Journal of Building Engineering 82 (2024) 108357

Contents lists available at ScienceDirect

Journal of Building Engineering

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

Validation of low-cost air quality monitoring platforms using model-based control charts

Mikael Boulic [a,*], Robyn Phipps [a,1], Yu Wang [a,2], Matthieu Vignes [b], Nurudeen A. Adegoke [a,3]

[a] School of Built Environment, Massey University, Auckland, 0632, New Zealand
[b] School of Mathematical and Computational Sciences, Massey University, Palmerston North 4410, New Zealand

ARTICLE INFO

Keywords:
Low-cost indoor air quality monitoring
Reference platform
Reliability
Shewhart and cumulative sum
Validation

ABSTRACT

The SARS COVID-19 pandemic highlighted the importance of routine indoor air quality (IAQ) monitoring. Recent advances in IAQ sensors and remote logging technologies offer opportunities to use low-cost platforms to monitor indoor air. The sensor's accuracy and stability are critical for reliable monitoring and health protection. Data from our low-cost IAQ platform (SKOMOBO) was validated against a commercial platform for carbon dioxide, temperature, and relative humidity measurements to test the reliability of the low-cost instrument. The traditional statistical method to test the variability between two data sets is the coefficient of determination method. We identified that this traditional method did not detect drifts in measurements, when comparing data from two platforms, in a controlled and uncontrolled environment. In our paper, we propose two complementary methods to detect potential drifts in measurements (a modified Shewhart method and a cumulative sum control chart method). The traditional coefficient of determination method indicated strong consistency (between 0.70 and 0.99) in the measurements between SKOMOBO and the reference platforms for both tested environments. Our more sensitive methods detected 100 % data matching for the controlled environment between the SKOMOBO and the reference platform but detected some drifts for the uncontrolled environment (between 81 % and 100 % data matching). It was expected that the uncontrolled environment would create more drifts in measurements than the controlled environment. Our new statistical methods achieved two important results; namely it advanced the validation process and proved the reliability of our low-cost platform for IAQ monitoring and assurance.

1. Introduction

The SARS-COVID-19 pandemic highlighted the criticality of Indoor Air Quality (IAQ) to protect the occupants from exposure to bioaerosols and indoor pollutants [3]. A significant limitation of commercially sourced instruments to measure IAQ is their high cost. Consequently, financial constraints can limit the number of commercially manufactured IAQ instruments that can practically be

* Corresponding author.
E-mail address: M.Boulic@massey.ac.nz (M. Boulic).
[1] Present address: Wellington Faculty of Architecture and Design Innovation, Victoria University of Wellington, 139 Vivian Street, Wellington, 6011, New Zealand.
[2] Present address: Building Research Association of New Zealand (BRANZ), 1222 Moonshine Road, Judgeford, Porirua, 5381, New Zealand.
[3] Present address: Melanoma Institute Australia, The University of Sydney, Sydney, Australia.

https://doi.org/10.1016/j.jobe.2023.108357
Received 17 August 2023; Received in revised form 11 December 2023; Accepted 18 December 2023
Available online 19 December 2023

installed in a building, or the length of time that these expensive instruments were installed. If monitoring is only conducted in a few locations of a building, or for short periods of time, then there is a significant risk of under-ventilated areas of the building or under-ventilated periods of time not being detected. Recent advancements in the development of low-cost carbon dioxide (CO_2) sensors, and remote monitoring technologies have opened the production and use of low cost IAQ monitoring platforms. Low-cost monitoring platforms can be developed for a fraction of the cost of commercially available reference platform. However, there is a need to inform on the reliability of the platform in monitoring our built environment [4]. Accurate and reliable IAQ data are crucial for estimating occupants' exposure to variables like temperature, relative humidity, and CO2. Temperature and relative humidity are important variables for defining thermal comfort in buildings. CO2 level is useful for estimating ventilation effectiveness [2] but it should not be considered as an overall indicator of IAQ [1]. Efficient ventilation protects building occupants from exposure to airborne viral aerosols such as SARS-COVID-19 [3]. Consequently, the accuracy and validation of the low-cost IAQ platforms is critical to provide reliable assessment of ventilation and health risks, when these low-cost platforms are used for routine monitoring of IAQ, or data quality control when the platforms are used for research purposes. As air quality can vary throughout the day, and in different parts of an indoor space, it is important to have a good distribution of sensors and have these installed for long periods of time to provide good data resolution.

In the literature, low-cost platforms are validated against calibrated commercial platforms using the coefficient of determination (R2) [5–7]. Ali et al. [6] compared the measurements from an in-house-made platform with a commercial equivalent, with R2 values ranging from 0.87 to 0.99. Penza et al. [7] also reported strong correlations between the measurements from the low-cost and the commercial platforms. However, the R2 value only measures the strength of the relationship between the measurements from the two types of platforms and fails to detect faults and drifts. Specifically, a perfect linear correlation between the two measurements does not indicate the existence of drifts in the measurements. Therefore, methods for detecting drifts in measurements are required to validate the accuracy of low-cost monitoring platforms.

To account for drift detection in the measurements between the low-cost and commercial platforms, we aim to propose new validation approaches inspired from control chart methods. This methodology tests the percentage of low-cost platform measurements within certain boundaries (an upper control limit and a lower control limit), calculated from the reference platform measurements. This methodology uses the Shewhart [8] and cumulative sum [9] procedures to detect large and small variations between measurements, respectively. Unlike classical control chart methods, our proposed method allows for serial correlations in the measurements and calculation of the boundaries. Any measurement from the low-cost platform that lies outside the boundaries proves that it differs from the reference platform. In addition, appropriate actions can track the cause of out-of-control points.

The proposed methodology assessed the performance of a low-cost platform manufactured at Massey University and called SKOMOBO. We will validate the CO2, relative humidity, and temperature measurements collected from the SKOMOBO against those from a commercial reference platform (TSI Q-Trak 7575, TSI Incorporated, Shoreview, Minnesota, USA). Detailed descriptions of the SKOMOBO platform, including the printed circuit board design, platform selection, and programming, were described by Wang et al. [10], Weyers et al. [11] and Wang et al. [12].

The remainder of this paper is organised as follows. Section 2 extends the classical charts (described in Section A of the Supplementary Method) to handle serial correlations in time-series datasets using two complementary-model-based control chart methodologies (the modified control chart and the residual-based control chart). When used jointly, these two methodologies complement each other and can determine the variation in the entire dataset. In Section 3, we describe the experimental designs for SKOMOBO validation. In Section 4, we apply the methodologies (developed in Section 2) to validate SKOMOBO measurements against reference platform measurements. Finally, Section 5 presents the conclusions of our study.

2. The proposed model-based methodologies for autocorrelated datasets

We developed innovative methodologies for the detection of anomalies within serially correlated datasets by employing classical control charts traditionally designed for independent data (see Supplementary Method, Section A). The inherent serial correlation within datasets is known to trigger an increased rate of false alarms when standard control chart techniques are used [13,14]. Therefore, it is imperative to address serial correlation prior to the application of control chart methods for such data [15]. Serial correlation emerging from the repeated sampling rate can be eliminated by sampling less frequently [16]. The disadvantage of this approach is that less information is available for the validation process, making it difficult to distinguish between measurement variations [17].

Here, we modelled serial correlations through the application of Autoregressive Integrated Moving Average (ARIMA) models, which are well-suited for time-series data [21,22]. ARIMA models excel in encapsulating diverse autocorrelation structures. By applying an ARIMA model to a dataset, we can separate the true signal (as represented by the model's fitted values) from stochastic noise (the residuals). Within the framework of control charts, these fitted values and the corresponding residuals (presumed to be free of autocorrelation if the ARIMA model is aptly specified) are monitored for indications of process deviation. This enables the identification of data shifts or heightened variability that deviates from the established autocorrelation pattern represented by the ARIMA model.

Our research developed two distinct model-based techniques to eliminate the impact of serial correlation on control chart efficacy [16–19]. The first technique involves a modified control chart procedure that compensates for the effects of serial correlations. This adjustment was accomplished by initially applying a suitable time-series model to encapsulate the autocorrelated structure of the dataset. The fitted values from the serially correlated datasets were then integrated into the classical Shewhart or cumulative sum control charts. The control limits of these charts were subsequently recalibrated to reflect the autocorrelation influence. This method

was designed to ensure the stability of a system that presents autocorrelation, marking a departure from the conventional Shewhart or cumulative sum charts which presuppose independent observations.

The second method employs residuals derived from time-series model fitting [16,20]. Similar to the first method, after the dataset is characterised for its autocorrelation through a time-series model, the residuals, which are expected to be uncorrelated if the model is well fitted, are utilised in Shewhart or cumulative sum charts. This approach guarantees that the chart is responsive to shifts in uncorrelated residuals, which are indicative of changes in the original autocorrelated measurements.

The proposed framework (Fig. 1) was executed in two phases. Phase I is dedicated to fitting the ARIMA model to reference measurements and estimating the model parameters, which were then used to establish control limits for the modified control charts. In Phase II, we employed predefined control charts to monitor SKOMOBO measurements, comparing the model-fitted values or residuals against the control limits set during Phase I. Observations of SKOMOBO measurements that exceed these predetermined limits are indicative of a potential measurement drift. Fig. 1 provides our robust framework for identifying drifts in serially correlated datasets.

This framework outlines a two-phase methodology for drift detection in time-series data characterised by serial correlation, integrating ARIMA modelling with Shewhart or cumulative sum control chart techniques. In Phase I (Model Fitting and Control Limit Estimation), the process begins with the acquisition of a serially correlated time-series dataset, followed by ARIMA modelling to capture the autocorrelation structure of the data, and then control limits are derived and stored using the model's parameters. Phase II (Real-time Monitoring and Drift Detection) involves the introduction of a new dataset for monitoring against the established control limits, analysis of each measurement for fitted values and residuals through the ARIMA model, and drift evaluation using a Monitoring Helper (Shewhart or cumulative sum chart) that assesses measurements against control limits. Measurements that deviated from the control limits were flagged as potential anomalies, and a comprehensive list of these anomalies was generated for further investigation.

2.1. Autoregressive Integrated Moving Average (ARIMA) models

For stationary time-series data with a constant true mean μ_0, a mixed ARIMA (p, d, q) model, where p is the autoregressive model term, d is the degree of differencing and q is the moving average term, is given as

$$Y_i = \mu_0 + \varphi_1 Y_{i-1} + \varphi_2 Y_{i-2} + \ldots + \varphi_p Y_{i-p}$$

$$+\varepsilon_i - \theta_1 \varepsilon_{i-1} - \theta_2 \varepsilon_{i-2} - \ldots - \theta_q \varepsilon_{i-q} \text{ for } i \in N. \quad \text{(Equation 1)}$$

Equation (1) may be written as $\varphi(B)(Y_i) = \mu_0 + \theta(B)\varepsilon_i$, where B is a backward shift operator given as $BY_i = Y_{i-1}$, $\varphi(B)$, and $\theta(B)$ are polynomials of degrees p and q, respectively. φ_i are the autoregressive model coefficients and θ_i represent the moving average coefficients. The quantity ε_i is a sequence of independent and identically distributed random variables with zero mean and variance σ_ε^2, i.e., $\varepsilon_i \sim N(0,\sigma_\varepsilon^2)$. The roots of $\varphi(.)$ lie outside the unit circle for a stationary time series. The orders of the parameters p and q of ARIMA (p, d, q) can be obtained by plotting the autocorrelation and partial autocorrelation functions, respectively, of the time-series dataset [23]. The autocorrelation function and partial autocorrelation function plots guide how the observations are interrelated and provide practical tools for identifying the appropriate models. In particular, the autocorrelation function plot helps to determine the order q of the moving average term. In contrast, the partial autocorrelation function plot helps identify the order p of the autoregressive terms of the model. Section B of the Supplementary Method includes the forms of the autoregressive model (1), moving average (1), and ARIMA (1,0,1), as well as their means and variances.

For a nonstationary time-series dataset (with a nonconstant mean function μ_t), it is important to make the series stationary by differencing successive observations before fitting an appropriate model to the differenced and stationary datasets. Stationarity can usually be achieved by taking successive differences in the observations one or more times [23]. The number of successive differences is denoted by d, and the notation $(1-B)^d Y_t$ indicates that d successive differences were used. The ARIMA models of the order (p,d,q) are as follows:

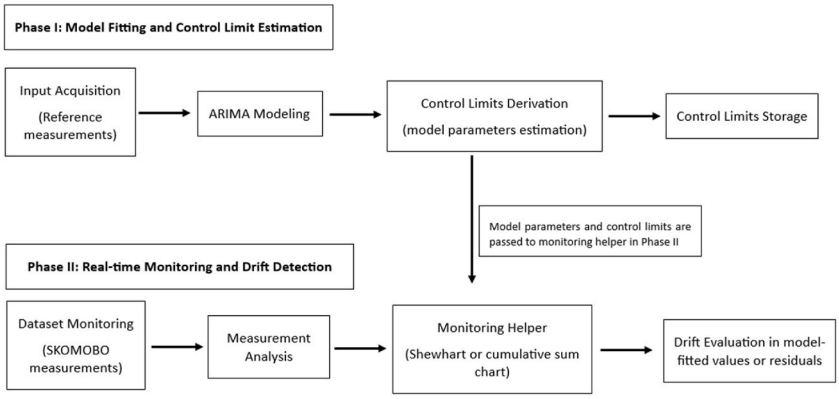

Fig. 1. The proposed two-phase drift detection method using ARIMA-based control charts.

$$\varphi(B)(1-B)^{d}(Y_i)=\mu_0+\varphi(B)\varepsilon_i \quad \text{(Equation 2)}$$

2.2. The modified control charts (method 1)

The modified control charts were obtained by adapting the classical chart limits to offset the impact of autocorrelation. We assumed that the serial correlation in the reference and new datasets could be modelled by an appropriate time-series ARIMA (p,d,q) model obtained from the reference dataset. We selected the parameters p and q from the autocorrelation and partial autocorrelation function plots of the time-series dataset. For the stationary series, $d=0$. However, d was chosen as the smallest number of successive differences for the nonstationary time-series data, making the series stationary.

Fig. 2 shows the two-phase analysis employed, namely Phases I and II. In Phase I:

1) We obtained an ARIMA model that characterised the reference dataset and then estimated the model parameters. The estimated parameters are mean ($\overline{Y}$) and standard deviation S.
2) The modified Shewhart and cumulative sum control chart limits were obtained from the model-estimated parameters, as follows:

i) The modified Shewhart lower control limit (LCL) and upper control limit (UCL) for monitoring serially correlated processes from these models are given by $LCL=\overline{Y}-Z_{1-\alpha/2}S$, and $UCL=\overline{Y}+Z_{1-\alpha/2}S$, respectively.

ii) Similarly, the upward and downward cumulative sum statistics are defined as $C_i^{+}=\max(0,C_{i-1}^{+}+(Y_i-\overline{Y}))-k)$, and $C_i^{-}=\max(0,C_{i-1}^{-}+(\overline{Y}-Y_i)-k)$, respectively, where $k=\frac{\delta}{2}S$.

In Phase II, the mean may shift from μ_0 to $\mu_0+\delta\sigma$.

1) We fitted the SKOMOBO measurements using the ARIMA model.
2) We compared the model-fitted values against the limits estimated from Phase I.
3) The charts provide signals whenever the fitted values are outside this limit.

Fig. 2 shows the proposed scheme based on the modified Shewhart and cumulative sum charts.

2.3. The residual-based control charts (method 2)

Here, monitoring was based on the residuals from a fitted time-series dataset. Residual control charts assume that a mean shift in a serially correlated dataset result in a mean shift in the residuals from the time-series model characterising the datasets. The residual values from a well-chosen time-series model are uncorrelated, thus providing theoretical evidence for using classical control charts designed for independent and identically distributed values. The residual-based chart is based on charting e_i given as

$$e_i=Y_i-\widehat{Y}_i$$

where $\widehat{Y}_i$ is the model-fitted value and is given as follows:

$$\widehat{Y}_i=\widehat{\mu}_0+\widehat{\varphi}_1Y_{i-1}+\widehat{\varphi}_2Y_{i-2}+\ldots+\widehat{\varphi}_pY_{i-p}$$

$$+\varepsilon_i-\widehat{\theta}_1\varepsilon_{i-1}-\widehat{\theta}_2\varepsilon_{i-2}-\ldots-\widehat{\theta}_q\varepsilon_{i-q},\ i\in N. \quad \text{(Equation 3)}$$

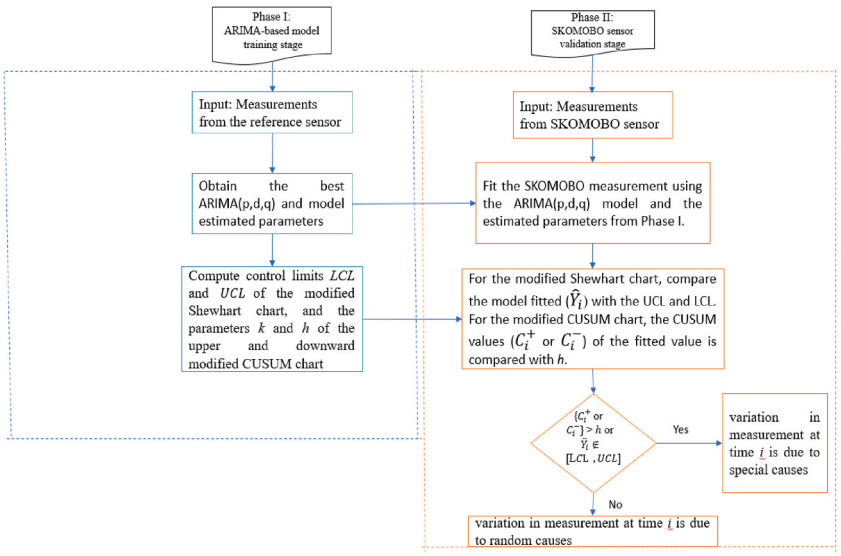

Fig. 2. Proposed scheme based on modified Shewhart and cumulative sum (CUSUM) charts (LCL: lower control limit; UCL: upper control limit).

where $\widehat{\mu}_0$, the estimated autoregressive model parameters $\{\widehat{\varphi}_1, \widehat{\varphi}_2, \ldots, \widehat{\varphi}_p\}$, and the estimated moving average parameters $\{\widehat{\theta}_1, \widehat{\theta}_2, \ldots, \widehat{\theta}_q\}$ are obtained from the model based on the reference dataset. The quantities e_i satisfy $e_i = Y_i - \widehat{Y}_i$, which is the realisation of the residual $\varepsilon_i \sim N(0, \sigma_\varepsilon^2)$.

The residual-based chart was employed in a two-phase procedure. In Phase I:

1) We obtained a model that characterised the reference dataset.
2) The residual e_i used to validate the measurement at time i was obtained as the difference between the true and model-fitted values of the reference measurements. The estimated value of σ_ε, denoted by, $\widehat{\sigma}_\varepsilon$, was calculated from the residual.
3) The control chart limits were obtained from the estimated residual values as follows:

i. The residual based Shewhart control chart limits: lower control limit and upper control limit are given as $LCL = -Z_{1-\alpha'2}\widehat{\sigma}_\varepsilon$, and $UCL = Z_{1-\alpha'2}\widehat{\sigma}_\varepsilon$, respectively.
ii. Similarly, the upward and downward cumulative sum statistics are defined as $C_i^+ = \max(0, C_{i-1}^+ + e_i - k)$, and $C_i^- = \max(0, C_{i-1}^- - e_i - k)$, respectively, where $k = \frac{\delta}{2}\widehat{\sigma}_\varepsilon$.

In Phase II, where the mean might have shifted from μ_0 to $\mu_0 + \delta\sigma_\varepsilon$:

1) We fitted the SKOMOBO measurements using the model-estimated parameters.
2) The fitted model values and residuals were obtained.
3) We compared the model residual values against the control limits obtained from Phase I.
4) The chart provides signals whenever residual values are outside the limit.

The proposed scheme based on residual Shewhart, and cumulative sum (CUSUM) charts is presented in Fig. 3.

3. Data collection and work validation designs

Co-location tests were conducted before (pre-fieldwork co-location) and after (post-fieldwork co-location) a two-year deployment period to validate the SKOMOBO platform against two reference commercial platforms (TSI Q-Trak 7575).

3.1. Pre-fieldwork co-location test

During the pre-fieldwork preparation, six SKOMOBOs were co-located with two reference platforms: 1) a controlled environment and 2) an uncontrolled environment. The controlled environment consisted of a small unoccupied office. All platforms were located on a desk and logged for CO_2, temperature, and relative humidity levels every minute for two days (Fig. 4). The office door was locked, and the window was closed during the experiment. Visits from the researcher for equipment verification were logged.

The second validation (uncontrolled environment) occurred in a lobby adjacent to seven offices. This lobby was used primarily from 8 a.m. to 6 p.m. on weekdays by seven researchers and their visitors undertaking office activities. This lobby had no window but was air-conditioned from 8 a.m. to 5 p.m. on weekdays. Six SKOMOBOs and two reference platforms were located on a table in the middle of the lobby (Fig. 5), and CO_2, temperature, and relative humidity were monitored every minute for three days.

3.2. Post-fieldwork co-location test

Following a two-year deployment, nine SKOMOBO platforms used in the two-year monitoring were co-located with three reference

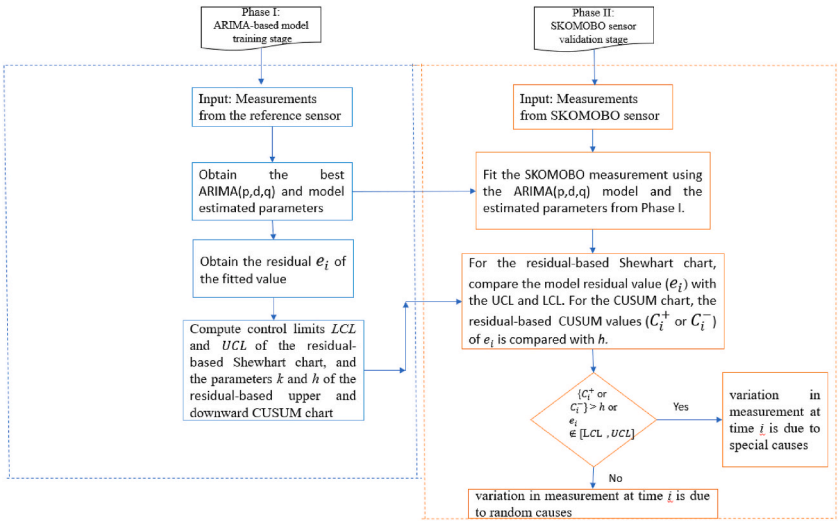

Fig. 3. Proposed scheme based on the residual-based Shewhart and cumulative sum (CUSUM) charts (LCL: lower control limit; UCL: upper control limit).

Fig. 4. Experimental setup for the controlled location of the pre-fieldwork validation study of SKOMOBO platforms with reference platforms. Six SKOMOBO sensors were enclosed in a clear acrylic frame and reference platforms (two TSI Q-Trak 7575, TSI Incorporated, Shoreview, Minnesota, USA) co-located in controlled office environments with a common area.

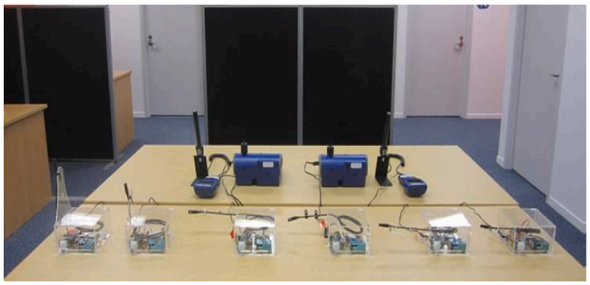

Fig. 5. Experimental setup for the uncontrolled location of the pre-fieldwork validation study of SKOMOBO platforms with reference platforms. Six SKOMOBO sensors enclosed in a clear acrylic frame and reference platforms (two TSI Q-Trak 7575, TSI Incorporated, Shoreview, Minnesota, USA) co-located uncontrolled office environments with a common area.

platforms (TSI Q-Trak 7575, TSI Incorporated, Shoreview, Minnesota, USA) in an unoccupied seminar room. The SKOMOBOs and reference platforms monitored the CO_2, temperature, and relative humidity levels every minute for two days. The seminar room was not locked, but the windows remained closed. All the tested SKOMOBOs in the pre- and post-fieldwork co-location tests were manufactured using the same sensors, design, and manufacturing processes. All the reference platforms used in the pre- and post-fieldwork testing were purchased simultaneously and factory-calibrated.

4. Validation of SKOMOBO measurements

We compared the measurements of temperature, relative humidity, and CO_2 from the SKOMOBOs (pre- and post-fieldwork co-location tests) with those from the reference platforms using 1) the coefficient of determination (R^2); the classical testing method reported in the literature, and 2) our proposed statistical methodologies described in Section 2.

4.1. Comparing SKOMOBO and the reference platform data using the coefficient of determination - R^2

Supplementary Tables 1 and 2 present the coefficient of determination values of CO2, temperature, and relative humidity measurements obtained from the SKOMOBOs and reference platforms for pre-and post-fieldwork environments, respectively. These R2 values demonstrated a strong correlation between the two measurement platforms. For the pre-fieldwork-controlled environment, the R2 values ranged between 0.996 and 0.999 for CO2, 0.988 and 0.997 for temperature, and 0.946 and 0.978 for relative humidity. In the pre-fieldwork-uncontrolled environment, they varied from 0.703 to 0.942 for CO2, 0.928 to 0.992 for temperature, and 0.772 to 0.974 for relative humidity. For the post-fieldwork scenario, the values ranged between 0.801 and 0.924 for CO2, 0.993 and 0.999 for temperature, and 0.992 and 0.999 for relative humidity. These findings suggest that the SKOMOBO platform produced reliable and consistent CO2, temperature, and relative humidity measurements compared to the reference platform measurements.

However, it is crucial to recognise that the coefficient of determination values primarily indicate the degree of linear association between the SKOMOBO measurements and the reference platform measurements without providing any insight into the accuracy of the SKOMOBO measurements in replicating the reference platform measurements. Thus, the coefficient of determination did not directly validate the magnitude of the SKOMOBO measurements against the reference platform readings. Supplementary Fig. 1a and 1b shows the coefficients of determination for the simulated random samples Y1 and Y2, respectively. Although the coefficient of determination was 1, indicating perfect linear correlation, there was a noticeable difference in the magnitude of Y2. Specifically, Supplementary Figure 1a presents Y2 derived by adding 0.03 to each observation in Y1, while Supplementary Figure 1b displays Y2 obtained by adding 5 to each observation in Y1. This demonstrates that even a perfect linear relationship can have significant variance in the actual magnitudes. This highlights the limitations of using the coefficient of determination values as the sole means of validation

because they do not necessarily reflect the congruence in measurement magnitudes between SKOMOBOs and reference platforms. These results showed that additional requirements are required to effectively detect drifts and validate low-cost platform reliability for monitoring IAQ variables.

4.2. Comparing SKOMOBO and the reference platforms data using the proposed control charts methodologies

Using the methodology described in Section 2, we investigated the ability of SKOMOBO to provide measurements of temperature, relative humidity, and CO_2 that were within the limits obtained from the reference platforms and provided some guidance on the percentages of the SKOMOBO measurements within limits. The limits are calculated such that if the SKOMOBO measurements lie within them, it is assumed that the difference between the SKOMOBO and reference platform measurements results from unavoidable causes inherent to the platforms. Such random variations might be caused by unimportant fluctuations or background noise in platforms that cannot be eliminated.

The proposed methodology was employed in a two-phase procedure. In Phase I, the reference platform measurements were used as training sets to obtain an appropriate time-series model. The model building starts by applying the augmented Dickey-Fuller test (ADF) [24] and the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) tests [25] to verify the stationarity of the measurements. The ADF and KPSS tests showed that the series was stationary and did not require trend differencing. Hence, we considered $d = 0$ in all the ARIMA models examined.

We investigated different autoregressive and moving average orders p and q, respectively. Following Murat et al. [23], we considered p and q in the range [0, 3]. We examined the potential of different ARIMA models using autocorrelation and partial autocorrelation function plots to obtain the best time series model. We chose the model that minimised the mean absolute error (MAE), root mean square error (RMSE), and Bayesian information criterion (BIC) as the best [23,26–28]. The MAE and RMSE are defined in Qi et al. [29].

$$MAE = \frac{1}{n}\sum_{i=i}^{n}|e_i|$$

and

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=i}^{n}e_i^2}$$

respectively, where n is the number of monitoring periods, and $e_i = Y_i - \widehat{Y}_i$ is the difference between the true (Y_i) and model-fitted values ($\widehat{Y}$) of the ith measurement. These quantities have the same measurement unit as the measured datasets. BIC [26,30] is expressed as follows:

$$BIC = klog(n) + nlog\left(1 - R^2\right)$$

where k is the number of parameters that are fitted in the model, log(.) is the natural logarithm, n is the number of observations in the series, and SSE is the sum of squared errors, given as $SSE = \sum_{i=i}^{n} e_i^2$.

Supplementary Tables 3 and 4 list the reference platform measurement models and their estimated parameters for the pre- and post-fieldwork locations, respectively. The SKOMOBO and the reference platform measurements were normalised using square-root transformations. We found consistency in the choice of ARIMA models for the measurements of reference platforms 1 and 2 in both pre- and post-fieldwork experiments. For example, the best ARIMA models for CO_2 measurements of the reference platforms 1 and 2 for the pre-fieldwork controlled environment experiment were with parameters $\widehat{\varphi}_1 = 0.6873$, $\widehat{\mu}_0 = 30.8437$, and $\widehat{\sigma}_\varepsilon^2 = 21.11428$) and with parameters $\widehat{\varphi}_1 = 0.6825$, $\widehat{\mu}_0 = 30.9083$, and $\widehat{\sigma}_\varepsilon^2 = 21.090$, respectively. Also, the best ARIMA models for the CO_2 measurements of the reference platforms 1 and 2 for the pre-fieldwork uncontrolled environment experiment were with parameters $\widehat{\varphi}_1 = 1.1813$, $\widehat{\varphi}_2 = -0.3717$, $\widehat{\mu}_0 = 21.2472$, and $\widehat{\sigma}_\varepsilon^2 = 0.1181$ and with parameters $\widehat{\varphi}_1 = 1.071$, $\widehat{\varphi}_2 = -0.2427$, $\widehat{\mu}_0 = 21.5345$, and $\widehat{\sigma}_\varepsilon^2 = 0.1408$ respectively.

Using the model-estimated parameters in Supplementary Tables 3 and 4 and the fitted model residuals, the control limits of the modified and residual-based charts used to validate the fitted and residual values of the SKOMOBO platform, respectively, were obtained for use in Phase II. The limits were calculated from the distributional properties of the fitted and residual values of the reference platform measurements. For example, the modified Shewhart chart's lower control limit (LCL) and upper control limit (UCL) of the CO_2 fitted values with parameters $\widehat{\varphi}_1 = 0.6873$, $\widehat{\mu}_0 = 30.8437$, and $\widehat{\sigma}_\varepsilon^2 = 21.11428$ of the reference platform 1 and the pre-fieldwork controlled environment are given as $LCL = \widehat{\mu}_0 - Z_{1-\alpha'2}\sqrt{\frac{\widehat{\sigma}_\varepsilon^2}{(1-\varphi_1^2)}}$, and $UCL = \widehat{\mu}_0 + Z_{1-\alpha'2}\sqrt{\frac{\widehat{\sigma}_\varepsilon^2}{(1-\varphi_1^2)}}$, respectively, where, $Z_{1-\alpha'2}$ is the $(1-\alpha'2)^{\text{th}}$ quantile point of standard normal distribution. By substituting $\alpha = 0.0027$, $\widehat{\mu}_0$, $\widehat{\varphi}_1$ and $\widehat{\sigma}_\varepsilon^2$ into the LCL and UCL equations, the LCL and UCL used to validate the fitted value in Phase II were 12.023 and 49.976, respectively. Similarly, the upward and downward CUSUM statistics are defined as $C_i^+ = \max(0, C_{i-1}^+ + (\widehat{Y}_i - \widehat{\mu}_0) - k)$, and $C_i^- = \max(0, C_{i-1}^- + (\widehat{Y}_i - \widehat{\mu}_0) - k)$, respectively, where $k = k_1\sqrt{\frac{\widehat{\sigma}_\varepsilon^2}{(1-\varphi_1^2)}}$, and $\widehat{Y}_i$ is the ith fitted value of the SKOMOBO platform measurement.

The LCL and UCL of the residual-based Shewhart chart for the reference platform 1 (pre-fieldwork-controlled environment) are

given as $LCL = -Z_{1-\alpha/2}\widehat{\sigma}_{\varepsilon}$, and $UCL = Z_{1-\alpha/2}\widehat{\sigma}_{\varepsilon}$, respectively. By substituting the values of $\widehat{\sigma}_{\varepsilon}^2 = 21.114$ and $\alpha = 0.0027$ in the LCL and UCL equations, the LCL and UCL used to monitor the residuals of the fitted values in Phase II were - 13.784 and 13.784, respectively. Similarly, the upward and downward CUSUM statistics are defined as $C_i^+ = \max(0, C_{i-1}^+ + \widehat{e}_i - k)$, and $C_i^- = \max(0, C_{i-1}^- - \widehat{e} - k)$, respectively, where $k = k_1\sqrt{\widehat{\sigma}_{\varepsilon}^2}$ and $\widehat{Y}_i$ is the residual of the ith fitted values of the SKOMOBO measurements. For both the modified and residual cumulative sum charts, we chose $k_1 = 0.125$ and used a binary search algorithm similar to that used in Refs. [31,32] to obtain the value of h that fixes the in-control average run-length (ARL) performance of the charts at 370. We refer the interested reader to Section C of the Supplementary Method for a detailed simulation study of the chart run-length performance and an explanation of the false alarm signal after every 370 samples.

In Phase II, the measurements from the SKOMOBO platforms were fitted using time-series models that characterised the measurements from the reference platforms given in Supplementary Tables 3 and 4 The fitted and residual values were then compared against the limits. Figs. 6–8 show the modified Shewhart control charts of the six SKOMOBO platform hourly averaged CO_2, temperature, and relative humidity measurements for the pre-fieldwork-controlled environment experimental study compared to the boundaries calculated from the measurements of the reference platform 1. We included the fitted line (or values) of the reference platform 1 measurements to show that all the platform-fitted values were within the boundaries calculated from the estimated parameters of the fitted line. Our interest lies in comparing the fitted values from the SKOMOBO platform measurements with the control limits of the charts.

As shown in Figs. 6–8, none of the fitted points of the SKOMOBO platform was outside the boundaries calculated from the reference measurements for all six SKOMOBO platforms. The modified Shewhart chart results demonstrated no variation between the measurements from these platforms for the pre-fieldwork controlled-environment experimental study. Figs. 6–8 show that the variations between the measurements from the SKOMOBO and the reference platforms were due to chance or random variations alone.

We also investigated the small and persistent variations between the platform measurements by comparing the measurements from the six SKOMOBO platforms and the reference platform 1 (controlled environment experimental study) using the modified cumulative sum charts. Supplementary Figures 2–4 show the modified cumulative sum charts of the six SKOMOBO platform hourly averaged CO_2, temperature, and relative humidity measurements. As shown in Supplementary Figures 2 and 3, the lower and upper cumulative sum statistics of the SKOMOBO fitted points for the CO_2 and temperature variables are within the upper limits calculated from the reference

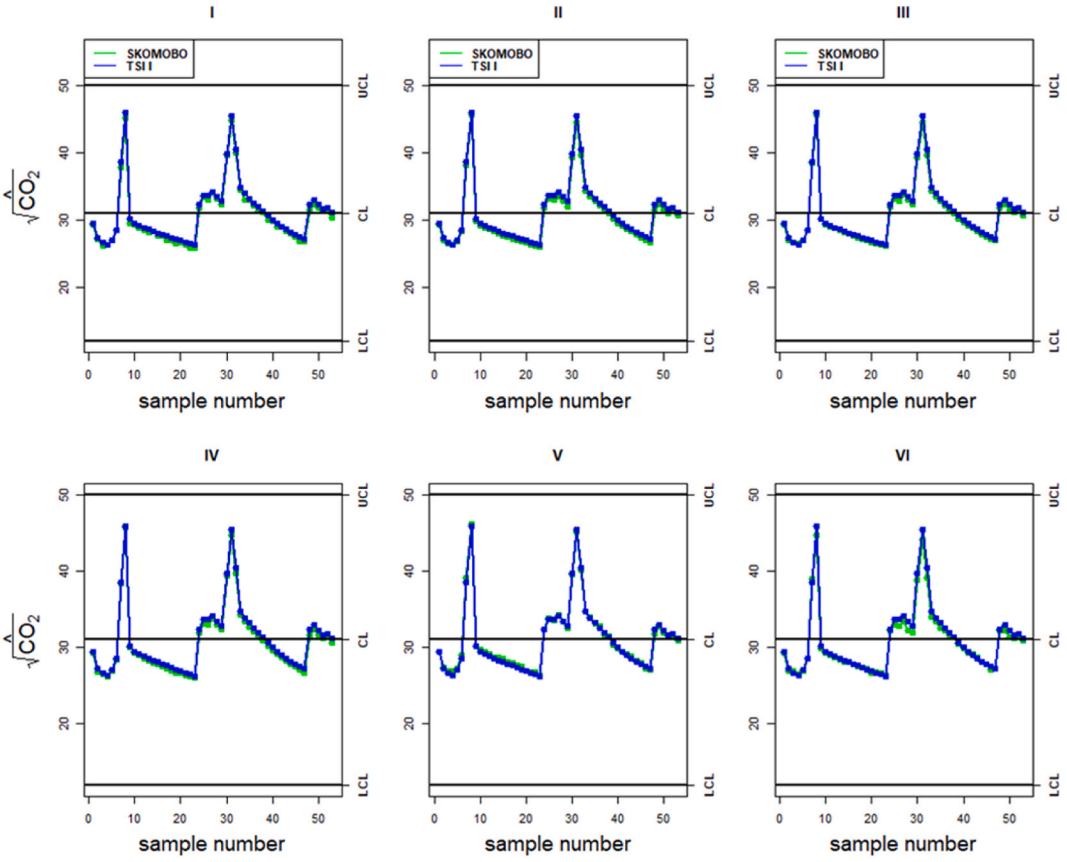

Fig. 6. The modified Shewhart charts of the hourly average CO_2 measurements of the six SKOMOBO platforms for the pre-fieldwork experimental-controlled environment compared to the boundaries calculated from the reference platform 1 measurements (LCL: lower control limit; CL: control limit; UCL: upper control limit).

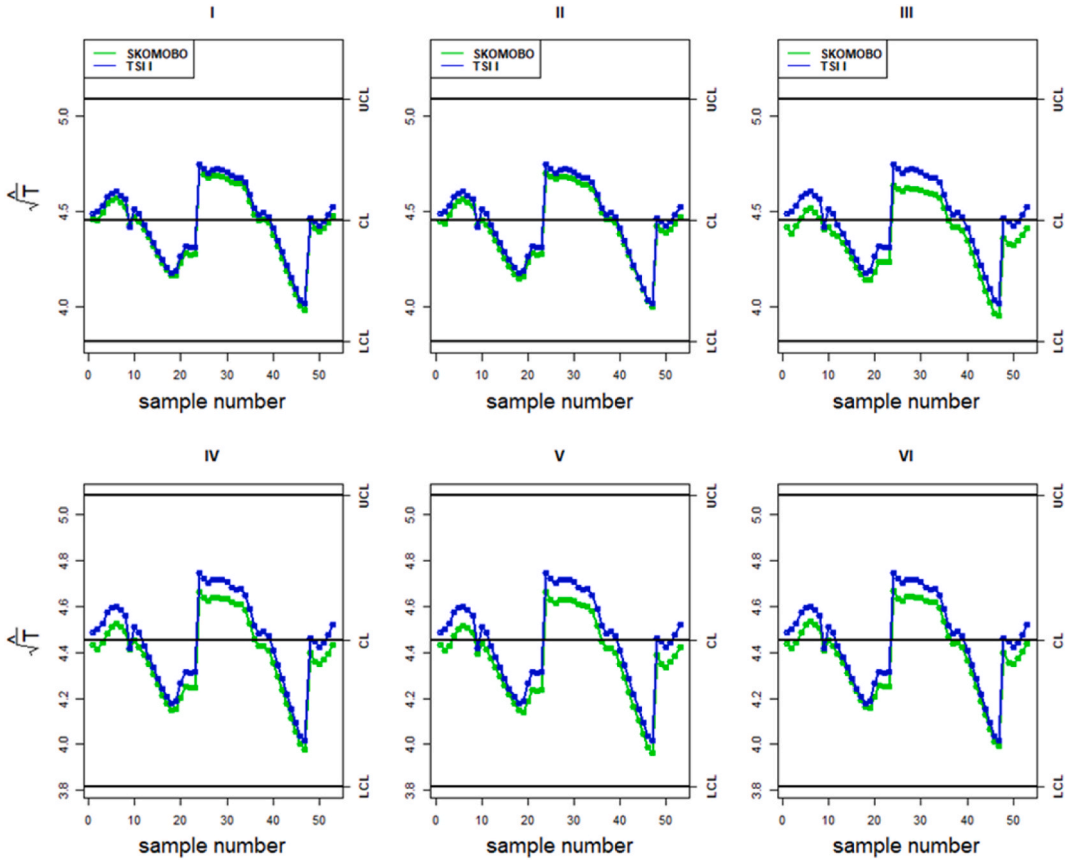

Fig. 7. The modified Shewhart charts of the hourly average temperature measurements of the six SKOMOBO platforms for the pre-fieldwork experimental-controlled environment compared to the boundaries calculated from the reference platform 1 measurements (LCL: lower control limit; CL: control limit; UCL: upper control limit).

platform measurements. Hence, the modified cumulative sum charts also showed no variation between the platforms.

In contrast, the SKOMOBO fitted points for the relative humidity (Supplementary Figure 4) provided some out-of-control points for the four SKOMOBO platforms. These demonstrate that the variations differed from the background noise between the relative humidity measurements of the reference platforms and SKOMOBO for the pre-fieldwork controlled experimental study. The modified cumulative sum charts for the relative humidity variable (Supplementary Figure 4) showed small variations between the measurements from the SKOMOBO platforms and the reference platform 1, which were not detected by the modified Shewhart chart (Fig. 8) because of their inability to detect small changes.

The residual-based Shewhart charts (Supplementary Figures 5–7) and cumulative sum charts (Supplementary Figures 8–10) for the CO_2, temperature, and relative humidity measurements (from the pre-fieldwork-controlled environment experiment), when compared to the limits calculated from the fitted values of the reference platform 1 measurements, showed no points outside the limits.

Tables 1 and 2 provide the average percentage of points from the SKOMOBO platforms within the boundaries of the modified and residual-based (Shewhart and cumulative sum) charts for the pre- and post-fieldwork locations, respectively. The average percentage was obtained by first calculating the percentages (i.e., true positive percentages) of SKOMOBO measurements within the chart boundaries from all SKOMOBO platforms and then obtaining the average of the percentages. For example, the average percentage of SKOMOBO platform CO_2 measurements in Fig. 6 is 100 %, which indicates that the modified Shewhart charts of all SKOMOBO platform CO_2 measurements (from the pre-fieldwork-controlled environment test) show no variation between the CO_2 measurements of the SKOMOBO platforms and the reference platform 1, in addition to some random variations. This implies that the modified Shewhart chart provides a 100 % match between the CO_2 measurements from the SKOMOBO platforms and reference platform 1 (pre-fieldwork-controlled environment experimental study).

The results of the modified Shewhart charts for the pre- and post-fieldwork location experiments (Tables 1 and 2) show that on average, there are more than 93 % matches between the CO_2 measurements, more than 99 % matches between the temperature measurements, and 100 % matches between the relative humidity measurements of the SKOMOBO platforms and the reference platforms. On average, the modified cumulative sum charts (Tables 1 and 2) show more than 81 % matches between the CO_2 measurements, more than 85 % matches between the temperature measurements, and more than 85 % matches between the relative humidity measurements of the SKOMOBO platforms and the reference platforms.

Similarly, the residual-based Shewhart charts for the post-fieldwork environment experiments (Tables 1 and 2) showed that on

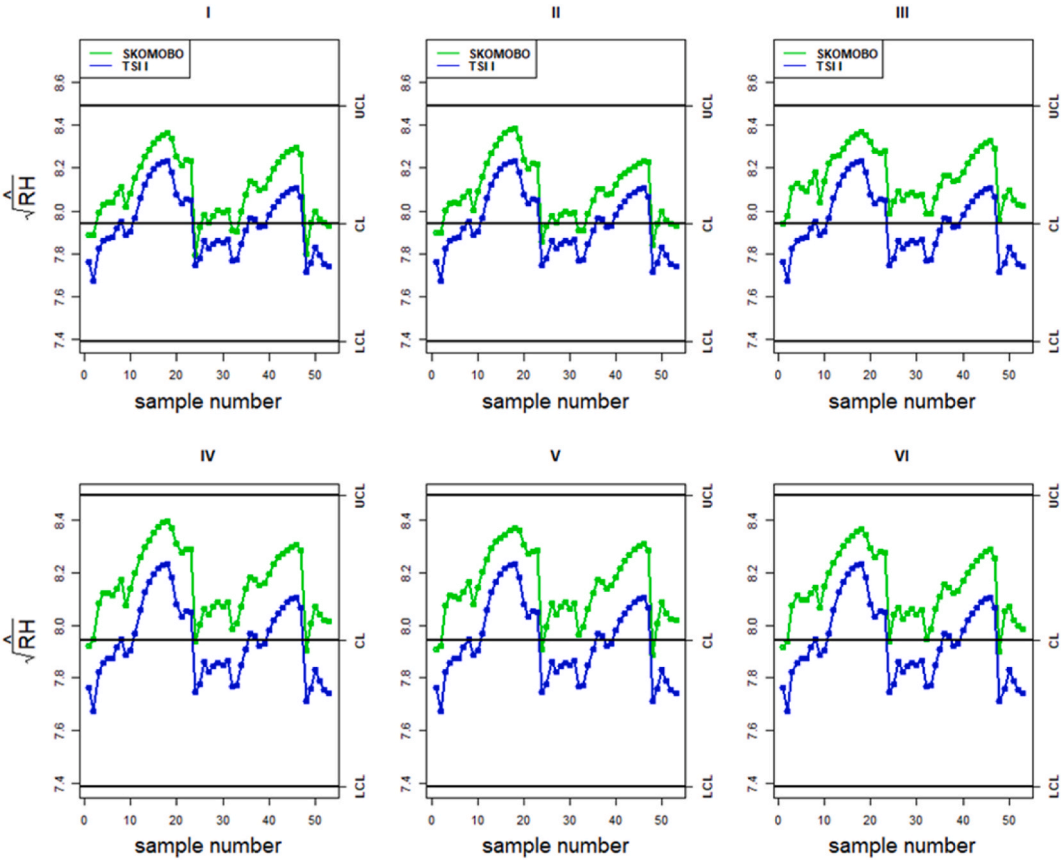

Fig. 8. The modified Shewhart charts of the hourly average relative humidity measurements of the six SKOMOBO platforms for the pre-fieldwork experimental-controlled environment compared to the boundaries calculated from the reference platform 1 measurements (LCL: lower control limit; CL: control limit; UCL: upper control limit).

Table 1
The average percentage of measurements from the SKOMOBO sensors that lie within the boundaries of the modified and residual-based Shewhart and cumulative sum (CUSUM) charts for both controlled and uncontrolled environments of the pre-fieldwork experimental study.

			Locations					
			Controlled			Uncontrolled		
Methods	Charts	Reference platforms	CO_2 (%)	T (%)	RH (%)	CO_2 (%)	T (%)	RH (%)
Modified-chart Scheme	Shewhart	RP1	100	100	100	97.3	99.9	100
		RP2	100	100	100	98.1	100	100
	CUSUM	RP1	100	100	90.8	81.6	85.8	85.3
		RP2	100	100	100	82.8	88.4	89.8
Residual-based Scheme	Shewhart	RP1	100	100	100	98.4	96.7	98.3
		RP2	100	100	100	98.4	97.3	98.3
	CUSUM	RP1	100	100	100	80.6	100	85.6
		RP2	100	100	100	82.1	100	81.2

average, more than 94 % matches between the CO_2 measurements, more than 96 % matches between the temperature measurements, and more than 97 % matches between the relative humidity measurements of the SKOMOBO platforms and the reference platforms. The residual-based cumulative sum charts (Tables 1 and 2) showed, on average, more than 80 % matches between the CO_2 measurements, 100 % matches between the temperature measurements, and more than 81 % matches between the relative humidity measurements of the SKOMOBO platforms and those of the reference platforms.

The results demonstrate the ability of the proposed methodologies to validate measurements from low-cost platforms against those from the reference platforms in terms of the percentage of low-cost platform measurements that lie within limits calculated from the distributional properties of the fitted values or residuals of the fitted values of the reference platforms.

Table 2
The average percentage of measurements from the SKOMOBO sensors that lie within the boundaries of the modified and residual-based Shewhart and cumulative sum (CUSUM) charts for post-fieldwork experimental studies.

Methods	Charts	Reference platforms	CO_2 (%)	T (%)	RH (%)
Modified-chart Scheme	Shewhart	RP1	97.3	100	100
		RP2	97.7	100	100
		RP3	93.8	100	100
	CUSUM	RP1	93.5	100	100
		RP2	95.6	100	95.5
		RP3	100	100	100
Residual-based Scheme	Shewhart	RP1	95.3	99.9	97.8
		RP2	95.4	99.9	97.6
		RP3	94.5	99.9	97.6
	CUSUM	RP1	98.7	100	100
		RP2	99.0	100	86.2
		RP3	100.0	100	100

4.3. Summary of the results

4.3.1. Results from the coefficient of determination method (Classical testing method):

- For the pre-fieldwork-controlled environment, the values for the coefficient of determination were above 0.946, 0.988 and 0.996 for the relative humidity, temperature, and CO_2 respectively,
- For the pre-fieldwork-uncontrolled environment, the values for the coefficient of determination were above 0.703, 0.772 and 0.928 for the CO_2, relative humidity, and temperature respectively.
- For the post-fieldwork environment, the values for the coefficient of determination were above 0.801, 0.992 and 0.993 for the CO_2, relative humidity, and temperature respectively,
- These values demonstrated a strong linear association between the two measurement platforms (SKOMOBO vs reference platform), but do not reflect the congruence in measurement magnitudes between SKOMOBOs and the reference platforms. This highlights the limitations of using the traditional statistical method, namely the coefficient of determination values, as the sole means of platform validation. We identified that a modified Shewhart method and a cumulative sum control chart method (reported below) were more sensitive and were able detect drift between the data sets.

4.3.2. Results from our proposed control chart methodologies

- As expected, the pre-fieldwork-controlled environment showed the best results. The modified Shewhart chart (to detect large variations) showed no variations (i.e., 100 % matching data) between SKOMOBO measurements and the reference platform measurements for CO_2, temperature or relative humidity. When using the modified cumulative sum chart (to detect small and persistent variations), only the relative humidity measurements showed small variations which were still falling in the acceptable limits defined by the reference platforms. No small variations were detected for the CO_2 or temperature measurements.
- Regarding the uncontrolled environment (pre fieldwork or post fieldwork), there were more fluctuations; however, the modified Shewhart charts showed more than 93 %, 99 % and 100 % matches for the CO_2, temperature, relative humidity measurements respectively when comparing the SKOMOBO platforms and the reference platforms. The modified cumulative sum charts show more than 81 %, 85 % and 85 % matches for the CO_2, temperature and relative humidity measurements when comparing the SKOMOBO platforms and the reference platforms.

Validating that the low-cost IAQ platform is reliable and accurate, enables building users and managers to confidently use these sensors in a larger number of areas of their building to routinely assess the quality and healthiness of indoor air, without the financial constraints incurred when commercial platforms are used. This provides a higher resolution of the data and reduces the potential of not detecting of a space where the ventilation system is malfunctioning. Remoted logging of data overcomes the limitation of running out of data storage, or the frequent need to download data and clear the memory of platforms with only fixed data storage capability. This is an important feature where these instruments are being used in longitudinal research studies, to minimise disruption to study participants.

5. Conclusions

Using a new methodology, we assessed the accuracy of a low-cost IAQ monitoring platform (SKOMOBO) by validating its CO_2, temperature, and relative humidity measurements with measurements from a factory-calibrated commercial reference (TSI Q-Trak 7575, TSI Incorporated, Shoreview, Minnesota, USA) platform. The SKOMOBO measurements were obtained from experimental studies before (pre-fieldwork) and after (post-fieldwork) a two-year continuous measurement of IAQ in primary school environments in different regions of New Zealand.

We first analysed the relationships between the measurements from SKOMOBO and the reference platforms using the coefficient of determination. The results showed that the SKOMOBO platforms provided reliable measurements of CO_2, temperature, and relative

humidity, consistent with the measurements from the reference platforms. However, the coefficient of determination failed to detect any faults and drifts in the measurements.

We then used our proposed methodology. We provided the average percentage of measurements from the SKOMOBO platforms within the boundaries of the modified or residual-based (Shewhart and cumulative sum) charts for pre-and post-fieldwork location experimental studies. The results showed that the modified or residual-based Shewhart charts provided more than 93 % matches between the SKOMOBO platforms and the reference platforms for CO_2, temperature, and RH relative humidity measurements, indicating a high agreement regarding the magnitudes of the measurements of the two platforms. The cumulative sum-based charts provide more than 80 % matches between the measurements from the platforms, indicating some small and persistent differences between the measurements from SKOMOBO and the reference platforms. The cumulative sum-based charts show the highest variation between the measurements of the platforms in the pre-fieldwork uncontrolled environment. We consider that the difference between the measurements of the two platforms from this location could result from many uncontrolled features such as airflow disturbance due to office occupant activities. However, further studies on the performance of the SKOMOBO platform in uncontrolled environments are required for a detailed understanding of platform responses to different uncontrolled environmental situations.

Recent advances in IAQ sensors and remote logging technologies offer opportunities to use low-cost platforms to monitor indoor air. The sensor's accuracy and stability are critical for reliable monitoring and health protection. Our study demonstrated that our proposed methods are reliable in detecting measurement drifts to support a low-cost platform validation against a reference platform. This study highlighted the limitations of using the coefficient of determination values as the sole means of validation. These findings have direct practical implications for the IAQ community as our innovative methods will serve as guidance to validate any monitoring platform. We recommend using our rigorous analytical approach in addition to the traditional coefficient of determination validation approach (which only detects the strength of the relationship between the measurements but fails to detect and drifts in measurements).

The validation of the low-cost platform will enable building users and managers to confidently rely on low-cost platforms to monitor the quality of the air in multiple spaces and for longer periods in their building. Compared to a commercial platform, using a well-validated, low-cost platform will potentially increase the monitoring capacity for a similar budget. Our research results will positively influence the use of reliable low-cost platform and empower the field of IAQ monitoring.

The SKOMOBO platform is currently the preferred monitoring platform across several Massey University research projects. For instance, we deployed SKOMOBOs in seven New Zealand passive houses to investigate the indoor climate achieved during the winter season and the summer season (risk of overheating). Another research project is researching the impact of New Zealand classrooms on students' cognitive performance.

In the future, we could increase the number of monitored variables of the SKOMOBO platform to investigate our indoor environment better. The recent advancements in data processing showed that it could be possible to include some artificial intelligence technology like forecasting in the SKOMOBO platform. This process could use the previously monitored concentration of CO_2 to predict the following one and inform on the need to open a window, for instance. This approach could assist in designing an alerting system based on the forecasted values and their compliance with different metrics. This will contribute to decision-making for real-world applications.

Novelty about our research

The results of this study demonstrate the effectiveness of the proposed model-based methodologies in detecting drifts in low-cost platforms and validating the reliability of the low-cost platform for monitoring indoor air quality variables (carbon dioxide, temperature and relative humidity).

CRediT authorship contribution statement

Mikael Boulic: Conceptualization, Data curation, Funding acquisition, Methodology, Project administration, Resources, Supervision, Validation, Writing – original draft, Writing – review & editing. **Robyn Phipps:** Conceptualization, Funding acquisition, Methodology, Supervision, Writing – review & editing. **Yu Wang:** Conceptualization, Data curation, Methodology, Validation, Writing – review & editing. **Matthieu Vignes:** Formal analysis, Visualization, Writing – review & editing. **Nurudeen A. Adegoke:** Conceptualization, Formal analysis, Methodology, Software, Validation, Visualization, Writing – original draft, 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.

Data availability

Data will be made available on request.

Acknowledgements

This SKOMOBO development was supported by a New Zealand Lottery Health Research Equipment Grant (#LHR-2019-102271) and a Building Research Association of New Zealand Research Grant (#LR0513).

The authors would like to thank Ryan Weyers and Alfred Moses for their technical support during the SKOMOBO development. The authors would also like to thank the anonymous reviewers for their valuable comments that improved the manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jobe.2023.108357.

References

[1] A. Persily, Quit blaming ASHRAE standard 62.1 for 1000 ppm CO_2, in: The 16th Conference of the International Society of Indoor Air Quality & Climate, Indoor Air, Seoul, Korea, 2020.
[2] S. Batterman, Review and extension of CO_2 based methods to determine ventilation rates with application to school classrooms, Int. J. Environ. Res. Publ. Health 14 (2) (2017 Feb 4) 145, https://doi.org/10.3390/ijerph14020145.
[3] Z. Noorimotlagh, N. Jaafarzadeh, S.S. Martínez, S.A. Mirzaee, A systematic review of possible airborne transmission of the COVID-19 virus (SARS-CoV-2) in the indoor air environment, Environ. Res. 193 (2021 Feb), 110612, https://doi.org/10.1016/j.envres.2020.110612.
[4] H. Chojer, P.T.B.S. Branco, F.G. Martins, M.C.M. Alvim-Ferraz, S.I.V. Sousa, Development of low-cost indoor air quality monitoring devices: recent advancements, Sci. Total Environ. 727 (2020), 138385.
[5] OSBSS, Open-Source Building Science Sensors, 2015. http://www.osbss.com/. July 28, 2023.
[6] A.S. Ali, Z. Zanzinger, D. Debose, B. Stephens, Open-Source Building Science Sensors (OSBSS): a low-cost Arduino-based platform for long-term indoor environmental data collection, Build. Environ. 100 (2016) 114–126.
[7] M. Penza, D. Suriano, M.G. Villani, L. Spinelle, M. Gerboles, Towards air quality indices in smart cities by calibrated low-cost sensors applied to networks, Proc . IEEE Sens. (2014) 2012–2017.
[8] W.A. Shewhart, Economic Control of Quality of Manufactured Product, vol. 509, ASQ Quality Press, 1931.
[9] E.S. Page, Continuous inspection schemes, Biometrika 41 (1/2) (1954) 100–115.
[10] Y. Wang, J. Jang- Jaccard, M. Boulic, et al., Deployment issues for integrated open-source - based indoor air quality school monitoring box (SKOMOBO), in: IEEE Sensors Applications Symposium (SAS), 2018, pp. 1–4.
[11] R. Weyers, J. Jang-Jaccard, A. Moses, et al., Low-cost indoor air quality (IAQ) platform for healthier classrooms in New Zealand: engineering Issues, in: Proc - 2017 4th Asia-Pacific World Congr Comput Sci Eng APWC CSE 2017, 2018, pp. 208–215.
[12] Y. Wang, M. Boulic, R. Phipps, et al., Integrating open-source technologies to build a school indoor air quality monitoring box (SKOMOBO), in: Proc - 2017 4th Asia-Pacific World Congr Comput Sci Eng APWC CSE 2017, 2018, pp. 216–223.
[13] N.F. Zhang, A statistical control chart for stationary process data, Technometrics 40 (1) (1998) 24–38.
[14] J.E.J. Wieringa, Statistical Process Control for Serially Correlated Data, 1999.
[15] R. Osei-Aning, S.A. Abbasi, M. Riaz, Monitoring of serially correlated processes using residual control charts, Sci. Iran. 24 (3) (2017) 1603–1614.
[16] R. Osei-Aning, S.A. Abbasi, M. Riaz, Mixed EWMA-CUSUM and mixed CUSUM-EWMA modified control charts for monitoring first order autoregressive processes, Qual. Technol. Quant. Manag. 14 (4) (2017) 429–453.
[17] X. Zhong, Efficient Sampling Plans for Control Charts when Monitoring an Autocorrelated Process, PhD dissertation. Virginia Polytechnic Institute and State University, 2006. http://scholar.lib.vt.edu/theses/available/etd-03012006-103618/.
[18] L.C. Alwan, Effects of autocorrelation on control chart performance, Commun. Stat. Theor. Methods 21 (4) (1992) 1025–1049.
[19] S. Psarakis, G.E.A. Papaleonida, SPC Procedures for monitoring autocorrelated processes, Qual. Technol. Quant. Manag. 4 (4) (2007) 501–540.
[20] R. Osei-Aning, S.A. Abbasi, M. Riaz, Monitoring of serially correlated processes using residual control charts, Sci. Iran. 24 (3) (2017) 1603–1614.
[21] G.E.P. Box, G. Jenki, Time Series Analysis: Forecasting and Control, Revised edition, Holden-Day, San Francisco., 1976.
[22] SM.Wu S. Pandit, Time Series and System Analysis with Application, Wiley, New York, 1983.
[23] M. Murat, I. Malinowska, M. Gos, J. Krzyszczak, Forecasting daily meteorological time series using ARIMA and regression models, Int. Agrophys. 32 (2) (2018) 253–264.
[24] D.A. Dickey, W.A. Fuller, Distribution of the estimators for autoregressive time series with a unit root, J. Am. Stat. Assoc. 74 (366a) (1979) 427–431.
[25] D. Kwiatkowski, P.C.B. Phillips, P. Schmidt, Y. Shin, Testing the null hypothesis of stationarity against the alternative of a unit root. How sure are we that economic time series have a unit root? J. Econom. 54 (1–3) (1992) 159–178.
[26] G. Schwarz, others, Estimating the dimension of a model, Ann. Stat. 6 (2) (1978) 461–464.
[27] J.E. Cavanaugh, A.A. Neath, Generalising the derivation of the Schwarz information criterion, Commun. Stat. Theor. Methods 28 (1) (1999) 49–66.
[28] R. Cardinael, T. Eglin, B. Guenet, C. Neill, S. Houot, C. Chenu, Is priming effect a significant process for long-term SOC dynamics? Analysis of a 52-years old experiment, Biogeochemistry 123 (1–2) (2015) 203–219.
[29] M. Qi, G.P. Zhang, An investigation of model selection criteria for neural network time series forecasting, Eur. J. Oper. Res. 132 (3) (2001) 666–680.
[30] Tiee-Jian Wu, A. Sepulveda, The weighted average information criterion for order selection in time series and regression models, Stat. Probab. Lett. 39 (1) (1998) 1–10.
[31] M.A. Mahmouda, P.E. Maravelakisb, The performance of the MEWMA control chart when parameters are estimated, Commun. Stat. Simulat. Comput. 39 (9) (2010) 1803–1817.
[32] C.W. Champ, L.A. Jones-Farmer, S.E. Rigdon, Properties of the T2 control chart when parameters are estimated, Technometrics 47 (4) (2005) 437–445.