Engineering Science and Technology, an International Journal 54 (2024) 101729

Contents lists available at ScienceDirect

Engineering Science and Technology, an International Journal

journal homepage: www.elsevier.com/locate/jestch

Machine-learning-based precise cost-efficient NO_2 sensor calibration by means of time series matching and global data pre-processing

Slawomir Koziel [a,b,*], Anna Pietrenko-Dabrowska [b], Marek Wojcikowski [b], Bogdan Pankiewicz [b]

[a] Engineering Optimization & Modeling Center, Reykjavik University, 102 Reykjavik, Iceland
[b] Faculty of Electronics, Telecommunications and Informatics, Gdansk University of Technology, 80-233 Gdansk, Poland

ARTICLE INFO

Keywords:
Air quality monitoring
Nitrogen dioxide
Cost-efficient sensors
Sensor correction
Machine learning
Time series matching
Neural networks

ABSTRACT

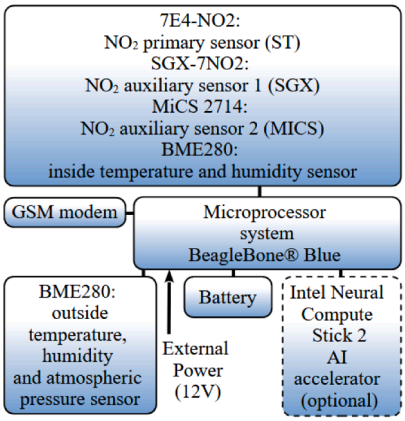
Air pollution remains a considerable contemporary challenge affecting life quality, the environment, and economic well-being. It encompasses an array of pollutants—gases, particulate matter, biological molecules—emanating from sources such as vehicle emissions, industrial activities, agriculture, and natural occurrences. Nitrogen dioxide (NO_2), a harmful gas, is particularly abundant in densely populated urban areas. Given its detrimental impact on health and the environment, precise monitoring of NO_2 levels is crucial for devising effective strategies to mitigate risks. However, precise measurement of NO_2 presents challenges as it traditionally relies on expensive and heavy (therefore, stationary) equipment. This has led to the pursuit of more affordable alternatives, though their dependability is frequently questionable. This study introduces an innovative technique for precise calibration of low-cost NO_2 sensors. Our methodology involves statistical preprocessing of sensor measurements to align their distributions with reference data. The core of the calibration model is an artificial neural network (ANN), trained to synchronize sensor and reference time series measurements. It incorporates environmental variables such as temperature, humidity, and atmospheric pressure, along with readings from redundant NO_2 sensors for cross-referencing, and short time series of primary sensor NO_2 measurements. This enables efficient learning of typical sensor changes over time in relation to these factors. Additionally, an interpolative kriging model serves as an auxiliary surrogate to enhance the correction process's reliability. Validation using an autonomous monitoring platform from Gdansk University of Technology, Poland, and public reference station data gathered over five months shows remarkable calibration accuracy, with a correlation coefficient close to 0.95 and RMSE of 2.4 $\mu g/m^3$. These results position the corrected sensor as an attractive and cost-effective alternative to conventional NO_2 measurement methods.

1. Introduction

Nitrogen dioxide (NO_2) pollution, originating from sources such as vehicle emissions, industrial processes, and combustion, remains a critical environmental issue. As part of nitrogen oxides (NO_x), this gas significantly affects air quality, leading to health issues (e.g., respiratory problems, aggravating asthma, exacerbating lung diseases), but also environmental damage. NO_2 reacts in the atmosphere to form harmful particles and ozone impacting ecosystems, and even contributing to climate change [1–8]. NO_x emissions notably contribute to the formation of photochemical smog, the onset of acid rain, and ecological deterioration in water reservoirs [9]. Additionally, elevated NO_x levels can raise O_3 levels, adversely impacting agriculture, also causing harm to materials, such as accelerating the corrosion of various metals [10]. Monitoring and mitigating NO_2 levels are crucial for addressing their detrimental impact on both human health and the environment [13]. Rigorous regulations, like the CAFE Directive, aim to control NO_2 concentrations, setting an annual average below 40 $\mu g/m^3$ and restricting hourly levels to not surpass 200 $\mu g/m^3$ for more than 18 h per year [11]. Even stricter limits have been suggested by the World Health Organization (WHO) [12]. Despite these measures, nearly fifteen percent of European monitoring stations report NO_2 levels exceeding these limits, notably in urban areas and along transportation routes. The economic repercussions of air pollution, including NO_2, are substantial [2,12]. An interesting study on raising awareness of air quality issues and the necessity for developing collective air monitoring platforms can be found

* Corresponding author at: Engineering Optimization & Modeling Center, Reykjavik University, 102 Reykjavik, Iceland.
E-mail addresses: koziel@ru.is (S. Koziel), anna.dabrowska@pg.edu.pl (A. Pietrenko-Dabrowska), marwojci1@pg.edu.pl (M. Wojcikowski), bogpanki@pg.edu.pl (B. Pankiewicz).

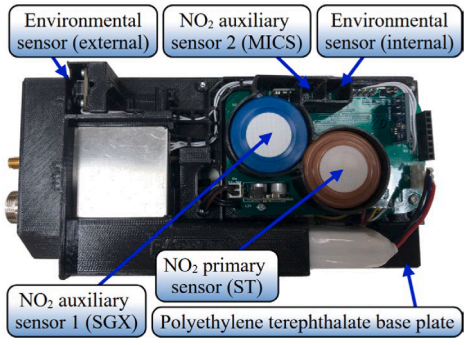
https://doi.org/10.1016/j.jestch.2024.101729
Received 11 March 2024; Received in revised form 10 May 2024; Accepted 27 May 2024
Available online 30 May 2024



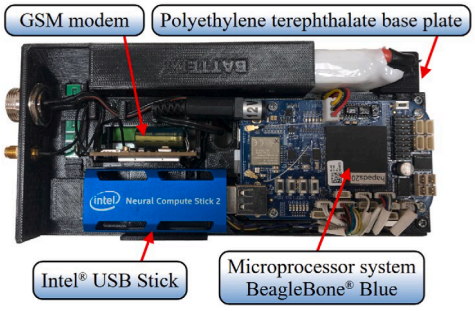
(a)

Sensor	Purpose
SGX-7NO2	NO_2 electrochemical sensor (SGX Sensortech [54]), denoted as SGX
7E4-NO2	NO_2 electrochemical sensor (SemaTech [55]), denoted as ST
MiCS 2714	Compact MOS air quality sensor (SGX Sensortech [56]), for detecting NO_2 and hydrogen, denoted as MICS
BME280	Environmental sensor (Bosch Sensortech [57]) for measuring ambient temperature, humidity and atmospheric pressure; mounted on PCB near the NO_2 sensors targeted to measure the operating conditions of the NO_2 sensors (internal)
BME280	Environmental sensor (Bosch Sensortech [57]) placed at the edge of the enclosure to measure the external parameters of air surrounding the system

(b)



(c)



(d)

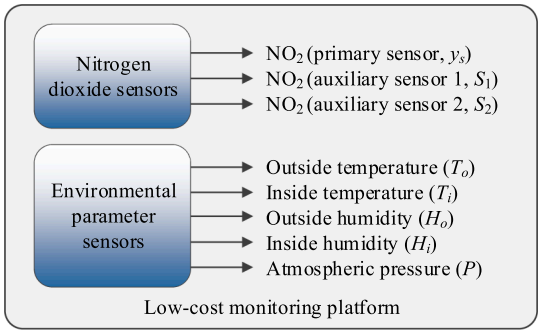
(e)

Sensor	SGX-7NO2	7E4-NO2	MiCS 2714
Measurement range	0 – 20 ppm[$]	0 – 5 ppm[$]	0.05 – 10 ppm[$]
Life expectancy	>24 months in the air	>24 months in the air	>2 years in the air
Operating temperature range	–30°C to +50°C	–40°C to +50°C	–30°C to +85°C
Operating pressure range	800 to 1200 mbar	1 ± 0.1 atm	Not provided
Operating humidity range	15% to 90% RH	15% to 90% RH	5% to 95% RH

[$]Conversion factor: at 760 mmHg and 25 degrees Celsius, 1 ppm = 1.882 mg/m^3.

(f)

Fig. 1. Autonomous air monitoring platform designed at Gdansk University of Technology, Poland: (a) block diagram, (b) included sensors, (c) internals (top view), (d) internals (bottom view), (e) the system mounted in weather-proof enclosure, (f) major specifications of the NO_2 sensors. (See above-mentioned references for further information.)



(a)

Data type		Symbol[#]
Nitrogen dioxide (NO_2) sensors	Primary sensor: ST	$y_s^{(j)}$, j = 1, …, N
	Auxiliary sensor 1: SGX	$S_1^{(j)}$, j = 1, …, N
	Auxiliary sensor 2: MICS	$S_2^{(j)}$, j = 1, …, N
Environmental parameters sensors	Outside temperature	$T_o^{(j)}$, j = 1, …, N
	Inside temperature	$T_i^{(j)}$, j = 1, …, N
	Outside humidity	$H_o^{(j)}$, j = 1, …, N
	Inside humidity	$H_i^{(j)}$, j = 1, …, N
	Atmospheric pressure	$P^{(j)}$, j = 1, …, N

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

(b)

Fig. 2. Outputs of the monitoring platform of Section 2.1: (a) NO_2 reading from the low-cost sensor under calibration (y_s). The sensor also produces auxiliary outputs: auxiliary NO_2 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; N stands for the total number of data samples obtained from the platform, further divided into training and testing sets.

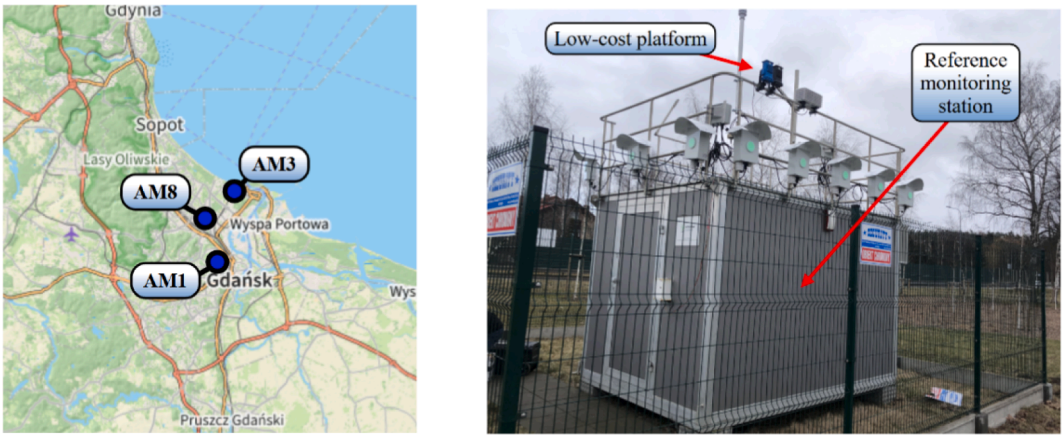


Fig. 3. Reference monitoring stations used to gather reference data: (a) station locations in the city of Gdansk, (b) a photograph of the selected station with the low-cost platform mounted in the vicinity.

Data division:

1. The overall data (N samples) is divided into the training set (N_0 samples), and the testing set (N_t samples);
2. N_t is set to approximately ten percent of the total number of samples N.

Training set:

- Reference (time series of length N_t): $\{\boldsymbol{y}_{r0.Nt}^{(j)}\}$, $j = 1, \ldots, N_0$;

1. Low-cost sensor:
   - Sensor's NO_2 measurements (time series of length N_t): $\{\boldsymbol{y}_{s0.Nt}^{(j)}\}$, $j = 1, \ldots, N_0$;
   - Auxiliary data: $\{\boldsymbol{z}_{s0}^{(j)}\}$, $j = 1, \ldots, N_0$, where $\boldsymbol{z}_{s0} = [T_{o0}^{(j)}\ T_{i0}^{(j)}\ H_{o0}^{(j)}\ H_{i0}^{(j)}\ P_0^{(j)}\ S_{10}^{(j)}\ S_{20}^{(j)}]^T$;

Testing set:

1. Reference (individual measurements): $\{y_{rt}^{(j)}\}$, $j = 1, \ldots, N_t$;
2. Low-cost sensor:
   - Sensor's NO_2 measurements (time series of length N_t): $\{\boldsymbol{y}_{st.Nt}^{(j)}\}$, $j = 1, \ldots, N_t$;
   - Auxiliary data: $\{\boldsymbol{z}_{st}^{(j)}\}$, $j = 1, \ldots, N_t$, where $\boldsymbol{z}_{st} = [T_{ot}^{(j)}\ T_{it}^{(j)}\ H_{ot}^{(j)}\ H_{it}^{(j)}\ P_t^{(j)}\ S_{1t}^{(j)}\ S_{2t}^{(j)}]^T$.

Fig. 4. Division of the reference and low-cost sensor data into training and testing set.

in Kosmidis et al. [71].

Conventional NO_2 measurement techniques rely on stationary and heavy equipment, requiring controlled environments and frequent upkeep. Commonly employed methods include photofragment chemiluminescence [14], cavity ring spectroscopy [15], long-range differential optical absorption spectroscopy [16], and laser-induced fluorescence [17]. While these techniques offer high sensitivity, some come with limitations such as unsuitability for localized monitoring [16] or the need for sophisticated hardware (e.g., a vacuum system and a pulsed laser [17]). The drawbacks associated with traditional measurement techniques have driven the development of more affordable and user-friendly alternatives. Significant efforts have been focused on creating portable platforms, particularly beneficial in urban areas characterized by uneven pollutant distribution [18–20]. However, inexpensive sensors, while cost-effective, often suffer from notable inaccuracies [21–23], owing to factors like instability [24], sensitivity to other gases [25–27], manufacturing variations [28,29], and susceptibility to environmental conditions such as temperature and humidity [30,31]. Despite these challenges, integrating low-cost sensors alongside sparsely positioned reference stations [32] can be valuable [72,73], especially in establishing comprehensive sensor networks [33,34], whether deployed on ground or aerial vehicles [35,36,74].

Enhancing the dependability of cost-efficient sensors remains crucial for reliable and portable air quality monitoring. Extensive research has been dedicated to developing effective calibration methods, broadly categorized into laboratory-based and field-based approaches [37]. While laboratory methods are theoretically more reliable, they often lack practicality as sensor performance in controlled lab settings differs from real-world conditions [21,22]. As a result, field-based methodologies are currently prevalent, relying on reference data obtained from high-performance public monitoring stations. Calibration itself commonly involves conventional regression models or more advanced machine learning techniques, often leveraging environmental factors such as temperature and humidity. For instance, in Nowack et al. [38], multivariate linear regression (MLR), support vector regression (SVR), and random forest regression (RFR) were utilized to calibrate electrochemical NO and NO_2 sensors. Another study, [39], employed ridge regression, random forest regression (RFR), Gaussian process regression (GPR), and MLR to correct low-cost NO_2 and PM_{10} sensors. Furthermore, in Jain et al. [40], MLR was used for the calibration of a chemiluminescence NO-NO_2-NO_x analyzer. Numerous other research works have explored a variety of regression models for sensor calibration, as documented in the literature, e.g., [41–46,75,76].

Recently, artificial intelligence methods, particularly neural networks (NNs) and various machine learning techniques, have garnered increased attention for achieving more accurate correction of low-cost sensors. For instance, in Han et al. [32], single linear regression (SLR), multivariate linear regression (MLR), random forest regression (RFR), and long short-term memory networks (LSTM) were employed to calibrate CO, NO_2, O_3, and SO_2 sensors. Comparative analysis highlighted

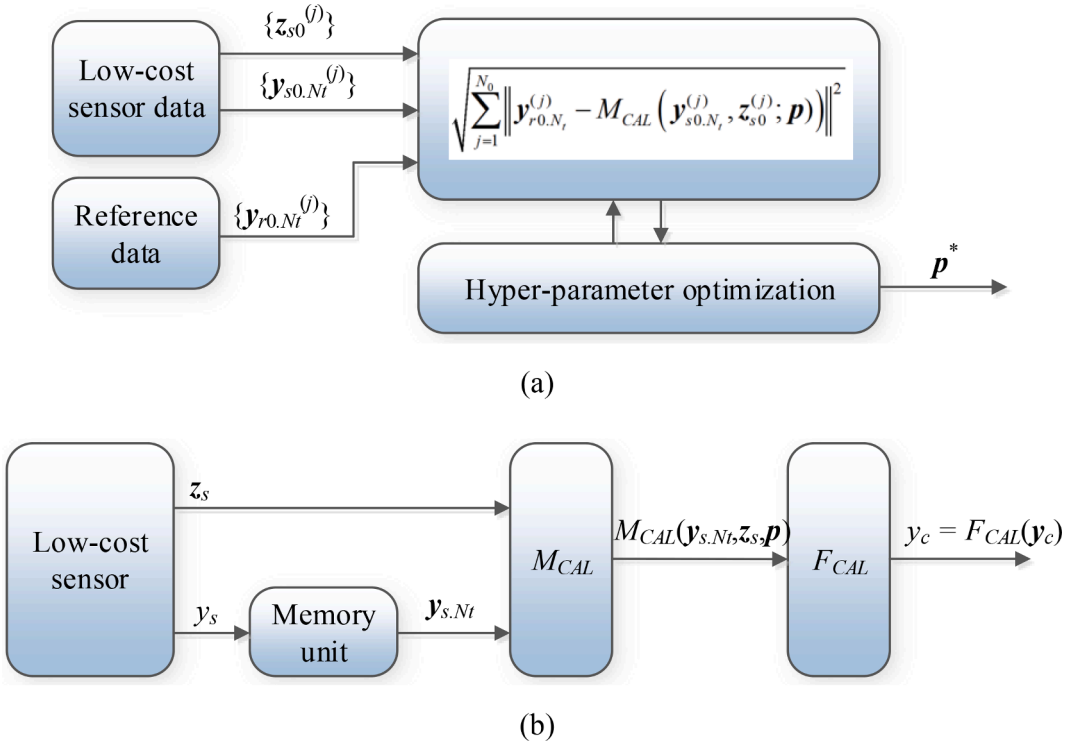


Fig. 5. Low-cost sensor calibration: (a) identification of the calibration model by solving a nonlinear regression problem that aims at matching the corrected sensor time series with the respective reference ones; (b) the overall flow of sensor calibration. Auxiliary data and sensor time series $\boldsymbol{y}_{s.Nt}$ are used to evaluate the calibration model output $M_{CAL}(\boldsymbol{y}_{s.Nt},\boldsymbol{z}_s;\boldsymbol{p})$, which is then used to compute the corrected sensor output y_c.

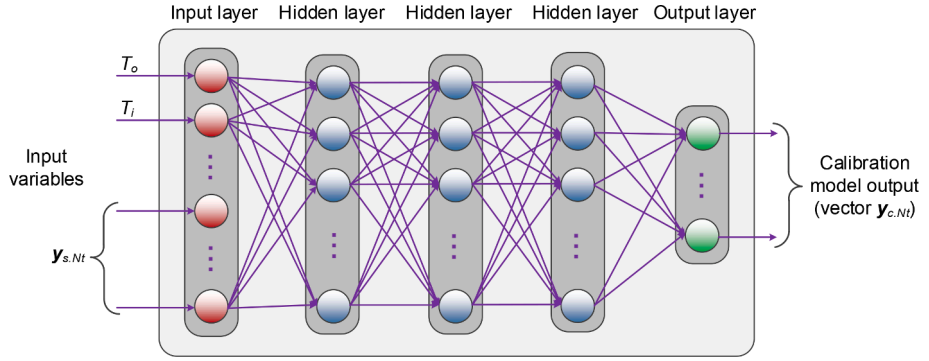


Fig. 6. ANN surrogate used as the core calibration model. Here, we employ a multi-layer perceptron (MLP) with three fully-connected hidden layers.

the superior performance of LSTM over traditional regression methods. Additionally, in Yu et al. [18], convolutional neural networks (CNNs) and recurrent neural networks (RNNs) were used to correct CO and O_3 sensors, leveraging temperature and humidity data, showcasing advantages over various regression techniques. In another study, [47], a multi-layer perceptron was employed for temporal correction of $PM_{2.5}$ sensors, demonstrating its efficiency compared to other methods. Moreover, in [48], calibration of a three-electrode NO_2 electrochemical sensor revealed the superior effectiveness of neural network-based correction over diverse regression methods. The literature contains numerous studies demonstrating the application of various NN surrogates, such as Bayesian NNs, shallow NNs, or dynamic NNs, for low-cost sensor calibration [49–52], as well as combinations of ANN with other correction mechanisms, e.g., global response correction [70]. A comprehensive study involving a comparison of different low-cost sensor calibration techniques, both regression-based (MLR, RFR, SVM) and AI-based (ANN), has been provided in Suriano and Penza [77].

Within this study, a novel technique for accurate calibration of low-cost NO_2 sensors is introduced. This method involves statistical preprocessing of sensor measurements to initially match their distributions with reference data. The primary calibration model is an artificial neural network (ANN), trained to align sensor and reference (short) time series measurements. It utilizes inputs encompassing environmental variables like temperature, humidity, and pressure, along with readings from supplementary NO_2 sensors to facilitate cross-referencing. Incorporating and matching prior NO_2 readings enables the ANN metamodel to understand typical temporal sensor changes as a function of these factors. Additionally, an interpolative kriging model serves as an auxiliary surrogate, enhancing the overall reliability of the correction process. The proposed calibration approach underwent validation by applying it to an NO_2 sensor integrated into an autonomous monitoring platform developed at Gdansk University of Technology, Poland. This platform incorporates electronic circuitry for managing air monitoring processes and facilitating wireless data transfer. Reference data used for validation was collected from high-precision public stations situated in Gdansk. The results underscore the remarkable dependability of our calibration

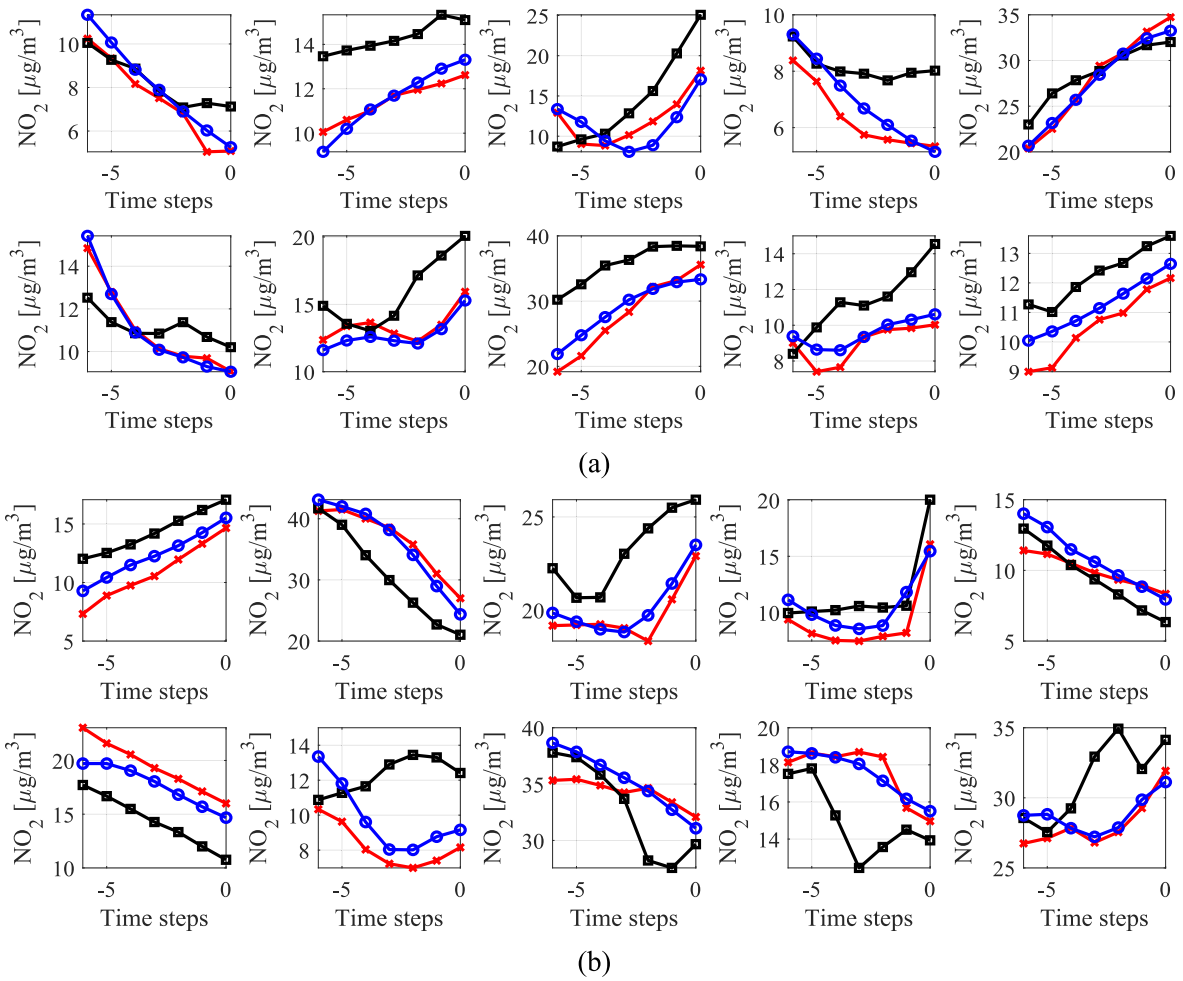


Fig. 7. Time series alignment obtained using the ANN surrogate for $N_t = 6$. Shown are the raw (uncorrected) low-cost sensor data (black), reference time series (red), and ANN-predicted sensor time series (blue). Significant improvement of the data alignment can be observed: (a) selected training points, (b) selected testing points.

technique. Achieving a correlation coefficient of 0.95 with reference data and an RMSE of 2.4 μg/m^3 across a wide range of NO_2 measurements (from zero to sixty μg/m^3) demonstrates the robustness of our method. Extensive comparative experiments highlight the significance of all correction mechanisms integrated into our framework, affirming their substantial contribution to the quality of sensor correction.

2. Hardware: NO_2 monitoring platform

The calibration methodology outlined in this study will be demonstrated through an autonomous monitoring platform developed at Gdansk University of Technology, Poland. Section 2.1 details the hardware specifications, while Section 2.2 delves into the output data generated by the sensors installed on this platform.

2.1. Hardware

The considered monitoring platform is an integrated setup equipped with various sensors to measure environmental variables (temperature, humidity, atmospheric pressure), as well as nitrogen dioxide levels using a primary sensor and two redundant sensors. Furthermore, it incorporates a GSM modem to transmit measurement data wirelessly to the cloud. Coordinating the air quality monitoring protocols are off-the-shelf components managed by the BeagleBone® Blue microprocessor system [53], hosting a 1 GHz ARM® Cortex-A8 processor, 512 MB DDR3 RAM, and 4 GB eMMC memory, running on the Linux operating system. The platform includes a rechargeable 7.4 V/4400 mA battery, which allows for sustaining operations for at least twenty hours without the necessity of using external power.

Fig. 1 displays the platform's block diagram and specifications of the employed sensor (Fig. 1(b)). Wireless data transmission is facilitated through the GSM modem, allowing the measurement data to be accessible online. The system is installed on a polyethylene terephthalate base plate, as shown in Fig. 1(c). Gas sensors (ST, SGX, MICS) are positioned in close proximity to internal environmental sensor (cf. Fig. 1(c)), ensuring the monitoring of their operational conditions. An additional environmental sensor is positioned at the edge of the device. The inclusion of auxiliary sensors is aimed at managing disparities between external and internal temperatures and humidity, largely influenced by heat produced by the electronic circuitry. Additionally, an Intel USB Stick module is integrated for potential on-board execution of calibration procedures. The entire platform is housed within a weatherproof enclosure, as depicted in Fig. 1(e).

2.2. Autonomous monitoring platform: Output data

The monitoring platform described in Section 2.1 acquires NO_2 measurements from the primary sensor and two auxiliary sensors, along with environmental parameters (temperature, humidity, atmospheric pressure). These outputs are visually depicted in Fig. 2(a). Fig. 2(b) provides the notation used across the paper. It is important to highlight that the platform captures both internal (near the NO_2 sensors) and external (at the platform's edge) environmental data. The internal and external temperature and humidity differ due to the heat produced by

The function of interest is assumed to be of the form

$$f(\boldsymbol{x}) = \boldsymbol{g}(\boldsymbol{x})^T \boldsymbol{\beta} + Z(\boldsymbol{x}) \quad \text{(K.1)}$$

where

- $\boldsymbol{g}(\boldsymbol{x}) = [g_1(\boldsymbol{x})\ \ g_2(\boldsymbol{x})\ \ \ldots\ \ g_K(\boldsymbol{x})]^T$ are known (here, EM-simulated) system responses;
- $\boldsymbol{\beta} = [\beta_1\ \beta_2\ \ldots\ \beta_K]^T$ are unknown hyperparameters;
- $Z(\boldsymbol{x})$ is a realization of a normally distributed Gaussian process with zero mean and variance σ^2.

Remark: The regression component $\boldsymbol{g}(\boldsymbol{x})^T\boldsymbol{\beta}$ serves as a trend function for f, whereas $Z(\boldsymbol{x})$ represent localized variations from the trend.

The covariance matrix of $Z(\boldsymbol{x})$ is

$$Cov[Z(\boldsymbol{x}^{(i)})Z(\boldsymbol{x}^{(j)})] = \sigma^2 \boldsymbol{R}([R(\boldsymbol{x}^{(i)}, \boldsymbol{x}^{(j)})]) \quad \text{(K.2)}$$

where $\boldsymbol{R}$ is a $N_0 \times N_0$ correlation matrix with $R_{ij} = R(\boldsymbol{x}^{(i)},\boldsymbol{x}^{(j)})$, and $R(\boldsymbol{x}^{(i)}, \boldsymbol{x}^{(j)})$ is the correlation function between data samples $\boldsymbol{x}^{(i)}$ and $\boldsymbol{x}^{(j)}$.

In this work, we use a Gaussian correlation function of the form

$$R(\boldsymbol{x}, \boldsymbol{y}) = \exp\left[-\sum_{k=1}^{n} \theta_k \,|\, x_k - y_k \,|^2\right] \quad \text{(K.3)}$$

where θ_k are the unknown correlation parameters, and x_k and y_k are the k^{th} elements of the vectors $\boldsymbol{x}$ and $\boldsymbol{y}$, respectively.

Kriging model definition:

$$s(\boldsymbol{x}) = \boldsymbol{g}(\boldsymbol{x})^T \boldsymbol{\beta} + \boldsymbol{r}^T(\boldsymbol{x})\boldsymbol{R}^{-1}(\boldsymbol{h} - \boldsymbol{G}\boldsymbol{\beta}) \quad \text{(K.4)}$$

where

$$\boldsymbol{r}(\boldsymbol{x}) = [R(\boldsymbol{x}, \boldsymbol{x}^{(1)}) \ldots R(\boldsymbol{x}, \boldsymbol{x}^{(p)})] \quad \text{(K.5)}$$

$$\boldsymbol{f} = [f(\boldsymbol{x}^{(1)})\ \ f(\boldsymbol{x}^{(2)})\ \ \ldots\ \ f(\boldsymbol{x}^{(p)})]^T \quad \text{(K.6)}$$

and $\boldsymbol{G}$ is a $p \times N$ matrix with $G_{ij} = P_j(\boldsymbol{x}^{(i)})$.

Parameter identification: Regression coefficients are found as

$$\boldsymbol{\beta} = (\boldsymbol{G}^T\boldsymbol{R}^{-1}\boldsymbol{G})^{-1}\boldsymbol{G}^T\boldsymbol{R}^{-1}\boldsymbol{f} \quad \text{(K.7)}$$

Hyperparameters θ_k are found by maximizing the expression (maximum likelihood estimation)

$$-[p\ln(\sigma^2) + \ln|\,\boldsymbol{R}\,|]/2 \quad \text{(K.8)}$$

where σ^2 and $\boldsymbol{R}$ are functions of θ_k.

Fig. 8. Surrogate modelling using kriging interpolation.

the electronic circuitry. Given their impact on sensor operation, incorporating both sets of data could bolster the reliability of the calibration process. Additionally, while the auxiliary NO_2 sensors may lack accuracy, their readings still offer indirect but valuable insights into factors influencing the primary sensor, particularly its cross-sensitivity to other gases.

3. Reference data. Public monitoring stations

The data used for calibrating the low-cost NO_2 sensor originates from high-precision public monitoring stations operated by the ARMAG Foundation [58] in Gdansk, Poland. These stations' locations are indicated in Fig. 3(a). Housed within air-conditioned containers, these stations are equipped with air monitoring instruments.

Specifically, AM1 and AM3 utilize Thermo Environmental 43C chemiluminescent NO_x analyzers, while AM8 uses API Teledyne 200E chemiluminescent NO_x analyzer. ARMAG provides open access to the monitoring data (https://armaag.gda.pl/en/). These stations conduct measurements hourly, and the data is accessible on the foundation's website for three days. To extend data collection, a custom script automates the acquisition of this information into a text file hosted on a dedicated server.

The low-cost measurement platforms were placed in close proximity to the corresponding reference stations to ensure that they work under the same environmental conditions and that the NO_2 concentration is similar. In most cases, the portable measurement platform was placed on the roof of the reference station as illustrated in Fig. 3(b). The platform was placed on the fence surrounding the reference station in one case. The low-cost platforms were placed around two meters above the ground level in all cases.

4. Calibration using ANN surrogates and time series matching

In this part of the article, we introduce the proposed cost-efficient sensor calibration methodology. The calibration problem is formulated in Section 4.1. Sections 4.2 and 4.3 elucidate an artificial neural network (ANN) surrogate and kriging interpolation being the primary and supplementary calibration models, respectively. Section 4.4 discusses statistical data pre-processing, which is an auxiliary step carried out before launching the main calibration step, and greatly improving the reference and low-cost measurement alignment. The complete correction workflow is summarized in Section 4.5.

4.1. Formulation of the calibration problem

Calibration of the low-cost sensor is based on datasets from two sources. The first one includes NO_2 readings obtained from the reference stations outlined in Section 3. The reference samples are denoted as $y_r^{(j)}$, $j = 1, \ldots, N$, where N is the total number of points. The second dataset encompasses measurements provided by the autonomous platform of Section 2 (cf. Fig. 2). The NO_2 readings from the sensor $y_s^{(j)}, j = 1, \ldots, N$, accompanied by environmental parameters and auxiliary nitrogen dioxide measurements, jointly marked as $\boldsymbol{z}_s = [T_o^{(j)}\ T_i^{(j)}\ H_o^{(j)}\ H_i^{(j)}\ P^{(j)}\ S_1^{(j)}\ S_2^{(j)}]^T, j = 1, \ldots, N$. In addition to that we also consider short time series of prior NO_2 measurements from the reference stations and the low-cost sensors, denoted as $\boldsymbol{y}_{r.Nt}$ and $\boldsymbol{y}_{s.Nt}$, respectively, where N_t stands for the time series length. More specifically, we have

$$\boldsymbol{y}_{r.N_t} = [y_r(-N_t\Delta t)\ \ y_r(-(N_t-1)\Delta t)\ \ \ldots\ \ y_r(-2\Delta t)\ \ y_r(-\Delta t)\ \ y_r(0)]^T \quad (1)$$

and

$$\boldsymbol{y}_{s.N_t} = [y_s(-N_t\Delta t)\ \ y_s(-(N_t-1)\Delta t)\ \ \ldots\ \ y_s(-2\Delta t)\ \ y_s(-\Delta t)\ \ y_s(0)]^T \quad (2)$$

where Δt is the time interval between measurements, $y_r(0)$ and $y_s(0)$ denote the most recent reading, whereas $y_r(-k\Delta t)$ and $y_s(-k\Delta t)$ denote readings taken k time intervals earlier. In particular, the vector $\boldsymbol{y}_{r.Nt}^{(j)}$ contains the reading $y_r^{(j)}$ as well as N_t readings taken before it.

Data partitioning into training and testing sets has been illustrated in

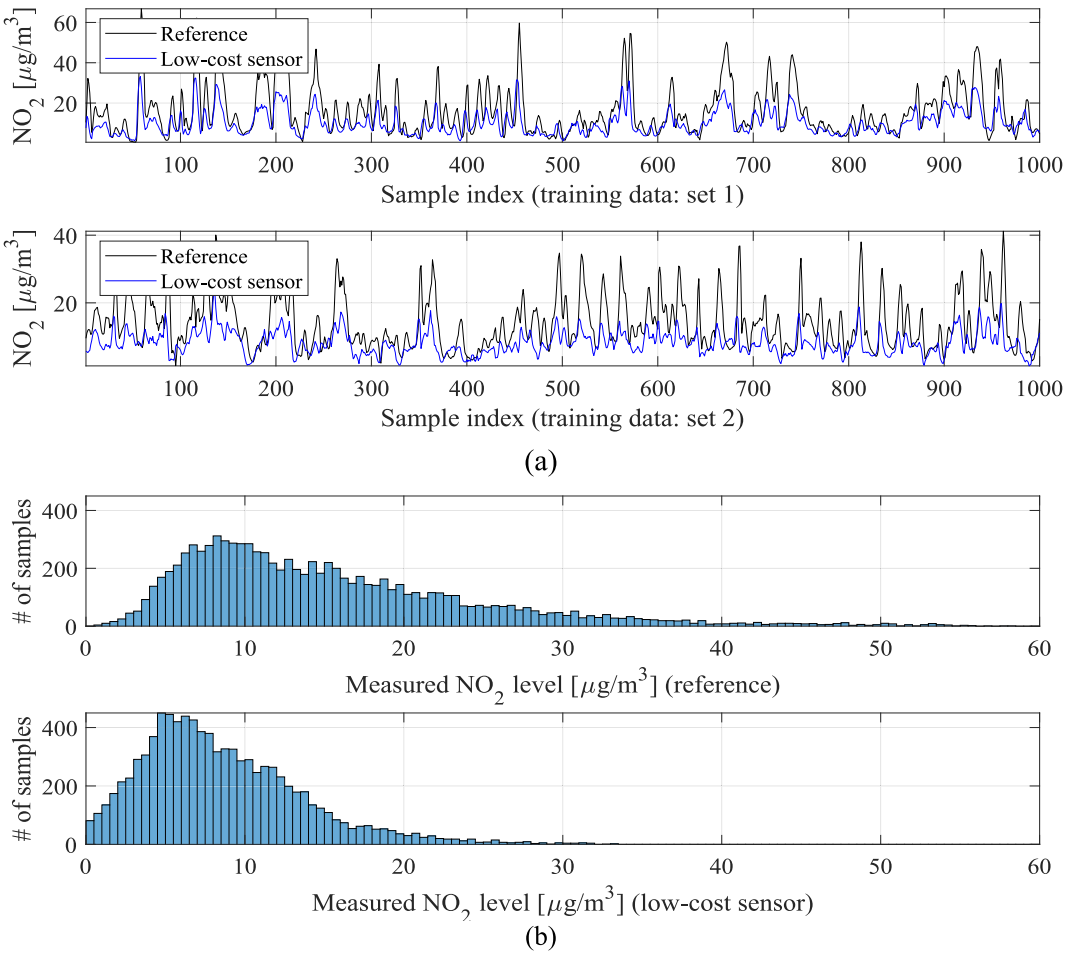


Fig. 9. Reference versus low-cost sensor data: (a) selected subsets of reference and low-cost sensor training data showing considerable disparity between typical NO_2 reading levels; (b) histograms of the reference NO_2 readings (top) and raw (uncorrected) low-cost sensor NO_2 measurements (bottom), obtained for the complete training datasets. The statistical distribution for the low-cost sensor is shifted towards lower values, which corroborates that the typical readings are lower than for the reference.

Pre-processing formulation: Low-cost sensor pre-processing is realized as a nonlinear transformation of the form

$$P(y_s, \boldsymbol{s}) = P(y_s, [s_1\ s_2\ s_3]^T) = s_1 + s_2 y_s + s_3 y_s^2 \tag{3}$$

which is applied to all sensor measurements simultaneously. A vector notation for P takes the form

$$P(\boldsymbol{y}, \boldsymbol{s}) = P\left([y_1 \ldots y_N]^T, [s_1\ s_2\ s_3]^T\right) = \left[s_1 + s_2 y_1 + s_3 y_1^2 \quad \ldots \quad s_1 + s_2 y_N + s_3 y_N^2\right]^T \tag{4}$$

Coefficient determination: Vector $\boldsymbol{s}$ is determined to improve the alignment of the smoothed histograms of the reference and low-cost sensor (cf. Fig. 10). The latter is defined as

$$H(\boldsymbol{y}) = \left[\boldsymbol{z}\ \ S(\boldsymbol{N}_y)\right] \tag{5}$$

where

$$\boldsymbol{z} = [z_1\ z_2 \ldots z_M]^T \tag{6}$$

is a vector of histogram bins (i.e., intervals splitting the horizontal axis in Fig. 8(b) into respective compartments), whereas

$$\boldsymbol{N}_y = \left[n_{y.1}\ n_{y.2} \ldots n_{y.M}\right]^T \tag{7}$$

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

$$\boldsymbol{s}^* = \arg\min_{\boldsymbol{s}} \left\| H(\boldsymbol{y}_r) - H(P(\boldsymbol{y}_s, \boldsymbol{s})) \right\| \tag{8}$$

where $\boldsymbol{y}_r$ and $\boldsymbol{y}_s$ stand for the aggregated reference and low-cost sensor NO_2 readings.
Remark: Note that if the histogram bins $\boldsymbol{z}$ are identical for the reference and the sensor (which is assumed here), the functional in (8) boils down to comparing the respective $S(\boldsymbol{N}_y)$ vectors.

Fig. 10. Statistical pre-processing of low-cost sensor data.

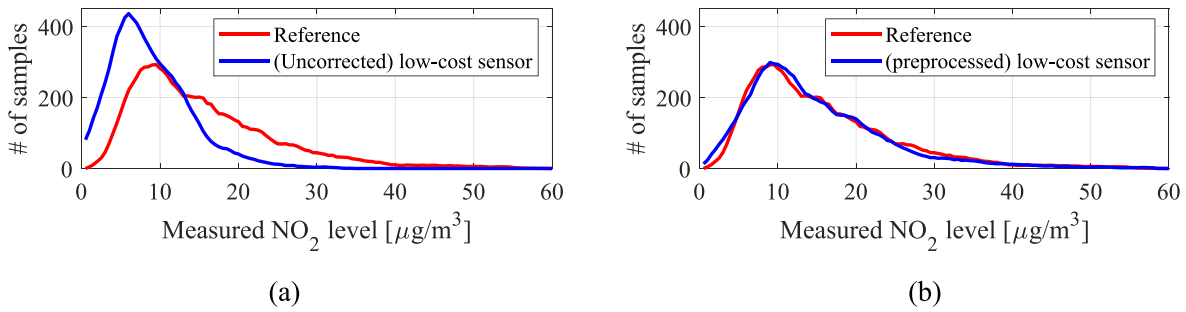


Fig. 11. Smoothened histograms: (a) reference and raw low-cost sensor, (b) reference and pre-processed low-cost sensor. Note that pre-processing aligns the measurement distributions of the low-cost sensor, thereby making ~~is~~it better prepared for further calibration.

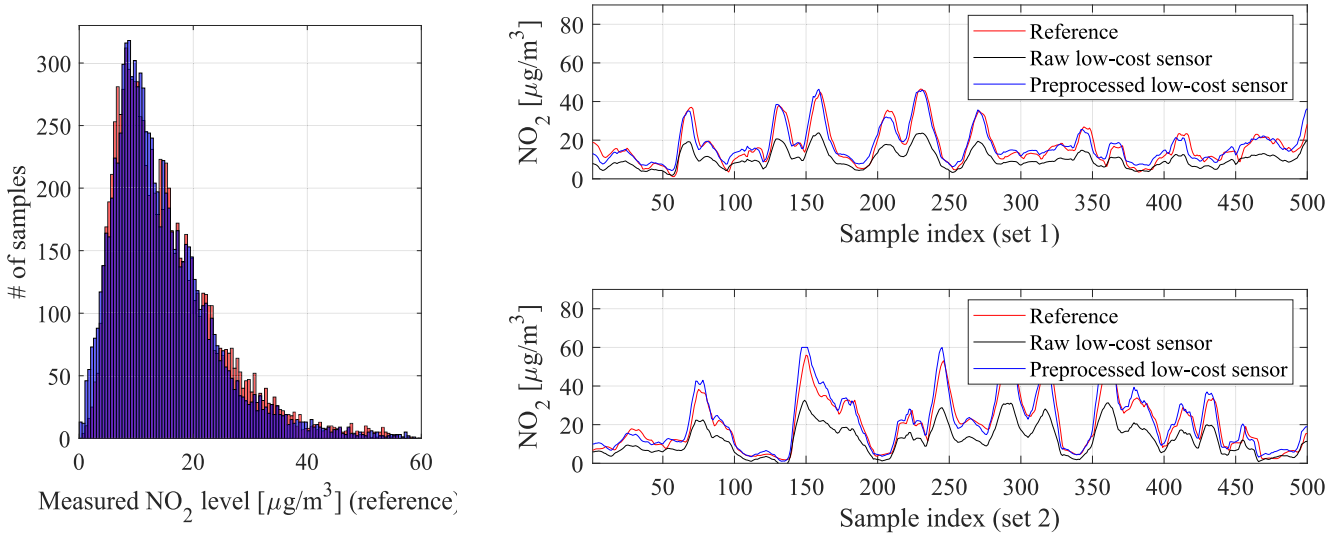


Fig. 12. The effects of statistical pre-processing: (a) comparison between the reference data (red) and pre-processed (blue) low-cost sensor histogram (overlapping data marked purple); good alignment between the two datasets can be observed; (b) improvement of NO_2 reading alignment illustrated for two selected subsets of the training data.

Fig. 4. The testing set encompasses several two-week sequences gathered at various time intervals over the five-month measurement campaign, as described in Section 5. The calibration model to be discussed in Section 4.3 is identified using the the training datasets $\{\boldsymbol{y}_{r0.Nt}^{(j)}\}$, $\{\boldsymbol{y}_{s0.Nt}^{(j)}\}$, and $\{\boldsymbol{z}_s^{(j)}\}$, $j = 1, \ldots, N_0$ (cf. Fig. 4). The calibration model output is denoted as $\boldsymbol{y}_{c.Nt} = M_{CAL}(\boldsymbol{y}_{s.Nt},\boldsymbol{z}_s;\boldsymbol{p})$, where $\boldsymbol{p}$ denotes the calibration model hyper-parameters, i.e., it returns the complete N_t-long time series, which is necessary for model identification. However, the final outcome if the calibration process is only $y_c(0)$, i.e., the prediction of the current corrected NO_2 reading. It will be denoted as $y_c = F_{CAL}(M_{CAL}(\boldsymbol{y}_{s.Nt},\boldsymbol{z}_s;\boldsymbol{p})) = F_{CAL}(\boldsymbol{y}_c)$; thus F_{CAL} simply extract the last component of the vector $\boldsymbol{y}c$ produced by M_{CAL}.Fig. 5.

Using this terminology, the calibration problem is posed as a nonlinear regression task of the form

$$\boldsymbol{p}^* = \arg\min_{\boldsymbol{p}} \sqrt{\sum_{j=1}^{N_0} \left\| \boldsymbol{y}_{r0.N_t}^{(j)} - M_{CAL}\left(\boldsymbol{y}_{s0.N_t}^{(j)}, \boldsymbol{z}_{s0}^{(j)};\boldsymbol{p}\right)\right\|^2} \quad (1)$$

In other words, the calibration model is trained to match the corrected low-cost sensor time series with those of the reference across the entire training set. The matching is understood in the least-square sense.

4.2. Primary calibration model: Artificial neural network

The primary calibration model utilized in this study employs an artificial neural network (ANN) surrogate. Although recurrent neural networks (RNNs) are commonly used for handling time series data [59], the fixed length N_t of vectors $\boldsymbol{y}_{r.Nt}$ and $\boldsymbol{y}_{s.Nt}$ in our case favours the suitability of feedforward networks. Specifically, we utilize a multi-layer perceptron (MLP) [60,61] comprising three fully connected hidden layers, each incorporating twenty neurons with a sigmoid activation function, as depicted in Fig. 6. Training of the model involves the use of a backpropagation Levenberg-Marquardt algorithm [62] (setup: 1000 learning epochs, assessment of performance via mean-square error (MSE), randomized division of training/testing data). The simplicity of the ANN's structure allows for efficient identification and faster processing. With a substantial training dataset (approximately 10,000 samples), the surrogate acts effectively as a regression model, facilitating noise reduction inherent in both reference and low-cost sensor outputs.

As elucidated in Section 4.1, the calibration model inputs are auxiliary parameters (vector $\boldsymbol{z}_s$) and the low-cost sensor time series $\boldsymbol{y}_{s.Nt}$. The output of the ANN surrogate is the predicted time series $M_{CAL}(\boldsymbol{y}_{s.Nt}, \boldsymbol{z}_s;\boldsymbol{p})$. In Section 5, we will analyse the impact on calibration efficacy by exploring different subsets of the vector $\boldsymbol{z}_s$ as inputs, aiming to illustrate how the input configuration influences the calibration process.

For the sake of illustration, Fig. 7 shows the selected time series of the raw (uncorrected) low-cost sensor, as well as the corresponding reference time series, and predictions made by the ANN calibration model. The data shown corresponds to $N_t = 6$. As indicated in the pictures, surrogate model predictions greatly improve the alignment between the reference and the low-cost sensor for the training data (Fig. 7(a)), which

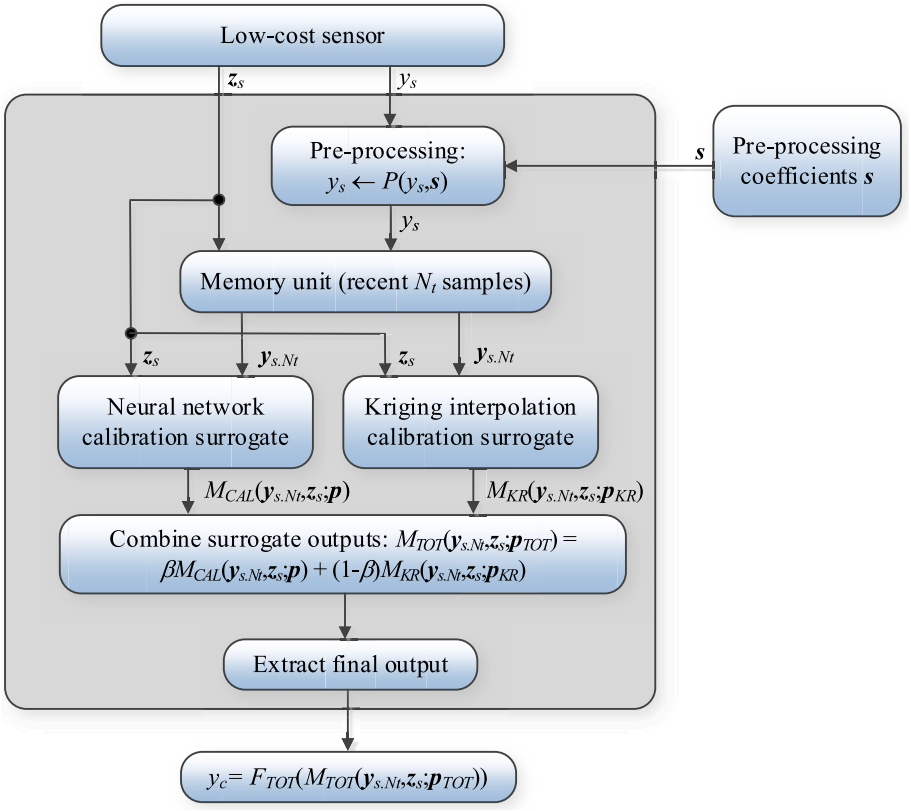


Fig. 13. Low-cost sensor calibration procedure as proposed in this study. Pre-processing of the sensor readings is followed by generating prediction of the calibrated time series M_{TOT} using the combination of ANN and kriging surrogates. The final output y_c is extracted as the last entry of the vector produced by the surrogates.

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 × 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.

Fig. 14. Details of the training and testing data gathered to calibrate the low-cost sensor of Section 2.

carries over to the testing data as well (Fig. 7(b)). This alignment, combined with other correction mechanism to be discussed later, ensures excellent performance of the calibration process.

4.3. Secondary calibration Model: Kriging interpolation

The ANN surrogate outlined in Section 4.2 is supported by a kriging interpolation model [63], utilized as a supplementary correction tool. Kriging is an immensely popular approach to behavioural modelling with countless engineering and scientific applications, e.g., [64–69]. Fig. 8 provides a brief formulation of kriging (here, using Gaussian correlation functions) assuming scalar outputs. In our case, independent models are rendered for each component of the vector $\boldsymbol{y}_{c.Nt}$. As indicated in Fig. 8, kriging encapsulates a regression part $\boldsymbol{g}(\boldsymbol{x})^T\boldsymbol{\beta}$ (typically, based on zero-, first- or second-order polynomials), and a stochastic part, which represents local departures from the regression model. The

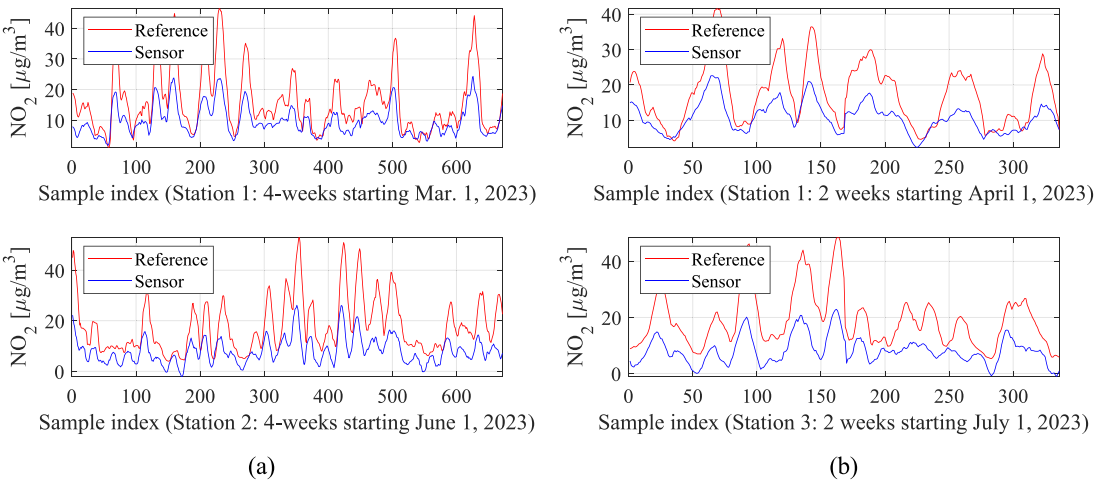


Fig. 15. Selected subsets of NO_2 readings from the reference stations and the raw (uncorrected) low-cost sensors: (a) training data, (b) testing data.

kriging hyper-parameters θ_k are obtained by means of maximum likelihood estimation.

The kriging model inputs are identical to those utilized by the ANN surrogate (cf. Section 4.2). The output is denoted as $M_{KR}(\boldsymbol{y}_{s.Nt},\boldsymbol{z}_s;\boldsymbol{p}_{KR})$. Recall that kriging is interpolative, i.e., it ensures perfect approximation of the training data. However, its generalization capability is limited. Combining the two models enables better control over the balance between approximation and generalization. In practical realization, we use a convex combination of the NN and kriging surrogates, defined as

$$M_{TOT}(\boldsymbol{y}_{s.Nt},\boldsymbol{z}_s,\boldsymbol{p}_{TOT}) = \beta M_{CAL}(\boldsymbol{y}_{s.Nt},\boldsymbol{z}_s,\boldsymbol{p}) + (1-\beta)M_{KR}(\boldsymbol{y}_{s.Nt},\boldsymbol{z}_s,\boldsymbol{p}_{KR}) \quad (2)$$

where $\boldsymbol{p}_{TOT} = [\boldsymbol{p}\ \boldsymbol{p}_{KR}]$ is the overall parameter vector. Our initial experiments indicate that the recommended setup of the convex combination coefficient, ensuring the best calibration quality, is $\beta = 0.7$.

4.4. Auxiliary correction by statistical pre-processing of low-cost sensor readings

In this study, we investigate an added correction method: statistical pre-processing of the low-cost sensor measurements. Fig. 9(a) indicates notable discrepancies in NO_2 readings between the reference and low-cost sensors, evident in the histograms depicted in Fig. 9(b). Specifically, the statistical distribution of the low-cost sensor data skews towards lower values. The aim of this introduced pre-processing scheme is to mitigate this misalignment by initially adjusting the scaling of the sensor outputs. The specifics of the procedure are outlined in Fig. 10. Fig. 11 displays the smoothed histograms before and after pre-processing, demonstrating significant improvement in alignment.

Direct comparisons between raw (non-smoothed) histograms are illustrated in Fig. 12(a), while Fig. 12(b) highlights the impact of pre-processing on selected subsets of the training data. Pre-processing serves as the initial calibration step, succeeded by surrogate-assisted correction.

4.5. NO_2 monitoring using calibrated low-cost sensor

This section encapsulates the operational flow of NO_2 monitoring utilizing the calibrated low-cost sensor. The measurement procedure incorporates the mechanisms delineated in Sections 4.2, 4.3, and 4.4. The initial phase involves pre-processing, as elucidated in Section 4.4, employing nonlinear scaling $P(y_s,\boldsymbol{s})$ derived from aligning reference and low-cost sensor histograms. Following this, the ANN and kriging surrogates predict the (local) time series of corrected model outputs. Both predictions are blended in the form of convex combination. The final result—the NO_2 sensor reading at the current time point is extracted as the last entry of the vector produced by the surrogates. A flow diagram detailing the complete process is presented in Fig. 13.

The calibration models, including ANN and kriging, were implemented in Matlab [78], the primary programming environment for conducting numerical experiments and generating the output data and visualizations. In particular, build-in Matlab's capabilities (specifically, the Deep Learning Toolbox [78]) were used. No input data normalization was employed, i.e., all calibration model inputs were presented to the respective calibration models without any modifications (e.g., normalization, etc.).

5. Results and discussion

This section showcases the implementation of the suggested calibration methodology for the low-cost NO_2 sensor integrated into the autonomous monitoring platform detailed in Section 2. The content in this section of the article is structured as follows: Section 5.1 elaborates on the reference and sensor datasets, followed by the presentation of results for various calibration model setups in Section 5.2. Section 5.3 encapsulates a summary of the study's findings.

5.1. Reference and low-cost sensor datasets

The validation of the proposed calibration procedure utilized datasets sourced from the reference stations (as detailed in Section 3) and the monitoring platforms (outlined in Section 2). The data was collected hourly spanning from March to August 2023, cf. Fig. 14. Fig. 15 showcases selected subsets of the reference and uncorrected low-cost sensor training and testing data, highlighting notable disparities between the readings from the reference and the sensor. These disparities pose a significant challenge to the calibration process.

5.2. Results

This section compiles the calibration outcomes for the low-cost NO_2 sensor integrated into the monitoring platform from Section 2. Multiple calibration setups are examined to showcase the importance of specific correction mechanisms. These include constraints on the number and nature of input parameters in the calibration model, as well as the selective activation/deactivation of the auxiliary (kriging interpolation) surrogate and statistical data pre-processing. All configurations under review are listed in Table 1. Each configuration underwent ten independent training cycles of the ANN surrogate, and the model with the best set of hyper-parameters was selected as the definitive surrogate.

The calibration setups are split into four groups. Group G1 utilizes various combinations of calibration model inputs as indicated in Table 1.

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

Calibration setup	Calibration model	Calibration input variables			Pre-processing
		Auxiliary data	NO_2 reading from primary sensor (y_s)	Time series $\boldsymbol{y}_{s.Nt}$	
G1.1	ANN	Restricted (T_o, T_i, H_o, and H_i)	×	×	×
G1.2	ANN	Restricted ($\boldsymbol{z}_s$ without P)	×	×	×
G1.3	ANN + Kriging	Restricted ($\boldsymbol{z}_s$ without P)	×	×	×
G1.4	ANN	Restricted ($\boldsymbol{z}_s$ without P)	✓	×	×
G1.5	ANN + Kriging	Restricted ($\boldsymbol{z}_s$ without P)	✓	×	×
G1.6	ANN	Complete $\boldsymbol{z}_s$	✓	×	×
G1.7	ANN + Kriging	Complete $\boldsymbol{z}_s$	✓	×	×
G1.8	ANN + global scaling	Complete $\boldsymbol{z}_s$	✓	×	×
G1.9	ANN + Kriging + global scaling	Complete $\boldsymbol{z}_s$	✓	×	×
G2.1	ANN	Complete $\boldsymbol{z}_s$	✓	✓(N_t = 1)	×
G2.2	ANN	Complete $\boldsymbol{z}_s$	✓	✓(N_t = 2)	×
G2.3	ANN	Complete $\boldsymbol{z}_s$	✓	✓(N_t = 3)	×
G2.4	ANN	Complete $\boldsymbol{z}_s$	✓	✓(N_t = 4)	×
G2.5	ANN	Complete $\boldsymbol{z}_s$	✓	✓(N_t = 5)	×
G2.6	ANN	Complete $\boldsymbol{z}_s$	✓	✓(N_t = 6)	×
G2.7	ANN	Complete $\boldsymbol{z}_s$	✓	✓(N_t = 7)	×
G2.8	ANN	Complete $\boldsymbol{z}_s$	✓	✓(N_t = 8)	×
G2.9	ANN	Complete $\boldsymbol{z}_s$	✓	✓(N_t = 9)	×
G2.10	ANN	Complete $\boldsymbol{z}_s$	✓	✓(N_t = 10)	×
G3.1	ANN	Complete $\boldsymbol{z}_s$	✓	✓(N_t = 1)	✓
G3.2	ANN	Complete $\boldsymbol{z}_s$	✓	✓(N_t = 2)	✓
G3.3	ANN	Complete $\boldsymbol{z}_s$	✓	✓(N_t = 3)	✓
G3.4	ANN	Complete $\boldsymbol{z}_s$	✓	✓(N_t = 4)	✓
G3.5	ANN	Complete $\boldsymbol{z}_s$	✓	✓(N_t = 5)	✓
G3.6	ANN	Complete $\boldsymbol{z}_s$	✓	✓(N_t = 6)	✓
G3.7	ANN	Complete $\boldsymbol{z}_s$	✓	✓(N_t = 7)	✓
G3.8	ANN	Complete $\boldsymbol{z}_s$	✓	✓(N_t = 8)	✓
G3.9	ANN	Complete $\boldsymbol{z}_s$	✓	✓(N_t = 9)	✓
G3.10	ANN	Complete $\boldsymbol{z}_s$	✓	✓(N_t = 10)	✓
G4.1	ANN + Kriging	Complete $\boldsymbol{z}_s$	✓	✓(N_t = 1)	✓
G4.2	ANN + Kriging	Complete $\boldsymbol{z}_s$	✓	✓(N_t = 2)	✓
G4.3	ANN + Kriging	Complete $\boldsymbol{z}_s$	✓	✓(N_t = 3)	✓
G4.4	ANN + Kriging	Complete $\boldsymbol{z}_s$	✓	✓(N_t = 4)	✓
G4.5	ANN + Kriging	Complete $\boldsymbol{z}_s$	✓	✓(N_t = 5)	✓
G4.6	ANN + Kriging	Complete $\boldsymbol{z}_s$	✓	✓(N_t = 6)	✓
G4.7	ANN + Kriging	Complete $\boldsymbol{z}_s$	✓	✓(N_t = 7)	✓
G4.8	ANN + Kriging	Complete $\boldsymbol{z}_s$	✓	✓(N_t = 8)	✓
G4.9	ANN + Kriging	Complete $\boldsymbol{z}_s$	✓	✓(N_t = 9)	✓
G4.10	ANN + Kriging	Complete $\boldsymbol{z}_s$	✓	✓(N_t = 10)	✓

Table 2
Sensor calibration performance: correlation coefficients and RMSE.

Calibration setup	Training data		Testing data	
	Correlation oefficient r	RMSE [μg/m^3]	Correlation coefficient r	RMSE [μg/m^3]
G1.1	0.82	4.0	0.702	5.6
G1.2	0.89	3.0	0.811	4.3
G1.3	0.95	2.2	0.819	4.4
G1.4	0.91	2.8	0.838	4.0
G1.5	0.95	2.0	0.849	3.9
G1.6	0.93	2.5	0.859	3.9
G1.7	0.96	1.8	0.861	3.8
G1.8	0.94	2.4	0.878	3.6
G1.9	0.96	1.7	0.883	3.5
G2.1	0.95	2.0	0.915	2.9
G2.2	0.96	1.8	0.928	2.8
G2.3	0.96	1.8	0.930	2.7
G2.4	0.96	1.9	0.932	2.7
G2.5	0.96	1.9	0.932	2.7
G2.6	0.96	2.0	0.934	2.7
G2.7	0.96	2.0	0.934	2.7
G2.8	0.96	2.0	0.932	2.7
G2.9	0.96	2.0	0.932	2.7
G2.10	0.96	2.0	0.930	2.7
G3.1	0.95	2.0	0.920	2.9
G3.2	0.96	1.9	0.933	2.8
G3.3	0.96	1.8	0.933	2.7
G3.4	0.96	1.9	0.936	2.6
G3.5	0.96	1.9	0.935	2.6
G3.6	0.96	1.9	0.937	2.6
G3.7	0.96	1.9	0.939	2.5
G3.8	0.96	2.0	0.937	2.6
G3.9	0.96	1.9	0.938	2.6
G3.10	0.96	1.9	0.935	2.6
G4.1	0.96	2.0	0.929	2.7
G4.2	0.97	1.6	0.939	2.5
G4.3	0.98	1.4	0.939	2.5
G4.4	0.97	1.5	0.940	2.5
G4.5	0.98	1.4	0.942	2.4
G4.6	0.98	1.5	0.945	2.4
G4.7	0.97	1.5	0.945	2.4
G4.8	0.97	1.5	0.940	2.5
G4.9	0.97	1.6	0.940	2.5
G4.10	0.97	1.6	0.938	2.7

Also, the sensor correction is arranged as affine scaling of the type $y_c = A(y_s + D)$, with the correction coefficients predicted by the calibration surrogates. The last two setups (G1.8 and G1.9) additionally utilize

Correlation coefficient:

$$r = \frac{\sum_{j=1}^{N_t}\left(y_c^{(j)} - \overline{y}_c\right)\left(y_{rt}^{(j)} - \overline{y}_{rt}\right)}{\sqrt{\sum_{j=1}^{N_t}\left(y_c^{(j)} - \overline{y}_c\right)^2 \sum_{j=1}^{N_t}\left(y_{rt}^{(j)} - \overline{y}_{rt}\right)^2}} \tag{9}$$

Root-mean-square error:

$$RMSE = \sqrt{\sum_{j=1}^{N_t} \frac{\left(y_c^{(j)} - y_{rt}^{(j)}\right)^2}{N_t}} \tag{10}$$

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

Fig. 16. Definitions of the correlation coefficient r and RMSE.

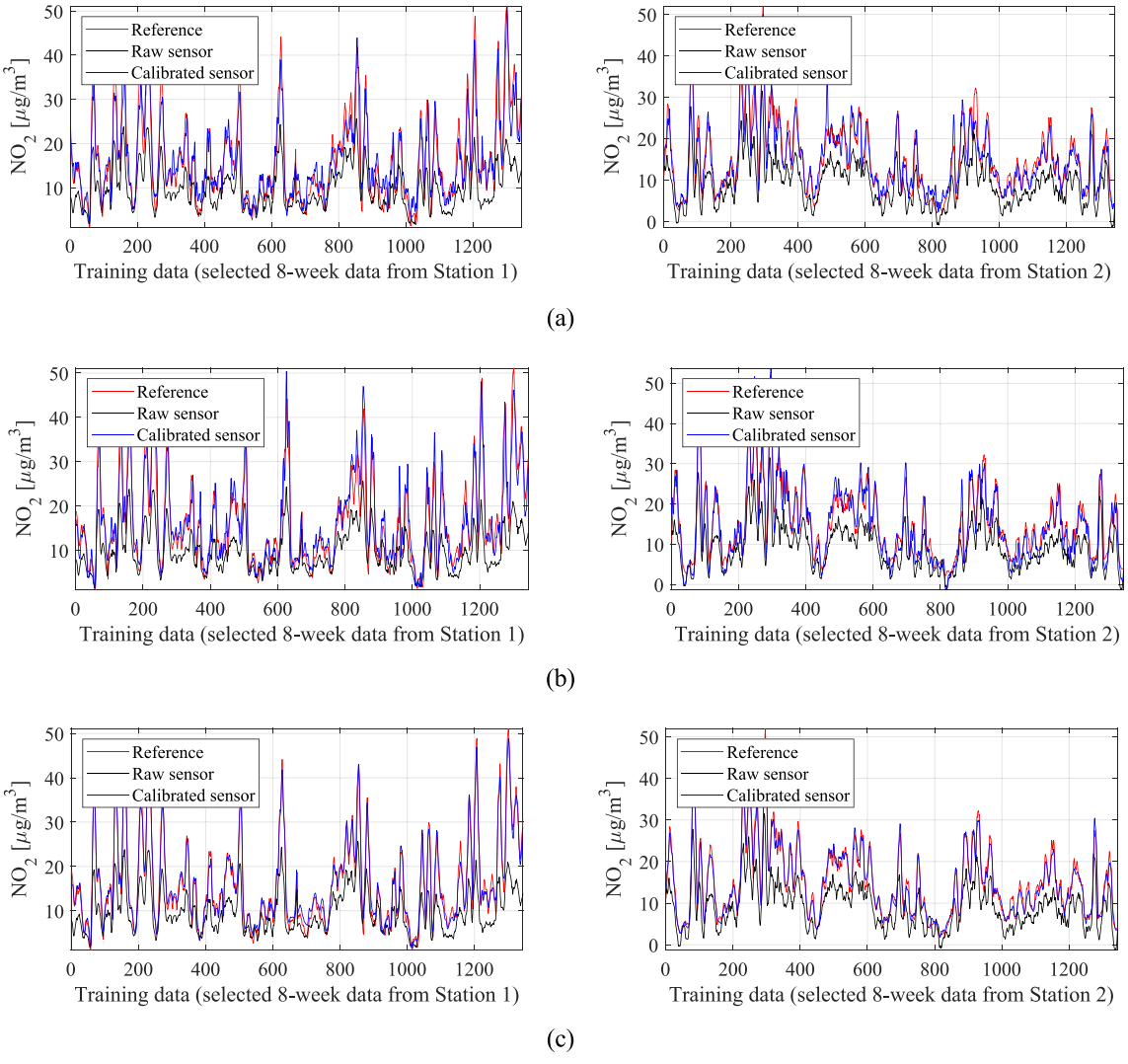


Fig. 17. Sensor calibration performance for selected subsets of the training data: (a) setup G1.2, (b) setup G1.8, (c) setup G4.6.

global data scaling of the form $y_c = A_G(y_s + D_G)$, where the coefficients A_G and D_G are established to improve the symmetry of the scatter plot for the complete training dataset. The remaining three groups of setups employ the calibration procedure described in this article, with different lengths of the time series N_t from one to ten. Group G2 only utilizes the ANN surrogate and does not employ pre-processing. Group G3 uses ANN and pre-processing, whereas group G4 employs all correction mechanisms (ANN combined with kriging, and pre-processing). Experimenting with different N_t values enables us to identify the most effective time series length.

The numerical data is available in Table 2, featuring correlation coefficients and modeling error values (RMSE) for both training and testing datasets. Definitions for these metrics can be found in Fig. 16. Visual representation of the data is presented for three specific calibration setups: G1.2, G1.8, and G4.6. Fig. 17 exhibits the reference, raw low-cost sensor, and calibrated sensor NO_2 measurements (training

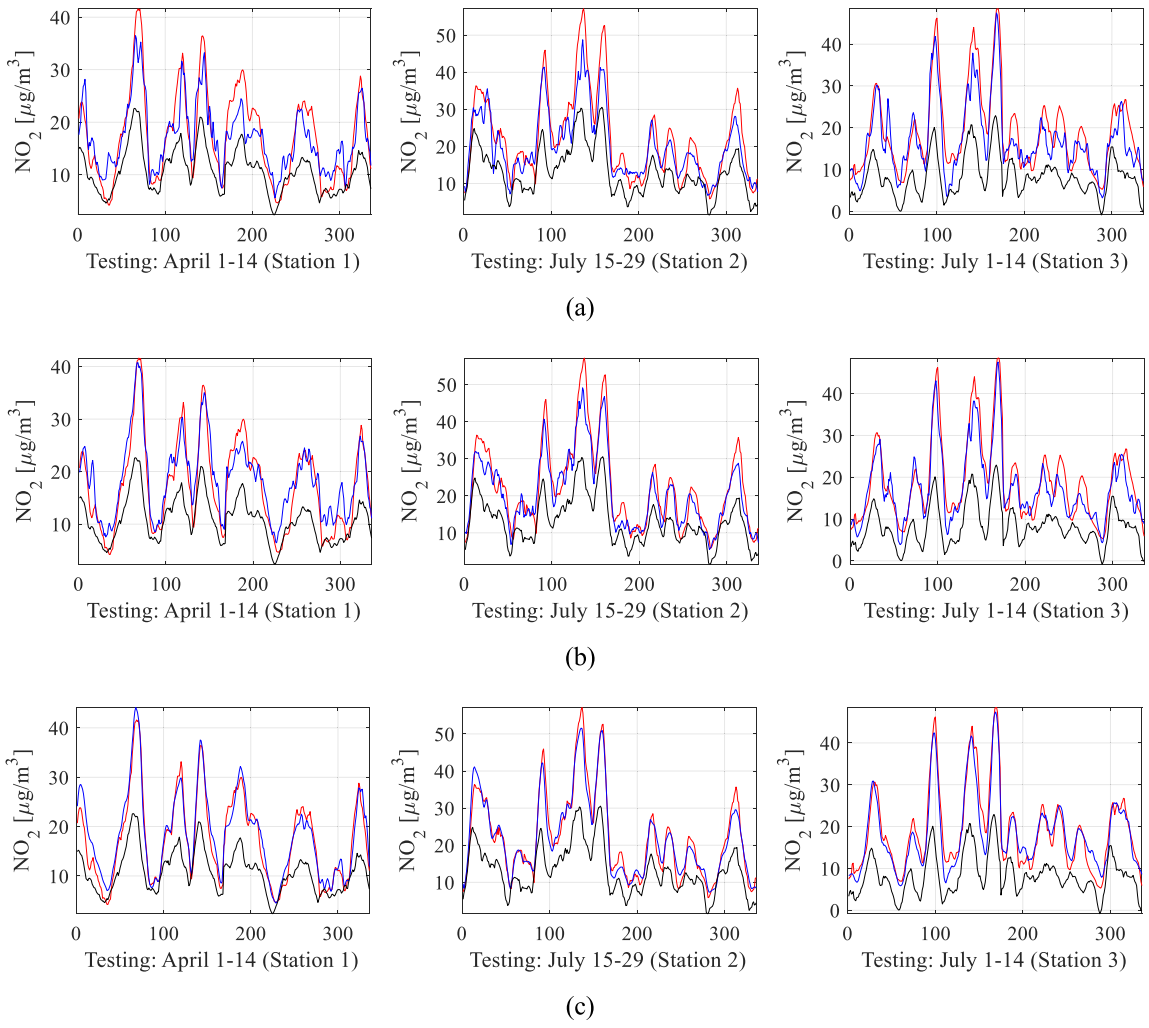


Fig. 18. Sensor calibration performance for selected subsets of the testing data: (a) setup G1.2, (b) setup G1.8, (c) setup G4.6. Reference – red, raw sensor – black, calibrated sensor – blue.

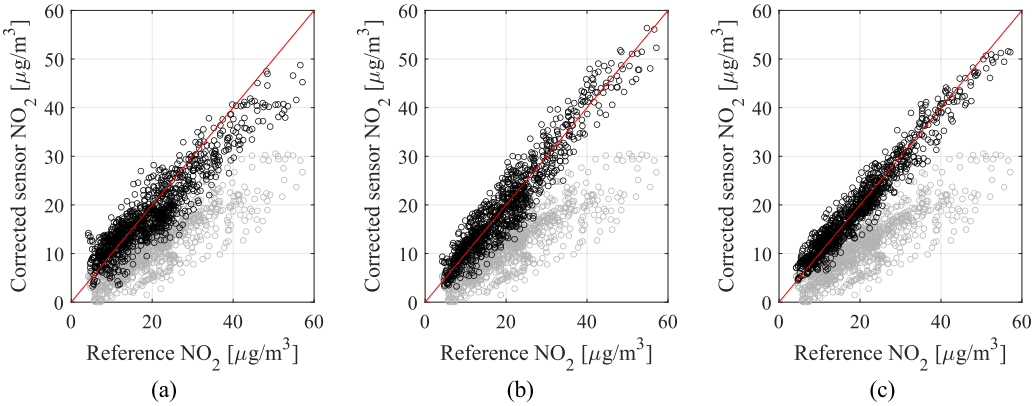


Fig. 19. Scatter plots for the testing data (uncorrected – gray, corrected – black): (a) setup G1.2, (b) setup G1.8, (c) setup G4.6.

data) for two selected eight-week periods. Fig. 18 illustrates similar information for testing data over three two-week periods, while Fig. 19 demonstrates scatter plots for the testing data.

5.3. Discussion

The experiments detailed in Section 5.2 aimed to validate the overall effectiveness of the calibration process, particularly within group G4 setups that encompass all correction mechanisms introduced in this study. Furthermore, we delved into the significance of specific components (pre-processing, time series, integration of supplementary kriging surrogate) and their impact on key performance indicators, notably the correlation coefficient and sensor error (RMSE). As the proposed calibration methodology leverages matching of time series of prior NO_2

readings from the reference and the low-cost sensor, investigating the effects of the time series length N_t was another point of interest. Given the substantial initial misalignment between the low-cost sensor and reference measurements, as well as the broad range of NO_2 levels (from almost zero to sixty μg/m^3) and their rapid changes, the calibration task proved to be highly challenging.

The outcomes summarized in Table 2 underscore the remarkable performance achieved by the calibration methodology introduced in this study. The most effective configurations, specifically G4.5, G4.6, and G4.7, encompass all correction methodologies outlined in Section 4. These setups incorporate statistical data pre-processing, a comprehensive set of input variables (including environmental parameters and readings from primary and redundant NO_2 sensors), alongside medium-length time series of prior nitrogen dioxide measurements. Notably, for configuration G4.6, the correlation coefficient nears 0.95, and RMSE stands as low as 2.4 μg/m^3 for the testing data, equivalent to a relative RMS error of approximately ten percent across the entire NO_2 level spectrum. The calibrated low-cost sensor's precision is evident through its excellent alignment with the reference data, observed in both the training (Fig. 17(c)) and testing data (Fig. 18(c)), as well as in the scatter plot (Fig. 19(c)).

The analysis of the results obtained from different calibration setups corroborates the significance of all incorporated correction mechanisms. Notably, the assessment emphasizes the pivotal role played by the testing data. Within this dataset, comparisons among various configurations within group A reveal substantial effects on both the correlation coefficient, increasing from around 0.7 to nearly 0.89, and the RMSE, declining from 5.6 to 3.5 μg/m^3. These changes occur with the inclusion of additional calibration model inputs, integration of primary sensor readings, and the use of supplementary kriging surrogates. Comparative evaluations across setups G2, G3, and G4 highlight that statistical data pre-processing consistently enhances the correlation coefficient by at least 0.02 and reduces RMSE by an average of nearly 0.8 μg/m^3. Simultaneous incorporation of the auxiliary kriging surrogate further bolsters the correlation coefficient by an additional 0.01 and decreases RMSE by 0.2 μg/m^3.

These improvements manifest vividly in Figs. 17, 18, and 19, where the transition from simpler configurations (G1.2 and G1.8) to the optimal setup (specifically, G4.6) notably enhances the alignment between the reference and corrected low-cost sensor readings. This advancement also visibly centres the scatter plots closer to the identity function, indicating a more accurate representation of the sensor data.

A distinct investigation was conducted within the configurations of group G2, G3, and G4, aimed at pinpointing the most advantageous length, N_t, for the time series of prior NO_2 measurements employed in the calibration process. The analysis of both the correlation coefficient and RMSE consistently indicates an optimal range between four and seven. For very short time series ($N_t = 1$ or 2), the benefits derived from incorporating additional data segments are less prominent. Similarly, using $N_t > 8$ marginally reduces the performance enhancements compared to N_t within the optimal range. However, this effect is less conspicuous within group G4, suggesting that when all correction mechanisms are active, the calibration process becomes less sensitive to the specific length of the time series.

The analysis conducted in this section affirms the remarkable efficacy of the proposed calibration technique. The corrected data from the low-cost sensor exhibits strong correlation with the reference data, showcasing minimal measurement error even across a broad spectrum of nitrogen dioxide levels, notably in the optimal configurations (such as G4.5 through G4.7). Implementing sensor correction can be achieved offline, serving as an intermediary step between data retrieval and presentation to the user, or can be integrated within the platform using its built-in computational capabilities.

6. Conclusion

In this research, we introduced an innovative approach to accurately calibrate affordable nitrogen dioxide sensors. Our method utilizes statistical pre-processing of sensor data to align it with reference data, incorporating two data-driven surrogate models for predicting corrected sensor NO_2 readings. The primary model, an artificial neural network (ANN), works in tandem with a supplementary kriging interpolation model. Input variables encompass environmental parameters like temperature, humidity, pressure, redundant NO_2 sensor readings, and short time series of previous NO_2 readings from the main sensor.

The proposed approach was put to the test on an independent monitoring platform developed at Gdansk University of Technology, Poland. This platform integrates primary and secondary NO_2 detectors, environmental sensors, and custom electronic systems for implementing data transmission and monitoring protocols. The validation relied on data from public monitoring stations in Gdansk collected over five months. Extensive comparative studies across various calibration model configurations highlighted the importance of all integrated correction mechanisms. Particularly, the most advanced setups, encompassing all aforementioned algorithmic components, exhibited remarkable reliability, achieving a correlation coefficient near 0.95 between reference and corrected sensor data, with a low RMSE of only 2.4 μg/m^3. This level of effectiveness reaffirms the practical feasibility of employing low-cost NO_2 monitoring, at least within the specific hardware context discussed in this study.

Future efforts will prioritize enhancing the accuracy of calibrated low-cost NO_2 monitoring. One possible approach involves integrating additional gas detectors such as SO_2, CO, and O_3 into the monitoring platform. Utilization of their readings may be used as supplementary data sources, refining the calibration model, especially concerning cross-sensitivity factors. Exploration of advanced machine learning methodologies, in particular, convolutional neural networks (CNNs) and recurrent neural networks (RNNs), is another part of the future research plan. RNNs, specialized in handling time series of varying lengths, could significantly boost monitoring reliability by leveraging such data.

CRediT authorship contribution statement

Slawomir Koziel: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. **Anna Pietrenko-Dabrowska:** Data curation, Formal analysis, Investigation, Validation, Writing – review & editing. **Marek Wojcikowski:** Funding acquisition, Investigation, Project administration, Resources, Supervision, Writing – review & editing. **Bogdan Pankiewicz:** Data curation, Software, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The research leading to these results has received funding from the Norway Grants 2014-2024 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 239858. 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.

References

[1] T.-M. Chen, W.G. Kuschner, J. Gokhale, S. Shofer, Outdoor air pollution: Nitrogen dioxide, sulfur dioxide, and carbon monoxide health Effects, Am. J. Med. Sci. 333 (4) (2007) 249–256.

[2] S. Zhao, S. Liu, Y. Sun, Y. Liu, R. Beazley, X. Hou, Assessing NO_2-related health effects by non-linear and linear methods on a national level, Sci. Total Environ. 744 (2020) 140909.

[3] C. Guerriero, L. Chatzikiakou, J. Cairns, D. Mumovic, The economic benefits of reducing the levels of nitrogen dioxide (NO_2) near primary schools: The case of London, J. Environ. Manage. 181 (2016) 615–622.

[4] F.J. Kelly, J.C. Fussell, Air pollution and airway disease, Clin. Exp. Allergy 41 (8) (2011) 1059–1071.

[5] D. Schwela, Air pollution and health in urban areas, Rev. Environ. Health 15 (1–2) (2000) 13–42.

[6] H. Salonen, T. Salthammer, L. Morawska, Human exposure to NO_2 in school and office indoor environments, Environ. Int. 130 (2019) 104887.

[7] P. Huangfu, R. Atkinson, Long-term exposure to NO_2 and O_3 and all-cause and respiratory mortality: A systematic review and meta-analysis, Environ. Int. 144 (2020) 105998.

[8] E. Samoli, E. Aga, G. Touloumi, K. Nisiotis, B. Forsberg, A. Lefranc, J. Pekkanen, B. Wojtyniak, C. Schindler, E. Niciu, R. Brunstein, M. Dodič Fikfak, J. Schwartz, K. Katsouyanni, Short-term effects of nitrogen dioxide on mortality: An analysis within the APHEA project, Eur. Resp. J. 27 (2006) 1129–1138.

[9] D.L. Mauzerall, B. Sultan, N. Kim, D.F. Bradford, NOx emissions from large point sources: Variability in ozone production, resulting health damages and economic costs, Atmos. Environ. 39 (16) (2005) 2851–2866.

[10] S. Oesch, M. Faller, Environmental effects on materials: the effect of the air pollutants SO_2, NO_2, NO and O_3 on the corrosion of copper, zinc and aluminium. A short literature survey and results of laboratory exposures, Corros. Sci. 39 (9) (1997) 1505–1530.

[11] J. Agras, D. Chapman, The Kyoto protocol, cafe standards, and gasoline taxes, Contemp. Econ. Policy 17 (3) (1999) 296–308.

[12] W.H. Organization, Air Quality Guidelines: Global Update 2005 : Particulate Matter, Ozone, Nitrogen Dioxide, and Sulfur Dioxide. World Health Organization, 2006.

[13] OECD, The Economic Consequences of Outdoor Air Pollution, OECD Publishing, Paris, 2016.

[14] M.O. Rodgers, J.D. Bradshaw, D.D. Davis, "Photofragmentation - laser induced fluorescence detection of NO2," *Topical Meeting on Spectroscopy in Support of Atmospheric Measurements (1980), paper TuP17*, Optica Publishing Group, 1980.

[15] G. Berden, R. Peeters, G. Meijer, Cavity ring-down spectroscopy: Experimental schemes and applications, Int. Rev. Phys. Chem. 19 (2010) 565–607.

[16] U. Platt, "Air monitoring by differential optical absorption spectroscopy," in *Encyclopedia of Analytical Chemistry*, John Wiley & Sons, Ltd, pp. 1–28, 2017.

[17] J. Matsumoto, J. Hirokawa, H. Akimoto, Y. Kajii, Direct measurement of NO_2 in the marine atmosphere by laser-induced fluorescence technique, Atmos. Environ. 35 (16) (2001) 2803–2814.

[18] H. Yu, Q. Li, R. Wang, Z. Chen, Y. Zhang, Y. Geng, L. Zhang, H. Cui, K. Zhang, A deep calibration method for low-cost air monitoring sensors with multilevel sequence modeling, IEEE Trans. Instrum. Meas. 69 (9) (2020) 7167–7179.

[19] J. Bi, A. Wildani, H.H. Chang, Y. Liu, Incorporating low-cost sensor measurements into high-resolution PM2.5 modeling at a large spatial scale, Environ. Sci. Technol. 54 (2020) 2152–2162.

[20] N. Castell, F.R. Dauge, P. Schneider, M. Vogt, U. Lerner, B. Fishbain, D. Broday, A. Bartonova, Can commercial low-cost sensor platforms contribute to air quality monitoring and exposure estimates? Environ. Int. 99 (2017) 293–302.

[21] L. Spinelle, M. Gerboles, M.G. Villani, M. Aleixandre, F. Bonavitacola, Field calibration of a cluster of low-cost available sensors for air quality monitoring. Part A: Ozone and nitrogen dioxide, Sens. Actuat. B-Chem. 215 (2015) 249–257.

[22] J. Fonollosa, L. Fernández, A. Gutièrrez-Gálvez, R. Huerta, S. Marco, Calibration transfer and drift counteraction in chemical sensor arrays using direct standardization, Sens. Actuat. B-Chem. 236 (2016) 1044–1053.

[23] A.C. Rai, P. Kumar, F. Pilla, A.N. Skouloudis, S.D. Sabatino, C. Ratti, A. Yasar, D. Rickerby, End-user perspective of low-cost sensors for outdoor air pollution monitoring, Sci. Total Environ. 607 (2017) 691–705.

[24] H. Kim, M. Müller, S. Henne, C. Hüglin, Long-term behavior and stability of calibration models for NO and NO_2 low-cost sensors, Atmos. Meas. Tech. 15 (2022) 2979–2992.

[25] Z. Wang, Y. Li, X. He, R. Yan, Z. Li, Y. Jiang, X. Li, Improved deep bidirectional recurrent neural network for learning the cross-sensitivity rules of gas sensor array, Sens. Actuators B: Chem. 401 (2024) 134996.

[26] N. Zimmerman, A.A. Presto, S.P.N. Kumar, J. Gu, A. Hauryliuk, E.S. Robinson, A. L. Robinson, R. Subramanian, A machine learning calibration model using random forests to improve sensor performance for lower-cost air quality monitoring, Atmos. Meas. Tech. 11 (2018) 291–313.

[27] A. Gorshkova, M. Gorshkov, N. Tripathi, K. Tukmakov, V. Podlipnov, D. Artemyev, P. Mishra, V. Pavelyev, V. Platonov, N.A. Djuzhev, Enhancement in NO_2 sensing properties of SWNTs: A detailed analysis on functionalization of SWNTs with Z-Gly-OH, J. Mater. Science: Mater. Electron. 34 (2023) 102.

[28] S. Poupry, K. Medjaher, C. Béler, Data reliability and fault diagnostic for air quality monitoring station based on low cost sensors and active redundancy, Measurement 223 (2023) 113800.

[29] M.C. Carotta, G. Martinelli, L. Crema, C. Malagù, 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 (1–3) (2001) 336–342.

[30] W. Jiao, G. Hagler, R. Williams, R. Sharpe, R. Brown, D. Garver, R. Judge, M. Caudill, J. Rickard, M. Davis, L. Weinstock, S. Zimmer-Dauphinee, K. Buckley, 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 (2016) 5281–5292.

[31] A.C. Lewis, J.D. Lee, P.M. Edwards, M.D. Shaw, M.J. Evans, S.J. Moller, K.R. Smith, J.W. Buckley, M. Ellis, S.R. Gillot, A. White, Evaluating the performance of low cost chemical sensors for air pollution research, Faraday Discuss. 189 (2016) 85–103.

[32] P. Han, H. Mei, D. Liu, N. Zeng, X. Tang, Y. Wang, Y. Pan, Calibrations of low-cost air pollution monitoring sensors for CO, NO_2, O_3, and SO_2, Sensors 21 (2021) 256.

[33] M. Müller, P. Graf, J. Meyer, A. Pentina, D. Brunner, F. Perez-Cruz, C. Hüglin, L. Emmenegger, Integration and calibration of non-dispersive infrared (NDIR) CO_2 low-cost sensors and their operation in a sensor network covering Switzerland, Atmos. Meas. Tech. 13 (2020) 3815–3834.

[34] A.A. Shusterman, V.E. Teige, A.J. Turner, C. Newman, J. Kim, R.C. Cohen, The BeErkeley atmospheric CO_2 observation network: Initial evaluation, Atmos. Chem. Phys. Discuss. 16 (2016) 13449–13463.

[35] T. Andersen, B. Scheeren, W. Peters, H. Chen, A UAV-based active AirCore system for measurements of greenhouse gases, Atmos. Meas. Tech. 11 (2018) 2683–2699.

[36] M. Kunz, J. Lavric, R. Gasche, C. Gerbig, R.H. Grant, F.-T. Koch, M. Schumacher, B. Wolf, M. Zeeman, Surface flux estimates derived from UAS-based mole fraction measurements by means of a nocturnal boundary layer budget approach, Atmos. Meas. Tech. 13 (2020) 1671–1692.

[37] A. Bigi, M. Mueller, S.K. Grange, G. Ghermandi, C. Hueglin, Performance of NO, NO_2 low cost sensors and three calibration approaches within a real world application, Atmos. Meas. Tech. 11 (2018) 3717–3735.

[38] P. Nowack, L. Konstantinovskiy, H. Gardiner, J. Cant, Machine learning calibration of low-cost NO_2 and PM10 sensors: Non-linear algorithms and their impact on site transferability, Atmosph. Meas. Tech. 14 (2021) 5637–5655.

[39] G. D'Elia, M. Ferro, P. Sommella, S. De Vito, S. Ferlito, P. D'Auria, G. di Francia, Influence of concept drift on metrological performance of low-cost NO_2 sensors, IEEE Trans. Instrum. Meas. 71 (2022) 1–11, article no. 1004811.

[40] S. Jain, A.A. Presto, N. Zimmerman, Spatial modeling of daily PM2.5, NO_2, 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) (2021) 8631–8641.

[41] M.-E. Ionascu, N. Castell, O. Boncalo, P. Schneider, M. Darie, M. Marcu, Calibration of CO, NO_2, and O_3 using Airify: A low-cost sensor cluster for air quality monitoring, Sensors 21 (2021) 7977.

[42] J. Bi, J. Stowell, E.Y.W. Seto, P.B. English, M.Z. Al-Hamdan, P.L. Kinney, F. R. Freedman, Y. Liu, Contribution of low-cost sensor measurements to the prediction of PM2.5 levels: A case study in Imperial County, California, USA, Environ. Research 180 (2020) 108810.

[43] V. van Zoest, F.B. Osei, A. Stein, G. Hoek, Calibration of low-cost NO_2 sensors in an urban air quality network, Atmos. Environ. 210 (2019) 66–75.

[44] J.A. Miech, L. Stanton, M. Gao, P. Micalizzi, J. Uebelherr, P. Herckes, M.P. Fraser, Calibration of low-cost NO_2 sensors through environmental factor correction, Toxics 9 (11) (2021) 281.

[45] B. Mijling, Q. Jiang, D. de Jonge, S. Bocconi, Field calibration of electrochemical NO_2 sensors in a citizen science context, Atmos. Meas. Techn. 11 (3) (2018) 1297–1312.

[46] R. Mawrence, S. Munniks, J. Valente, Calibration of electrochemical sensors for nitrogen dioxide gas detection using unmanned aerial vehicles, Sensors 20 (2020) 7332.

[47] M. Casari, L. Po, L. Zini, AirMLP: A multilayer perceptron neural network for temporal correction of PM2.5 values in Turin, Sensors 23 (2023) 9446.

[48] A. Panjević, T. Uzunović, B. Can Ustundag, "Development of correction models for three-electrode NO_2 electrochemical sensor," *Int. Conf. Information, Communication and Automation Technologies (ICAT)*, Sarajevo, Bosnia and Herzegovina, pp. 1-7, 2022.

[49] S. De Vito, P. Delli Veneri, E. Esposito, M. Salvato, V. Bright, R.L. Jones, O. Popoola, "Dynamic multivariate regression for on-field calibration of high speed air quality chemical multi-sensor systems," *XVIII AISEM Annual Conf.*, Trento, Italy, 2015, pp. 1-3.

[50] N. Masson, R. Piedrahita, M. Hannigan, Quantification method for electrolytic sensors in long-term monitoring of ambient air quality, Sensors 15 (2015) 27283–27302.

[51] E. Esposito, S. De Vito, M. Salvato, V. Bright, R.L. Jones, O. Popoola, Dynamic neural network architectures for on field stochastic calibration of indicative low cost air quality sensing systems, Sens. Actuators B: Chemical 231 (2016) 701–713.

[52] Z. Wang, C. Xie, B. Liu, Y. Jiang, Z. Li, H. Tai, X. Li, Self-adaptive temperature and humidity compensation based on improved deep BP neural network for NO_2 detection in complex environment, Sens. Actuators B: Chem. 362 (2022) 131812.

[53] BeagleBone® Blue, BeagleBoard, https://www.beagleboard.org/boards/beaglebone-blue.

[54] SGX-7NO2 Datasheet, Industrial Nitrogen Dioxide (NO2) Sensor', SGX Sensortech: https://www.sgxsensortech.com/content/uploads/2021/10/DS-0338-SGX-7NO2-datasheet.pdf.

[55] Four electrode NO2 sensor, SemaTech (7E4-NO2-5) (PN: 057-0400-200), SemeaTech Inc., https://www.semeatech.com/uploads/datasheet/7series/057-0400-200_EN.pdf.

[56] Datasheet MiCS-2714 1107 rev 6, SGX Sensortech, https://www.sgxsensortech.com/content/uploads/2014/08/1107_Datasheet-MiCS-2714.pdf.

[57] Humidity Sensor BME280, Bosch Sensortec: https://www.bosch-sensortec.com/products/environmental-sensors/humidity-sensors-bme280/.
[58] ARMAG Foundation: Home: https://armaag.gda.pl/en/index.htm.
[59] F.M. Salem, Recurrent Neural Networks. From Simple to Gated Architectures, Springer, New York, 2022.
[60] R. Vang-Mata (Ed.), *Multilayer perceptrons*, Nova Science Pub. Inc., 2020.
[61] S. Dlugosz, Multi-Layer Perceptron Networks for Ordinal Data Analysis, Logos Verlag, 2008.
[62] M.T. Hagan, M. Menhaj, Training feed-forward networks with the Marquardt algorithm, IEEE Trans. Neural Networks 5 (6) (1994) 989–993.
[63] A.I.J. Forrester, A.J. Keane, Recent advances in surrogate-based optimization, Prog. Aerospace Sci. 45 (2009) 50–79.
[64] A. Bingler, S. Bilicz, M. Csörnyei, Global sensitivity analysis using a kriging metamodel for EM design problems with functional outputs, IEEE Trans. Magn. 58 (9) (2022) 1–4. Art no. 7501004.
[65] M. Diago-Mosquera, A. Aragón-Zavala, L. Azpilicueta, R. Shubair, F. Falcone, A 3-D indoor analysis of path loss modeling using kriging techniques, IEEE Ant. Wireless Propag. Lett. 21 (6) (2022) 1218–1222.
[66] D. Zhan, H. Xing, A fast kriging-assisted evolutionary algorithm based on incremental learning, IEEE Trans. Evol. Comp. 25 (5) (2021) 941–955.
[67] S. Yu, Y. Li, Active learning kriging model with adaptive uniform design for time-dependent reliability analysis, IEEE Access 9 (2021) 91625–91634.
[68] A. Sinha, V. Shaikh, Solving bilevel optimization problems using kriging approximations, IEEE Trans. Cybernetics 52 (10) (2022) 10639–10654.
[69] Z. Song, H. Wang, C. He, Y. Jin, A kriging-assisted two-archive evolutionary algorithm for expensive many-objective optimization, IEEE Trans. Evol. Comp. 25 (6) (2021) 1013–1027.
[70] S. Koziel, A. Pietrenko-Dabrowska, M. Wojcikowski, B. Pankiewicz, On memory-based precise calibration of cost-efficient NO_2 sensor using artificial intelligence and global response correction, Knowl. Based Syst. 290 (2024) 111564.
[71] E. Kosmidis, P. Syropoulou, S. Tekes, P. Schneider, E. Spyromitros-Xioufis, M. Riga, P. Charitidis, A. Moumtzidou, S. Papadopoulos, S. Vrochidis, I. Kompatsiaris, I. Stavrakas, G. Hloupis, A. Loukidis, K. Kourtidis, A.K. Georgoulias, G. Alexandri, hackAIR: Towards raising awareness about air quality in europe by developing a collective online platform, ISPRS Int. J. Geo-Inf. 7 (5) (2018) 187.
[72] I. Christakis, G. Hloupis, I. Stavrakas, O. Tsakiridis, "Low cost sensor implementation and evaluation for measuring NO_2 and O_3 pollutants," *Int. Conf. Modern Circuits and Systems Technologies* (MOCAST), Bremen, Germany, pp. 1-4, 2020.
[73] I. Christakis, E. Sarri, O. Tsakiridis, I. Stavrakas, Identification of the safe variation limits for the optimization of the measurements in low-cost electrochemical air quality sensors, Electrochem 5 (2024) 1–28.
[74] I. Christakis, O. Tsakiridis, D. Kandris, I. Stavrakas, Air pollution monitoring via wireless sensor networks: the investigation and correction of the aging behavior of electrochemical gaseous pollutant sensors, Electronics 12 (2023) 1842.
[75] I. Christakis, O. Tsakiridis, K. Kandris, I. Stavrakas, Kalman filter scheme for the optimization of low-cost gas sensor measurements, Electronics 13 (2024) 25.
[76] I. Christakis, E. Sarri, O. Tsakiridis, I. Stavrakas, Investigation of LASSO regression method as a correction measurements' factor for low-cost air quality sensors, Signals 5 (2024) 60–86.
[77] D. Suriano, M. Penza, Assessment of the performance of a low-cost air quality monitor in an indoor environment through different calibration models, Atmosphere 13 (2022) 567.
[78] Matlab (R2023), The MathWorks Inc., Natick, MA, USA.