electrochem

Article

Identification of the Safe Variation Limits for the Optimization of the Measurements in Low-Cost Electrochemical Air Quality Sensors

Ioannis Christakis *, Elena Sarri, Odysseas Tsakiridis and Ilias Stavrakas

Department of Electrical and Electronics Engineering, Faculty of Engineering, University of West Attica, Thivon Av. 250, GR-12241 Athens, Greece; elena_s@uniwa.gr (E.S.); odytsak@uniwa.gr (O.T.); ilias@uniwa.gr (I.S.)
* Correspondence: jchr@uniwa.gr

Abstract: Nowadays, the study of air quality has become an increasingly prominent field of research, particularly in large urban centers, given its significant impact on human health. In many countries, government departments and research centers use official high-cost scientific instruments to monitor air quality in their regions. Meanwhile, concerned citizens interested in studying the air quality of their local areas often employ low-cost air quality sensors for monitoring purposes. The optimization and evaluation of low-cost sensors have been a field of research by many research groups. This paper presents an extensive study to identify the safe percentage change limits that low-cost electrochemical air quality sensors can have, in order to optimize their measurements. For this work, three low-cost air quality monitoring stations were used, which include an electrochemical sensor for nitrogen dioxide (NO_2) (Alphasense NO2-B43F) and an electrochemical sensor for ozone (O_3) (Alphasense OX-B431). The aim of this work is to explore the variance of the aforementioned sensors and how this variability can be used to optimize the measurements of low-cost electrochemical sensors, closer to real ones. The analysis is conducted by employing diagrams, boxplot and violin curves of the groups of sensors used, with satisfactory results.

Keywords: ozone (O_3); nitrogen dioxide (NO_2); air quality IoT; low-cost sensing systems; optimization low-cost sensor measurements

check for updates

Citation: Christakis, I.; Sarri, E.; Tsakiridis, O.; Stavrakas, I. Identification of the Safe Variation Limits for the Optimization of the Measurements in Low-Cost Electrochemical Air Quality Sensors. *Electrochem* **2024**, *5*, 1–28. https://doi.org/10.3390/electrochem5010001

Academic Editors: Michael EG Lyons and Masato Sone

Received: 1 November 2023
Revised: 11 December 2023
Accepted: 19 December 2023
Published: 21 December 2023

1. Introduction

Low-cost air pollution sensors are based on technologies that promise increasing trends in air quality monitoring, while their development in the field can provide great spatial and temporal data resolution, thus answering scientific questions and end-user applications. For this reason, research groups and government agencies interested in air quality are focusing on developing their own framework to evaluate and use low-cost sensors (LCS) [1]. The operation of electrochemical sensors is based on a chemical reaction that can be oxidation or reduction between the working electrode of a sensor and a gas target, depending on the concentration of the gas. The result of this reaction is the generation of a corresponding electrical signal [1,2]. Low-cost electrochemical gas pollutant sensors and LCS in general are affordable, small in size, portable, as well as exhibit low power consumption, compared to official instruments, which allows a larger spatial coverage of air quality monitoring of an area [2–6]. In Europe, low-cost air monitoring technologies have gained recognition and are proposed to be in the Air Quality Directive [7]. There is an extensive existing body of research in this field providing valuable insights. The research team [8] describes the role of LCS in the future of air quality monitoring. In IoT systems, a research team [9] placed an array of four electrochemical LCS and, after applying, regression methods (linear regression, multiple linear regression), and a machine learning method, presented satisfactory results. Many research team articles report on the development and applications of air quality, low-cost monitors, and their networks [10–16].

These articles report on the behavior of a set of low-cost environmental monitoring sensors, such as particulate matter [15,17] and gaseous pollutants [16,18]. In another work [19], the difference between laboratory and field experiments was investigated, and for two different types of low-cost electrochemical gas sensors for ozone and nitrogen dioxide. In other case, researchers study the data correlation between two low-cost particle concentration optical sensors in relation to reference data [20]. In another study [21] the performance, similarity, and correlation of the primary measurements (mV) of electrodes of the electrochemical gas sensor of ozone and nitrogen dioxide were presented, and an application of a common equation as a correction method focusing on the homogenous measurements was proposed. In another research work, a report on the impact of LCS on environmental conditions such as humidity and temperature [22,23] and corrective equations [24,25] have been published for each category of sensors, including temperature and humidity for the value corrections. The work [26] referred to the behavior of three particulate matter sensors, and a reference instrument in a laboratory environment, another work [27] presents a case study of a large number of LCS from different manufacturers when they were evaluated at the same time. On the other hand, a reliable approach for the evaluation of measurements from LCS is the application of statistical methods, such as linear regression and multiple linear regression, to study the reliability of sensors through their measurements. In research work [28] an aging treatable of measurements of the low-cost electrochemical sensors by applying aging factors per month was proposed; in relation to the lifetime of the sensor, this application offered excellent measurement reliability. This work involves an extensive study to determine the tolerance range of uncorrected values produced by low-cost electrochemical sensors, with the objective of optimizing their alignment with reference values. The analysis includes data from low-cost air quality monitoring stations [5] and specifically refers to low-cost electrochemical sensors for ozone (O_3) and nitrogen dioxide (NO_2). Furthermore, after the identification of the limits variation of the measurements correction, that is covering the relative difference or bias measurements of the LCS in respect to the reference measurements, a statistical linear regression (LR) and multiple linear regression (MLR) analysis for optimizing the results from the LCS is presented.

The objective of this study is to ascertain the variance range and secure percentage thresholds in low-cost electrochemical sensors for ozone and nitrogen dioxide in order to obtain values close to the actual measurements. This will facilitate the application of variation to the sensor measurements, ultimately resulting in optimized measurements when compared to the reference data.

Finally, the evaluation of the optimized values was performed by applying the Mean Absolute Deviation (MAD), Mean Square Error (MSE), Mean Absolute Percentage Error (MAPE), and Root Mean Square Error (RMSE) statistical models.

2. Materials and Methods

This work utilized low-cost air quality monitoring stations [5,28] designed and constructed at the Electronic Devices and Materials Laboratory (EDML) of the Electrical and Electronics Engineering Department of the University of West Attica. Three different types of low-cost air quality monitoring stations were constructed in relation to the communication mode and the power supply.

- wireless connectivity;
- mobile network connectivity;
- autonomous power supply.

The three categories are presented in Figures 1–3.

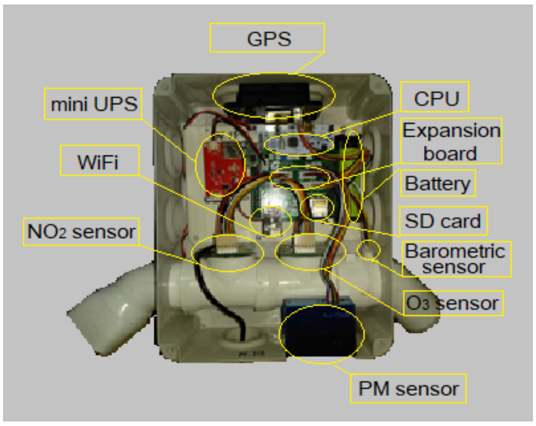


Figure 1. Wireless connectivity (Wi-Fi).

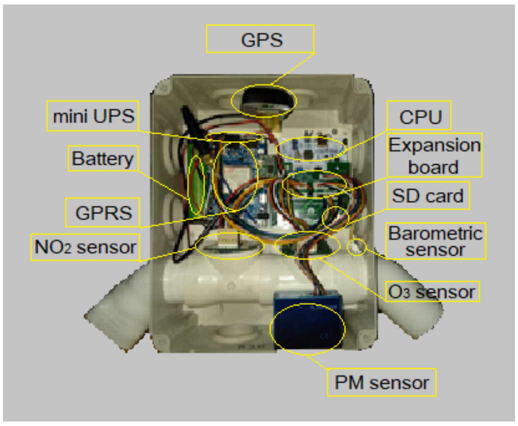


Figure 2. Mobile network connectivity (GPRS).

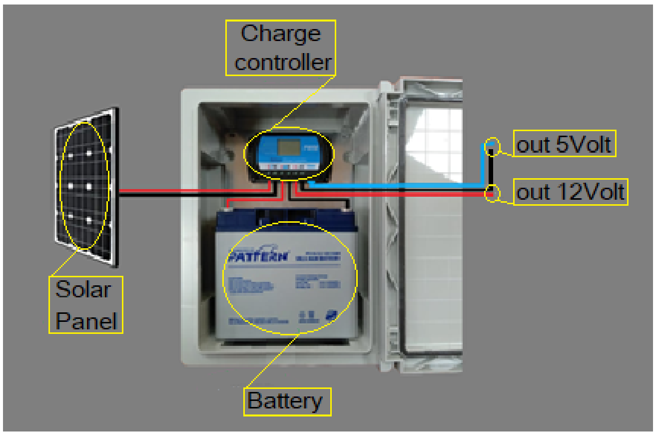


Figure 3. Autonomous power supply.

Each station consists of a durable plastic box (DEBFLEX) IP55 standard (dimensions 21 cm L × 17 cm W × 8 cm H) including a STM@Nucleof091rc processing unit which ensures high processing power and low consumption, an expansion board (Figure 4) which includes the Wi-Fi module holder, the SD card slot and the interconnections between the processing unit, peripherals and the sensors, a mini ups supported by a 18,650 battery (3.7 V/3400 mA) in case of a short length power failure, GPS unit (u-blox Neo-6) for the location and timestamp, SD card for data storage and operating conditions, and, depending on the case of communication, it is provided with a wireless network module (Wi-Fi ESP8266) or with a mobile network module (GPRS SIM-808). The station is powered by

220 V AC, through a power supply device (DC 5 V/2 A). In the case of an autonomous power supply, an accompanying device containing a solar panel (50 W), a charger device, and the battery (12 V/18 A), must be installed. Barometric conditions such as temperature, humidity, and pressure are obtained from a Bosch BME 280 barometric sensor. Plantower's sensor for the particles' concentration measurements was used. Alphasense's gas sensors were used for pollution measurements of nitrogen dioxide NO_2 and ozone O_3.

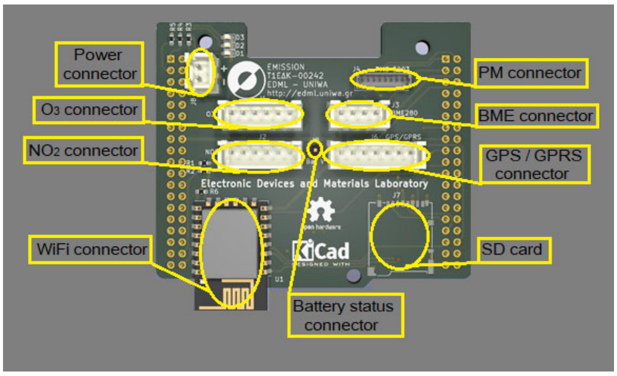


Figure 4. Expansion board.

In order to improve the reliability of the measurements [29], the gas pollutant sensors and the barometric sensor were placed inside a tube (Figure 5), facilitating the constant air flow within the tube by the fan positioned at its end to support the airflow. The particle sensor executes a similar function, however, from the construction of the sensor, it is acquired with an embedded fan, for the maintenance of a constant airflow.

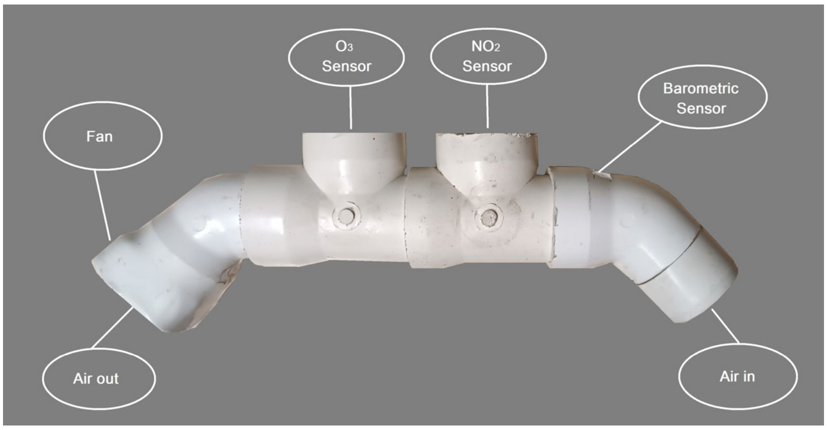


Figure 5. Sensor holder tube.

Their calibration and evaluation are presented in [5,28]. For this experiment, the three low-cost air quality monitoring stations were installed at the same place in the center of Athens, Greece. Specifically, the low-cost stations are installed on a pole on the roof of a building, one below the other, at Nymfon Hill in the Thisio area. The installation point height of LCS is 7 m. The distance between the low-cost stations is 10 cm. The reference data obtained by the monitoring instruments were installed at the facilities of the Greek Ministry of Energy and Environment [30]. The air quality department of the ministry has installed a network of stations in the Attica area. Specifically for ozone the HORIBA APOA-360 and for nitrogen dioxide, HORIBA APNA-360 devices were installed. Pollutants are continuously measured throughout the 24 h period. The response time of the automatic analyzers is one minute, while the average hourly pollution values are calculated every hour. The measurement methods of gaseous pollutants are performed for ozone with ultraviolet absorption, while for nitrogen dioxide with chemiluminescence. Field calibration of automated analyzers is performed by dynamic dilution every month and after each repair. The concentration of calibration gas for ozone is 180 ppb, while for nitrogen dioxide it is 450 ppb. The calibration duration is done with a measurement stability criterion of less than 1 ppb. The distance between the low-cost air quality monitoring stations and the reference instruments was 880 m.

2.1. Low-Cost Sensing Stations

Data were collected from low-cost air quality monitoring stations [5,28]. Each station is characterized by a unique id, and for this work, the nodes (GPRS) with ids N1, N2, N3, were used. Electrochemical sensors were used for gas sensors, the recommended O_3 gas sensor is the Alphasense OX_B431 [25] and for the NO_2 gas sensor, we used the Alphasense NO2-B43F sensor [25]. According to the work [28], the low-cost electrochemical sensor is affected by the aging of their lifetime, the measurements can be corrected by aging correction factors which are changed each month. For this reason, data for ozone (O_3), and nitrogen dioxide (NO_2) were collected from 14 April 2021 to 20 May 2021. The data measurements were taken on an hourly basis.

Electrochemical Sensor Correction

The Alphasense OX-B431 Ozone sensor and Alphasense NO2-B43F Nitrogen dioxide sensor, generally are electrochemical sensors consisting of four (4) electrodes, supported by an Individual Sensor Board (ISB), providing the output measurement in mV. In order to translate the mV recordings into actual gas concentration, a two-step procedure must be followed. During the first step, the differential voltage level between the working and auxiliary electrodes must be used with respect to the environmental temperature and the sensor sensitivity. This work is conducted following the proposed Equation (1) [31] that includes the electrodes' voltage, zero voltage calibration from ISB, and temperature:

$$WE_c = (WE_u - WE_e) - n_T \cdot (AE_u - AE_e) \quad (1)$$

where WE_c is the corrected working electrode value, WE_u is the working electrode reading value, AE_u is the auxiliary electrode reading value, n_T is the slope value of temperature T, according to the Alphasence Application note (AAN-803) [31], WEe is the working electrode electronic zero value, AEe is the auxiliary electrode electronic zero value.

Sequentially, according to [28], the final setup of the calibration of the calculated concentrations are leveled up and scaled by two factors (C1, C2) that are estimated after the calibration period when the stations are placed either in a controlled environment or near official and high-cost instruments. The formula that is used to get the final corrected values is in the following general form:

$$Gas_{concentration} = [(WE_c/Sensor_{sensitivity}) + C1]/C2 \quad (2)$$

where $Sensor_{sensitivity}$ is provided by the manufacturer to transform voltage into gas concentration in ppb, C1 is the level up factor, and C2 is the scaling factor.

2.2. Methodology

The data analysis is presented in four sections. In the first section, a correlation coefficient (R^2) of the measurable values of the electrochemical low-cost ozone and nitrogen dioxide sensors in relation to the reference values is calculated. In the second section, the average and median method was applied to investigate the percentage change between the LCS values and reference values. In the third section, the average percentage change, using both the average and the median methods, was applied to the LCS in order to optimize the collected measurements, in relation to the reference measurements. In the fourth section, we perform a statistical analysis, focusing on the application of statistical models to the presented data. The use of the statistical models of linear regression (LR) and multiple linear regression (MLR) was considered as a suitable approach since they both provide the required results and at the same time constitute simple and practical solutions. In addition, the MAD, MSE, MAPE, and RMSE methods were used to confirm the evaluation and the correctness of the results.

2.2.1. Coefficient of Determination R^2

In statistics, the coefficient of determination, denoted as R^2 or R squared, shows the percentage variation of the dependent variable (x) which is explained by all the independent variables (y); it shows the relationship degree between two variables x and y. The higher the degree the better; the values of the coefficient of determination degree are always between 0 and 1.

2.2.2. Linear Regression and Multiple Linear Regression

Linear regression and multiple linear regression methods can be used for sensor calibration [32]. The following Equation (3) describes the linear regression and the general Equation (4) describes the multiple linear regression.

$$Y = \alpha + \beta X \tag{3}$$

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_k X_k + \varepsilon \tag{4}$$

3. Results and Discussion

In this section, the results of the measurements' correlation between the three low-cost air quality monitoring stations in relation to the official data are presented. The report of the results is divided into two parts. In the first part, the results are presented according to the evaluation method of the LCS, while in the second part, there is a study of statistical analysis of the data from the LCS in relation to the reference data.

3.1. Evaluation Method

For the experimental setup, three sets of electrochemical sensors, ozone (O_3), and nitrogen dioxide (NO_2), were installed at the center of Athens, Greece. The sensor evaluation took place in an earlier time period when the sensor was placed next to official instruments, based on the method as described at work [5]. The data analysis and results were processed using the MATLAB environment. The measurements of LCS are presented as non-corrected, where, in their extraction, firstly Equation (1) is applied for the WEc calculation, and secondly Equation (2) for the concentration calculation.

3.1.1. Gases Evaluation

In this section, the behavior of the non-corrected measurements of the low-cost sensors is presented with respect to the reference measurements of the concentration of each gaseous pollutant, in hourly base time according to the ISO 9169:2006 [33]. The NO_2 time series of three LCS measurements (non-corrected) and the reference measurements are shown in Figure 6 and the correlation coefficient (R^2) of each sensor's measurements (non-corrected) with the reference measurements is shown by the scatter plots in Figure 7.

From Figure 6 it is evident that the time series of nitrogen dioxide concentration measurements, both of the LCS and of the reference, do not behave well. Specifically, it should be mentioned that the time series of the concentration measurements among the LCS are consistent, while, in relation to the reports, they do not show a good response. Figure 7 shows the behavior of the (non-corrected) nitrogen dioxide concentration measurement values, from where a large dispersion of the measurements and a relatively small degree of correlation are observed (the smallest degree of correlation is 0.33 and the largest degree of correlation is 0.42, while the average degree of correlation is 0.39), it should be it is reported that, according to the literature, the degree of correlation for electrochemical nitrogen dioxide sensors ranges in these values.

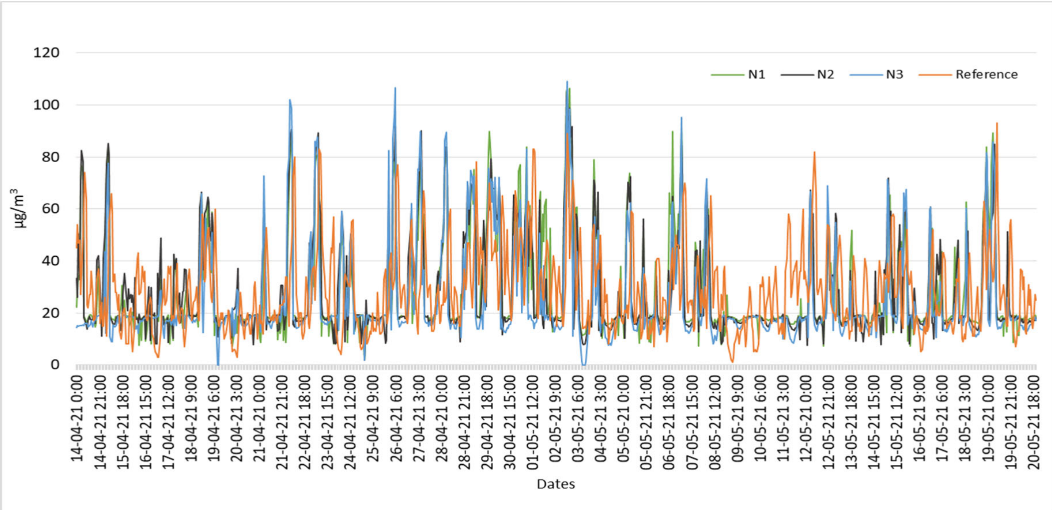
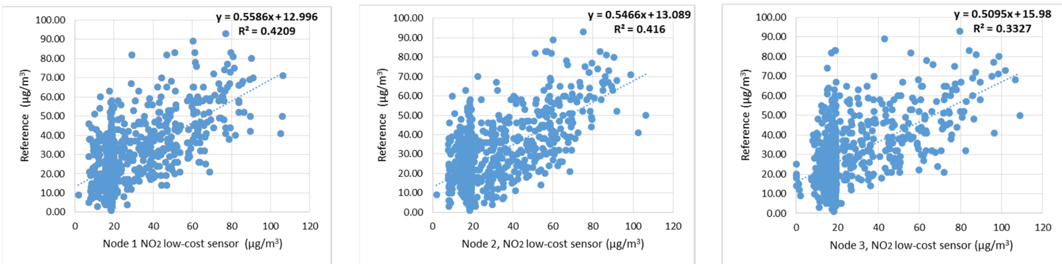


Figure 6. NO_2 time series of three LCS and reference.



(**a**) (**b**) (**c**)

Figure 7. Scatter plot of values of each LCS NO_2 (N1, N2, N3) and reference. (**a**) Scatter plot of measurements of Node 1 NO_2 LCS and reference, (**b**) Scatter plot of measurements of Node 2 NO_2 LCS and reference, (**c**) Scatter plot of measurements of Node 3 NO_2 LCS and reference.

The O_3 time series of three LCS measurements (non-corrected) and the reference measurements are shown in Figure 8 while the correlation coefficient (R^2) of each sensor measurements (non-corrected) with reference measurements are shown by the scatter plots in Figure 9.

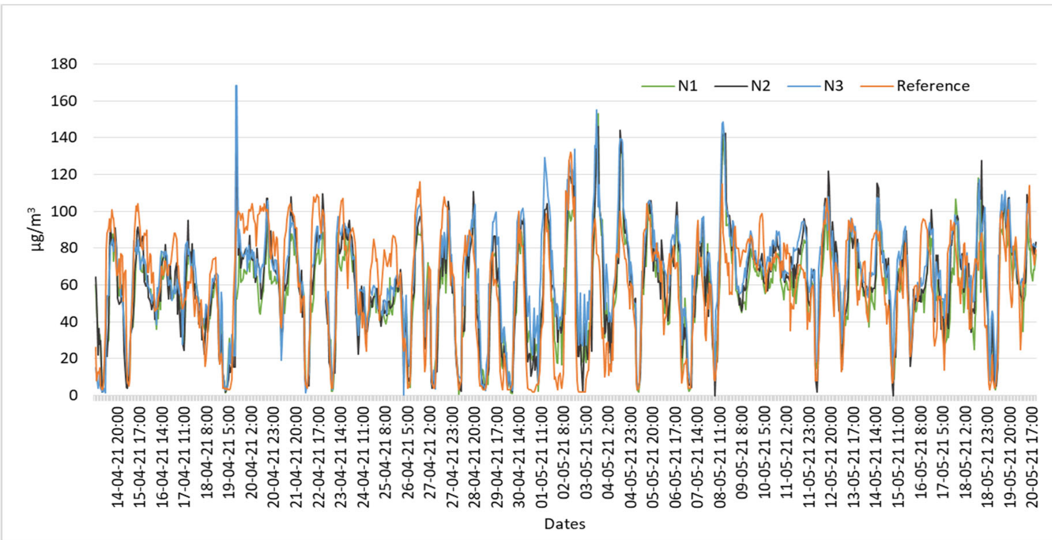
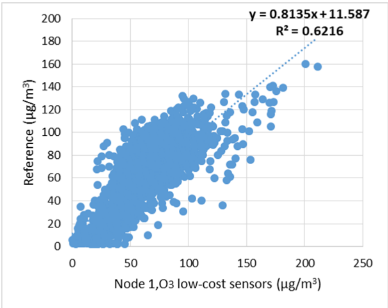
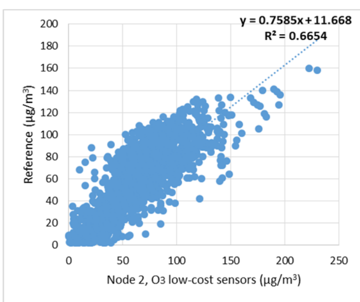


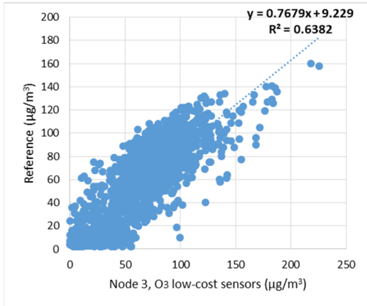
Figure 8. O_3 time series of three LCS and reference.



(**a**)



(**b**)



(**c**)

Figure 9. Scatter plot of values of each LCS O_3 (N1, N2, N3) and reference. (**a**) Scatter plot of measurements of Node 1 O_3 LCS and reference, (**b**) Scatter plot of measurements of Node 2 O_3 LCS and reference, (**c**) Scatter plot of measurements of Node 3 O_3 LCS and reference.

From Figure 8 it is evident that the time series of the ozone concentration measurements, both of the LCS and of the reference, follow each other. Figure 9 shows the behavior of the (not corrected) ozone concentration measurement values, in this case, a smaller dispersion of the measurements and a relatively high degree of correlation is observed (the smallest degree of correlation is 0.62 and the largest degree of correlation is 0.66, while the average degree of correlation is 0.64), in accordance with the literature, the degree of correlation for electrochemical ozone sensors ranges in these values.

3.1.2. Average and Median Methods' Percentage Change of Gases

The average values of concentration and the median values of concentration were extracted from the measurements from three low-cost nitrogen dioxide (NO_2) sensors, then the degree of correlations of the average values with the reference values and the median values with the reference values for nitrogen dioxide gas pollutant concentration were investigated.

In general, the average formula and median formula are presented in Equations (5) and (6), respectively. In the median formula, the equation changes according to the number of observations, if it is odd or even.

$$Average = \frac{\sum_1^n x}{n} \tag{5}$$

$$(\text{ODD})\ Median = x\left(\frac{n+1}{2}\right) \tag{6}$$

$$(\text{EVEN})\ Median = \frac{x_{(n/2)} + x_{((n/2)+1)}}{2}$$

Where x is the ordered list of values in the data set, n is the number of values of the data set.

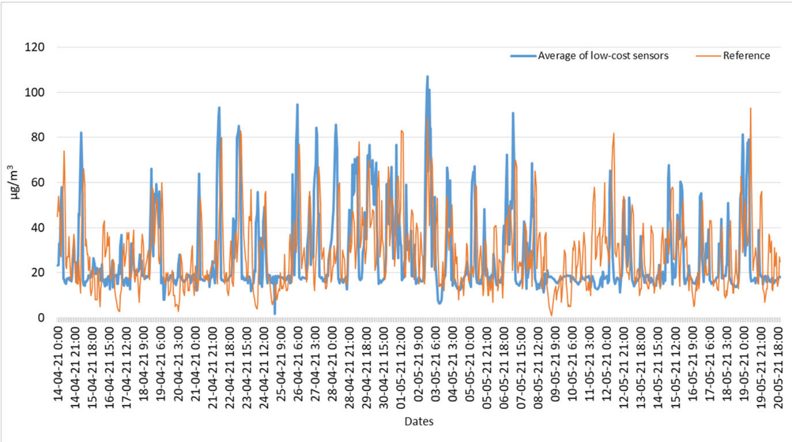
In our experiment, three (n = 3) low-cost air quality monitoring stations were used. The measurements of each sensor are: Ni, with the values were arranged in ascending order, for this work the equations of the average and median formulas at same time point (t) are presented in Equations (7) and (8), respectively.

$$Average_{(t)} = \frac{\sum_{i=1}^3 N_i(t)}{3} \tag{7}$$

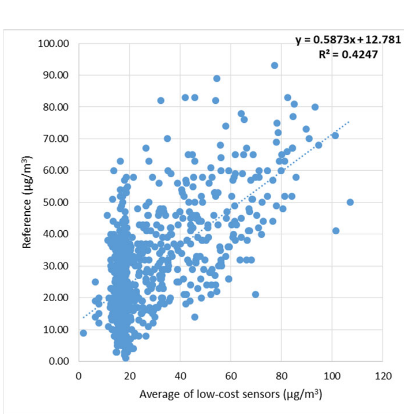
$$Median_{(t)} = N_i(\text{t})\ \text{2th term} \tag{8}$$

According to Equations (7) and (8), the measurements of three LCS, of each gas, are taken every minute, sixty times an hour, where the average hourly measurement is calculated. For each gas, the average hourly measurements from the LCS were used to calculate both the average value and median value. This procedure yields a series of average values and a series of median values that are used next to determine the variance of the LCS readings relative to the reference readings. Then, the percentage of deviation of each value, both of the average and the median, were calculated with respect to the reference values for each time point. From this procedure, the common percentage average value of the average method and the common percentage median value of the median method are extracted, where they were applied as a correction factor to the measurements of the LCS, of the same time point, and the results are shown below.

Figure 10 shows the average values of three low-cost nitrogen dioxide sensors, Figure 10a shows the time series of the output of the average and reference values of nitrogen dioxide, and the scatter plot Figure 10b shows the correlation between the average and reference values of nitrogen dioxide. Figure 11 shows the median values of low-cost nitrogen dioxide sensors. Figure 11a shows the time series of the output of the median and reference values of nitrogen dioxide, and the scatter plot Figure 11b shows the correlation between the median and reference values of nitrogen dioxide.

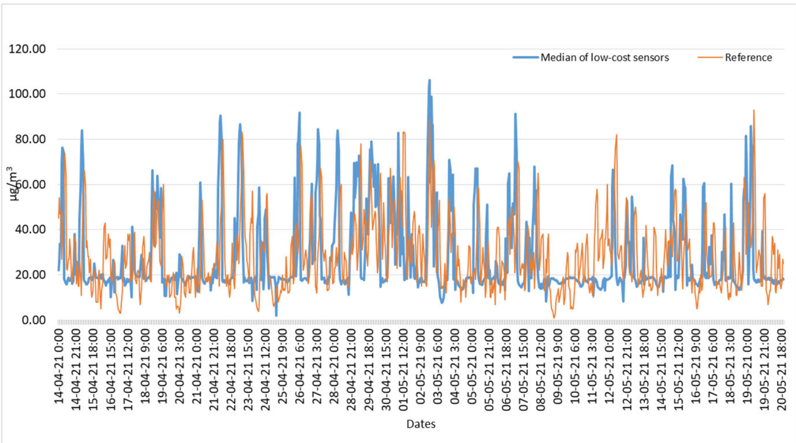


(**a**)

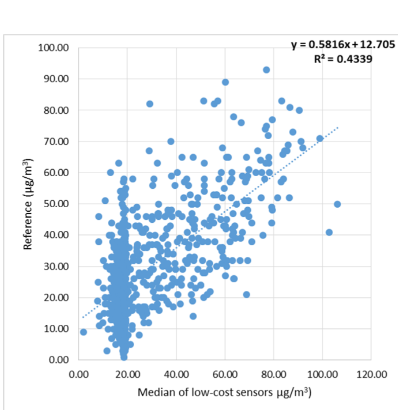


(**b**)

Figure 10. Performance of average values of three low-cost nitrogen dioxide sensors and the reference values. (**a**) Time series of the average values and reference values of nitrogen dioxide concentration, (**b**) Scatter plot between the average and reference values of nitrogen dioxide concentration.



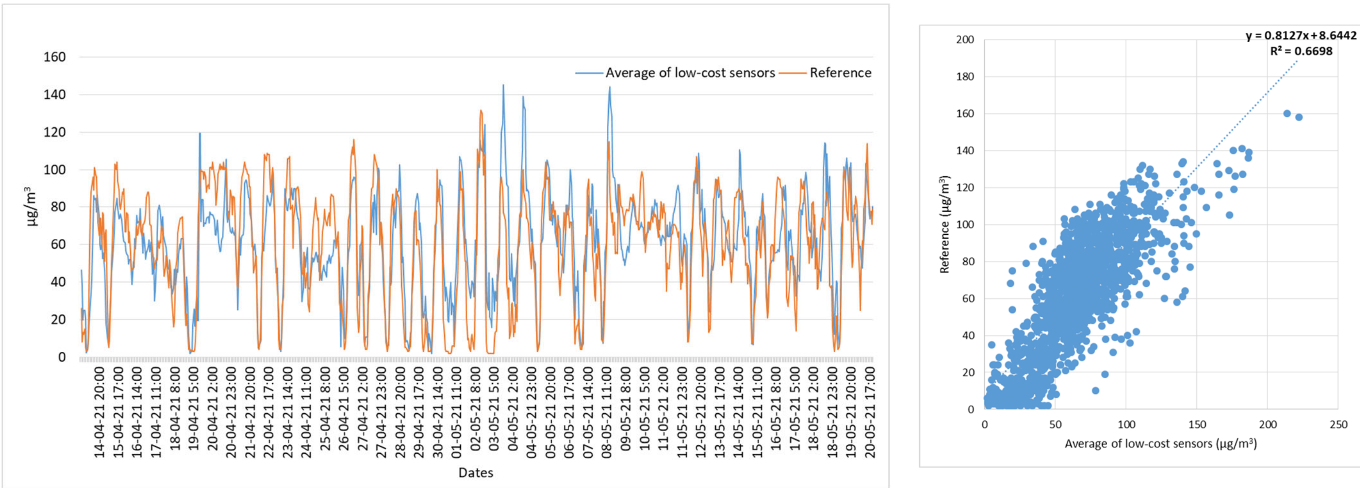
(**a**)



(**b**)

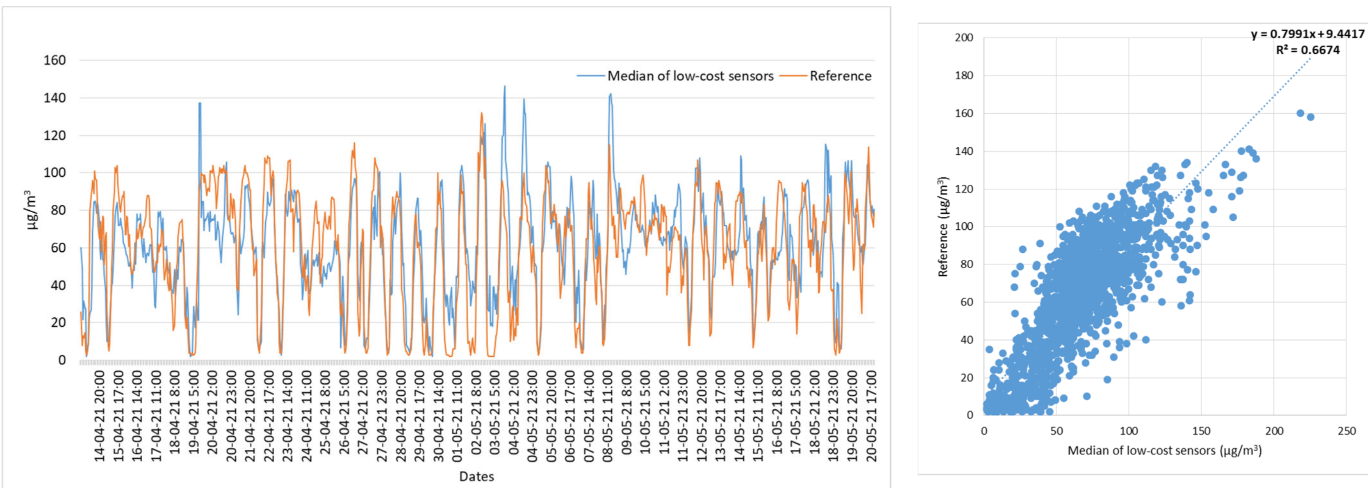
Figure 11. Performance of median values of three low-cost nitrogen dioxide sensors and the reference values. (**a**) Time series of the median values and reference values of nitrogen dioxide concentration, (**b**) Scatter plot between the median and reference values of nitrogen dioxide concentration.

Accordingly, the average values and the median values were extracted from the measurements from three low-cost ozone sensors. Subsequently, an examination of the degree of correlation between the average and median values and the reference values for ozone gas was conducted. Figure 12 shows the performance of the average values of three low-cost ozone sensors. Figure 12a shows the time series of the output of the average and reference values of ozone. Figure 12b shows the correlation between the average and reference values of ozone. Figure 13 shows the median values of low-cost ozone sensors. Figure 13a shows the time series of the median values and reference values of ozone. Figure 13b shows the correlation between the median and reference values of ozone.



(**a**) (**b**)

Figure 12. Performance of average values of three low-cost ozone sensors and the reference values. (**a**) Time series of the average values and reference values of ozone concentration, (**b**) Correlation (R^2) between the average and reference values of ozone concentration.



(**a**) (**b**)

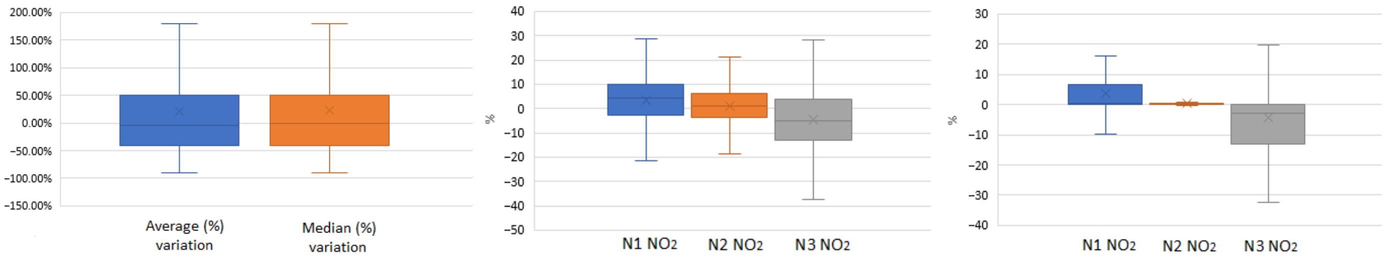
Figure 13. Performance of median values of three low-cost ozone sensors and the reference values. (**a**) Time series of the median values and reference values of ozone concentration, (**b**) Correlation (R^2) between the median and reference values of ozone concentration.

3.1.3. Percentage Change of LCS Gases Measurements

For the research purpose, an extended study took place in relation to the percentage change of the collected measurements (non-corrected) of the LCS in respect to reference measurements, for each gas. Percentage change (%) is defined as the difference between LCS measurements and the reference measurements normalized by the reference value.

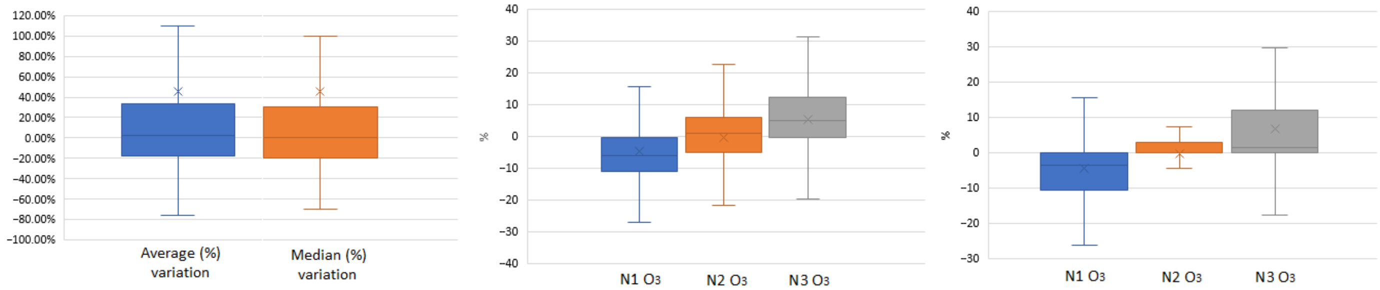
This procedure involves the comparison of the set of average values and non-corrected values, as well as the set of median values and non-corrected values for each sensor. This results in the variation of each sensor in relation to the average values, and the variation of each sensor in relation to the median values.

A boxplot (box and whisker plot) displays the data distribution based on a five-number summary (minimum, first quartile, median, third quartile, and maximum). It identifies extremal points in the figure and their corresponding values. It can also show if values are symmetrical, how tightly the values are grouped, and how skewed they are. Figures 14 and 15 present the boxplots of variation for the three low-cost NO_2 sensors values and three low-cost O_3 sensors in relation to their average and median values. Figure 14a presents details of the total variation, including the values of all three low-cost NO_2 sensors, both the average and the median, Figure 14b shows the variation of each low-cost NO_2 sensor in relation to the percentage average values and reference values for each sensor, and Figure 14c shows the variation of each low-cost NO_2 sensor in relation to the percentage median values and reference values for each sensor. Figure 15a shows the total variation, including the values of all three low-cost O_3 sensors, both the average and the median, Figure 15b shows the variation of each low-cost O_3 sensor in relation to the percentage average values and reference values for each sensor, and Figure 15c shows the variance of each low-cost O_3 sensor against the median average variance values versus the reference values for each sensor.



(**a**) (**b**) (**c**)

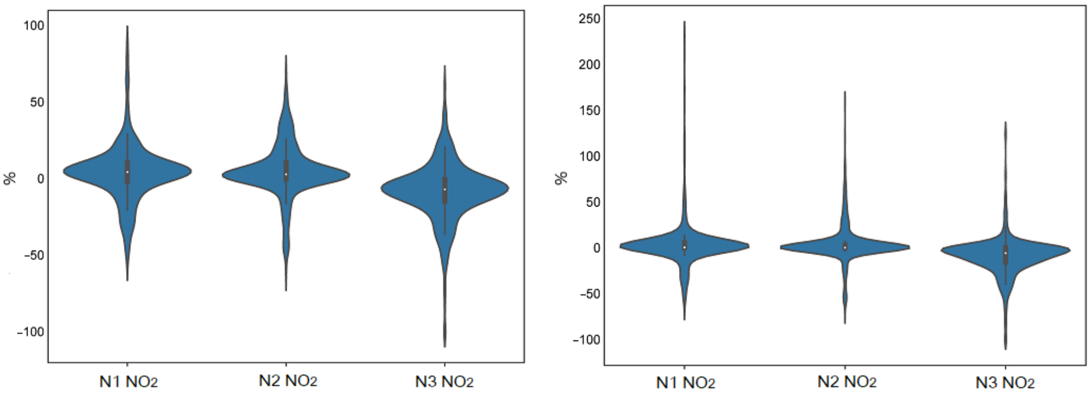
Figure 14. Boxplots of NO_2 LCS. (**a**) Total variation including the values of all three low-cost NO_2 sensors, both the average and the median. (**b**) Variation of each low-cost NO_2 sensor in relation to the percentage average values and reference values for each sensor. (**c**) Variation of each low-cost NO_2 sensor in relation to the percentage median values and reference values for each sensor.



(**a**) (**b**) (**c**)

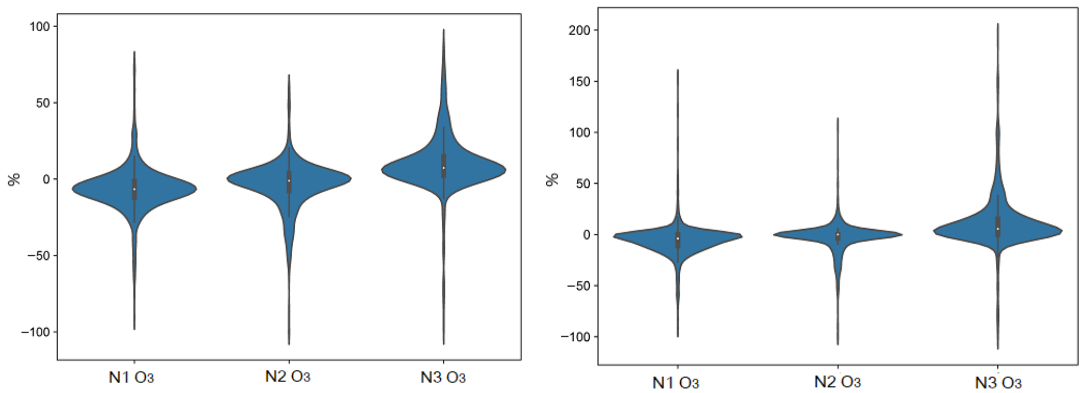
Figure 15. Boxplots of O_3 LCS. (**a**) Total variation including the values of all three low-cost O_3 sensors, both the average and the median. (**b**) Variation of each low-cost O_3 sensor in relation to the percentage average values and reference values for each sensor. (**c**) Variance of each low-cost O_3 sensor in relation to the median average variance values versus the reference values for each sensor.

In addition, the visualization of the percentage change of the data of the electrochemical sensor measurements using violin curves will help to draw more comprehensive conclusions. The violin curves on the vertical axis depict the percentage change of values and, on the horizontal axis, they depict, proportional to the width of the curve, the frequency of repetition of the percentage change of value. Figure 16 shows the violin curves of the distribution of the percentage change of the measurements of the NO_2 electrochemical sensors, Figure 16a depicts the NO_2 measurements distribution of the average method, while Figure 16b shows the NO_2 measurements distribution of the median method. Violin curves presented in Figure 17 show the distribution of the percentage change of the measurements of the O_3, Figure 17a shows the O_3 measurements distribution of the average method and Figure 17b depicts the O_3 measurements distribution of the median method.



(**a**) (**b**)

Figure 16. Violin curves of NO_2 LCS. (**a**) Percentage change distribution of average method of NO_2 measurements. (**b**) Percentage change distribution of median method of NO_2 measurements.



(**a**) (**b**)

Figure 17. Violin curves of O_3 LCS. (**a**) Percentage change distribution of average method of O_3 measurements. (**b**) Percentage change distribution of median method of O_3 measurements.

From Figures 14 and 15 it is possible to observe the dynamic range in which the changes of the measured quantities move and, as a result, the ability of the proposed algorithm to operate in large ranges of values is determined. Examination of the boxplots and violin curves for NO_2 (Figures 14 and 16) and O_3 (Figures 15 and 17) allows for the determination of the safe percentage change range for the average and median methods across all three electrochemical sensors. The NO_2 sensors fall within a safe range of −40% to +50% for both the average and median methods, while the O_3 sensors are presenting a safe variation between −17% and 33% for the average method and −20% and 30% for the median method. Individually, for each NO_2 sensor the average and median methods present a maximum safe variation between −12% and 10% and −13% and +6%, respectively, while individually for each O_3 sensor the average and median methods are presenting a maximum safe variation between −11% and +13% and −10% and +12%, respectively.

3.1.4. Evaluation Procedure

For the evaluation purpose on the set of sensors, per case of NO_2 and O_3 sensors, the common percentage average value of the average method and the common percentage median value of the median method were applied separately to each non-corrected measurement of each sensor. Figure 18 shows the corrected values of each NO_2 sensor using the average variation as a correction factor in relation to the reference values. Figure 19 shows the correlation (R^2) of corrected values (average variation) in relation to the reference values of each sensor. Figure 19a shows the correlation (R^2) of corrected values in relation to the reference value of the NO_2 sensor of Node 1. Figure 19b shows the correlation (R^2) of corrected values in relation to the reference values of the NO_2 sensor of Node 2. Figure 19c shows the correlation (R^2) of corrected values in relation to the reference values of the NO_2 sensor of Node 3.

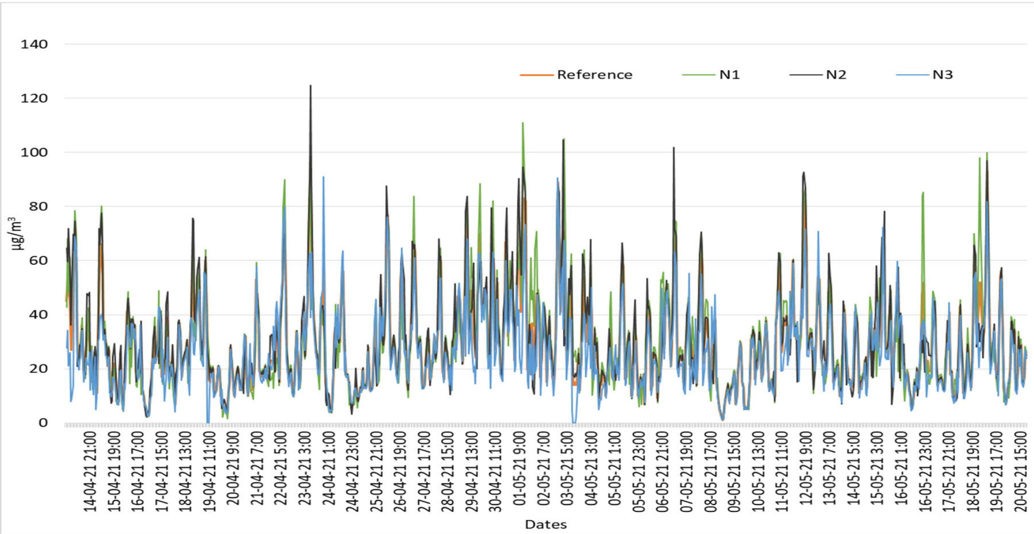
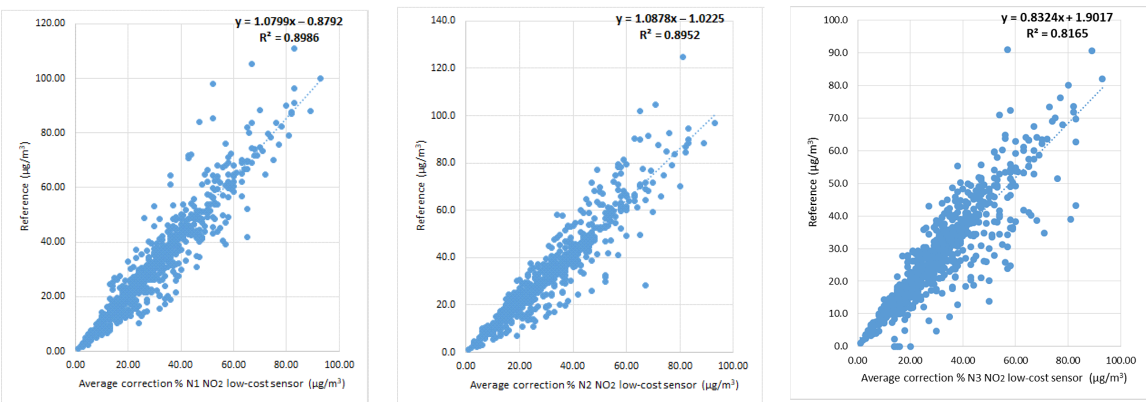


Figure 18. Corrected values of each low-cost NO_2 sensor using the average variation as correction factor in relation to the reference values.



(**a**) (**b**) (**c**)

Figure 19. Scatter plots of corrected values (average variation) in relation to the reference values of each NO_2 sensor. (**a**) Scatter plot of corrected values in relation to the reference values of NO_2 sensor of Node 1. (**b**) Scatter plot of corrected values in relation to the reference values of NO_2 sensor of Node 2. (**c**) Scatter plot of corrected values in relation to the reference values of NO_2 sensor of Node 3.

Figure 20 shows the corrected values of each NO_2 sensor using the median variation as a correction factor in relation to the reference values. Scatter plots by Figure 21 show the correlation (R^2) of corrected values (median variation) in relation to the reference values presented for each sensor. Figure 21a–c display these scatter plots for the NO_2 sensors of Node 1, Node 2, and Node 3, respectively.

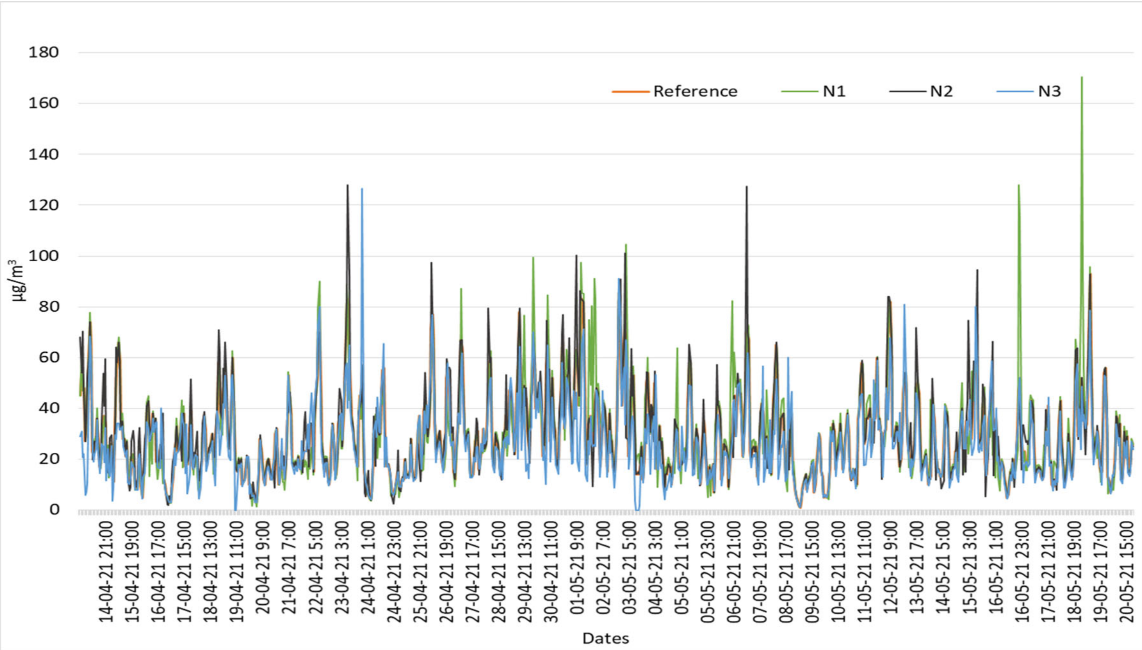
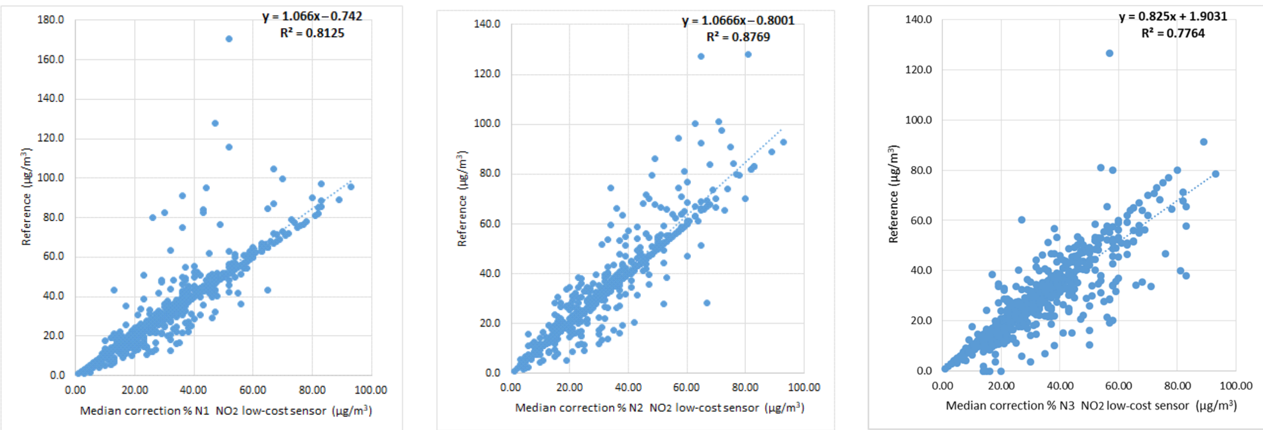


Figure 20. Corrected values of each low-cost NO_2 sensor using the median variation as correction factor in relation to the reference values.



(**a**) (**b**) (**c**)

Figure 21. Scatter plots of corrected values (median variation) in relation to the reference values of each NO_2 sensor. (**a**) Scatter plot of median variation corrected values in relation to the reference values of NO_2 sensor of Node 1. (**b**) Scatter plot of median variation corrected values in relation to the reference values of NO_2 sensor of Node 2. (**c**) Scatter plot of median variation corrected values in relation to the reference values of NO_2 sensor of Node 3.

Figure 22 shows the corrected values of each O_3 sensor using the average variation as a correction factor in relation to the reference values. Figure 23, by scatter plots, shows the correlation (R^2) of corrected values (average variation) in relation to the reference values presented for each sensor. Figure 23a–c display these scatter plots for the O_3 sensors of Node 1, Node 2, and Node 3, respectively.

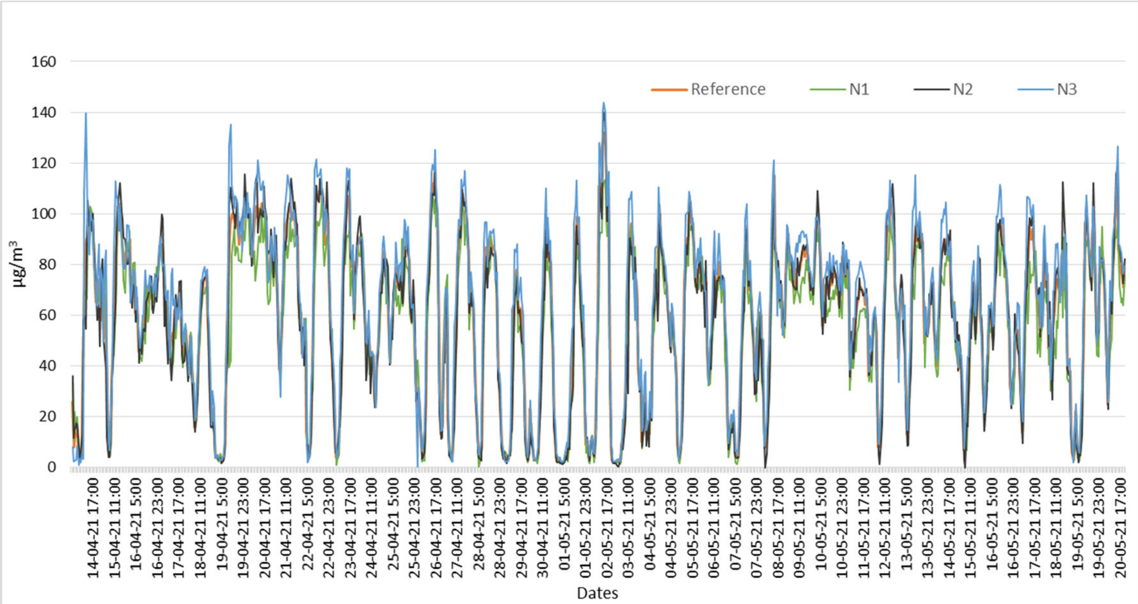


Figure 22. Corrected values of each low-cost O_3 sensor using the average variation as correction factor in relation to the reference values.

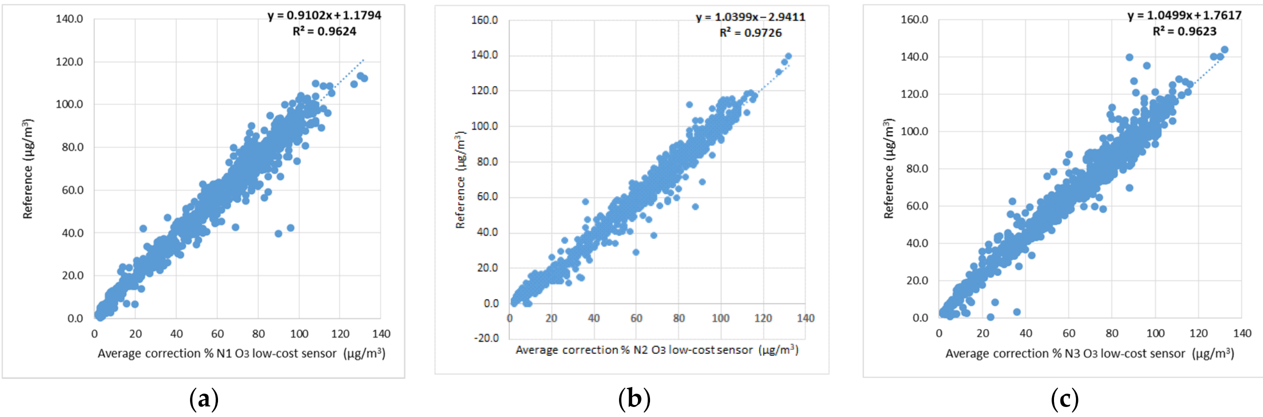


Figure 23. Scatter plots of corrected values (average variation) in relation to the reference values of each O_3 sensor. (**a**) Scatter plot of average variation corrected values in relation to the reference values of O_3 sensor of Node 1. (**b**) Scatter plot of average variation corrected values in relation to the reference values of O_3 sensor of Node 2. (**c**) Scatter plot of average variation corrected values in relation to the reference values of O_3 sensor of Node 3.

Figure 24 shows the corrected values of each O_3 sensor using the median variation as a correction factor in relation to the reference values. Figure 25, by scatter plots, shows the correlation (R^2) of corrected values (median variation) in relation to the reference values presented for each sensor. Figure 25a–c display these scatter plots for the O_3 sensors of Node 1, Node 2, and Node 3, respectively.

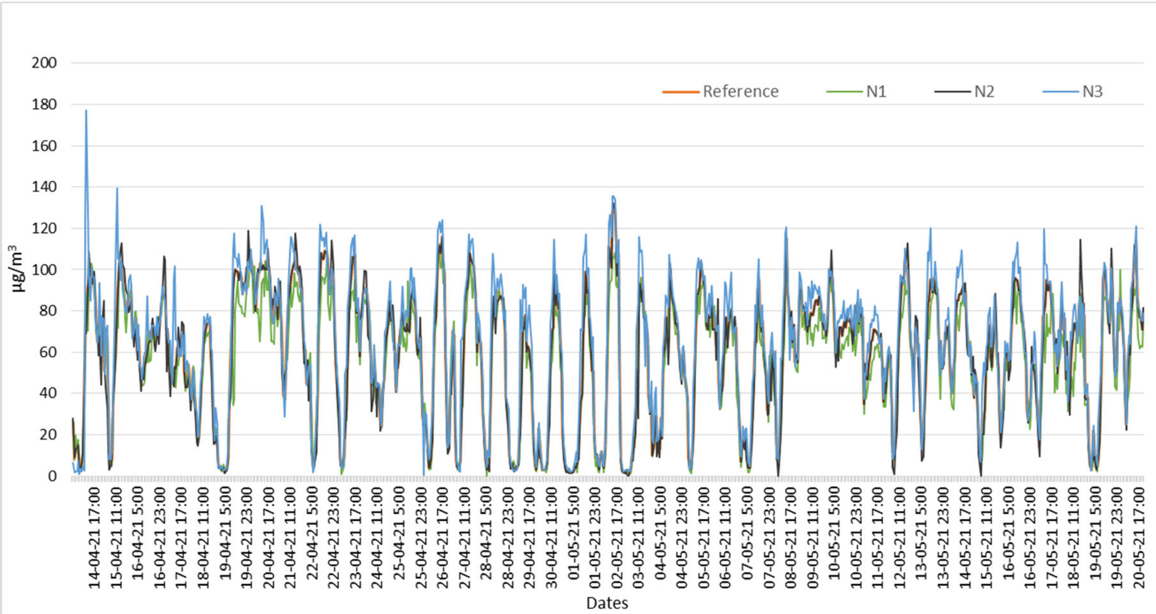
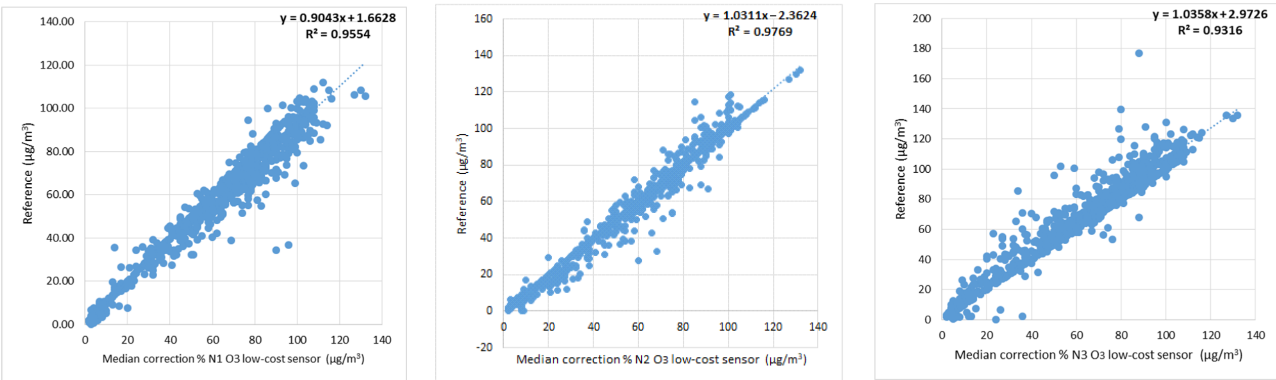


Figure 24. Corrected values of each low-cost O_3 sensor using the median variation as correction factor in relation to the reference values.



(**a**) (**b**) (**c**)

Figure 25. Scatter plots, of corrected values (median variation) in relation to the reference values of each O_3 sensor. (**a**) Scatter plot of median variation corrected values in relation to the reference values of O_3 sensor of Node 1. (**b**) Scatter plot of median variation corrected values in relation to the reference values of O_3 sensor of Node 2. (**c**) Scatter plot of median variation corrected values in relation to the reference values of O_3 sensor of Node 3.

3.2. Statistical Analysis

The statistical analysis of the data is based on regression statistics of linear regression and multiple linear regression, and some statistical terms are interpreted so that the results are understandable. Multiple R is the multiple correlation coefficient between three or more variables. R-squared represents the correlation coefficient between the response variable of a regression model that can be explained by the fitted variables. Adjusted R-squared is similar to R-squared but it adjusts for the number of predictors in a regression model.

3.2.1. Linear Regression (LR)

Linear regression is an approach to model the relation between a scalar response and one explanatory variable. In this work, the linear regression method was applied for each case of gas sensors values for all the low-cost sensing stations.

Liner Regression (LR) NO_2

Table 1 shows the LR of reference values and, non-corrected by average or median variation, values of LCS N1, N2, N3 of NO_2.

Table 1. LR, reference values and N1, N2, N3 (non-corrected values) of NO_2.

Regression Statistics	N1	N2	N3
Multiple R	0.648	0.645	0.577
R Square	0.420	0.416	0.333
Adjusted R Square	0.420	0.415	0.332
Standard Error	12.192	12.241	13.082
Observations	886	886	886
Significance F	8.45×10^{-107}	2.89×10^{-105}	1.02×10^{-79}
Intercept (p-value)	6.72×10^{-60}	2.84×10^{-60}	4.35×10^{-83}
NO_2 Sensor (p-value)	8.45×10^{-107}	2.89×10^{-105}	1.02×10^{-79}

Table 2 shows the LR of reference values and average and median values of non-corrected values (average or median variation) of LCS N1, N2, N3 of NO_2.

Table 2. LR, reference values and the average, and median values of non-corrected values (average or median variation) of NO_2 (N1, N2, N3).

Regression Statistics	Average N1, N2, N3	Median N1, N2, N3
Multiple R	0.652	0.659
R Square	0.425	0.434
Adjusted R Square	0.424	0.434
Standard Error	12.142	12.045
Observations	886	886
Significance F	2.19×10^{-108}	1.723×10^{-111}
Intercept (p-value)	4.98×10^{-57}	1.23×10^{-57}
Aver. N1, N2, N3 (p-value)	2.19×10^{-108}	
Med. N1, N2, N3 (p-value)		1.72×10^{-111}

Table 3 shows the LR of reference values and corrected by average variation values of LCS N1, N2, N3 of NO_2.

Table 3. LR, reference values and average variation corrected values of NO_2 (N1, N2, N3).

Regression Statistics	N1	N2	N3
Multiple R	0.941	0.946	0.896
R Square	0.886	0.895	0.802
Adjusted R Square	0.886	0.894	0.802
Standard Error	5.412	5.199	7.117
Observations	886	886	886
Significance F	0	0	0
Intercept (p-value)	3.66×10^{-19}	9.35×10^{-26}	1.86×10^{-25}
Aver. Corr. (p-value)	0	0	0

Table 4 shows the LR of reference values and corrected by median variation values of LCS N1, N2, N3 of NO_2.

Table 4. LR, reference values and median variation corrected values of NO_2 (N1, N2, N3).

Regression Statistics	N1	N2	N3
Multiple R	0.894	0.939	0.874
R Square	0.799	0.882	0.763
Adjusted R Square	0.799	0.882	0.763
Standard Error	7.172	5.496	7.790
Observations	886	886	886
Significance F	0	0	6.17×10^{-279}
Intercept (p-value)	4.07×10^{-29}	3.89×10^{-22}	1.61×10^{-31}
Med. Corr. (p-value)	0	0	6.17×10^{-279}

Linear Regression (LR) O_3

Table 5 shows the LR of reference values and, non-corrected by average or median variation, values of LCS N1, N2, N3 of O_3.

Table 5. LR, reference values and N1, N2, N3 (non-corrected values) of O_3.

	Regression Statistics	N2	N3
Multiple R	0.748	0.785	0.764
R Square	0.559	0.617	0.584
Adjusted R Square	0.559	0.617	0.584
Standard Error	20.146	18.776	19.569
Observations	886	886	886
Significance F	1.95×10^{-159}	1.68×10^{-186}	1.32×10^{-170}
Intercept (*p*-value)	7.40×10^{-7}	6.87×10^{-11}	0.0438
O_3 Sensor (*p*-value)	1.95×10^{-159}	1.68×10^{-186}	1.32×10^{-170}

Table 6 shows the LR of reference values and average and median values of non-corrected values (average or median variation) of LCS N1, N2, N3 of O_3.

Table 6. LR, reference values and the average, and median values of non-corrected values (average or median variation) of O_3 (N1, N2, N3).

Regression Statistics	Average N1, N2, N3	Median N1, N2, N3
Multiple R	0.786	0.782
R Square	0.618	0.612
Adjusted R Square	0.618	0.611
Standard Error	18.748	18.900
Observations	887	887
Significance F	3.68×10^{-187}	4.67×10^{-184}
Intercept (*p*-value)	0.005	0.0002
Aver. N1, N2, N3 (*p*-value)	3.68×10^{-187}	
Med. N1, N2, N3 (*p*-value)		4.67×10^{-184}

Table 7 shows the LR of reference values and corrected by average variation values of LCS N1, N2, N3 of O_3.

Table 7. LR, reference values and average variation corrected values of O_3 (N1, N2, N3).

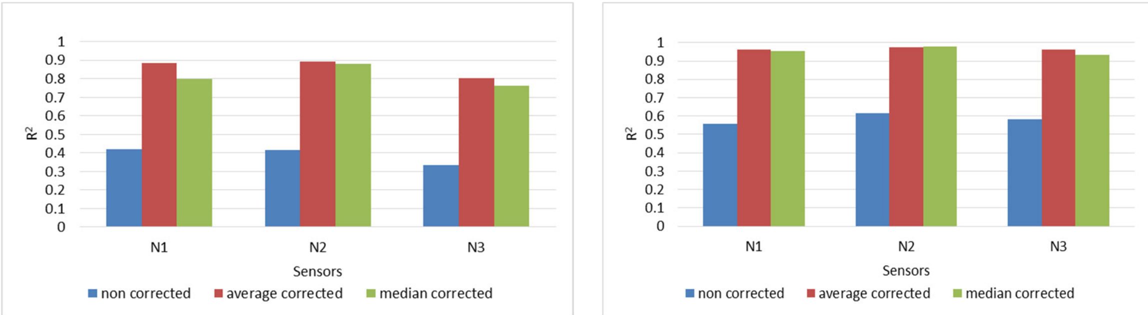
Regression Statistics	N1	N2	N3
Multiple R	0.981	0.986	0.981
R Square	0.963	0.973	0.962
Adjusted R Square	0.963	0.973	0.962
Standard Error	5.863	4.989	5.883
Observations	886	886	886
Significance F	0	0	0
Intercept (*p*-value)	0.03	1.34×10^{-32}	0.15
Aver. Corr. (*p*-value)	0	0	0

Table 8 shows the LR of reference values and corrected by median variation values of LCS N1, N2, N3 of O_3.

Table 8. LR, reference values and median variation corrected values of O_3 (N1, N2, N3).

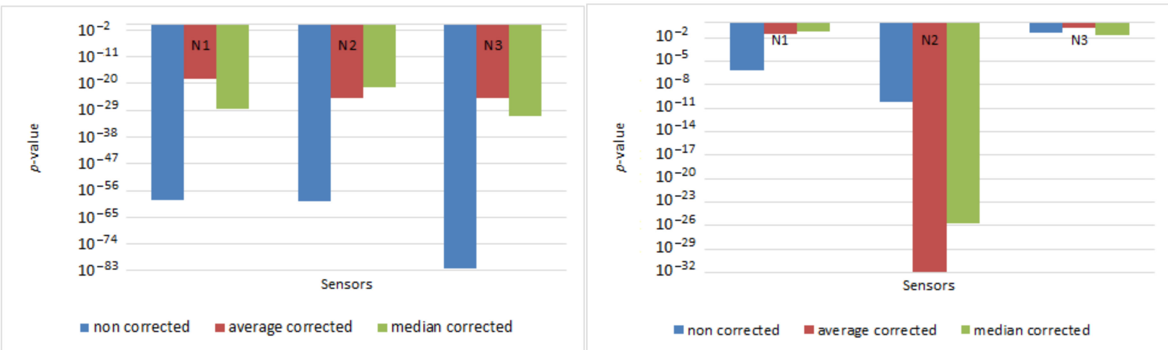
Regression Statistics	N1	N2	N3
Multiple R	0.978	0.988	0.965
R Square	0.956	0.977	0.932
Adjusted R Square	0.956	0.977	0.932
Standard Error	6.378	4.591	7.930
Observations	886	886	886
Significance F	0	0	0
Intercept (*p*-value)	0.069	1.85×10^{-26}	0.018
Med. Corr. (*p*-value)	0	0	0

From the above tables of linear regressions we derived the follow figures which represent, in Figure 26, the degree of correlations (R^2), and in Figure 27, the value of intercept (*p*-value), of the non-corrected measurements, corrected measurements by the average method, and corrected measurements by the median method, of each sensor and each gas.



(**a**) (**b**)

Figure 26. LR correlations (R^2) of non-corrected, corrected by average method, and corrected by median method, measurements in respect to the reference measurements. (**a**) LR correlations (R^2) of NO_2 gas concentrations, (**b**) LR correlations (R^2) of O_3 gas concentrations.



(**a**) (**b**)

Figure 27. Value of LR interception (*p*-value) of non-corrected, corrected by average method, and corrected by median method, measurements in respect to the reference measurements. (**a**) LR intercept (*p*-value) of NO_2 gas concentrations, (**b**) LR interception (*p*-value) of O_3 gas concentrations.

The results of the linear regression are shown in Figure 26. It can be seen for both nitrogen dioxide and ozone that the correlation coefficient improves both for the average method and at the median method. For nitrogen dioxide, with the non-corrected values, the correlation coefficient shows a minimum value of 0.33 and a maximum value of 0.42, and for the corrected values with the average method, the correlation coefficient shows a minimum value of 0.80 and a maximum value of 0.89, and for the corrected values with the median method, the correlation coefficient appears with a minimum value of 0.76 and a maximum value of 0.88, while for ozone for the non-corrected values the correlation coefficient shows a minimum value of 0.56 and a maximum value of 0.62, for the corrected values with the average method the correlation coefficient shows a minimum value of 0.96 and a maximum value of 0.97, and for the corrected values with the median method the correlation coefficient shows a minimum value of 0.96 and a maximum value of 0.98. On the other hand, in Figure 27, the reduction of the intercept (*p*-value) can be observed after the implementation of the average and median methods in relation to the non-corrected, which results in better reliability of the corrected measurements. Based on the linear regression results, valuable insights can be obtained. Table 1 shows that the uniformity between the sensors can be distinguished, as well as the significance F but also the small degree of the *p*-value for all three nitrogen dioxide sensors. Table 2 shows the linear regression of both the average and the median per set of sensors, and in this case the significance F shows a very low value, while the *p*-value values of both the mean and the median are equally very small. Table 3 shows the regression between the corrected values with the variance of the average relative to the reference values, the degree of significance F is excellent as is the *p*-value of the corrected values. Similar behavior is observed in Table 4, as it shows a high degree of correlation of the corrected values with the variation of the median in relation to the reference values, the degree of significance is very high as is the *p*-value of the corrected values. Table 5 shows the results of linear regression of three ozone sensors; the sensors show similar behavior in terms of the degree of correlation, while the degree of significance and *p*-value show very low values. Table 6 shows the behavior of the average and the median by sensor category, the degree of significance F is very low, the degree of correlation is satisfactory, while the *p*-values of both the average and the median are very low. Table 7 shows the linear regression of the corrected values with the average variation, where the degree of correlation is excellent, as well as the significance F but also the *p*-value of the corrected values with the average variation. Table 8 shows the correction of values with the median variation, it shows an excellent degree of correlation, as well as the significance F, and the *p*-value of the corrected values with the median variation.

3.2.2. Multiple Linear Regression (MLR)

Multiple linear regression (MLR) is a model of the linear relationship between the explanatory variables and response variable. In this work, multiple linear regression was applied for each type of gas sensor, in order to evaluate whether each sensor is affected by the percentage change of the average and median values. The significance level is equal 5% ($\alpha = 0.05$).

Multiple Liner Regression (MLR) NO_2

Table 9 shows the MLR of reference values and the affection of average and median percentage changes in the non-corrected values (average or median variation) of LCS N1, N2, N3 of NO_2.

Table 9. MLR, reference values, and the affection of average and median percentage changes in non-corrected values (average or median variation) of NO_2 (N1, N2, N3).

Regression Statistics	Average N1, N2, N3	Median N1, N2, N3
Multiple R	0.789	0.795
R Square	0.622	0.632
Adjusted R Square	0.622	0.631
Standard Error	9.847	9.724
Observations	886	886
Significance F	1.90×10^{-187}	2.84×10^{-192}
Intercept (*p*-value)	4.57×10^{-129}	1.20×10^{-131}
Aver. N1, N2, N3 (*p*-value)	3.73×10^{-164}	
% Average—Ref (*p*-value)	1.24×10^{-82}	
Med. N1, N2, N3 (*p*-value)		3.91×10^{-169}
% Median—Ref (*p*-value)		2.29×10^{-84}

Table 10 shows the MLR of reference values and corrected by average variation values (using average variation % as correction factor) of LCS N1, N2, N3 of NO_2.

Table 10. MLR, reference values, and average variation corrected values of NO_2 (N1, N2, N3).

Regression Statistics	N1	N2	N3
Multiple R	0.943	0.946	0.896
R Square	0.889	0.895	0.803
Adjusted R Square	0.8885.348	0.895	0.802
Standard Error	886	5.194	7.114
Observations		886	886
Significance F	0	0	0
Intercept (*p*-value)	1.67×10^{-10}	1.16×10^{-19}	2.87×10^{-24}
Aver. Corr. (*p*-value)	0	0	0
% Average—Ref (*p*-value)	2.45×10^{-6}	0.10	0.21

Table 11 shows the MLR of reference values and corrected by median variation values (using median variation % as correction factor) of LCS N1, N2, N3 of NO_2.

Table 11. MLR, reference values, and median variation corrected values of NO_2 (N1, N2, N3).

Regression Statistics	N1	N2	N3
Multiple R	0.897	0.940	0.874
R Square	0.805	0.882	0.763
Adjusted R Square	0.804	0.883	0.763
Standard Error	7.078	5.486	7.794
Observations	886	886	886
Significance F	0	0	4.00×10^{-277}
Intercept (*p*-value)	9.38×10^{-17}	3.01×10^{-16}	5.98×10^{-27}
Median Corr. (*p*-value)	0	0	7.94×10^{-275}
% Median—Ref (*p*-value)	8.86×10^{-7}	0.037	0.80

Multiple Linear Regression (MLR) O_3

Table 12 shows the MLR of reference values and the affection of average and median percentage changes in the non-corrected values (average or median variation) of LCS N1, N2, N3 of O_3.

Table 12. MLR, reference values, and the affection of average and median percentage change in non-corrected values (average or median variation) of O_3 (N1, N2, N3).

Regression Statistics	Average N1, N2, N3	Median N1, N2, N3
Multiple R	0.839	0.837
R Square	0.704	0.701
Adjusted R Square	0.703	0.701
Standard Error	16.514	16.589
Observations	887	887
Significance F	2.03×10^{-234}	1.10×10^{-232}
Intercept (*p*-value)	1.10×10^{-23}	2.92×10^{-27}
Aver. N1, N2, N3 (*p*-value)	7.60×10^{-190}	
% Average—Ref (*p*-value)	6.55×10^{-51}	
Med. N1, N2, N3 (*p*-value)		1.03×10^{-189}
% Median—Ref (*p*-value)		2.80×10^{-52}

Table 13 shows the MLR of reference values and corrected by average variation values (using average variation % as correction factor) of LCS N1, N2, N3 of O_3.

Table 13. MLR, reference values, and average variation corrected values of O_3 (N1, N2, N3).

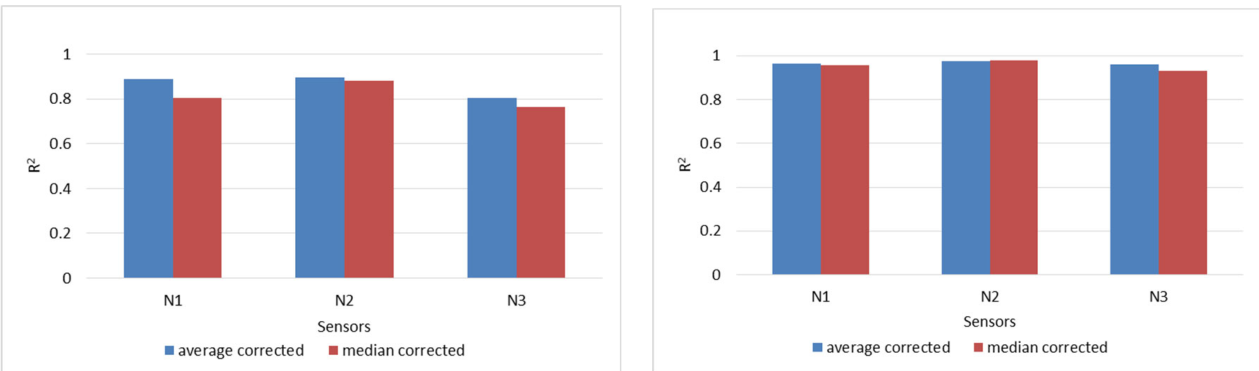
Regression Statistics	N1	N2	N3
Multiple R	0.981	0.987	0.981
R Square	0.963	0.974	0.962
Adjusted R Square	0.963	0.974	0.962
Standard Error	5.856	4.889	5.882
Observations	886	886	886
Significance F	0	0	0
Intercept (*p*-value)	0.005	1.97×10^{-38}	0.067
Aver. Corr. (*p*-value)	0	0	0
% Average—Ref (*p*-value)	0.08	1.35×10^{-9}	0.257

Table 14 shows the MLR of reference values and corrected by median variation values (using median variation % as correction factor) of LCS N1, N2, N3 of O_3.

Table 14. MLR, reference values, and median variation corrected values of O_3 (N1, N2, N3).

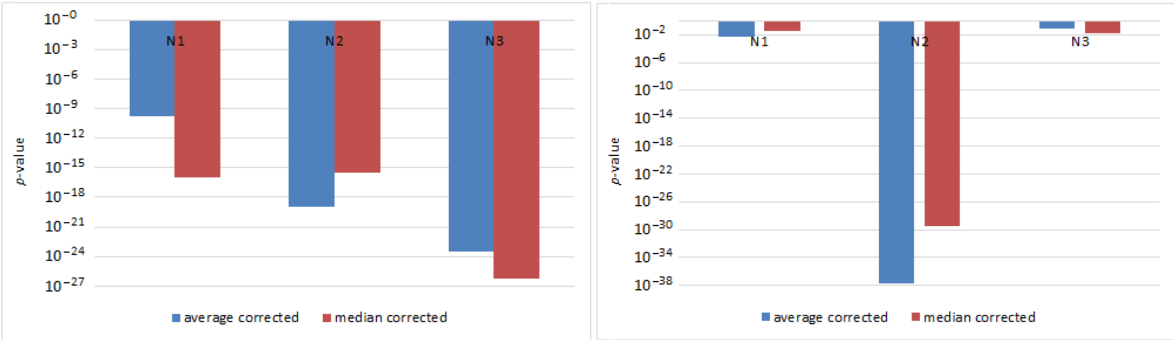
Regression Statistics	N1	N2	N3
Multiple R	0.978	0.989	0.965
R Square	0.956	0.977	0.932
Adjusted R Square	0.956	0.977	0.932
Standard Error	6.377	4.525	7.931
Observations	886	886	886
Significance F	0	0	0
Intercept (*p*-value)	0.031	2.70×10^{-30}	0.018
Median Corr. (*p*-value)	0	0	0
% Median—Ref (*p*-value)	0.236	3.00×10^{-7}	0.369

From Tables 9–14 of multiple linear regressions, we derived the follow figures which represent, in Figure 28, the degree of correlations (R^2), in Figure 29, the value of intercept (*p*-value), and in Figure 30, the % value (*p*-value) of the average and median methods as a correction factor, of the corrected measurements by the average method and corrected measurements by the median method, of each sensor and for each gas.



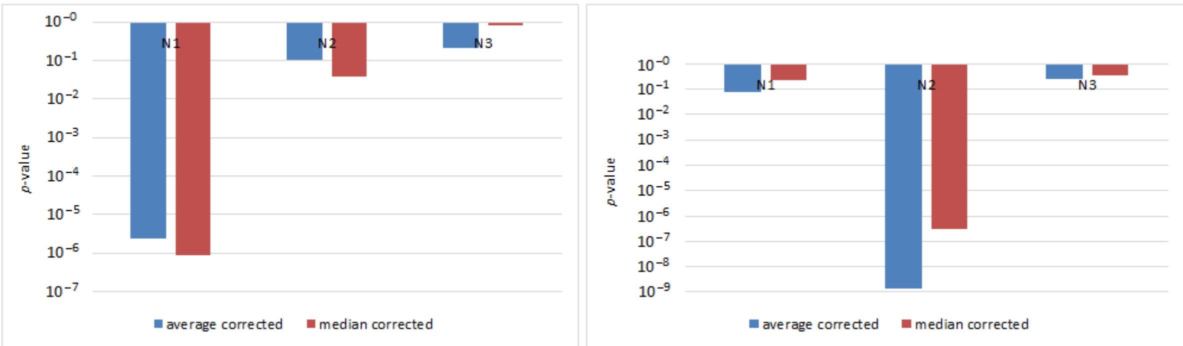
(**a**) (**b**)

Figure 28. MLR correlations (R^2) of corrected by average method and corrected by median method, measurements in respect to the reference measurements. (**a**) MLR correlations (R^2) of NO_2 gas concentrations, (**b**) MLR correlations (R^2) of O_3 gas concentrations.



(**a**) (**b**)

Figure 29. Value of MLR interception (*p*-value) corrected by average method and corrected by median method, measurements in respect to the reference measurements. (**a**) MLR intercept (*p*-value) of NO_2 gas concentrations, (**b**) MLR interception (*p*-value) of O_3 gas concentrations.



(**a**) (**b**)

Figure 30. Value of MLR percentage change (%) as a correction factor (*p*-value) of corrected by average method and corrected by median method, measurements in respect to the reference measurements. (**a**) MLR percentage change (%) as a correction factor (*p*-value) of NO_2 gas concentrations, (**b**) MLR percentage change (%) as a correction factor (*p*-value) of O_3 gas concentrations.

The results of the multiple linear regression are summarized in Figures 28–30. From Figure 28, it can be seen for both nitrogen dioxide and ozone that the correlation coefficient is improved in both the average and the median methods. For nitrogen dioxide, for the corrected values with the average method, the correlation coefficient shows a minimum value of 0.80 and a maximum value of 0.89, for the corrected values with the median method, the correlation coefficient shows a minimum value of 0.76 and a maximum value of 0.88, while for ozone, for the corrected values with the average method, the correlation coefficient shows a minimum value of 0.96 and a maximum value of 0.97, for the corrected values with the median method, the correlation coefficient shows a minimum value of 0.93 and a maximum value of 0.98. Figures 29 and 30 show the degree of *p*-value, both of the intercept and the percentage change (%), respectively, and in both cases, the small degree of the *p*-value is observed after the application of the average and median methods, which means that both the intercept and the percentage change have a direct and significant effect on the correction of measurements. The results of the Multiple Linear Regression yield intriguing findings. Table 9 shows the multiple linear regression of both the average and the median, of the three nitrogen dioxide (NO_2) LCS measurements, in relation to the reference measurements; the homogeneity is shown by the satisfactory degree of correlation while the significance F presents a very low value, the values of the *p*-value both of the average and of the median are respectively very small, also a very low *p*-value appears in the variation of both the average—reference and the median—reference. Tables 10 and 11 show the results of the multiple linear regression between the corrected values of measurements with the average variation relative to the reference measurements, and the corrected values of measurements with the median variation in relation to the reference measurements, respectively. In these two cases, the degree of significance F is excellent, as well as the *p*-value of the corrected values in each case; also the percentage change of the average—reference and the percentage change of the median—reference, shows a very low *p*-value. Table 12 shows the average and median results of multiple linear regression of the three ozone (O_3) LCS measurements, in relation to the reference measurements. The similarity is shown by the correlation of regression that presents a great degree while the significant F presents a satisfactory value. The *p*-values both of the average and of the median are respectively very small, while a very low *p*-value appears in the variation of both the average—reference and the median—reference. Tables 13 and 14 show the results between the corrected values of measurements with the average variation relative to the reference measurements, and the corrected values of measurements with the median variation in relation to the reference measurements of the multiple linear regression, respectively. Both the regression of the average variation and the median variation presented an excellent degree of significance F while the *p*-values of the average corrected values and median corrected values were quite low. In addition, very low *p*-values are observed for the variation of the average—reference and the variation of the median—reference.

Table 15 shows the correlation degree of the proposed methodology in this work in relation to correlation degree from another published research works that have used different calibration methods, such as artificial neural networks (ANN), linear (LINEAR), and multiple linear regression (MLR).

Table 15. Comparison of correlation degree of our work and published works.

Gas Pollutant	Calibration Method	References	R^2
NO_2	ANN	Spinelle et al. [1], Wastine et al. [34]	0.94
	LINEAR	Wastine et al. [34], Castell et al. [35], Cross et al. [29]	0.17
	MLR	Spinelle et al. [1], Karagulian et al. [36], Wei et al. [37]	0.81
	Average/Median	Our work	0.78–0.87
O_3	ANN	Spinelle et al. [1], Wastine et al. [34]	0.89
	LINEAR	Wastine et al. [34], Castell et al. [35], Cross et al. [29]	0.53
	MLR	Spinelle et al. [1], Karagulian et al. [36], Wei et al. [37]	0.91
	Average/Median	Our work	0.93–0.96

Observing Table 15, it becomes evident that the goal of the work has been achieved as the degree of correlation of the concentration of the gaseous pollutants nitrogen dioxide and ozone presents a high value which is comparable to other research works.

3.2.3. MAD, MSE, MAPE, RMSE Statistical Modes

The evaluation of the measurements of the experiment was done by applying statistical models MAD, MSE, MAPE, RMSE to confirm the robustness of the results. These statistical models were applied to the non-corrected measurements, the average variation corrected measurements, and the median variation corrected measurements. The results of MAD, MSE, MAPE, RMSE models are presented in Tables 16–19, respectively.

Table 16. MAD of non-corrected, average variation corrected, and median variation corrected values of each LCS.

MAD	N1 NO_2	N2 NO_2	N3 NO_2	N1 O_3	N2 O_3	N3 O_3
Non-corrected	2.1	3.2	2.6	15.3	18.6	17.1
Average corrected	9.8	10.3	8.9	19.5	21.8	21.1
Median corrected	9.7	10.2	9.0	18.6	21.8	21.4

Table 17. MSE of non-corrected, average variation corrected, and median variation corrected values of each LCS.

MSE	N1 NO_2	N2 NO_2	N3 NO_2	N1 O_3	N2 O_3	N3 O_3
Non-corrected	216.4	223.2	265.0	418.8	375.3	438.3
Average corrected	37.8	38.7	59.6	54.2	29.3	64.8
Median corrected	67.2	39.0	72.6	59.3	24.0	100.3

Table 18. MAPE of non-corrected, average variation corrected, and median variation corrected values of each LCS.

MAPE	N1 NO_2	N2 NO_2	N3 NO_2	N1 O_3	N2 O_3	N3 O_3
Non-corrected	49.0	52.7	50.8	62.5	56.2	81.1
Average corrected	12.9	11.7	14.8	11.2	9.9	13.9
Median corrected	10.5	9.1	13.81	9.1	7.0	14.8

Table 19. RMSE of non-corrected, average variation corrected, and median variation corrected values of each LCS.

RMSE	N1 NO_2	N2 NO_2	N3 NO_2	N1 O_3	N2 O_3	N3 O_3
Non-corrected	0.35	0.48	1.63	0.68	0.30	0.30
Average corrected	0.06	0.04	0.10	0.12	0.19	0.31
Median corrected	0.04	0.00	0.06	0.09	0.00	0.25

The mean absolute deviation (MAD) in both types of sensors (NO_2 and O_3) exhibits more consistent behavior in the non-corrected values. It is noteworthy that the MAD values are very close in both the corrected values with the average variation, and in the corrected values with the median variance. For this reason, the mean absolute deviation (MAD) model was also applied to the reference values where the resulting value for NO_2 was 10, while for O_3 it was 19.8. This confirms the reliability of the results, after the application of the measurement values' correction, in both cases (average correction, median correction).

The mean square error (MSE) shows significant improvement both in the corrected values with the average variation, and in the corrected values with the median variation, in relation to the non-corrected values. That means the improvement has a direct relation to the corrected values (average, median) relative to reference values.

Equally satisfactory are the results that appear in the mean absolute percentage error (MAPE), in both cases (average and median) the corrected values show a significant improvement in relation to the uncorrected values.

Finally, the results shown by the root mean square error (RMSE) method are exceptional, as the application of this method to the experiment shows a very low degree, both in the corrected values of average variation, as well as in the corrected values of median variation, in relation to the high degree (RMSE) of the non-corrected values.

4. Conclusions

In today's era of rapid scientific and technological advancements, the development of low-cost sensors has become a reality. The use of low-cost sensors has flooded the market. Owing to their affordability, a growing number of individuals are opting to acquire low-cost sensors for collecting air quality data, recognizing the substantial impact of air quality on human health.

This article is based on the data analysis of low-cost electrochemical air quality sensors; the sensors that were used are from the Alphasence, specifically, the nitrogen dioxide (NO_2) sensors (NO2-B43F), and ozone sensors (O_3) (OX-B431) were studied. The goal was to define the percentage change of measurements from these sensors relative to reference measurements. This can be applied to similar low-cost sensor measurements, providing more accurate results. In addition, the investigation of the safe limits of measurement variability of electrochemical air quality sensors took place. The application of the average and median methods was performed to identify the aforementioned. The results showed, for the non-corrected measurements of the low-cost sensors in relation to the reference measurements, a degree of correlation (R^2) for nitrogen dioxide 0.33–0.42, while for ozone it was 0.62–0.66. The average and median of the measurements of the non-corrected measurements of the nitrogen dioxide sensors relative to the reference measurements showed a degree of correlation (R^2) of 0.42 and 0.43, respectively, while for ozone the degree of correlation (R^2) of the average and of the median was 0.67.

The common average of the percentage change of the average of the sensor measurements, and the common average of the percentage change of the median of the sensor measurements were calculated. From the distribution of the percentage change, the safe limits of change that the sensors can have were identified; the results extracted from Figures 18–21 (boxplots and violin curves) for both methods show that the safe limits can have a change of measurements of the nitrogen dioxide sensors from −13% to +10% while for measurements of the ozone sensors it is from −13% to +13%.

Ascertaining the percentage change or the deviation between the measurements from low-cost sensors in relation to the reference measurements can be used as a correction factor in measurements from low-cost sensors. The outcome of the proposed methodology is considered successful, as the extracted results of Table 15 show a degree of correlation similar to other research works that apply different calibration methods.

For the purpose of evaluation, the common average of the percentage change of the average of the sensors, and the common average of the percentage change of the median of the sensors, were applied to the non-corrected measurements of the low-cost sensors, and the results for both cases showed a correlation degree (R^2) of nitrogen dioxide from 0.78 to 0.89, and a degree of correlation (R^2) of ozone from 0.93 to 0.97. The results were analyzed with the methods of linear regression and multiple linear regression, where both the significance F of the hypothesis, as well as the participation (p-value) of the average and the median coefficients can be seen. The correctness of the results is confirmed by the MAD, MSE, MAPE, and RMSE methods which confirm the methodology implemented in this work. An extension of this work would involve multiple, more than three, low-cost sensors and evaluation at other locations with different atmospheric characteristics.

Low-cost electrochemical air quality sensors are employed for environmental monitoring, and the reliability of the measurements can be improved, as seen in this work, through data processing methods, ensuring the accuracy of information provided to the public.

Author Contributions: Conceptualization, I.S. and I.C.; methodology, I.S., E.S., O.T. and I.C.; software, I.C.; validation, I.S., E.S., O.T. and I.C.; formal analysis, E.S. and I.C.; investigation, I.C.; resources, I.C. and E.S.; data curation, O.T. and I.C.; writing—original draft preparation, I.C. and E.S.; writing—review and editing, O.T., E.S. and I.S.; visualization, I.C.; supervision, O.T. and I.S.; project administration, I.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All of the data created in this study are presented in the context of this article.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. 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.* **2015**, *215*, 249–257. [CrossRef]
2. Mead, M.I.; Popoola, O.A.M.; Stewart, G.B.; Landshoff, P.; Calleja, M.; Hayes, M.; Baldovi, J.J.; McLeod, M.W.; Hodgson, T.F.; Dicks, J.; et al. The Use of Electrochemical Sensors for Monitoring Urban Air Quality in Low-Cost, High-Density Networks. *Atmos. Environ.* **2013**, *70*, 186–203. [CrossRef]
3. Ikram, J.; Tahir, A.; Kazmi, H.; Khan, Z.; Javed, R.; Masood, U. View: Implementing Low Cost Air Quality Monitoring Solution for Urban Areas. *Environ. Syst. Res.* **2012**, *1*, 10. [CrossRef]
4. Jiao, W.; Hagler, G.; Williams, R.; Sharpe, R.; Brown, R.; Garver, D.; Judge, R.; Caudill, M.; Rickard, J.; Davis, M.; 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.* **2016**, *9*, 5281–5292. [CrossRef] [PubMed]
5. Christakis, I.; Hloupis, G.; Stavrakas, I.; Tsakirdis, O. Low Cost Sensor Implementation and Evaluation for Measuring NO_2 and O_3 Pollutants. In Proceedings of the 2020 9th International Conference on Modern Circuits and Systems Technologies (MOCAST), Bremen, Germany, 7–9 September 2020. [CrossRef]
6. Arroyo, P.; Gómez-Suárez, J.; Suárez, J.I.; Lozano, J. Low-Cost Air Quality Measurement System Based on Electrochemical and PM Sensors with Cloud Connection. *Sensors* **2021**, *21*, 6228. [CrossRef] [PubMed]
7. Borrego, C.; Coutinho, M.; Costa, A.M.; Ginja, J.; Ribeiro, C.; Monteiro, A.; Ribeiro, I.; Valente, J.; Amorim, J.H.; Martins, H.; et al. Challenges for a New Air Quality Directive: The Role of Monitoring and Modelling Techniques. *Urban Clim.* **2015**, *14*, 328–341. [CrossRef]
8. Morawska, L.; Thai, P.K.; Liu, X.; Asumadu-Sakyi, A.; Ayoko, G.; Bartonova, A.; Bedini, A.; Chai, F.; Christensen, B.; Dunbabin, M.; et al. Applications of Low-Cost Sensing Technologies for Air Quality Monitoring and Exposure Assessment: How Far Have They Gone? *Environ. Int.* **2018**, *116*, 286–299. [CrossRef] [PubMed]
9. Ferrer-Cid, P.; Barcelo-Ordinas, J.M.; Garcia-Vidal, J.; Ripoll, A.; Viana, M. Multi-Sensor Data Fusion Calibration in IoT Air Pollution Platforms. *IEEE Internet Things J.* **2020**, *7*, 3124–3132. [CrossRef]
10. Kosmidis, E.; Syropoulou, P.; Tekes, S.; Schneider, P.; Spyromitros-Xioufis, E.; Riga, M.; Charitidis, P.; Moumtzidou, A.; Papadopoulos, S.; Vrochidis, S.; et al. hackAIR: Towards Raising Awareness about Air Quality in Europe by Developing a Collective Online Platform. *ISPRS Int. J. Geo-Inf.* **2018**, *7*, 187. [CrossRef]
11. Kumar, P.; Morawska, L.; Martani, C.; Biskos, G.; Neophytou, M.; Di Sabatino, S.; Bell, M.; Norford, L.; Britter, R. The Rise of Low-Cost Sensing for Managing Air Pollution in Cities. *Environ. Int.* **2015**, *75*, 199–205. [CrossRef]
12. Gäbel, P.; Koller, C.; Hertig, E. Development of Air Quality Boxes Based on Low-Cost Sensor Technology for Ambient Air Quality Monitoring. *Sensors* **2022**, *22*, 3830. [CrossRef]
13. Kumar, P.; Skouloudis, A.N.; Bell, M.; Viana, M.; Carotta, M.C.; Biskos, G.; Morawska, L. Real-Time Sensors for Indoor Air Monitoring and Challenges Ahead in Deploying Them to Urban Buildings. *Sci. Total Environ.* **2016**, *560–561*, 150–159. [CrossRef]
14. Borghi, F.; Spinazzè, A.; Rovelli, S.; Campagnolo, D.; Cattaneo, A.; Cavallo, D. Miniaturized Monitors for Assessment of Exposure to Air Pollutants: A Review. *Int. J. Environ. Res. Public Health* **2017**, *14*, 909. [CrossRef] [PubMed]
15. Clements, A.L.; Griswold, W.G.; Abhijit, R.S.; Johnston, J.E.; Herting, M.M.; Thorson, J.; Collier-Oxandale, A.; Hannigan, M. Low-Cost Air Quality Monitoring Tools: From Research to Practice (a Workshop Summary). *Sensors* **2017**, *17*, 2478. [CrossRef] [PubMed]
16. McKercher, G.R.; Salmond, J.A.; Vanos, J.K. Characteristics and Applications of Small, Portable Gaseous Air Pollution Monitors. *Environ. Pollut.* **2017**, *223*, 102–110. [CrossRef] [PubMed]
17. Jovašević-Stojanović, M.; Bartonova, A.; Topalović, D.; Lazović, I.; Pokrić, B.; Ristovski, Z. On the Use of Small and Cheaper Sensors and Devices for Indicative Citizen-Based Monitoring of Respirable Particulate Matter. *Environ. Pollut.* **2015**, *206*, 696–704. [CrossRef]
18. Baron, R.; Saffell, J. Amperometric Gas Sensors as a Low Cost Emerging Technology Platform for Air Quality Monitoring Applications: A Review. *ACS Sens.* **2017**, *2*, 1553–1566. [CrossRef] [PubMed]
19. Laref, R.; Losson, E.; Sava, A.; Siadat, M. Field Evaluation of Low Cost Sensors Array for Air Pollution Monitoring. In Proceedings of the 2019 10th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS), Metz, France, 18–21 September 2019. [CrossRef]
20. Sayahi, T.; Butterfield, A.; Kelly, K.E. Long-Term Field Evaluation of the Plantower PMS Low-Cost Particulate Matter Sensors. *Environ. Pollut.* **2019**, *245*, 932–940. [CrossRef] [PubMed]
21. Christakis, I.; Hloupis, G.; Tsakiridis, O.; Syropoulou, P.; Papadakis, N.; Stavrakas, I. Performance comparison of low-cost gas and particles sensors. In Proceedings of the Eighth International Conference on Environmental Management, Engineering, Planning & Economics, Thessaloniki, Greece, 20–24 July 2021; pp. 332–341.

22. Hojaiji, H.; Kalantarian, H.; Bui, A.A.T.; King, C.E.; Sarrafzadeh, M. Temperature and Humidity Calibration of a Low-Cost Wireless Dust Sensor for Real-Time Monitoring. In Proceedings of the 2017 IEEE Sensors Applications Symposium (SAS), Glassboro, NJ, USA, 13–15 March 2017. [CrossRef]
23. DeSouza, P. A Nairobi Experiment in Using Low Cost Air Quality Monitors. *Clean Air J.* **2017**, *27*, 12–42. [CrossRef]
24. Levy Zamora, M.; Xiong, F.; Gentner, D.; Kerkez, B.; Kohrman-Glaser, J.; Koehler, K. Field and Laboratory Evaluations of the Low-Cost Plantower Particulate Matter Sensor. *Environ. Sci. Technol.* **2018**, *53*, 838–849. [CrossRef]
25. Alphasense. Alphasense UK—Browse Gas Sensors & Air Quality Monitors. Available online: http://www.alphasense.com (accessed on 20 December 2023).
26. He, M.; Kuerbanjiang, N.; Dhaniyala, S. Performance Characteristics of the Low-Cost Plantower PMS Optical Sensor. *Aerosol Sci. Technol.* **2019**, *54*, 232–241. [CrossRef]
27. Bulot, F.M.J.; Russell, H.S.; Rezaei, M.; Johnson, M.S.; Ossont, S.J.J.; Morris, A.K.R.; Basford, P.J.; Easton, N.H.C.; Foster, G.L.; Loxham, M.; et al. Laboratory Comparison of Low-Cost Particulate Matter Sensors to Measure Transient Events of Pollution. *Sensors* **2020**, *20*, 2219. [CrossRef] [PubMed]
28. Christakis, I.; Tsakiridis, O.; Kandris, D.; Stavrakas, I. Air Pollution Monitoring via Wireless Sensor Networks: The Investigation and Correction of the Aging Behavior of Electrochemical Gaseous Pollutant Sensors. *Electronics* **2023**, *12*, 1842. [CrossRef]
29. Cross, E.S.; Williams, L.R.; Lewis, D.K.; Magoon, G.R.; Onasch, T.B.; Kaminsky, M.L.; Worsnop, D.R.; Jayne, J.T. Use of Electrochemical Sensors for Measurement of Air Pollution: Correcting Interference Response and Validating Measurements. *Atmos. Meas. Tech.* **2017**, *10*, 3575–3588. [CrossRef]
30. Ministry of Environment & Energy Greece. Air Pollution Measurement Data. Available online: https://ypen.gov.gr/perivallon/poiotita-tis-atmosfairas/dedomena-metriseon-atmosfairikis-rypansis/ (accessed on 20 December 2023).
31. *AAN 803-01*; Alphasense Application Note AAN 803-01 Correcting for Background Currents in Four Electrode Toxic Gas Sensors. Alphasense Limited: Essex, UK, 2014. Available online: https://zueriluft.ch/makezurich/AAN803.pdf (accessed on 16 November 2023).
32. Moltchanov, S.; Levy, I.; Etzion, Y.; Lerner, U.; Broday, D.M.; Fishbain, B. On the Feasibility of Measuring Urban Air Pollution by Wireless Distributed Sensor Networks. *Sci. Total Environ.* **2015**, *502*, 537–547. [CrossRef] [PubMed]
33. *ISO 9169:2006*; Air quality—Definition and Determination of Performance Characteristics of an Automatic Measuring System. Available online: https://www.iso.org/obp/ui/#iso:std:iso:9169:ed-2:v1:en (accessed on 6 December 2023).
34. Wastine, B. Essai d'Aptitude AirSensEUR du 12-janv au 22-fev 2018 Réalisé par Atmo Normandie Pour L'exercise D'intercomparaison n 1 du LCSQA. Available online: https://db-airmontech.jrc.ec.europa.eu/download/181114_ASE_ICP_1_v4.pdf (accessed on 6 December 2023).
35. Castell, N.; Dauge, F.R.; Schneider, P.; Vogt, M.; Lerner, U.; Fishbain, B.; Broday, D.; Bartonova, A. Can Commercial Low-Cost Sensor Platforms Contribute to Air Quality Monitoring and Exposure Estimates? *Environ. Int.* **2017**, *99*, 293–302. [CrossRef] [PubMed]
36. Karagulian, F.; Borowiak, A.; Barbiere, M.; Kotsev, A.; Van den Broecke, J.; Vonk, J.; Signorini, M.; Gerboles, M. *JRC Technical Reports. Calibration of AirSensEUR Boxes during a Field Study in the Netherlands*; Report no. JRC116324; European Commission: Luxembourg, 2020.
37. Wei, P.; Ning, Z.; Ye, S.; Sun, L.; Yang, F.; Wong, K.; Westerdahl, D.; Louie, P. Impact Analysis of Temperature and Humidity Conditions on Electrochemical Sensor Response in Ambient Air Quality Monitoring. *Sensors* **2018**, *18*, 59. [CrossRef]