

УДК 004.89:697.9

<https://doi.org/10.31713/vt2202614>

Borodii Ivan, ⁽¹⁾

PhD student

Osukhivska Halyna, ^(1:ORCID: 0000-0003-0132-1378)

Candidate of Technical Sciences, Associate Professor

Sverstiuk Andrii, ^(2:ORCID: 0000-0003-3146-9664)

Doctor of Technical Sciences, Professor

Voloshchuk Andrii, ^(1:ORCID: 0009-0007-1478-1601)

PhD

Zharovskyi Ruslan, ^(ORCID: 0000-0002-1159-3417)

Candidate of Technical Sciences

¹Ternopil Ivan Puluj National Technical University, Ternopil

² I. Horbachevsky Ternopil National Medical University, Ternopil

INTEGRATED ASSESSMENT OF INDOOR AIR QUALITY AND MICROCLIMATE ACCOUNTING FOR SENSOR DATA COMPLETENESS

A method for the integrated assessment of indoor air quality and microclimate was developed by combining sensor data validation with hierarchical fuzzy aggregation. Aerosol, ventilation and microclimatic components were derived from PM_{2.5}, PM₁₀, CO₂, temperature and relative humidity measurements. Each measurement was assigned a validity status, and each evaluation point received an OK, PARTIAL or FAILED completeness status without imputing missing values. The method was evaluated using 169,331 records collected over 59 days, covering 16,992 scheduled evaluation points. Of these, 89.70% received an OK status, 5.91% a PARTIAL status and 4.39% a FAILED status. Component averaging produced a more favourable class in 97.6% of cases with a component in the least favourable class, whereas the proposed aggregation preserved its influence. The median processing time of 1.62 s on a Raspberry Pi 5 supported the method's suitability for continuous monitoring.

***Keywords:* indoor air quality; environmental monitoring; low-cost sensors; Internet of Things; fuzzy logic; data analysis; assessment method.**

Introduction. Indoor air quality and the indoor microclimate have a direct impact on people's health, well-being and productivity. As people spend most of the day indoors, even short-term deviations in indoor air

quality parameters, such as increased concentrations of particulate matter and carbon dioxide, as well as changes in temperature and relative humidity, can have a negative impact on people's well-being, productivity and health. Regulatory documents set out requirements for establishing and assessing appropriate indoor environmental conditions in buildings, particularly regarding the microclimate and air quality [1; 2].

Low-cost sensors, combined with single-board computers, enable data collection and processing in air pollution monitoring systems based on Internet of Things technologies [3]. Pre-processing data at the edge reduces latency and dependence on network connectivity compared to transmitting all measurements to cloud-based systems. At the same time, such sensors may, over time, produce distorted readings, miss measurements or record isolated anomalous values unrelated to changes in air quality. If this data is used without prior verification, conclusions about the condition of the premises may prove to be incorrect.

It is difficult to interpret several individual indicators simultaneously. Therefore, they are aggregated into a single composite indicator and a room condition class. However, the existing indices may mask a critical condition of an individual indicator due to averaging, cause the class to change abruptly near the control limits, or fail to indicate incompleteness in the input data. It is therefore essential to conduct an integrated assessment of indoor air quality and microclimate, in which the final result reflects the unfavourable status of individual indicators and includes information on the suitability and completeness of sensor data.

The aim of this study is to develop a method for the integrated assessment of indoor air quality and microclimate based on the validation of sensor data and the hierarchical fuzzy aggregation of component assessments, with the formation of an environmental condition class and verification of data completeness.

The scientific novelty of the study lies in the refinement of a method for the integrated assessment of indoor air quality and microclimate, accounting for the suitability and completeness of sensor data. With this method, where data sets are incomplete, the assessment is formed without replacing missing values and is marked with a completeness status. This has made it possible to distinguish between

complete and incomplete results without the use of artificially generated data.

A classification of environmental conditions is proposed, which combines the aerosol, ventilation and microclimatic components using a hierarchical fuzzy inference process, prioritising the most unfavourable class, while the width of the transition regions between classes is determined by the measurement error of the sensors. Unlike the averaging of component assessments and fuzzy classification, this ensures that the influence of the most critical component is preserved in the final assessment and reduces local variations in the index near the control limits.

An analysis of recent studies and publications. Existing indoor air quality indices differ in terms of the composition of indicators, threshold values, normalisation methods and aggregation rules. Therefore, the same measurements may result in different overall classes. In particular, the authors of [4] considered the definition of an index for residential buildings based on limit values for long-term and short-term exposure. This approach links the assessment to maximum permissible levels of health impact. However, in the context of continuous monitoring with frequent updates to the results, it is also important to take into account short-term changes in indicators and the availability of individual input channels.

The study [5] proposes an index for the continuous monitoring of residential premises, which takes into account the mass concentrations of particulate matter with aerodynamic diameters of up to $2.5\ \mu\text{m}$ ($\text{PM}_{2.5}$) and up to $10\ \mu\text{m}$ (PM_{10}), the concentration of carbon dioxide (CO_2), air temperature (T) and relative humidity (RH). The final result was determined by the most unfavourable partial assessment. This approach retains the influence of the worst indicator, but strict limits may cause a change in class due to a slight deviation and do not reflect measurement uncertainty.

The selection of reference values presents a separate challenge in constructing the index. Reference values for particulate matter concentrations are most often taken from WHO recommendations [6]. These levels are set for 24-hour and annual averaging and therefore cannot be directly applied to short measurement intervals. In [7], CO_2 concentration is considered primarily as an indicator of the adequacy of air exchange relative to the number of people in the room, rather than as an independent criterion of air quality. The permissible ranges for the indoor microclimate are set out in [1; 2] and depend on the time of

year and the category of the premises. Consequently, a single index combines indicators with different types of reference values, which complicates their direct conversion to a single scale and justifies the preliminary formation of separate component assessments.

The authors of [8] noted that air quality indices differ primarily in the way they aggregate component scores, and examined two main disadvantages of such aggregation: 'masking', where the overall index indicates a favourable condition despite an unfavourable value for an individual indicator, and ambiguity, where it indicates an unfavourable condition not supported by the individual indicators. Averaging the partial assessments is prone to 'masking', whilst selecting the maximum assessment eliminates this. However, when combined with rigid class boundaries, this leads to a sudden change in the result. Fuzzy inference allows the most unfavourable condition to be retained while ensuring smooth transitions between classes.

In [9], a fuzzy decision support system was applied to monitor air quality in an operating theatre. Unlike rigid classification, a fuzzy system allows a value to partially belong to two neighbouring classes near the class boundaries. In [10], the authors proposed an adaptive fuzzy inference tree, in which indicators are first combined into partial assessments and then into a final result. This hierarchical structure also allows the influence of each group of indicators on the final air quality assessment to be analysed separately.

The width of the transition regions between classes in fuzzy systems is usually chosen by researchers at their own discretion. At the same time, the measurement error of low-cost sensors, as stated by the manufacturers, may be large enough to influence the classification of the measured value into neighbouring classes. The measurement accuracy specifications are given in the technical documentation for the relevant sensors: the SPS30 for suspended particulate matter concentrations [11], the SCD41 for CO₂ concentration [12] and the BME688 for temperature and relative humidity [13].

Fuzzy evaluation was applied to combine indicators of air pollution and thermal conditions. In [14], these groups of indicators were combined into a single fuzzy indoor air quality index. The authors of [15] developed a fuzzy controller for a vehicle cabin heating, ventilation and air-conditioning system, which simultaneously maintains thermal comfort, as assessed by the PMV index, and CO₂ concentration at the desired levels.

The authors in [16] demonstrated that the measurement results from low-cost sensors are influenced by air temperature and humidity, cross-sensitivity to other gases and gradual drift. Therefore, such data require calibration and correction. At the same time, the result of data validation is not usually reflected in the integrated index, and the user does not know how many measurements and from which channels were used to form the class.

The authors of the study [17] compared methods for filling gaps of 6, 12 and 24 hours in $PM_{2.5}$ measurement series and found that the accuracy of extrapolating the last known value decreased as the duration of the gap increased. Generally, when data are incomplete, calculations are either not performed at all, which reduces the availability of results, or are based on estimated values, hiding the incompleteness of the measurements.

Existing approaches address the tasks of verifying the suitability of sensor data, aggregating indicators, creating smooth transitions between classes and handling incomplete data separately. However, they do not combine these within a single method for the comprehensive assessment of air quality.

Method description. The monitoring system was deployed on a Raspberry Pi 5 single-board computer (Broadcom BCM2712, 8 GB of RAM). When organising the loading and storage of sensor data, the results of a previous study on the impact of data format and volume on the efficiency of these operations in Data Lakehouse architectures [18] were taken into account. The measurement node was installed in a room with an area of approximately 20 m². $PM_{2.5}$ and PM_{10} concentrations were measured using an SPS30 sensor, CO₂ concentration using an SCD41 sensor, and temperature (T) and relative humidity (RH) using a BME688 sensor. The other outputs of the SPS30 and the redundant T and RH channels of the SCD41 sensor were used solely to verify data consistency. Records were generated every 30 seconds without prior averaging. The dataset covered 59 days during the warm season and comprised 169,331 records.

The first stage of the method converted the aggregated values of the direct inputs into the aerosol component A ($PM_{2.5}$, PM_{10}), the ventilation component V (CO₂) and the microclimatic component M (T, RH), the second stage combined these into an integrated index:

$$I = F_I(A, V, M), \quad (1)$$

where the function F_I denotes the second level of fuzzy inference according to the Mamdani method. The components and the index are

presented on a dimensionless scale from 0 to 100, where 0 corresponds to the most favourable state. The membership degrees A , V and M for the four classes were passed to the second level, and their numerical values, obtained using the centroid method, were used for comparative configurations and correlation analysis.

Prior to calculating the index, each measurement was assigned a data validity status of VALID, SUSPECT, INVALID or MISSING. The status VALID indicated a valid value; SUSPECT indicated a physically plausible deviation requiring confirmation; INVALID indicated a technical error or a value outside the permissible range; MISSING indicated a missing record.

The technical check covered the presence of a record, numerical format, device identifiers, whether the value fell within the sensor's technical range, and the physical consistency of concentrations $PM_1 \leq PM_{2.5} \leq PM_4 \leq PM_{10}$ with a 20% acceptable range. Local anomalies were identified using the Hampel filter [19] based on the current value and the ten preceding values, provided that all had passed the technical check. The preliminary status 'SUSPECT' was also assigned following the detection of five identical values in a row or five consecutive changes in the same direction with a total magnitude of at least three times the sensor's uncertainty.

SUSPECT values were only included in the calculation after confirmation, which, for all parameters, required the change to be maintained for at least two consecutive measurements. Changes in $PM_{2.5}$ and PM_{10} were additionally considered confirmed if at least one auxiliary output of the SPS30 or the concentration of particles in the outdoor air changed in the same direction. Temperature and relative humidity were compared with the duplicate channels of the SCD41, with a tolerance equal to the sum of the stated uncertainties of the two sensors (1.3°C and 9%). Changes in CO_2 concentration were confirmed solely on the basis of their persistence over time, as there is no duplicate channel for this parameter.

After checking the individual records for each direct input i , the data coverage was determined within a time window ending at time t :

$$\rho_{i,w}(t) = \frac{N_{i,w}^{usable}(t)}{N_{i,w}^{expected}(t)}, \quad (2)$$

where $\rho_{i,w}(t)$ is the coverage coefficient for direct input i over an interval of duration w ending at time t . The value of $\rho_{i,w}(t)$ lies in the interval from 0 to 1 and indicates the completeness of the usable data. $N_{i,w}^{usable}(t)$ is the number of observations with a VALID status or a confirmed SUSPECT status, while $N_{i,w}^{expected}(t)$ is the number of records expected based on the measurement frequency and window duration.

The input indicator was passed to the fuzzy inference unit only when the coverage coefficient ρ_i was at least 0.80. For a window duration of 15 minutes and a recording interval of 30 seconds, the expected number of records is 30. Therefore, this threshold corresponds to at least 24 usable records. If the condition was not met, the reading was considered unavailable within that window.

The index was updated every 5 minutes based on measurements from the preceding 15 minutes. This window reduced the impact of short-term fluctuations. For each available reading i within a window of duration w , a time-weighted average was calculated:

$$\bar{c}_{i,w}(t) = \frac{\sum_{k=1}^n c_i(t_k) \Delta t_k u_i(t_k)}{\sum_{k=1}^n \Delta t_k u_i(t_k)}, \quad (3)$$

where $i \in \{PM_{2.5}, PM_{10}, CO_2, T, RH\}$ denotes a direct input; w – aggregation window duration; t is the moment of calculation; k – numbers the observations within this interval; $c_i(t_k)$ is the measured value of parameter i at time t_k ; Δt_k – the duration of the interval during which this value was considered representative; $u_i(t_k)$ – the observation validity indicator; n – the number of records in the time window. The value $u_i(t_k) = 1$ was assigned to the states VALID and confirmed SUSPECT, while the value $u_i(t_k) = 0$ was assigned to INVALID and MISSING. Aggregation took into account only intervals with valid measurements.

Direct inputs, component assessments and the integrated index were classified into four ordered classes: FAVOURABLE, ACCEPTABLE, DEGRADED and CRITICAL. The common structure of the classes ensured a consistent interpretation of the component and summary assessments. The reference values for $PM_{2.5}$ concentrations were 15, 25 and 50 $\mu g/m^3$, and for PM_{10} , – 45, 75 and 100 $\mu g/m^3$.

CO_2 concentration was considered an indirect indicator of air exchange, the interpretation of which depends on the number of people, the intensity of ventilation and the characteristics of the room [7]. Values of 800, 1,000 and 1,500 ppm were taken as working reference

points for the ventilation component, rather than as universal regulatory limits.

The RH control zones have been established in accordance with DBN V.2.5-67:2013 [1]. Temperature limits were adopted as working parameters for classification, using the numerical guidelines in Table A.2 of DSTU B EN 15251:2011 [2], which contains winter minimums and summer maximums for the operational temperature of premises. The resulting classes characterised deviations from the accepted reference ranges, without establishing normative compliance of the microclimate.

Table 1

Class boundaries for input parameters and the integrated index

Parameter	FAVOURABLE	ACCEPTABLE	DEGRADED	CRITICAL
PM _{2.5} , µg/m ³	< 15	15 ≤ PM _{2.5} < 25	25 ≤ PM _{2.5} < 50	≥ 50
PM ₁₀ , µg/m ³	< 45	45 ≤ PM ₁₀ < 75	75 ≤ PM ₁₀ < 100	≥ 100
CO ₂ , ppm	< 800	800 ≤ CO ₂ < 1000	1000 ≤ CO ₂ < 1500	≥ 1500
RH, %	30 ≤ RH ≤ 50	25 ≤ RH < 30 or 50 < RH ≤ 60	20 ≤ RH < 25 or 60 < RH ≤ 70	RH < 20 or RH > 70
T, °C	21 ≤ T ≤ 25.5	20 ≤ T < 21 or 25.5 < T ≤ 26	18 ≤ T < 20 or 26 < T ≤ 27	T < 18 or T > 27
Index I	0 ≤ I < 25	25 ≤ I < 50	50 ≤ I < 75	75 ≤ I ≤ 100

The width of the transition region for each parameter is set to a fixed operating value, taking into account the accuracy characteristics of the relevant sensor: 5 µg/m³ for PM_{2.5} and 25 µg/m³ for PM₁₀ according to the SPS30 documentation [11], 70 ppm for CO₂ according to the SCD41 documentation [12] (50 ppm + 2.5% of the reading at concentrations around 800 ppm), 3% for RH and 0.5°C for T according to the BME688 documentation [13]. The transition region is situated symmetrically about the reference value, so at the boundary point itself, the membership degrees for both neighbouring classes are 0.5. The integrated index is not linked to any sensor channel, so the transition regions between its classes are defined by a fixed width of 2 points.

The classes were defined using trapezoidal, triangular and shoulder membership functions with values ranging from 0 to 1. Their parameters were determined by the thresholds and the width of the transition regions, within which a value could partially belong to two neighbouring classes. The trapezoidal function took the form:

$$\mu(x; a, b, c, d) = \begin{cases} 0, & x \leq a, \\ \frac{x-a}{b-a}, & a < x < b, \\ 1, & b \leq x \leq c, \\ \frac{d-x}{d-c}, & c < x < d, \\ 0, & x \geq d, \end{cases} \quad (4)$$

where $\mu(x; a, b, c, d)$ is a degree of membership of the value x in the corresponding linguistic category. The parameters a and d are the points beyond which membership of the class is 0, and the indices b and c are the boundaries of the region of full membership, which is 1. The extreme classes were defined by step functions.

The indices T and RH were assessed in both directions, as both low and high values could be unfavourable. The direction of increasing concentration was used only for PM_{2.5}, PM₁₀ and CO₂.

The component scores A , V , M and integrated index I were calculated according to formula (1). Two-input fuzzy inference using the Mamdani method was applied to the aerosol and microclimatic components. The ventilation component had a single direct input, CO₂, so the degrees of its membership in the four linguistic classes were directly transferred to the corresponding output classes of component V .

The classes FAVOURABLE, ACCEPTABLE, DEGRADED and CRITICAL corresponded to ranks 0–3. The rank of a rule's consequence was equal to the highest rank of its premises. The rule bases for the aerosol and microclimatic components each contained 16 rules, whilst the second level, comprising the full set of components, contained 64 rules.

The activation strength of a rule characterised the degree to which all its preconditions were simultaneously fulfilled. For each rule, the activation strength α_r was defined as the minimum degree of membership among the preconditions:

$$\alpha_r = \min_j \{ \mu_{L_{rj}}(x_j) \}, \quad (5)$$

where α_r – the activation strength of fuzzy rule r ; x_j – the value of the j input; L_{rj} – the linguistic class of the j input in the antecedent of rule r ; $\mu_{L_{rj}}$ – the degree of membership of the value x_j in this category. The output fuzzy sets of all activated rules were combined using the maximum operator:

$$\mu_{\text{out}}(z) = \max_r \left[\min \left(\alpha_r, \mu_{B_r}(z) \right) \right], \quad (6)$$

where $\mu_{out}(z)$ – the aggregated output membership function; B_r – the output linguistic class of rule r ; μ_{B_r} – the degree of membership of value z in the output category B_r . The output variable z was defined on a dimensionless scale from 0 to 100, common to the component scores A , V , M and the integrated index I . The numerical value of I was obtained using the centre of gravity method for the area under the aggregated output membership function:

$$I = \frac{\int_0^{100} z \mu_{out}(z) dz}{\int_0^{100} \mu_{out}(z) dz} \quad (7)$$

The integrated index class was determined according to the intervals in Table 1. A value that exactly matched the transition point was assigned to the less favourable class.

A component was considered accessible provided there was sufficient coverage of all its inputs. The presence of three components corresponded to the 'OK' status, and two components to the 'PARTIAL' status. For the 'PARTIAL' status, a two-input rule base was applied, prioritising the most unfavourable class, without substituting for the missing component. The result contained a list of inaccessible inputs and characterised only the evaluated components of the environment. If no more than one component was present, the status was set to FAILED, and no index or class was generated. The sequence for generating the result is shown in Figure 1.

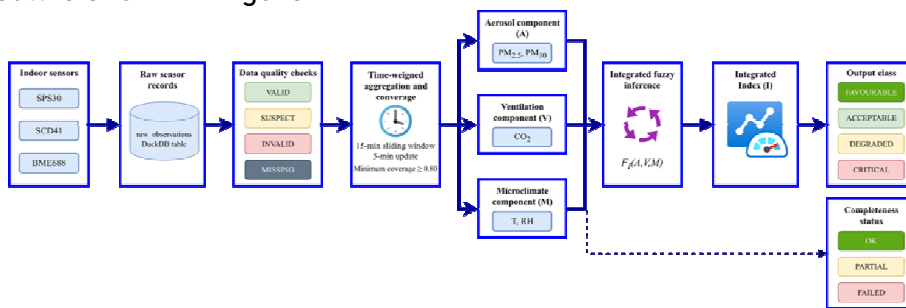


Figure 1. Workflow of sensor data validation, component assessment, and determination of the integrated index, environmental class and data completeness status

Comparative evaluation methodology. The proposed PROPOSED-HFIS method was compared with three configurations. CRISP-CLASS-MAX identified the worst-performing class of inputs based on the crisp boundaries in Table 1, and used the midpoints of the class intervals (12.5, 37.5, 62.5 and 87.5) for numerical comparison. WEIGHTED-MEAN

calculated $(A + V + M)/3$, and FUZZY-COMPONENT-MAX determined $\max(A, V, M)$ to verify the contribution of the second level of fuzzy inference.

The study comprised three tasks. The first involved determining the distributions of statuses and classes, the contribution of components, and the frequency of a drop in the critical class during aggregation. Real-world data and a regular grid of A , V and M with a step size of 10, comprising 1,331 combinations, were used. For the 15,242 results with an 'OK' status, Pearson's correlations were calculated between I and M , and between I and $P_{\max} = \max(A, V)$. A pollution estimate I_P was calculated separately using the same second-level inference, but only for the components A and V .

The second task involved checking 64 second-level rules, the regular grid and the neighbourhoods of the control limits. The third task involved the artificial removal of T or RH from complete calculated moments, in 360 cases each. Evaluations without the unavailable component were compared with the substitution of the most favourable, average and last valid values dating back up to 10 minutes. These substitutions were used solely as comparative strategies.

Research results. Among 16,992 calculation points, the status 'OK' was assigned to 15,242 (89.70%), 'PARTIAL' to 1,004 (5.91%) and 'FAILED' to 746 (4.39%). The proportion of values with a 'SUSPECT' status was 3.8% for $PM_{2.5}$, 3.6% for PM_{10} , 7.6% for CO_2 , 7.7% for RH and 27.0% for T, whereas the proportion of 'INVALID' values for all channels did not exceed 4.7%.

The distribution of classes by 'OK' status is shown in Table 2. The 'CRITICAL' class was assigned to 6,841 integral results (44.9%) and 521 (3.4%) I_P estimates. The correlation between I and M was 0.974, while the correlation between I and $P_{\max}$ was 0.296, indicating that the microclimatic component had a predominant influence in the dataset under study.

The WEIGHTED-MEAN configuration produced a more favourable class in 6,798 out of 6,963 cases with a critical component on real data (97.6%) and in 699 out of 819 corresponding regular grid combinations (85.3%). Figure 2 illustrates the difference in class rankings between PROPOSED-HFIS and WEIGHTED-MEAN. The horizontal and vertical axes represent the A and V scores on a scale of 0–100 in increments of 10, whilst the three panels correspond to fixed values of M . A cell value of 0

indicates class agreement, 1 or 2 indicates that the WEIGHTED-MEAN class is more favourable by the corresponding number of levels.

Table 2

Class distribution for results with OK status

Assessment	FAVOURABLE	ACCEPTABLE	DEGRADED	CRITICAL
Integrated index I	23.2%	13.0%	18.9%	44.9%
Pollution assessment I_p	78.8%	14.7%	3.1%	3.4%
Microclimate component M	26.0%	11.5%	18.3%	44.1%

The WEIGHTED-MEAN configuration produced a more favourable class in 6,798 out of 6,963 cases with a critical component on real data (97.6%) and in 699 out of 819 corresponding regular grid combinations (85.3%). Figure 2 illustrates the difference in class rankings between PROPOSED-HFIS and WEIGHTED-MEAN. The horizontal and vertical axes represent the A and V scores on a scale of 0–100 in increments of 10, whilst the three panels correspond to fixed values of M . A cell value of 0 indicates class agreement, 1 or 2 indicates that the WEIGHTED-MEAN class is more favourable by the corresponding number of levels.

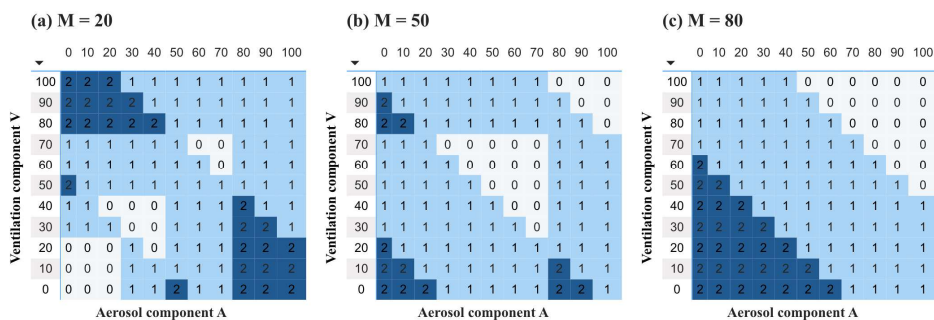


Figure 2. Difference in class ranks between PROPOSED-HFIS and WEIGHTED-MEAN as a function of the aerosol (A) and ventilation (V) components: (a) $M = 20$, (b) $M = 50$ and (c) $M = 80$

In the example where $A = 20$, $V = 20$ and $M = 80$, the average value is 40 and corresponds to the ACCEPTABLE class, while PROPOSED-HFIS generates the CRITICAL class. Therefore, the corresponding cell in the

panel (c) contains the value 2. For $A = V = M = 80$, both configurations generate the CRITICAL class, which corresponds to the value 0.

Near the 33 control limits across 99 scenarios, the average maximum change in the index between neighbouring input values was 1.74 points for PROPOSED-HFIS and 14.90 points for CRISP-CLASS-MAX. The same value of 1.74 points was obtained for FUZZY-COMPONENT-MAX, indicating that first-level fuzzy membership functions ensure a reduction in local variations. The PROPOSED-HFIS and FUZZY-COMPONENT-MAX classifications coincided completely across 15,242 real-world results, whilst on a grid comprising 101 values along each axis, the agreement was 98.34%. All 17,101 discrepancies occurred only when the largest component was exactly 75. Thus, the second level did not provide an advantage in terms of class-based results. Its role was to form an assessment based on a common set of rules for both complete and incomplete sets of components. The median processing cycle time on the Raspberry Pi 5 was 1.62 s, or 0.54% of the update interval.

The results of the comparison of strategies under artificial removal of the input are presented in Table 3. Removal of T or RH rendered the microclimatic component unavailable. The assessment of the remaining components differed from the full assessment by an average of 43.60 points and failed to reproduce 89.1% of its critical classes. Therefore, the PARTIAL result cannot be interpreted as a complete assessment of the environment. Substituting a favourable RH value preserved the critical classes, while a similar substitution of T failed to reproduce 89.1% of them, indicating the decisive influence of temperature in the selected cases. In terms of membership degrees, temperature was associated with a critical class in 72.6% of critical results. Substituting the last suitable T value reduced this proportion to 1.8%, with an average absolute error of 0.89 points.

Conclusions. A method has been developed for the integrated assessment of indoor air quality and microclimate, which combines the validation of sensor data, hierarchical fuzzy aggregation of aerosol, ventilation and microclimatic components, and the determination of an environmental condition class together with the completeness status of the result. Based on data from 59 days of monitoring (169,331 records), out of 16,992 calculations, 89.70% received an 'OK' status, 5.91% a 'PARTIAL' status and 4.39% a 'FAILED' status.

Table 3

Comparison of strategies after artificial removal of an input (360 cases per input)

Input	Strategy	Mean absolute error, index points	Class agreement, %	Missed critical cases, %
RH	Excluding the unavailable component	43.60	20.6	89.1
	Favourable value	0.26	98.9	0.0
	Mean value	0.26	98.9	0.0
	Last valid value	0.01	99.4	0.0
T	Excluding the unavailable component	43.60	20.6	89.1
	Favourable value	42.01	21.9	89.1
	Mean value	21.08	25.6	89.1
	Last valid value	0.89	97.2	1.8

Generating an assessment based on incomplete data sets without replacing missing values, in combination with the 'OK', 'PARTIAL' and 'FAILED' statuses, enabled the distinction between complete and incomplete results without the use of artificially generated data. A comparative strategy of carrying forward the last valid temperature value reduced the proportion of unreproducible critical classes from 89.1% to 1.8% , but did not restore the actual completeness of the measurements. Therefore, this result also requires a 'PARTIAL' status.

Classification of the environmental state, prioritising the most unfavourable class and with transition region widths determined by sensor measurement error, preserved the influence of the most critical component in the final assessment. In contrast, averaging the component assessments resulted in a more favourable class in 97.6% of cases with a critical component in the actual data and in 85.3% of the corresponding combinations of the regular grid. Compared with crisp classification, the average maximum local change in the index near the control limits decreased from 14.90 to 1.74 points.

In the studied premises, the critical values of the integrated index were determined predominantly by the microclimatic component: the proportion of the CRITICAL class was 44.9% for the integrated index and 3.4% for the air pollution assessment, whereas the correlation coefficient between the index and the microclimatic component was 0.974. The median duration of a full processing cycle on the Raspberry Pi 5 was 1.62 s, confirming the method's suitability for continuous monitoring on a peripheral computing node.

Further research is aimed at refining the method in terms of fuzzy classification parameters, handling incomplete data and adapting it to different indoor operating conditions.

1. DBN V.2.5-67:2013. Heating, ventilation and air conditioning. Kyiv : Ministry of Regional Development, Construction, Housing and Communal Services of Ukraine, 2013. [In Ukrainian]. **2.** DSTU B EN 15251:2011. Indoor environmental input parameters for design and assessment of energy performance of buildings addressing indoor air quality, thermal environment, lighting and acoustics (EN 15251:2007, IDT). Kyiv : Ministry of Regional Development, Construction, Housing and Communal Services of Ukraine, 2012. [In Ukrainian]. **3.** Palamar A., Karpinski M., Palamar M., Osukhivska H., Mytnyk M. Remote air pollution monitoring system based on Internet of Things // CEUR Workshop Proceedings. 2022. Vol. 3309. P. 194–204. URL: <https://ceur-ws.org/Vol-3309/paper14.pdf>. **4.** Cony Renaud-Salis L., Ramalho O., Abadie M. Towards the definition of an indoor air quality index for residential buildings based on long- and short-term exposure limit values // International Journal of Ventilation. 2020. Vol. 19, No. 3. P. 189–200. DOI: 10.1080/14733315.2019.1615218. **5.** Wei W., Sivanantham S., Desvignes V., Dassonville C., Grégoire A., Lafaurie E., Samri D., Ramalho O. Development of an indoor air quality index for continuous monitoring in French dwellings // E3S Web of Conferences. 2025. Vol. 672. Article 07009. DOI: 10.1051/e3sconf/202567207009. **6.** World Health Organization. WHO global air quality guidelines: particulate matter (PM_{2.5} and PM₁₀), ozone, nitrogen dioxide, sulfur dioxide and carbon monoxide. Geneva : World Health Organization, 2021. URL: <https://www.who.int/publications/i/item/9789240034228>. **7.** Persily A. Development and application of an indoor carbon dioxide metric // Indoor Air. 2022. Vol. 32, No. 7. Article e13059. DOI: 10.1111/ina.13059. **8.** Fortuna F., Naccarato A., Terzi S. Functional association measures among air quality index components // Stochastic Environmental Research and Risk Assessment. 2025. Vol. 39, No. 7. P. 3173–3186. DOI: 10.1007/s00477-025-03015-y. **9.** Colella Y., Valente A. S., Rossano L., Trunfio T. A., Fiorillo A., Improta G. A fuzzy inference system for the assessment of indoor air quality in an operating room to prevent surgical site infection // International Journal of Environmental

Research and Public Health. 2022. Vol. 19, No. 6. Article 3533. DOI: 10.3390/ijerph19063533. **10.** Saini J., Dutta M., Marques G. ADFIST: adaptive dynamic fuzzy inference system tree driven by optimized knowledge base for indoor air quality assessment // *Sensors*. 2022. Vol. 22, No. 3. Article 1008. DOI: 10.3390/s22031008. **11.** Sensirion AG. Datasheet SPS30: particulate matter sensor for air quality monitoring and control. Version 2.0-D1. Staefa : Sensirion AG, 2023. URL: https://sensirion.com/media/documents/8600FF88/64A3B8D6/Sensirion_PM_Sensors_Datasheet_SPS30.pdf. **12.** Sensirion AG. SCD4x: breaking the size barrier in optical CO2 sensing : datasheet. Version 1.7. Staefa : Sensirion AG, 2025. URL: https://sensirion.com/media/documents/48C4B7FB/67FE0194/CD_DS_SCD4x_Datasheet_D1.pdf. **13.** Bosch Sensortec GmbH. BME688: digital low power gas, pressure, temperature & humidity sensor with AI : datasheet. Revision 1.3. Document No. BST-BME688-DS000-03. Reutlingen : Bosch Sensortec GmbH, 2024. URL: <https://www.bosch-sensortec.com/media/boschsensortec/downloads/datasheets/bst-bme688-ds000.pdf>. **14.** Dionova B. W., Mohammed M. N., Al-Zubaidi S., Yusuf E. Environment indoor air quality assessment using fuzzy inference system // *ICT Express*. 2020. Vol. 6, No. 3. P. 185–194. DOI: 10.1016/j.icte.2020.05.007. **15.** Rajeswari Subramaniam K., Cheng C.-T., Pang T. Y. Fuzzy logic controlled simulation in regulating thermal comfort and indoor air quality using a vehicle heating, ventilation, and air-conditioning system // *Sensors*. 2023. Vol. 23, No. 3. Article 1395. DOI: 10.3390/s23031395. **16.** Hayward I., Martin N. A., Ferracci V., Kazemimanesh M., Kumar P. Low-cost air quality sensors: biases, corrections and challenges in their comparability // *Atmosphere*. 2024. Vol. 15, No. 12. Article 1523. DOI: 10.3390/atmos15121523. **17.** Lee J.-Y., Han S.-H., Kang J.-G., Lee C.-Y., Lee J.-B., Kim H.-S., Yun H.-Y., Choi D.-R. Comparison of models for missing data imputation in PM-2.5 measurement data // *Atmosphere*. 2025. Vol. 16, No. 4. Article 438. DOI: 10.3390/atmos16040438. **18.** Borodii I., Osukhivska H. Research on the efficiency of data loading and storage in Data Lakehouse architectures for the formation of analytical data systems // *Information Technology: Computer Science, Software Engineering and Cyber Security*. 2025. No. 4. P. 28–36. DOI: 10.32782/IT/2025-4-4. **19.** Roos-Hoefgeest Toribio M., Garnung Menéndez A., Roos-Hoefgeest Toribio S., Álvarez García I. A novel approach to speed up Hampel filter for outlier detection // *Sensors*. 2025. Vol. 25, No. 11. Article 3319. DOI: 10.3390/s25113319.

REFERENCES:

1. DBN V.2.5-67:2013. Heating, ventilation and air conditioning. Kyiv : Ministry of Regional Development, Construction, Housing and Communal Services of Ukraine, 2013. [In Ukrainian].
2. DSTU B EN 15251:2011. Indoor environmental input parameters for design and assessment of energy performance of buildings addressing indoor air quality, thermal environment, lighting and acoustics (EN 15251:2007, IDT). Kyiv : Ministry of Regional Development, Construction, Housing and Communal Services of Ukraine, 2012. [In Ukrainian].
3. Palamar A., Karpinski M., Palamar M., Osukhivska H., Mytnyk M. Remote air pollution monitoring system based on Internet of Things // CEUR Workshop Proceedings. 2022. Vol. 3309. P. 194–204. URL: <https://ceur-ws.org/Vol-3309/paper14.pdf>.
4. Cony Renaud-Salis L., Ramalho O., Abadie M. Towards the definition of an indoor air quality index for residential buildings based on long- and short-term exposure limit values // International Journal of Ventilation. 2020. Vol. 19, No. 3. P. 189–200. DOI: 10.1080/14733315.2019.1615218.
5. Wei W., Sivanantham S., Desvignes V., Dassonville C., Grégoire A., Lafaurie E., Samri D., Ramalho O. Development of an indoor air quality index for continuous monitoring in French dwellings // E3S Web of Conferences. 2025. Vol. 672. Article 07009. DOI: 10.1051/e3sconf/202567207009.
6. World Health Organization. WHO global air quality guidelines: particulate matter (PM_{2.5} and PM₁₀), ozone, nitrogen dioxide, sulfur dioxide and carbon monoxide. Geneva : World Health Organization, 2021. URL: <https://www.who.int/publications/i/item/9789240034228>.
7. Persily A. Development and application of an indoor carbon dioxide metric // Indoor Air. 2022. Vol. 32, No. 7. Article e13059. DOI: 10.1111/ina.13059.
8. Fortuna F., Naccarato A., Terzi S. Functional association measures among air quality index components // Stochastic Environmental Research and Risk Assessment. 2025. Vol. 39, No. 7. P. 3173–3186. DOI: 10.1007/s00477-025-03015-y.
9. Colella Y., Valente A. S., Rossano L., Trunfio T. A., Fiorillo A., Improta G. A fuzzy inference system for the assessment of indoor air quality in an operating room to prevent surgical site infection // International Journal of Environmental Research and Public Health. 2022. Vol. 19, No. 6. Article 3533. DOI: 10.3390/ijerph19063533.
10. Saini J., Dutta M., Marques G. ADFIST: adaptive dynamic fuzzy inference system tree driven by optimized knowledge base for indoor air quality assessment // Sensors. 2022. Vol. 22, No. 3. Article 1008. DOI: 10.3390/s22031008.
11. Sensirion AG. Datasheet SPS30: particulate matter sensor for air quality monitoring and control. Version 2.0-D1. Staefa : Sensirion AG, 2023. URL: https://sensirion.com/media/documents/8600FF88/64A3B8D6/Sensirion_PM_Sensors_Datasheet_SPS30.pdf.
12. Sensirion AG. SCD4x: breaking the size barrier in optical CO₂ sensing : datasheet. Version 1.7. Staefa : Sensirion AG, 2025. URL: https://sensirion.com/media/documents/48C4B7FB/67FE0194/CD_DS_SCD4x_Datasheet_D1.pdf.
13. Bosch Sensortec GmbH. BME688: digital low power

gas, pressure, temperature & humidity sensor with AI : datasheet. Revision 1.3. Document No. BST-BME688-DS000-03. Reutlingen : Bosch Sensortec GmbH, 2024. URL: <https://www.bosch-sensortec.com/media/boschsensortec/downloads/datasheets/bst-bme688-ds000.pdf>. **14.** Dionova B. W., Mohammed M. N., Al-Zubaidi S., Yusuf E. Environment indoor air quality assessment using fuzzy inference system // ICT Express. 2020. Vol. 6, No. 3. P. 185–194. DOI: 10.1016/j.icte.2020.05.007. **15.** Rajeswari Subramaniam K., Cheng C.-T., Pang T. Y. Fuzzy logic controlled simulation in regulating thermal comfort and indoor air quality using a vehicle heating, ventilation, and air-conditioning system // Sensors. 2023. Vol. 23, No. 3. Article 1395. DOI: 10.3390/s23031395. **16.** Hayward I., Martin N. A., Ferracci V., Kazemimanesh M., Kumar P. Low-cost air quality sensors: biases, corrections and challenges in their comparability // Atmosphere. 2024. Vol. 15, No. 12. Article 1523. DOI: 10.3390/atmos15121523. **17.** Lee J.-Y., Han S.-H., Kang J.-G., Lee C.-Y., Lee J.-B., Kim H.-S., Yun H.-Y., Choi D.-R. Comparison of models for missing data imputation in PM-2.5 measurement data // Atmosphere. 2025. Vol. 16, No. 4. Article 438. DOI: 10.3390/atmos16040438. **18.** Borodii I., Osukhivska H. Research on the efficiency of data loading and storage in Data Lakehouse architectures for the formation of analytical data systems // Information Technology: Computer Science, Software Engineering and Cyber Security. 2025. No. 4. P. 28–36. DOI: 10.32782/IT/2025-4-4. **19.** Roos-Hoefgeest Toribio M., Garnung Menéndez A., Roos-Hoefgeest Toribio S., Álvarez García I. A novel approach to speed up Hampel filter for outlier detection // Sensors. 2025. Vol. 25, No. 11. Article 3319. DOI: 10.3390/s25113319.

Бородій І. І., ⁽¹⁾

аспірант

Осухівська Г. М., ^(1:ORCID: 0000-0003-0132-1378)

к.т.н., доцент

Сверстюк А. С., ^(2:ORCID: 0000-0003-3146-9664)

д.т.н., професор

Волощук А. В., ^(1:ORCID: 0009-0007-1478-1601)

PhD

Жаровський Р. О., ^(:ORCID: 0000-0002-1159-3417)

к.т.н.

¹Тернопільський національний технічний університет імені І. Пулюя,
м. Тернопіль

²Тернопільський національний медичний університет імені І.Я.
Горбачевського, м. Тернопіль

ІНТЕГРАЛЬНЕ ОЦІНЮВАННЯ ЯКОСТІ ПОВІТРЯ ТА МІКРОКЛІМАТУ ПРИМІЩЕННЯ З УРАХУВАННЯМ ПОВНОТИ СЕНСОРНИХ ДАНИХ

Було розроблено метод інтегрального оцінювання якості повітря та мікроклімату приміщення шляхом поєднання перевірки сенсорних даних з ієрархічним нечітким агрегуванням. Аерозольні, вентиляційні та мікрокліматичні компоненти було отримано на основі вимірювань показників $PM_{2.5}$, PM_{10} , CO_2 , температури та відносної вологості. Кожному вимірюванню було присвоєно статус достовірності, а кожна точка оцінки отримала статус повноти «OK», «PARTIAL» або «FAILED» без заміщення відсутніх значень. Метод було оцінено на основі 169 331 запису, зібраних протягом 59 днів, що охоплювали 16 992 заплановані точки оцінки. З них 89,70 % отримали статус «OK», 5,91 % — статус «PARTIAL» та 4,39 % — статус «FAILED». Усереднення компонентів давало більш сприятливий клас у 97,6 % випадків, коли компонент належав до найменш сприятливого класу, тоді як запропонована агрегація зберігала його вплив. Медіанний час обробки, що становив 1,62 с на Raspberry Pi 5, підтвердив придатність методу для безперервного моніторингу.

Ключові слова: якість повітря в приміщеннях; моніторинг навколишнього середовища; низьковартісні сенсори; Інтернет речей; нечітка логіка; аналіз даних; метод оцінювання.

Отримано: 29 березня 2026 року
Прорецензовано: 24 квітня 2026 року
Прийнято до друку: 29 травня 2026 року



© 2026 [Borodii Ivan, Osukhivska Halyna, Sverstiuk Andrii, Voloshchuk Andrii, Zharovskyi Ruslan]. Licensee [NUWEE]. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution-NonCommercial (CC BY-NC) license ([creativecommons.org](https://creativecommons.org/licenses/by-nc/4.0/)).