

# ESTABLISHING AN AIR QUALITY INDEX BASED ON PROXY DATA FOR URBAN PLANNING: FINALIZATION, VALIDATION AND APPLICATION TO A CASE STUDY

C. Falzone <sup>a,b</sup>, A.C. Romain <sup>a</sup>

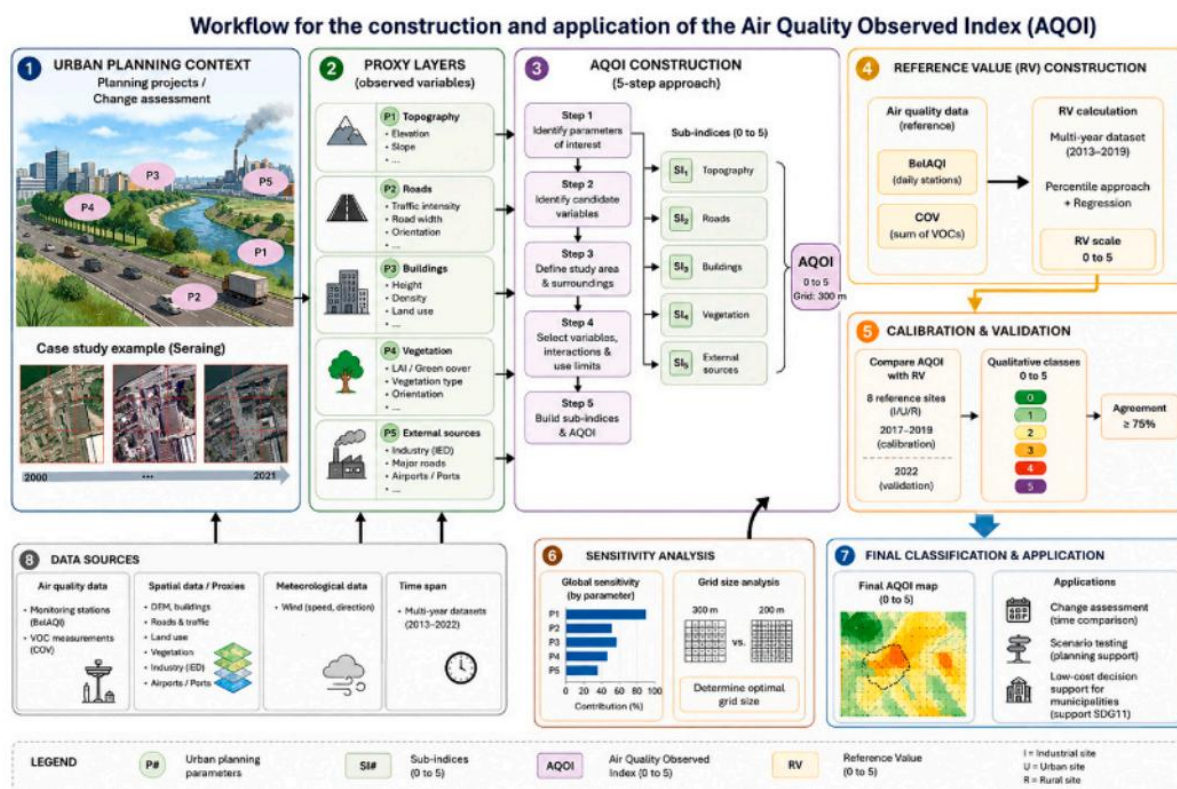
<sup>a</sup> Research Unit SPHERE, Department of Environmental Sciences and Management, University of Liege, 4000, Liège, Belgium

<sup>b</sup> University of Liege, UR SPHERES, Sensing of Atmospheres and Monitoring, Belgium

## HIGHLIGHTS

- Proxy-based AQ index for 300 m urban grids.
- AQOI integrates five proxies into six AQ classes.
- Health reference merges BelAQI with total VOCs.
- Eight-site validation shows 75% health- based agreement.
- Seraing 2000–2021 case tracks redevelopment impacts

## GRAPHICAL ABSTRACT



## Keywords:

Air quality index, Urban planning, Air pollution, Proxy data, Local scale

## Abstract:

Urban planning must increasingly integrate air quality considerations in alignment with Sustainable Development Goal 11. However, existing tools for assessing the impact of urban developments on air quality are either technically demanding or dependent on post-construction measurements. This study presents the finalization, validation, and application of a spatially explicit air quality index based on proxy data. The index aggregates five normalized parameters—topography, roads, buildings, vegetation, and external sources—into a single annual score ranging from 0 (excellent) to 5 (very poor) on a 300 m grid. A novel reference value (RV), combining the BelAQI index and total VOC concentrations, was used to calibrate index thresholds via regression and percentile optimization. Validation on eight monitoring sites yielded a 75% agreement rate with the RV classification. The index was then applied to the city of Seraing, Belgium, revealing a slight decrease in structural pressure on air quality over two decades despite significant urban changes, while remaining in the highest impact class. It offers a low-cost, accessible decision-support tool for municipalities, providing preemptive insight into the air quality impacts of land use planning. It stands as a complementary indicator to traditional air quality metrics, with strong potential for adaptation and future refinement.

## 1. Introduction

Over the years, air pollution has become a significant environmental and public health issue in several regions of the world (World Health Organization, 2023). Since 2016, the International Agency for Research on Cancer has classified air pollution as carcinogenic to humans (Vol. 109) (IARC Working Group on the Evaluation of Carcinogenic Risks to Humans, 2016). Following critical incidents linked to air quality (Fowler et al., 2020), a global air quality monitoring network was established. At the European level, this network is governed by Directive 2024/2881, which particularly incorporates the guideline values issued by the World Health Organization concerning the health impacts of various atmospheric pollutants (European union, 2024; *WHO Global Air Quality Guidelines*, 2021).

This network comprises monitoring stations dedicated to real-time air quality assessment. Each country is responsible for maintaining its own network through governmental agencies tasked with measurement, data quality assurance, and communication. These agencies also produce spatial representations of pollutant dispersion through air quality models (*Air Quality Models*, n.d.). These models provide real-time or short-term forecasts of air quality across various territorial scales and are based on measurement data, land use information, and meteorological parameters. In parallel, a rapidly growing body of work uses advanced machine-learning architectures such as LSTM and GRU networks, CNN–RNN hybrids and federated learning frameworks to forecast air-quality indices or pollutant concentrations in smart-city contexts, focusing mainly on short-term prediction from

monitoring data rather than on planning-oriented proxy indicators (Alwabli, 2024; Chatterjee et al., 2024; Drewil and Al-Bahadili, 2022; Faraji et al., 2022).

In recent years, miniaturized air quality monitoring stations have been added to the existing network. Although the technologies used are less expensive, they tend to be less robust and less accurate than instruments integrated into official monitoring systems (AirParif and AIRLAB, 2023). Nevertheless, their deployment contributes to refining the monitoring grid by offering localized insights into air quality at the neighborhood level (pollution niches) (ISSEP, n.d.; *Micro-capteurs*, 2022).

Since 2015, the 17 Sustainable Development Goals (SDGs) have served as a guiding framework for many countries worldwide (Neshovski, n.d.). Among these, SDG 11 focuses on sustainable cities and communities. One of its specific targets states: “By 2030, reduce the adverse per capita environmental impact of cities, including by paying special attention to air quality and municipal and other waste management” (*Cities - United Nations Sustainable Development Action 2015*, 2015). To meet this objective, cities receive support from various financial mechanisms to improve air quality, among other priorities (*Sustainable Finance Hub*, 2025).

Numerous strategies have been implemented, each with varying economic and social implications. Some focus on traffic management, including speed reductions, restrictions based on vehicle age (low-emission zones), and the promotion of public and alternative transport modes (Bigazzi and Rouleau, 2017; Pelgrims et al., 2024). Other strategies emphasize reducing pollutant emissions from buildings, particularly through improved insulation or heating systems (Mahmoud et al., 2021). In such cases, citizens play a central role in the transition. In heavily industrialized cities, some initiatives aim to rehabilitate disused industrial sites and transform them into neighborhoods that align with SDG 11 requirements (*Réhabiliter les friches*, 2022).

Despite the availability of tools such as measurement networks and modelling systems, municipalities face significant challenges in meeting SDG 11's sixth sub-goal. Air quality monitoring networks demand continuous maintenance and costly instrumentation. Meanwhile, air quality modelling requires specialized resources and expertise, which are not always accessible to local administrations, often forcing them to outsource studies related to their environmental impact. Consequently, despite the availability of effective technical tools, urban planners often struggle to make informed decisions regarding neighborhood redevelopment aligned with SDG 11 objectives. Real-time measurement systems are not suited for evaluating unbuilt projects, and air quality modelling capabilities are typically lacking within municipal teams.

In this context, the idea emerged to develop a practical and accessible decision-support tool to help urban planners design projects with minimal impact on air quality. The aim of the present research is to develop a decision-making tool, hereafter referred to as an air quality observed index based on proxy data for urban planning, intended for use in new urban development projects by municipal authorities. The objective is to provide a low-cost, user-friendly indicator that does not require specific expertise in air quality modelling or atmospheric chemistry, and that reflects the structural pressure exerted by urban form, transport infrastructure, vegetation and major emission sources at neighborhood scale. The methodological development of this index was described in a previous paper (Falzone and Romain, 2022), which demonstrated promising results. The present paper complements this work by introducing model upgrades, performing a formal calibration and validation against a measurement-

based reference, and applying the index to a long-term urban redevelopment case study (Falzone, 2025).

## 2. Material and methods

### 2.1. Data

#### 2.1.1. Data used for AQOI calculations

In Europe, the INSPIRE directive (European union, 2007) mandates that public agencies provide access to spatial data related to environmental monitoring. According to Directive 2024/1785, industrial emissions data are available through the European Environment Agency (EEA) website (European Environment Agency, 2025a, 2025b). These datasets can be easily visualized and extracted using Geographic Information System (GIS) tools and thus used to calculate a proxy-based air quality index developed in previous work (Falzone and Romain, 2022).

Data on wind in the study area, including direction and frequency over a year, are available from official weather stations. The *worldmet* package in R (Carslaw, 2025), provides access to a large amount of validated meteorological data from these stations. All proxy variables listed in Table 1 are encoded following a structured set of rules (orientation and position matrices, value classes and thresholds) detailed in Supplementary Material S2, with the aim of ensuring reproducibility and enabling a semi-automated implementation within a GIS workflow.

#### 2.1.2. Data used for reference value (RV)

To develop a qualitative scale to facilitate the interpretation of the proposed index and to validate its performance, two additional official data sets were used to construct a composite reference value specifically for classification and calibration purposes from the official air quality measurement network:

- The BelAQI index, the official Belgian daily air quality index, comprises ten classes and is based on measurements of  $PM_{10}$ ,  $PM_{2.5}$ ,  $O_3$ , and  $NO_2$  (*BelAQI Index*, 2022).
- VOC concentration measurements, based on the list of regulated compounds provided in Annex VI-B of Directive 2024/2881 (European union, 2024).

The index was initially developed from air quality data collected in Wallonia, Belgium, by official monitoring stations. The selection of stations was based on the availability of data on VOCs,  $O_3$ ,  $NO_2$ ,  $PM_{10}$ , and  $PM_{2.5}$ . Thirteen stations were included in the study: one in an industrial area, four in urban or suburban areas, and eight in rural areas (S1. Reference stations - Table S1 and Fig. S1). Data from four specific years (2017, 2018, and 2019 for model development, and 2022 for validation) were used and are publicly available at [www.wallonair.be](http://www.wallonair.be). The calibration dataset thus comprises 13 monitoring stations located in Wallonia, with three years (2017–2019) used for model development and one year (2022) reserved for validation. As a result, the present analysis mainly assesses internal consistency within this regional context, rather than providing a comprehensive validation across all possible urban morphologies and climates.

**Table 1**

Variables comprising each parameter associated with their impact on air quality.

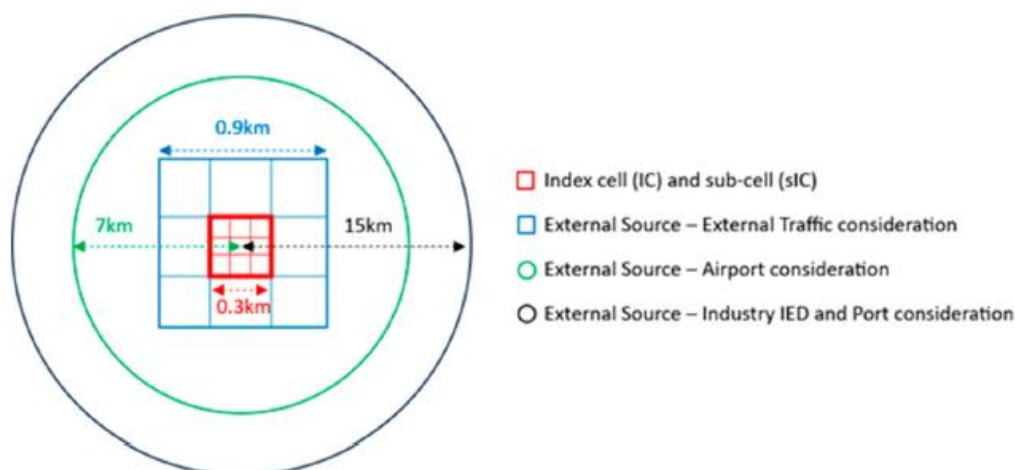
| Parameters & variables                     | Impacts                                                                                    |
|--------------------------------------------|--------------------------------------------------------------------------------------------|
| <b>Topography (<math>P_1</math>)</b>       |                                                                                            |
| Relief                                     | Dispersion of pollutants                                                                   |
| Navigable Water Network                    | Emission of pollutants by boats                                                            |
| <b>Road (<math>P_2</math>)</b>             |                                                                                            |
| Road - Geometry                            | Dispersion of pollutants (street canyon)                                                   |
| Orientation                                | Dispersion of pollutants (street canyon)                                                   |
| Class                                      | Emission of pollutants based on traffic volume                                             |
| Stop                                       | Emission of pollutants                                                                     |
| Electric Park                              | Emission of pollutants                                                                     |
| Traffic management                         | Emission of pollutants                                                                     |
| <b>Buildings (<math>P_3</math>)</b>        |                                                                                            |
| Building - Geometry                        | Dispersion by screen effect                                                                |
| Position in IC                             | Dispersion by screen effect                                                                |
| Heating                                    | Emission of pollutants based on the technology used                                        |
| Gallery                                    | Dispersion                                                                                 |
| Green walls                                | Sink                                                                                       |
| Assignment                                 | Pollutant emissions based on traffic imported into IC according to building classification |
| <b>Vegetation (<math>P_4</math>)</b>       |                                                                                            |
| Vegetation - Canopy                        | Vertical Dispersion                                                                        |
| Geometry & Disposition                     | Sink                                                                                       |
| Type & Deciduous                           | Dispersion by screen effect                                                                |
| <b>External Sources (<math>P_5</math>)</b> |                                                                                            |
| Port                                       | Emission of pollutant                                                                      |
| Airport                                    | Emission of pollutant                                                                      |
| IED                                        | Emission of pollutant                                                                      |
| Road                                       | Emission of pollutant                                                                      |
| <b>Other</b>                               |                                                                                            |
| Wind direction & frequencies               | Dispersion                                                                                 |

## 2.2. Air quality observed index (AQOI)

Based on the Land Used Regression model (LUR) (Briggs et al., 1997), the index is calculated on a 300 m square cell (90,000m<sup>2</sup>- Fig. 1 bold red). This resolution is consistent with neighborhood-scale land-use regression applications and was selected as a pragmatic compromise between capturing local contrasts and ensuring that each cell contains enough roads, buildings and vegetation elements for the proxy-based algorithms to be meaningful; previous methodological tests with finer grids (100–250 m) showed limited additional differentiation compared with 200–300 m, while substantially increasing data preparation effort and sometimes leaving too few elements per cell. A direct comparison between

300 m and 200 m cells indicated that qualitative AQOI class assignments remained identical or very similar for about 90% of the cases, suggesting that the 300 m grid provides a stable and operationally feasible resolution for planning applications. The perimeter considered for variables related to external sources is a 900 m square cell for external road parameters. It is centered on the main cell (Fig. 1 blue), 7 km for airport (Fig. 1 green) and 15 km for the port and IED industries (Fig. 1 black). These buffer distances are consistent with previous methodological work, where a 900 m neighborhood was used to capture the influence of major external roads, a 7 km radius was adopted for airports based on reported impact ranges, and a 15 km radius was derived from Gaussian plume considerations to encompass the main downwind influence of IED-regulated industries and ports under typical stability conditions.

The index is made up of 5 parameters: “Topography” ( $P_1$ ), “Roads” ( $P_2$ ), “Buildings” ( $P_3$ ), “Vegetations” ( $P_4$ ) and “External sources” ( $P_5$ ). Each parameter is defined by an algorithm made up of variables related to air quality. For the building parameter ( $P_3$ ), geometric and positional variables are explicitly designed to reflect key dispersion mechanisms identified in urban flow studies. Buildings are classified into height categories that approximate the expected extent of their wake and turbulence zones, and these height classes are combined with the position of each building within the 300 m cell and with annual wind-direction frequencies through a position–wind matrix. This encoding distinguishes situations where continuous, tall frontages aligned with prevailing winds act as screens that trap pollutants in the cell from configurations where more fragmented or recessed structures create ventilation corridors and facilitate exchange with the overlying boundary layer. Dedicated variables for galleries and facade green walls capture local porosity and the potential for enhanced air exchange or localized pollutant removal near building surfaces, so that  $P_3$  integrates both emission-related aspects (heating, traffic imported by building assignment) and physically grounded dispersion effects. These variables reflect pollutants' dispersion, sources or sinks. They are either qualitative, quantitative or Boolean. The variables used in the index are listed in Table 1. The primary formulations of the index were described in a previous methodological paper (Falzone and Romain, 2022). The updated algorithms for the five parameters are fully specified in Supplementary Material S2, where all equations, variable definitions and encoding matrices are provided to enable direct implementation of the index.



**Fig. 1.** Schematic view of the areas useful to calculate the index within the centre of the Index cell – IC

For the road parameter in particular, road class, geometry and traffic management are used as structured proxies for traffic-related emissions and accumulation potential. Higher-order roads (e.g. national or regional routes) typically combine higher traffic volumes and speed limits than local residential streets, so that road class implicitly encodes contrasted traffic and speed regimes without requiring explicit traffic counts or fleet data for each grid cell. Detailed traffic measurements can be mobilized in dedicated validation exercises to refine these coding rules, but they are not required for routine application of the index, which was designed as a low-cost, easy-to-apply planning tool.

These five parameters were selected to capture the main structural drivers that modulate pollutant dispersion and accumulation at neighborhood scale: valley morphology and water bodies, road network configuration and traffic management, building density and functional assignment, urban greenery, and major external emission sources. For the external sources parameter (P5), industrial facilities regulated under the Industrial Emissions Directive are encoded using both their spatial relationship to the 300 m grid cell and basic information on emitted pollutants. Each installation is characterized by its position relative to the cell and by annual wind-direction frequencies through an orientation–distance matrix, which identifies whether the site lies predominantly upwind or downwind and accounts qualitatively for dilution with increasing distance (up to 15 km). In parallel, reported emission loads and pollutant toxicity classes from the IED database are combined into a source term that differentiates facilities according to the magnitude and hazardousness of their emissions, rather than treating all industrial sites as equivalent point sources. Stack-height details, plume-rise dynamics and hourly operating schedules are not modelled explicitly; instead, their aggregate influence is captured indirectly through this distance–orientation weighting and the use of annual emission totals, consistent with the structural, planning-oriented scope of the AQOI. Most of these drivers evolve slowly compared with meteorological conditions or short-term emission changes, which makes them particularly relevant for assessing the long-term pressure of urban form on air quality in the context of land-use planning. Dynamic atmospheric processes (e.g. seasonal circulation patterns, boundary-layer height, photochemistry) are accounted for indirectly through the calibration against the reference value derived from measured pollutants, rather than being explicitly modelled within the index formulation.

The proxy variables listed in Table 1 are encoded following a structured protocol (orientation and position matrices, value classes and thresholds) described in Supplementary Material S2. This protocol was applied consistently across all sites to enhance reproducibility and to limit subjective interpretation during manual encoding. In the vegetation parameter (P4), this encoding explicitly distinguishes between beneficial and ventilation-blocking greenery. A qualitative canopy variable differentiates between wide and narrow tree canopies at the grid-cell scale: wide canopies are assumed to hinder vertical air exchange within the urban canopy layer and are therefore encoded as having a negative impact on air quality, whereas narrow canopies are treated as ventilation-neutral. In addition, orientation–position matrices combine the spatial arrangement of tall vegetation (e.g. linear roadside plantings versus dense clusters) with prevailing wind directions to determine whether vegetation elements act as barriers that trap pollutants within the cell or as protective screens that limit pollutant inflow from upwind sources. As a result, vegetation is not treated as uniformly beneficial, but contributes to the index as a context-dependent combination of sinks (via deposition) and modifiers of local ventilation.

The final value of the index corresponds to the sum of the normalized equations (S2–S6; Supplementary Material S2). Normalization brings all five parameters to a common 0–1 scale so that no single parameter dominates purely because of its numerical range (Eq. (1)). An equal weighting scheme is adopted at the parameter level, so that each of the five parameters can contribute up to one fifth of the maximum index value. Within each parameter, individual elements (e.g. road segments, buildings, vegetation patches) are already weighted by their relative surface area in the 300 m grid cell through a “relative occupancy” term, so that extensive features such as large building blocks or continuous road corridors have a much stronger influence on the corresponding sub-index than small, isolated elements, while sparse vegetation keeps the vegetation sub-index close to zero. In this way, part of the heterogeneity in contributions is handled at the intra-parameter level, while equal weighting operates only at the higher aggregation level between parameters. The minimum and maximum values used correspond to the 2.5th and 97.5th percentiles from the distributions obtained for each parameter, based on the 2017–2019 dataset described in Section 2.1. This percentile-based approach reduces the influence of extremes and stabilizes the index scale but may induce saturation effects for very high or very low raw sub-index values.

$$(P_i - P_{\min[i]}) / (P_{\max[i]} - P_{\min[i]}) \quad (1)$$

$P_i$  as the value of each  $i$  parameter, with  $P_{\min[i]}$  and  $P_{\max[i]}$  as mini-mum and the maximum value respectively (Table 2).

The index value is expressed in arbitrary units between 0 (very good air quality) and 5 (poor air quality).

### 2.3. Reference value development (RV)

AQOI reflects the state of air quality on an annual basis using a 300 m resolution grid. Therefore, for classification and validation purposes based on a reference value ( $RV$ ), it is essential to consider several pollutants representative of ambient air quality. The BelAQI index was selected for this purpose, as it integrates daily measurements of four key pollutants:  $O_3$ ,  $PM_{10}$ ,  $PM_{2.5}$ , and  $NO_2$ .

However, since the proposed index accounts for a broader range of pollution sources and is not limited to these four pollutants, it was decided to incorporate VOC (volatile organic compounds) concentrations into the reference value as well.

The reference value ( $RV$ ) is calculated by combining the BelAQI index and the VOC concentrations. To estimate the annual air quality status, both BelAQI and VOC values are first aggregated separately and then multiplied together, as shown in Eqs. (2) to (4).

$$S.BelAQI = \sum_{i=1}^{365} BelAQI_i \quad (2)$$

where  $BelAQI_i$  represents the daily index value (ranging from 1 to 10), and  $i$  is the Julian day of the year.

$$S.VOC = \sum_{j=1}^{365} VOC_{[ij]} \quad (3)$$

where  $VOC_{[ij]}$  corresponds to the concentration of each VOC measured on day  $j$ , and  $i$  is the number of VOC species included.

$$RV = S.BelAQI \times S.VOC \quad (4)$$

The resulting reference value ( $RV$ ) thus represents the annual VOC concentration (in  $mg/m^3$ ), weighted by the annual sum of the BelAQI index, which is directly proportional to the concentrations of  $PM_{2.5}$ ,  $PM_{10}$ ,  $O_3$ , and  $NO_2$ . As a result, the  $RV$  increases when both the aggregated BelAQI and total VOC concentrations increase, and can be interpreted as a composite indicator of pollution burden for calibration purposes, without a direct physical unit.

This composite reference value is not intended to represent a physically meaningful or regulatory metric. Rather, it serves as a pragmatic intermediate construct to rank sites on an ordinal scale that reflects both particulate/oxidant pollution (via the BelAQI) and the presence of VOC precursors, for the sole purpose of calibrating the index thresholds and assessing classification consistency. In this sense, the  $RV$  is health-informed, because it is anchored in pollutants and indices that are themselves health-relevant, rather than a direct quantitative health-risk metric.

**Table 2**

Minimum and maximum values used in Eq. (1).

|         | Topography | Roads | Buildings | Vegetations | External sources |
|---------|------------|-------|-----------|-------------|------------------|
| Minimum | −1.00      | +0.00 | −0.00     | −0.60       | +0.00            |
| Maximum | +1.52      | +0.26 | +15.00    | +0.16       | +0.11            |

To construct the index quality scale, a corresponding reference scale must be defined based on the  $RV$ . This scale, defined on an annual basis, is derived from the classification thresholds used for both the BelAQI and VOC concentrations.

## 2.4. Threshold definition for the index and reference value - methodology

The threshold values for both the  $RV$  and the index were developed through a seven-step procedure.

### *Step 1: Determination of S.BelAQI thresholds*

The first step involved identifying threshold values associated with the annual BelAQI index ( $S.BelAQI$ ). To define these thresholds, a theoretical approach was adopted: scenarios were constructed where a single BelAQI value remained constant throughout the year for a given location. For instance, if the daily BelAQI index is equal to 2 for all 365 days, then the annual  $S.BelAQI$  equals 730 ( $2 \times 365$ ). Similarly, a constant daily BelAQI of 3 results in a  $S.BelAQI$  of 1095, and so forth up to a value of 10. These calculations are presented in Table 3, Column B, for all 10 BelAQI levels.

### Step 2: Integration of VOC concentration classes

The second step consisted of linking the *S.BelAQI* values with total VOC concentrations. According to Directive 2024/2881, VOCs are considered precursors of ozone formation rather than pollutants themselves, and their monitoring is thus considered preventive in nature. Although the toxicological values for individual VOCs are well established, defining a threshold for total VOCs is complex due to the variable composition of VOC mixtures another (Zhou et al., 2023).

As no regulatory standards exist for total VOC concentrations in ambient air, guidance values for total volatile organic compounds (TVOC) in indoor environments were taken from the German “TVOC concept” developed by the Ad-hoc Working Group on Indoor Air Guideline Values and the German Environment Agency (Umweltbundesamt), which defines hygienic classes for TVOC with thresholds at 0.3, 1 and 3 mg/m<sup>3</sup> (Umweltbundesamt, 2007). In addition, the U.S. Green Building Council's LEED v4.1 Indoor Air Quality Assessment credit considers a TVOC concentration of 500 µg/m<sup>3</sup> (0.5 mg/m<sup>3</sup>) as an upper acceptable limit for newly constructed buildings before occupancy. Based on these established guideline values, the following classification was adopted in the present study:

- Low: <0.3 mg/m<sup>3</sup>
- Acceptable: [0.3–0.5] mg/m<sup>3</sup>
- Marginal: [0.5–1] mg/m<sup>3</sup>
- High: [1–3] mg/m<sup>3</sup>

Although not a regulatory standard for outdoor air, this classification provides indicative concentration ranges that can be used to structure a qualitative interpretation of total VOC levels. In this study, these TVOC classes – originally developed for indoor air quality assessment – are not interpreted as health-based thresholds for ambient air. They are used solely as convenient categories to discretize the distribution of annual *S. VOC* values when constructing the composite reference value RV, in a context where no established TVOC guidelines exist for outdoor environments.

**Table 3**

Calculation of reference values based on *BelAQI* and VOC categories.

| A             | B     | C     | D      | E    | F      |
|---------------|-------|-------|--------|------|--------|
| <i>BelAQI</i> | A*365 | B*0.3 | B*0.5  | B*1  | B*3    |
| 1             | 365   | 109.5 | 182.5  | 365  | 1095   |
| 2             | 730   | 219   | 365    | 730  | 2190   |
| 3             | 1095  | 328.5 | 547.5  | 1095 | 3285   |
| 4             | 1460  | 438   | 730    | 1460 | 4380   |
| 5             | 1825  | 547.5 | 912.5  | 1825 | 5475   |
| 6             | 2190  | 657   | 1095   | 2190 | 6570   |
| 7             | 2555  | 766.5 | 1277.5 | 2555 | 7665   |
| 8             | 2920  | 876   | 1460   | 2920 | 8760   |
| 9             | 3285  | 985.5 | 1642.5 | 3285 | 9855   |
| 10            | 3650  | 1095  | 1825   | 3650 | 10,950 |

Each *S.BelAQI* value was then multiplied by the upper limit of the corresponding VOC category to compute a combined *RV* for each scenario (Eq. (4)). The results of this step are shown in Table 3, Columns C to F, corresponding to each VOC class (from “Low” to “High”).

#### *Step 3: Establishing the index-Reference Value Relationship*

A linear regression model was then established to define the relationship between the index values and the corresponding reference values computed in Step 2, based on data from selected monitoring sites (Supplementary Material S1). This regression is used as an operational calibration tool to obtain a monotonic mapping between the composite reference value and the index, rather than as a mechanistic representation of atmospheric processes. A linear form was chosen for transparency and ease of interpretation, in line with the intended use of the index by non-expert practitioners.

#### *Step 4: Preliminary class grouping*

Six preliminary AQOI classes were defined based on aggregated ranges of reference values. BelAQI categories 1 and 2 were grouped to form the first class, with other classes following accordingly.

#### *Step 5: Calculation of the index thresholds*

The index values corresponding to the calculated reference thresholds were determined using the regression model established in Step 3.

#### *Step 6: Adjustment by percentile optimization*

To refine class boundaries, the reference thresholds were adjusted using the percentile that maximized the classification agreement between the index classes and reference value classes. An additional constraint was imposed to ensure that all classes collectively cover the full range of the index values (0 to 5). The detailed percentile calculation is described in the Results section.

#### *Step 7: Qualitative interpretation*

Each index class was assigned a qualitative label to enhance communication with the public. The final scale includes the following categories, ranked from best to worst air quality (“Excellent”, “Very good”, “Good”, “Moderate”, “Insufficient” and “Very poor”).

## **2.5. Validation**

Validation of the index was performed using data from the year 2022, by comparing the reference value (*RV*) with the corresponding index scores at the following monitoring stations: Engis, Herstal, Lodelinsart, Mons, Dourbes, Havinnes, Sainte-Ode, and Vielsalm (Supplementary Material S1). In addition, the AQOI is currently defined on an annual time scale, which makes it suited to characterize long-term structural pressures rather than seasonal or episodic exposure patterns (e.g. winter heating peaks, summer ozone episodes or short-lived stagnation events).

Between 2017 and 2022, only wind direction frequencies and IED (industries subject to the directive) emissions data exhibited variations; all other input parameters remained unchanged. This implementation therefore treats most input parameters as structurally stable over the 2017–2022 period, focusing on slowly evolving features such as buildings, road layout and major industrial sources, and does not explicitly capture short-term changes in traffic intensity or heating practices. The validation step therefore relies on 8 monitoring sites located in Wallonia for a single year (2022). Consequently, the reported agreement between the index and the *RV* should be interpreted as an

internal consistency check within this specific regional setting, rather than as a full validation over a wide range of urban configurations.

## 2.6. Case study

The city of Seraing (50°36'N, 5°32'E) is located in eastern Wallonia, within the province of Liège (Fig. 2). Situated along the Meuse River, the city spans an area of 35.2 km<sup>2</sup> and lies at an elevation of 61 m a.s.l., with a population density of approximately 1837 inhabitants/km<sup>2</sup> (IWEPS, 2025). The local climate is classified as warm oceanic, with no distinct dry season. Average summer and winter temperatures are around 20 °C and 8 °C, respectively, and annual precipitation of about 600–700 mm (Institut Royal Météorologique, 2026).

The Seraing site has undergone substantial transformation over the last two decades (Fig. 2). Urban redevelopment projects have included the construction of a new administrative center, the creation of pedestrian-friendly public spaces, the renovation of buildings, and the reconfiguration of the road network to optimize traffic flow and promote alternative modes of transportation (e.g., expanded roadways, directional changes, and dedicated bus lanes). All of these initiatives are consistent with the objectives of SDG 11. The urban development of the city center between 1994 and 2021 was selected as the case study for the index application (Fig. 2).

All data were retrieved from the *Projet Informatique de Cartographie Continue* (PICC) database (Service public de Wallonie, 2025) in shapefile format. Industrial Emissions Directive (IED) data were obtained from the EEA Datahub (European Environment Agency, 2025a, 2025b). Data visualization and extraction were performed using QGIS software (QGIS Development Team, 2021), and certain parameters (e.g., distances) were measured directly using QGIS tools. Wind direction data were sourced from the Bierseix Airport meteorological station, located 5.6 km from the Seraing site, using the *worldmet* package in R.

## 3. Results and discussions

### 3.1. Development of reference value and the index class thresholds

Like any environmental index, the proposed index must communicate information in a format that is easily interpretable by the general public (European Environment Agency, 1999). To achieve this, index values are categorized according to a qualitative scale.

The construction of this scale relies on a reference value (RV), which itself is derived from a qualitative classification based on measured field data. This ensures that index values are grounded in real-world concentrations of air pollutants.

#### 3.1.1. Threshold determination using percentile analysis

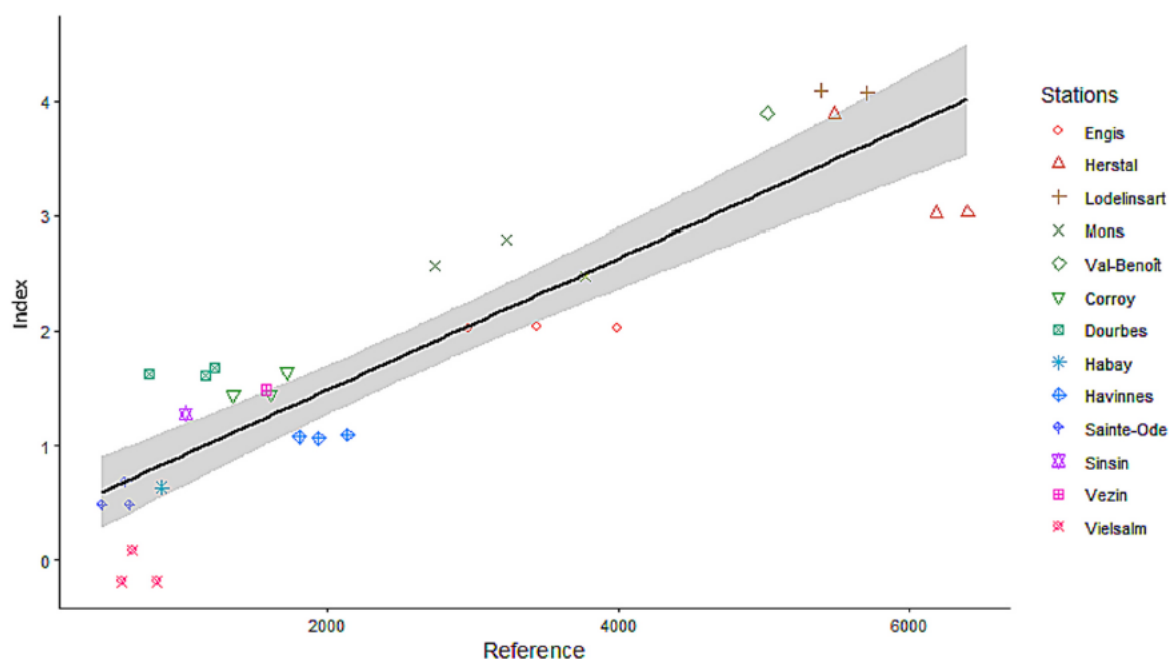
Several comparison tests were performed using a linear regression model (Fig. 3) to evaluate the match between index-based site classifications and those based on the RV. The fitted relationship showed a coefficient of determination ( $R^2 = 0.80$ ), providing a basic indication of goodness-of-fit for this

operational use of the regression. While more complex, potentially non-linear models could in principle capture additional structure in the relationship between the proxy-based index and the measurement-based reference value, the linear formulation was retained as a pragmatic choice given the limited size and diversity of the calibration dataset and the primary objective of deriving stable, interpretable class thresholds. The best correspondence was obtained using the 80th percentile of the five *RV* vectors presented in Table 3 (summarized in Supplementary Material S3 - Table S3).

As shown in Tables S4 and S5 (Supplementary Material S3), percentiles 10, 80, 85, and 100 provided the highest classification match:



**Fig. 2.** a. Localization of the city of Seraing in Belgium; b–d. Evolution of Seraing visualized through orthophotos used in AQOI cell calculations: b. 1994–2000 (Service public de Wallonie (SPW), 2015), c. 2009–2010 (Service public de Wallonie (SPW), 2019), d. 2021 (Service public de Wallonie, 2021). Data visualized in QGIS (QGIS Development Team, 2021). IC cell size: 300 m (external grid) (outer grid); sIC cell size: 100 m (internal grid).



**Fig. 3.** Linear regression between the reference values and index values calculated for the years 2017, 2018, and 2019 ( $Index = 5.7 \cdot 10^{-4} \times RV + 0.34$ ;  $R^2 = 0.80$ ).

- P10, P80, and P85: Each yielded a match in 22 out of 30 datasets (73% agreement).
- P100: Achieved the highest concordance, with 26 out of 30 datasets correctly classified.

However, only the 80th percentile ensured a balanced distribution across the six index classes, covering index values from 0 to  $>4.55$ . In contrast:

- P10 distributed the six classes over index values from 0 to  $>0.97$ , resulting in highly compressed thresholds.
- P85 extended the range up to  $>5.79$ , but the final class was never triggered.
- P100 spanned up to  $>6.59$ , but the first two classes clustered nearly all index values below  $<4.10$ , leading to skewed class distribution.

The 80th percentile was therefore selected as a pragmatic compromise, combining relatively high class-level agreement with a balanced use of the full index range. This choice is empirical and tailored to the current calibration dataset, and may be revisited as additional sites and years become available. Out of the 30 datasets used for class construction (Supplementary Material S1 - Table S2), 22 showed consistent qualitative class assignments between the calculated reference value and the proposed index, representing an overall concordance of 73% (Supplementary Material S3 – Table S6).

### 3.1.2. Explanation of classification differences

The observed differences between index and RV classifications can be attributed to the following factors:

- **Data dependency:** Each index parameter is derived from encoded proxy indicators, whereas the reference value is based on actual field measurements. Missing data due to instrumentation errors, data acquisition failures, or other operational issues can affect the completeness and accuracy of the reference value.
- **Model generalization:** Unlike air quality monitoring stations that record pollutant concentrations directly, the index relies on generalized representations of environmental conditions. For example, two industrial buildings may be assigned identical index values, regardless of differences in their activity types or emission intensities.

As a consequence, the index should not be interpreted as providing a causal mapping between each proxy variable (e.g. road class, building geometry) and pollutant concentrations at the monitoring sites. Instead, the measurement-based reference value is used at an aggregate level to anchor the proxy-based index on a qualitative scale that is consistent with observed air-quality conditions. The calibration framework therefore ensures overall coherence between proxy information and measured air quality, without implying one-to-one causal validation of each individual variable. For the road parameter in particular, this means that the index does not explicitly resolve differences in traffic volume, fleet composition or congestion between roads of the same class. Instead, it relies on the typical traffic and speed characteristics associated with each road category as a pragmatic surrogate for emission intensity, in line with the low-cost and proxy-based positioning of the AQOI.

## 3.2. Validation

Validation was conducted at eight sites in 2022 where both VOC concentrations and the pollutants constituting the BelAQI index were measured. Table 4 presents the *S.BelAQI*, *S.VOC*, *RV*, and proposed index values along with their respective classifications. Only the Lodelinsart and Mons sites showed discrepancies in categorization, resulting in an overall classification agreement of 75%, which is consistent with the value obtained during the definition of classification thresholds (73%). This agreement is quantified in a class-by-class contingency table (Table 4), which highlights where the index and the RV diverge and for which site types (industrial, urban, rural). Given the categorical nature of the index and the small number of validation sites, class-level agreement was prioritized over continuous error metrics such as RMSE or MAE, which are less directly interpretable for the intended planning applications. This level of agreement is encouraging but must be considered in light of the limited number of stations and the spatial focus on a single region; additional validation on larger and more diverse datasets will be required to evaluate the general applicability of the index to other urban settings.

**Table 4**

Calculation and classification of the 8 sites for the validation.

| Stations |             | S.BelAQI | S.CO <sub>V</sub> | RV   | Index | Classification |           |
|----------|-------------|----------|-------------------|------|-------|----------------|-----------|
|          |             |          |                   |      |       | RV             | Index     |
| I        | Engis       | 1779     | 1771              | 3151 | 1.95  | Good           | Good      |
| U        | Herstal     | 1861     | 2877              | 5354 | 3.49  | Moderate       | Moderate  |
| U        | Lodelinsart | 1836     | 2043              | 3751 | 3.58  | Good           | Moderate  |
| U        | Mons        | 1848     | 1351              | 2497 | 2.83  | Very good      | Moderate  |
| R        | Dourbes     | 1620     | 589               | 955  | 1.63  | Very good      | Very good |
| R        | Havannes    | 1658     | 950               | 1576 | 0.96  | Very good      | Very good |
| R        | Sainte-Ode  | 1578     | 394               | 621  | 0.64  | Excellent      | Excellent |
| R        | Vielsalm    | 1409     | 465               | 656  | 0.01  | Excellent      | Excellent |

A closer inspection of the normalized sub-indices at the two misclassified sites, building on the detailed parameter-level analysis reported in the thesis of anonymous (2025), shows that the discrepancies are mainly driven by specific AQOI components rather than by a global bias. At Lodelinsart, the roads (P2) and buildings (P3) sub-indices are systematically higher than at other urban sites, reflecting dense road canyons, large enclosed asphalted areas and the presence of industrial-type buildings, while the local vegetation layout (P4) can, under certain wind directions, hinder dispersion instead of acting as an effective sink. As a result, the AQOI tends to overestimate structural pollution pressure compared with the reference value, which is closer to the upper “Good” class. For Mons, the 2022 discrepancy arises mainly from the combination of P2 and P4: the prevailing circulation increases the frequency of canyon-like configurations and enhances the screening effect of tall vegetation around the monitoring site, leading to a slightly higher AQOI class (“Moderate”) than the reference (“Very good/Good”). These two sites thus highlight that future refinements should focus on the coding of roads, buildings and vegetation in dense mixed-use urban fabrics, for example by better distinguishing large industrial blocks and by representing vegetation structures that may either enhance or impede dispersion depending on wind conditions.

A closer examination of the two divergent sites reveals that the *RV* values for Lodelinsart in 2018 and 2019 were both classified as *Moderate*, while in 2022, the classification improved to *Good*, indicating enhanced air quality. A similar trend is observed for Mons, where the *RV* classification was *Good* from 2017 to 2019, improving to *Very Good* in 2022.

Regarding the 2022 index values, the Lodelinsart site was classified as *Moderate*, reflecting an improvement compared to *Insufficient* ratings in 2018 and 2019. However, this positive trend was not observed for Mons, which remained in the *Moderate* class in 2022, similar to 2018.

The discrepancies observed may be attributed to several factors:

- Data quality issues, such as gaps or inconsistencies in measurement records;
- Manual data encoding, which can introduce errors through subjective interpretation or transcription;
- The simplified nature of the proposed index, which models a complex environmental reality using proxy variables and assumptions;

- The composition of the reference value (*RV*), which relies exclusively on measured concentrations used to compute the BelAQI and VOCs values, without incorporating other potential influencing factors

### 3.3. Case study: Seraing

Among the five index parameters (P1–P5), only Topography (P1 – Table 5) remained unchanged over the period 2000–2021. The Meuse Valley's morphology, hydrographic network, and riverbanks have not undergone any significant alteration during this time.

**Table 5**

Comparison of non-normalized and normalized sub-index (SI) values for each index parameter, illustrating their respective contributions to the overall index.

| Years |                 | P <sub>1</sub> | P <sub>2</sub> | P <sub>3</sub> | P <sub>4</sub> | P <sub>5</sub> | Index | Classes   |
|-------|-----------------|----------------|----------------|----------------|----------------|----------------|-------|-----------|
| 2000  | SI              | 1.20           | 0.25           | 514            | 0.14           | 0.11           | 4.8   | Very poor |
|       | SI <sub>n</sub> | 0.87           | 0.96           | 1.00           | 0.97           | 1.00           |       |           |
| 2010  | SI              | 1.20           | 0.25           | 502            | 0.08           | 0.12           | 4.7   | Very poor |
|       | SI <sub>n</sub> | 0.87           | 0.95           | 1.00           | 0.90           | 1.00           |       |           |
| 2021  | SI              | 1.20           | 0.22           | 19.49          | 0.04           | 0.13           | 4.6   | Very poor |
|       | SI <sub>n</sub> | 0.87           | 0.85           | 1.00           | 0.85           | 1.00           |       |           |

In contrast, Road Infrastructure (P2 – Table 5) experienced notable modifications. The implementation of pedestrian zones and areas reserved for public transportation contributed to a reduction in traffic-related impacts. Additionally, the demolition of certain buildings led to a widening of the main street, thereby mitigating the canyon effect that previously intensified pollutant concentrations.

Over the past two decades, Building Architecture (P3 – Table 5) has undergone substantial transformation. The raw value associated with this parameter declined by a factor of 26—from 514 in 2000 to 19.49 in 2021. However, the normalized value of P3 remained unchanged. This apparent contradiction is due to the upper threshold value used during the normalization process, which is capped at 15. All raw values exceeding this threshold are limited to a normalized value of 1. A broader dataset incorporating a wider range of urban morphologies could justify an upward revision of this limit. This illustrates a saturation effect induced by the current percentile-based normalization, which limits the sensitivity of the normalized building sub-index to substantial structural improvements once raw values exceed the upper threshold.

Vegetation (P4 – Table 5) also evolved over the 20-year period. Urban greening has been characterized by the introduction of moderately sized, widely spaced trees and low ground vegetation. This change is reflected in a marked decline in the normalized sub-index for vegetation.

The External Sources (P5 – Table 5) parameter has shifted in two key ways. First, the number of facilities listed under the European Pollutant Release and Transfer Register (E-PRTR) and regulated by the IED directive fell from 13 in 2007 to 6 in 2021. Moreover, the number of pollutants reported

decreased from 50 in 2007 to 15 in 2021. As a result, the contribution of IED industries to the non-normalized sub-index declined from 0.0123 in 2007 to 0.0014 in 2021.

Conversely, the road infrastructure surrounding the study grid evolved in a way that increased the influence of external sources. The expansion of Rue Cockerill and the construction of Rue Potier in the western sector of the grid contributed to an increase in the non-normalized sub-index value for external roads, which rose from 0.1035 to 0.1204.

Taken together, these changes suggest that urban development over the past two decades has reduced the structural pressure exerted by roads (P2), buildings (P3) and vegetation (P4) on air quality, as reflected in the evolution of the corresponding raw sub-index values. The most pronounced decrease in structural pressure is observed for the Buildings parameter.

Despite these favorable developments, the overall index value decreased only slightly—by 4%, from 4.8 to 4.6. This reduction, although notable, was not sufficient to shift the index into a lower qualitative class—remaining in the “Very Poor” category ( $\text{Index} > 4.55$ ), just above the threshold for “Insufficient” ( $3.64 < \text{Index} \leq 4.55$ ). In this saturated part of the index scale, such a small change should therefore be interpreted as a limited decrease in structural pressure rather than as evidence of a marked improvement in ambient pollutant levels.

This limited impact is primarily due to the normalization process, which relies on the 2.5th and 97.5th percentiles of the distribution of raw sub-index values from a reference set of sites located near air quality monitoring stations. The relatively small and non-representative sample—lacking sites equivalent to Seraing in terms of industrial and urban density—likely resulted in underestimation of the upper limit values for normalization. Expanding the dataset to include more urbanized environments could improve percentile calibration and, consequently, the index's sensitivity to environmental improvements.

This single case study should therefore be interpreted as an illustrative application of the index to a long-term urban redevelopment project, highlighting both its ability to track structural changes and its current limitations in terms of sensitivity under saturated normalization settings, rather than as a comprehensive validation exercise. No systematic comparison with long-term trends in measured or modelled pollutant concentrations in or around Seraing is provided here, so the observed AQOI evolution should not be taken as a formal demonstration of air-quality improvement.

Although the AQOI is applied here to a single 300 m × 300 m grid cell in Seraing, the case study is intended primarily as an illustrative application of the index to a long-term urban redevelopment process rather than as a city-wide validation exercise. Extending the computation of the AQOI to multiple grid cells or to a larger spatial coverage (for instance the full redevelopment area of Seraing) would allow the construction of a city-scale map and a more systematic assessment of spatial patterns in structural pressure on air quality. Such a city-wide implementation lies beyond the scope of the present work, mainly due to time and data constraints, but it represents a logical next step to fully exploit the spatial planning potential of the index.

### 3.4. Limitations related to atmospheric dynamics

By construction, the proxy-based index provides a simplified representation of a fundamentally dynamic atmospheric system. Meteorological and chemical processes such as inversions, seasonal changes in wind regimes or secondary aerosol formation are not explicitly simulated in the algorithm. They are instead reflected indirectly through the reference value, which is derived from time series of measured concentrations and from an official multi-pollutant air-quality index, and therefore already embeds the response of pollutants to meteorological variability.

In addition, the AQOI is calibrated against a composite reference value that combines the BelAQI index, aggregating  $PM_{2.5}$ ,  $PM_{10}$ ,  $NO_2$  and  $O_3$ , with annual total VOC concentrations, so that it primarily reflects an integrated particulate–oxidant–VOC pollution burden rather than pollutant-specific risks. As a consequence, the index is not intended to resolve detailed spatial gradients for individual species (e.g. near-road  $NO_2$  peaks or ozone titration patterns), but to summarize the long-term structural pressure of urban form and major emission sources on a set of key pollutants represented jointly in the reference value. Pollutant-specific assessments, for regulatory or health impact purposes, should therefore continue to rely on dedicated concentration fields from monitoring networks and dispersion or chemical transport models, with the AQOI used as a complementary planning tool to compare alternative urban configurations at neighborhood scale.

In the context of a planning support tool, the development focused on a robust description of long-term morphological and functional characteristics rather than on day-to-day dynamics. Because the AQOI is constructed from annual proxy data and an annual, measurement-based reference value, it does not resolve intra-annual variability and should therefore not be used to analyze seasonal or episode-scale exposure differences. This implies that the index is not intended to replace chemical transport models or operational forecasting systems, but to complement them in situations where such tools are not available or cannot be mobilized at the early stages of urban projects. Future extensions could nevertheless integrate additional meteorological indicators – such as wind speed, atmospheric stability or seasonal descriptors – where consistent data are available at the spatial scale of the grid.

## 4. Conclusions

Aligned with the objectives of Sustainable Development Goal 11 (SDG 11), which promotes inclusive, safe, resilient, and sustainable cities, air quality must be considered a central component in land-use planning and decision-making. The proposed air quality index was developed precisely for this purpose.

Existing tools for assessing the impact of the built environment on air quality often require specialized expertise or rely on post-implementation concentration measurements. In contrast, the AQOI is designed as a user-friendly index that can be implemented with standard GIS tools and a spreadsheet-based, open-source interface. Because each sub-index relies on simple spatial queries and algebraic operations, the total computational cost increases approximately linearly with the number of spatial elements per grid cell, allowing the method to be applied efficiently to large urban areas in standard GIS environments. It is built from proxy variables that are publicly accessible and straightforward to encode, and aggregates five parameters – topography, roads, buildings, vegetation

and external sources – into a single score classified into six qualitative categories from 0 (excellent) to 5 (very poor), mirroring commonly used public air-quality communication scales.

The relationship with measured pollutants is established through a composite reference value that combines the annual sum of the health-based BelAQI index with total VOC concentrations. This reference was used to calibrate class thresholds and should be regarded as a context-dependent calibration device rather than as a new health metric or a mechanistic representation of VOC–NO<sub>2</sub>–O<sub>3</sub> chemistry. Within the current calibration dataset (13 stations), an equal-weight formulation of the five parameters provided the most stable agreement with the reference value and the most transparent interpretation for non-expert end-users, with little added benefit from more complex weighting schemes.

Validation performed at eight sites for 2022 yielded a 75% agreement between AQOI classes and the reference value, supporting both the conceptual design of the index and the value assignments used for its proxy variables. Discrepancies observed at two urban sites point to the need for further refinement of the road, building and vegetation encodings in dense mixed-use contexts, and underline the importance of expanding calibration and validation to larger and more diverse regions and urban typologies.

The longitudinal application of the index to a 300 m × 300 m grid cell in Seraing over two decades demonstrates its ability to capture structural changes in the built environment and their implications for air quality. Marked improvements in the building and vegetation sub-indices reflect a shift towards a less obstructive and potentially more breathable urban form, even though the overall index value decreased only slightly and remained in the same qualitative class. This behavior reveals saturation effects linked to the current percentile-based normalization and highlights the need for periodic recalibration of normalization bounds and class thresholds as the calibration dataset grows.

In its current form, the AQOI is best interpreted as a health-informed structural pressure indicator for early-stage screening and comparative assessment between neighborhoods or development scenarios, particularly where detailed modelling resources are not available. It is not intended to replace regulatory monitoring, exposure or health-risk assessments, nor process-based dispersion or chemistry-transport models, but to complement them by providing structurally grounded information at the design stage.

Future developments could focus on several directions: enlarging the calibration and validation database to encompass more diverse climates and urban forms; testing alternative formulations of the reference value, exploring data-driven weighting schemes and hybrid approaches that combine spatial proxies with supervised learning models; refining the treatment of external sources and traffic when detailed emission and traffic data are available; where multi-year traffic counts, fleet composition data and emission factors are available, the road parameter could also be refined by replacing or complementing the current class-based proxies with more explicit traffic and emission indicators, while maintaining the class-based formulation as a low-cost option for data-scarce municipalities; and, ultimately, benchmarking the index against dispersion models or other independent air-quality datasets in multiple urban settings. Future validation efforts on larger datasets could also incorporate additional performance metrics, such as kappa statistics and, where appropriate, continuous error indicators (e.g. RMSE or MAE), in order to better quantify the robustness and uncertainty of the proposed index. More detailed, seasonal variants of the AQOI could in principle be developed, but

would require season-specific reference values and consistently available seasonal proxy datasets; this is identified as a promising avenue for future work once suitable data become available.

The current implementation also relies on fixed buffer distances for external sources (900 m for external roads, 7 km for airports and 15 km for IED-regulated industries and ports), chosen on the basis of typical dispersion scales; their influence on AQOI classes, particularly in highly industrialized or airport-proximate settings, should be assessed systematically. At this stage, the AQOI is implemented as a deterministic index and results are reported as point estimates; a full uncertainty propagation analysis (for example using Monte Carlo simulations based on plausible input ranges) has not yet been performed and is identified as a priority for future developments of the method. As urban infrastructure evolves, periodic updates of proxy encodings and normalization settings will be required, but the underlying concept – a practical, accessible and health-informed index to support urban planning decisions on air quality – will remain central.

Nevertheless, its conceptual basis remains valid and future adaptations will continue to support its central aim as a practical, accessible and health-informed indicator for urban planning.

## Credit authorship contribution statement

**C. Falzone:** Writing – original draft, Visualization, Validation, Software, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **A.C. Romain:** Writing – review & editing, Supervision, Funding acquisition.

## Informed consent statement

Not applicable.

## Institutional review board statement

Not applicable.

## Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work the author(s) used ChatGPT in the writing process (grammar correction and improvement of the text for readability). After using this tool, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the published article.

## Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Anne-Claude ROMAIN reports financial support was provided by

Wallonia Region. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.scitotenv.2026.182044>.

## Data availability

The PICC dataset is available on <https://geoportail.wallonie.be>; The E-PRTR dataset is available on <https://industry.eea.europa.eu/download>; The air quality data are available on <https://www.wallonair.be/fr>. The internal values related to the different calculations used to obtain the AQOI of the different sites are not published.

## References

- Air Quality Models [WWW Document]*, n.d. . irCELine. URL <https://www.irceline.be/en/documentation/models> (accessed 7.25.25).
- AirParif, AIRLAB, 2023. Résultats du challenge AIRLAB Microcapteurs 2023 [WWW Document]*. Airlab. URL. <https://airlab.solutions/actualites/resultats-du-challenge-airlab-microcapteurs-2023-203> (accessed 7.25.25).
- Alwabli, A., 2024. Federated learning for privacy-preserving air quality forecasting using IoT sensors. *Eng. Technol. Appl. Sci. Res.* 14, 16069–16076. <https://doi.org/10.48084/etasr.7820>.
- BelaAQI Index [WWW Document]*, 2022. irCELine. URL. <https://www.irceline.be/en/air-quality/measurements/belaqi-air-quality-index/information> (accessed 7.28.25).
- Bigazzi, A.Y., Rouleau, M., 2017. Can traffic management strategies improve urban air quality? A review of the evidence. *J. Transp. Health* 7, 111–124. <https://doi.org/10.1016/j.jth.2017.08.001>.
- Briggs, D.J., Collins, S., Elliott, P., Fischer, P., Kingham, S., Lebre, E., Pryl, K., Van Reeuwijk, H., Smallbone, K., Van Der Veen, A., 1997. Mapping urban air pollution using GIS: a regression-based approach. *Int. J. Geogr. Inf. Sci.* 11, 699–718. <https://doi.org/10.1080/136588197242158>.
- Carslaw, D., 2025. worldmet: Import Surface Meteorological Data from NOAA Integrated Surface Database (ISD).
- Chatterjee, K., Kumar, S.S., Kumar, R.P., Bandyopadhyay, A., Swain, S., Mallik, S., Al-Rasheed, A., Abbas, M., Soufiene, B.O., 2024. Future air quality prediction using long short-term memory based on hyper heuristic multi-chain model. *IEEE Access* 12, 123678–123693. <https://doi.org/10.1109/ACCESS.2024.3441109>.
- Cities - United Nations Sustainable Development Action 2015*, 2015. U. N. Sustain. Dev. URL. <https://www.un.org/sustainabledevelopment/cities/> (accessed 7.25.25).
- Drewil, G.I., Al-Bahadili, R.J., 2022. Air pollution prediction using LSTM deep learning and metaheuristics algorithms. *Meas. Sens.* 24, 100546. <https://doi.org/10.1016/j.measen.2022.100546>.

- European Environment Agency, 1999. *Environmental Indicators: Typology and Overview* (Technique No. 25/1999). Copenhagen, Denmark.
- European Environment Agency, 2025a. *European Industrial Emissions Portal* [WWW Document]. Eur. Ind. Emiss. Portal (accessed 7.28.25a). <https://industry.eea.europa.eu>.
- European Environment Agency, 2025b. *European Industrial Emissions Portal | Download* [WWW Document]. Eur. Ind. Emiss. Portal (accessed 7.29.25b). <https://industry.eea.europa.eu/download>.
- European Union, 2007. Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE). OJ L 108, 1–14. <http://data.europa.eu/eli/dir/2008/50/oj/eng>.
- European Union, 2024. Directive (EU) 2024/2881 of the European Parliament and of the Council of 23 October 2024 on ambient air quality and cleaner air for Europe. OJ L 1–70. <https://eur-lex.europa.eu/eli/dir/2024/2881/oj/eng>.
- Falzone, C., 2025. Développement d'un indice pour le diagnostic de la qualité de l'air en zone urbanisée pour des projets à l'échelle du quartier. Université de Liège, Arlon Campus. <https://hdl.handle.net/2268/331171>.
- Falzone, C., Romain, A.-C., 2022. Establishing an air quality index based on proxy data for urban planning part 1: methodological developments and preliminary tests. *Atmosphere* 13, 1470. <https://doi.org/10.3390/atmos13091470>.
- Faraji, M., Nadi, S., Ghaffarpasand, O., Homayoni, S., Downey, K., 2022. An integrated 3D CNN-GRU deep learning method for short-term prediction of PM<sub>2.5</sub> concentration in urban environment. *Sci. Total Environ.* 834, 155324. <https://doi.org/10.1016/j.scitotenv.2022.155324>.
- Fowler, D., Brimblecombe, P., Burrows, J., Heal, M.R., Grennfelt, P., Stevenson, D.S., Jowett, A., Nemitz, E., Coyle, M., Liu, X., Chang, Y., Fuller, G.W., Sutton, M.A., Klimont, Z., Unsworth, M.H., Vieno, M., 2020. A chronology of global air quality. *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.* 378, 20190314. <https://doi.org/10.1098/rsta.2019.0314>.
- IARC Working Group on the Evaluation of Carcinogenic Risks to Humans, 2016. *Outdoor air pollution*. In: *IARC Monographs on the Evaluation of Carcinogenic Risks to Humans*. International Agency for Research on Cancer, World Health Organization, Lyon, France.
- Institut Royal Météorologique, 2026. IRM - Climat dans votre commune [WWW Document]. KMI. <https://www.meteo.be/fr/climat/climat-de-la-belgique/climat-dans-votre-commune> (accessed 7.9.26).
- ISSeP, n.d. *Micro-capteurs* [WWW Document]. Wallonair. URL <https://www.wallonair.be/fr/en-savoir-plus/projets/176-projet-microcapteurs.html> (accessed 7.25.25).
- IWEPS, 2025. WalStat - Détail de l'entité SERAING (Commune) [WWW Document]. WalStat. [https://walstat.iweps.be/walstat-fiche-entite.php?entite\\_id=62096](https://walstat.iweps.be/walstat-fiche-entite.php?entite_id=62096) (accessed 7.9.26).
- Mahmoud, M., Ramadan, M., Naher, S., Pullen, K., Olabi, A.-G., 2021. The impacts of different heating systems on the environment: a review. *Sci. Total Environ.* 766, 142625. <https://doi.org/10.1016/j.scitotenv.2020.142625>.
- Micro-capteurs: quelle utilité pour connaître l'exposition à la pollution de l'air ? [WWW Document], 2022. Anses - Agence Natl. Sécurité Sanit. L'alimentation L'environnement Trav. URL. <https://www.anses.fr/fr/content/micro-capteurs-que-lle-utilit%C3%A9-pour-conna%C3%Aetre-l%E2%80%99exposition-%C3%A0-la-pollution-de-l%E2%80%99air> (accessed 7.25.25).

- Neshovski, R., n.d. *Sustainable Development Goals* [WWW Document]. U. N. Sustain. Dev. URL <https://www.un.org/sustainabledevelopment/> (accessed 7.25.25).
- Pelgrims, I., Pauwels, A., De Clercq, E., 2024. *Overview of air Quality Measures Aiming to Reduce Emissions From Road Traffic and Mitigate Health Impact in Belgium* [WWW Document]. sciensano.be. URL. <https://www.sciensano.be/fr/biblio/overview-air-quality-measures-aiming-reduce-emissions-road-traffic-and-mitigate-health-impact> (accessed 7.25.25).
- QGIS Development Team, 2021. *QGIS Geographic Information System*.
- Réhabiliter les friches : opportunités et méthodologie [WWW Document], 2022. Cerema. URL. <https://www.cerema.fr/fr/actualites/rehabiliter-friches-opportunites-met-hodologie> (accessed 7.25.25).
- Service public de Wallonie, 2025. *Projet Informatique de Cartographie Continue (PICC)* [WWW Document]. Geoportail Wallonie. URL. <https://geodata.wallonie.be/id/b795de68-726c-4bdf-a62a-a42686aa5b6f> (accessed 7.29.25).
- Sustainable Finance Hub [WWW Document], 2025. U. N. Dev. Programme. URL. <https://sdgfinance.undp.org/about-us> (accessed 7.25.25).
- Umweltbundesamt, 2007. *Beurteilung von Innenraumluftkontaminationen mittels Referenz- und Richtwerten: Handreichung der Ad-hoc-Arbeitsgruppe der Innenraumlufthygiene-Kommission des Umweltbundesamtes und der Obersten Landesgesundheitsbehörden*. Bundesgesundheitsbl. Gesundheitsforsch. Gesundheitsschutz 50, 990–1005. <https://doi.org/10.1007/s00103-007-0290-y>.
- WHO Global Air Quality Guidelines: Particulate Matter (PM<sub>2.5</sub> and PM<sub>10</sub>), Ozone, Nitrogen Dioxide, Sulfur Dioxide and Carbon Monoxide, 1st ed. ed, 2021. World Health Organization, Geneva.
- World Health Organization (Ed.), 2023. *WHO Ambient Air Quality Database, 2022 Update: Status Report*, 1st ed. World Health Organization, Geneva.
- Zhou, Xihe, Zhou, Xiang, Wang, C., Zhou, H., 2023. *Environmental and human health impacts of volatile organic compounds: a perspective review*. *Chemosphere* 313, 137489. <https://doi.org/10.1016/j.chemosphere.2022.137489>.