Building and Environment 209 (2022) 108668

Contents lists available at ScienceDirect

Building and Environment

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

A methodology for the selection of pollutants for ensuring good indoor air quality using the de-trended cross-correlation function

Maria Justo Alonso [a,*], Sebastian Wolf [b], Rikke Bramming Jørgensen [c], Henrik Madsen [b], Hans Martin Mathisen [a]

[a] Department of Energy and Process Engineering, NTNU, Kolbjørn Hejes v 1B, Trondheim, Norway
[b] Department of Applied Mathematics and Computer Science, DTU, Asmussens Allé, Kgs Lyngby, Denmark
[c] Department of Industrial Economics and Technology Management, NTNU, Sem Sælands vei 5, Trondheim, Norway

ARTICLE INFO

Keywords:
IAQ
DCV
CO_2
Air temperature
RH
$PM_{2.5}$
Formaldehyde

ABSTRACT

CO_2 is customarily used to control ventilation as it is a proxy for bio-effluents and pollutants related to the presence and activity of people in the room. However, CO_2 could not be a satisfactory indicator for pollutants that do not have a metabolic origin, i.e., emissions from building materials or emissions from traffic. A methodology to select pollutants besides or instead of CO_2 is presented in this article. This methodology sets to study (i) the suitable location to measure air pollutants and (ii) which parameters to measure. The answers to these two questions are based on correlation analysis between pollutants and indoor/outdoor ratios.

Measurements of CO_2, air temperature, relative humidity, formaldehyde, and particulate matter have been taken in an office, an industrial kitchen, and a gym and are used to show how to apply the methodology. Correlations were studied in detrended (pre-whitened) time series. Studying correlations in detrended time series via cross-correlation functions is recommended because correlation coefficients may be overestimated because of the trends in the time series. In contrast to Pearson's correlation coefficient, the cross-correlation function studies the correlation between pollutants concurrently (as Pearson) but also at different time lags.

From the measurements we can conclude on the need to measure at least one parameter representing: 1) pollutants related to human activities 2)pollutants that infiltrate from processes like combustion or traffic outdoors, 3)pollutants related to combustion indoors, 4)pollutants related to degassing from building materials, 5) pollutants related to other "non-combustion-related activities" indoors and moisture loads.

1. Introduction

Buildings have evolved from having high rates of uncontrolled and unfiltered leakages to very tight envelopes with very reduced leakages to save energy [1–3]. Ventilation and filtering of air are necessary to secure the minimum requirements for indoor pollutants levels and thermal comfort in modern buildings [4,5]. The indoor environment is among the essential factors for a person's cumulative air pollutant intake [6]. Outdoor air pollutants enter the indoor air via infiltrations and ventilation systems. Pollutants are generated also indoors as a result of different activities [7]. All adverse airborne pollutants, disregarding their origin, must be ventilated away to ensure good indoor air quality. The World Health Organization defines the maximum threshold concentrations for various contaminants based on health effects [8]. These guidelines intend to inform national policymakers on the selection of appropriate targets for healthy air quality. However, national thresholds vary among countries and standards define different requirements of VR. In the USA and many countries in Asia, HVAC system sizing and VR are chosen to provide comfort, not health, though ASHRAE Standard 62.1 defined the acceptable indoor air quality to be without any known contaminants at harmful concentrations [4]. Logue [9] proved that in residences in the US and countries with similar lifestyles, air pollutant concentrations indoors exceed health-based standards for chronic and acute exposures in many measured cases. The WHO concluded that about 3.8 million people die annually due to household air pollution [10].

Thus there is a growing interest in monitoring IAQ by using low-cost sensors and developing platforms that can integrate sensing with actuating at low cost [11]. Guyot et al. [12] analyzed literature related to smart residential ventilation. In their review, they refer to ventilation controls using CO_2, temperature, relative humidity, and total volatile

* Corresponding author.
E-mail address: maria.j.alonso@ntnu.no (M. Justo Alonso).

https://doi.org/10.1016/j.buildenv.2021.108668
Received 22 October 2021; Received in revised form 5 December 2021; Accepted 7 December 2021
Available online 11 December 2021
0360-1323/

Nomenclature

AR	Autoregressive model
ARIMA	Autoregressive integrated moving average model
ARMA	Autoregressive moving average
CAV	Constant air volume
CCF	Cross-correlation function
CO_2	Carbon Dioxide
DCV	Demand-Controlled Ventilation
HVAC	Heating, ventilation, and air-conditioning
HOCH	Formaldehyde
IAQ	Indoor Air Quality
I/O	Ratio between indoor and outdoor pollutant concentrations
IoT	Internet of things
MA	Moving average model
MET	Metabolism
NOx	Nitrogen oxides
O_3	Ozone
PCP	Pentachlorophenol
PM	Particulate matter
PAH	Polycyclic Aromatic Hydrocarbon
RH	Relative humidity
SBS	Sick Building Syndrome
TVOC	Total Volatile organic compound
UFP	Ultrafine particles
VR	Ventilation Rate
VOC	Volatile organic compound
WHO	World Health Organization

organic compounds (mostly in bathrooms). Chiesa [11] developed an IoT application that controlled ventilation based on CO_2, volatile organic compounds, atmospheric pressure, RH, and temperature. They concluded that the proposed system defined proper airflow rates so that IAQ indexes are maintained. This article builds upon the possibility of using several parameters for ventilation control and develops a method that will be helpful to unveil correlations among pollutants to choose which ones are necessary and which ones are "nice to have". The same methodology can also be used to know where sensors should be placed so that they are useful for a ventilation control.

1.1. CO_2 as a marker for demand-controlled ventilation

CO_2 is often monitored as a proxy for occupancy in rooms [13,14]. People produce CO_2 proportionally to their body mass and metabolic rate. CO_2 concentrations are also understood as an indicator of the hygiene of the indoor air.

DCV is a ubiquitous choice to save energy in buildings where the occupancy varies throughout the day, e.g., in office buildings. The demand is defined from the level of one or several parameters. CO_2-DCV targets keep CO_2 below a set point concentration. If CO_2 indoor levels are below the defined threshold, VR can be reduced [15]. The airflow rate decrease is the mechanism by which CO_2-DCV realizes energy savings [16].

Carrer et al. [17] questioned using CO_2 as a measure of the ventilation's ability to dilute and remove pollutants. More than 50% of the pollutants present in offices are not emitted by humans [18]. In addition, the air supplied to the room can be taken from outdoors (via mechanical ventilation plus filter or infiltrating via cracks), it can be recirculated from the extracted air or infiltrated from other rooms. Depending on the air's origin or its pollutants concentration, it has different "dilution power".

Ramalho et al. [19] investigated correlations of CO_2 concentrations and selected indoor pollutants (formaldehyde, acetaldehyde, benzene, $PM_{2.5}$, PM_{10}) in 567 dwellings and 310 educational buildings (nurseries, kindergartens, and schools). They concluded that the correlations between CO_2 and pollutant concentrations were weak or very weak. Their study concluded that the probability of exceeding pollutant health guideline values correlates with high CO_2 concentration, but the possibility of exceedance is still high at low CO_2 levels. Choe et al. [20] found that air cleaners could reduce PM concentrations while CO_2 concentrations were still high. Wu et al. [21] presented measurements in green buildings with one-to three-star ratings. In their case, CO_2 and PM where lower than in ordinary buildings but VOC was higher. Therefore, some authors specify that CO_2 should only be used as a signal of occupant-related pollutants [22,23]. Others suggest that CO_2 should be observed as an IAQ indicator and a pollutant impacting health and cognitive functions [24,25]. Some authors suggest also controlling other parameters [17,19,23,26]. However, to the knowledge of the authors, there is no clear guideline why or when several pollutants should be measured in addition to CO_2 and temperature. Morawaska et al. confirm that there are no ventilation guidelines to specifically control the concentration of benzene, carbon monoxide, formaldehyde, and other chemicals, indoors [27] In this article we set the goal of developing a methodology to know which parameters should be measured in different types of rooms based on detrended correlation studies. Sun et al. [28] show the need of increasing the number of pollutants measured when correlating health outcomes and concentrations of pollutants. In their case they propose to use weights for the correlations. Here we propose a stepwise approach,1) measuring several pollutants, 2) study de-trended correlations and 3) Parameters that are correlated don't need to be further measured as correlation equations can be deployed. Uncorrelated pollutants need to be continuously measured. In the next section, selected pollutants that can be measured with low-cost sensors will be discussed.

1.2. Other (selected) indoor air pollutants: sources

Fine particles and UFP (<0.1 μm) can infiltrate buildings through leakages [29] and ventilation (mechanical or natural) openings. Mechanical ventilation using filters can reduce the I/O of $PM_{2.5}$ compared to natural ventilation [30]. The chosen filters in the HVAC systems, the precision of the mounting and their condition will also affect the I/O ratio. Chen & Zhao [31] concluded that the I/O varied importantly also due to the cracks geometry in building envelopes, and the air exchange rates. The principal indoor sources of $PM_{2.5}$ are smoking, cooking, fuel combustion for heating, human activities, hair, skin, and burning incense [32]. Indoor UFP can be generated from candles, cleaning and aerosol products, cooking, and other sources [33]. Morawska et al. [34] assessed that 10–30% of the total burden of disease from PM exposure was due to indoor-generated particles.

Particle's chemical composition, size, shape, deposition, and resuspension and hygroscopic growth appear to depend on RH [35]. RH affects also the rate of degassing of formaldehyde and VOC from indoor materials [36], the formation of molds and allergens and pathogens [37]. Gładyszewska-Fiedoruk [38] claimed that if humans are the most significant contributors to moisture generation, CO_2 and RH are also highly correlated, at least in naturally ventilated buildings. For air conditioning, where the air is cooled or dehumidified, correlations cannot be determined [38].

Salthammer summarized sources and intensity of formaldehyde in European housing [39] Formaldehyde is widely used in the manufacture of building materials and numerous household products. It is also a by-product of combustion from candles, incense sticks, mosquito coils, cigarettes, wood-burning fireplaces [39] and a preservative in some food packing [40]. Air cleaning devices, textiles, cooking, carpets and surface coatings, plywood, MDF are also sources for formaldehyde [39]. Huang et al. [41] concluded that formaldehyde, acetaldehyde, or benzene can be derived from cooking activities.

Sources of VOC in indoor air could be building materials, furnishings, cooking, household products, cleaning products, products for personal hygiene, etc. [42].

Nitrogen oxides (NO_x), Ozone (O_3), Pentachlorophenol (PCP), Polycyclic Aromatic Hydrocarbon (PAH), bio effluents, tobacco smoke are also main indoor pollutant substances, but they will not be further studied in this article as low-cost sensors for measuring them were not found.

Thus, additionally to occupancy measured by CO_2, and thermal loads measured by temperature, the following parameters should be monitored:1) pollutants that infiltrate from processes like combustion or traffic outdoors, 2) pollutants related to combustion indoors, 3) pollutants related to degassing from building materials, 4) pollutants related to other "non-combustion-related activities" indoors and moisture loads so that the main sources for pollutants are covered.

1.3. Other (selected) indoor air pollutants: main health effects

Indoor air humidity, defined as the perceived dry air or dryness (usually of eyes, upper airways, mucosae, or skin), is essential due to the associated health effects [37]. Fewer tears are produced, and precorneal and epithelial damage has been observed at low RH [37]. Dry air perception can be connected to mucous membrane irritation of eyes and upper airways in the presence of sensory irritants [43]. The reported "stuffy or dry air" may be affected by alteration of the composition, dynamics, deposition and resuspension of inhaled particles, possibly in concert with sensitive eyes or mucous membranes in the upper airways at low RH [37]. Cain et al. [44] claimed that temperature and RH altered VOC emission profiles and this correlated to the perception of IAQ.

Moisture and microbial contamination in the building structure and HVAC systems have adverse health effects [45]. The growth of microorganisms (fungi, bacteria, viruses) and the occurrence of allergens were linked to high RH [46]. Thus, indoor RH should be kept below mold-or-mites growth thresholds by ventilation or air conditioning [45]. However, too low indoor temperatures and low RH were associated with increased occurrence of respiratory tract infections. Influenza virus increased survival rate and transmission efficiency at low RH [37,47]. Contrarily, RH>40% dramatically reduced the infectivity of some other virus [48]. Coronavirus seemed to decay faster close to 60% RH than at other levels [49]. In general, there is a good agreement in the literature that many viruses decay faster in the range 40–60% [50–52].

Multiple studies with varying populations and regions showed consistent correlations between PM and cardiovascular problems. The data demonstrated a dose-dependent relationship between PM in ambient air and human disease [53]. Chronic $PM_{2.5}$ exposure affects the respiratory and cardiovascular systems [54]. Chronic bronchitis, stroke, heart disease, and thickening of arterial walls, diabetes, and reduced lung function were also connected to $PM_{2.5}$ exposures [55–57].

The relations between indoor particulate matter (PM_{10}, $PM_{2.5}$) and associated health risks are less known [53,58]. Venn et al. [59] proved an increasing risk of wheeze with increasing proximity for children living within 150 m of a main road. Peters et al. [60] concluded that decreases in peak expiratory flow, feeling ill during the day, and coughing were associated with the concentration of fine and UFP on asthmatics. PM impacts the IAQ and health, but may also be the carrier of viruses such as influenza [37].

According to the INDEX project results [61], the EU's risk assessment of IAQ agrees on prioritizing: formaldehyde, carbon monoxide, nitrogen dioxide, benzene, and naphthalene. Formaldehydes exist in the indoor air at a concentration that is larger than the outdoor air [39]. Formaldehyde has been classified as a potential human carcinogen by the US EPA and International Agency for Research on Cancer as a Class 2A carcinogen. Also, it irritates humans mostly in the upper airways, mucosae, and eyes [62]. Formaldehyde is a sensitizing agent that can cause an immune system response and sensory irritation [63].

VOCs at typical indoor environment concentrations may yield adverse health effects, depending on their composition. VOC concentrations indoors are generally below thresholds for sensory irritation in eyes and airways, but above odor thresholds [64]. Even if there is confirmation of a variety of dangerous effects probably linked to VOC, established scientific knowledge about direct health risks of VOCs is absent [65].

To sum up, when it comes to health effects the following parameters should be measured as exposures as these pollutants have important health effects: 1) RH as it affects the perception of IAQ and mostly the survival of viruses, 2) PM as the exposure to them is connected to cardiovascular and breathing problems and 3) formaldehyde as it is known as an irritant and a potential human carcinogen.

1.4. Exposure vs. concentration measurements

Most of the epidemiological studies discuss the relations between exposures and sickness. The NAS report [66] defines personal exposure as $E = \frac{1}{T}\int C(t)dt$ where E is personal exposure, C(t) is the time-variant concentration, and t is the time that the person experiences a specific concentration. Children and adults may be exposed differently as the particles have different spatial positions and particle size distribution [67–69]. Wilson & Suh [70] concluded that the relevant epidemiologic parameter was the concentration of the ambient particles that have penetrated the indoor microenvironment and remained suspended. The settling velocity is directly proportional to the particle diameter (to the square) and the density of the particle. Particles smaller than 10 μm can remain suspended for longer periods [71,72]. Guak & Lee [73] studied the relationship between personal exposure and ambient concentration of PM_{10} and $PM_{2.5}$ for different time-activity patterns. They concluded that personal exposure and $PM_{2.5}$ were highly correlated.

Therefore, in this study, it is assumed that when measuring concentration, an imperfect indicator of exposures is obtained, but that there is a correlation between concentrations and exposures.

1.5. Objectives of the study

Today, DCV deploys CO_2 and temperatures as control parameters as they are linked to comfort and productivity and sensors are highly available. From the conclusions of chapters 1.2 and 1.3 RH, $PM_{2.5}$ and formaldehyde should be measured additionally to CO_2 and temperature to account for the main pollutants from non-metabolic activities and their health effects. We hypothesize that the other pollutants may be at adversely high concentrations, despite CO_2 and temperature values being below thresholds.

The main objectives of this article were:

1) Development of a methodology for selecting which pollutants to use as control parameters for flow rates and to control the share of outdoor air in the supplied air:
   a. A methodology to determine which pollutants can be proxy for others was deployed using CCF. With CCF the study can focus on the i) present time correlation: looking for a mutual relationship between two pollutants at the same point in time, and ii) determining the correlation between the variables at different time lags. The CCF pattern is affected by the underlying time series structures, by the autocorrelation or trends of each of the two variables. Thus, it is helpful to de-trend the data by pre-whitening.
   b. An I/O -study approach was used to allow a deeper insight into the origin of the pollutants (indoor or outdoor). Based on the

origin of the pollutant, increasing ventilation with outdoor air would be either beneficial or harmful for the IAQ.

2) *Examine the suitability of the methodology with 3 case studies.* RH, $PM_{2.5}$ and formaldehyde have been measured for at least one week, additionally to CO_2 and temperature in an office, a gym and a canteen/kitchen. The results were used to evaluate the suitability of the method and not to do a thorough mapping of correlations for the studied types of rooms and situations.

To the knowledge of the authors: i) pre-whitening of data for studying correlations in detrended time series has not been applied for selection of pollutants to control DCV before. ii) the same combination of pollutants has not been previously evaluated.

2. Methods

2.1. Methodology for data analysis

Fig. 1 summarizes the methodology of this work.

The data analysis focused on correlations between the selected parameters: CO_2, RH, temperature, formaldehyde, and $PM_{2.5,}$ as well as between the location where parameters were measured: the corresponding breathing height in each room (see Fig. 2) and the supply air terminals.

Previous studies have based their criteria for correlation on Pearson's analyzes with non-pre-whitened time series [74–78]. The correlation coefficient between two time series following the same trend often suggests a high correlation. However, the high value may be due to auto-correlation in the respective time series, rather than due to a real correlation between the two series [79]. The correlation patterns are affected by underlying time series structures of each of the two variables and the trends that each series has. Thus, it is helpful to de-trend the data. By pre-whitening data, detrended time series are attained. Pre-whitening of data should always be done before deriving correlations in trended time series [80,81]. Unless studying de-trended correlations, nothing can be assured about the causal relationship of these two-time series.

In this article, the calculation of the correlation of two time series is expressed by the linear correlation of different time lags between the two series [82] via the cross-correlation function (CCF). The correlation that the CCF shows is not pure inter-series-correlation, it is also affected by the autocorrelation of each of the two series (intra--series-correlation). By using CCF instead of simple Pearson coefficients, "time-shifted" correlations can be also studied. The Pearson correlation only studies the contemporaneous relationship between the two-time series and not how the variation of one parameter may affect another in time. Let's consider formaldehyde, the emission of this pollutant can be affected by RH and temperature. If we only look at Pearson coefficients, there may be no relationship, but if we studied several lags of time an effect of the variation of RH may be a predecessor of a peak in formaldehyde. In addition, not all the sensors have the same response time, or not all the reactions happen equally fast, thus studying the cross-correlation function is more complete.

One approach to isolate the correlation between the time series is to remove the autocorrelation, i.e., to detrend the series [83]. This approach is called pre-whitening.

Considering two time series, x, and y, of equal length. The three steps of pre-whitening are:

1. Determine a time series model for x, in this case, an Autoregressive Integrated Moving-Average Model (ARIMA) was used for trend removal [82]. The goal of this step is to describe x up to residuals that are white noise, e.g., a time series without autocorrelation.
2. Transform (filter) y by using the model used for x (using the same coefficients).
3. Calculate the CCF between the residuals from step 1 and the filtered y-values from step 2.

The cross-correlation that is left in step 3 corresponds to the correlation between the time series. It is proportional to the impulse response function between x and y. If pre-whitening was not done, then the CCF would have been affected by the autocorrelations in the signals. Y has (normally/always) autocorrelation, but this is not a problem if x is "white".

ARIMA models belong to the class of linear time series models. Hence, it was assumed that the concentration of all pollutants behaved linearly over time, i.e., each measurement could be represented as a linear combination of its past values. ARIMA models were chosen because they are the most general form of linear time series models, and they include simpler models such as AR (Auto-regressive), MA (moving average), or ARMA (autoregressive moving average) models. Note that it was not our foremost goal to identify a perfect model to describe the time series, but rather a model whose residuals are close to white noise, to remove autocorrelation from the series.

Some pollutants were generated in the considered space and some infiltrated from outdoors. In a room with a high concentration of outdoor produced pollutants, increasing the airflow rate would not be beneficial for diluting their concentration. For example, in a room where the concentration of pollutants from traffic was too high, increasing the ventilation airflow rate from outdoors would further increase the concentration of these pollutants (supposing that these pollutants were not filtered. Filter efficiency plays a big role in the concentration of pollutants). To define the best ventilation procedure to dilute measured pollutants I/O has been evaluated. When an I/O was below one, it meant that the pollutant was produced outside the room. In this case, it would not be useful to increase outdoor air ventilation rates to dilute the outdoors generated pollutant. For example, in a room where plastics and old papers are stored, the formaldehyde values can be high. In this case, the I/O of formaldehyde may be over 1 and increasing the ventilation rate of outdoor air would reduce the formaldehyde concentration.

2.2. Measurement spaces

The criteria for selecting the measured spaces were as follows:

1) Similar exposure to outdoor pollutants as they were all placed in the same building. East-oriented with similar airtightness,
2) Same ventilation solution, constant air volume (CAV) and equal ceiling-mounted diffuser whose jets generated strong mixing of the air in the room assumed to be, fully mixed

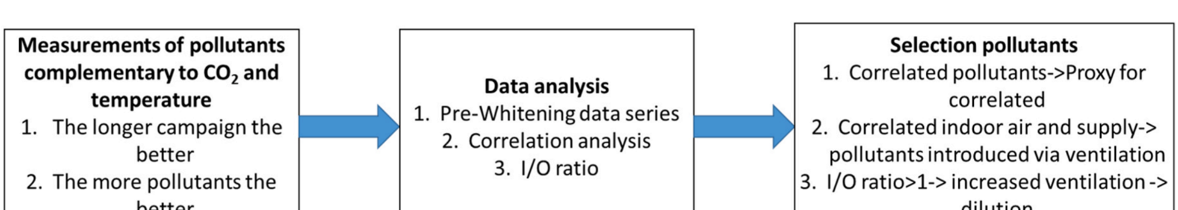

Fig. 1. Summary of methodology for the selection of the pollutants to be used to control ventilation.

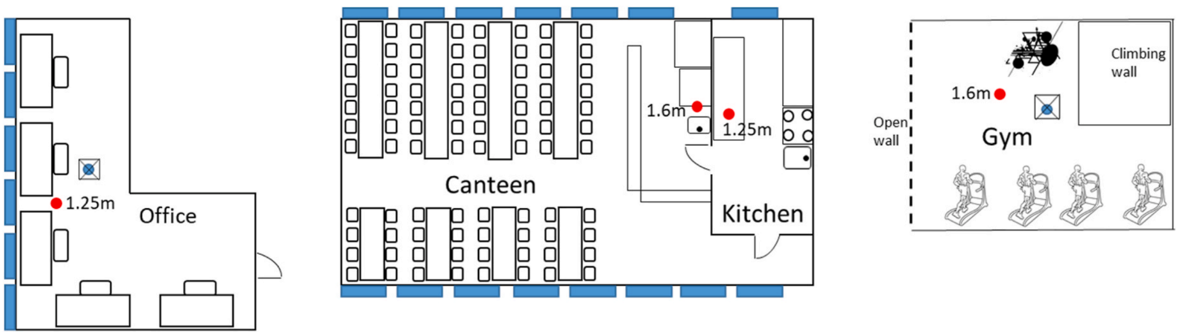

Fig. 2. Placement of measuring equipment. The blue dot shows the positioning in the supply air terminals and the red dots the positioning of the measurement equipment and the breathing height. Blue rectangles show the windows. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

3) Different activities performed in the room and
4) Connection to the same air handling unit (AHU). The AHU was East oriented, on the sixth floor, about 30 m away from the street measured in a straight line

Owing to these criteria and to validate the Methodology for data analysis, measurements were performed at the supply terminal of an office and a gym and breathing height in an office, a canteen, a kitchen and a gym. All the measured spaces were in the same building in Trondheim, Norway less than 40 m away from the road (measured in a straight line). Fig. 2 shows the layout of the rooms and the placement of sensors. The red dots represent the location and height of the measurement at "breathing height" in the three rooms and the blue dots represent the measurement at the supply air terminal. Table 1 summarized the equipment and characteristics of the three measured locations.

The 36 m^2 office on the second floor was dimensioned for five occupants and had a constant airflow rate of 300 m^3/h from 06:00 to 20:00 during working days. Outside this period or during weekends, the ventilation was off. The measurements lasted three weeks during June. The room was renovated with new walls, painting, windows, and a ventilation system about one year before the measurements.

Measurements in the 100 m^2 canteen and the 15 m^2 kitchen lasted one week. These two rooms were placed on the 6th floor. The kitchen was used by one cook that was responsible for baking bread, general cooking, and washing. The canteen was occupied from 11:30 to 13:30. Up to 50 people could be sitting at the most crowded periods. The canteen measurements were done close to the kitchen door. The supply and exhaust airflow rates were not measured, but the ventilation was constant during the same periods as the office. The canteen and kitchen had not been renovated in the last years.

The 50 m^2 gym was in the basement of the building. Measurements at the gym lasted twenty days. The occupancy was irregular from 15 people to long-vacant periods (users reported use of the room in a diary). The supply and exhaust airflow rates were not measured, but they were constant from 06:00 to 20:00 during weekdays. This gym was open to a large corridor (no wall, see dashed line in Fig. 2). The gym was built one year before measurements, and the ventilation system was not upgraded.

Table 1
Description of equipment and parameters measured.

	Office	Kitchen/ Canteen	Training room
Area (m^2)	36	15 + 100 (two rooms)	50
Occupants #	5 desks. 2–3 occupants during measurements	Variable, 0–50 persons	Variable, 0–8 persons
Ventilation principle	CAV mix ventilation 350 m^3/h	CAV mix ventilation	CAV mix ventilation + one wall open to the lab
Floor	Second	Sixth	Underground −1
Duration	Three weeks	One week	20 days

2.3. Equipment

The activities in these rooms were very different and the pollutants produced were expected to be different in quantity and type. Table 2 shows the expected contaminants based on the different types of activities.

Fig. 3 shows pictures of the installation of the sensors.

Measurements were done with low-cost sensors. Low-cost sensors were preferred as they could economically replace the "normal" CO_2-temperature sensor typically installed in these types of rooms. Table 3 summarizes the sensors, the accuracy, the measuring range, and the response time. More information about the calibration can be found in Ref. [84] (under publication). Demanega et al. discuss as well the performance of the particle sensors Sensirion SPS30 [85], Tryner et al. discussed the use of SCD30 CO_2 sensor [86] that measures as well humidity and temperature. Measurements were taken every 1 min.

In all measurements, the sensors were protected from direct contact with the users, direct disturbance from the ventilation supply air and solar irradiation. Measurements happened at a single point to mimic the normal measurement procedure when measuring CO_2 and temperature. For the reduced size of the rooms (not applying for the canteen where the representativeness is more limited), we assume that the single measurement was representative for the occupied breathing zone as the ventilation was mixing air ventilation.

3. Results

3.1. Correlation between different variables in each room

Correlations were sought for the whole measured period.

The pre-whitening process and CCF described in section "Correlations Analysis" were carried out to find correlations between two time series as described in the methodology chapter. In the plots of CCF, the x-axis (lag) represented the offset between both series, its sign determined in which direction the series were shifted. The y-axis showed the Pearson correlation coefficient of the two respective time lags. The larger the y-value, the larger was the correlation. The lag i value returned by CCF (x, y) estimates the correlation between x [t + i] and y [t] [87]. A negative correlation value CCF (x, y) < 0 meant that if one parameter increased, the other decreased. The lag times showed how long it took for one perturbance to propagate in the other series. A positive time index between two pollutants at lag i (i > 0) represented that the current value of a pollutant (current meaning at time [t]) was correlated with the future value of the other pollutant at time [t + i]. Equally, a negative value at lag i (i < 0) meant that the previous value of

Table 2
Expected pollutants generated in the studied rooms based on the different activities.

Type of room	Type of activity	Expected indoor air quality sources				
		CO_2	Air temperature	Formaldehyde	$PM_{2.5}$	RH
Office	Sitting and PC working	Breathing	Heat gain from MET, sun and computers	Beauty products, papers, wall painting furniture	Infiltration	MET
Gym	Physical activity	Breathing: proportional to occupancy and activity: ↑ MET: ↑exhalation	Heat relative to occupancy and activity level: ↑ MET: ↑sweat and ↑exhalation	Apparels and carpets	Infiltration Friction treadmill Climbing wall dust	↑ MET: ↑sweat and ↑exhalation
Industrial kitchen	Cooking	Breathing limited to the chef	Heat gain from cooking, oven, dishwasher, sun	Trash and cleaning products	From cooking, oven	From cooking, oven, dishwasher
Canteen	Eating	Breathing: proportional to occupancy	Heat gain from ↑ occupancy short period, sun	Food wrapping?	↑ occupancy + food remaining	↑ occupancy + food vapor

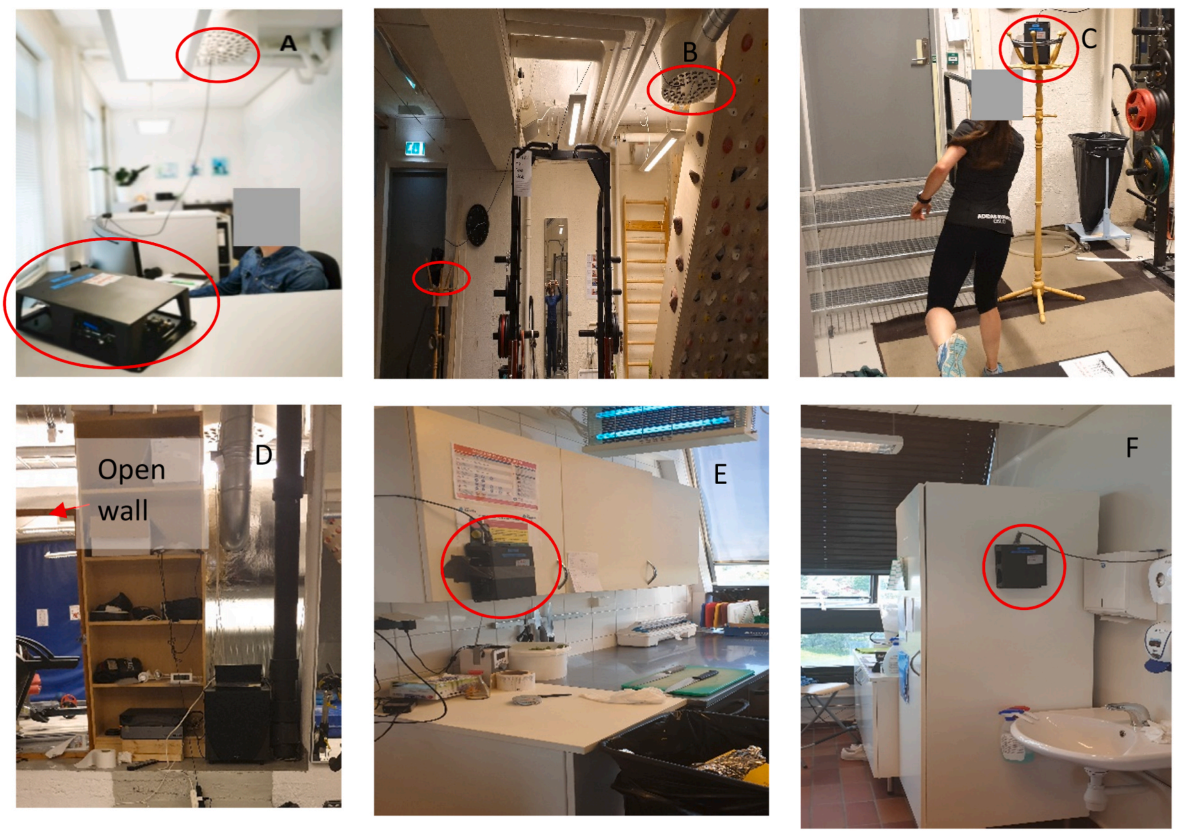

Fig. 3. Installations of measurement equipment: (A) in the office; (B), (C) and (D): in the gym; (E) in the kitchen; (F) in the canteen.

Table 3
Properties of deployed sensors.

Parameter	Sensor type	Accuracy	Measurement range	Response time
Relative humidity	Capacitive	±3% at 25 °C	0–100%	8s
CO_2	Nondispersive infrared (NDIR)	±30 ppm ± 3% of reading (500–1500 ppm)	400–10000 ppm	20s
Temperature	10K NTC Thermistor	±0.4 °C + 0.023 (t [°C] - 25 °C)	−40 °C–70 °C	>10s
Particle concentration	Optical sensor	±10 μg/m^3 (0–100 μg/m^3) ±10% (100–1000 μg/m^3)	0–1000 μg/m^3	20 ms
Formaldehyde	Electrochemical sensor (MOS)	≤0.02 ppm formaldehyde equivalent < ± 2% repeatability	0.03–2 ppm	<40S

one pollutant at time [t + i] was correlated with the present value of the other pollutants at time [t]. This can be read as one pollutant was a predecessor for the current value of the other. The dashed blue line represented the 95% confidence bound (blue dashed lines) for a significant correlation.

3.1.1. Correlations between pollutants from the gym

As shown in Fig. 4, both supply and room levels of absolute humidity and temperature followed similar trends. CO_2 levels were mostly below 700 ppm for the presented period. Mostly, one person trained at the time, and no group training was presented in Fig. 4. In general, during training activities, the MET increased, the temperature rose quickly, and

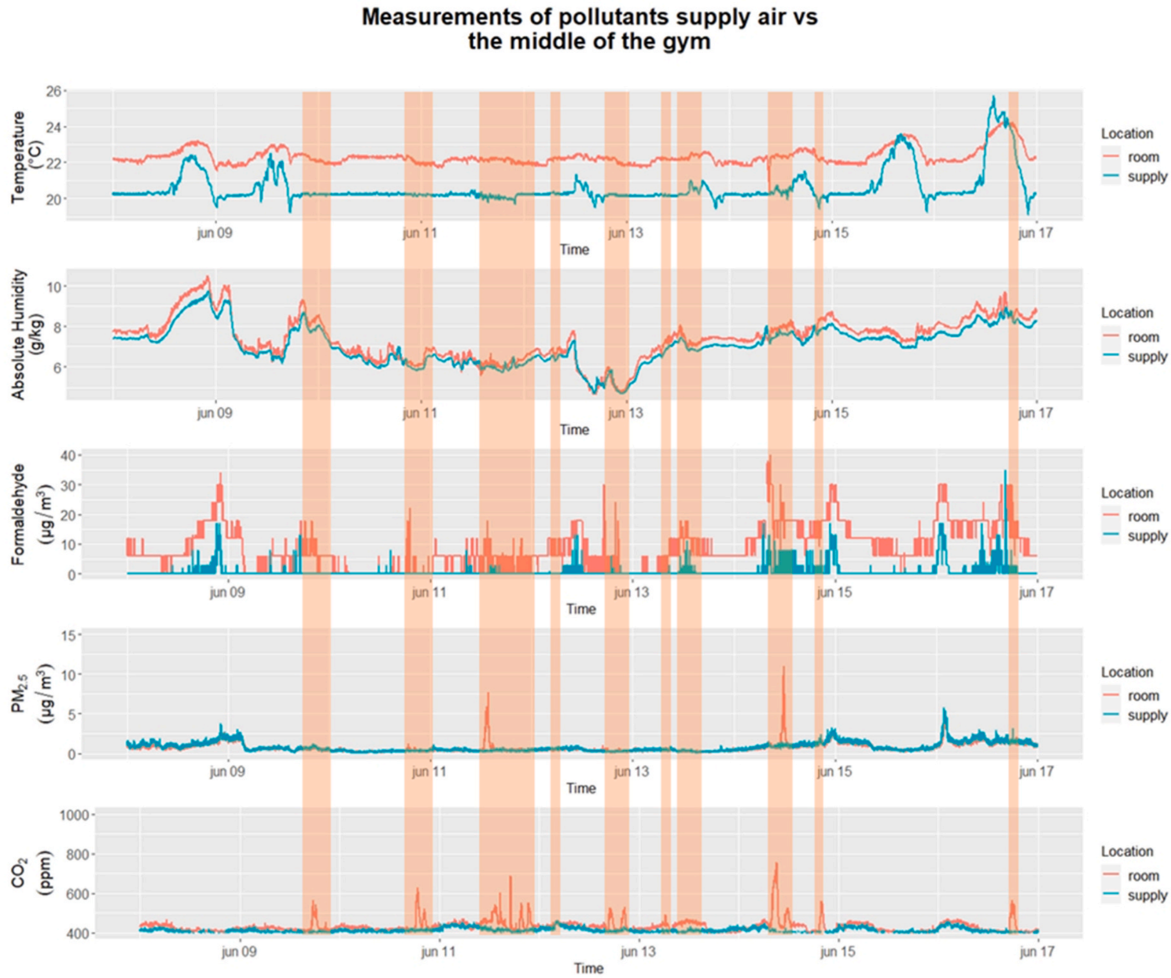

Fig. 4. Evolution of the concentration of different pollutants in the supply and room air in the gym. Training periods are shaded in orange. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

the production of CO_2 was much higher. Formaldehyde concentrations rose towards the end of the day while ventilation was shut down and decreased every morning when ventilation started. Formaldehyde levels also seem to rise together with the periods of training read as peaks of CO_2. Regarding $PM_{2.5}$, the level of particles was continuously very low, the levels only rose during climbing activities (read from gym diary).

Fig. 5 shows the cross-correlation functions of the detrended time series of each two pollutants in the gym (room values). There were very few significant correlations between $PM_{2.5}$ and temperature or absolute humidity or formaldehyde, and those that were over the 95% confidence bands are very small in value. There is a significant correlation between $PM_{2.5}$ and CO_2 probably corresponding to the particle creation by gym users. However, judging by the module of the correlation and the few time lags with significant correlations, CO_2 does not seem to be describing $PM_{2.5}$ concentrations unquestionably. There was a weak correlation between CO_2 and temperature or absolute humidity in higher lags. The lags were related to the simultaneity, higher lags can be read as a delay in the effect. Occupants would come into the gym and CO_2 will rise faster than temperature and absolute humidity or $PM_{2.5}$. However, the absolute value of the correlation factors was still low, which meant a low effect. There was a strong correlation between temperature and absolute humidity at lag zero, i.e., simultaneously. Formaldehyde is correlated to CO_2 and absolute humidity. Many points were over the 95% confidence bands, and even if they were small, a large number of smaller values will add up to a large total effect. Even if we cannot conclude on a causal relationship, we can conclude that using these two parameters is important to describe formaldehyde. That is exactly another strength of this method that describes the effects of several parameters. Formaldehyde values were correlated positively to temperature in lag zero. When the temperature rose, the formaldehyde also rose.

The CCF when using non-detrended time series is presented in Fig. S1 in the appendix. In the trended analysis, the correlations appeared to be stronger than the detrended and all variables are significant (outside the 95% confidence interval). Mathematically, the assumptions to use covariance methods as Pearson's are: 1) that the cases should be independent to each other, 2) that the two variables should be linearly related to each other and 3) the residuals scatterplot should be roughly rectangular-shaped. This is not the case when using a non-de-trended time series.

3.1.2. *Correlations between pollutants from the office*

Occupants of the office reported that the door was only opened and closed to allow them to enter. They kept a diary of when they arrived and left the office, but they did not record short vacancies. Fig. 6 presents the room air and supply air concentrations of the five pollutants. The occupancy of the office was low, CO_2 stayed mostly below 650 ppm. The ventilation of the office was dimensioned to constantly deliver 100% design airflow rate for five persons from 06:00 to 20:00, but the maximum registered occupancy was 3 persons. CO_2 peaks seem to be simultaneous to formaldehyde and $PM_{2.5}$ peaks but not to the ones of temperature and absolute humidity.

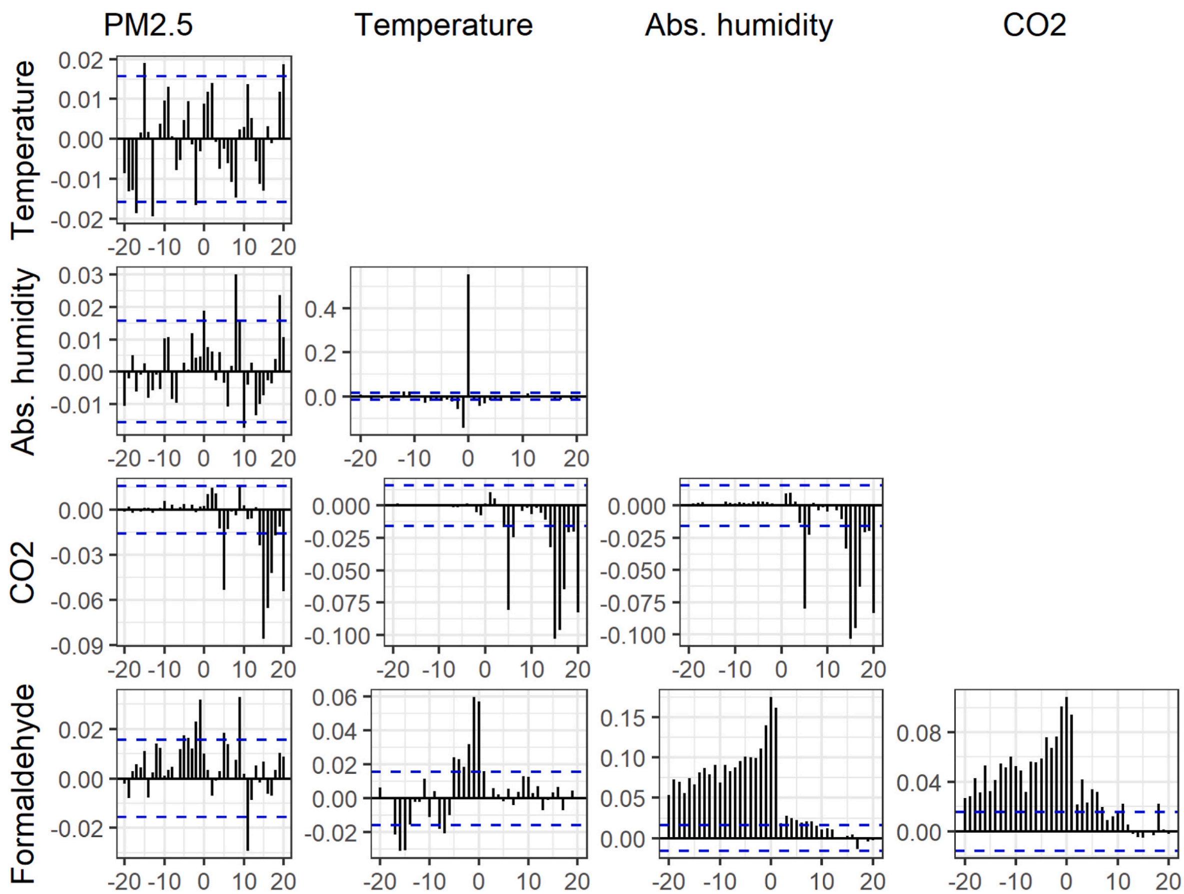

Fig. 5. Cross-correlation function between the different pollutants for the gym measurements (room measurements). The x-axis (lag) represents the offset between both series, its sign determines in which direction the series are shifted. The y-axis shows the Pearson correlation coefficient of the two respective time lags.

The levels of formaldehyde in supply air were almost always below indoor levels. Formaldehyde is produced continuously in small amounts by occupants, paper, and many other sources [48]. The emissions from the furniture and paintings were very low in this case as the university confirmed using low emitting materials and paintings and the renovation was done one year before the measurements. Most of the peaks happened while the ventilation was on, thus not necessarily because of the emission from materials but more related to occupants' activities. Therefore, a correlation with CO_2 could be justified. The same applies to $PM_{2.5}$ given that the outdoor air is very little polluted. Road construction works were happening, and the traffic was limited. The concentration of $PM_{2.5}$ in Trondheim's air is generally low in June. The spring cleaning of the roadway finished in May, and the studded snow tires are no longer used [88].

Fig. 7 shows the results for the cross-correlation function of the de-trended data series. Only temperature and absolute humidity correlate at a high value. There is also a correlation between CO_2 and temperature and CO_2 and absolute humidity that happened at low lags. However, the values were very small in the range (ca. 0.075) and had few occurrences. The absolute value of the function may depend on the number of observations, and probably a more extended sample should be analyzed to have stronger conclusions. Formaldehyde correlated significantly with absolute humidity, temperature and CO_2 at low lags, and these seem important as many different lags are significant.

3.1.3. *Correlations between pollutants from the kitchen*

Fig. 8 shows the development of the pollutants in the canteen and kitchen for three consecutive working days. In this case, due to the high temperature in the kitchen, the cook ran an additional personal air-cooling system. Besides, the kitchen had solar shading, which justifies the temperature difference with the canteen despite being both rooms connected.

The absolute humidity in both rooms depended mostly on the activities. During dishwashing, baking, or floor mopping, humidity levels rose. Regarding CO_2, both rooms followed each other as they were communicating through a large opening. During busy periods, up to fifty people can sit for lunch. These high occupancies are followed by peaks of CO_2. After 14:00, there was seldom anyone in the kitchen beside the cook that left around 15:00.

The high concentration levels of formaldehyde after the room was vacant were probably connected to the trash bin being left open in the room (emptied every third day). Thus, the first two days had high concentrations from probably the trash bin, and the third did not show the same pattern. The door separating the canteen and kitchen was left open during the first night and closed during the second.

$PM_{2.5}$ levels were generally low during the measurement period. During the food preparation and bread baking, some $PM_{2.5}$ spikes were recorded. Otherwise, the levels were almost consistently below 2.5 μg/m^3 despite the cooking activities. The kitchen was on the sixth floor thus, hardly exposed to traffic-related sources.

Fig. 9 shows the CCF for the de-trended time series. In this case, only the measurements inside the kitchen were analyzed. There was a higher correlation in the low lags between absolute humidity and $PM_{2.5}$ or formaldehyde. Formaldehyde and $PM_{2.5}$ are also correlated in low lags. Baking and cooking yielded formaldehyde, $PM_{2.5}$ and humidity. Absolute humidity and temperature were also correlated in low lags. CO_2 correlated temperature and absolute humidity in low lags, but not with $PM_{2.5}$ or formaldehyde.

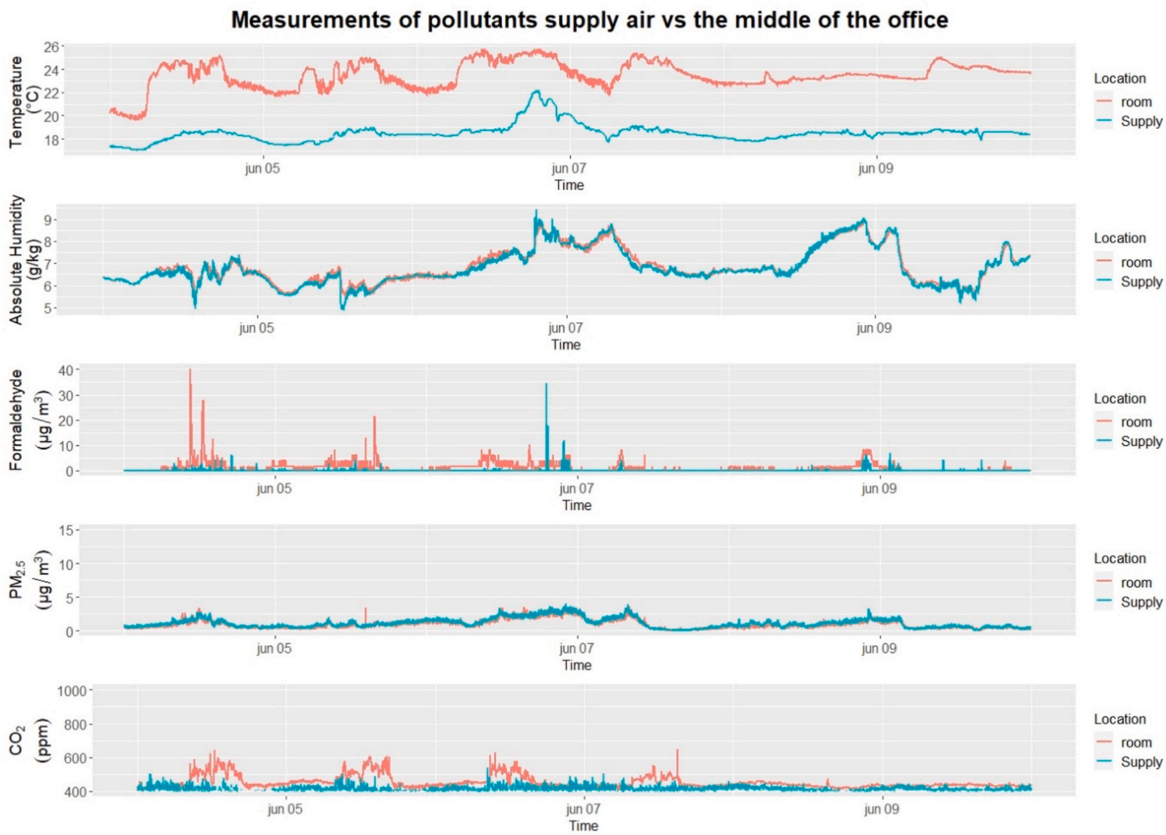

Fig. 6. Evolution of concentration of different pollutants in the office room.

3.1.4. Results using pearson correlations assessment

The widely used procedure for analyzing correlation in measurements of air pollutants is to use Pearson's correlation. The Pearson coefficients are indicators of a linear correlation between two sets of data. The Pearson correlation has two assumptions: i) the two variables are normally distributed and ii) the relationship between the two variables is linear.

However, if the time series show trends, it is crucial to remove these first to avoid autocorrelation interferences. A common way to de-trend is not to use the absolute values of the series but their relative changes over time. Hence, one applies Pearson's correlation coefficient on the differenced time series [89]. In this case, the de-trend of the values was done using ARIMA models as explained in steps 1 and 2 of the chapter of methodology for data analysis.

Fig. S2 to Fig. S7 presented in the Appendix show the correlation values for the pollutants in the three different rooms, first using the raw data and then using the de-trended time series. In the kitchen using the raw data would have induced us to conclude a correlation between absolute humidity and CO_2 that is not confirmed when using the de-trended data. Looking at the de-trended values in the kitchen there is only a correlation between the absolute humidity and the $PM_{2.5}$ or temperature. In the gym, there is a strong correlation between absolute humidity and temperature and a smaller correlation between absolute humidity and CO_2. In the office, there is only a correlation between absolute humidity and temperature. Using de-trended values is necessary to avoid overestimating correlations.

Pearson's correlations analyze correlation only the lag 0, whereas the cross-correlation functions analyze the whole time series. Thus, when using Pearson's correlations delays in response from different sensors would not be reflected in the result. For instance, the correlation in the kitchen between absolute humidity and formaldehyde or $PM_{2.5}$ are neglected when only looking at lag zero as these happen at lag -2 and -3.

3.2. Correlation of the different pollutants between supply and breathed air

This chapter presents the CCF between the same pollutants in the supplied air and breathed air and the I/O. No analysis of the kitchen/canteen is presented as the measurements were not at supply and breathing height but only at breathing height.

3.2.1. Correlation between room and supply air at the office

There was a high correlation between the room and the supply air in the variables $PM_{2.5}$, temperature, and absolute humidity in the office. Most of the particles in an office are derived from infiltrations from the outdoors. Thus, it was expected to see a correlation in $PM_{2.5}$. For absolute humidity and temperature, given that the measurements were taken under summer conditions for low occupancy periods, there were no large sources of moisture or heat, and correlation was expected. The CO_2 levels of supply and breathed air correlate weakly, which was plausible, given that indoors the largest source of CO_2 was human exhalation, but the ventilation rate is very high. Formaldehyde had a low correlation as most sources happened indoors. These findings agreed with the plots on Fig. S8 in the appendix.

From chapter 3.1.2 it was concluded that control to modify the supply airflow rates in this office should use CO_2 (representing temperature and absolute humidity indoors), formaldehyde and $PM_{2.5}$. Absolute humidity, temperature, formaldehyde and $PM_{2.5}$ correlate between indoor and outdoor, thus in this case measuring only indoor should be sufficient. Fig. 10 right shows the I/O. For all the presented pollutants, the I/O is larger than one meaning that increasing ventilation to remove these pollutants is an effective solution.

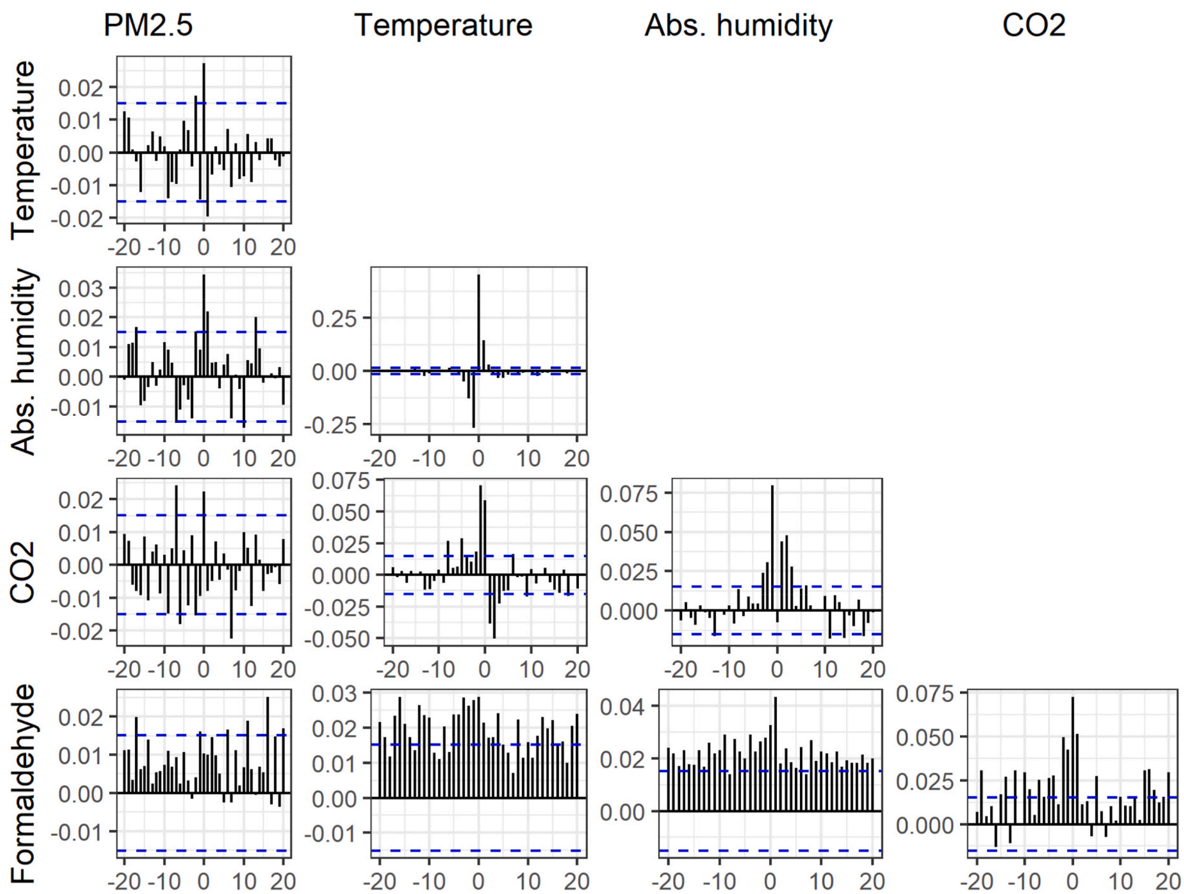

Fig. 7. Cross-correlation function between the different pollutants for the office measurements (room measurements).

3.2.2. Correlation between room and supply air at the gym

The gym had no window to the road that could let in $PM_{2.5}$ caused by traffic. $PM_{2.5}$ was mostly brought to the room via activities such as climbing. Hence, correlations for $PM_{2.5}$ deviate from the ones at the office as shown in Fig. S9 in the Appendix. Formaldehyde is brought to the room via the activities; thus, we do not see correlations with the supply air. The same applies to CO_2. For absolute humidity and temperature, as there is no heating of the air neither humidification there is a correlation between the supply and the room values.

In this case, most of the I/O ratios are over one for $PM_{2.5}$, formaldehyde and CO_2 as shown in Fig. 11. For $PM_{2.5}$, formaldehyde and CO_2 it is efficient to increase the airflow rates. For absolute humidity and temperature increasing the ventilation (given that there is no cooling nor dehumidification) may not be a good way of reducing over-temperature or too high absolute humidity.

For absolute humidity and temperature, the values are correlated between room and supply, but as the values can be below 1, both indoor and outdoor should be measured. As the I/O ratios are over 1 for formaldehyde, $PM_{2.5}$ and CO_2, increasing ventilation airflow rates to remove the pollutants is a good measure. Given the lack of correlation between indoor and outdoor and that the I/O is over one for CO_2, formaldehyde and $PM_{2.5}$ probably measuring only indoors is enough. However, for more conclusions regarding the removal of sensors longer measurement periods are recommended.

4. Discussion

CCF in de-trended time series is more accurate than CCF in non-detrended time series as autocorrelation could suggest stronger correlations [80,81]. The differences between Fig. 5 and Fig. S1 in the Appendix show the large effect of having trends in the data when analyzing with CCF. If only Fig. S1 in the Appendix was used for the analysis, overestimations of the correlations would be assumed. Using CCF instead of simple Pearson's correlation at lag zero studies both the contemporaneous relationship and delayed correlations. Fig. S2-Fig. S7 in the Appendix prove the need of expanding the analysis to CCF instead of only simultaneous correlations. When there is a risk for delayed effects, using Pearson would not suffice.

In this article, ARIMA models have been used to remove the autocorrelation. The assumption of linearity when using an ARIMA model has not been proven for these pollutants. However, in this work, we limit ourselves to linear models. With longer measurement periods the linearity could be tested as well.

In general, due to health hazards and possibilities for energy savings connected to reductions of VR, measuring several parameters, and using them for control of ventilation is recommended. Formaldehyde, $PM_{2.5}$, moisture and VOC were selected, additional to CO_2, because they represent the most plausible sources of pollutants in the measured indoor environment. Other pollutants, as ozone, bioaerosols, bacteria, NO_x, or SO_x could have been additionally measured but no available/reliable low-cost sensors were found for them.

The robustness of the conclusions from the measurements is limited, as the measurement campaign was too short and only in one season. Seasonal variations of outdoor-generated pollutants such as PM are expected. These are not reflected when measuring only for such a short period. For this methodology, the larger the data sample, the more robust conclusions. A one-week measurement period was sufficient only to demonstrate the methodology and this is how these results should be read. For the data presented for the rooms considered, the following can be concluded.

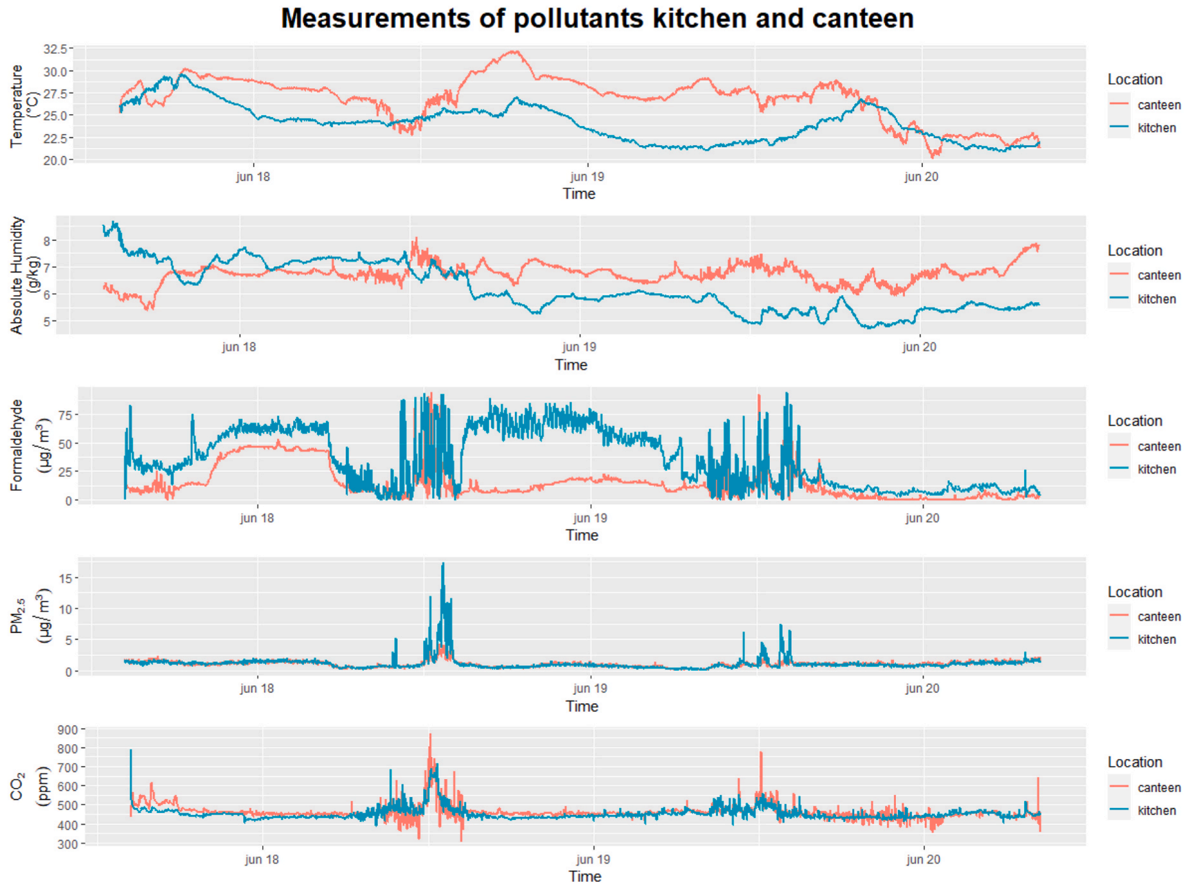

Fig. 8. Evolution of the concentration of various pollutants in the industrial kitchen and canteen.

- For most of the measured cases, the absolute humidity and room temperature were correlated (see Fig. 5, Figs. 7 and 9). Consequently, using only one to be representative of the other may be sufficient. Note that in these measurements, there were no large sources of humidity as it may happen in a bathroom.
- CO_2 did not capture most of the peaks in $PM_{2.5}$ (Fig. 4, Figs. 6 and 8), for the three measured spaces CO_2 and $PM_{2.5}$ did not correlate (see Fig. 5, Figs. 7 and 9), thus using CO_2 would not have been a good proxy to control $PM_{2.5}$. $PM_{2.5}$ was not strongly correlated to any of the other pollutants and should, therefore, be measured both in supply and room air. However, in the measurement period, the concentrations of $PM_{2.5}$ were very low indoors and outdoors. More data would be required representing a period with higher occupancy and higher outdoor $PM_{2.5}$ to have a better background. If the values were to be high and still not represented by the measurements of CO_2 as in the measurement period, they should be included in the ventilation control.
- Formaldehyde measurements may be exacerbated due to cross sensitivities. However, in some rooms such as the kitchen or gym where sources were available, it should at least be monitored to avoid surpassing safe limits. In the measured office and gym, formaldehyde was correlated to relative humidity, temperature, and CO_2.

These three cases show the need to measure $PM_{2.5}$ additionally to temperature and CO_2 to map all the different sources of pollutants. During the measurements in this article, the risk of exceedance of health guidelines was low as most of the rooms were ventilated with very high airflow rates. But the correlation between CO_2 and $PM_{2.5}$ was weak as for Ramalho et al. [19] measurements.

The proposed methodology provides a reliable way to select which parameters to use in the ventilation control. Correlation between the parameters does not induce causation. Causal relationships between the parameters are not subject of this study. In this study, correlation is used to determine which parameters can serve as a proxy for others. In practice, the correlations between parameters should be considered in the logic of IAQ control, either by monitoring all the parameters or by developing correlation equations that would ensure maintaining non-measured pollutants within a satisfactory range. Even if parameters are correlated, their absolute values are still important. For instance, formaldehyde should be at least measured to develop correlation equations. Formaldehyde peaks may be described by measurements of CO_2, temperature and absolute humidity. This is justified as in the measured cases; people and their activities are the largest reason for formaldehyde emissions. If we compare this methodology to the one proposed by Sun et al. [28], in their case, using weighting factors, would not allow for developing descriptive relations for the pollutants that are correlated. However, more measurements are needed as the ones presented here do not represent the design occupancy of the rooms or different seasons. The development of the control strategy must be case-and-space-dependent. An example of the protocol to formulate VR and a ventilation control based on the parameters that are not correlated in an office can be seen in the article from the authors [90] where a traffic pollutant and CO_2 are used for control.

Several reviews have studied existing knowledge about low-cost sensors. Coulby et al. [91] concluded on several sensor having varying accuracy compared to reference devices, but most of them responding similarly to environmental changes. Thus, several sensors were deemed to have high precision but reduced accuracy. In such cases, calibration can increase the accuracy [91]. However, Giordano et al. concluded that low cost sensors are subjected to the biases and calibration dependencies and the correction of such can range from simple linear regressions to very complex machine learning algorithms [92]. Therefore, low-cost

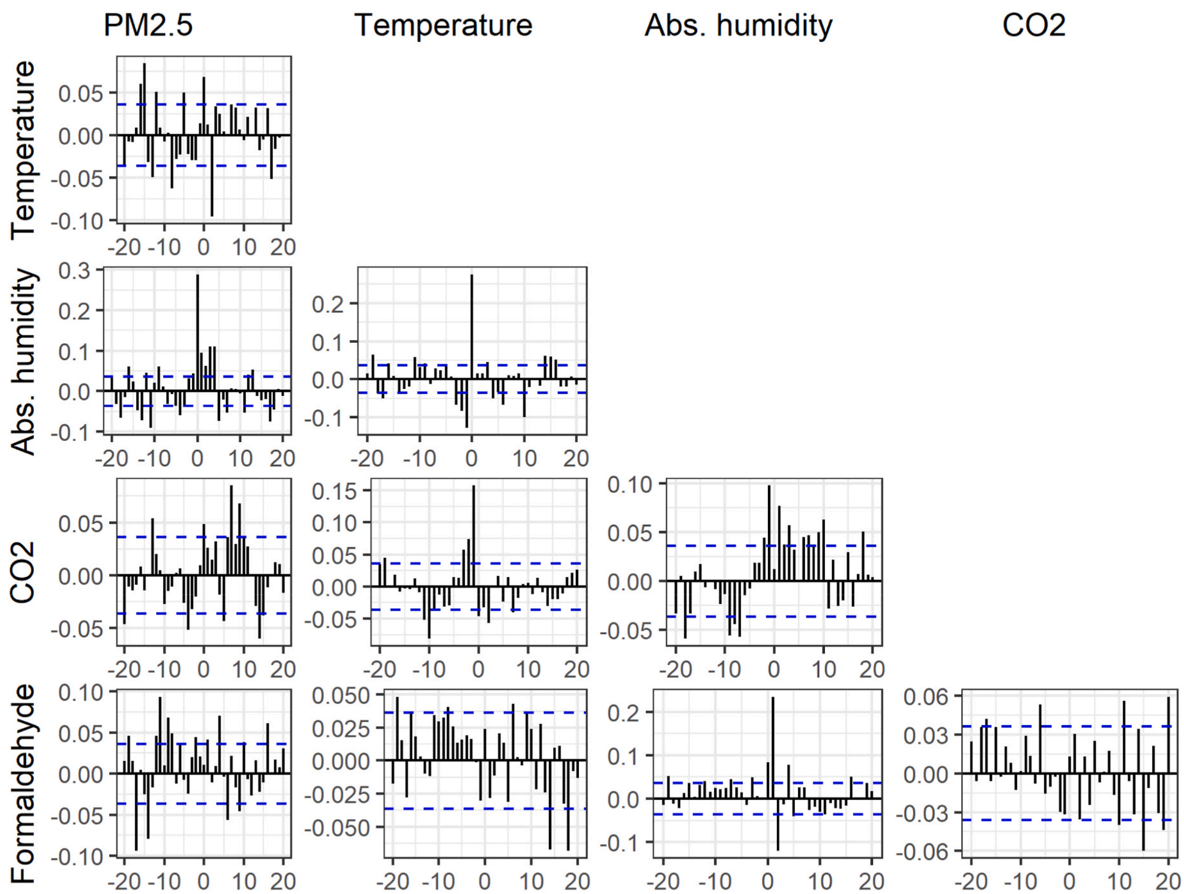

Fig. 9. Cross-correlation function between the different pollutants for the kitchen measurements.

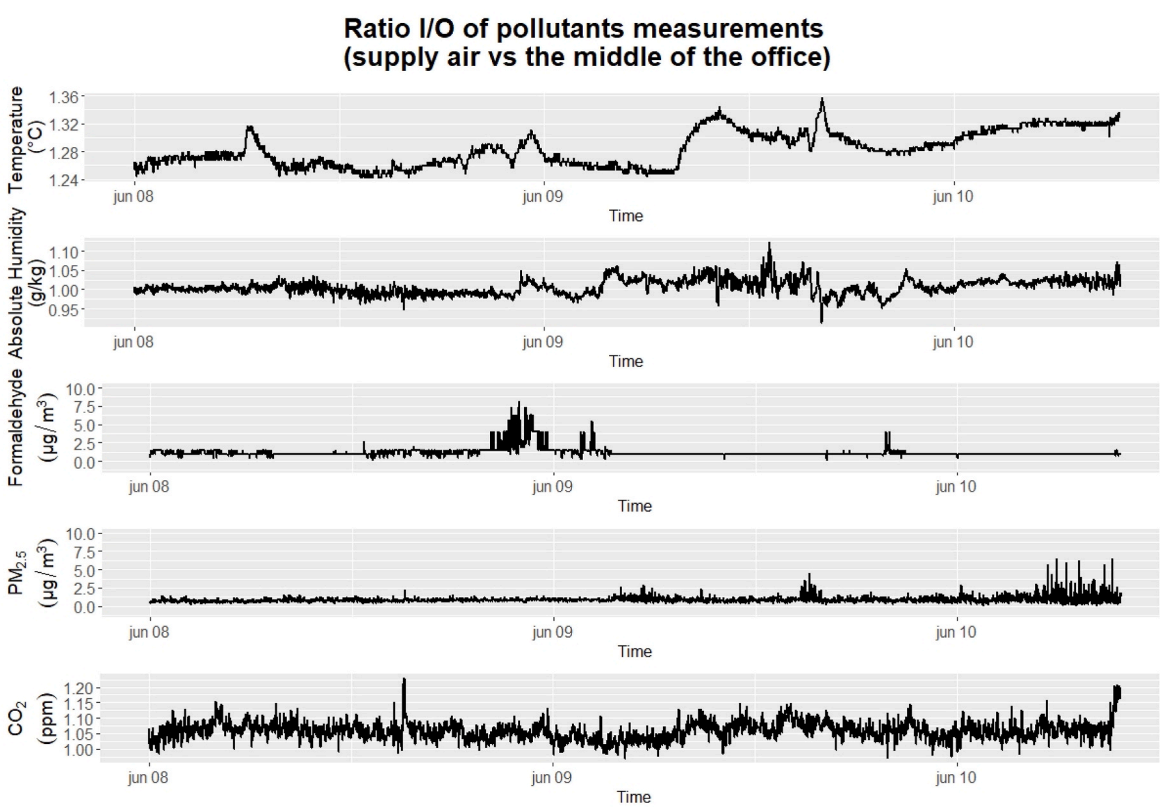

Fig. 10. I/O of pollutants measured in supply and room air at the office.

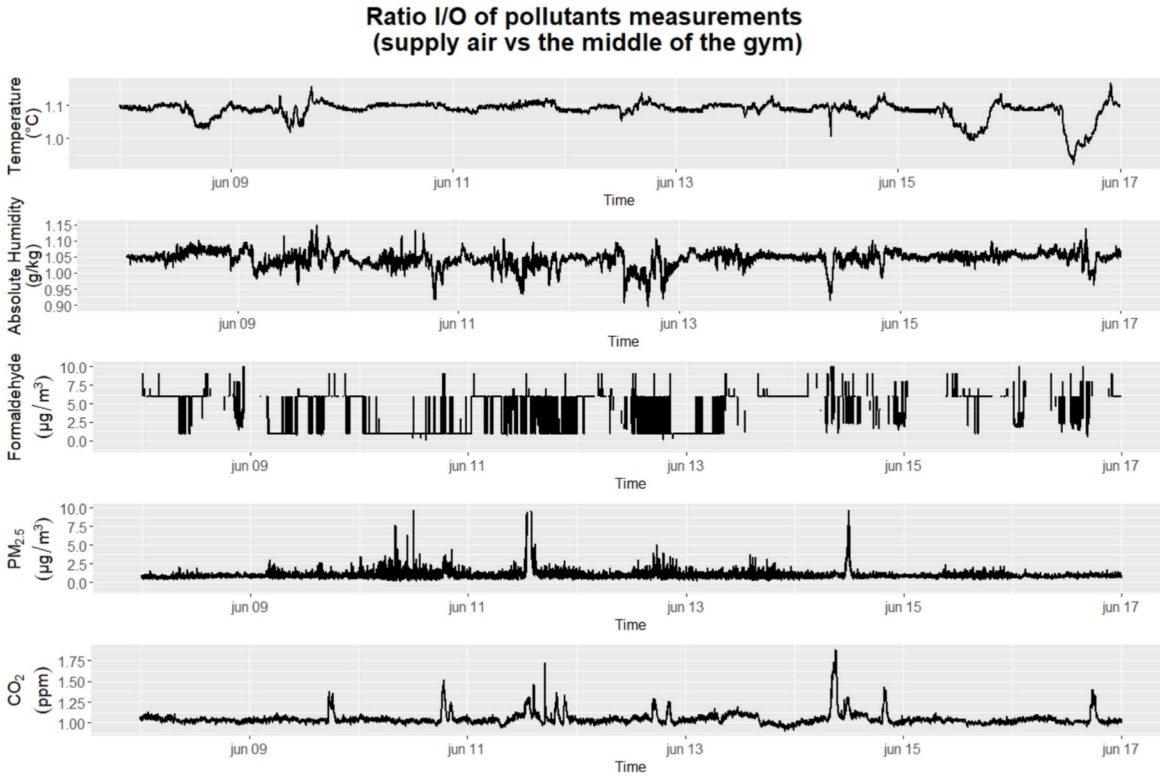

Fig. 11. Left: CCF of supply and room air concentrations of pollutants at the gym, Right I/O of pollutants measured in supply and room air at the gym.

sensors seem to be acceptable for this kind of study, but they must be calibrated always. When used for control of air quality using changes in concentration may be more suitable than using absolute values. Uncertainty of the measurements may jeopardize the selection of the main pollutants.

4.1. Limitations

- Short time measurements: The measurements in this study have been taken to map correlations between pollutants in different spaces. In observational studies like this, there is a potential for bias based on the representativeness and the short duration of the measurements. The difference in sampling duration is owed to the accessibility of the rooms. The presented measurements are in this case proof of concept of the methodology. Due to the small sample size and having only one room of each type, the representability of the results is limited. If these measurements would be used to develop a ventilation control, a longer measurement period would be needed to draw more robust conclusions. We would recommend having at least one week of measurements in every season at different occupancy levels so that the results are more representative. The results in this article should be read as the validation of the methodology. A generalization from the conclusions is not recommended.
- In practice, the deployment of this methodology would be very labor-intensive as first many sensors would need to be installed leaving the building "free float" with constant ventilation, and then the ventilation control strategy would need to be developed. However, with the development of deep learning algorithms and with a collection of data from different buildings, it is expected that the load of work to migrate to such a more holistic control strategy would be mitigated.
- Use of low-cost sensors: Manufacturers have started marketing low-cost air quality sensors to measure air pollution. The availability of such sensors will likely continue to grow [98]. Provided that they could produce reliable data, they could improve current ambient air monitoring. The providers of many of these sensors report limited information about sensor reliability and accuracy. Yet, due to their "low cost" and ease of use, they are used more and more. However, preliminary tests performed in the U.S. [99], [100] and in Europe [101], [102] suggest uncertain reliability, some do not perform well under ambient conditions, and do not correlate with data from "standard" measurement methods employed by regulatory agencies. They may also stop communicating with the system or may just have a shorter lifetime or show incorrect measurements Therefore, it is urgent to characterize the actual performance of IAQ sensors and to educate the public and users about the potential and limitations of these devices [98]. Drift was not followed up as calibration was done before the measurements, and the sampling duration was short. However, when using this methodology in practice, if low-cost sensors are used, their correct performance should be followed up. Several authors recommend using differential rather than absolute values in control [92,93].

Placement of the sensor and using single point: With the universalization of the use of low-cost sensors, better recommendations regarding placement should be delivered together with the sensor datasheets as the users are less expert on IAQ measurements. In this article, the placement of the sensor is the same as for the customarily CO_2 - temperature sensors. However, this affects, among others, the measurement of $PM_{2.5}$, which depending on the size, may distribute differently, or the formaldehyde that is heavier than air. The discussion of the optimal placement of sensors in the different types of rooms (e.g., with different functions, areas, height, and sources) is not taken in this article. However, to follow "standard control strategies", the sensor prototype measuring the five parameters at the same time is placed in breathing height (1.2 or 1.8 m high depending on the normal tasks of the occupants). The location of the sensor, low-cost or "standard", is very important as this measurement must be representative of the room. In the measured cases the air is provided via terminals that encourage full mixing. The placement of the sensor was protected from direct disturbances such as sunlight, heat from radiators, close to a trash bin, etc.

Using a single point to represent the whole occupied volume is a bold assumption. Even more in this case where we have not proven that the air is fully mixed. This is a practical and technical limitation related to the cost of sensors and the limitations on the disturbance to the users of the sampled rooms. We acknowledge that using one single point is an imperfect indicator of exposures and that it cannot provide high-resolution spatiotemporal data that would be important for an accurate evaluation of a dynamic indoor environment. However, this is standard practice when measuring temperature and CO_2 in DCV-systems and we have decided to follow it for practical reasons.

- The formaldehyde sensor has known cross-sensitivity issues with methanol, ethanol, isopropanol, carbon monoxide, phenol, acetaldehyde H_2, H_2S, and SO_2 [94]. Many low-cost $PM_{2.5}$ sensors are affected by RH and temperature [95]. Additionally, converting the light spreading to concentrations of PM depends on chemical and physical properties, size, and shape of particles and others that are not measured. Also, the air intake affects the particles entering the equipment by entraining smaller particles along. In general, these sensors are recommended for cases where the particle types are known and remain unchanged [96] what may not be the case here.
- Low concentration of pollutants: This building had very high ventilation rates per person during the measured period (due to vacancies as some measurements were done in summer periods). It would be very interesting to repeat the measurements when lower ventilation rates per person would be supplied. However, once again the measurements here are to be seen as an illustration of the methodology.

5. Conclusions

This paper presented a methodology to select the pollutants that should be used to control ventilation.

1. A methodology for the selection of pollutants *to use as control parameters* for supply flow rates and to control the share of outdoor air in the supplied air was developed. This is based on the study of CCF in pre-whitened data series. Additionally, an I/O -study approach was used to allow a deeper insight into the origin of the pollutants (indoor or outdoor). This methodology sets to study (i) Where to measure, supply, or/and breathing height and (ii) Which parameters to measure. The methodology should be used to give answers that are case-and-space-dependent.
2. Time series were detrended and correlations due to autocorrelation were removed. Studying correlations in detrended (pre-whitened) time series instead of Pearson's coefficients is superior as autocorrelation on the time series could imply stronger correlations and using CCF allows for studying the correlations at different time lags.
3. The methodology was studied with three case studies, an office, a gym and a kitchen. Measuring the five selected parameters (CO_2, $PM_{2.5}$, temperature, RH and formaldehyde) seems to give a more complete picture of the IAQ in the studied rooms than using only CO_2 and temperature. For most of the measured cases, the absolute humidity and temperature were correlated; CO_2 or temperature did not capture most of the peaks in $PM_{2.5.}$ and formaldehyde was correlated to temperature and CO_2.
4. From the measurements we can conclude on the need to measure at least one parameter representing: 1) pollutants related to human activities 2) pollutants that infiltrate from processes like combustion or traffic outdoors, 3) pollutants related to combustion indoors, 4) pollutants related to degassing from building materials, 5) pollutants related to other "non-combustion-related activities" indoors and moisture loads. These are not undoubtedly covered using only CO_2 and temperature.
5. In conclusion, this is a promising methodology that should be used further.

CRediT authorship contribution statement

Maria Justo Alonso: Writing – original draft, Visualization, Validation, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Sebastian Wolf:** Formal analysis, Methodology, Software, Visualization, Writing – review & editing. **Rikke Bramming Jørgensen:** Writing – review & editing, Formal analysis. **Henrik Madsen:** Formal analysis. **Hans Martin Mathisen:** Writing – review & editing, Supervision.

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.

Acknowledgments

The authors gratefully acknowledge the support from the Research Council of Norway (project no. 257660) and several partners through the Research Centre on Zero Emission Neighborhoods in Smart Cities (FME ZEN). Even Johan Christiansen and Olav Aleksander Myrvang at the Department of Electronic Systems at NTNU for their help with the Arduinos.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.buildenv.2021.108668.

References

[1] European Commission, Directive 2010/31/EU of the European Parliament and of the Council of 19 May 2010 on the Energy Performance of Buildings (Recast), *Off. J. Eur. Union*, 2010, pp. 13–35.
[2] Passive House Institute, Criteria for the Passive House, EnerPHit and PHI Low Energy Building Standard, 2016.
[3] DIBK, Energy Performance of Buildings - Calculation of Energy Needs and Energy Supply, 2016.
[4] ASHRAE, Standard 62.1 Ventilation for Acceptable Indoor Air Quality, *ASHRAE Stand.*, 2019.
[5] DIBK, Byggteknisk forskrift (TEK17), 2017.
[6] V. V Tran, D. Park, Y.-C. Lee, Indoor air pollution, related human diseases, and recent trends in the control and improvement of indoor air quality, Int. J. Environ. Res. Publ. Health 17 (8) (2020).
[7] V.G. Mihucz, G. Záray, Indoor air pollution, Compr. Anal. Chem. (2016).
[8] WHO, Guidelines for indoor air quality - selected Pollutants, WHO Guidel. 9 (2010) 454.
[9] J.M. Logue, T.E. Mckone, M.H. Sherman, B.C. Singer, Hazard assessment of chemical air contaminants measured in residences, Indoor Air 21 (2) (2011) 92–109.
[10] A. Pruss-Ustun, C. Corvalan, Preventing Disease through Healthy Environment: A Global Assessment of the Burden of Disease from Environmental Risks, *World Heal. Organ.*, 2006.
[11] G. Chiesa, S. Cesari, M. Garcia, M. Issa, S. Li, Multisensor IoT Platform for Optimising IAQ Levels in Buildings through a Smart Ventilation System, *Sustain.*, 2019.
[12] G. Guyot, et al., Residential Smart Ventilation : a Review & Savoie Mont Blanc University 2 Building Technologies and Urban Systems Division, 2017.
[13] L. Chatzidiakou, D. Mumovic, A. Summerfield, Is CO2 a good proxy for indoor air quality in classrooms? Part 1: the interrelationships between thermal conditions, CO2 levels, ventilation rates and selected indoor pollutants, Build. Serv. Eng. Res. Technol. (2015).
[14] A. Persily, Challenges in Developing Ventilation and Indoor Air Quality Standards: the Story of ASHRAE Standard 62, *Building and Environment*, 2015.
[15] A. Persily, J. Braun, S. Emmerich, K. Mercer, T. Lawrence, Recommendations for Application of CO2-Based Demand Controlled Ventilation, Including Proposed Guidance for ASHRAE Standard 62 and California's Title 2, 2004, pp. 1–8.
[16] S.J. Emmerich, A.K. Persily, State-of-the-Art Review of CO2 Demand Controlled Ventilation Technology and Application, 2001.

[17] P. Carrer, et al., What does the scientific literature tell us about the ventilation-health relationship in public and residential buildings? Build. Environ. 94 (2015) 273–286. P1.
[18] P.O. Fanger, Introduction of the olf and the decipol units to quantify air pollution perceived by humans indoors and outdoors, Energy Build. 12 (1) (1988) 1–6.
[19] O. Ramalho, et al., Association of Carbon Dioxide with Indoor Air Pollutants and Exceedance of Health Guideline Values, *Build. Environ.*, 2015.
[20] Y. Choe, et al., Inadequacy of Air Purifier for Indoor Air Quality Improvement in Classrooms without External Ventilation, *Build. Environ.*, 2022.
[21] P. Wu, et al., Comparative Analysis of Indoor Air Quality in Green Office Buildings of Varying Star Levels Based on the Grey Method, *Build. Environ.*, 2021.
[22] R. Maddalena, et al., Effects of ventilation rate per person and per floor area on perceived air quality, sick building syndrome symptoms, and decision-making, Indoor Air (2015).
[23] W.J. Fisk, How Home Ventilation Rates Affect Health: A Literature Review, Indoor Air, 2018.
[24] U. Satish, et al., Impact of CO_2 on Human Decision Making and Productivity vol. 2011, in *Indoor Air*, 2011.
[25] J. Allen, P. Macnaughton, U. Satish, S. Santanam, J. Vallarino, J.D. Spengler, Associations of cognitive function scores with carbon dioxide, ventilation and volatile organic compound exposure in office workers, Environ. Health Perspect. 124 (6) (2016) 805–812.
[26] K.W. Tham, Indoor air quality and its effects on humans—a review of challenges and developments in the last 30 years, Energy Build. 130 (2016) 637–650.
[27] L. Morawska, et al., A paradigm shift to combat indoor respiratory infection, Science (2021).
[28] C. Sun, X. Huang, J. Zhang, R. Lu, C. Su, C. Huang, The New Model for Evaluating Indoor Air Quality Based on Childhood Allergic and Respiratory Diseases in Shanghai, *Build. Environ.*, 2022.
[29] H. Zhao, B. Stephens, "Using Portable Particle Sizing Instrumentation to Rapidly Measure the Penetration of Fine and Ultrafine Particles in Unoccupied Residences, Indoor Air, 2017.
[30] J. Pantelic, M. Dawe, D. Licina, Use of IoT Sensing and Occupant Surveys for Determining the Resilience of Buildings to Forest Fire Generated PM2.5, *PLoS One*, 2019.
[31] C. Chen, B. Zhao, Review of Relationship between Indoor and Outdoor Particles: I/O Ratio, Infiltration Factor and Penetration Factor, *Atmospheric Environment*, 2011.
[32] Z. Li, Q. Wen, R. Zhang, Sources, health effects and control strategies of indoor fine particulate matter (PM2.5): a review, Sci. Total Environ. 586 (2017) 610–622.
[33] L. Stabile, E.R. Jayaratne, G. Buonanno, L. Morawska, Charged particles and cluster ions produced during cooking activities, Sci. Total Environ. (2014).
[34] L. Morawska, et al., Indoor aerosols: from personal exposure to risk assessment, Indoor Air (2013).
[35] Z.J. Wu, et al., Particle hygroscopicity and its link to chemical composition in the urban atmosphere of Beijing, China, during summertime, Atmos. Chem. Phys. (2016).
[36] A.V. Arundel, E.M. Sterling, J.H. Biggin, T.D. Sterling, Indirect health effects of relative humidity in indoor environments, Environ. Health Perspect. (1986).
[37] P. Wolkoff, Indoor air humidity, air quality, and health – an overview, Int. J. Hyg Environ. Health 221 (3) (2018) 376–390.
[38] K. Gładyszewska-Fiedoruk, Concentration of carbon dioxide measured in University classrooms, Civ. Environ. Eng. (2010).
[39] T. Salthammer, Formaldehyde Sources, Formaldehyde Concentrations and Air Exchange Rates in European Housings, *Building and Environment*, 2019.
[40] U. S. N. C. for B. NHI, Compound Summary: Formaldehyde, 2019 [Online]. Available: https://pubchem.ncbi.nlm.nih.gov/compound/Formaldehyde.
[41] Y. Huang, S.S.H. Ho, K.F. Ho, S.C. Lee, J.Z. Yu, P.K.K. Louie, Characteristics and health impacts of VOCs and carbonyls associated with residential cooking activities in Hong Kong, J. Hazard Mater. 186 (1) (2011) 344–351.
[42] A. Mozaffar, Y.L. Zhang, Atmospheric Volatile Organic Compounds (VOCs) in China: a Review," Current Pollution Reports, 2020.
[43] R.L. Doty, J.E. Cometto-Muñiz, A.A. Jalowayski, P. Dalton, M. Kendal-Reed, M. Hodgson, Assessment of upper respiratory tract and ocular irritative effects of volatile chemicals in humans AU -, Crit. Rev. Toxicol. 34 (2) (2004) 85–142. Jan.
[44] W.S. Cain, R. Schmidt, B.P. Leaderer, J.F. Gent, D. Bell, L.G. Berglund, Emission of VOCs from Materials Used in Buildings: Analytical and Sensory Aspects, in *ASHRAE Transactions*, 2002.
[45] J. K, O. Seppänen, Moisture control and ventilation, in: WHO Guidelines for Indoor Air Quality: Dampness and Mould, 2009.
[46] P. Wolkoff, S.K. Kjærgaard, The Dichotomy of Relative Humidity on Indoor Air Quality, " *Environment International*, 2007.
[47] M.M. Derby, et al., Update of the scientific evidence for specifying lower limit relative humidity levels for comfort, health, and indoor environmental quality in occupied spaces (RP-1630), Sci. Technol. Built Environ. 23 (1) (2017) 30–45.
[48] T.A. Myatt, M.H. Kaufman, J.G. Allen, D.L. MacIntosh, M.P. Fabian, J.J. McDevitt, Modeling the airborne survival of influenza virus in a residential setting: the impacts of home humidification, Environ. Heal. A Glob. Access Sci. Source (2010).
[49] D.H. Morris, et al., Mechanistic Theory Predicts the Effects of Temperature and Humidity on Inactivation of SARS-CoV-2 and Other Enveloped Viruses, *bioRxiv*, 2020.
[50] L.M. Casanova, S. Jeon, W.A. Rutala, D.J. Weber, M.D. Sobsey, Effects of air temperature and relative humidity on coronavirus survival on surfaces, Appl. Environ. Microbiol. (2010).
[51] K.A. Kormuth, K. Lin, Z. Qian, M.M. Myerburg, L.C. Marr, S.S. Lakdawala, Environmental Persistence of Influenza Viruses Is Dependent upon Virus Type and Host Origin, *mSphere*, 2019.
[52] K. Lin, L.C. Marr, Humidity-dependent decay of viruses, but not bacteria, in aerosols and droplets follows disinfection kinetics, Environ. Sci. Technol. 54 (2) (2020) 1024–1032. Jan.
[53] J.O. Anderson, J.G. Thundiyil, A. Stolbach, Clearing the air: a review of the effects of particulate matter air pollution on human health, J. Med. Toxicol. (2012).
[54] N. Martinelli, O. Olivieri, D. Girelli, Air particulate matter and cardiovascular disease: a narrative review, Eur. J. Intern. Med. (2013).
[55] C.A. Pope, et al., Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution, J. Am. Med. Assoc. 287 (9) (2002) 1132–1141.
[56] N. Kunzli, et al., Ambient air pollution and atherosclerosis in Los Angeles, Environ. Health Perspect. 113 (2) (2005) 201–206.
[57] R.T. Burnett, M. Smith-doiron, D. Stieb, S. Cakmak, J.R. Brook, Effects of particulate and gaseous air pollution on cardiorespiratory hospitalizations, Arch. Environ. Health 54 (2) (1999) 130–139.
[58] H. Fromme, et al., Particulate matter in the indoor air of classrooms—exploratory results from Munich and surrounding area, Atmos. Environ. 41 (4) (2007) 854–866. Feb.
[59] A.J. Venn, S.A. Lewis, M. Cooper, R. Hubbard, J. Britton, Living near a main road and the risk of wheezing illness in children, Am. J. Respir. Crit. Care Med. 164 (12) (2002) 2177–2180.
[60] A. Peters, D.W. Dockery, J. Heinrich, H.E. Wichmann, Short-term effects of particulate air pollution on respiratory morbidity in asthmatic children, Eur. Respir. J. 10 (4) (1997) 872–879.
[61] K. Koistinen, et al., The INDEX project: executive summary of a European Union project on indoor air pollutants, Allergy: Eur. J. Aller. Clin. Immunol. (2008).
[62] S. Norliana, A.S. Abdulamir, F. Abu Bakar, A.B. Salleh, The health risk of formaldehyde to human beings, Am. J. Pharmacol. Toxicol. (2009).
[63] P. Wolkoff, Indoor Air Pollutants in Office Environments: Assessment of Comfort, Health, and Performance, *International Journal of Hygiene and Environmental Health*, 2013.
[64] C. van Thriel, Odor Impact on Health and Work Performance from Indoor Air Exposure, *NIVA course*, 2018.
[65] US-EPA, Technical overview of volatile organic compounds, Indoor Air Qual. (2016) 1–5.
[66] NAS, Human Exposure Assessment for Airborne Pollutants: Advances and Opportunities, " *Choice Rev. Online*, 1991.
[67] Z.G. Cao, et al., Particle size: a missing factor in risk assessment of human exposure to toxic chemicals in settled indoor dust, Environ. Int. (2012).
[68] Z. Cao, G. Yu, B. Wang, J. Huang, S. Deng, Human exposure assessment of indoor dust: importance of particle size and spatial position, Environ. Health Perspect. (2013).
[69] M.V. Ruby, Y.W. Lowney, Selective soil particle adherence to hands: implications for understanding oral exposure to soil contaminants, Environ. Sci. Technol. (2012).
[70] W.E. Wilson, H.H. Suh, Fine particles and coarse particles: concentration relationships relevant to epidemiologic studies, J. Air Waste Manag. Assoc. (1997).
[71] L.C. Marr, Transmission of Viruses in Droplets and Aerosols, 2020.
[72] E.J. Farrell, D.J. Sherman, A new relationship between grain size and fall (settling) velocity in air, Prog. Phys. Geogr. (2015).
[73] S. Guak, K. Lee, Different relationships between personal exposure and ambient concentration by particle size, Environ. Sci. Pollut. Res. (2018).
[74] H.W. Davies, J.J. Vlaanderen, S.B. Henderson, M. Brauer, Correlation between co-exposures to noise and air pollution from traffic sources, Occup. Environ. Med. (2009).
[75] Y.H. Dai, W.X. Zhou, Temporal and Spatial Correlation Patterns of Air Pollutants in Chinese Cities, " *PLoS One*, 2017.
[76] Z. Andualem, Z. Gizaw, L. Bogale, H. Dagne, Indoor bacterial load and its correlation to physical indoor air quality parameters in public primary schools, Multidiscip. Respir. Med. (2019).
[77] J. Taylor, C. Shrubsole, P. Symonds, I. Mackenzie, M. Davies, Application of an Indoor Air Pollution Metamodel to a Spatially-Distributed Housing Stock, " *Sci. Total Environ.*, 2019.
[78] W. Liang, B. Zhao, J. Liu, J. Pei, Can carbon dioxide be a good indicator for formaldehyde in residences?—Monte Carlo modeling for a whole year, Sci. Technol. Built Environ. (2020).
[79] C. Lewis, N.R. Farnum, L.W. Stanton, Quantitative forecasting methods, J. Oper. Res. Soc. (1990).
[80] Sio-Long Ao, Applied Time Series Analysis, 2010.
[81] A.I. McLeod, H. Yu, E. Mahdi, Time Series Analysis with R, in *Handbook of Statistics*, 2012.
[82] H. Madsen, Time Series Analysis, 2008.
[83] S. Bisgaard, M. Kulahci, Quality Quandaries: Studying Input-Output Relationships, Part I, " *Qual. Eng.*, 2006.
[84] M. Justo Alonso et al., "Performance Assessment of a Low-Cost Arduinos-Based Sensor Station under Representative Indoor Air Conditions.".
[85] I. Demanega, I. Mujan, B.C. Singer, A.S. Anđelković, F. Babich, D. Licina, Performance Assessment of Low-Cost Environmental Monitors and Single Sensors under Variable Indoor Air Quality and Thermal Conditions, " *Build. Environ.*, 2021.
[86] J. Tryner, et al., Design and Testing of a Low-Cost Sensor and Sampling Platform for Indoor Air Quality, " *Build. Environ.*, 2021.
[87] G.E.P. Box, G. Jenkins, An Unexpected Route to Time Series, " *This Week's Cit. Class.*, 1976.
[88] Luftkvalitet info, Luftkvaliteten I Trondheim: Juni 2019, 2019.
[89] T. Fawcett, Avoiding Common Mistakes with Time Series, 2015.

[90] M. Justo Alonso, Dols, W. Stuart, and H. M. Mathisen, "Using Co-simulation between EnergyPlus and CONTAM to evaluate recirculation-based, demand-controlled ventilation strategies in an office building," Build. Environ.
[91] G. Coulby, A. Clear, O. Jones, A. Godfrey, A scoping review of technological approaches to environmental monitoring, Int. J. Environ. Res. Publ. Health (2020).
[92] M.R. Giordano, et al., From low-cost sensors to high-quality data: a summary of challenges and best practices for effectively calibrating low-cost particulate matter mass sensors, J. Aerosol Sci. (2021).
[93] N. Castell, et al., Can commercial low-cost sensor platforms contribute to air quality monitoring and exposure estimates? Environ. Int. (2017).
[94] Dart, "Dart Sensors WZ-S Formaldehyde Module Operation Manual.".
[95] H.Y. Liu, P. Schneider, R. Haugen, M. Vogt, Performance assessment of a low-cost PM 2.5 sensor for a near four-month period in Oslo, Norway, Atmosphere (2019).
[96] P. Lilienfeld, Dual-wavelength nephelometry for PM2.5 monitoring, J. Aerosol Sci. (1998).