AiR'QUALITY SENSOR	**EVALUATION PROTOCOL FOR SENSOR SYSTEMS FOR AMBIENT AIR QUALITY MONITORING AT FIXED SITE**

TABLE OF CONTENTS

1 INTRODUCTION

In recent years, it was observed an increase on the market of sensor systems for air quality monitoring. However, there is currently no national or European normative framework for the evaluation of the performances of these different systems. However, this framework is under construction within the CEN/TC264/WG42 standardization working group.

This document is part of the "Air Quality Sensor" certification process and establishes the technical evaluation requirements for candidate sensor systems, as defined in the PR-1053 specific rules of certification. It describes a protocol that aims to evaluate the performances of sensor systems used to monitor the concentrations of gaseous and particulate pollutants in ambient air. For a given pollutant, the evaluation protocols described here are applied under the same conditions to the different types of sensor systems that will be subject to the certification process.

This evaluation protocol is divided in two independent parts:

- a first part related to the laboratory evaluation of metrological parameters of the sensor systems (under controlled conditions);
- a second part related to the performance evaluation of these sensor systems under outdoor conditions and with an enhanced ambient air matrix with gas or particulate matter.

The systems will be evaluated for an use in ambient air quality monitoring at fixed site, i.e. measurements taken outdoors using a stationary sensor system. For particulate matter (PM), as the type of aerosol may vary depending on the measurement site typology, this certification covers background site type. In a first step, this protocol concerns only the measurement of nitrogen dioxide (NO_2) and the $PM_{2.5}$ particulate fraction. This evaluation applies to "complete" sensor systems as defined in the PR-1053 specific rules of certification, i.e. "commercially available ready-to-use products" including: sensitive element (sensor), data processing algorithms, associated electronics (signal processing, acquisition) and usage requirements (instruction manual, wiring diagram, etc.). Thus, they will be evaluated as a "ready-to-use system" which is intended to provide directly usable measurements of air pollutant concentrations. The provision of a minimum of 3 identical sensor systems (so-called replicas) is required in order to be able to assess the reproducibility of manufacturing processes for the systems.

These replicas should be identical in every aspect, i.e.:

- identical measuring principles and technology used;
- control electronics of measurement component from the same generation;
- the same version of internal software;
- same measurement time interval, ideally one measurement per minute;
- if used, same version of data correction algorithm;
- if required, same version of the data recovery platform.

2 ACRONYMS

APS	Aerosol Particle Sizer
DQO	Data Quality Objectives defined by Directive 2008/50/EC
FDMS	Filter Dynamics Measurement Systems
INERIS	French National Institute for Industrial Environment and Risks (Institut national de l'environnement industriel et des risques)
LCSQA	French national reference laboratory for monitoring air quality (Laboratoire Central de Surveillance de la Qualité de l'Air)
LD	Detection limit
LNE	National Metrology and Testing Laboratory (Laboratoire national de métrologie et d'essais)
LV	Limit value defined by the Directive 2008/50/EC
MAPE	Mean Absolute Percentage Error
PE	Span point or *Point échelle*
RSS	Residual Sum of Squares
SMPS	Scanning Mobility Particle Sizer
TEOM	Tapered Element Oscillating Microbalance

3 CRITERIA FOR ASSESSING AND ASSIGNING PERFORMANCE RATINGS

Performance evaluation will be conducted on each replica individually and then the scores obtained will be grouped together (see 3.3) to provide the final rating. By its definition, reproducibility will be assessed directly on all replicas.

In a similar way to the approach adopted by the standardization working group CEN/TC264/WG42 " Air quality — Performance evaluation of air quality sensors", the first two ratings are compatible with the measurement requirements specified in Directive 2008/50/EC concerning ambient air quality and clean air for Europe[1] in accordance with the following correspondence:

Division A	Category of data quality objectives defined in **this evaluation protocol** and compliant with the data quality objectives (uncertainty, minimum data capture) of **Indicative Measurement** as described in the Directive 2008/50/EC.
Division B	Category of data quality objectives defined in **this evaluation protocol** and compliant with the data quality objectives (uncertainty, minimum data capture) of **Objective Estimation** as described in the Directive 2008/50/EC.
Division C	Category of data quality objectives defined in **this evaluation protocol** but that are out of the scope of the Directive 2008/50/EC. For this division, the level of requirements on terms of uncertainty is only sufficient for citizen science studies, educational action, etc., defined as **Awareness Studies.**

The following paragraphs describe the evaluation criteria to be achieved based on the associated rating and the proposed overall scoring system.

[1] Directive 2008/50/EC of the European Parliament and the Council of 21 May 2008 on Ambient Air Quality and Cleaner Air for Europe, https://eur-lex.europa.eu/eli/dir/2008/50/2015-09-18.

3.1 Evaluation criteria for NO_2 measurement

		NO_2 Division A	Division B	Division C
LABORATORY	**Accuracy (slope)**	$0.7 \leq p \leq 1.3$	$0.5 \leq p < 0.7$ or $1.3 < p \leq 1.5$	$p < 0.5$ or $p > 1.5$
	Linearity (from 0 to 300 µg/m³)	$R^2 \geq 0.75$	$0.5 \leq R^2 < 0.75$	$R^2 < 0.5$
	Detection limit	$LD \leq 19$ µg/m³	19 µg/m³ < LD ≤ 29 µg/m³	LD > 29 µg/m³
	Repeatability at 200 µg/m³	r ≤ 7.6 µg/m³	7.6 µg/m³ < r ≤ 11.5 µg/m³	r > 11.5 µg/m³
	Influence of relative humidity (15% and 80%) at 200 µg/m³	*Deviation* ≤ 20 µg/m³	20 µg/m³ < *Deviation* ≤ 40 µg/m³	*Deviation* > 40 µg/m³
	Influence of ozone at 200 µg/m³	*Deviation* ≤ 20 µg/m³	20 µg/m³ < *Deviation* ≤ 40 µg/m³	*Deviation* > 40 µg/m³
	Drift at zero within 3 weeks	$d_{zero} \leq 20$ µg/m³	20 µg/m³ < $d_{zero} \leq 30$ µg/m³	$d_{zero} > 30$ µg/m³
	Drift at span (PE) within 3 weeks at 200 µg/m³	$d_{PE} \leq 10$ %	10% < $d_{PE} \leq 15$ %	$d_{PE} > 15$ %
FIELD	**Reproducibility** (u(bs,s))	u(bs,s) < 7.6 µg/m³	u(bs,s) < 15 µg/m³	u(bs,s) < 31 µg/m³
	Slope	$0.7 \leq p \leq 1.3$	$0.5 \leq p < 0.7$ or $1.3 < p \leq 1.5$	$p < 0.5$ or $p > 1.5$
	Linearity	$R^2 \geq 0.75$	$0.5 \leq R^2 < 0.75$	$R^2 < 0.5$
	MAPE	< 50%	from 50% to 100%	> 100%
DIR 2008/50/EC	**Minimum data capture**	≥ 90 %	from 14% to 90%	< 14%
	Field uncertainty (DQO@ 200µg/m³)	U ≤ 25 % (U ≤ 50µg/m³)	25% < U ≤ 75 % (50 < U ≤ 150µg/m³)	75% < U ≤ 200% (150 < U ≤ 400µg/m³)

3.2 Evaluation criteria for $PM_{2.5}$ measurement

		$PM_{2.5}$		
		Division A	**Division B**	**Division C**
LABORATORY	**Accuracy (slope)**	$0.7 \leq p \leq 1.3$	$0.5 \leq p < 0.7$ or $1.3 < p \leq 1.5$	$p < 0.5$ or $p > 1.5$
	Linearity (from 0 to 120 µg/m³)	$R^2 \geq 0.75$	$0.5 \leq R^2 < 0.75$	$R^2 < 0.5$
	Detection limit	LD ≤ 5 µg/m³	5 µg/m³ < LD ≤ 10 µg/m³	LD > 10 µg/m³
	Repeatability at 80 µg/m³	r ≤ 5 µg/m³	5 µg/m³ < r ≤ 10 µg/m³	r > 10 µg/m³
	Influence of relative humidity (15% and 80%) at 80 µg/m³	*Deviation* ≤ 10 µg/m³	10 µg/m³ < *Deviation* ≤ 15 µg/m³	*Deviation* > 15 µg/m³
	Drift at zero within 3 weeks	d_{zero} ≤ 5 µg/m³	5 µg/m³ < d_{zero} ≤ 10 µg/m³	d_{zero} > 10 µg/m³
	Drift at span (PE) within 3 weeks at 80 µg/m³	d_{PE} ≤ 10 %	10% < d_{PE} ≤ 15 %	d_{PE} > 15 %
FIELD	**Reproducibility** (u(bs,s))	u(bs,s) < 7.5 µg/m³	u(bs,s) < 15 µg/m³	u(bs,s) < 30 µg/m³
	Slope	$0.7 \leq p \leq 1.3$	$0.5 \leq p < 0.7$ or $1.3 < p \leq 1.5$	$p < 0.5$ or $p > 1.5$
	Linearity	$R^2 \geq 0.75$	$0.5 \leq R^2 < 0.75$	$R^2 < 0.5$
	MAPE	< 50%	from 50% to 100%	> 100%
DIR 2008/50/EC	**Minimum data capture**	≥ 90 %	from 14% to 90%	< 14%
	Field uncertainty (DQO@ 50µg/m³)	U ≤ 50% (U ≤ 25µg/m³)	50 < U ≤ 100% (25 < U ≤ 50µg/m³)	100 < U ≤ 200% (50 < U ≤ 100µg/m³)

3.3 Rule for assigning the performance division

To assign the performance division, it is proposed to evaluate each replica individually according to 3 assessment aspects that gather all of the criteria referred to above: laboratory evaluation, field evaluation (with ambient air matrix), and a combined assessment of the measurement uncertainty and minimum data capture which are the two criteria described in the Directive 2008/50/EC. All the scores obtained for each sensor and each criterion are concatenated according to the following process to result in assigning a division (A, B or C) and summarised in the Figure below:

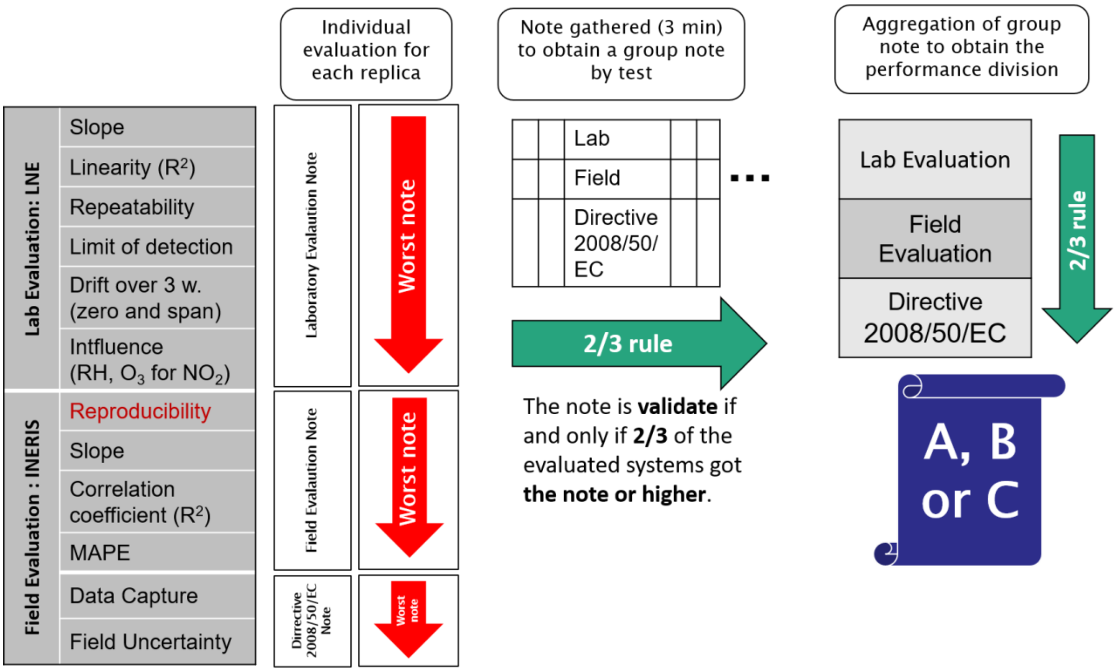

Figure 1: Rules for assigning the performance division

Phase 1: Reproducibility evaluation of the systems: As previously indicated, the reproducibility is evaluated for all the replicas as a single score resulting in an A, B, or C or criteria not met. If the score is 'criteria not met', the difference in reproducibility is considered too important for a certificate to be issued. The certification process stops at this stage. If the score is A, B or C, this score is considered to be the reproducibility criterion for phase 2 for each replica;

Phase 2: Individual replica results evaluation: each **replica** is individually evaluated against all of the criteria presented in this document. Each result is assigned a **score** (A, B or C) based on the criteria tables above;

Phase 3 ("worst note" on Figure 1)**: Grouping of scores by assessment cluster** (laboratory, field and "Directive 2008/50/EC") **for each of the replicas:** the scores obtained (A, B or C) by criteria in the previous phase are grouped in order to obtain a **"laboratory" score**, a **"field" score** and a **"Directive 2008/50/EC" score for each replica**. The **grouping** of the scores is done by **the lowest score achieved rule.** This phase makes it possible to create a trio of individual scores for each replica;

Phase 4 ("2/3 rule ➲" on Figure 1)**: Grouping of scores of all the replicas for each assessment cluster: for each assessment cluster** (laboratory, field and "Directive 2008/50/EC"), the **scores** obtained **for each replica** are **aggregated in line with the 2 out of 3 rule**, i.e. a minimum of 2 out of 3 of the scores must be identical (or higher) to justify its assignment (for example: A, B, C ➔ B because an A is at least a B);

Phase 5 ("2/3 rule ⮋" on Figure 1)**: Grouping the three overall scores by assessment cluster to obtain the performance division:** the previous phase made it possible to generate **a set of 3 scores** that are then **combined** using the same **2 out of 3 rule** to give the performance division. However, if the score obtained at this stage is higher than the **"Directive 2008/50/EC (uncertainty and minimum data capture)"** score, the performance division will be **lowered** to the level of the **"Directive 2008/50/EC (uncertainty and minimum data capture)"** score.

4 LABORATORY EVALUATION OF THE SENSOR SYSTEMS

This part describes the operating procedure used by LNE to evaluate the metrological parameters of sensor systems in laboratory (7 weeks per sensor system). It is designed based on the following studies:

- Protocole de détermination des caractéristiques de performance métrologique des micro-capteurs, (Protocol to define the metrological performance characteristics of microsensors) N. Redon, F. Delcourt, S. Crunaire, N. Locoge, LCSQA/IMT Lille Douai report, March 2017

- Mise en œuvre d'un protocole pour l'évaluation en laboratoire de micro-capteurs pour la mesure des concentrations massiques PM, (Implementation of a protocol for laboratory assessment of microsensors for measuring mass concentrations) N. Redon, S. Crunaire, B. Herbin, E. Morelle, F. Gaie-Levrel, T. Amodéo, LCSQA report, July 2018

- Développement d'un protocole pour l'évaluation en laboratoire des capteurs de PM, (Development of a protocol for the laboratory assessment of PM sensors) F. Gaie-Levrel, L. Brégonzio-Rozier, A. Bescond, T. Macé, LCSQA/LNE report, December 2019

4.1 Test bench schematic

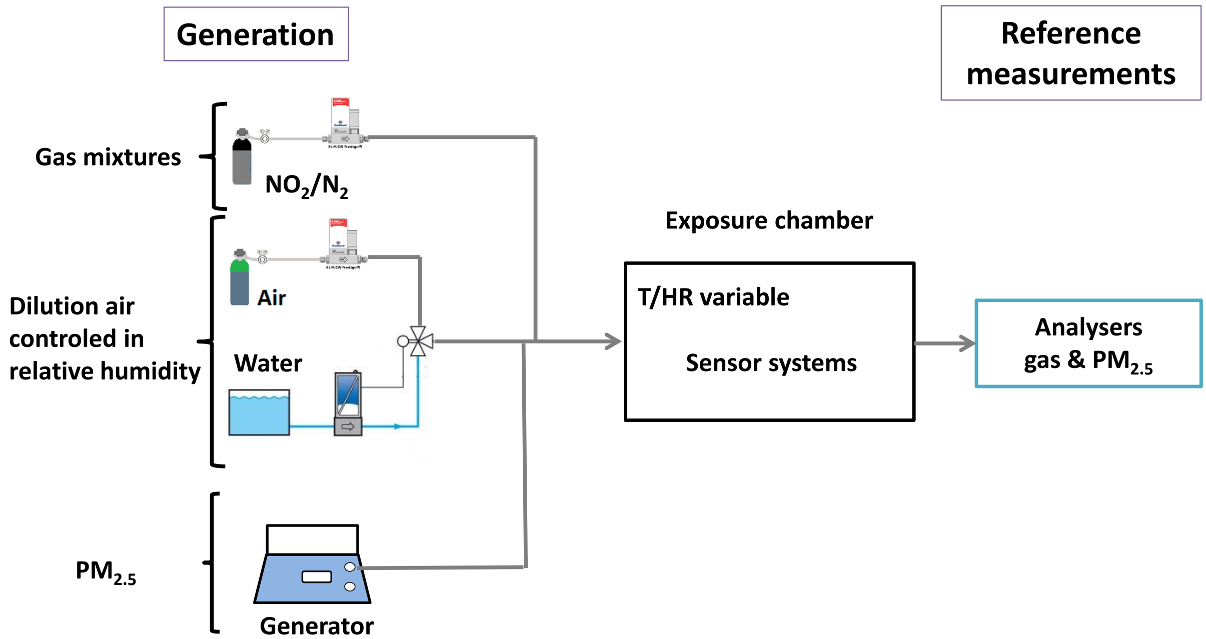

Figure 2: Schematic of the test bench for the metrological parameters evaluation of the sensor systems in laboratory

4.2 Equipment description

4.2.1 Exposure chamber

The exposure chamber consists in a stainless-steel enclosure which has been subject to prior stabilisation and is equipped with:

- injection modules for the introduction of gases and particulate matter,
- sampling modules in order to connect the different instruments.

Specifications:

Temperature	20.0°C ± 0.5°C
Relative humidity	10% to 85% (±2%)
Pressure	Ambient pressure

4.2.2 Instrumentation associated with generation

4.2.2.1 Dilution air controlled for relative humidity

Different relative humidity levels are generated in the exposure chamber using a quantity of water regulated using a liquid mass flow controller (MFC) and vaporised in an air stream.

The tests are carried out at 3 relative humidity levels:

- 15 % RH (±2%)
- 50 % RH (±2%)
- 80 % RH (±2%)

4.2.2.2 Gas mixtures (NO_2)

A nitrogen dioxide (NO_2) gas mixture is generated by dynamic dilution of a NO_2/N_2 gas mixture in a cylinder with dilution air controlled for relative humidity. Gas flows are controlled using calibrated MFCs.

The NO_2 mass concentrations of the gas mixtures generated are between 0 and 300 µg/m^3.

4.2.2.3 Particulate matter ($PM_{2.5}$)

The generation of particulate matter ($PM_{2.5}$) is carried out by dry or wet process, which supplies an aerosolization nozzle with small quantities of sample.

The $PM_{2.5}$ mass concentrations of the aerosols generated range from 0 to 120 µg/m^3.

4.2.3 Instrumentation associated with reference measurements

4.2.3.1 Monitoring of exposure conditions

A calibrated thermo-hygrometer is used to measure the temperature and relative humidity in the exposure chamber.

4.2.3.2 Measurement of gas concentrations (NO_2)

A chemiluminescence-based analyser (compliant with NF EN 14211) is used to perform NO_2 measurements in the exposure chamber. The instrument is calibrated using gas measurement standards generated by dynamic dilution of gas mixtures of NO/NOx with a higher molar fraction.

4.2.3.3 Measurement of particulate matter concentrations ($PM_{2.5}$)

The granulometric distribution of the particles generated in the exposure chamber is measured using two instruments, the Scanning Mobility Particle Sizer (SMPS) and the Aerosol Particle Sizer (APS). The SMPS allows to measure the particle number distribution in a range of electrical mobility diameters from 10 nm to 1 µm, while the APS allows to define the particulate concentration in number based on the aerodynamic diameter for a size range between 0.6 and 20 µm. Both instruments are size-calibrated with reference suspensions (e.g. polystyrene spheres).

Particulate mass concentrations are measured using a TEOM 50°C microbalance with good temporal resolution (30s) and traceable to the international system via a working standard (particle generator).

4.3 Test descriptions

Both the installation and start-up of the devices will be carried out in compliance with the operating instructions of the manufacturer/supplier of the sensor system. The manufacturer/supplier will be able to come on-site to give recommendations for its implementation. They should ensure that the system is functioning correctly and that data acquisition is operational (for the compound NO_2, the values should be expressed in µg/m^3, ppb or µmol/mol; for $PM_{2.5}$, the values should be expressed in µg/m^3). The manufacturer/supplier will also have to train LNE staff to use the data acquisition or agree a method to ensure the correct functioning of the sensor system being assessed for the testing period. The manufacturer/supplier should also ensure that a contact person is available, when necessary, throughout the evaluation.

4.3.1 Tests for accuracy, linearity and detection limit

The sensor systems are placed in the exposure chamber.

These 3 performance characteristics are evaluated by generating an increasing mass concentration of NO_2 and $PM_{2.5}$ at 20°C and 50% of relative humidity in the exposure chamber.

- NO_2: up to a mass concentration of 300 µg/m^3;
- $PM_{2.5}$: up to a mass concentration of 120 µg/m^3.

The responses provided by the sensor systems are used during tests for increasing and decreasing concentrations.

Individual responses from each sensor system and the reference values are continuously measured.

4.3.2 Repeatability tests

The evaluation of this performance characteristic is carried out at the concentration level (NO_2: 200 µg/m^3 and $PM_{2.5}$: 80 µg/m^3), at 20°C and 50% of relative humidity in the exposure chamber.

When a stable concentration plateau has been reached, the individual responses of each sensor system are measured for one hour (at least 10 points).

4.3.3 Tests on the influence of relative humidity

These influence tests are carried out at a span point (NO_2: 200 µg/m^3 and $PM_{2.5}$: 80 µg/m^3), at 20°C and at three levels of relative humidity in the exposure chamber:

- 15 %HR (±2%)
- 50 % RH (±2%)
- 80 % RH (±2%)

For each test, the individual responses of each sensor system are measured and for each sensor, the worst-case value will be taken into account.

4.3.4 Tests on the effect of ozone on the responses of the NO_2 sensor systems.

The sensor systems are placed in the exposure chamber.

This interference test is conducted at an NO_2 concentration of 200 µg/m^3 with and without an ozone concentration of 200 µg/m^3, at 20°C and 50% of relative humidity in the exposure chamber.

For each test, the individual responses of each sensor system are measured with and without ozone.

4.3.5 Drift within 3 weeks

The drift within 3 weeks of the sensor systems is evaluated 3 weeks after the start of the linearity testing (t_0). During these 3 weeks, the sensor systems undergo all the tests described in paragraphs 4.3.1 to 4.3.4.

The evaluation of this performance characteristic is carried out at zero and at a span point (NO_2: 200 µg/m^3 and $PM_{2.5}$: 80 µg/m^3), at 20°C and 50% of relative humidity.

For each test, the individual responses of each sensor system are measured.

4.4 Description of data processing for the evaluation of the parameters

4.4.1 Accuracy

The accuracy of the sensor system is evaluated by calculating the slope (*b*) and the intercept (*a*) of the regression line between the responses of the sensor systems and the reference values.

They are calculated according to the following formulae:

$$b = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sum(x_i - \bar{x})^2}$$

$$a = \bar{y} - b\bar{x}$$

Where: x_i the reference value i

$\bar{x}$ the mean reference value

y_i the i response from the sensor system

$\bar{y}$ the mean sensor system response

The value obtained for *b* is compared to the criteria in § 3.1 or §3.2.

4.4.2 Linearity

The linearity of each sensor system is evaluated by calculating the regression coefficient R^2 of the linear regression line between the sensor system responses and the reference values.

It is calculated using the following formula:

$$R^2 = \frac{(\sum(x_i - \bar{x})(y_i - \bar{y}))^2}{\sum(x_i - \bar{x})^2 \sum(y_i - \bar{y})^2}$$

Where: x_i the reference value i

$\bar{x}$ the mean reference value

y_i the i response from the sensor system

$\bar{y}$ the mean sensor system response

The value obtained for R^2 is compared to the criteria in § 3.1 or §3.2.

4.4.3 Repeatability

Repeatability (*r*) is evaluated by calculating standard deviation for the responses of the sensor systems obtained.

$$r = \sqrt{\frac{\sum(y_i - \bar{y})^2}{n}}$$

Where: y_i the i response from the sensor system

$\bar{y}$ the mean sensor system i response

The value obtained for *r* is compared to the criteria in § 3.1 or §3.2.

4.4.4 Detection limit (LD)

The detection limit of the sensor system is calculated according to the following formula:

$$LD = \frac{|a| + 3u_b}{|b|}$$

Where: $|a|$ the absolute y-intercept

u_b standard uncertainty for the slope

b regression line slope

The value obtained for *LD* is compared to the criteria in § 3.1 or §3.2.

4.4.5 Influence of relative humidity

The deviation determining the influence of 15% or 80% of relative humidity on the sensor system responses is calculated according to the following formula:

RH deviation (15% or 80%) $= \bar{y}\ (15\%RH\ ou\ 80\%RH) - \bar{y}\ (50\%RH)$

Where: $\bar{y}\ (15\%RH\ ou\ 80\%RH)$ the mean sensor system response obtained at 15% or 80% of relative humidity

$\bar{y}\ (50\%RH)$ the mean sensor system response obtained at 50% of relative humidity

The value obtained for this *deviation* is compared to the criteria in § 3.1 or §3.2.

4.4.6 Tests on the effect of ozone on the responses of the NO_2 sensor systems.

The deviation determining the ozone influence on NO_2 sensor system responses is calculated according to the following formula:

Ozone influence deviation= $\left(\bar{y}\ (with\ ozone) - \bar{y}\ (without\ ozone)\right)$

Where: $\bar{y}\ (with\ ozone)$ the mean sensor system response obtained with ozone

$\bar{y}\ (without\ ozone)$ the mean sensor system response obtained without ozone

The value obtained for this *deviation* is compared to the criteria in § 3.1 or §3.2.

4.4.7 Drift within 3 weeks

Zero drift within 3 weeks is calculated using the following formula:

$$D_{zero} = \left(\bar{y}\,(t_0 + 3\,weeks) - \bar{y}\,(t_0)\right)$$

Where: $\bar{y}\,(t_0 + 3\,weeks)$ the mean sensor system response obtained under testing at t_0 + 3 weeks at zero

$\bar{y}\,(t_0)$ the mean sensor system response obtained under linearity testing at t_0 at zero

Span point drift within 3 weeks is calculated using the following formula:

$$D_{PE} = \frac{\bar{y}\,(t_0+3\,weeks)-\bar{y}\,(t_0)}{200\,(NO_2)\;or\;80\,(PM_{2,5})} \times 100$$

Where: $\bar{y}\,(t_0 + 3\,weeks)$ the mean sensor system response obtained under testing at t_0 + 3 weeks at span point

$\bar{y}\,(t_0)$ the mean sensor system response obtained under linearity testing at t_0 at span point

The value obtained for D_{zero} and D_{PE} is compared to the criteria in § 3.1 or §3.2.

4.5 Conclusion

At the end of these tests, **a "laboratory" score is established in accordance with the rules set out in paragraph 3.3.**

5 REAL AND ENHANCED AIR MATRIX EVALUATION OF THE SENSOR SYSTEMS

This chapter describes the protocol that is implemented when evaluating the sensor systems under real atmospheric conditions in association with an enhanced air matrix system. This evaluation is carried out for all replicas over a period of a minimum of 4 weeks up to a maximum of 6 weeks.

5.1 Evaluation system schematic

The system used (Figure 3) allows for the homogeneous and simultaneous exposure of several sensor systems and reference devices to a mixture of external air enriched with artificially generated particulate matter (INERIS Report DRC-16-152318-06089A, Development of a PM doping system for inter-laboratory comparisons of PM automatic analysers, June 2016), by nebulising a mixture of diluted salts. Gas mixture generation is ensured by a dynamic dilution from a high concentration gas cylinder.

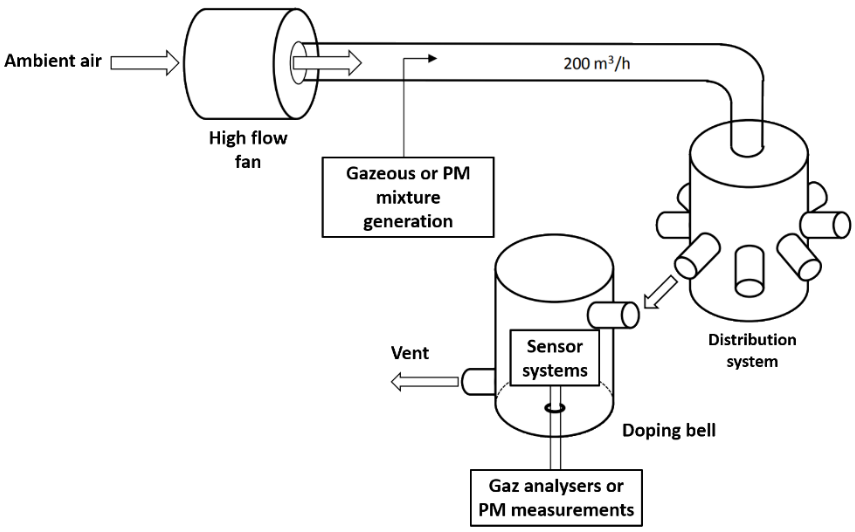

Figure 3: Schematic of the system used in evaluating the sensor systems with an enhanced air matrix.

In the initial phase, this system includes a high-volume fan with adjustable speeds which channels outside air into a circuit consisting of a distribution system that evenly distributes the enhanced air into eight cylinders (Figure 4). Of these eight cylinders, three are dedicated to reference measurement devices, leaving five cylinders available to install sensor systems in them.

Figure 4: Test trailer (left) and doping cylinder (right) used in inter-laboratory comparison exercises of PM automated analysers.

5.2 Equipment description

Pollutant measurements are carried out using so called reference analysers and compliant with the requirements of European Directives 2008/50/EC and 2015/14/80:

- Nitrogen dioxide NO_2: the nitrogen dioxide will be measured using a chemiluminescence analyser (according to NF EN 14211);
- $PM_{2.5}$ particulate matter: the controls of concentration will be carried out using two techniques. A TEOM FDMS 1400AB microbalance + 8500C MODULE is used to provide the hourly value for mass concentration using a method that is equivalent to the reference method (NF EN 12341). However, because the measurement time interval does not allow for providing dynamic monitoring of particulate matter concentrations, a FIDAS 200S (automatic analyser equivalent to the reference method) will also be used to provide a response with a time interval similar to that of the PM sensors. As for the laboratory tests, a particle size dispersion checks will be carried out using an APS and an SMPS before the PM doping phase;

Each gas and PM analyser will be calibrated before the evaluation in compliance with current best practice.

5.3 Description of tests with enhanced air matrix

This test procedure will be performed in 5 successive phases:

- **Preliminary phase - getting started/installation:**

Both the installation and the start-up of the devices will be carried out in compliance with advice from the sensor system's manufacturer/supplier. The manufacturer/supplier may allow Ineris staff to perform the installation in accordance with the documents provided when the certification file was created: for example, the user manual or the installation manual. They will also have the possibility to come on-site to install the system and give recommendations for use if the sensor systems require specific skills. In the first case, an inspection by the manufacturer/supplier will be organised before the beginning of the tests to ensure the installation compliance and validate it.

In any case, the manufacturer/supplier will need to ensure that the sensor systems are functioning correctly and that data acquisition is active. It will also be needed to train the Ineris staff to use the data acquisition or agree a method to ensure the correct functioning of the sensor systems being assessed during the testing period. The manufacturer/supplier should also ensure that a contact person is available, where necessary, throughout the assessment.

- **Phase 1 – Sensor system adaptation:**

Upon request by the manufacturer/supplier, the sensor systems may be installed and then left in the outside air for a maximum of 2 weeks in order to give the sensor systems the time to adapt to the environmental conditions of the site. A mandatory 48 hr minimum shall in all cases be observed prior to the start of assessment testing in order to ensure the best worjing condition of the sensor systems, or to allow the sensor components to function optimally. This phase is carried out with the final assembly (cylinder mounting base only, cylinders cap removed). When installed, the manufacturer/supplier should ensure that the mounting devices do not cause any malfunctions in the systems being evaluated. This period will also be used to check that the data acquisition system is operating correctly.

- **Phase 2 – Evaluation in the outside air:**

During this stage, the sensor systems are left in the outdoor air for two weeks, installed on the cylinder supports. Reference values are registered to ensure continuous performance monitoring of the sensor systems being evaluated.

- **Phase 3 – Evaluation of the sensor systems with gaseous enhanced air matrix:**

Two series of 4 levels (only outdoor air without doping + 3 concentrations) are performed. Tests with ozone, a major gas interferent for NO_2 sensor systems, will also be carried out in order to evaluate the sensitivity of the sensor system in the presence of the interferent alone and the mixture of interferent and relevant pollutant, this sensitivity will be taken into account in the uncertainty calculation.

Concentration range for added nitrogen dioxide NO_2: 40-150 µg/m^3.

Phase 4 – Evaluation of the sensor systems with $PM_{2.5}$ enhanced air matrix

A series of four levels of particulate matter concentrations (only outdoor air without doping + 3 concentrations) are carried out.

Concentration range: $PM_{2.5}$: 50-150 µg/m^3.

5.4 Description of parameter evaluation methods

As described in the introduction, the aim of this evaluation is carried out to characterise the response the sensor system in the context of fixed-site outdoor air quality monitoring. This means that different metrological criteria are assessed.

5.4.1 Reproducibility coefficient

The measurement reproducibility of the sensor systems is calculated using the entire field campaign including the phases of gas and particulate matter enhanced air matrix by comparison between each replica of the sensor system. This data dispersion shows the difference in reproducibility between different replicas of the same system. Reproducibility is the only evaluation parameter that is considered to be eliminatory.

This variation is calculated based on the uncertainty between replicas ("mean deviation") using the following formula across all replicas:

$$\boldsymbol{u(bs,s)} = \left(\frac{\sum_{i=1}^{n}\sum_{j=1}^{p}\left(y_{i,j}-y_m\right)^2}{n(p-1)}\right)^{\frac{1}{2}}$$

where

$y_{i,j}$	data of sensor system *j* for period *i*;
y_m	mean sensor system data for period *i*;
n	amount of data;
p	number of replicas.

The value obtained is compared to the criteria in § 3.1 or § 3.2.

5.4.2 Minimum data capture

The **minimum data capture** for each sensor system compared to the respective reference methods is calculated in accordance with B. Fishbain et al.[2] "An evaluation tool kit of air quality micro-sensing units". The calculation proposed by B. Fishbain et al. is consistent with the requirements of Directive 2008/50/EC[3] as well as with those described in the *Methodological Guide for the Calculation of Air Quality Statistics*[4].

Apart from the system's ability to provide continuous information, minimum data capture is used to assess the quality of the data sending and sharing system. In the best-case scenario, i.e. when there is a level of 100%, the sensor system has not been disconnected or suffered loss of data and there is enough data available for use.

[2] Science of the Total Environment 575 (2017) 639-648

[3] Directive 2008/50/EC of the European Parliament and the Council of 21 May 2008 on Ambient Air Quality and Cleaner Air for Europe, https://eur-lex.europa.eu/eli/dir/2008/50/2015-09-18

[4] Guide méthodologique pour le calcul des statistiques relatives à la qualité de l'air, (Methodological guide for calculating statistics relating to air quality) L. Malherbe, M. Beauchamp, Rapport LCSQA/Ineris, June 2016

The value obtained is compared to the criteria in § 3.1 or § 3.2.

5.4.3 Accuracy and linearity in real conditions

These criteria are evaluated by plotting the concentrations measured by the sensor systems agaisnt the concentrations measured by the reference instruments. This graph is then used to plot the correlation curve using a linear regression as a correlation model.

Thus, **the slope value** characterises the sensor system's capacity to produce accurate measurements in relation to the reference instrument. If both measurements are the same, the ideal slope will be equal to 1.

R^2, also called the **coefficient of determination**, defines the sensor's ability to produce measurements that are consistent with reference measurements. This characteristic is called "measurement dispersion" and its ideal value of 1 means that the sensor and the reference method are in perfect synchronicity.

The value obtained for R^2 is compared with the criteria of § 3.1 or §3.2.

5.4.4 MAPE

MAPE (the Mean Absolute Percentage Error) is the mean absolute value deviation between measurements modelled according to the linear model of chapter 5.4.3 and the reference measurements relating to reference values and is calculated in the following way:

$$MAPE = \frac{1}{N}\sum_{k=1}^{N} \frac{|reference_k - modeled_k|}{reference_k}$$

It shows error dispersion once correction is made according to the linear regression model.

The value obtained is compared to the criteria in § 3.1 or § 3.2.

5.4.5 Calculation of uncertainty at the reference value from "field" data

An overall uncertainty integrating all the factors influencing sensor measurement during the field test period (ambient air + doping) is also calculated according to the following formula.

$$\boldsymbol{U(C_i)} = \mathbf{k}\sqrt{\frac{\mathbf{RSS}}{(\boldsymbol{n}-\mathbf{2})} - \mathbf{u}^2_{\boldsymbol{bs,RM}} + [\mathbf{a} + (\mathbf{b}-\mathbf{1}) \times \mathrm{RV}]^2}$$

where

$U(C_i)$	extended uncertainty of the measurement system C_i expressed in µg/m³;
RSS	residual sum of squares resulting from the linear regression;
$u_{bs,RM}$	uncertainty of reproducibility of the reference method, in µg/m³;
a	value of the intercept
b	linear regression slope
RV	reference value
n	amount of data over the time period
k	coverage factor reflecting the number of degrees of freedom. Given the large number of experimental results available, a coverage factor of 2 can be used

The reference value is taken equal to the hourly limit value of 200µg/m³ for NO_2, as described in Directive 2008/50/EC, and equal to 50µg/m³ for $PM_{2.5}$, which corresponds to the daily PM_{10} limit value because there is no daily equivalence for $PM_{2.5}$.

The residual sum of squares resulting from linear regression is calculated based on the formula:

$$RSS = \sum_{i=1}^{n} (y_i - a - bx_i)^2$$

where

x_i data from the reference method in µg/m³

y_i sensor system data in µg/m³

The uncertainty value obtained is compared to the criteria in § 3.1 or § 3.2.

5.5 Post-assessment correction of the gradient and offset at the origin

As part of the current CEN/TC264/WG42 normative work, it is possible to apply a **correction** of slope and/or intercept in order to reduce the uncertainty of field measurement **in cases of systematic error aver all of the replicas.** These corrections are dependent on the environment where the system is being used, therefore if **different correction factors are necessary** for the different replicas tested as part of these tests, the system will be **evaluated based on the uncorrected data.**

In the event of a correction, **it remains essential for this to be identical for all the sensor systems over the entire dataset.** In addition, if a correction of the slope and/or intercept is made in order to reduce the measurement variation, the system must then be able to apply such a correction and users must be informed of the obligation to apply these corrections in order to maintain the division achieved in the course of this evaluation. The evaluation report will then make reference to the two scores obtained before and after correction.

Depending on the type of correction applied (slope only, intercept only, or slope and intercept), additional terms are added to the variation calculation formula as shown below:

- correction of intercept:

$$U(C_i, \mathrm{corr}) = k\sqrt{\left(\frac{\mathrm{RSS}}{(n-2)} - u_{bs,RM}^2\right) + [c + (d-1) \cdot \cdot RV]^2 + u_a^2}$$

- correction of slope:

$$U(C_i, \mathrm{corr}) = k\sqrt{\left(\frac{\mathrm{RSS}}{(n-2)} - u_{bs,RM}^2\right) + [c + (d-1) \cdot \cdot RV]^2 + \left(RV^2 . u_b^2\right)}$$

- correction of slope and intercept:

$$U(C_i, \mathrm{corr}) = k\sqrt{\left(\frac{\mathrm{RSS}}{(n-2)} - u_{bs,RM}^2\right) + [c + (d-1) \cdot \cdot RV]^2 + u_a^2 + \left(RV^2 . u_b^2\right)}$$

where

$U(C_i, \mathrm{corr})$ extended uncertainty of the measurement system C_i expressed in µg/m³

u_a^2 square of the uncertainty of the intercept before correction

u_b^2 square of the slope uncertainty before correction

c value of the intercept after correction

d linear regression slope after correction