Microsensors CHALLENGE 2025

With the participation of:

Last revision: 04.04.2025

Table of contents

1 Introduction

Remarkable progress in sensor technologies over the last two decades has opened the door to a wide range of potential new applications based on air quality measurements. This has in turn led to the emergence of a very dynamic and arguably volatile market of integrated air quality monitoring solutions based on low-cost sensors. While important steps have been made in recent years for developing norms and performance standards, the large array of commercially available devices – very few of which have been subject to any data quality certification process – can lead to confusion even among expert users when faced with the challenge of selecting an appropriate measurement platform for a particular application.

The AIRLAB Microsensors Challenge aims to address the increasing demand for an independent and objective evaluation of the performance of microsensor-based air quality monitoring solutions. This edition builds on previous efforts carried out internationally, including the pioneering work of the South Coast Air Quality Management District through its Air Quality Sensor Performance Evaluation Center (AQ-SPEC) [1]. For the 2025 edition, AQ-SPEC joins the Challenge as an expert partner, alongside other leading air quality sensor evaluation initiatives such as the Air Quality Sensor Evaluation and Training Facility for West Africa (Afri-SET), the India Sensor Evaluation and Training center (Indi-SET) at CSTEP (Center for Study of Science, Technology and Policy), and NILU (Norwegian Institute for Air Research). Their collective expertise strengthens the Challenge's ability to assess sensor performance across a range of environments and use cases. This collaboration also reinforces AIRLAB's holistic approach, which extends beyond metrological performance to consider aspects such as utility, usability, portability, and cost.

Additionally, together with Ile-de-France Mobilités and SNCF Gares & Connexions, the 2025 Challenge will introduce an evaluation of microsensors in underground railway stations, marking an important step in assessing air quality monitoring solutions for enclosed transit environments. The Challenge is structured as a periodic event in which all participating sensor platforms are evaluated concurrently, providing a snapshot of the state of commercially available microsensors at a given point in time.

The current iteration of the Challenge seeks to consolidate and refine its evaluation process by leveraging the experience accumulated over the first four editions (2018, 2019, 2021, and 2023) and to push the inherent limitations of a challenge format in terms of the generalizability of its results by diversifying evaluation environments. To this end, the 5th edition of the AIRLAB Microsensors Challenge will include two outdoor evaluation sites outside of metropolitan France, expanding to Accra, Ghana, and Bengaluru, India, which both experience significantly different climatic and pollution conditions. The 2025 Challenge will investigate the impact of new parameters on microsensors' performance, including higher pollution levels, diverse emissions profiles, and varying meteorological conditions, particularly higher temperatures and humidity. These new sites have been chosen due to their distinct environmental characteristics and their strong local capacities in air quality monitoring and microsensor-based research.

The scope of this document is to present in detail the Challenge evaluation criteria as well as the associated measurement protocols. It serves as a complement to the Microsensors Challenge Rules (full title: *Microsensors Challenge 2025 – Terms and Conditions, Regulations, and Guidelines*). It is provided for informational purposes and is published alongside the call for participation. The protocol may be subject to modifications during the course of the Challenge due to material constraints (e.g., replacement of reference analyzers with different models), adjustments based on specific pollutant evaluation demands depending on actual candidate submissions, or any other changes deemed necessary by the Challenge Steering Committee to ensure the quality of the evaluation process. All modifications to the protocol will be incorporated into revised versions of this document, and Challenge participants will be duly notified of each revision.

2 Method

In order for their respective air quality sensors to be evaluated within the Challenge, candidates need to submit their applications in accordance with the Microsensors Challenge Terms and Conditions. Once all applications are received, the Jury convenes to validate them and to perform a preliminary selection of candidate solutions based on the completeness and pertinence of the submissions.

In the two months following the preliminary selection, three units of the microsensor platforms per testing environment are received and a set of initial technical tests are performed (e.g., data recovery verifications, mechanical robustness for usage category, etc.), in order to evaluate the technical soundness of the candidate solutions. On the basis of the experience gathered during this phase, the Jury reconvenes to determine which candidates can proceed to the next phase of the Challenge.

Thereafter, the metrological quality of the candidate sensors is tested in accordance with the list of use-case categories specified in their application either in a metrology laboratory or in the field. These tests provide the necessary measurement data and usage information for evaluating the candidate solutions according to the *accuracy* and *usability* criteria. The remaining criteria (i.e. *utility*, *portability* and *cost*) are evaluated mostly based on the verification of manufacturer specifications.

In this section we present the methodology used for performing the evaluation of the candidate solutions. We start by defining the different usage categories considered by the Challenge, we then describe the evaluation environments, and, finally, we present in detail each evaluation criterion and its calculation.

2.1 Categories

In the context of the Challenge, we define a sensor's *category* as its type of use or intended application. The categories in this Challenge are six; some of them are freely adapted from the World Meteorological Organization (WMO) technical report No. 1215 on microsensors [42]. They are grouped based on the targeted application domain: **Outdoor Air** and **Indoor Air** (see Figure 1).

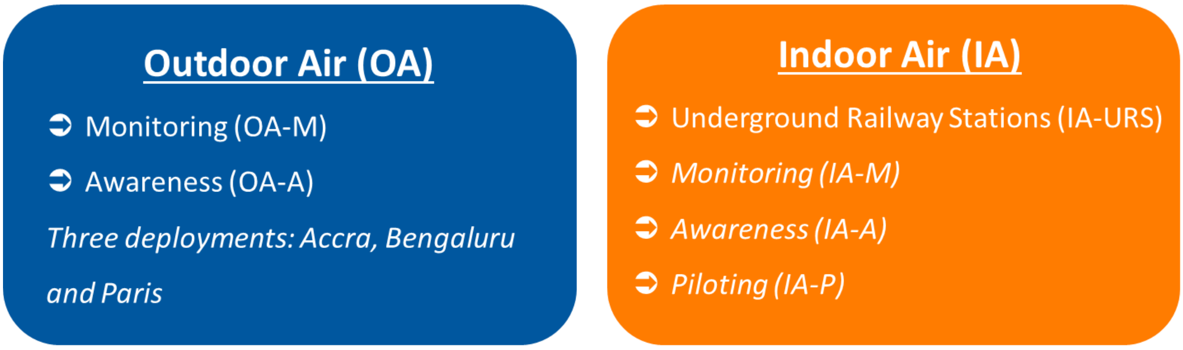

Figure 1: The six Challenge categories grouped by their application domain.

In this context, the six usage categories are defined as follows:

- **Outdoor Air (OA):**
  - **Awareness (OA-A)** – Promote the information and the awareness of the public or users through outdoor air data. The requirements for this type of application are lower on the quality of the data. These sensors aim only at coherence to reference devices and not at equivalence. The panel of pollutants to be monitored may be reduced.
  - **Monitoring (OA-M)** – Target the complementary integration into regulatory networks for monitoring of compliance to national or transnational standards of air quality for a given outdoor location. This implies very high requirements for the quality of data produced and their traceability to reference devices. The main regulated and problematic pollutants are to be measured.

- **Indoor Air (IA):**
  - **Underground Railway Stations (IA-URS)** – Air quality measurement inside underground railway infrastructures, specifically on station platforms. As these data can complement existing air quality monitoring systems in underground railway stations, high-quality data, comparable to those produced by reference instruments, are expected, along with the monitoring of key pollutants in this environment—particulate matter.
  - **Awareness (IA-A)** – Promote the information and the awareness of the public or users through indoor air data. The requirements for this type of application are lower on the quality of the data. These sensors aim only at coherence to reference devices and not at equivalence. The panel of pollutants to be monitored may be reduced.
  - **Monitoring (IA-M)** – The support of the verification of compliance to national air quality standards in childcare establishments under 6 years old (nurseries, day-care centers, etc.), kindergartens and elementary schools. This implies a high quality of data by meeting the accreditation requirements LAB REF30 or the specifications of the INERIS on this subject[1]. The measurement process follows fully prescribed methods and best practices.
  - **Piloting (IA-P)** – Controlling, managing, and regulating indoor air quality for building or installations with the help of a multi-parameter sensor. The requirements for this type of application are lower on the quality of the data. These sensors aim only at coherence to reference devices and not at equivalence, while at the same time being continuously available and easily interoperable with the domotics system, including the managing or user interface.

2.2 Evaluation environments

To enable the different Challenge categories, the 2025 Challenge considers five different types of evaluation settings: three outdoor static environments and two indoor environments. They are listed in Table 1 and described in detail in the remainder of this subsection.

Table 1 : The evaluation environments of the 2025 AIRLAB Microsensors Challenge

	Evaluation Environment	Location of deployment	Categories
Outdoor Air	Temperate climate – Europe	Paris, France	OA-M, OA-A
	Tropical climate – West Africa	Accra, Ghana	OA-M, OA-A
	Tropical climate – South Asia	Bengaluru, India	OA-M, OA-A
Indoor Air	Underground Railway Stations	Paris, France	IA-URS
	Non-Specific Spaces	Paris, France	IA-M, IA-A, IA-P

Figure 2: Outdoor evaluation site – BP-Est monitoring station exterior (left) and interior (right) views.

[1] Evaluation of the conformity of kits for the realization of indicative measurements of formaldehyde, benzene and carbon dioxide in the indoor air of establishments receiving children – INERIS, 2017

The **Outdoor Temperate climate – Europe evaluation environment** is represented by the regulatory monitoring station Boulevard Périphérique Est (BP-Est) which is located near the Porte Dorée city gate (see Figure 2). The elevation at this site is 48 m, with WGS 84 coordinates: 48° 50’ 19”N (latitude) and 2° 24’ 46”E (longitude). This monitoring station has a particularity in that, depending on the direction of the wind, it will behave as either a near-traffic station (East wind) or a background station (West wind). It thus has a large dynamic range which makes it an excellent site for testing new measurement equipment. Challenge categories that are evaluated at the BP-Est station are **OA-A** and **OA-M**.

The following analyzers are used at BP-Est as reference measurements for the purpose of the Microsensors Challenge:

- **Particulate matter** is measured with two Met One BAM-1020[2] beta attenuation monitors (BAM) for **PM_{10}** and **$PM_{2.5}$** (Reference standard for ambient air: NF EN 16450: 2017) and a PALAS Fidas 200 analyzer[3], an optical spectrometer, for particle counts (**PN**). **Hourly averages** are used for both the BAMs and the Fidas, although finer time resolutions are possible for the latter.

- **Nitrogen oxides** are measured by a Thermo Scientific 42i[4] chemiluminescence analyzer, which provides measurements for **NO** and **NO_2** (Reference standard for ambient air: NF EN 14211: 2012). In the context of the Challenge **hourly averages** are used, however the data from this analyzer is also available at finer resolutions (e.g., 15 min averages, 10 s scans).

- **Ozone** is measured with a Teledyne API T400[5] UV fluorescence analyzer (Reference standard is ambient air: NF EN 14625: 2013). **Hourly averages** are used, but finer resolution data is also available (e.g., 15 min averages, 10 s scans).

- **Aerosol Black Carbon** is measured with the Magee Scientific AE33 aethalometer[6]. **Hourly averages** are used, but time resolutions as rapid as 1s are possible.

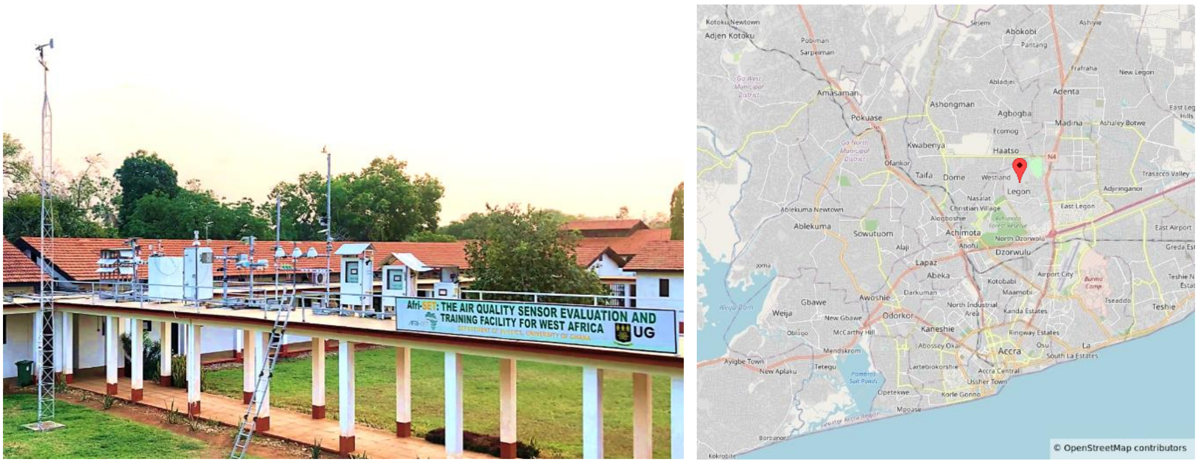

Figure 3: Outdoor evaluation site – Afri-SET station (left) and location of the University of Ghana campus (right).

The **Outdoor Tropical climate – West Africa evaluation environment** is represented by the Afri-SET urban background monitoring station located within the campus of the University of Ghana in Legon, Accra, Ghana, at 5.65136° N, 0.18566° W at an elevation of 108 m (see Figure 3). Accra is the capital and largest city in Ghana. It has a tropical climate with a rainy season (April–October) and a dry season (November–March). Humidity is high year-round but eased by sea breezes. The Harmattan (December–February) brings dry, dusty air, while heavy rains peak in April–July and September–October. Usage categories evaluated at this site are **OA-A** and **OA-M**.

[2] https://metone.com/air-quality-particulate-measurement/regulatory/bam-1020/
[3] https://www.palas.de/en/product/fidas200
[4] https://www.thermofisher.com/order/catalog/product/42I#/42I
[5] http://www.teledyne-api.com/products/oxygen-compound-instruments/t400
[6] https://www.aerosolmageesci.com/products/aerosol-magee-scientific-aethalometer/

The following analyzers will be used at Afri-SET as references for the 2025 Challenge:

- **Particulate matter** is measured using the Teledyne T640 analyzer[7], an optical spectrometer (measurement principle described by NF EN 16450: 2017) for PM_{10} and $PM_{2.5}$. **Hourly averages** are used, but finer time resolutions are also possible.
- **Nitrogen oxides** are measured by using an EcoTech Serinus 40 NOx chemiluminescence analyzer[8], which provides measurements for **NO** and NO_2 (Reference standard for ambient air: NF EN 14211: 2012). In the context of the Challenge **hourly averages** are used, however the data from this analyzer is also available at finer resolutions (e.g., down to one-minute scans).

Figure 4: Outdoor evaluation site – Indi-SET station in Bengaluru exterior (left) and interior (right) views.

The **Outdoor Tropical climate – South Asia evaluation environment** is represented by the first Indi-SET center established by CSTEP in its Bengaluru campus. It constitutes an urban background site located at 13° 2' 54.6" N, 77° 34' 46.5" E and at approximately 10 meters above ground level (see Figure 4). Bengaluru is the capital of the Karnataka state and India's third most populous city. Bengaluru, at 920 m (3,020 ft) elevation, has a tropical savanna climate with moderate temperatures year-round. The dry season (November–April) is warm and sunny, while the rainy season (May–October) brings monsoon showers, peaking from June to September. Humidity is moderate, and evenings tend to be cooler due to the city's altitude. Usage categories evaluated at this site are **OA-A** and **OA-M**.

The following Indi-SET analyzers will be used in Bengaluru as references for the 2025 Challenge:

- **Particulate matter** is measured using the Vasthi Instruments Vair-9009[9] working on the principle of Beta ray attenuation method and Palas Fidas 200S[10] based on the optical light scattering by single particles. For the Challenge, we make use of $PM_{2.5}$ and PM_{10} from Vair-9009 and the particle counts (**PN**) from the Fidas 200S. **Hourly averages** are used, but finer time resolutions for FIDAS are available.
- **Nitrogen oxides** are measured by using a Kentek Mezus 210 chemiluminescence analyzer[11], which provides measurements for **NO** and NO_2. In the context of the Challenge **hourly averages** are used, however the data from this analyzer is also available at finer resolutions (e.g., down to one-minute scans).
- **Ozone** is measured with the 2.B Technologies Model 108-L[12] UV absorbance analyzer. **Hourly averages** are used, but finer resolution data is also available (e.g., 15 min averages, 10 s scans).

[7] https://www.atecorp.com/products/teledyne-api/t640
[8] https://metone.com/products/serinus-40-nitrogen-oxides-analyzer/
[9] https://vasthi.com/product/22/pm10-and-pm2.5-caaqms
[10] https://www.palas.de/en/product/fidas200s
[11] https://emin.asia/kentekmezus-210-kentek-mezus-210-nox-analyzer-0-1-0-5-1ppm-151761/pr.html
[12] https://2btech.io/items/industrial-ozone-monitors/model-108-l-ozone-monitor/

- **Aerosol Black Carbon** is measured with the Magee Scientific AE33 aethalometer. **Hourly averages** are used, but time resolutions as rapid as 1s are possible.
- **Carbon monoxide** is measured using the Horiba Ambient CO monitor APMA- 370[13], using a non-dispersion cross modulation infrared analysis method. **Hourly averages** are used, but 3-minute data is also possible.

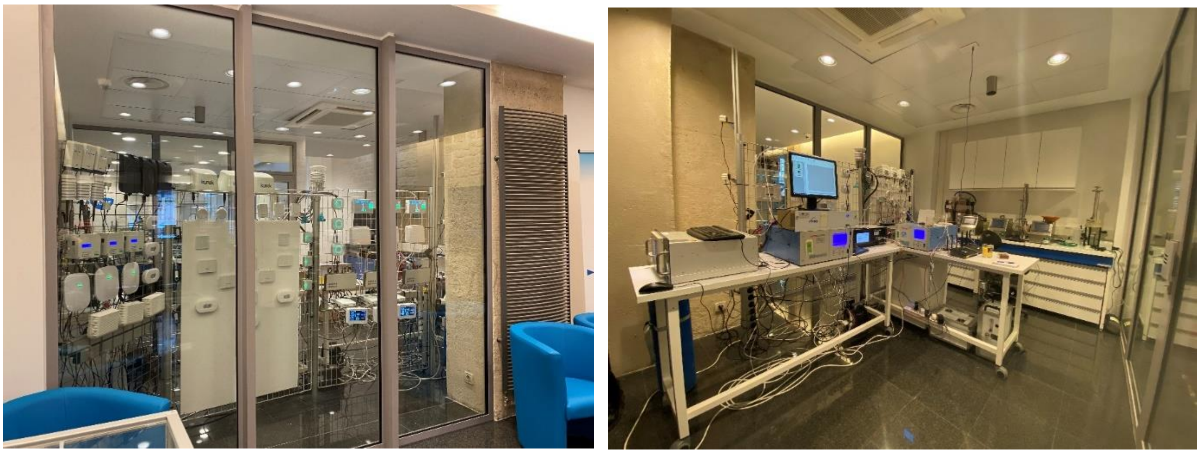

Figure 5: Front view of the indoor air sensor deployment (left) and back view showing the reference analyzers (right).

For the evaluation of indoor air quality sensors for non-specific spaces, a dedicated extension of Airparif's metrology laboratory is used as **Indoor Air – Non-Specific Spaces evaluation environment** (see Figure 5), which was built in January 2023 for large volume (i.e. room size) exposure testing of measurement systems. It provides easy access to the necessary reference analyzers and gas circuitry, air conditioning of the room which allows a certain degree of control over the environment, controlled access to the deployment, and, last but not least, the availability of a glass wall which allowed the tests to be showcased for communication purposes.

The usage categories that are evaluated in this space are **IA-A**, **IA-M**, and **IA-P**. For this purpose, the following material is used as reference:

- **Particulate matter** is measured using the PALAS Fidas 200 analyzer, an optical spectrometer (measurement principle described by NF EN 16450: 2017). The Fidas provides multiple outputs, including particle granulometry. Currently, for the Challenge, we make use of its **PM10**, **PM2.5**, **PM1**, and particle count (**PN**) outputs with a **10 min average** temporal aggregation.
- **Carbon dioxide** is measured using the Thermo Scientific 410i[14], a non-dispersive infrared (NDIR) analyzer, with a **10 min average** temporal aggregation.
- **Nitrogen oxides** are measured by Thermo Scientific 42i chemiluminescence analyzer, which provides measurements for **NO** and **NO_2** (measurement principle described by NF EN 14211: 2012). For the indoor evaluations **10 min averages** are used.
- **Volatile organic compounds (VOCs)** concentrations are measured using the Syntech Spectras GC 955[15], which is based on the method of automatic pumped sampling with in-situ gas chromatography (described by NF EN 14662-3). It is configured for **15 min-averaged** measurements.
- **Formaldehyde** concentrations are measured using the method of pumped sampling on Tenax sorbent tubes followed by thermal desorption and gas chromatography analysis

[13] https://www.horiba.com/int/process-and-environmental/products/detail/action/show/Product/apma-370-453/
[14] https://www.thermofisher.com/order/catalog/product/410I#/410I
[15] https://www.synspec.nl/products/gc-955.html

(measurement principle described by NF EN 14662-1). The Tenax tubes are exposed with an **8 hour periodicity**.

In order to investigate the response of the sensors, a number of specific stimulation scenarios are considered for the indoor evaluations by using regular and electronic cigarettes, candles, incense, cleaning products, and cooking.

Figure 6: Indoor Air-URS evaluation site – Avenue Foch RER C station exterior view (left) and the reference monitoring station deployed on the platform (right).

The **Indoor Air – Underground Railway Stations evaluation environment** is installed on the platform of the Avenue Foch Station, an RER C station operated by the SNCF (see Figure 6). It is located in the 16th Arrondissement of Paris at 48°52’14”N (latitude) and 2°16’31”E (longitude). The Challenge categories evaluated at the Avenue Foch Station is **IA-URS**.

A long-term monitoring station is already present on the station’s platform which will be used as reference for the Challenge and includes the following analyzers:

- **Particulate matter** is measured using the Thermo Scientific 1405-F monitor which uses a tapered element oscillating microbalance (TEOM) and a Filter Dynamics Measurement System (FDMS)[16] to provide **PM_{10}, $PM_{2.5}$, PM_1** data (Reference standard for ambient air: NF EN 16450: 2017) and a PALAS Fidas 200 analyzer, an optical spectrometer, for particle counts (**PN**). **15 min averages** are used for the Challenge, but finer time resolutions are possible with this type of analyzer.
- **Carbon dioxide** is measured using the Thermo Scientific 410i, a non-dispersive infrared (NDIR) analyzer, with a **15 min average** temporal aggregation.

2.3 Sensor data pre-processing and validation

To ensure the consistency and transparency of the data used in the evaluation process, it is important to define clear guidelines for data pre-processing and technical validation. This section outlines the procedures for handling data invalidation, addressing technical issues, and managing potential data loss. While we apply technical validation to ensure the reliability of the data, it is important to note that environmental criteria will not be used to invalidate any data. Environmental invalidation would require the use of expert knowledge, which may not be available to all users of sensor systems. Therefore, we rely exclusively on technical validation to maintain fairness and consistency throughout the Challenge. Data pre-processing and validation will be conducted according to the following principles:

- **Invalidation based on manufacturer-side algorithms** – Data invalidation will be handled using the **pre-existing algorithms** implemented by the manufacturer, available to all users. These algorithms will be applied consistently across all data, and no additional invalidation requests from participants that would be specific to the Challenge campaign will be accepted.

[16] https://www.thermofisher.com/order/catalog/product/fr/en/TEOM1405F

- **Technical invalidation due to outliers –** The objective of this step is to remove outliers and other aberrant data that result from non-systematic sources of error or critical sensor failures. However, systematic errors – such as those arising from cross-sensitivities, environmental factor sensitivities, or other inherent limitations of the sensor technology – are not subject to invalidation.
  - **Spurious, non-systematic erroneous data** will be removed based on technical criteria.
  - **Negative values will be retained** unless clearly caused by sensor malfunctions.
  - **Aberrant data resulting from identifiable sensor critical failures** (e.g., values stuck at zero or a fixed maximum level) will be removed.
  - **Minimum valid data presence requirement**: A minimum of 75% valid data must be available within each **aggregation window** (e.g., for calculating hourly means in outdoor evaluations).
- **Data loss due to deployment infrastructure issues –** Data loss resulting from deployment infrastructure issues – such as power or communication outages, or errors in sensor installation or configuration that are our responsibility – will not affect the **Presence sub-criterion** in the evaluation score. These issues will be clearly documented, and the impact on the data will be accounted for appropriately.

2.4 Performance criteria

The AIRLAB Microsensors Challenge adopts a holistic approach to air quality sensor evaluation by combining, for the 2025 edition[17], *accuracy*, *utility*, *usability*, and *cost* criteria. In this subsection we present each of these criteria and detail their calculations.

1. **Accuracy** – The accuracy performance criterion is defined based of the Sensor Evaluation Toolkit (SET) index from Fishbain et al. [43], enriched with additional criteria for trueness and precision. The SET Global Method Index includes seven evaluation metrics:

- The *Root-Mean-Square Error* (RMSE) is a frequently used error metric for numerical deviations. While being an excellent and popular general-purpose error measure, it is sensitive to outliers, and when used on its own it can be disproportionally penalizing for signals that contain large sporadic errors. This shortcoming can be counterbalanced by the use of correlation coefficients.
- The *Pearson correlation coefficient* (ρ) characterizes the presence of a linear relationship between two signals (e.g., reference and candidate sensor). It is the most commonly used correlation criterion.
- The *Kendall correlation coefficient* (τ) and the *Spearman correlation coefficient* (S) are two different rank correlation coefficients which are used to test for the presence of a non-linear relationship between two variables.
- The *Presence* ($s_{presence}$) metric represents the evaluation of the completeness of the data, highlighting sensor failures, operational or data transmission problems.
- The *Source analysis* (s_{source}) characterizes the capacity of the device to identify and localize a source (perception of the variations of pollutant level as a function of wind direction).
- The *Match score* (s_{match}) relates to the common use of air quality grading schemes (e.g., the Air Quality Index) in the context of applications that do not require precise absolute measurements such as citizen science projects or general risk estimations. It consists in the

[17] Previous editions of the Challenge included also the criterion of portability, associated with mobile evaluation categories. For the 2025 edition no mobile applications are considered.

division of the reference and candidate sensor dynamic ranges into equal number of bins and quantifying the bin-classification agreement for reference-candidate measurement pairs.

- The *Lower Frequencies Energy* (LFE) metric characterizes the acquired sensor signal rather than a comparison with a reference instrument and reflects the sensor's ability to capture the temporal variability of the targeted pollutant.

The additional criteria used in the Challenge are:

- The *Slope* (*b*) and *Intercept* (*a*) of a linear regression model of the relationship between the reference measurements and the microsensor measurements. The values of these two parameters permit an evaluation of the trueness of the candidate solution. A partial score s_b is assigned as a function of the value of the slope, by splitting its possible values into three groups as detailed in Table 2. The split is based on the approach adopted by the LNE and the INERIS for their "AIR Quality Sensor" Certification[18].

Table 2: Assignment of the s_b score for the slope.

	Group A	Group B	Group C
Slope (*b*)	0.7 ≤ b ≤ 1.3	0.5 ≤ b < 0.7 or 1.3 < b ≤ 1.5	b < 0.5 or b > 1.5
Score (s_b)	1	0.5	0

A second partial score, s_a, is calculated based on the intercept using the following formula:

$$s_a = \begin{cases} 1 - \dfrac{|a|}{\text{median}(Y_{REF})}, & if\ |a| \leq \text{median}(Y_{REF}) \\ 0, & if\ |a| > \text{median}(Y_{REF}) \end{cases}$$

where Y_{REF} is denotes the reference data.

The final score for the trueness criterion, s_t, is the average between s_b and s_a.

- *Reproducibility* is an expression of the precision of the candidate solution and is calculated across the microsensor samples of a candidate solution, it includes both the variability due to causes intrinsic to one sensor unit (e.g., measurement noise) and inter-device variability (e.g., due to the manufacturing process). It is calculated[19] according to the CEN TS 17660 standard[20] as the standard deviation of the reproducibility normalized by the measurement average and expressed as a percentage, s_R.

The final *Integrated Performance Index* (IPI) aggregates the eight metrics and has a value between 0 and 1 (1 being equivalent to the reference method):

$$IPI = average\{(1 - NRMSE), \rho, \tau, S, s_{presence}, s_{source}, s_{match}, LFE, s_t, 1 - s_R/100\}$$

where *NRMSE* represents the normalized *RMSE*, which we calculate as the ratio between the error and the measurement range:

$$NRMSE = \frac{RMSE}{\max(Y_{REF}) - \min(Y_{REF})}$$

In the context of the Challenge, the accuracy of the candidate sensors is calculated by using a reference measurement and the presented method for calculating the IPI with data acquired over a time interval of at a minimum ten days. An example of a result of the complete accuracy calculation is illustrated in Table 3.

[18] https://prestations.ineris.fr/en/certification/certification-sensors-system-air-quality-monitoring

[19] The calculation is made only for complete triplets of data across the three units under test (i.e. periods of data loss are discarded).

[20] CEN TS 17660 standard: Air quality – Performance evaluation of air quality sensor systems

Table 3: Example of accuracy result.

	SET method									
	Match	RMSE	Pearson	Kendall	Spearman	Presence	LFE	Trueness	Repro	IPI
Sensor #1	0.44	0.85	0.83	0.62	0.82	0.96	0.99	0.84	0.89	**0.80**

2. **Utility** – This criterion reflects the capacity of a sensor system to provide the essential functionalities for accomplishing the targeted goal. The criteria considered to evaluate Utility vary based on the category of use, with two criteria always present: *targeted pollutants* and *data recovery*:

- The *targeted pollutants* sub-criterion rewards a good match between the type of measurements provided by the sensor platform and the pollutant of interest for a given environment. It is calculated based on Table 4, by adding the corresponding coefficients for the pollutants targeted by the candidate solution, up to a maximum score of 1.

Table 4: Evaluation grid for calculation of the targeted pollutants sub-criterion.

	OA			IA	
	West Africa	South Asia	Europe	Non-Specific	URS
NO2	0.4	0.4	0.4	0.3	-
CO2	-	-	-	0.4	0.4
PM10	0.4	0.4	0.4	0.3	0.4
PM2.5	0.4	0.4	0.4	0.4	0.4
PM1	0.2	0.2	0.2	0.2	0.1
O3	0.3	0.3	0.3	-	-
CH2O	-	-	-	0.4	-
VOC	-	-	-	0.4	0.2
BC	0.3	0.3	0.3	-	-
Benzene	0.2	0.2	0.2	0.4	-
SO2	0.3	0.1	0.1	-	-
CO	0.3	0.2	0.1	0.4	-
CP	0.2	0.2	0.2	0.2	0.3
H2S	0.1	0.1	0.1	0.2	-
NH3	0.3	0.3	0.3	0.2	-
NO	0.2	0.2	0.2	0.2	-

- The *data recovery* characterizes the communication options that allow the recuperation of measurement data from the sensor for inspection, analysis or further processing. It is calculated based on Table 5, by adding the available communication options for the candidate solution, up to a maximum score of 1.

Table 5: Evaluation grid for calculation of the data recovery sub-criterion.

	Outdoor Air	Indoor Air
Physical connection	0.25	0.25
Short-distance wireless (e.g., Bluetooth, WiFi etc.)	0.25	0.75
Long-distance wireless (e.g., WWAN, LoRa)	0.75	0.75

Other five sub-criteria that, depending on the category of use can be part of the utility or the usability criterion calculation are:

- The energy *autonomy* of a microsensor product can play an important role in its usability, particularly for settings that do not provide access to a standard power supply connection (e.g., mobile applications, remote fixed locations, etc.) or if the power supply, although available, is of poor quality (e.g., prone to outages). This characteristic is graded according to the grid in Table 6.

Table 6: Evaluation grid for the autonomy sub-criterion.

	Autonomy in hours
1	> 72
0.9	72
0.8	48
0.7	24
0.6	16
0.5	8
0.4	6
0.3	4
0.2	2
0.1	1
0	Power socket

- *Data interoperability* characterizes a system's ability to allow for data exchange with other systems without technical restrictions. In the context of the Challenge, we consider the ability of the candidate microsensor systems to allow for the use of its data by other systems. The score for this sub-criterion is calculated according to Table 7.

Table 7: Evaluation grid for the data interoperability sub-criterion.

Data not accessible	0.00
Proprietary data format	0.25
Open data format	0.75
Open format respecting INSPIRE[21]	1.00

- *Data visualization* plays an important role in the understanding of measurement data. More so for devices that are designed for non-expert users as is often the case for air quality microsensors. In this context we consider a two-dimensional partitioning of the possible visualization solutions (see Table 8). On the horizontal axis we consider the availability of either a real-time or offline (subsequent to the measurement experiment) visualization solution. On the second axis we consider whether a display (or alert indicator) is integrated directly on the sensor, constitutes a remote solution (e.g., leveraging a mobile phone display or a computer display through a cloud service), or no possibility of displaying the data exists. The presence of an integrated display has an impact only for awareness raising applications for indoor applications, as this feature might actually be undesirable for other types of applications.

Table 8: Evaluation grid for the data visualization sub-criterion.

	Real-time		Offline	
	IA-A	OA, IA-M, IA-P, IA-URS	IA-A	OA, IA-M, IA-P, IA-URS
Integrated display	1.00	1.00	0.75	0.75
Remote display	0.75	1.00	0.50	0.75
None	0.00	0.00	0.00	0.00

[21] https://inspire.ec.europa.eu

- The *measurement time step* sub-criterion refers to the sampling period of the microsensor solutions. A relatively higher temporal resolution is typically needed for mobile sensing applications. The score for this sub-criterion is given according to Table 9.

Table 9: Evaluation grid for the measurement time step sub-criterion.

Sampling period	Score
≤ 1 min	1.00
5 min	0.80
15 min	0.60
60 min	0.40
120 min	0.20
> 240 min	0.00

- *Real-time notifications* – relates to whether the considered microsensor system permits the transmission of notifications to the operator/user either directly through an integrated screen/indicator or through remote messaging (e.g., SMS, e-mail). The score of this sub-criterion is evaluated according to Table 10.

Table 10: Evaluation grid for the real-time notifications sub-criterion.

Type of notification	Score
None	0.00
Visual	0.50
Remote	0.75
Visual and remote	1.00

The way in which these five sub-criteria are assigned to the utility criterion depending on the category of use is presented in Table 11. The overall utility criterion score is calculated as the average over all considered sub-criteria.

Table 11: Sub-criteria forming the utility criterion depending on category of use.

OA-M	Targeted pollutants	Data recovery	*Data interoperability*	*Real-time notifications*	-
OA-A	Targeted pollutants	Data recovery	*Data visualization*	*Real-time notifications*	-
IA-M	Targeted pollutants	Data recovery	*Data interoperability*	*Real-time notifications*	*Measurement time step*
IA-A	Targeted pollutants	Data recovery	*Data visualization*	*Real-time notifications*	-
IA-P	Targeted pollutants	Data recovery	*Data interoperability*	*Measurement time step*	-
IA-URS	Targeted pollutants	Data recovery	*Data interoperability*	*Measurement time step*	*Autonomy*

3. **Usability** – This criterion characterizes the ability of the candidate solution to provide the conditions for its users to perform the tasks safely, effectively, and efficiently while enjoying the experience. The criteria considered to evaluate Usability vary based on the category of use (detailed in the Challenge Protocol), with two criteria always present – the Ease-of-Use and the Statistical summary:

- The *ease-of-use* sub-criterion is calculated using a timed start-up test. The candidate sensor is unpackaged and a timer is started. If available, its user manual is consulted and the sensor is switched on. The timer is stopped when its correct operation can be confirmed (e.g., measurement values read on an integrated display or through a cloud interface). Two grades are given following this test. The first one, $s_{impression}$, is based on the general impression of the test operator who can give one of 4 possible qualifiers. These are subsequently scored based on the evaluation grid in Table 12.

Table 12: Evaluation grid for general impression score.

Score	Qualifier
0	Unsatisfactory
0.50	Average
0.75	Satisfactory
1	Excellent

The second grade, s_{time}, is based on the time elapsed for performing the test and considers the presence and quality of the user manual through an additive bonus-malus factor, which is applied if the start-up test takes longer than 5 min. The s_{time} grade, including any potential bonus-malus factor, is bounded between 0 and 1. The principle of calculating this grade is summarized in Table 13. The final grade for the ease-of-use sub-criterion is the average between $s_{impression}$ and s_{time}.

Table 13: Evaluation grid for elapsed time score.

	Elapsed time [min]
1	≤ 5
0.9	10
0.8	15
0.7	25
0.6	30
0.5	45
0.4	> 50

	Manual presence/quality
-0.3	No manual
-0.1	Unsatisfactory
+0.1	Average
+0.2	Satisfactory
+0.3	Excellent

- *Statistical summary* – this sub-criterion characterizes whether or not the microsensor product provides statistical options for the measured data. These statistical options relate on the one hand to the possibility of accessing historical data, either in raw form or through customizable aggregations, and on the other hand to the availability of statistical summaries. We classify statistical summaries into two broad classes: indicative or comparative. Indicative statistics are direct calculations exclusively on the basis of the measured data (e.g., mean, median, minimum, maximum), while comparative statistics highlight relationships with respect to specific external benchmarks (e.g., limit levels, statistics over a population, a region, or a historical period). The method of evaluating the analytics criterion is presented in Table 14.

Table 14: Evaluation grid for the statistical summary sub-criterion.

		Statistics		
		None	Indicative	Comparative
History	None	0.00	0.25	0.50
	Raw only	0.25	0.50	0.75
	Customizable aggregations	0.50	0.75	1.00

The rest of the sub-criteria considered in the calculation of the usability criterion and their assignment as a function of the category of use is presented in Table 15. The overall usability criterion score is calculated as the average over all considered sub-criteria.

Table 15: Sub-criteria forming the usability criterion depending on category of use.

OA-M	Ease of use	Statistical summary	*Autonomy*	*Data visualization*
OA-A	Ease of use	Statistical summary	*Autonomy*	-
IA-M	Ease of use	Statistical summary	*Autonomy*	*Data visualization*
IA-A	Ease of use	Statistical summary	*Autonomy*	*Measurement time step*
IA-P	Ease of use	Statistical summary	*Real-time notifications*	*Data visualization*
IA-URS	Ease of use	Statistical summary	*Real-time notifications*	*Data visualization*

4. The **Environmental Footprint** criterion provides an assessment of the sensor system's lifecycle impact, considering factors such as material efficiency, manufacturing location, maintenance requirements, expected lifetime, and Corporate Social Responsibility (CSR) initiatives. While this evaluation offers a partial estimate of the system's environmental impact and does not claim to be exhaustive, it serves as a valuable approximation, as even a partial assessment is a meaningful step towards understanding its broader environmental effects. The score is based on the following five sub-criteria:

- *Material Efficiency* (20%) – Evaluates the material footprint of the sensor based on its mass, volume, and use of sustainable materials. It is calculated as the average of two sub-scores: the *form factor* and the *sustainable materials* score:
    - The *form factor* score is calculated based on two characteristics of the microsensor product: its mass and its volume. Considering the number of sensors integrated into the sensor system, $N_{sensors}$, the score is calculated by first attributing a normalized mass score, $s_{\overline{mass}}$, and a normalized volume score, $s_{\overline{volume}}$ (see Table 16), and then computing the form factor score as follows:

$$s_{form\ factor} = \sqrt{s_{\overline{mass}} \cdot s_{\overline{volume}}},$$

where

$$\overline{mass} = \frac{Mass}{N_{sensors}}$$

and

$$\overline{volume} = \frac{Volume}{N_{sensors}}$$

Table 16: Evaluation grids for mass (left) and volume (right) scores for the form factor score.

	$\overline{mass}$ in g
1	< 100
0.9	250
0.8	500
0.7	750
0.6	1000
0.5	2000
0.4	4000
0.3	6000
0.2	8000
0.1	100000
0	> 10000

	$\overline{volume}$ in cm^3
1	< 10
0.9	50
0.8	100
0.7	200
0.6	400
0.5	800
0.4	1600
0.3	3200
0.2	6400
0.1	12800
0	> 12800

- The *sustainable materials* score is calculated based on the percentage of recycled or responsibly sourced materials used in the production of the sensor system (see Table 17).

Table 17: Evaluation grid for the sustainable materials score.

Recycled or responsibly sourced materials	Score
≥ 50%	1.00
40%	0.80
30%	0.60
20%	0.40
10%	0.20
< 10%	0.00

- *Manufacturing Impact*[22] (20%) – Assesses emissions linked to the production and transport of the sensor. It is calculated based on the average of two scores: the *manufacturing location* and the *renewable energy* scores, calculated according to Table 18.

Table 18: Evaluation grids for the manufacturing location (left) and renewable energy (right) scores for the manufacturing impact sub-criterion.

Distance between manufacturing and deployment location	Score
≤ 500 km	1.00
501-2000 km	0.75
2001-4500 km	0.50
4501-6000 km	0.25
> 6000km	0.00

Share of renewable energy in manufacturing	Score
≥ 50%	1.00
25%-49%	0.5
< 25%	0.0

- *Maintenance & Serviceability* (20%) – Evaluates frequency and environmental impact of maintenance and is calculated as a weighted average of the *reduced maintenance* score ($s_{reduced\ maintenance}$) and the *eco-friendly maintenance* practices score ($s_{eco-friendly\ maintenance}$), as follows:

$$s_{Maintenance} = 0.75 \cdot s_{reduced\ maintenance} + 0.25 \cdot s_{eco-friendly\ maintenance}$$

- The *reduced maintenance* score is a measure of the periodicity of necessary maintenance operations. Its value represents the average of the user maintenance score and the professional maintenance score. The former refers to the maintenance operations to be performed by the microsensor platform user. Examples of such operations are: cleaning of inlets, filter changes, basic calibration (e.g., for zero levels), battery replacement, etc. The latter is defined as a measure of the periodicity of all maintenance operations that cannot be performed directly by the user and need the intervention of a specialized technician. The values of these scores are determined according to the grids presented in Table 19.

[22] This evaluation pertains solely to the environmental impact at the system assembly level and does not account for the environmental impact of individual subcomponents within the sensor.

Table 19: Evaluation grids for the user maintenance (left) and professional maintenance (right) scores for the maintenance & serviceability sub-criterion.

User Maintenance	
Periodicity	**Score**
Hourly	0
Daily	0.1
Weekly	0.2
Monthly	0.4
Trimestral	0.6
Annually	0.8
> Annually	1

Professional Maintenance	
Periodicity	**Score**
< Monthly	0
Monthly	0.1
Trimestral	0.3
Annually	0.6
18 Months	0.8
> 18 Months	1

- The *eco-friendly maintenance practices* score rewards the implementation of responsible maintenance and service practices (e.g., modular components, remote diagnostics, waste-free servicing) and is calculated according to Table 20.

Table 20: Evaluation grid for the eco-friendly maintenance practices score.

Number of implemented practices	Score
≥ 3	1.00
2	0.66
1	0.33
None	0.00

- *Product Lifetime & Durability* (25%) – Rewards sensors with a longer operational life and responsible end-of-life management and is the weighted average of the *expected lifetime* score ($s_{expected\ lifetime}$) and the end-of-life management score ($s_{end-of-life}$):

$$s_{Lifetime\ \&\ durability} = 0.6 \cdot s_{expceted\ lifetime} + 0.4 \cdot s_{end-of-life}$$

Table 21: Evaluation grids for the expected lifetime (left) and end-of-life management (right) scores for the product lifetime & durability sub-criterion.

Expected lifetime	Score
≥ 8 years	1.00
5-7 years	0.66
3-4 years	0.33
< 3 years	0.00

End-of-life management	Score
Take-back recycling program & biodegradable parts	1.00
Either take-back/recycling or biodegradable parts	0.50
None	0.00

- *CSR & Environmental Initiatives* (15%) – Based on the number and scope of CSR initiatives reported and calculated according to Table 22:

Table 22: Evaluation grid for the eco-friendly maintenance practices score.

Number of CSR initiatives	Score
≥ 4	1.00
3	0.75
2	0.50
1	0.25
None	0.00

The final Environmental Footprint score is calculated as the following weighted average:

$$s_{Environmental\ Footprint} = 0.2 \cdot s_{Material\ Efficiency} + 0.2 \cdot s_{Manufacturing\ Impact} + 0.2 \cdot s_{Maintenance} + 0.25 \cdot s_{Lifetime\ \&\ durability} + 0.15 \cdot s_{CSR\ initiatives}$$

5. **Cost** is an important selection criterion for any product. In the context of the Challenge, we consider the cumulated investment and running costs (e.g., for subscriptions, sensitive element replacement, etc.) over the first 3 years of the microsensor system use, divided by the number of sensors integrated in the sensor system, $N_{sensors}$. The cost criterion is graded similarly to the other criteria, on a scale from 0 (most expensive) to 1 (least expensive), as presented in Table 23, with costs falling between two consecutive classes being graded through a linearization between the two corresponding score indexes.

Table 23: Evaluation grid for the cost criterion.

	Cost
1.0	<100 €
0.9	200 €
0.8	500 €
0.7	1 000 €
0.6	2 000 €
0.5	5 000 €
0.4	8 000 €
0.3	10 000 €
0.2	20 000 €
0.1	30 000 €
0.0	> 30000 €

2.5 Criteria weighting

The performance criteria presented in the previous subsection are naturally more or less relevant depending on the use category targeted by the platform. For instance, the Accuracy of a microsensor that is competing in categories that target support for monitoring applications (e.g. OA-M, IA-M) is more important than for applications that target only raising awareness (i.e., OA-A and IA-A). Likewise, the cost of a platform to be used for regulatory purposes is less important than for the other categories.

In order to reflect this type of considerations, we use a criteria-weighting mechanism which modulates the criteria evaluation scores as follows:

- Once the different performance criteria are evaluated for a microsensor candidate platform, its obtained scores are then weighted for each of the categories that it is competing in, by using the weighting factors summarized in Table 24.
- The resulting weighted scores are then averaged to obtain an overall grade on a star scoring scheme, between 0 and 5 stars, for each of the categories in which the microsensor platform is competing.
- Sensors competing in the fixed Outdoor Air categories (OA-M and/or OA-A) receive distinct evaluation scores for each of the three respective evaluation environments (i.e. Accra, Bengaluru and Paris).

Table 24: Performance criteria weighting as a function of usage categories.

		Accuracy	Utility	Usability	Environmental Footprint	Cost	
						Paris	Accra/Bengaluru
OA	OA-M	5	5	3	2	3	4
	OA-A	3	4	4	3	5	
IA	IA-M	5	5	3	2	3	
	IA-A	3	4	4	3	5	
	IA-P	3	5	4	2	4	
	IA-URS	5	5	3	2	3	

2.6 Deliverables

The results of the Challenge are published through a dedicated interactive Web interface which enables an interactive user experience, allowing for searches by specific criteria, and side-by-side comparisons of different candidate solutions. For the current edition we will seek to consolidate and further refine this interactive interface by providing new features. However, no major overhaul of the interface is envisioned. The main principles of clarity and accessibility that were at the core of previous Challenge editions deliverables will continue to guide its design.

In the previous edition (2023) several new features were added to this interface which will be retained for the current edition:

- **Zoom-in on performance details** – the ability to visualize further details of the evaluation criteria (e.g., for the accuracy criterion a possibility to see the scores of its sub-criteria).
- **Time series graphs** – the inclusion of images of sample graphs of the temporal evolution of the sensors versus the reference.
- **Additional accuracy metrics** – for information purposes, a number of alternative accuracy metrics, used by other sensor evaluation initiatives, with no direct impact on the calculation of the AIRLAB Challenge Accuracy score (e.g., R-squared, Mean Bias Error, Mean Absolute Error).

A final version of the protocol will be published once the new design of the results interface will be finalized. Since all of the features of the previous edition will still be available for the 2025 edition, we present the 2023 Challenge Results interface for reference below.

The results portal is available with bilingual language support (i.e. in English and French). It consists of tabbed view with four selectable options labeled *Search*, *Reports*, *Comparison*, and *About.*

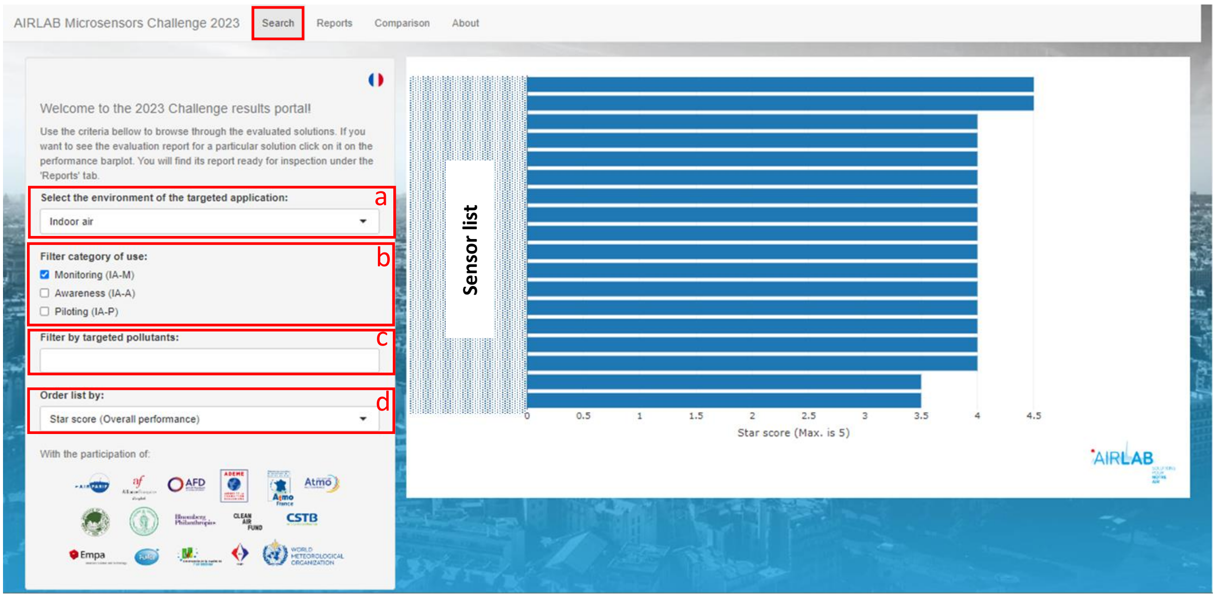

Figure 7: Overview of the *Search* tab.

The **Search** tab (see Fig. 7) allows the user to select the deployment environment for the targeted application (a), the category of use (b), to filter only for the candidate solutions that provide measurements for pollutants from a configurable list (c), and to order the resulting list of sensors depending on a criterion (d). The sensor list is interactive, allowing the user to click on a sensor of interest, which will set the sensor to be viewed in the Reports tab, and the first sensor in the Comparison tab.

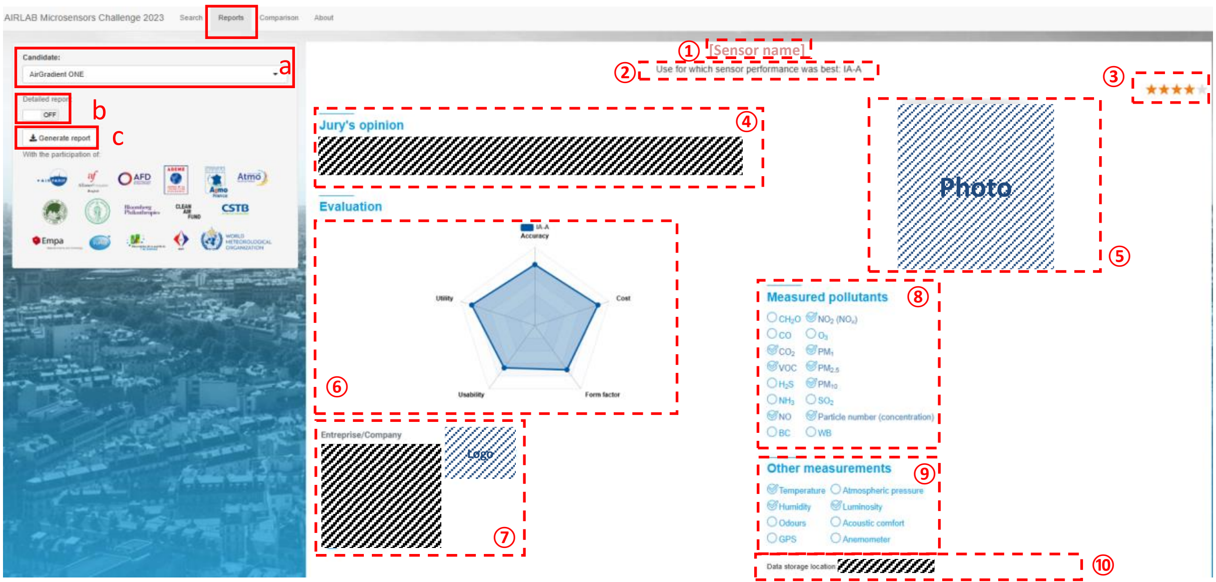

Figure 8: Overview of *Reports* tab

The **Reports** tab (see Fig. 8) gives access to the results report for each candidate solution. By default, if a selection was made in the research tab, it will be reflected here in the selected sensor report for display. From the side panel of this tab, the candidate sensor selection can be changed (a), the display of the detailed report can be turned on or off (b) and a PDF of the displayed report can be downloaded.

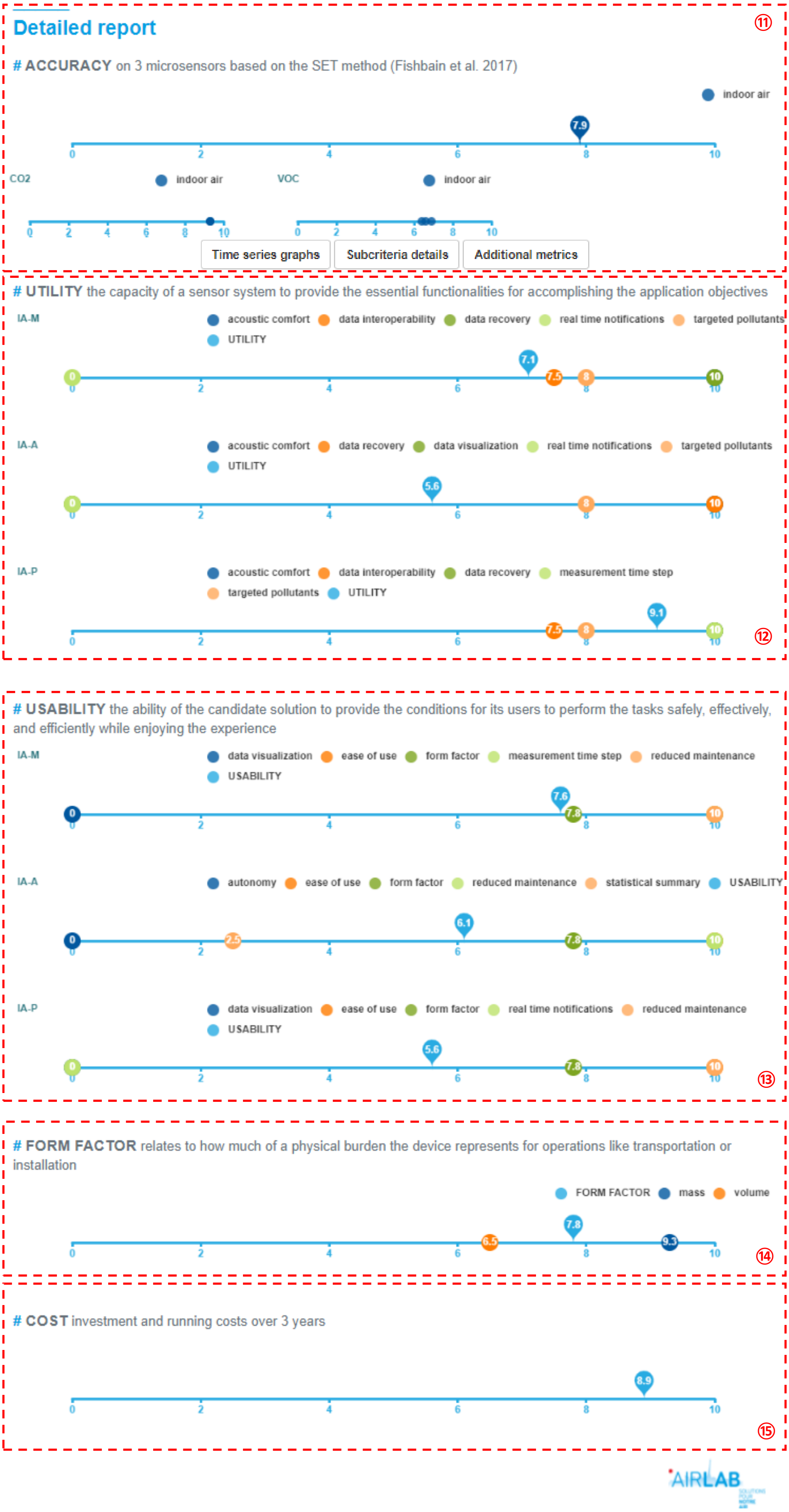

Figure 9: Detailed report overview

The short version of the report contains the following elements:

1) The name of the microsensor product.
2) Sentence stating the category in which the candidate performed the best.
3) The overall star score (0 to 5) of the microsensor for the category in which it performed best.
4) Short paragraph presenting the overall review of the microsensor by the Challenge jury.
5) The photo of the microsensor product.
6) Radar chart showing the five main criteria scores of the microsensor for each use case category. Clicking on the legend elements allows to hide/show the different use categories.
7) Name, logo and coordinates of the company commercializing the microsensor product.
8) Check list marking the pollutants targeted by the microsensor solution. Hovering over the targeted pollutants with the mouse reveals a tooltip regarding their level of processing, as defined in Schneider et al., 2019 [4].
9) Check list marking additional measured environmental parameters.
10) The data storage location.

When the detailed report switch is turned on, the following additional information is displayed (see Fig. 9):

11) Detailed graphical representation of the accuracy results breakdown: by category of use, by targeted pollutant, and for each of the three provided microsensor samples. It includes also the three new features introduced in the 2023 edition: time series graphs, subcriteria details and additional metrics, all accessible through push buttons.
12) Graphical representation of the utility criterion result and its sub-criteria by category of use.
13) Graphical representation of the usability criterion result and its sub-criteria by category of use.
14) Graphical representation of the portability or, for static solutions, the form factor criterion results.
15) Graphical representation of the cost criterion result.

The **Comparison** tab (see Fig. 10) allows a side-by-side comparison of two device reports and follows the same logic of report composition. Once the first candidate solution is selected, the second one can be chosen only from a list of devices that share at least one evaluation setting in common with the first.

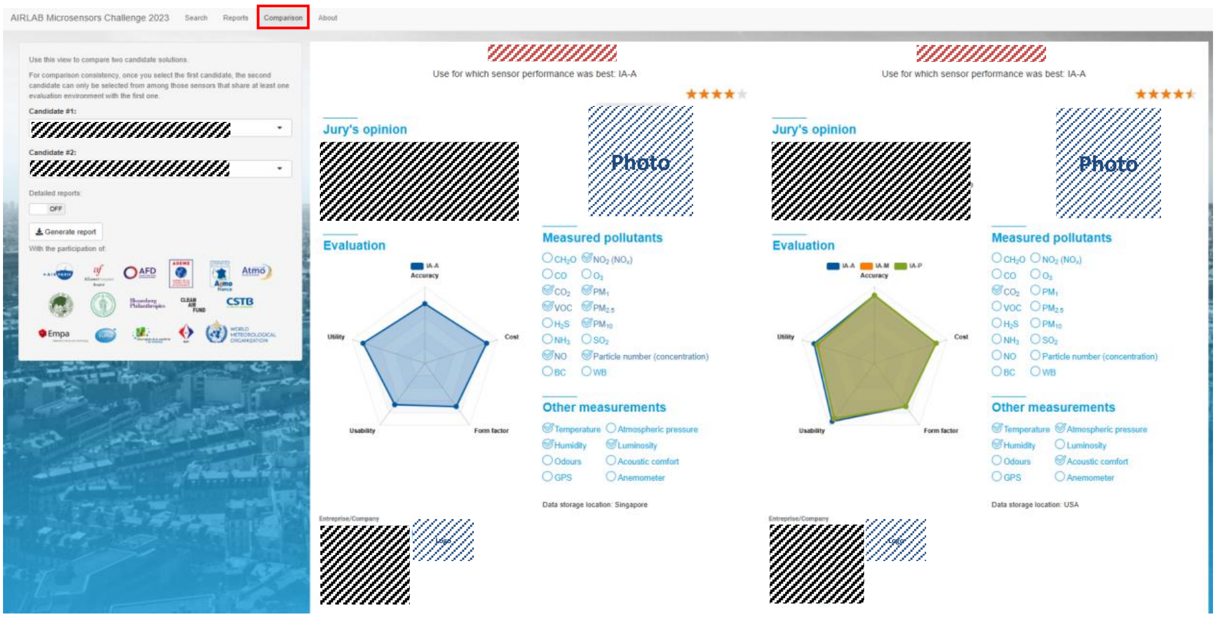

Figure 10: Overview of *Comparison* tab

The **About** tab provides general information on the context of the AIRLAB Microsensors Challenge, its scope and limitations.

All results summaries are made publicly available on the AIRLAB website[23].

[23] http://www.airlab.solutions/

3 Discussion

The AIRLAB Microsensors Challenge represents a periodic evaluation of the state of the art of commercially available microsensors solutions for air quality monitoring. Its main goal is to create a large and publicly available information repository for the benefit of all potential users (i.e. academics, industry, and the general public).

The results of the AIRLAB Microsensors Challenge are published for information purposes only and do not constitute a guaranty of product performance. In this section we discuss the limitations of the Challenge and the possible directions for improvement in future editions.

3.1 Limitations

A correct interpretation of the Challenge results should consider the inherent limitations of such an evaluation format. The first limitation derives from the relatively short time in which the candidate microsensors are available for evaluation. The length of the evaluation period represents a compromise between the desire for a time period that is as representative as possible for the evaluation site and the material constraints of the project (e.g., man-hours, instrumentation and consumables costs, etc.). This reasoning takes also into account the fact that, since candidate solutions are temporarily lent by the participants, the duration of their use for the Challenge needs to be kept within reasonable bounds.

A second inherent limitation lies in the choice of the location of the Challenge evaluation site which, in the case of outdoor measurements, plays a decisive role in constraining the dynamic range of the observable pollutant concentrations. To a certain extent this choice will also influence observable indoor concentration levels. This limitation implies that the performance scores of microsensor platforms when evaluated under the specific geographical, urban, and climatic conditions of Accra, of Bengaluru or of Paris might differ significantly when evaluated at locations with greatly divergent conditions.

For specific pollutants, like SO_2 and CO, the observable concentrations in the Parisian region are very low with a dynamic range that typically stays bounded within the uncertainty interval of the reference analyzer. Thus, for microsensor devices that target the monitoring of these pollutants in outdoor environments any meaningful evaluation is technically impossible.

The principle of simultaneously testing all candidate microsensors is fundamental to the AIRLAB Microsensors Challenge philosophy. This however requires a large degree of volume flexibility for the indoor evaluation site thus eliminating the possibility of performing controlled concentration testing, which are typically performed inside relatively small exposure chambers.

The use of an entire room for the indoor evaluation site implies safety concerns for particularly toxic pollutants, like CO. Microsensor platforms that target the monitoring of this pollutant indoors are not currently evaluated for this feature.

3.2 Outlook

While the intrinsic limitations which are dictated by the Challenge format cannot be completely eliminated, a number of measures can be envisioned to further improve the representativeness of the Challenge results. The temporal limitation could be tackled by increasing the evaluation period and/or considering evaluation intervals that capture seasonal effects.

The 2025 AIRLAB Microsensors Challenge has taken a significant step forward in enhancing the geographical representativity of its results compared to previous editions, building on the progress made in 2023. This edition includes two new evaluation deployments in Accra and Bengaluru, further expanding the diversity of measurement environments beyond metropolitan France. The geographical representativity could be improved further by incorporating additional types of evaluation sites. Considering monitoring sites with different typologies (e.g., rural, industrial) could enhance the dynamic range for specific pollutants (e.g., ozone, sulfur dioxide). Moreover, continued collaborations with regional associations and monitoring bodies will further strengthen the representativity of the Challenge's evaluations.

4 Bibliography

1. Polidori, A., Feenstra, B., Papapostolou, V., and Zhang, H., 2017. Field Evaluation of Low-cost Air Quality Sensors. *South Coast Air Quality Management District (SCAQMD)*, Diamond Bar, California, U.S.A. Accessed: 8 March 2019, [Online]. Available: http://www.aqmd.gov/docs/default-source/aq-spec/protocols/sensors-field-testing-protocol.pdf?sfvrsn=0

2. Lung, C., Jones, R., Zellweger, C., Karppinen, A., Penza, M., Dye, T., Hüglin, C., Ning, Z., Leigh, R., Hagan, D. et al., 2018. Low-Cost Sensors for the Measurement of Atmospheric Composition: Overview of Topic and Future Applications, *World Meteorological Organization (WMO)*, 2018, Geneva, Switzerland. Accessed: 8 March 2019, [Online]. Available: https://www.wmo.int/pages/prog/arep/gaw/documents/Low_cost_sensors_post_review_final.pdf

3. Fishbain, B., Lerner, U., Castell, N., Cole-Hunter, T., Popoola, O., Broday, D.M., Iñiguez, T.M., Nieuwenhuijsen, M., Jovasevic-Stojanovic, M., Topalovic, D. and Jones, R.L., 2017. An evaluation tool kit of air quality micro-sensing units. *Science of the Total Environment*, *575*, pp.639-648.

4. Schneider, P., Bartonova, A., Castell, N., Dauge, F.R., Gerboles, M., Hagler, G.S.W., Hüglin, C., Jones, R.L., Khan, S., Lewis, A.C., Mijling, B., Müller, M., Penza, M., Spinelle, L., Stacey, B., Vogt, M., Wesseling, J., and Williams, R.W., Toward a Unified Terminology of Processing Levels for Low-Cost Air-Quality Sensors, *Environmental Science & Technology* 2019 53 (15), 8485-8487