technology, Poland: (a) block diagram, (b) included sensors^{46–49}.

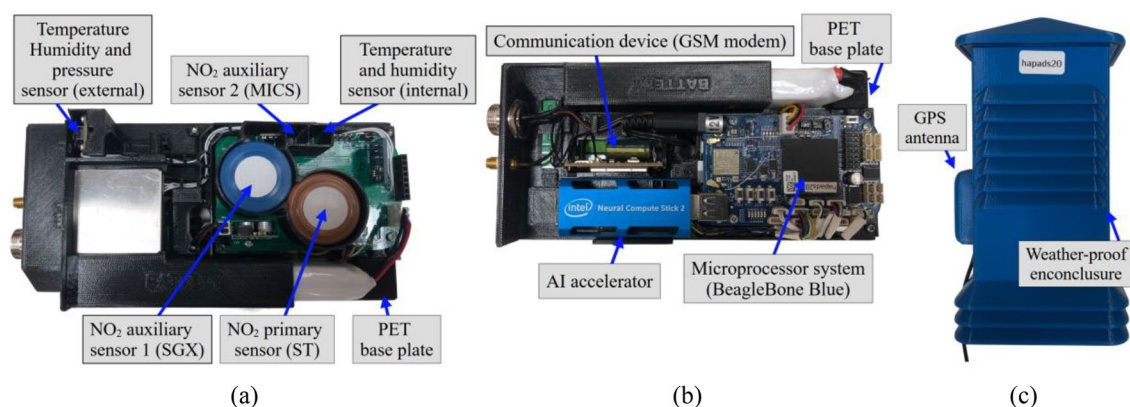
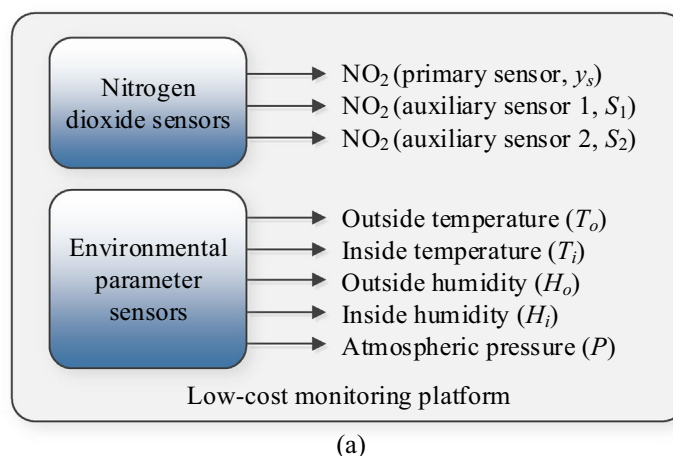


Figure 2. Autonomous monitoring platform designed at Gdansk University of Technology, Poland: (a) internals (top view), (b) internals (bottom view), (c) systems mounted in weather-proof enclosure.



Data type	Symbol [#]	
Nitrogen dioxide (NO ₂) sensors	Primary sensor: ST	$y_s^{(j)}, j = 1, \dots, N$
	Auxiliary sensor 1: SGX	$S_1^{(j)}, j = 1, \dots, N$
	Auxiliary sensor 2: MICS	$S_2^{(j)}, j = 1, \dots, N$
Environmental parameters sensors	Outside temperature	$T_o^{(j)}, j = 1, \dots, N$
	Inside temperature	$T_i^{(j)}, j = 1, \dots, N$
	Outside humidity	$H_o^{(j)}, j = 1, \dots, N$
	Inside humidity	$H_i^{(j)}, j = 1, \dots, N$
	Atmospheric pressure	$P^{(j)}, j = 1, \dots, N$

[#] N is the total number of samples from all measurement locations.

(b)

Figure 3. Outputs of the low-cost monitoring platform of Section "Hardware description": (a) NO₂ reading from the low-cost sensor under calibration (y_s). The sensor also produces auxiliary outputs: auxiliary NO₂ readings (S_1 and S_2), outside and inside temperature (T_o and T_i , respectively), outside and inside humidity (H_o and H_i , respectively), and atmospheric pressure (P); (b) symbols of data produced by the platform's sensors. The number N stands for the total number of data samples obtained from the platforms, further divided into training and testing sets (cf. Section "Precise sensor calibration using statistical pre-processing, ANN surrogates, and global data scaling").

the auxiliary NO₂ sensors within the platform is limited, their readings offer indirect yet valuable insights into the factors affecting the primary sensor, notably its cross-sensitivity to other gases.

Reference data. Public monitoring stations

The calibration process for the low-cost sensor will utilize reference data obtained from high-precision public monitoring stations strategically located in Gdansk, Poland, operated by the ARMAG Foundation⁵⁰. The geographical distribution of these stations is illustrated in Fig. 4a. The stations are housed within air-conditioned containers and are equipped with high-performance air monitoring instruments, detailed in Fig. 4b. The specific sensors used for NO-NO₂-NO_x measurements are listed in Fig. 4c. ARMAG provides open access to the generated data on their website (<https://armaag.gda.pl/en/>). Measurements are carried out hourly and are accessible on the foundation's website for a duration of three days. To enable extended data collection periods, a custom script has been prepared, which allows automated download of this information into a text file hosted on a dedicated server.

Precise sensor calibration using statistical pre-processing, ANN surrogates, and global data scaling

This section delineates the comprehensive methodology devised for the calibration of low-cost NO₂ sensors. The task of correcting the sensor is formulated in Section "Sensor calibration. Problem statement". Further details regarding the affine correction scheme are provided in Section "Additive and multiplicative low-cost sensor correction". Section "Statistical pre-processing of low-cost sensor measurements" delves into the statistical pre-processing of data, designed to enhance the initial alignment between the outputs of the reference and low-cost sensors. An in-depth exploration of the primary calibration model, an artificial neural network (ANN) surrogate, is presented in Section "Sensor calibration using neural network surrogate". The various configurations of inputs to the ANN model are elucidated in Section "Calibration model inputs". These encompass fundamental environmental parameters and redundant NO₂ sensor readings (Section "Calibration input configuration I: basic

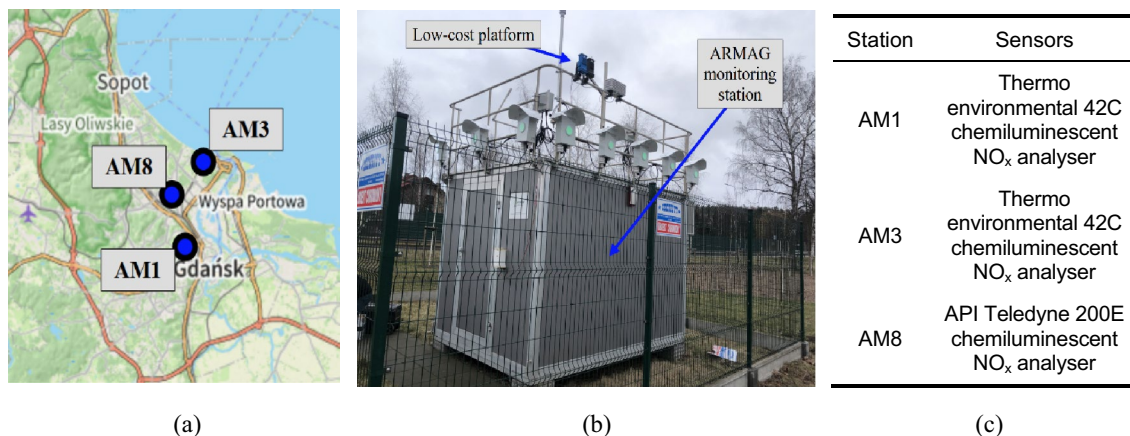


Figure 4. Reference monitoring stations of the ARMAG foundation used to acquire reference data: (a) station locations in the city of Gdansk, (b) photograph of the selected station with the proposed low-cost platform mounted in the vicinity, (c) NO_x sensors installed on the stations.

setup"), expanded sets incorporating differentials (Section "Calibration input configuration II: differentials"), and time-series-based inputs comprising prior NO₂ measurements from the primary sensor (Section "Calibration input configuration III: time series of prior NO₂ measurements"). Additionally, Section "Global data scaling" discusses an auxiliary calibration mechanism, specifically global data scaling. The comprehensive workflow for NO₂ monitoring utilizing the calibrated low-cost sensor is elucidated in Section "Operating flow of NO₂ monitoring by means of calibrated sensor".

Sensor calibration. Problem statement

Sensor calibration is based on two datasets. The first one comprises NO₂ readings obtained from the reference stations, as outlined in Section "Reference data. Public monitoring stations". The respective samples will be denoted as $y_r^{(j)}$, $j = 1, \dots, N$, where N is the total number of points. The datasets obtained from the autonomous platform described in Section "Autonomous NO₂ monitoring platform", i.e., $\{y_s^{(j)}\}$ and the respective environmental parameter vectors $\{z_s^{(j)}\}$ (cf. Fig. 3) is in correspondence with $\{y_r^{(j)}\}$, i.e., the respective outputs are collected at the same time intervals. Figure 5 elucidates the division of this data into training and testing sets. The testing set consists of several two-week sequences gathered at different time intervals during the five-month measurement campaign, as elaborated in Section "Results and discussion".

Sensor calibration is realized using the training datasets $\{y_{r0}^{(j)}\}$, $\{y_{s0}^{(j)}\}$, and $\{z_{s0}^{(j)}\}$, $j = 1, \dots, N_0$ (cf. Fig. 5). The correction coefficients are jointly denoted as $C(y_s, z_s; \mathbf{p})$, cf. Fig. 6, where \mathbf{p} stands for the combined calibration model hyper-parameters. The corrected sensor's output is denoted as $y_c = F_{CAL}(y_s, C(y_s, z_s; \mathbf{p}))$. Based on this terminology, the calibration problem is posed as a nonlinear minimization task.

$$\mathbf{p}^* = \arg \min_{\mathbf{p}} \sqrt{\sum_{j=1}^{N_0} \left(y_{r0}^{(j)} - F_{CAL} \left(y_{s0}^{(j)}, C \left(y_{s0}^{(j)}, z_{s0}^{(j)}; \mathbf{p} \right) \right) \right)^2} \quad (1)$$

<p>Data division:</p> <ol style="list-style-type: none"> The overall data (N samples) is divided into the training set (N_0 samples), and the testing set (N_t samples); N_t is set to approximately ten percent of the total number of samples N. <p>Training set:</p> <ol style="list-style-type: none"> Reference: $\{y_{r0}^{(j)}\}$, $j = 1, \dots, N_0$; Low-cost sensor: <ul style="list-style-type: none"> Sensor's NO₂ measurements: $\{y_{s0}^{(j)}\}$, $j = 1, \dots, N_0$; Auxiliary data: $\{z_{s0}^{(j)}\}$, $j = 1, \dots, N_0$, where $\mathbf{z}_{s0} = [T_{r0}^{(j)} \ T_{i0}^{(j)} \ H_{a0}^{(j)} \ H_{i0}^{(j)} \ P_0^{(j)} \ S_{10}^{(j)} \ S_{20}^{(j)}]^T$; <p>Testing set:</p> <ol style="list-style-type: none"> Reference: $\{y_{rt}^{(j)}\}$, $j = 1, \dots, N_t$; Low-cost sensor: <ul style="list-style-type: none"> Sensor's NO₂ measurements: $\{y_{st}^{(j)}\}$, $j = 1, \dots, N_t$; Auxiliary data: $\{z_{st}^{(j)}\}$, $j = 1, \dots, N_t$, where $\mathbf{z}_{st} = [T_{ot}^{(j)} \ T_{it}^{(j)} \ H_{at}^{(j)} \ H_{it}^{(j)} \ P_t^{(j)} \ S_{1t}^{(j)} \ S_{2t}^{(j)}]^T$.

Figure 5. Division of the reference and low-cost sensor data into training and testing set.

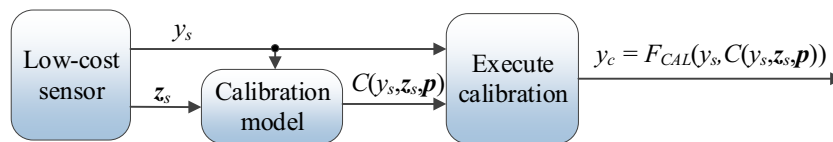


Figure 6. Overall flow of the low-cost sensor calibration. Auxiliary data and sensor output y_s are used to obtain the correction coefficients $C(y_s, z_s, \mathbf{p})$, which are then used to compute the corrected sensor output y_c , see Sections "Additive and multiplicative low-cost sensor correction" through Sect. "Global data scaling" for details. A more detailed procedure will be discussed in Section "Global data scaling".

The aim of (1) is to optimize the hyper-parameters of the calibration model to maximize the (L -square) alignment between the NO_2 readings from the reference and corrected low-cost sensors across the training set.

Additive and multiplicative low-cost sensor correction

Conventional correction methods often model the disparities between reference and low-cost sensor readings directly. In this study, we adopt an affine scaling approach that involves both additive and multiplicative correction. This method introduces additional degrees of freedom, enhancing the reliability of the calibration process. In our case, it is recommended to use a multiplicative scaling factor greater than one, as the typical amplitude variations in reference data are higher than those in low-cost sensor measurements, cf. Fig. 7. Details of this correction process are outlined in Fig. 8. It is essential to note that for $A^{(j)}$ to be greater than unity, the hyper-parameter α must be less than unity (cf. (8)). In practice, α can be optimized simultaneously with training the NN calibration model (see Section "Statistical pre-processing of low-cost sensor measurements"). Through preliminary experiments, a suitable value for α found to be 0.8 will be utilized in our validation studies discussed in Section "Results and discussion".

As indicated in Fig. 8, the ANN model is identified based on the training data in the form of the coefficients A and D computed for each training sample. In other words, the coefficients $A^{(j)}$ and $D^{(j)}$ are computed for each pair of the raw sensor data $y_{s,0}^{(j)}$ and $y_{r,0}^{(j)}$ so that perfect matching is ensured as shown in (5). Subsequently, the calibration ANN model is trained to render the values of A and D for any combination of auxiliary parameters z_s and primary sensor reading y_s . The information about the reference reading at this combination is encoded in the training pairs $A^{(j)}$, $D^{(j)}$ combined with their corresponding sensor output $y_{s,0}^{(j)}$.

Statistical pre-processing of low-cost sensor measurements

One of the keystones of the proposed calibration procedure is statistical pre-processing of the low-cost sensor readings. A potential usefulness of this procedure stems from the observations made in Section "Additive and multiplicative low-cost sensor correction", specifically, the observed discrepancies between typical measured NO_2 levels between the reference station and the low-cost sensor, as illustrated in Fig. 7. These discrepancies are well-represented on the histogram plots shown in Fig. 9. The statistical distribution of the measurements for the low-cost sensor is shifted towards lower values, which indicates that the typical readings are lower than for the reference.

The proposed pre-processing procedure aims at reducing the aforementioned misalignment by initial scaling of the low-cost sensor readings using a nonlinear transformation of the form

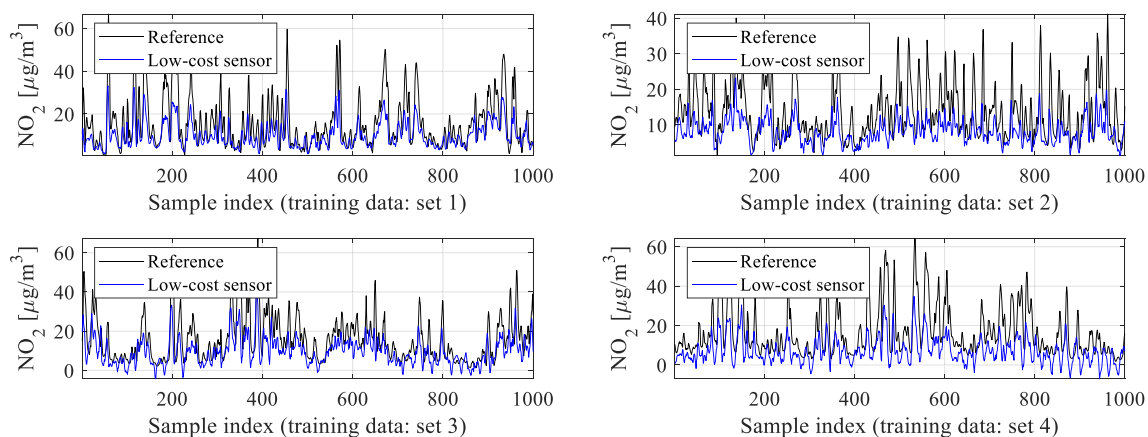


Figure 7. Selected reference and low-cost sensor training data subsets. A typical amplitude of low-cost sensor data variations is lower than for the reference, therefore, multiplicative scaling with coefficient $A > 1$ may be advantageous in improving the calibration process quality.

Basic sensor correction scheme:

$$y_c = A(y_s + D) \tag{2}$$

where y_c is the corrected sensor reading, whereas both correction coefficients are produced by the calibration function $C(y_s, \mathbf{z}_s, \mathbf{p})$.

Given (2), the calibration function takes the form of

$$C(y_s, \mathbf{z}_s, \mathbf{p}) = [A(y_s, \mathbf{z}_s, \mathbf{p}) + D(y_s, \mathbf{z}_s, \mathbf{p})]^T \tag{3}$$

and

$$y_c = F_{CAL}(y_s, C(y_s, \mathbf{z}_s, \mathbf{p})) = A(y_s, \mathbf{z}_s, \mathbf{p})[y_s + D(y_s, \mathbf{z}_s, \mathbf{p})] \tag{4}$$

Calibration coefficient identification: parameters A and D are computed for each training sample, so that we have

$$y_{r0}^{(j)} = A^{(j)}(y_{s0}^{(j)} + D^{(j)}) \tag{5}$$

for $j = 1, \dots, N_0$, and the problem (1) is replaced by

$$\mathbf{p}^* = \arg \min_p \|\mathbf{A}_0 \mathbf{D}_0 - \mathbf{C}_0(\mathbf{p})\|^2 \tag{6}$$

where

$$\mathbf{A}_0 = \begin{bmatrix} A^{(1)} \\ \vdots \\ A^{(N_0)} \end{bmatrix}, \quad \mathbf{D}_0 = \begin{bmatrix} D^{(1)} \\ \vdots \\ D^{(N_0)} \end{bmatrix}, \quad \mathbf{C}_0(\mathbf{p}) = \begin{bmatrix} C(y_{s0}^{(1)}, \mathbf{z}_{s0}^{(1)}, \mathbf{p})^T \\ \vdots \\ C(y_{s0}^{(N_0)}, \mathbf{z}_{s0}^{(N_0)}, \mathbf{p})^T \end{bmatrix} = \begin{bmatrix} A(y_{s0}^{(1)}, \mathbf{z}_{s0}^{(1)}, \mathbf{p}) & D(y_{s0}^{(1)}, \mathbf{z}_{s0}^{(1)}, \mathbf{p}) \\ \vdots & \vdots \\ D(y_{s0}^{(N_0)}, \mathbf{z}_{s0}^{(N_0)}, \mathbf{p})^T & D(y_{s0}^{(N_0)}, \mathbf{z}_{s0}^{(N_0)}, \mathbf{p})^T \end{bmatrix} \tag{7}$$

and

Uniqueness problem: Solution to (5) is not unique. Here, a regularization is implemented by introducing a calibration hyper-parameter $0 \leq \alpha \leq 1$, which 'distributes' the weights between additive and multiplicative correction. We have

$$D^{(j)} = \alpha [y_{r0}^{(j)} - y_{s0}^{(j)}] \quad \text{and} \quad A^{(j)} = \frac{y_{r0}^{(j)}}{y_{s0}^{(j)} + \alpha(y_{r0}^{(j)} - y_{s0}^{(j)})} \tag{8}$$

Figure 8. Fundamental output correction of the low-cost NO₂ sensor: affine scaling.

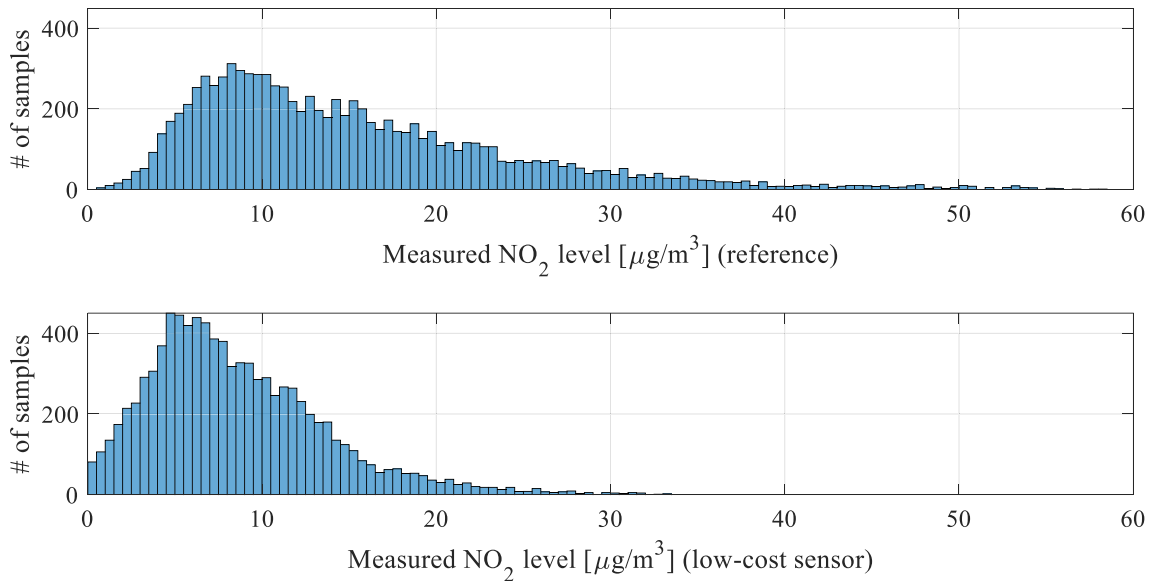


Figure 9. Histograms of the reference NO₂ readings (top) and raw (uncorrected) low-cost sensor NO₂ measurements (bottom), obtained for the complete training datasets. Note that the statistical distribution for the low-cost sensor is shifted towards lower values, which indicates that the typical readings are lower than for the reference, as also observed in Fig. 7.

$$P(y_s, \mathbf{s}) = P(y_s, [s_1 \ s_2 \ s_3]^T) = s_1 + s_2 y_s + s_3 y_s^2 \tag{9}$$

which is to be applied to all sensor measurements simultaneously. The second order polynomial has been chosen as the simplest nonlinear function that can be utilized to match the probability distributions represented by the histograms. The idea is as follows. Assuming that the probability distributions are generally similar, using affine

transformation (shift + linear scaling) is generally sufficient because it allows for matching the distribution means and standard deviations. The second order has been added in order to introduce a slight nonlinearity, thereby improving the quality of histogram matching. We will also use a vector notation for P , i.e.,

$$P(\mathbf{y}, \mathbf{s}) = P\left([y_1 \dots y_N]^T, [s_1 \ s_2 \ s_3]^T\right) = \begin{bmatrix} s_1 + s_2 y_1 + s_3 y_1^2 \\ \vdots \\ s_1 + s_2 y_N + s_3 y_N^2 \end{bmatrix} \tag{10}$$

The coefficient vector \mathbf{s} is determined to improve the alignment of the smoothed histograms shown in Fig. 10. The latter is defined as

$$H(\mathbf{y}) = [\mathbf{z} \ S(\mathbf{N}_y)] \tag{11}$$

where

$$\mathbf{z} = [z_1 \ z_2 \ \dots \ z_M]^T \tag{12}$$

is a vector of histogram bins (i.e., intervals splitting the horizontal axis in Fig. 9 into respective compartments), whereas

$$\mathbf{N}_y = [n_{y,1} \ n_{y,2} \ \dots \ n_{y,M}]^T \tag{13}$$

denotes the vector of the number of (training data) readings that fall within the respective intervals. The function $S(\cdot)$ represents a smoothing procedure.

Having defined the smoothed histogram, the pre-processing is accomplished by solving

$$\mathbf{s}^* = \arg \min_{\mathbf{s}} \|H(\mathbf{y}_r) - H(P(\mathbf{y}_s, \mathbf{s}))\| \tag{14}$$

where \mathbf{y}_r and \mathbf{y}_s stand for the aggregated reference and low-cost sensor NO_2 readings.

Note that if the histogram bins \mathbf{z} are identical for the reference and the sensor (which is assumed here), the functional in (14) boils down to comparing the respective $S(\mathbf{N}_y)$ vectors. Solving problem (14) is equivalent to matching the smoothed histograms of the reference and pre-processed low-cost sensor histograms. The unknown variables in this process are the scaling polynomial coefficients, that is, the vector \mathbf{s} defined in Eq. (9). Note that the matching is not performed for the number of observations falling into the reference bins as these are discrete numbers, and solving least-square regression problem would be problematic when using gradient-based routines. Instead, matching is performed upon smoothed histograms, which are continuous functions of the bin indices. The process (14) is effectively fitting the second-order polynomial that determines the histogram scaling.

Figure 10 shows the smoothed histograms before (top) and after pre-processing (bottom), indicating considerable improvement in terms of the alignment. Direct comparison between raw (non-smoothed) histograms can be found in Fig. 11. Figure 12 shows the effects of pre-processing for selected subsets of the training data. As

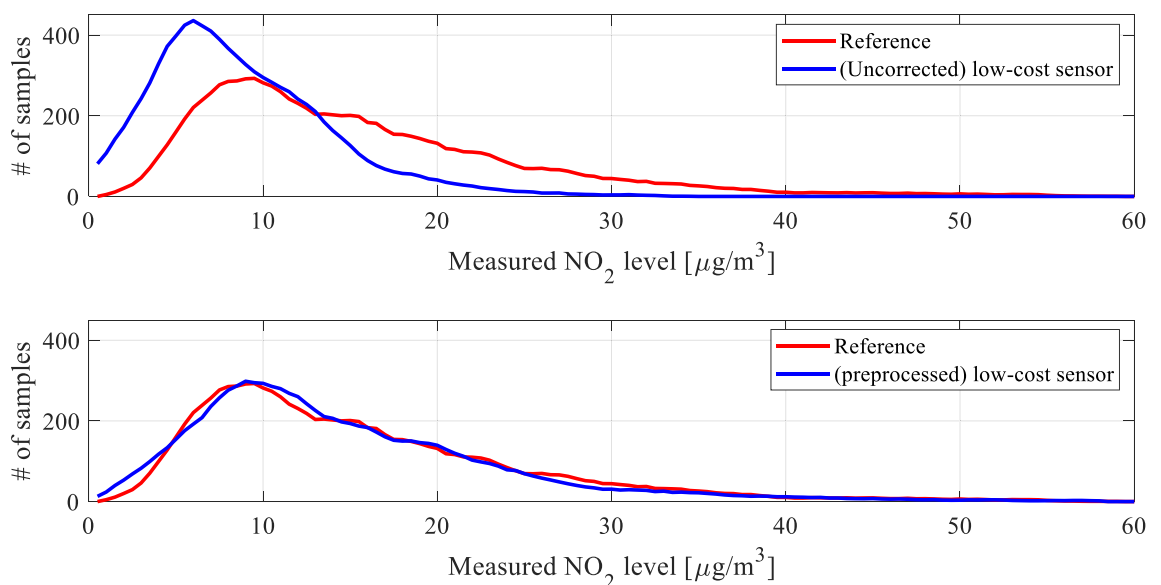


Figure 10. Smoothened histograms of the reference versus raw low-cost sensor (top) and the reference versus pre-processed low-cost sensor (bottom). As it can be observed, pre-processing aligns the measurement distributions of the low-cost sensor, thereby making it better prepared for further calibration.

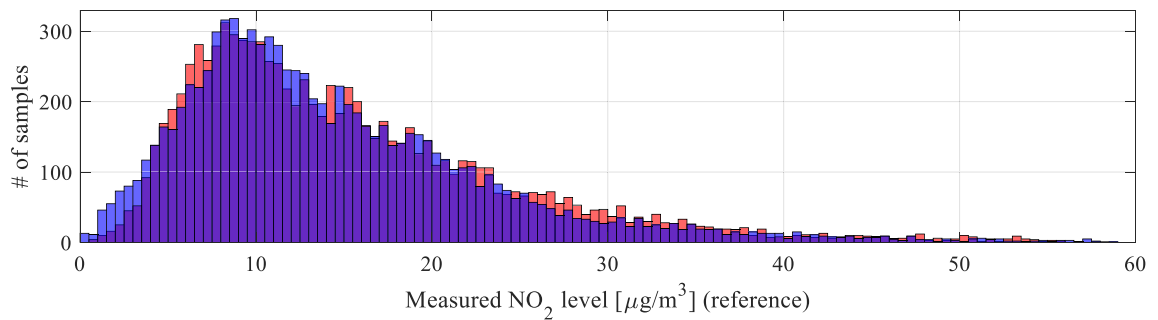


Figure 11. A comparison between the reference data (red) and pre-processed (blue) low-cost sensor histogram. Good alignment between the two datasets can be observed. Overlapping data marked purple.

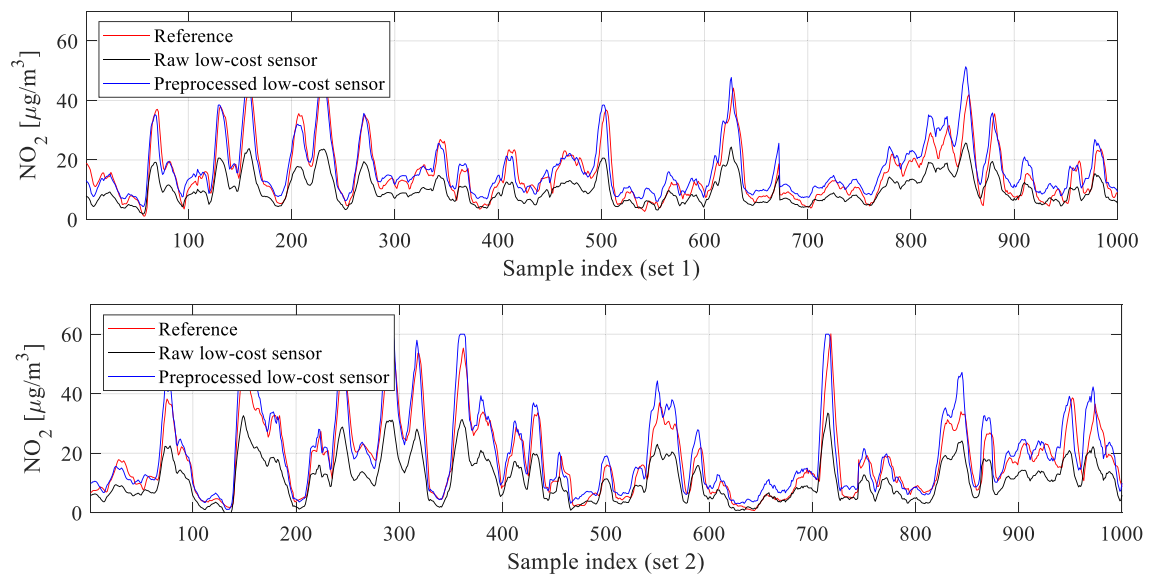


Figure 12. The effects of statistical pre-processing illustrated for two selected subsets of the training data. As it can be observed, pre-processing leads to a significant improvement of correlation between the reference and low-cost sensor readings.

mentioned earlier, pre-processing will be employed as the first calibration step, followed by surrogate-predicted correction to be discussed from Section "[Sensor calibration using neural network surrogate](#)" on.

Sensor calibration using neural network surrogate

The primary calibration model employed in this study is an artificial neural network (ANN) surrogate. Specifically, we have opted for a multi-layer perceptron (MLP) architecture^{51,52} featuring three fully connected hidden layers, each consisting of twenty neurons utilizing a sigmoid activation function, as illustrated in Fig. 13. The model's hyper-parameters are identified using a backpropagation Levenberg–Marquardt algorithm⁵³ (setup: 1000 learning epochs, performance evaluation using mean-square error (MSE), randomized training/testing data division). It should be emphasized that the aforementioned data division is pertinent to the training data itself (i.e., the training data is internally split into 'training' and 'validation' data for the purpose of ANN training in each epoch). The testing data as specified in Fig. 5 is kept separate and only used for model validation in the numerical experiments in Section "[Results and discussion](#)".

We deliberately chose a relatively simple ANN architecture to expedite the training process and prioritize its role as a regression model. Given the ample training samples available, the model's sensitivity to the number of layers and neurons is limited. Furthermore, this streamlined architecture effectively mitigates inherent noise present in both the reference and sensor readings.

The calibration model takes inputs comprising environmental factors (internal/external temperature, humidity, etc.) and NO₂ measurements from both the primary and auxiliary sensors. The outputs of the neural network (NN) model are the affine scaling coefficients A and D . In Section "[Calibration model inputs](#)", we delve into diverse extended input sets aimed at bolstering the calibration process's reliability. The effects of these expanded sets, alongside the consequences of restricting inputs to various subsets of the vector \mathbf{z}_s , will be analysed in Section "[Results and discussion](#)" to assess how input configuration impacts the efficacy of calibration.

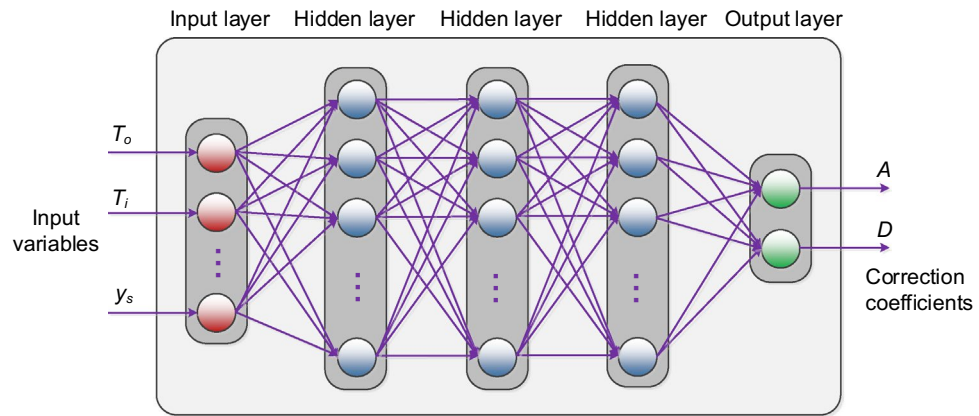


Figure 13. ANN surrogate used as the core calibration model. Here, we employ a multi-layer perceptron (MLP) with three fully-connected hidden layers. When statistical data pre-processing is utilized (cf. Section "Statistical pre-processing of low-cost sensor measurements"), then the input y_s of the primary sensor reading is not taken directly from the sensor. Instead, it is a pre-processed value.

Calibration model inputs

In this section, we discuss various input configurations of the ANN calibration model. Section "Calibration input configuration I: basic setup" recalls the basic parameter set discussed earlier. The extended input set, integrating differentials of environmental variables and primary NO₂ readings, is explored in Section "Calibration input configuration II: differentials".

Section "Calibration input configuration III: time series of prior NO₂ measurements" analyses the final setup that involves time series of prior NO₂ measurements from the low-cost sensor. In our investigations, we focus on potential benefits of particular setups in terms of improving the calibration process dependability.

Calibration input configuration I: basic setup

The fundamental configuration of the calibration model inputs includes the auxiliary data vector $z_s = [T_o, T_i, H_o, H_i, P, S_1, S_2]^T$. This set of values comprises external/internal temperature, humidity, atmospheric pressure, and NO₂ data from redundant sensors. These elements are augmented by the primary sensor's NO₂ measurements, y_s . Section "Results and discussion" will further investigate constrained variations of this arrangement to determine the individual elements' significance.

Calibration input configuration II: differentials

The basic input arrangement elucidated in Section "Calibration input configuration I: basic setup" can be extended by incorporating additional parameters representing local (temporal) fluctuations in environmental variables and NO₂ readings. More specifically, we define differentials

$$\Delta y_s^{(j)} = \frac{y_s^{(j)} - y_s^{(j)}(-\Delta t)}{\Delta t} \quad (15)$$

where Δt is the time interval between subsequent sensor readings; $y_s^{(j)}(-\Delta t)$ stands for the last measurement taken before $y_s^{(j)}$. Differentials of the environmental parameters are defined in a similar manner

$$\Delta T_o^{(j)} = \frac{T_o^{(j)} - T_o^{(j)}(-\Delta t)}{\Delta t}, \quad \Delta T_i^{(j)} = \frac{T_i^{(j)} - T_i^{(j)}(-\Delta t)}{\Delta t} \quad (16)$$

$$\Delta H_o^{(j)} = \frac{H_o^{(j)} - H_o^{(j)}(-\Delta t)}{\Delta t}, \quad \Delta H_i^{(j)} = \frac{H_i^{(j)} - H_i^{(j)}(-\Delta t)}{\Delta t} \quad (17)$$

$$\Delta P^{(j)} = \frac{P^{(j)} - P^{(j)}(-\Delta t)}{\Delta t} \quad (18)$$

Note that computing (15), (16), (17), (18) only requires storing one extra set of readings. The differentials, especially $\Delta y_s^{(j)}$, quantify local fluctuations in NO₂ level, which facilitates prediction of forthcoming alterations. Moreover, integrating differentials of environmental variables can provide explicit or implicit insights into the dynamics of relevant factors such as cross-sensitivity to other gases. This addition of differentials as supplementary inputs into the NN surrogate allows exploration of their potential contribution to enhancing the calibration quality.

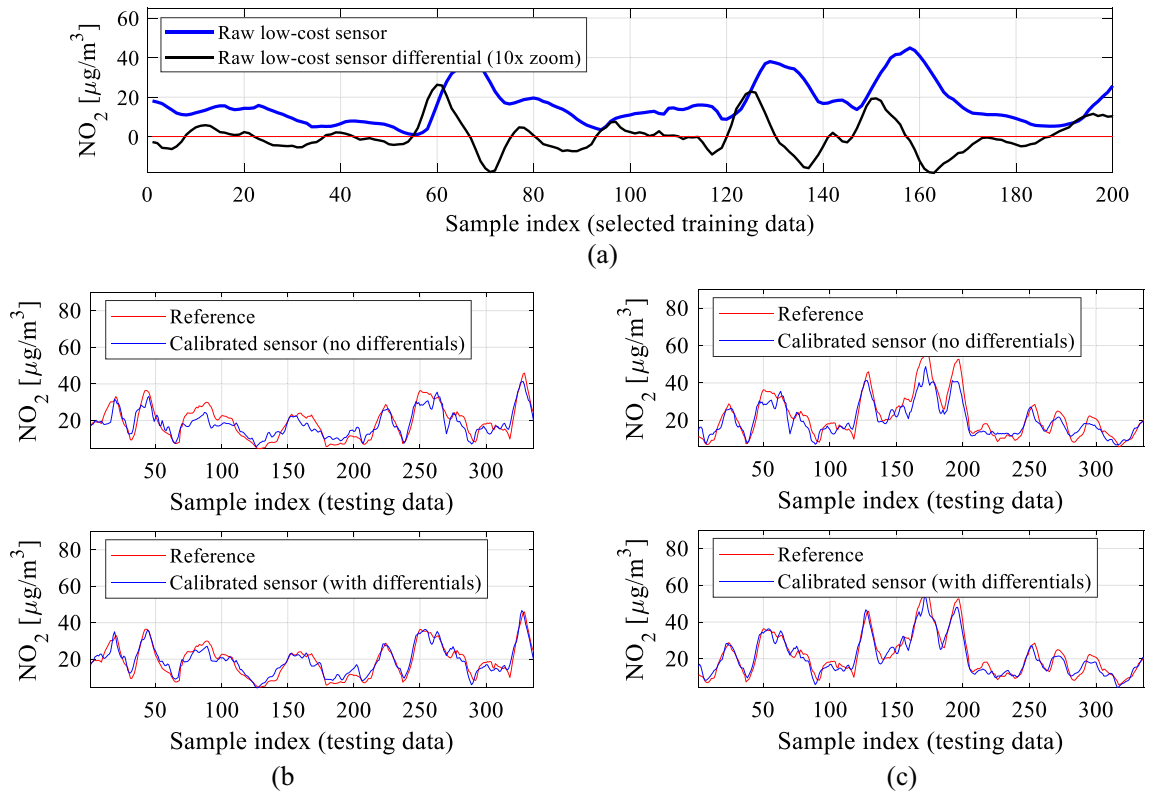


Figure 14. Differentials used as additional ANN surrogate inputs to enhance calibration dependability: (a) selected training data sequence (NO_2 readings from the low-cost sensor) and its corresponding differentials (15); (b) the effects of incorporating differentials shown for a selected sequence of testing data; (c) the effects of differentials shown for another testing data sequence. Note that including differentials (here, of all environmental variables and the primary NO_2 readings from the low-cost sensor) noticeably improves data alignment.

A visual illustration has been provided in Fig. 14. In particular, Fig. 14a shows—for a selected sequence of the training data—the NO_2 readings from the low-cost sensor alongside the respective differentials. Meanwhile, Fig. 14b and c, demonstrate the effects of incorporating the differentials as auxiliary calibration model inputs. The flow diagram of the modified calibration process involving differentials can be found in Fig. 15.

Calibration input configuration III: time series of prior NO_2 measurements

Expanding the concept of differentials might involve integrating an extended series of previous sensor measurements, which may not be suitable for mobile monitoring platforms but could significantly enhance the calibration of stationary systems, like the one discussed in Section "Autonomous NO_2 monitoring platform". The additional inputs for the calibration surrogate comprise

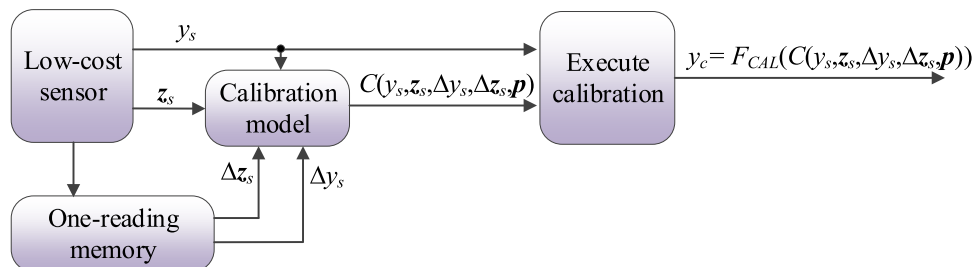


Figure 15. Calibration of the low-cost sensor with differentials used as additional calibration model inputs. Auxiliary data and sensor output y_s are used to obtain the correction coefficients $C(y_s, z_s, \Delta y_s, \Delta z_s, p)$, used to compute the corrected sensor output y_c . The pre-processing step is not shown for clarity.

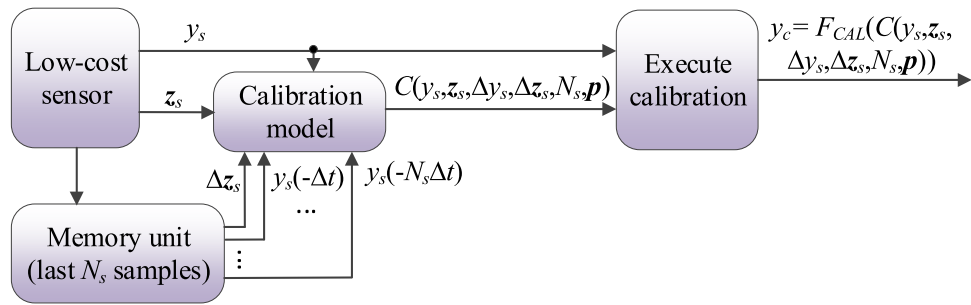


Figure 16. Calibration of the low-cost sensor with time series of prior measurements used as additional calibration model inputs. Auxiliary data are used to obtain the correction coefficients $C(y_s, z_s, \Delta y_s, \Delta z_s, N_s, p)$, used to compute the corrected sensor output y_c . The pre-processing step is not shown for clarity.

$$y_s^{(j)}(-s\Delta t), \quad s = 1, 2, \dots, N_s. \tag{19}$$

In (19), Δt is the reading time interval, whereas N_s is the number of prior measurements used as extra inputs. Although a natural choice for incorporating a time series such as (19) would be recurrent neural networks (RNN)⁵⁴, in our case, N_s will be fixed throughout making feedforward networks a sufficient representation. Note that $N_s = 1$ is equivalent to the incorporation of differentials described in Section "Calibration input configuration II: differentials".

The extended flow diagram of the calibration procedure involving the time series of length N_s has been shown in Fig. 16. Figure 17 demonstrates the advantages of including short time series as auxiliary calibration model inputs for $N_s = 3$. Section "Results and discussion" will carry out a comprehensive analysis of the effects of the length N_s on calibration process reliability.

Global data scaling

The last algorithmic component integrated into the proposed calibration process involves global data scaling. This approach adjusts the correction coefficients anticipated by the ANN surrogate based on the current values of environmental factors, NO₂ measurements from both primary and redundant sensors, potential differentials, and a time series of N_s -length primary NO₂ data. The surrogate aims to minimize the disparity between the reference and low-cost sensor data in the least-square sense (cf. (1)). Yet, resolving (1) might reveal certain systematic discrepancies reliant on the measured NO₂ level, as depicted in Fig. 18a and b for a specific subset of training data. This distinction becomes apparent when examining the data sorted by reference NO₂ levels and through the scatter plot's slight skew seen in the bottom panel of Fig. 18b.

The global data scaling aims at reducing the discussed offsets by means of an affine transformation of the smoothed sensor measurements. In plain words, it corresponds to a 'rotation' of the scatter plot rendering it less skewed with respect to the identity mapping. A rigorous formulation of the process has been explained in Fig. 19.

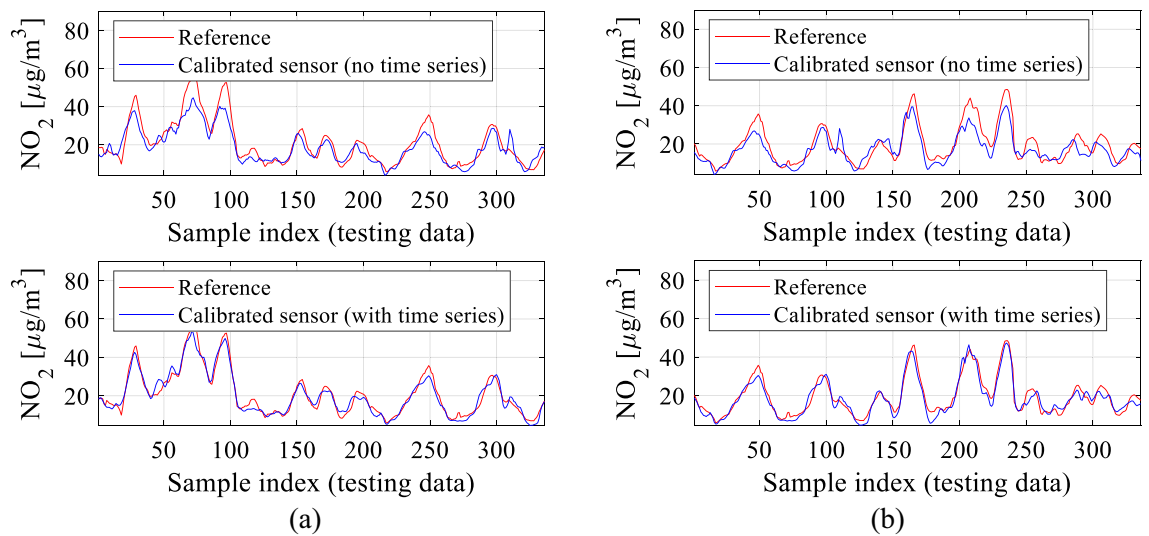


Figure 17. The effects of incorporating a time series of length $N_s = 3$ of prior NO₂ readings into the NN calibration model, along with the environmental parameters differentials. Shown are reference and calibrated low-sensor data without and with the mentioned time series, obtained for two selected sequences of the testing data: (a) first sequence, (b) second sequence.

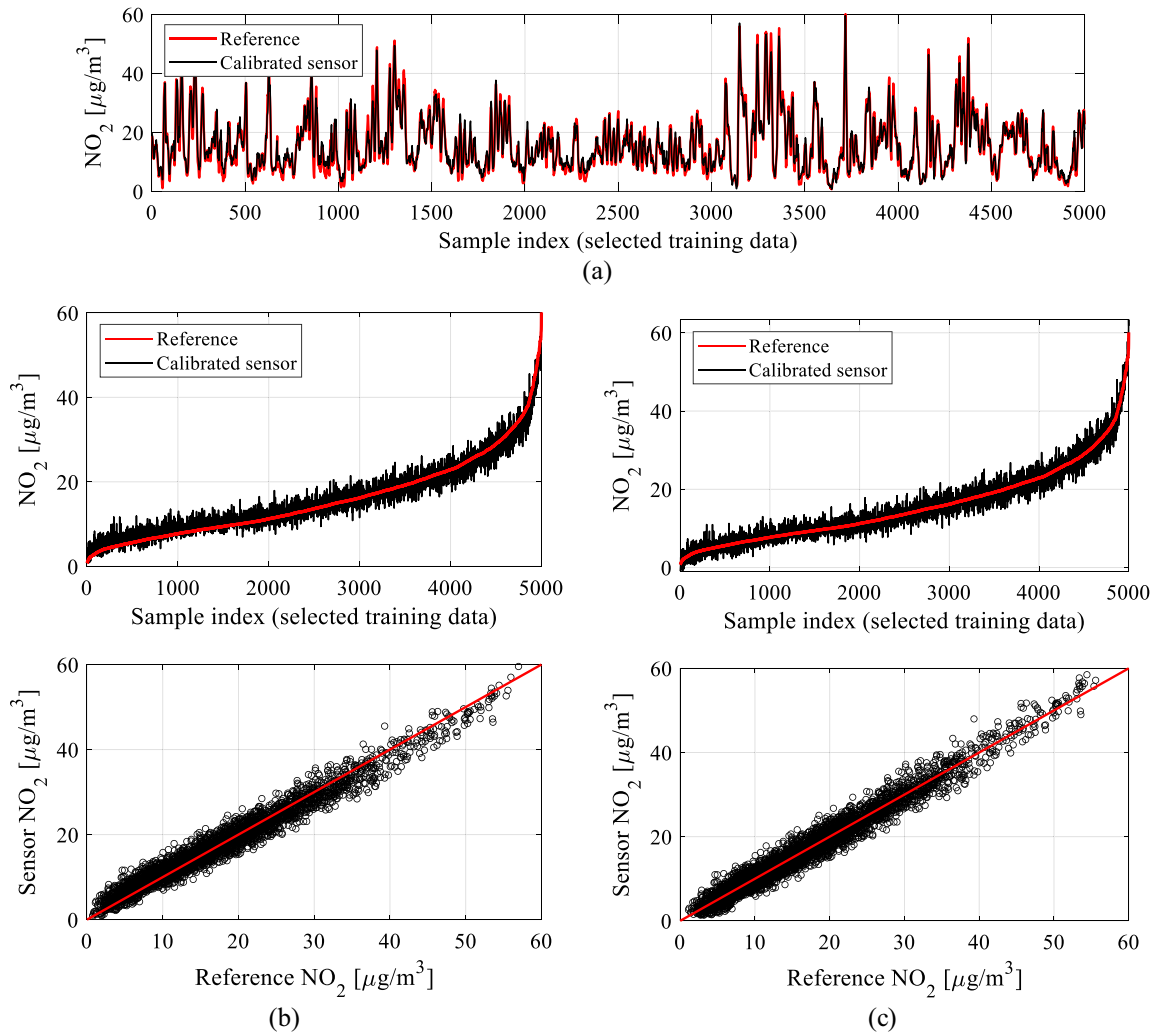


Figure 18. Global response correction: (a) a subset of selected training data; (b) the same data arranged based on increasing NO₂ reference readings (top) accompanied by the corresponding scatter plot (bottom). Despite the apparent alignment showcased in Fig. 18a, there is an observable systematic offset dependent on the level; (c) the same data after the application of global data scaling, showcasing a notable decrease in the systematic offset and an enhancement in the symmetry of the scatter plot. In this instance, global correction results in an improved correlation coefficient, rising from 0.93 to 0.95, and a reduction in RMSE from 2.1 to 1.8 μg/m³.

Notation:

- \mathbf{y}_r – ordered reference data vector (a red line in Fig. 18(b));
- \mathbf{y}_c – corresponding calibrated low-cost sensor data (a black line in Fig. 18(b));
- $S_m(\mathbf{y}_c)$ – smoothed vector \mathbf{y}_c (note that aggressive smoothing S_m would produce a monotonically increasing curve which represents (local) mean values for \mathbf{y}_c).

Global data scaling is realized by means of an affine transformation

$$\mathbf{y}_{c,G}^{(j)} = A_G(\mathbf{y}_c^{(j)} + D_G) \tag{20}$$

for $j = 1, \dots, N_0$, where A_G and D_G are found as

$$[A_G \ D_G] = \arg \min_{[A \ D]} \|\mathbf{y}_r - A(S_m(\mathbf{y}_c) + D)\| \tag{21}$$

Remark: Note that A_G and D_G are determined based on complete training data vectors and do not rely on any environmental or auxiliary parameters.

Figure 19. Global response correction through affine transformation of the ordered NO₂ data from the calibrated low-cost sensor.

Coefficients A_G and D_G are determined from the complete dataset; they are not functions of the environmental or auxiliary parameters.

The impact of implementing global data scaling is evident in Fig. 18c. In the depicted case, there is a noticeable reduction in the offset and an enhanced symmetry within the scatter plot. Simultaneously, the correlation coefficient improves from 0.93 to 0.95, while the RMSE decreases from 2.1 to 1.8 $\mu\text{g}/\text{m}^3$ based on the training data. Although its advantages might be somewhat constrained for the testing data, global data scaling still proves beneficial, as shown in Section "Results and discussion".

Again, it should be noted that the global data correction is a separate stage, which is applied after calibrating the sensor using the scaling coefficients A and D rendered by the ANN model. The inputs of the ANN model are the auxiliary parameters (vector \mathbf{z}_s), the primary sensor measurement y_s , and (optionally) the differentials and the time series of prior measurements.

The ANN model produces coefficients A and D being functions of these input variables and applies them to the low-cost sensor readings as in (2). The global correction (20) is applied afterwards using coefficients A_G and D_G obtained for the entire training dataset (i.e., not being functions of individual measurements). These coefficients are the same for all samples underdoing the global correction process.

Operating flow of NO₂ monitoring by means of calibrated sensor

Below, we summarize the operation of the complete calibration process of the low-cost sensor. The procedure combines the correction mechanisms detailed in Sections "Additive and multiplicative low-cost sensor correction" through "Global data scaling". The first step is pre-processing elucidated in Section "Statistical pre-processing of low-cost sensor measurements", where the overall distributions of the sensor and the reference data are aligned. Subsequently, the ANN surrogate predicts the (local) correction coefficients using the auxiliary vector \mathbf{z}_s and NO₂ reading y_s from the low-cost sensor, their differentials, as well as an N_s -long time series of prior NO₂ measurements from the primary sensor. The intermediate outcome y_c is obtained by applying the affine correction (2), (3). The last stage is global data scaling (20), (21), which produces the final corrected NO₂ reading. A flow diagram of the process has been shown in Fig. 20.

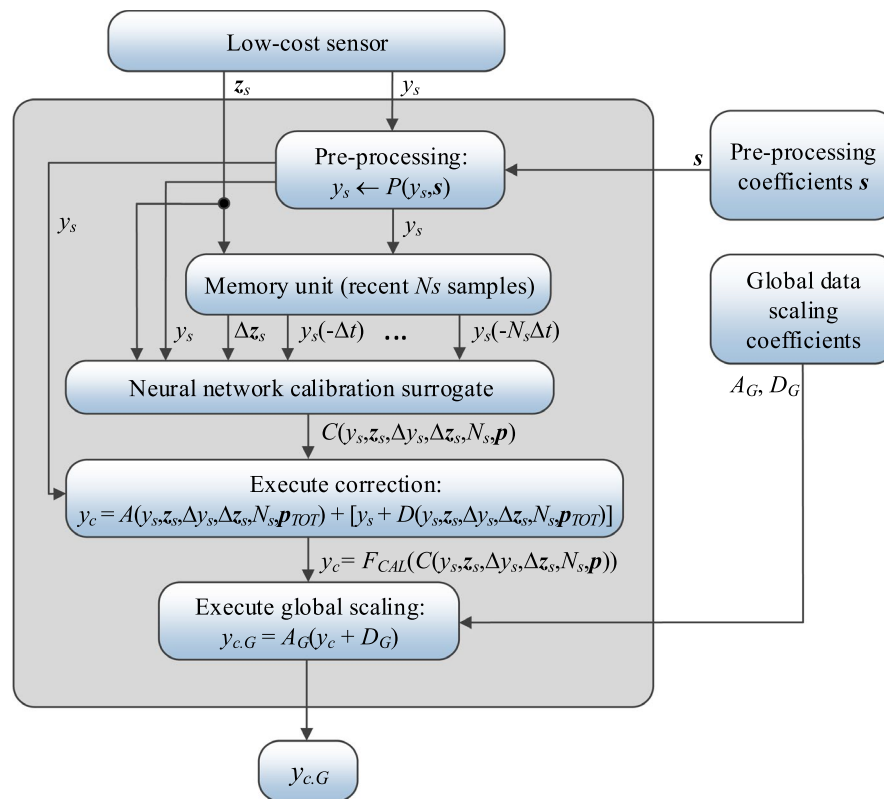


Figure 20. Low-cost sensor calibration procedure as proposed in this study. Pre-processing of the sensor readings is followed by generating (local) calibration coefficients using the ANN surrogate (based on the auxiliary vector \mathbf{z}_s , the actual NO₂ reading y_s from the low-cost sensor, their differentials, as well as a short-term time series of prior nitrogen dioxide readings from the primary sensor). The affine scaling is then applied to the sensor reading to produce the outcome y_c . Subsequently, global response correction is superimposed to produce the final corrected reading $y_{c.G}$.

Results and discussion

This section concentrates on validating the proposed calibration method for the low-cost sensor, applied to the autonomous monitoring platform detailed in Section "Autonomous NO₂ monitoring platform". The content is organized as follows. Section "Reference and low-cost sensor datasets" discusses the reference and low-cost sensor datasets. Section "Results" presents results obtained from various calibration setups explored in comparative experiments. Finally, Section "Discussion" summarizes findings and discusses the performance of the calibration process.

Reference and low-cost sensor datasets

The proposed calibration procedure has been validated using the datasets acquired from the reference stations (as outlined in Section "Reference data. Public monitoring stations") and the monitoring platforms (detailed in Section "Autonomous NO₂ monitoring platform"). The data was collected hourly between March and August 2023, cf. Figure 21. For the sake of illustration, Fig. 22 presents selected subsets of the reference and uncorrected low-cost sensor training and testing data. Significant disparities between the readings from the reference and the sensor can be observed, which poses a considerable challenge for the calibration process.

Data acquisition:

- Reference data acquired in the period from March to August 2023 from three reference stations allocated in the city of Gdansk, Poland (cf. Section 3);
- Low-cost sensor data acquired from cost-efficient platforms outlined in Section 2, allocated in close proximity of the respective reference stations;
- Data recorded every hour;
- Overall number of samples >10,000 (corrupted samples were removed).

Training data:

- Approximately ninety percent of the dataset was utilized to establish the calibration models (NN, global response correction scaling);
- The total number of training points $N_0 = 9325$.

Testing data:

- The remaining data, specifically $N_t = 3 \times 336 = 1008$ samples, were reserved for testing purposes;
- Test samples correspond to three distinct two-week periods at three different sensor locations
 - Station 1, April 1-15, 2023;
 - Station 2, July 15-29, 2023;
 - Station 3; July 1-14, 2023.

Figure 21. Characterization of the training and testing data acquired to carry out calibration of the low-cost sensor of Section "Autonomous NO₂ monitoring platform".

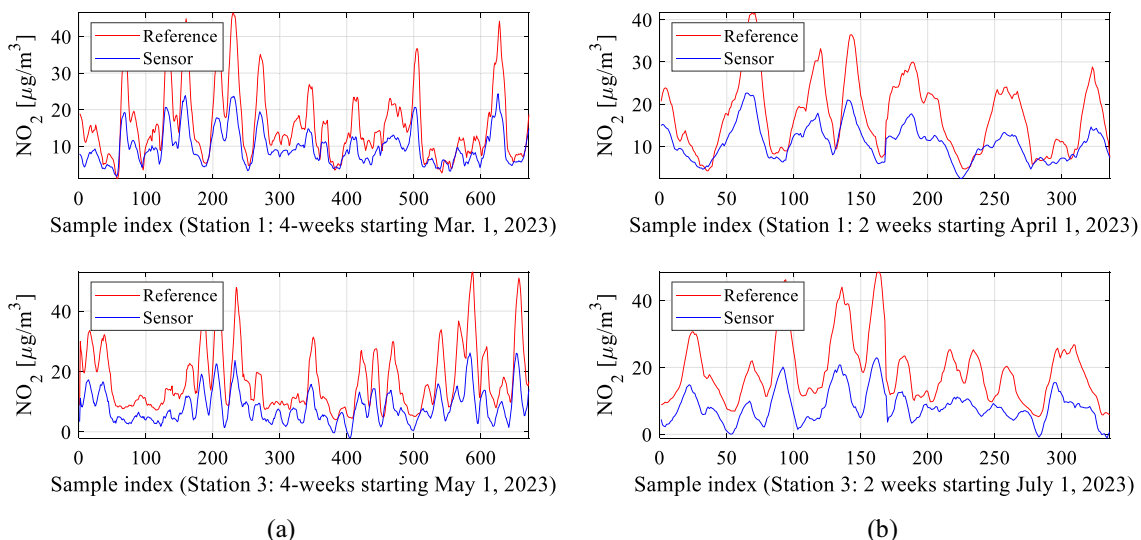


Figure 22. Selected subsets of NO₂ readings from the reference stations and the raw (uncorrected) low-cost sensors: (a) training data, (b) testing data.

Results

In this analysis, we delve into the calibration outcomes of the low-cost NO₂ sensor within the monitoring platform highlighted in Section "Autonomous NO₂ monitoring platform". We explore various setups of the calibration model inputs to assess the importance of specific algorithmic elements within the correction scheme. Additionally, we selectively enable or disable auxiliary mechanisms, i.e., pre-processing and global data scaling for some configurations. Table 1 presents all the scrutinized setups. Each configuration undergoes ten independent training cycles, and the model with the optimal set of hyper-parameters is chosen as the final model.

The calibration setups under examination are divided into four groups, denoted as A to D. The first group encompasses configurations that do not utilize the time series of previous NO₂ measurements. The second group involves setups that incorporate time series of past readings, varying in length (N_s), excluding global response correction. The third group combines time-series-based calibration with global data scaling. The final group incorporates pre-processing as detailed in Section "Statistical pre-processing of low-cost sensor measurements". Experimenting with different N_s values enables us to identify the most effective time series length.

The results from all calibration setups are consolidated in Table 2, encompassing the correlation coefficient and modeling error (RMSE) for both training and testing data (see Fig. 23 for definitions). To streamline the presentation, data visualization is provided for four specific calibration setups: B.4, and D.3. Figure 24 displays the reference, raw low-cost sensor, and calibrated sensor NO₂ measurements (training data) for two chosen eight-week periods. Figure 25 illustrates the same information for testing data across three two-week periods, while Fig. 26 showcases scatter plots for the testing data. Finally, Fig. 27 presents NO₂ measurements for setups B.4, and D.3 based on ascending reference readings.

Discussion

The experiments in Section "Results" aimed to verify the effectiveness of the proposed calibration process. One crucial aspect under examination was whether the correction strategy introduced could adequately align the reference and low-cost sensor readings, ensuring reliable monitoring of nitrogen dioxide. Furthermore, we aimed at verifying the relevance of correction mechanisms, specifically, the pre-processing and global data scaling procedures, and benefits of incorporating environmental parameter differentials, and time series of prior NO₂ readings from the low-cost sensor as additional calibration inputs. We were also interested in identifying the optimal length N_s of this series. It is also important to recall that the initial discrepancies between the low-cost sensor and the reference measurements are significant, whereas the NO₂ level changes considerably (from almost zero to sixty µg/m³) and often quickly, which make the calibration a challenging endeavour.

Calibration setup	Calibration input variables					Global response correction	Pre-processing
	Auxiliary data	NO ₂ reading from primary sensor (y_s)	Differentials		Time series $y_s(-\Delta t), \dots, y_s(-N_s \Delta t)$		
			Δz_s	Δy_s			
A.1	Restricted ($T_o, T_p, H_o,$ and H_i)	×	×	×	×	×	×
A.2	Restricted (z_s without P)	×	×	×	×	×	×
A.3	Restricted (z_s without P)	✓	×	×	×	×	×
A.4	Complete z_s	✓	×	×	×	×	×
A.5	Complete z_s	✓	×	×	×	✓	×
A.6	Complete z_s	✓	×	✓	×	✓	×
A.7	Complete z_s	✓	✓	✓	×	×	×
A.8	Complete z_s	✓	✓	✓	×	✓	×
B.1	Complete z_s	✓	✓	×	✓ ($N_s=0$)	×	×
B.2	Complete z_s	✓	✓	×	✓ ($N_s=2$)	×	×
B.3	Complete z_s	✓	✓	×	✓ ($N_s=4$)	×	×
B.4	Complete z_s	✓	✓	×	✓ ($N_s=6$)	×	×
B.5	Complete z_s	✓	✓	×	✓ ($N_s=8$)	×	×
C.1	Complete z_s	✓	✓	×	✓ ($N_s=0$)	✓	×
C.2	Complete z_s	✓	✓	×	✓ ($N_s=2$)	✓	×
C.3	Complete z_s	✓	✓	×	✓ ($N_s=4$)	✓	×
C.4	Complete z_s	✓	✓	×	✓ ($N_s=6$)	✓	×
C.9	Complete z_s	✓	✓	×	✓ ($N_s=8$)	✓	×
D.1	Complete z_s	✓	✓	×	✓ ($N_s=0$)	✓	✓
D.2	Complete z_s	✓	✓	×	✓ ($N_s=2$)	✓	✓
D.3	Complete z_s	✓	✓	×	✓ ($N_s=4$)	✓	✓
D.4	Complete z_s	✓	✓	×	✓ ($N_s=6$)	✓	✓
D.5	Complete z_s	✓	✓	×	✓ ($N_s=8$)	✓	✓

Table 1. Input setups of the calibration model considered in verification experiments.

Calibration setup	Training data		Testing data	
	Correlation coefficient r	RMSE [$\mu\text{g}/\text{m}^3$]	Correlation coefficient r	RMSE [$\mu\text{g}/\text{m}^3$]
A.1	0.82	4.0	0.70	5.6
A.2	0.89	3.0	0.81	4.3
A.3	0.91	2.8	0.84	4.0
A.4	0.93	2.5	0.86	3.9
A.5	0.94	2.4	0.878	3.6
A.6	0.93	2.6	0.886	3.4
A.7	0.93	2.4	0.896	3.3
A.8	0.94	2.3	0.903	3.17
B.1	0.92	2.6	0.892	3.3
B.2	0.93	2.4	0.907	3.1
B.3	0.93	2.4	0.907	3.1
B.4	0.93	2.5	0.903	3.2
B.5	0.93	2.5	0.900	3.3
C.1	0.92	2.5	0.905	3.1
C.2	0.93	2.4	0.921	2.9
C.3	0.93	2.4	0.923	2.8
C.4	0.93	2.4	0.917	2.9
C.5	0.93	2.5	0.912	3.1
D.1	0.94	2.0	0.932	2.6
D.2	0.95	1.9	0.944	2.4
D.3	0.96	1.8	0.947	2.4
D.4	0.96	1.8	0.947	2.4
D.5	0.96	1.8	0.945	2.4

Table 2. Sensor calibration performance: correlation coefficients and RMSE.

Correlation coefficient:

$$r = \frac{\sum_{j=1}^{N_t} (y_c^{(j)} - \bar{y}_c)(y_{rt}^{(j)} - \bar{y}_{rt})}{\sqrt{\sum_{j=1}^{N_t} (y_c^{(j)} - \bar{y}_c)^2 \sum_{j=1}^{N_t} (y_{rt}^{(j)} - \bar{y}_{rt})^2}} \quad (22)$$

Root-mean-square error:

$$RMSE = \sqrt{\frac{\sum_{j=1}^{N_t} (y_c^{(j)} - y_{rt}^{(j)})^2}{N_t}} \quad (23)$$

where $y_c^{(j)}$ and $y_{rt}^{(j)}$, $j = 1, \dots, N_t$, are the corrected sensor and reference NO_2 readings, whereas the quantities with bars are the respective mean values.

Figure 23. Definitions of the correlation coefficient r and RMSE.

The findings in Table 2 showcase the exceptional performance of the proposed calibration technique. Among the calibration setups assessed, the most effective configurations belong to group D, specifically D.3 and D.4. These setups integrate all correction mechanisms outlined in Section "Precise sensor calibration using statistical pre-processing, ANN surrogates, and global data scaling", encompassing pre-processing, global data scaling, and leveraging extended input variables covering environmental parameters, auxiliary NO_2 readings, differentials, and medium-length time series (N_s ranging between four and six). For instance, in setup D.3, the correlation coefficient reaches approximately 0.95, with an RMSE of 2.4 $\mu\text{g}/\text{m}^3$ for the testing data. Moreover, the average relative RMS error is merely around 11 percent. The precision of the calibrated sensor is evident in its excellent alignment with the reference data, as observed in both the training (Fig. 24d) and testing data (Fig. 25d). The reported numbers are particularly impressive when compared to the metrics of the raw (uncorrected) sensor, which are as follows: correlation coefficients 0.07 and 0.04 (training and testing data, respectively), and RMSE of 8.9 and 10.8 $\mu\text{g}/\text{m}^3$ (training and testing data, respectively).

A review of the results across various calibration setups underscores the significance of each incorporated correction mechanism. For instance, augmenting the inputs in the calibration model significantly impacts both the correlation coefficient and RMSE. Comparing configurations A.1, A.2, A.3, A.4, and A.7 (excluding global response correction) highlights this, where the correlation coefficient improves from 0.7 to 0.89, and RMSE drops

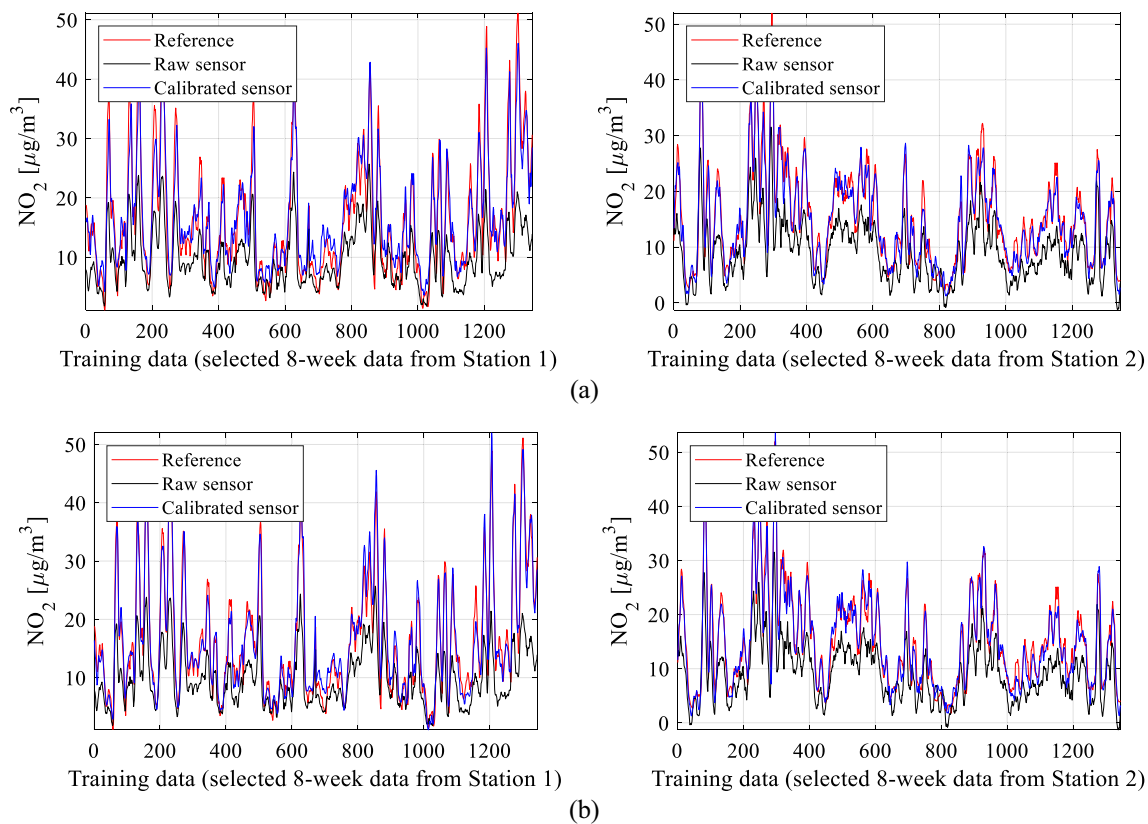


Figure 24. Sensor calibration performance for selected subsets of the training data: (a) setup B.4, (b) setup D.3.

from 5.6 to 3.4 $\mu\text{g}/\text{m}^3$. Consistent integration of global response correction consistently bolsters the correlation coefficient by nearly 0.02 and reduces RMSE by about 0.2 $\mu\text{g}/\text{m}^3$ (e.g., comparing setup A.5 versus A.4, or C.1 versus B.1).

Introducing time series data further enhances results, achieving up to a 0.03 improvement in correlation coefficient and a reduction of 0.3 $\mu\text{g}/\text{m}^3$ in RMSE (e.g., setups C.3 or C.4). Moreover, data pre-processing significantly contributes to calibration enhancements by adding up to 0.03 to the correlation coefficient and reducing RMSE by nearly 0.3 $\mu\text{g}/\text{m}^3$. These improvements are visually evident in Figs. 24, 25, and 26, where transitioning from simpler configurations to more advanced ones (e.g., B.4 and D.3) noticeably improves alignment between the reference and corrected low-cost sensor readings. Additionally, it centres the scatter plots closer to the identity function.

The enhancements in reliability are also visually highlighted in Fig. 27, where both training and testing data are arranged by ascending reference NO_2 levels. Moving from the simpler setup A.2 through intermediate stages (A.7 and B.4) to the advanced configuration D.3 significantly reduces deviations between the reference and calibrated sensor readings. An in-depth analysis of setups B and C reveals that the most favourable configuration in terms of the time series length is $N_s = 4$, showcasing the highest correlation coefficient and minimal RMSE. However, with the inclusion of pre-processing (setups D), the impact of N_s becomes less distinctive, suggesting that the calibration performance becomes more resilient to variations in this parameter.

Additional experiments were conducted to verify the effects of including auxiliary NO_2 sensor readings as supplementary calibration inputs. The considered setups are listed in Table 3. The results are encapsulated in Table 4. Note that setups E.1 and E.5 were previously considered as Cases E.1 and E.3 in Table 1. These are repeated to ensure completeness of the data in Tables 3 and 4. Note that incorporating auxiliary NO_2 sensor data does improve the calibration process dependability. Also, it can be observed that the second auxiliary sensor S_2 has a slightly higher impact, as it can be inferred from the values of correlation coefficient and RMSE. On the other hand, when the auxiliary sensors are not utilized, data alignment degrades noticeably (cf. setup E.2 versus E.3, E.4, or E.5). Furthermore, including the primary sensor measurements is also important.

For supplementary validation, the calibration approach introduced in this paper has been compared to several benchmark methods, specifically, linear regression, neural-network-based calibration, as well as calibration implemented using a convolutional neural network (CNN)⁵⁵. In the case of ANN/CNN, the neural network predicts the calibrated model output directly instead of rendering the correction coefficients. Linear regression is a model of the form

$$S(\mathbf{z}_s) = \alpha_0 + \alpha_1 T_o + \alpha_2 T_i + \alpha_3 H_o + \alpha_4 H_i + \alpha_5 S_1 + \alpha_6 S_2 \quad (22)$$

when using vector \mathbf{z}_s as calibration input, and

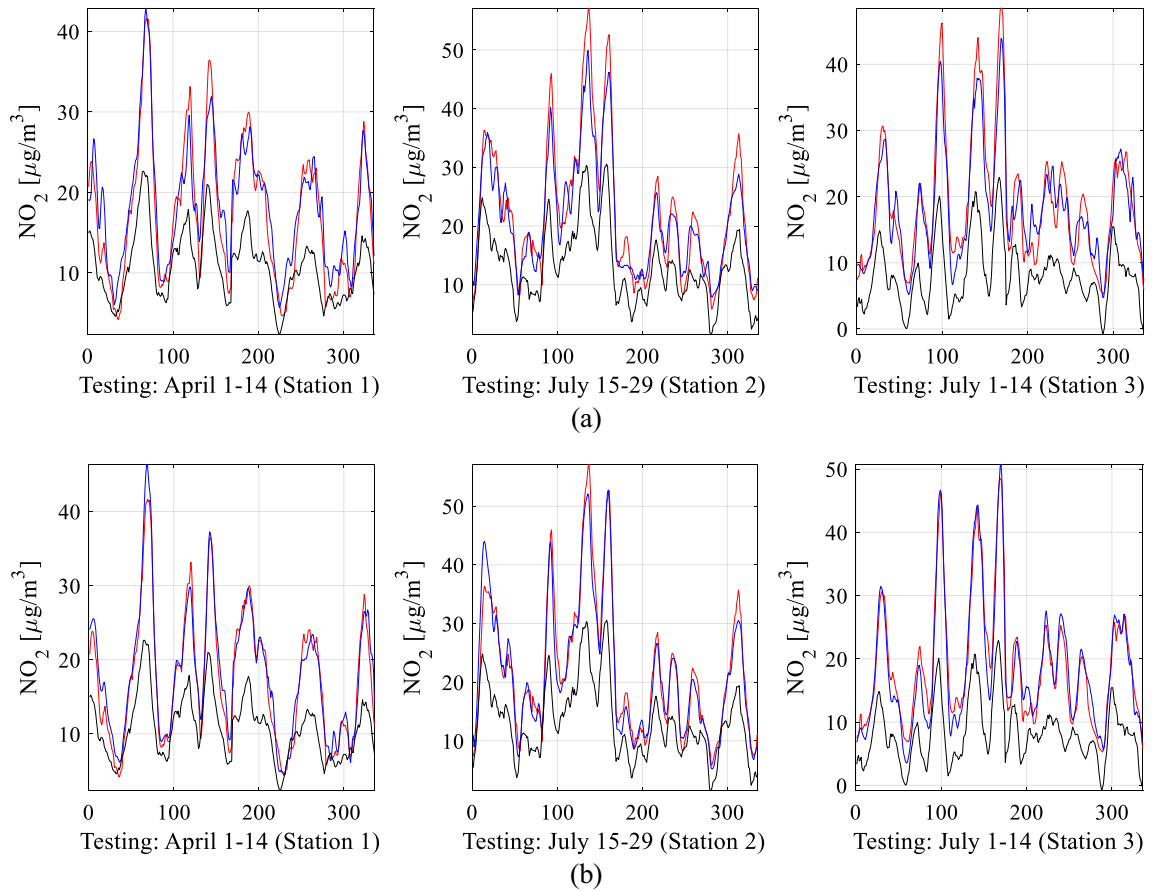


Figure 25. Sensor calibration performance for selected subsets of the testing data: (a) setup B.4, (b) setup D.3.

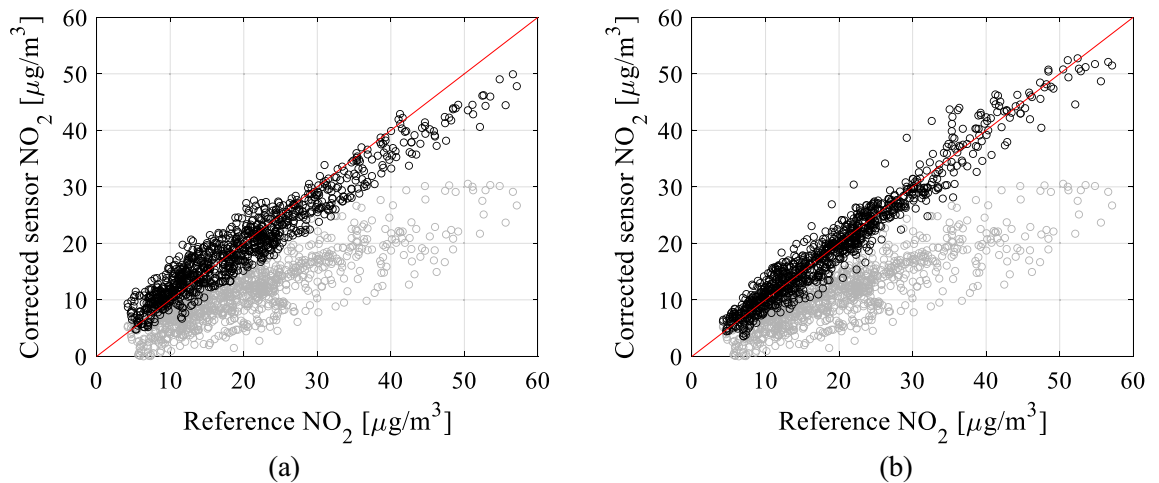


Figure 26. Scatter plots for the testing data (uncorrected—gray, corrected—black): (a) setup B.4, (b) setup D.3.

$$S_y(z_s, y_s) = \alpha_0 + \alpha_1 T_o + \alpha_2 T_i + \alpha_3 H_o + \alpha_4 H_i + \alpha_5 S_1 + \alpha_6 S_2 + \alpha_7 y_s \tag{23}$$

when using extended calibration inputs (i.e., primary sensor data). The coefficients in (22) and (23) are found through least-square regression based on the training data. The ANN uses the same architecture as described in Section "Precise sensor calibration using statistical pre-processing, ANN surrogates, and global data scaling". CNN architecture is uses filters of the size $4 \times 1 \times 1$, three convolution layers of spatial sizes 32, 16, and 8, followed by a fully connected layer of the size 64 neurons (version I), layers of sizes 64, 32, 16 (version II), and 126, 64, and 32 (version III), as well as batch normalization and ReLU layers in between the convolution layers. CNN is trained using the ADAM's algorithm with a mini batch size of 1000 [70]. Table 5 gathers the numerical results.

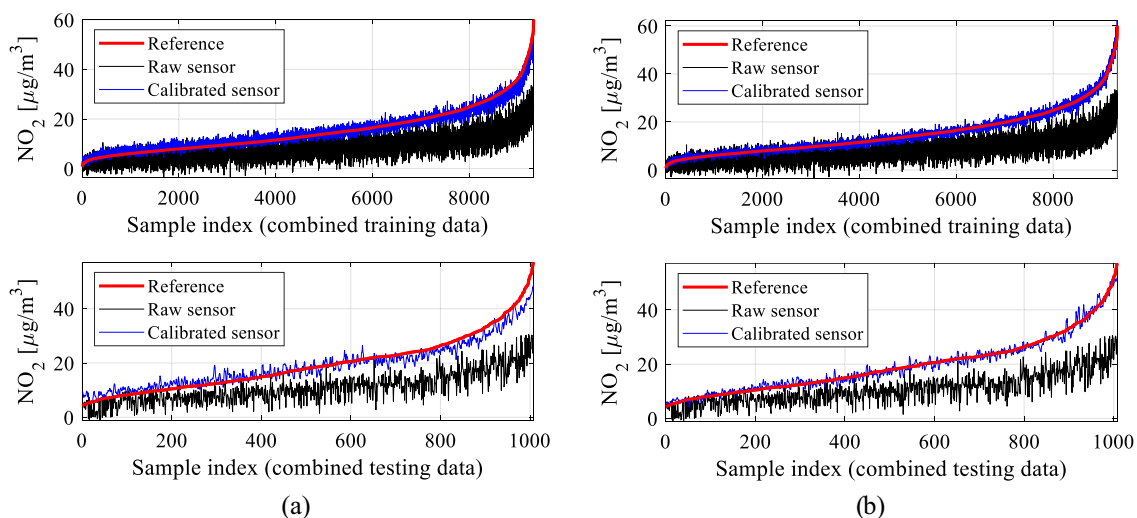


Figure 27. Performance of sensor calibration for: (a) setup B.4, (b) setup D.3. Shown are the entire training dataset (top) and testing dataset (bottom), arranged in ascending order according to NO₂ reference readings. Note substantial enhancement achieved through calibration, i.e., bringing the calibrated sensor readings much closer to their corresponding reference measurements compared to the raw data.

Case	Calibration model	Calibration input variables
E.1	ANN (= Case A.1 of Table 2)	Restricted auxiliary data (T_o , T_p , H_o , and H_i)
E.2	ANN	Restricted auxiliary data (T_o , T_p , H_o , and H_i) and y_s
E.3	ANN	Restricted auxiliary data (T_o , T_p , H_o , and H_i), S_1 and y_s
E.4	ANN	Restricted auxiliary data (T_o , T_p , H_o , and H_i), S_2 and y_s
E.5	ANN (= Case A.3 of Table 2)	Restricted auxiliary data (T_o , T_p , H_o , and H_i), S_1 , S_2 and y_s

Table 3. Verification case studies: calibration model setup.

Case	Training data		Testing data	
	Correlation coefficient r^2	RMSE [$\mu\text{g}/\text{m}^3$]	Correlation coefficient r^2	RMSE [$\mu\text{g}/\text{m}^3$]
E.1	0.82	4.0	0.70	5.6
E.2	0.88	3.2	0.81	4.4
E.3	0.90	3.0	0.84	4.1
E.4	0.89	3.0	0.85	4.0
E.5	0.91	2.8	0.84	4.0

Table 4. Sensor calibration performance for calibration scenarios listed in Table 3.

Calibration method	Training data		Testing data	
	Correlation coefficient r^2	RMSE [$\mu\text{g}/\text{m}^3$]	Correlation coefficient r^2	RMSE [$\mu\text{g}/\text{m}^3$]
Linear regression $S(z_s)$	0.28	7.8	0.07	9.9
Linear regression $S_p(z_s, y_s)$	0.66	5.4	0.56	6.8
Direct ANN-based prediction (z_s)	0.77	4.4	0.26	8.8
Direct ANN-based prediction (z_s and y_s)	0.83	3.8	0.61	6.4
Direct CNN-based prediction (z_s and y_s) (convolution layers: 32, 16, 8)	0.50	6.5	0.29	8.6
Direct CNN-based prediction (z_s and y_s) (convolution layers: 64, 32, 16)	0.72	4.8	0.45	7.6
Direct CNN-based prediction (z_s and y_s) (convolution layers: 128, 64, 32)	0.77	4.5	0.42	7.7

Table 5. Comparative studies: linear regression and direct ANN/CNN-based prediction.

It should be noted that the calibration methodology proposed in this study provides significantly better results, both in terms of correlation coefficients and RMSE. Utilization of affine correction (cf. Table 2) is superior to direct prediction of the calibrated sensor when using ANN of the same architecture as well as CNN.

In summary, the showcased calibration approach proves remarkably effective. The corrected low-cost sensor measurements closely align with the reference readings, particularly in the advanced configurations, such as D.3, representing the optimal calibration setup. In practical terms, this sensor correction can be integrated offline or implemented within the platform using its on-board computational resources, as outlined in Section "Autonomous NO₂ monitoring platform".

Conclusion

This article introduced an innovative methodology for high-efficiency calibration of affordable nitrogen dioxide sensors. The proposed technique integrates various correction mechanisms, encompassing data pre-processing, additive and multiplicative response adjustments executed by an artificial neural network (ANN) surrogate, and global data scaling. The pre-processing step focuses on aligning the distribution of low-cost sensor readings across the entire training dataset with reference measurements. Utilizing the ANN surrogate, the method predicts specific correction coefficients based on environmental parameters and additional NO₂ readings from redundant sensors. Additionally, the calibration model explores extended input parameters, including differentials of environmental variables and historical time series data from the primary sensor, proving their significance. Global data scaling acts as the final step, enhancing scatter plot symmetry and consequent reduction in prediction errors for the calibrated sensor.

Our technique was applied and validated on a monitoring platform developed at Gdansk University of Technology, Poland, comprising primary and secondary NO₂ detectors, environmental sensors, and custom-designed electronic systems for data transmission and monitoring protocols. The validation involved data from public monitoring stations in Gdansk, Poland. Extensive comparative experiments across diverse calibration model configurations underscored the importance of the integrated algorithmic components. The most comprehensive setup, encompassing all correction mechanisms, demonstrated exceptional reliability, achieving a correlation coefficient of 0.95 between reference and corrected sensor data, with an RMSE below 2.4 µg/m³ (an average relative RMS error of just eleven percent). This high efficacy underscores the practical viability of low-cost NO₂ monitoring.

Future endeavors will focus on refining the precision of calibrated low-cost NO₂ monitoring. One avenue involves integrating supplementary gas detectors like SO₂, CO, and O₃ into the measurement platform. This addition aims to leverage their readings as supplemental data sources to further refine the calibration model, particularly regarding cross-sensitivity considerations. Additionally, exploring advanced machine learning methodologies, including convolutional neural networks (CNNs) and recurrent neural networks (RNNs), is on the agenda. RNNs, adept at managing time series of varying lengths, may specifically enhance monitoring reliability by harnessing such data.

Data availability

The datasets used and/or analyzed during the current study available from the corresponding author on reasonable request.

Received: 6 February 2024; Accepted: 17 April 2024

Published online: 21 April 2024

References

- Chen, T.-M., Kuschner, W. G., Gokhale, J. & Shofer, S. Outdoor air pollution: Nitrogen dioxide, sulfur dioxide, and carbon monoxide health Effects. *Am. J. Med. Sci.* **333**(4), 249–256 (2007).
- Zhao, S. *et al.* Assessing NO₂-related health effects by non-linear and linear methods on a national level. *Sci. Total Environ.* **744**, 140909 (2020).
- Guerriero, C., Chatzikiakou, L., Cairns, J. & Mumovic, D. The economic benefits of reducing the levels of nitrogen dioxide (NO₂) near primary schools: The case of London. *J. Environ. Manag.* **181**, 615–622 (2016).
- Kelly, F. J. & Fussell, J. C. Air pollution and airway disease. *Clin. Exp. Allergy* **41**(8), 1059–1071 (2011).
- Schwela, D. Air pollution and health in urban areas. *Rev. Environ. Health* **15**(1–2), 13–42 (2000).
- Salonen, H., Salthammer, T. & Morawska, L. Human exposure to NO₂ in school and office indoor environments. *Environ. Int.* **130**, 104887 (2019).
- Mauzerall, D. L., Sultan, B., Kim, N. & Bradford, D. F. NO_x emissions from large point sources: variability in ozone production, resulting health damages and economic costs. *Atmos. Environ.* **39**(16), 2851–2866 (2005).
- Agras, J. & Chapman, D. The Kyoto protocol, cafe standards, and gasoline taxes. *Contemp. Econ. Policy* **17**(3), 296–308 (1999).
- World Health Organization. *Air Quality Guidelines: Global Update 2005: Particulate Matter, Ozone, Nitrogen Dioxide, and Sulfur Dioxide* (World Health Organization, 2006).
- OECD. *The Economic Consequences of Outdoor Air Pollution* (OECD Publishing, 2016).
- Rodgers, M. O., Bradshaw, J. D. & Davis, D. D. Photofragmentation—Laser induced fluorescence detection of NO₂. In *Topical Meeting on Spectroscopy in Support of Atmospheric Measurements (1980), Paper TuP17* (ed. Rodgers, M. O.) (Optica Publishing Group, 1980).
- Platt, U. Air monitoring by differential optical absorption spectroscopy. In *Encyclopedia of Analytical Chemistry* (ed. Platt, U.) 1–28 (Wiley, 2017).
- Matsumoto, J., Hirokawa, J., Akimoto, H. & Kajii, Y. Direct measurement of NO₂ in the marine atmosphere by laser-induced fluorescence technique. *Atmos. Environ.* **35**(16), 2803–2814 (2001).
- Berden, G., Peeters, R. & Meijer, G. Cavity ring-down spectroscopy: Experimental schemes and applications. *Int. Rev. Phys. Chem.* **19**, 565–607 (2010).
- Yu, H. *et al.* A deep calibration method for low-cost air monitoring sensors with multilevel sequence modeling. *IEEE Trans. Instrum. Meas.* **69**(9), 7167–7179 (2020).

16. Bi, J., Wildani, A., Chang, H. H. & Liu, Y. Incorporating low-cost sensor measurements into high-resolution PM_{2.5} modeling at a large spatial scale. *Environ. Sci. Technol.* **54**, 2152–2162 (2020).
17. Castell, N. *et al.* Can commercial low-cost sensor platforms contribute to air quality monitoring and exposure estimates?. *Environ. Int.* **99**, 293–302 (2017).
18. Spinelle, L., Gerboles, M., Villani, M. G., Aleixandre, M. & Bonavitacola, F. Field calibration of a cluster of low-cost available sensors for air quality monitoring. Part A: Ozone and nitrogen dioxide. *Sens. Actuators B-Chem.* **215**, 249–257 (2015).
19. Fonollosa, J., Fernández, L., Gutiérrez-Gálvez, A., Huerta, R. & Marco, S. Calibration transfer and drift counteraction in chemical sensor arrays using Direct Standardization. *Sens. Actuators B-Chem.* **236**, 1044–1053 (2016).
20. Rai, A. C. *et al.* End-user perspective of low-cost sensors for outdoor air pollution monitoring. *Sci. Total Environ.* **607**, 691–705 (2017).
21. Kim, H., Müller, M., Henne, S. & Hüglin, C. Long-term behavior and stability of calibration models for NO and NO₂ low-cost sensors. *Atmos. Meas. Tech.* **15**, 2979–2992 (2022).
22. Poupyry, S., Medjaher, K. & Béler, C. Data reliability and fault diagnostic for air quality monitoring station based on low cost sensors and active redundancy. *Measurement* **223**, 113800 (2023).
23. Carotta, M. C. *et al.* Nanostructured thick-film gas sensors for atmospheric pollutant monitoring: Quantitative analysis on field tests. *Sensors Actuators B Chem.* **76**(1–3), 336–342 (2001).
24. Wang, Z. *et al.* Improved deep bidirectional recurrent neural network for learning the cross-sensitivity rules of gas sensor array. *Sensors Actuators B Chem.* **401**, 134996 (2024).
25. Zimmerman, N. *et al.* A machine learning calibration model using random forests to improve sensor performance for lower-cost air quality monitoring. *Atmos. Meas. Tech.* **11**, 291–313 (2018).
26. Gorshkova, A. *et al.* Enhancement in NO₂ sensing properties of SWNTs: A detailed analysis on functionalization of SWNTs with Z-Gly-OH. *J. Mater. Sci. Mater. Electron.* **34**, 102 (2023).
27. Jiao, W. *et al.* Community Air Sensor Network (CAIRSENSE) project: Evaluation of low-cost sensor performance in a suburban environment in the southeastern United States. *Atmos. Meas. Tech.* **9**, 5281–5292 (2016).
28. Lewis, A. C. *et al.* Evaluating the performance of low cost chemical sensors for air pollution research. *Faraday Discuss.* **189**, 85–103 (2016).
29. Han, P. *et al.* Calibrations of low-cost air pollution monitoring sensors for CO, NO₂, O₃, and SO₂. *Sensors* **21**, 256 (2021).
30. Müller, M. *et al.* Integration and calibration of non-dispersive infrared (NDIR) CO₂ low-cost sensors and their operation in a sensor network covering Switzerland. *Atmos. Meas. Tech.* **13**, 3815–3834 (2020).
31. Shusterman, A. A. *et al.* The Berkeley atmospheric CO₂ observation network: Initial evaluation. *Atmos. Chem. Phys. Discuss.* **16**, 13449–13463 (2016).
32. Andersen, T., Scheeren, B., Peters, W. & Chen, H. A UAV-based active AirCore system for measurements of greenhouse gases. *Atmos. Meas. Tech.* **11**, 2683–2699 (2018).
33. Kunz, M. *et al.* Surface flux estimates derived from UAS-based mole fraction measurements by means of a nocturnal boundary layer budget approach. *Atmos. Meas. Tech.* **13**, 1671–1692 (2020).
34. Bigi, A., Mueller, M., Grange, S. K., Ghermandi, G. & Hueglin, C. Performance of NO, NO₂ low cost sensors and three calibration approaches within a real world application. *Atmos. Meas. Tech.* **11**, 3717–3735 (2018).
35. Nowack, P., Konstantinovskiy, L., Gardiner, H. & Cant, J. Machine learning calibration of low-cost NO₂ and PM₁₀ sensors: Non-linear algorithms and their impact on site transferability. *Atmosph. Meas. Tech.* **14**, 5637–5655 (2021).
36. D'Elia, G. *et al.* Influence of concept drift on metrological performance of low-cost NO₂ sensors. *IEEE Trans. Instrum. Meas.* **71**, 1–11 (2022).
37. Jain, S., Presto, A. A. & Zimmerman, N. Spatial modeling of daily PM_{2.5}, NO₂, and CO concentrations measured by a low-cost sensor network: Comparison of linear, machine learning, and hybrid land use models. *Environ. Sci. Technol.* **55**(13), 8631–8641 (2021).
38. Ionascu, M.-E. *et al.* Calibration of CO, NO₂, and O₃ using Airify: A low-cost sensor cluster for air quality monitoring. *Sensors* **21**, 7977 (2021).
39. Bi, J. *et al.* Contribution of low-cost sensor measurements to the prediction of PM_{2.5} levels: A case study in Imperial County, California, USA. *Environ. Res.* **180**, 108810 (2020).
40. van Zoest, V., Osei, F. B., Stein, A. & Hoek, G. Calibration of low-cost NO₂ sensors in an urban air quality network. *Atmos. Environ.* **210**, 66–75 (2019).
41. De Vito, S. *et al.* Dynamic multivariate regression for on-field calibration of high speed air quality chemical multi-sensor systems. In *XVIII AISEM Annual Conference* (ed. De Vito, S.) 1–3 (IEEE, 2015).
42. Masson, N., Piedrahita, R. & Hannigan, M. Quantification method for electrolytic sensors in long-term monitoring of ambient air quality. *Sensors* **15**, 27283–27302 (2015).
43. Esposito, E. *et al.* Dynamic neural network architectures for on field stochastic calibration of indicative low cost air quality sensing systems. *Sensors Actuators B Chem.* **231**, 701–713 (2016).
44. Wang, Z. *et al.* Self-adaptive temperature and humidity compensation based on improved deep BP neural network for NO₂ detection in complex environment. *Sensors Actuators B Chem.* **362**, 131812 (2022).
45. BeagleBone® Blue, BeagleBoard. <https://www.beagleboard.org/boards/beaglebone-blue>.
46. SGX-7NO₂ Datasheet, Industrial Nitrogen Dioxide (NO₂) Sensor, SGX Sensortech. <https://www.sgxsensortech.com/content/uploads/2021/10/DS-0338-SGX-7NO2-datasheet.pdf>.
47. Four electrode NO₂ sensor, SemaTech (7E4-NO2-5) (PN: 057-0400-200), SemeaTech Inc. https://www.sematech.com/uploads/datasheet/7series/057-0400-200_EN.pdf.
48. Datasheet MiCS-2714 1107 rev 6, SGX Sensortech. https://www.sgxsensortech.com/content/uploads/2014/08/1107_Datasheet-MiCS-2714.pdf.
49. Humidity Sensor BME280, Bosch Sensortec. <https://www.bosch-sensortec.com/products/environmental-sensors/humidity-sensors-bme280/>.
50. ARMAG Foundation: Home. <https://armaag.gda.pl/en/index.htm>.
51. Vang-Mata, R. (ed.) *Multilayer Perceptrons* (Nova Science Pub. Inc., 2020).
52. Dlugosz, S. *Multi-Layer Perceptron Networks for Ordinal Data Analysis* (Logos Verlag, 2008).
53. Hagan, M. T. & Menhaj, M. Training feed-forward networks with the Marquardt algorithm. *IEEE Trans. Neural Netw.* **5**(6), 989–993 (1994).
54. Salem, F. M. *Recurrent Neural Networks. From Simple to Gated Architectures* (Springer, 2022).
55. Aggarwal, C. C. *Neural Networks and Deep Learning* (Springer, 2018).

Acknowledgements

The research leading to these results has received funding from the Norway Grants 2014–2021 via the National Centre for Research and Development, grant NOR/POLNOR/HAPADS/0049/2019-00. This work was also supported in part by the Icelandic Centre for Research (RANNIS) Grant 217771. The authors would also like to

thank the Agency of Regional Atmospheric Monitoring Gdansk-Gdynia-Sopot (ARMAG) for providing free of charge data from the reference measurement stations.

Author contributions

S.K. - conceptualization, data curation, formal analysis, methodology, writing A.P.D. - formal analysis, investigation, software, writing M.W. - funding acquisition, software, resources, review & editing B.P. - funding acquisition, software, resources, review & editing.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to S.K.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024