



(ID Modèle = 454913)

Ineris - 213878 - 2779150 - v1.0

16/11/2023

**ESSAIS IN-SITU D’EVALUATION DE
PERFORMANCES DE SYSTEMES CAPTEURS
DEDIES A LA MESURE DES PARTICULES EN
ENCEINTES FERROVIAIRES SOUTERRAINES
(EFS)**

PRÉAMBULE

Le présent document a été réalisé au titre de la mission d'appui aux pouvoirs publics confiée à l'Ineris, en vertu des dispositions de l'article R131-36 du Code de l'environnement.

La responsabilité de l'Ineris ne peut pas être engagée, directement ou indirectement, du fait d'inexactitudes, d'omissions ou d'erreurs ou tous faits équivalents relatifs aux informations utilisées.

L'exactitude de ce document doit être appréciée en fonction des connaissances disponibles et objectives et, le cas échéant, de la réglementation en vigueur à la date d'établissement du document. Par conséquent, l'Ineris ne peut pas être tenu responsable en raison de l'évolution de ces éléments postérieurement à cette date. La mission ne comporte aucune obligation pour l'Ineris d'actualiser ce document après cette date.

Au vu de ses missions qui lui incombent, l'Ineris, n'est pas décideur. Les avis, recommandations, préconisations ou équivalents qui seraient proposés par l'Ineris dans le cadre des missions qui lui sont confiées, ont uniquement pour objectif de conseiller le décideur dans sa prise de décision. Par conséquent, la responsabilité de l'Ineris ne peut pas se substituer à celle du décideur qui est donc notamment seul responsable des interprétations qu'il pourrait réaliser sur la base de ce document. Tout destinataire du document utilisera les résultats qui y sont inclus intégralement ou sinon de manière objective. L'utilisation du document sous forme d'extraits ou de notes de synthèse s'effectuera également sous la seule et entière responsabilité de ce destinataire. Il en est de même pour toute autre modification qui y serait apportée. L'Ineris dégage également toute responsabilité pour chaque utilisation du document en dehors de l'objet de la mission.

Nom de la Direction en charge du rapport : DIRECTION SITES ET TERRITOIRES

Rédaction : BERTHELOT Brice -

Vérification : QUERON JESSICA; SPINELLE LAURENT; AMARA ANIS

Approbation : Document approuvé le 16/11/2023 par MORIN ANNE

Liste des personnes ayant participé à l'étude : FABLE Sébastien, RICHET Christophe

Table des matières

| | | |
|-------|---|----|
| 1 | Contexte et objectifs | 7 |
| 2 | Matériel & Méthodes..... | 8 |
| 2.1 | Terminologie | 8 |
| 2.2 | Sélection des modèles de systèmes capteurs de particules dits « candidats » | 8 |
| 2.3 | Méthodes prises pour « référence »..... | 9 |
| 2.4 | Réseaux ferroviaires investigués..... | 10 |
| 2.5 | Méthodologie pour le traitement des données | 10 |
| 2.5.1 | Métriques d'évaluation de la performance des méthodes dites « candidates » | 11 |
| 2.5.2 | Valeurs de comparaison pour les métriques de performance..... | 11 |
| 3 | Résultats..... | 13 |
| 3.1 | Déroulement des campagnes..... | 13 |
| 3.1.1 | Temps de recouvrement des campagnes de mesure | 13 |
| 3.1.2 | Teneurs particulières acquises par les méthodes prises pour référence au cours des campagnes de mesure | 15 |
| 3.1.3 | Distributions des concentrations acquises par les méthodes prises pour référence selon le moment de la journée et de la semaine | 19 |
| 3.2 | Performances de mesure des méthodes dites « candidates » | 20 |
| 3.2.1 | Séries temporelles des données acquises par les méthodes candidates | 20 |
| 3.2.2 | Estimation des métriques de performance | 25 |
| 3.2.1 | Bilan par méthode candidate..... | 26 |
| 4 | Conclusion et Perspectives | 28 |
| 5 | Références | 29 |
| 6 | Liste des annexes..... | 31 |
| | Annexe A : Formule d'estimation des métriques..... | 32 |
| 1.1 | Moyennes horaires et/ou journalières | 32 |
| 1.2 | Moyennes (horaires et/ou journalières) par microcapteur candidat | 32 |
| 1.3 | Précision | 32 |
| 1.4 | Biais et linéarité | 33 |
| 1.5 | Erreur..... | 33 |

Table des figures

| | |
|--|----|
| Figure 1. Présentation des deux types de configurations matérielles employées lors des essais. | 9 |
| Figure 2. Présentation des périodes de fonctionnement des dispositifs mis en place au cours de la campagne réalisée chez l'opérateur de type « petit réseau ». | 13 |
| Figure 3. Présentation des périodes de fonctionnement des dispositifs mis en place au cours de la campagne réalisée chez l'opérateur de type « réseau moyen ». | 14 |
| Figure 4. Présentation des périodes de fonctionnement des dispositifs mis en place au cours de la campagne réalisée chez l'opérateur de type « grand réseau ». | 14 |
| Figure 5. Suivi temporel des concentrations horaires en particules par mesures équivalentes PM _{2.5} pour chaque campagne de mesure. | 16 |
| Figure 6. Suivi temporel des concentrations horaires en particules par mesures équivalentes PM ₁₀ pour chaque campagne de mesure. | 17 |
| Figure 7. Distribution des valeurs des teneurs horaires PM _{2.5} pour les trois campagnes de mesure. . | 18 |
| Figure 8. Distribution des valeurs des teneurs horaires PM ₁₀ pour les trois campagnes de mesure. .. | 18 |
| Figure 9. Boîte à moustaches des populations de données PM _{2.5} et PM ₁₀ acquises par la méthode prise pour référence au cours des deux semaines de campagne chez l'opérateur n°1. | 20 |
| Figure 10. Suivi temporel des concentrations horaires en PM _{2.5} par la méthode candidate n°1. | 21 |
| Figure 11. Suivi temporel des concentrations horaires en PM ₁₀ par la méthode candidate n°1. | 22 |
| Figure 12. Suivi temporel des concentrations horaires en PM _{2.5} par la méthode candidate n°2. | 23 |
| Figure 13. Suivi temporel des concentrations horaires en PM ₁₀ par la méthode candidate n°2. | 24 |
| Figure 14. Illustration des distances considérées dans les deux types de régression linéaire considérés dans l'étude. | 33 |

Résumé

L'utilisation de systèmes capteurs personnels de suivi de la qualité de l'air par un nombre croissant d'utilisateurs et d'experts conduit à une multiplication des données qui peuvent notamment être acquises en enceinte ferroviaire souterraine (EFS). Compte-tenu des retours d'expérience disponibles ainsi que des sources multiples d'erreur d'interprétation et d'utilisation, la question de la qualité des données acquises par de tels équipements se pose.

Pour répondre à ce besoin de clarification des performances actuelles des dispositifs individuels disponibles sur le marché, des essais in-situ d'évaluation de performances de systèmes capteurs dédiés à la mesure des particules en EFS ont été organisés par l'Ineris. Trois systèmes capteurs ont été sélectionnés et cinq dispositifs de chaque marque ont été acquis puis déployés dans trois stations de trois réseaux ferroviaires différents ont été sélectionnées pour accueillir. La liste des dispositifs retenus pour cette étude vise à être représentative des systèmes disponibles sans pour autant être exhaustive. De manière analogue, les stations ferroviaires choisies reflètent la typologie des réseaux auquel les utilisateurs ont accès en France. Les campagnes de mesure toutes initiées en 2022 ont duré *a minima* deux semaines sur chaque site de mesure.

Inédite en EFS, cette étude renforce les constatations déjà émises dans la littérature au sujet des limites des systèmes capteurs pour l'exposition individuelle. La qualité métrologique comme la disparité des performances de mesure des dispositifs disponibles sur le marché sont autant de difficultés auxquelles le grand public peut se heurter. La comparaison des données générées par ce type de dispositifs à des valeurs sanitaires n'est pas envisageable. L'emploi de ces données qualitatives pour une hiérarchisation et/ou un suivi des évolutions temporelles des teneurs particulières sur quais pourrait être envisagé pour certains systèmes pour peu que des précautions métrologiques puissent être mises en place afin d'identifier, évaluer et maîtriser au mieux les sources d'incertitudes qui peuvent peser sur la donnée par le dispositif employé. Il revient ainsi aux opérateurs et utilisateurs de ce type de dispositif de s'assurer d'une démarche qualité en adéquation avec le contexte et les objectifs pour lesquels ce type de dispositif peut être déployé avec, par exemple, la mise en place de phases de calibration ou de vérification des données acquises en regard de valeurs de référence.

Sur la base de cette première étude, les perspectives d'amélioration liées aux technologies et aux usages pourront mener à des études ouvertes à l'ensemble des fabricants/intégrateurs/distributeurs volontaires de systèmes capteur potentiellement compatibles avec les objectifs de hiérarchisation des atmosphères rencontrées en EFS.

Abstract

The use of personal sensor systems to monitor air quality by a growing number of users and experts is leading to an increase in the amount of data that can be acquired, particularly in underground rail enclosures (EFS). Given the feedback available, and the many sources of error in interpretation and use, the question of the quality of data acquired by such equipment arises.

This need for clarification of the current performances of the devices leads Ineris to organize a study to evaluate the measurement performance of personal sensor-systems dedicated to monitoring mass concentrations of suspended particulates (PM_{2.5}, PM₁₀), using so-called reference methods. Three sensor systems were selected, and five devices of each brand were acquired and deployed in three stations of three different railway networks. The list of devices selected for this study is intended to be representative of the systems available, without being exhaustive. Similarly, the railway stations selected reflect the typology of networks to which users have access in France.

A first of its kind in an underground rail enclosure, this study reinforces the observations already made in the literature about the limitations of personal exposure sensors. The metrological quality and disparity of measurement performance of the market-available devices are just some of the difficulties that the public may encounter. Ultimately, it is not possible to compare the data generated by this type of device with any health values. The use of these qualitative data to prioritize and/or monitor the temporal evolution of particulate levels on platforms could be envisaged for certain systems, provided that metrological precautions can be put in place to identify, evaluate and control as effectively as possible the sources of uncertainty that may weigh on the data provided by the device used. It is therefore up to the operators and users of this type of device to ensure a quality approach in line with the context and objectives for which this type of device may be deployed, with, for example, the implementation of calibration phases or verification of acquired data against reference values.

On the basis of this initial study, the prospects for improvement linked to technologies and uses could lead to studies open to all willing manufacturers/integrators/distributors of sensor systems potentially compatible with the objectives of prioritizing the atmospheres encountered in EFS.

Pour citer ce document, utilisez le lien ci-après :

Institut national de l'environnement industriel et des risques, , Verneuil-en-Halatte : Ineris - 213878 - v1.0, 16/11/2023.

Mots-clés : Qualité de l'air, Enceintes Ferroviaires Souterraines (EFS), micro-environnement, systèmes capteurs, particules

1 Contexte et objectifs

Depuis 2013, un plan d'actions pour l'amélioration de la qualité de l'air intérieur, intégré au plan national santé environnement (PNSE) depuis sa 3ème édition, prévoit notamment l'amélioration de la surveillance de la pollution de l'air dans les Enceintes Ferroviaires Souterraines (EFS).

Dans le cadre du 4ème plan national santé environnement (PNSE4, [1]) et dans le but d'accompagner les opérateurs dans la caractérisation des niveaux de concentration de certains polluants pour évaluer l'exposition des usagers, l'Ineris a élaboré un guide de recommandations pour la réalisation de mesures harmonisées de la qualité de l'air dans les enceintes ferroviaires [2]. Sur la base d'une importante recherche bibliographique regroupant des données d'études françaises et internationales, ce guide recommande de caractériser la qualité de l'air des atmosphères des quais situés en souterrain et des habitacles des rames en caractérisant les niveaux de concentrations de polluants d'intérêt prioritaire auxquels les voyageurs sont exposés. Selon les données disponibles et les recommandations du guide, les niveaux de concentration en particules en suspension dans l'air (PM₁₀ et PM_{2.5}) sont principalement visés par ce type de caractérisation.

Si le guide mentionne les références normatives et guides associés à la mesure des particules, les préoccupations croissantes de la population vis-à-vis de la qualité de l'air conduisent, de plus en plus, à des initiatives privées sous formes de campagnes de mesure au cours desquelles des données sont acquises à l'aide de systèmes capteurs portatifs destinés au grand public. Si certaines caractéristiques techniques de ces dispositifs (e.g. meilleure couverture spatio-temporelle, portabilité, haute résolution temporelle, coût d'achat réduit, faible encombrement...) peuvent présenter de nombreuses opportunités dans le domaine de l'évaluation de l'exposition individuelle, ces appareils montrent rapidement leurs limites et suscitent des interrogations quant à la qualité des données acquises et leur utilisation [3].

Concernant la qualité des données, de nombreuses études montrent des limites concernant la qualité métrologique limitée de ce type de dispositifs (manque de sélectivité et sensibilité aux interférents...) et laissent circonspect quant à leur utilisation dans le domaine privé [4-8]. Dans le domaine de la recherche, domaine auquel est réservé actuellement l'évaluation des expositions individuelles à la pollution de l'air par les systèmes capteur, l'ANSES dans sa saisine n° « 2018-SA-2071 » souligne qu'à peine la moitié des publications sur ce sujet faisait explicitement référence à une qualification métrologique des dispositifs de mesure employés. L'agence insiste alors à juste titre sur le fait que, même s'il est envisageable de penser que les équipes de recherche se sont assurées de la validité des équipements, l'absence de données métrologiques limite la confiance à accorder aux résultats [8].

Concernant l'emploi des données issues des systèmes capteurs utilisés à titre privé, ces données peuvent induire en erreur l'utilisateur sur le sens de l'information acquise. En effet, cette dernière n'est qu'une indication de la concentration dans l'air à un instant précis et un lieu précis. Elle ne reflète ainsi pas l'exposition de l'individu intégrée sur une durée donnée. Par conséquent, une donnée de ce type générée en temps-réel ne peut pas être mise au regard d'une valeur guide (ou VTR ou autre) établie au niveau d'une population pour un pas de temps bien plus long (24h, annuel, ...).

Dans le cadre de ses missions d'appui au ministère en charge de l'environnement, l'Ineris a réalisé des essais comparatifs de systèmes capteurs de particules avec des méthodes prises pour référence afin de mieux appréhender les performances métrologiques de ces outils en EFS et d'évaluer la pertinence d'utiliser de tels systèmes pour évaluer l'exposition individuelle aux particules dans l'air des usagers des EFS. Confrontée au grand nombre et diversité des systèmes capteurs disponibles sur le marché, l'étude s'est concentrée sur les dispositifs commercialisés, personnels et portables à bas coûts. La mise à disposition de ces systèmes au grand public et leur facilité d'utilisation pour les utilisateurs des EFS laissent supposer une diffusion plus large dans le futur et les prédisposent donc à être retenus pour cette étude.

Ce rapport vise à reprendre les éléments méthodologiques (e.g. sélection du matériel testé, organisation des campagnes...) ainsi que les résultats en lien avec cette étude d'évaluation des performances métrologiques de systèmes capteurs destinés au grand public en EFS. Les systèmes capteur évalués et les réseaux ferroviaires volontaires dans lesquels se sont déroulées les campagnes expérimentales ont été anonymisés dans ce document.

2 Matériel & Méthodes

2.1 Terminologie

Les définitions retenues par le groupe de travail sont celles proposées par l'Agence française de normalisation (Afnor) dans son fascicule de documentation relatif aux capteurs pour la qualité de l'air (FD X43-121, 2021) :

- **élément sensible**, transducteur transformant une grandeur physique observée en une grandeur utilisable de type signal électrique dont l'amplitude est relative notamment à la concentration d'un polluant recherché dans l'air ;
- **capteur**, dispositif équipé d'au moins un élément sensible servant à la prise d'informations relatives à la grandeur à mesurer ainsi que d'un système électronique pour l'acquisition et le traitement de données ;
- **système capteur**, matériel intégrant au moins un capteur ou un élément sensible et un logiciel pour détecter une quantité et/ou mesurer une concentration de composés (gaz, aérosol) sur un pas de temps prédéfini.

Les systèmes capteurs portatifs correspondent aux dispositifs portés par l'individu. Les termes « micro-capteurs », « système capteur personnel » et « système capteur individuel » sont autant de termes qui peuvent être rencontrés dans la littérature.

2.2 Sélection des modèles de systèmes capteurs de particules dits « candidats »

Actuellement, le marché des systèmes capteurs de particules peut être schématiquement scindé en deux catégories de dispositifs avec :

- 1- des dispositifs personnels et portables à bas coûts, destinés au grand public et appariables à des *smartphones*.
- 2- des dispositifs plus onéreux vendus généralement pour des mesures en poste fixe par des entreprises du secteur de la surveillance de la qualité de l'air. Ces produits ciblent une clientèle métier de type exploitants, collectivités, bureaux d'études...

Dans le cadre de ces travaux, 3 dispositifs commercialisés des systèmes « grand public » ont été retenus sur la base des deux critères précédents. Cette sélection vise à être représentative des systèmes disponibles sans pour autant être exhaustive. Ces trois dispositifs seront qualifiés de « candidats (aux critères de performance définis dans cette étude) » dans la suite du rapport et seront identifiés par les appellations « Candidat n°1 », « Candidat n°2 », « Candidat n°3 ».

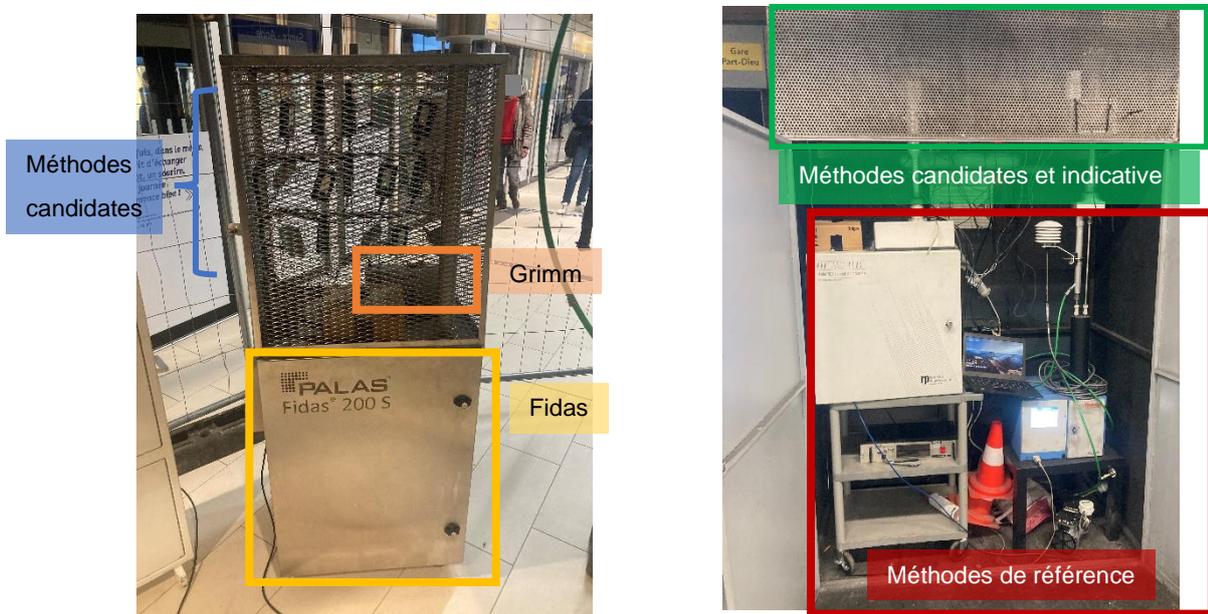
Afin d'appréhender les éventuelles dispersions de performances métrologiques dues à la production de chaque modèle de systèmes capteur, cinq dispositifs ont été acquis pour mener à bien les campagnes d'essais.

A ce stade, il faut noter que la collecte des données issues de l'ensemble des systèmes capteurs sélectionnés est réalisée à l'aide de *smartphones*, déployés lors des campagnes de mesure.

2.3 Méthodes prises pour « référence »

L'évaluation des performances des systèmes capteurs pour le suivi de la qualité de l'air en EFS est faite en comparant les données fournies par les méthodes « candidates » (systèmes capteurs personnels) à celles fournies par des méthodes prises pour référence (TEOM¹), voire des méthodes dites indicatives (e.g. compteurs optiques utilisés dans le cadre des hiérarchisations conduites dans le guide EFS) installées en parallèle sur les quais. Dans la mesure où il n'existe pas de référentiel précisant les objectifs de qualité des données pour les mesures en EFS, les méthodes prises pour référence correspondent à des méthodes reconnues comme équivalentes à la méthode de référence NF EN 12341 selon les recommandations de la norme NF EN 16450 dans le cadre de la surveillance de la qualité de l'air ambiant. Les méthodes indicatives sont déployées pour fournir des données complémentaires pouvant servir à interpréter les écarts entre les différents instruments de mesure.

Selon les opportunités et disponibilités des matériels des opérateurs participant à cette étude d'évaluation des performances de mesure, deux configurations d'essais ont été envisagées : configuration (a) et configuration (b). En ce qui concerne la première configuration, l'Ineris fournit, en parallèle d'une méthode prise pour référence, une plateforme mobile composée des méthodes candidates et diverses méthodes indicatives comme le Fidas® modèle 200S de Palas GmbH et le Dust Monitor 1.108 de Grimm Aerosol Technik. La plateforme présentée en Figure 1(a) montre un Fidas 200s sur lequel est montée une structure grillagée abritant les méthodes candidates et indicative.



a. Plateforme mobile mise à disposition au besoin par l'Ineris.

b. Installation des méthodes candidates dans une station de mesure d'un opérateur ferroviaire.

Figure 1. Présentation des deux types de configurations matérielles employées lors des essais.

La seconde configuration consiste à utiliser les structures déjà en place chez les opérateurs pour installer les méthodes candidates et indicatives. La Figure 1(b) présente l'installation réalisée chez un des opérateurs ferroviaires sélectionné dans le cadre de cette étude. Nous pouvons constater que les méthodes candidates et indicative ont été installées en partie haute de la station de mesure implantée sur quai. A titre d'information, deux des trois campagnes de mesure ont été réalisées avec cette configuration.

Comme évoqué précédemment, la collecte des données se fait au travers de smartphones dotées d'applications spécifiques pour chaque marque de systèmes capteurs. Afin de suivre le fonctionnement des appareils à distance, un modem 4G a de plus été installé. Ce dernier a en outre facilité la récupération des données en offrant la possibilité de collecter les mesures sur un serveur soit via un

¹ Tapered Element Oscillating Microbalance, méthode répondant aux exigences de la norme NF EN 16450 « Air ambiant — Systèmes automatisés de mesurage de la concentration de matière particulaire (PM10 ; PM2,5) ».

site dédié, soit via une *API* (*application programming interface* ou « interface de programmation d'application »).

2.4 Réseaux ferroviaires investigués

Plusieurs typologies de réseaux ont été investigués :

- un « Grand réseau » disposant de plus de 300 stations,
- un « Moyen réseau » disposant de 30 à 40 stations,
- un « Petit réseau » disposant de 20 stations.

Plusieurs exploitants de réseaux ferroviaires français ont répondu favorablement pour mettre à disposition leurs enceintes pour conduire ces essais. La sélection finale des opérateurs a été réalisée dans l'objectif de disposer de différentes typologies de réseaux et de manière à mener les expérimentations sur l'année 2022. Cet aspect méthodologique permet de tester les systèmes capteurs dans des environnements différents en termes de niveaux de concentrations de PM qui peuvent notamment être liés à l'âge des réseaux et du type de matériel roulant.

L'ensemble du matériel (dispositifs « candidats » et « pris pour référence ») a été déployé pour des campagnes durant **a minima deux semaines** chez les opérateurs volontaire retenus. Le Tableau 1 précise à ce titre les dates de début et de fin de campagne de mesure. Compte-tenu des contraintes d'exploitation des stations principalement liées à des problématiques de sécurité, les dispositifs de mesure ont été déployés à différentes positions relatives sur le quai. Le Tableau 1 reprend ces positionnements relatifs et précise également la configuration de déploiement optée pour chaque campagne de mesure.

Tableau 1. Organisation des campagnes d'essais chez les opérateurs ferroviaires volontaires.

| Opérateur | Typologie du réseau | Dates des campagnes de mesure | Positionnement sur le quai |
|---------------|---------------------|-------------------------------------|--|
| Opérateur n°1 | Petit | 7 octobre au 24 octobre 2022 | Extrémité de quai en configuration a (plateforme de mesure mobile) |
| Opérateur n°2 | Moyen | 17 novembre au 5 décembre 2022 | Milieu de quai en configuration b (station de mesure) |
| Opérateur n°3 | Grand | 8 décembre 2022 au 1er février 2023 | Milieu de quai en configuration b (station de mesure) |

2.5 Méthodologie pour le traitement des données

Bien que des travaux et des actions sont engagées par l'Agence française de normalisation (Afnor)² et le comité européen de normalisation (CEN)³ en ce sens, il n'existe pas à l'heure actuelle de protocoles d'essai, de critères de performance ou de protocoles de traitement des données normalisés pour les systèmes-capteurs dédiés à la mesure des particules dans l'air ambiant, les environnements intérieurs ou encore les micro-environnements.

La revue de la littérature mène à constater que plus d'une centaine de systèmes capteurs (ou éléments sensibles) ont déjà fait l'objet d'évaluations indépendantes [6]. Cependant, à notre connaissance, aucune étude d'évaluation des performances de mesure des systèmes capteurs dans les EFS françaises n'a été publiée à date de rédaction de ce rapport.

Le traitement des données et l'exploitation des résultats s'est appuyé sur différents travaux de la littérature [9-12], menés par l'Ineris (e.g. les essais nationaux d'aptitude des micro-capteurs pour la surveillance de la qualité de l'air [13]), de la certification « Air Quality Sensor » des systèmes capteurs pour la mesure de la qualité de l'air extérieur [14], les travaux normatifs internationaux du groupe de

² Avec la publication d'un fascicule de documentation relatif aux capteurs pour la qualité de l'air (FD X43-121, 2021)

³ En 2015, un groupe de travail (TC 264/WG 42 « Capteurs pour la qualité de l'air ») est constitué pour l'élaboration d'une spécification technique visant à fournir des lignes directrices pour l'évaluation des performances des systèmes capteurs utilisés pour la mesure indicative de polluants gazeux et particulaires de l'air ambiant. La publication de la première partie de cette norme relative aux polluants gazeux a eu lieu en 2022. Le pendant pour les polluants particulaires est attendu.

travail WG42 au sein du Comité européen de normalisation sur la qualité de l'air ou des organismes évaluateurs de différents pays européens ou américains (Airparif, Joint research Centre, Environmental Protection Agency, AQ-SPEC...) [15-21].

Les critères d'appréciation des performances sont repris dans les sous-sections suivantes.

2.5.1 Métriques d'évaluation de la performance des méthodes dites « candidates »

De nombreuses métriques peuvent être envisagées pour évaluer les performances de systèmes de mesure en comparaison à des valeurs vraies ou des mesures prises pour référence. Leur définition pouvant varier ou être sujette à discussion, le Tableau 2 reprend celles considérées dans le cadre de cette étude. Leur formalisation mathématique est reprise en Annexe de ce rapport (Annexe A : Formule d'estimation des métriques).

Il est à noter que les notions de *limite de détection* ou encore *de quantification* ne sont pas abordées dans le cadre de cette étude. Ce choix délibéré tient au fait que le protocole d'évaluation des performances ne tient pas compte des usages des instruments envisagés pour ce type de dispositifs comme la mesure en mobilité, le suivi d'exposition etc. La prise en compte de tels aspects mènerait à complexifier le protocole d'étude sans pour autant apporter d'éléments nouveaux sur la capacité des méthodes candidates à pouvoir fournir une donnée représentative des teneurs particulières des EFS dans lesquelles ils sont placés.

Tableau 2. Métriques considérées par le cadre de cette étude.

| Métrique | Description |
|-----------|---|
| Précision | Variation autour de la moyenne d'un ensemble de mesures rapportées simultanément par les cinq micro-capteurs du même type, installés dans les mêmes conditions d'échantillonnage. La précision est mesurée ici à l'aide de l'écart type (σ) et du coefficient de variation (CV). |
| Biais | Désaccord systématique (non aléatoire) entre les concentrations rapportées par les méthodes candidates et prises pour référence. Le biais est déterminé à l'aide d'une modélisation des données par régression linéaire orthogonale. |
| Linéarité | Mesure dans laquelle les mesures rapportées par un capteur peuvent expliquer les concentrations rapportées par l'instrument pris pour référence. La linéarité est déterminée ici à l'aide du coefficient de détermination (R^2). |
| Erreur | Mesure du désaccord entre les concentrations de polluants rapportées par le capteur et l'instrument pris pour référence. L'erreur est mesurée ici à l'aide de l'erreur quadratique moyenne (RMSE) et de l'erreur quadratique moyenne normalisée (NRMSE). |

2.5.2 Valeurs de comparaison pour les métriques de performance

Comme évoqué précédemment, les paramètres de performance et les valeurs cibles correspondantes ont été identifiés sur une sélection de critères constituant la base de l'état actuel de connaissance, des analyses documentaires, des conclusions d'autres organisations qui effectuent des évaluations de routine des capteurs ou encore des programmes de normalisation/certification des capteurs en cours d'élaboration par d'autres organisations. A des fins de comparaison avec les données disponibles au grand public, les valeurs cibles ont été notamment reprises de celles proposées par l'U.S. EPA [21]. A ce titre, le Tableau 3 reprend les différents critères et les valeurs cibles pour les protocoles d'essai envisagés dans le cadre de cette étude. Pour ce qui est du critère portant sur les erreurs de mesure, un dispositif dont les performances permettent de répondre à l'un des deux critères sera considéré comme y répondant.

Tableau 3. Mesures de performance et valeurs cibles considérées pour les capteurs de PM dans l'air des EFS.

| Métrique | | Valeurs repères |
|-----------|--|---|
| Précision | Ecart-type de mesure (σ) | $\leq 5 \mu\text{g}/\text{m}^3$ |
| | Coefficient de variation (CV) | $\leq 30\%$ |
| Biais | Pente (a) | $1,0 \pm 0,35$ |
| | Ordonnée à l'origine (b) | $-5 \leq b \leq 5 \mu\text{g}/\text{m}^3$ |
| Linéarité | Coefficient de détermination (R^2) | $\geq 0,70$ |
| Erreurs | RMSE | $\leq 7 \mu\text{g}/\text{m}^3$ |
| | NMSE | $\leq 30\%$ |

3 Résultats

3.1 Déroutement des campagnes

3.1.1 Temps de recouvrement des campagnes de mesure

Comme évoqué précédemment, les campagnes d'évaluation des performances de mesure ont été effectuées chez les opérateurs ferroviaires pour des durées minimales de deux semaines.

Les figures Figure 2, Figure 3 et Figure 4 présentent sous forme schématique les périodes de fonctionnement des différents dispositifs impliqués et ce, pour les trois campagnes de mesure. Il faut noter que pour la campagne menée chez l'opérateur de type « petit réseau », l'appareil servant à fournir des données prises pour référence est un *High Volume Aerosol Sampler DHA-80*. Se faisant, une tête de coupure est employée pour sélectionner les PM_{2.5} ou les PM₁₀. Ceci amène donc à considérer deux phases lors de la campagne : une phase où les PM_{2.5} sont prélevées et une autre phase où les PM₁₀ sont prélevées.

Pour chaque campagne, une distinction est faite entre les méthodes prises pour référence/indicative et les méthodes candidates. En outre, pour chacune de ces dernières, les cinq dispositifs testés sont identifiés à des fins de traçabilité par un numéro allant de 1 à 5, numéro qui a été conservé pour l'ensemble des campagnes de mesure.

La lecture de ces figures permet de constater qu'hormis certaines interruptions dans l'acquisition des données, les dispositifs des méthodes candidates n°1 et n°2 permettent d'acquérir des données sur l'ensemble des périodes des campagnes. Ils présentent en outre l'avantage de fonctionner sur batterie et ont donc la capacité d'enregistrer les données même lorsque les installations ne sont plus alimentées. Un exemple de cela est observable lors de la campagne chez l'opérateur de type « petit réseau ».

Pour ce qui est des interruptions dans l'acquisition des données lors des différentes campagnes pour les méthodes candidates n°1 et n°2, ils ont été généralement associés à des pertes d'appairage temporaires avec les smartphones associés.

Pour la méthode candidate n°3, l'exploitation des données montre que pour chaque campagne de mesure, il n'a pas été possible d'extraire ne serait-ce que 10% des données qui ont auraient dues être produites lors des trois campagnes de mesure. Dans ces conditions, une exploitation des données n'a pas pu être réalisée. Les données issues du déploiement de la méthode candidate n°3 ont donc été écartées de l'étude.

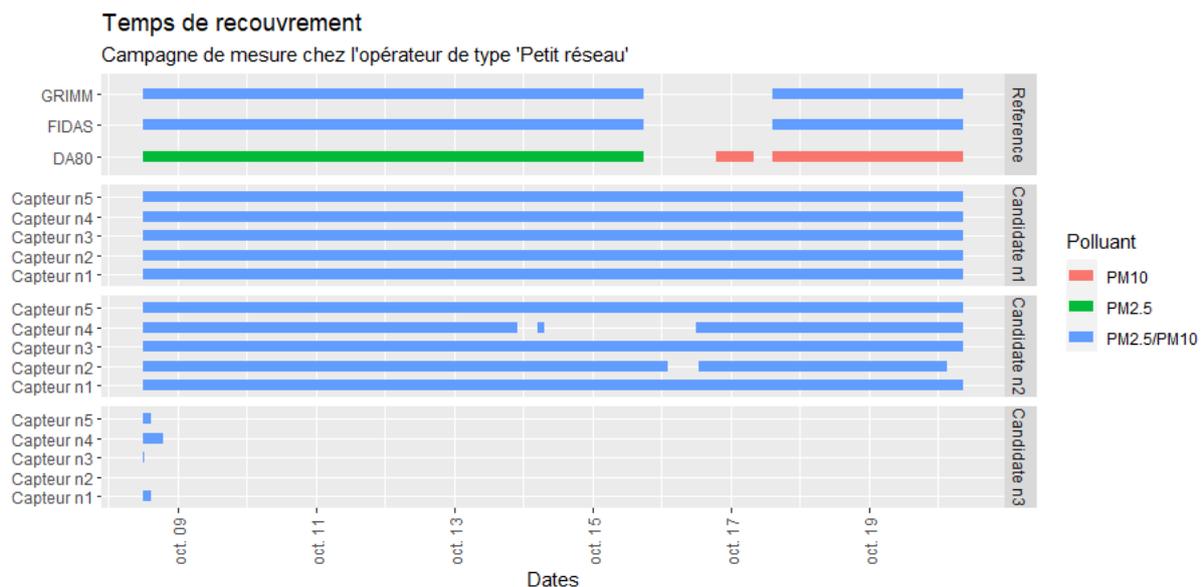


Figure 2. Présentation des périodes de fonctionnement des dispositifs mis en place au cours de la campagne réalisée chez l'opérateur de type « **petit réseau** ».

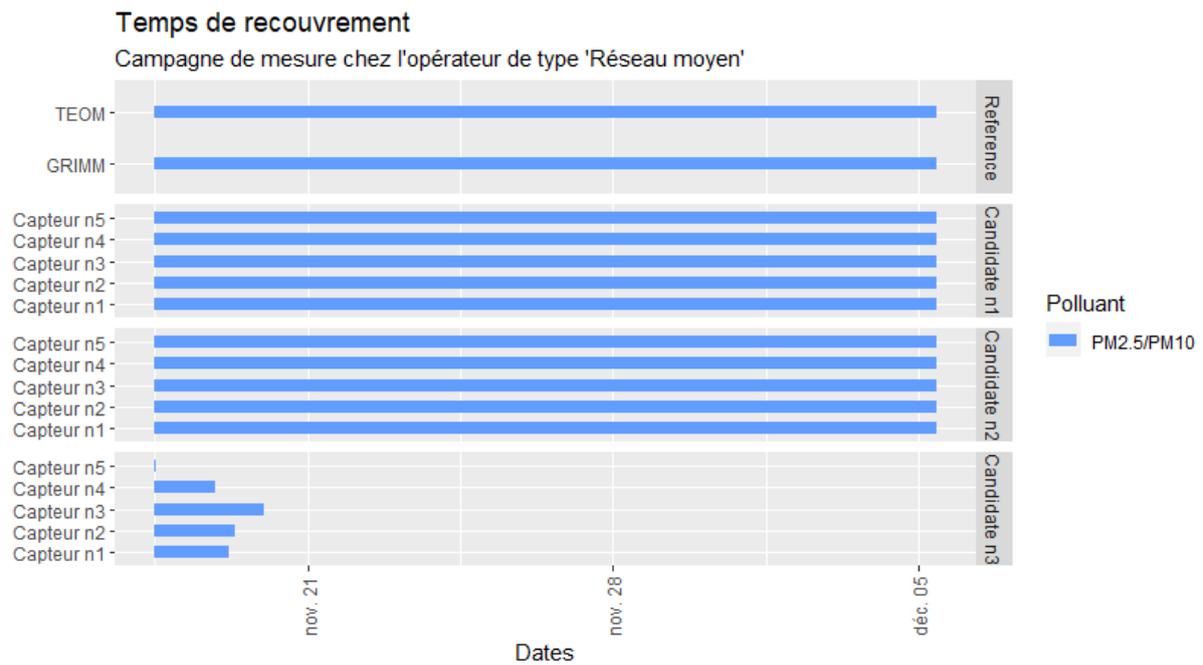


Figure 3. Présentation des périodes de fonctionnement des dispositifs mis en place au cours de la campagne réalisée chez l'opérateur de type « **réseau moyen** ».

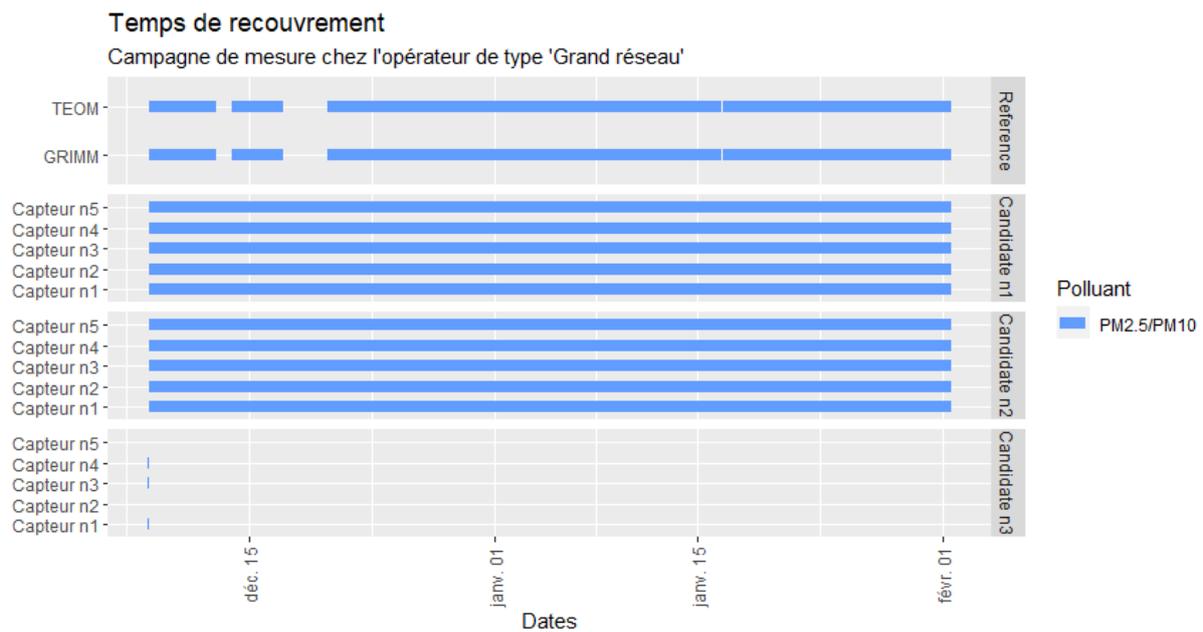


Figure 4. Présentation des périodes de fonctionnement des dispositifs mis en place au cours de la campagne réalisée chez l'opérateur de type « **grand réseau** ».

3.1.2 Teneurs particulières acquises par les méthodes prises pour référence au cours des campagnes de mesure

La disponibilité de mesures prises pour référence pour les $PM_{2.5}$ et PM_{10} permet dans un premier temps de s'intéresser aux teneurs auxquelles les systèmes capteurs ont été exposés tout au long des différentes campagnes de mesure. Les suivis temporels des concentrations horaires sont ainsi présentés en Figure 5 et Figure 6. Les intervalles de temps bleutés sur les graphiques représentent des intervalles de temps durant lesquels les mesures ne sont pas disponibles.

Des maximas et minimas locaux des séries temporelles des données de concentrations sont repérés respectivement de points rouge et vert. Les valeurs des concentrations à ces instants sont affichées et, quand l'affichage le permet, les horaires de survenue sont également indiqués entre parenthèses. Sur la base de ces informations, il est intéressant de constater le profil cyclique des concentrations, caractéristique de ce microenvironnement. Ce profil est par ailleurs plus marqué pour les $PM_{2.5}$ que les PM_{10} .

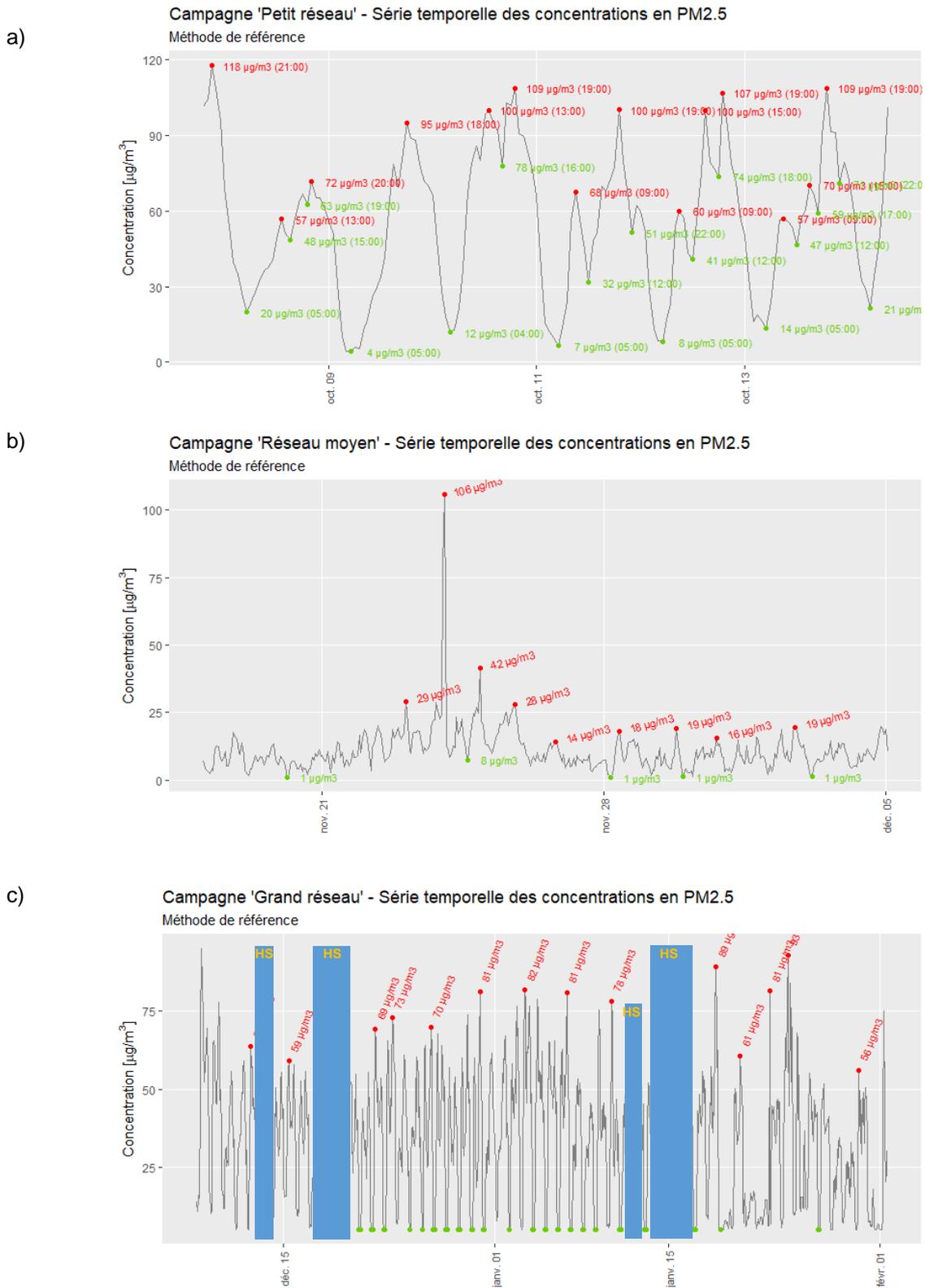
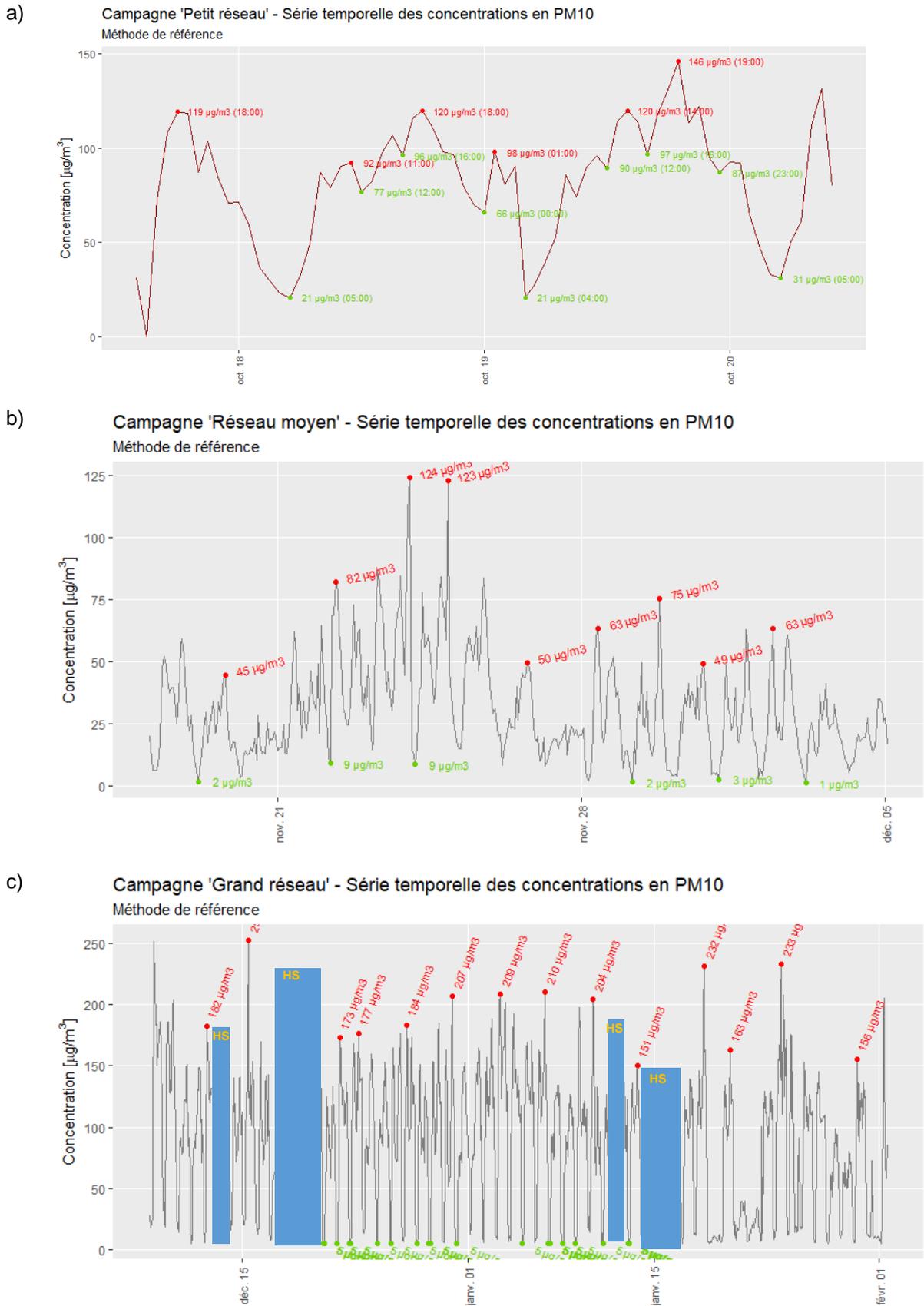


Figure 5. Suivi temporel des concentrations horaires en particules par mesures équivalentes PM_{2.5} pour chaque campagne de mesure.



Les Figure 7 et Figure 8 présentent respectivement les distributions des valeurs en $PM_{2.5}$ et PM_{10} . Il est intéressant de noter que les distributions des concentrations pour l'opérateur 3 ('Grand réseau') sont très nettement multi modales quand ce phénomène semble être plus atténué pour les deux autres opérateurs. Cette constatation pourrait traduire différentes situations d'émissions de particules. A ce stade, il n'est pas possible d'émettre d'hypothèses quant à la nature de ces situations sans données-métier complémentaires. Elles pourraient être liées au fonctionnement de l'EFS, à la fréquentation, la pénétration de la pollution extérieure, le matériel roulant... Dans le cadre de ce rapport, seule l'hypothèse d'une influence de la fréquentation en lien avec des périodes identifiables en distinguant les jours et heures ouvrés ou non-ouvrés est considérée et analysée.

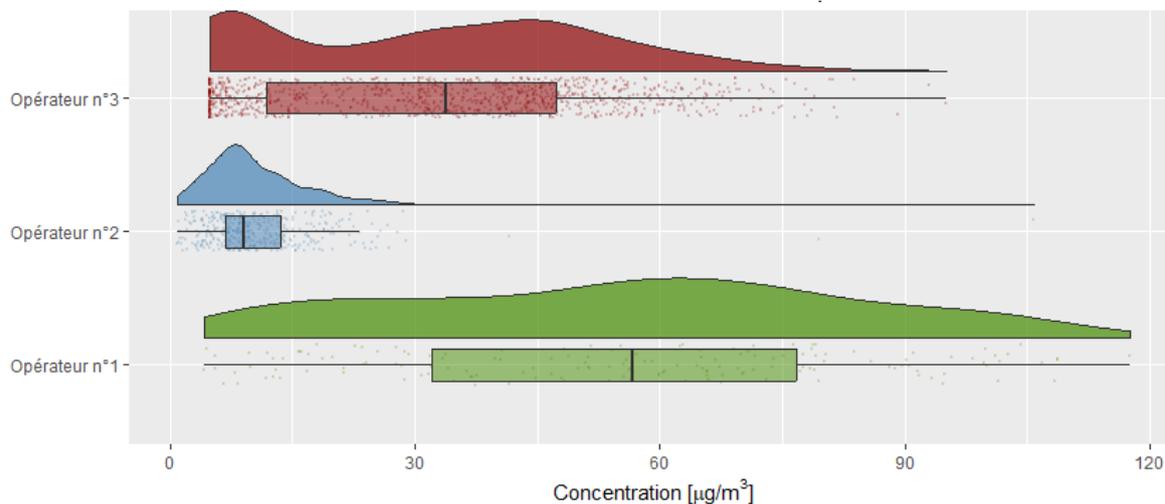


Figure 7. Distribution des valeurs des teneurs horaires $PM_{2.5}$ pour les trois campagnes de mesure.

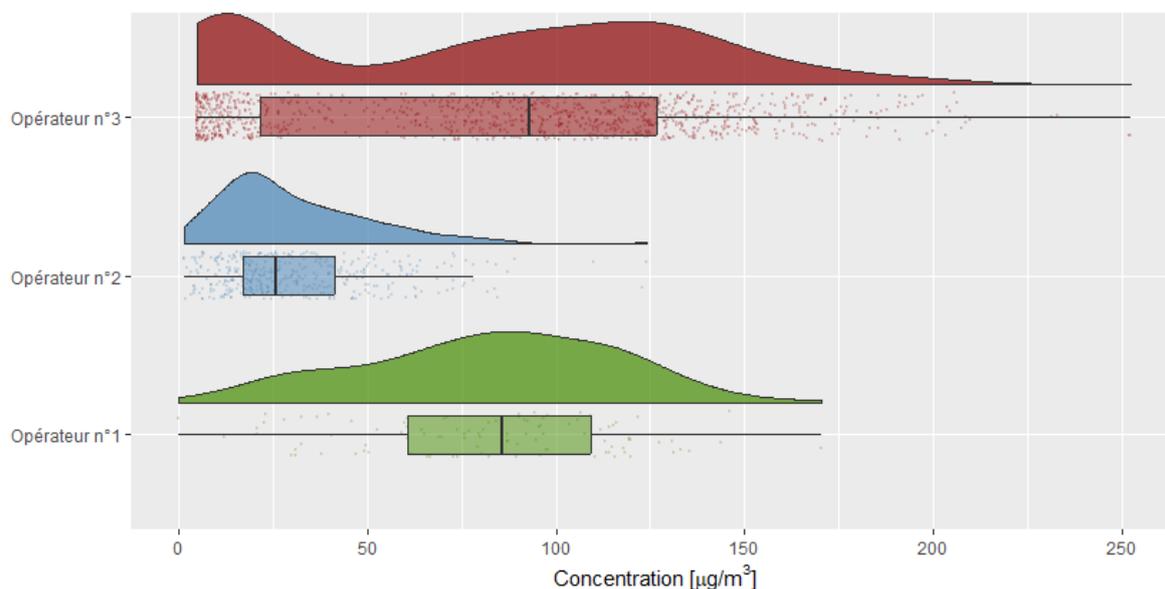


Figure 8. Distribution des valeurs des teneurs horaires PM_{10} pour les trois campagnes de mesure.

Les profils des 3 distributions permettent toutefois de valider la méthodologie choisie pour cette étude en considérant 3 types de réseau puisqu'ils diffèrent statistiquement les uns des autres, tant en termes de moyenne que de variance. Ceci a notamment été confirmé à l'aide de tests de rangs signés de Wilcoxon et de Levene⁴. Ces tests statistiques permettent de tester l'hypothèse selon laquelle les

⁴ https://fr.wikipedia.org/wiki/Test_des_rangs_sign%C3%A9s_de_Wilcoxon

distributions de chacun de deux groupes de données sont proches en détectant de manière non-paramétrique des changements sur la médiane entre deux populations.

Dans le cadre des comparaisons des populations menées sur les données PM_{2.5} et PM₁₀, les valeurs-p⁵ (toutes inférieures à 0,01) ont mené à infirmer les hypothèses nulles adoptées. Statistiquement, cela confirme ce qui peut être observé en Figure 7 et Figure 8, les distributions sont différentes et ainsi les sollicitations auxquelles sont soumises les capteurs sont différentes.

Ainsi, les méthodes candidates sont exposées :

- **chez l'opérateur 1 (de type 'Petit réseau')**, à des concentrations horaires variant sur un intervalle interquartile allant de 32 µg.m⁻³ à 77 µg.m⁻³ (médiane à 57 µg.m⁻³) pour les PM_{2.5} et de 61 µg.m⁻³ à 109 µg.m⁻³ (médiane à 86 µg.m⁻³) pour les PM₁₀
- **chez l'opérateur 2 (de type 'Réseau moyen')**, à des concentrations horaires variant sur un intervalle interquartile allant de 7 µg.m⁻³ à 14 µg.m⁻³ (médiane à 9 µg.m⁻³) pour les PM_{2.5} et de 17 µg.m⁻³ à 31 µg.m⁻³ (médiane à 26 µg.m⁻³) pour les PM₁₀
- **chez l'opérateur 3 (de type 'Grand réseau')**, à des concentrations horaires variant sur un intervalle interquartile allant de 12 µg.m⁻³ à 47 µg.m⁻³ (médiane à 34 µg.m⁻³) pour les PM_{2.5} et de 22 µg.m⁻³ à 126 µg.m⁻³ (médiane à 93 µg.m⁻³) pour les PM₁₀.

Ces valeurs sont cohérentes quand elles sont comparées à celles obtenues sur quai dans des études précédentes (notamment le guide recommandations pour la « Réalisation de mesures harmonisées de la qualité de l'air » de l'Ineris [1]).

3.1.3 Distributions des concentrations acquises par les méthodes prises pour référence selon le moment de la journée et de la semaine

Outre la construction des séries temporelles horaires, l'ensemble des données⁶ acquises par les méthodes prises pour référence ont été utilisées pour construire des sous-populations de teneurs en fonction du moment d'acquisition de la mesure dans la journée et la semaine. En effet, pour chaque polluant (PM_{2.5} et PM₁₀) et chaque campagne de mesure, différentes périodes ont pu être considérées : la totalité de la campagne, les jours de la semaine, les week-ends et les périodes de pointe en semaine.

Pour chaque campagne de mesure, des comparaisons des sous-populations entre elles ont pu être menées à l'aide de tests des rangs signés de Wilcoxon. Une illustration des comparaisons réalisées sur les données obtenues lors de la campagne menée chez l'opérateur 1 (de type 'Petit réseau') est proposée en Figure 9. Les conclusions étant les mêmes pour les 3 opérateurs, il a été décidé de ne traiter dans ce paragraphe que du cas de l'opérateur n°1 du fait de l'étendue des concentrations en PM_{2.5} obtenues lors de la campagne de mesure.

Si la méthodologie d'interprétation des résultats sur la base des valeurs p reste la même, les seuils considérés pour les p-value ont été symbolisés de la manière suivante sur la figure :

- $p \leq 0,01$ ou « **** » : très forte présomption contre l'hypothèse nulle ;
- $0,01 < p \leq 0,05$ ou « *** » : forte présomption contre l'hypothèse nulle ;
- $0,05 < p \leq 0,5$ ou « ** » : faible présomption contre l'hypothèse nulle ;
- $0,1 < p$ ou « * » voire « n.s. » : pas de présomption contre l'hypothèse nulle.

⁵ Grandeur statistique qui permet de quantifier la significativité statistique d'un résultat dans le cadre d'une hypothèse nulle (égalité des deux échantillons) et pour un taux d'erreur de première espèce fixé à 5%.

⁶ Ici, les données quart-horaires sont utilisées pour que l'échantillonnage temporel des valeurs des teneurs reste cohérent avec les intervalles de temps considérés s'étalant sur quelques heures.

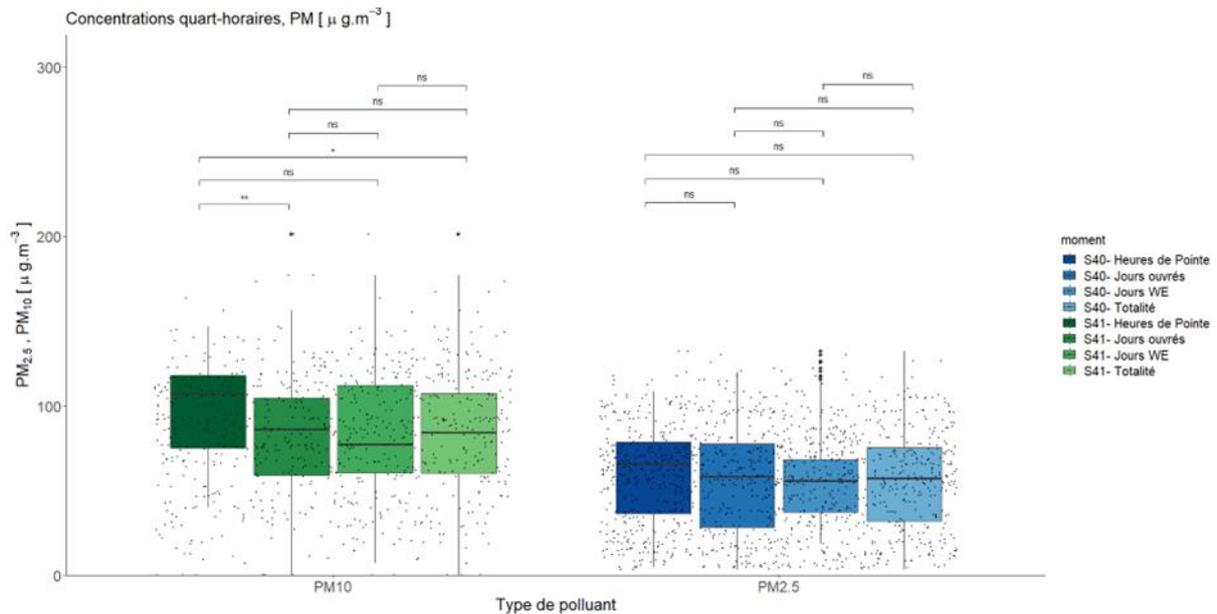


Figure 9. Boîte à moustaches des populations de données $PM_{2.5}$ et PM_{10} acquises par la méthode prise pour référence au cours des deux semaines de campagne chez l'opérateur n°1.

Il est intéressant de noter que pour les $PM_{2.5}$, les concentrations quart-horaires acquises par la méthode prise pour référence semblent issues d'une même population quel que soit la période considérée. Pour les PM_{10} , les teneurs en heures de pointe ont une tendance à être supérieures aux autres moyennes évaluées sans pour autant que cela soit bien marqué. En revanche, comme pour les $PM_{2.5}$, il ne semble pas y avoir de différence entre les jours ouvrés et la fin de semaine en termes de concentrations massiques horaires. Si les émissions en heures de pointe des jours ouvrés sont élevées, leur durée est courte et n'influe pas énormément sur les teneurs en moyenne.

Ces conclusions conduisent à considérer, pour chaque campagne, l'ensemble des données sans différenciation de la période de la journée ou la semaine. Cette considération dans le cadre de cette étude comparative était importante à établir puisqu'une évolution du mode fonctionnement des EFS pouvait engendrer des conditions de tests différentes et induire des biais d'interprétation des performances des méthodes candidates.

3.2 Performances de mesure des méthodes dites « candidates »

3.2.1 Séries temporelles des données acquises par les méthodes candidates

Les suivis temporels des mesures acquises par les différents systèmes capteurs testés dits « méthodes candidates » sont proposés en Figure 10, Figure 11, Figure 12 et Figure 13. A titre comparatif, chaque figure inclut le suivi temporel assuré par la méthode prise pour référence acquis sur la même période.

3.2.1.1 Commentaires qualitatifs sur le suivi des teneurs en $PM_{2.5}$ par la méthode candidate n°1

Hormis⁷ le capteur n°3 qui présente un comportement différent de ses homologues, la méthode candidate n°1 présente un suivi temporel qui semble conserver les dynamiques d'évolution des concentrations en $PM_{2.5}$ et fidèle à la méthode prise pour référence. Ils sous-estiment cependant fortement les teneurs et semblent présenter une forte erreur systématique. Une disparité des résultats entre les différents capteurs peut être observées avec par exemple pour la campagne chez l'opérateur

⁷ Pour la campagne n°3 (type 'Grand réseau'), les résultats du capteur n°5 ont été écartés car au moment de la récupération dans la mesure où les voies d'aspiration du capteur étaient obstruées. N'ayant pas d'informations/hypothèses sur l'origine, le début, la durée de l'obstruction, il a été décidé d'écartier les données obtenues pour ne pas biaiser les résultats.

de type « petit réseau », des maxima deux à huit fois plus importants que ceux des capteurs sans même parler du capteur 4 qui ne semble pas percevoir ces maxima.

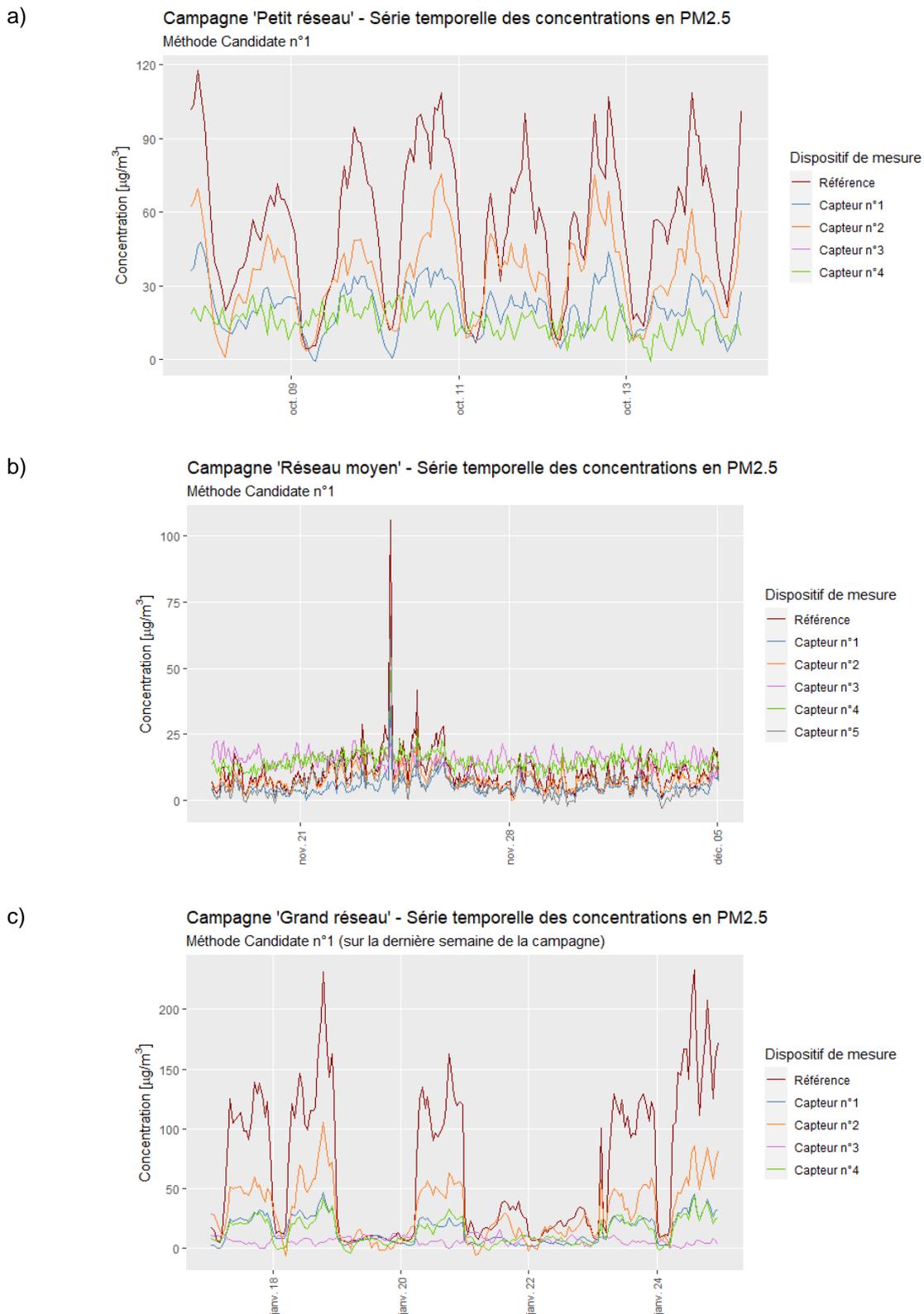


Figure 10. Suivi temporel des concentrations horaires en PM_{2.5} par la méthode candidate n°1.

3.2.1.2 Commentaires qualitatifs pour le suivi des teneurs en PM₁₀ par la méthode candidate n°1

Pour le suivi des teneurs en PM₁₀ avec la méthode candidate n°1, les conclusions sont similaires à ce qui a été évoqué pour les PM_{2,5} à savoir : (i) une sous-estimation des concentrations vraies (manque de justesse de la méthode), (ii) un comportement somme toute assez similaire pour les capteurs (hormis un appareil, la méthode semble fidèle), (iii) une bonne restitution des dynamiques d'évolution des concentrations au cours du temps.

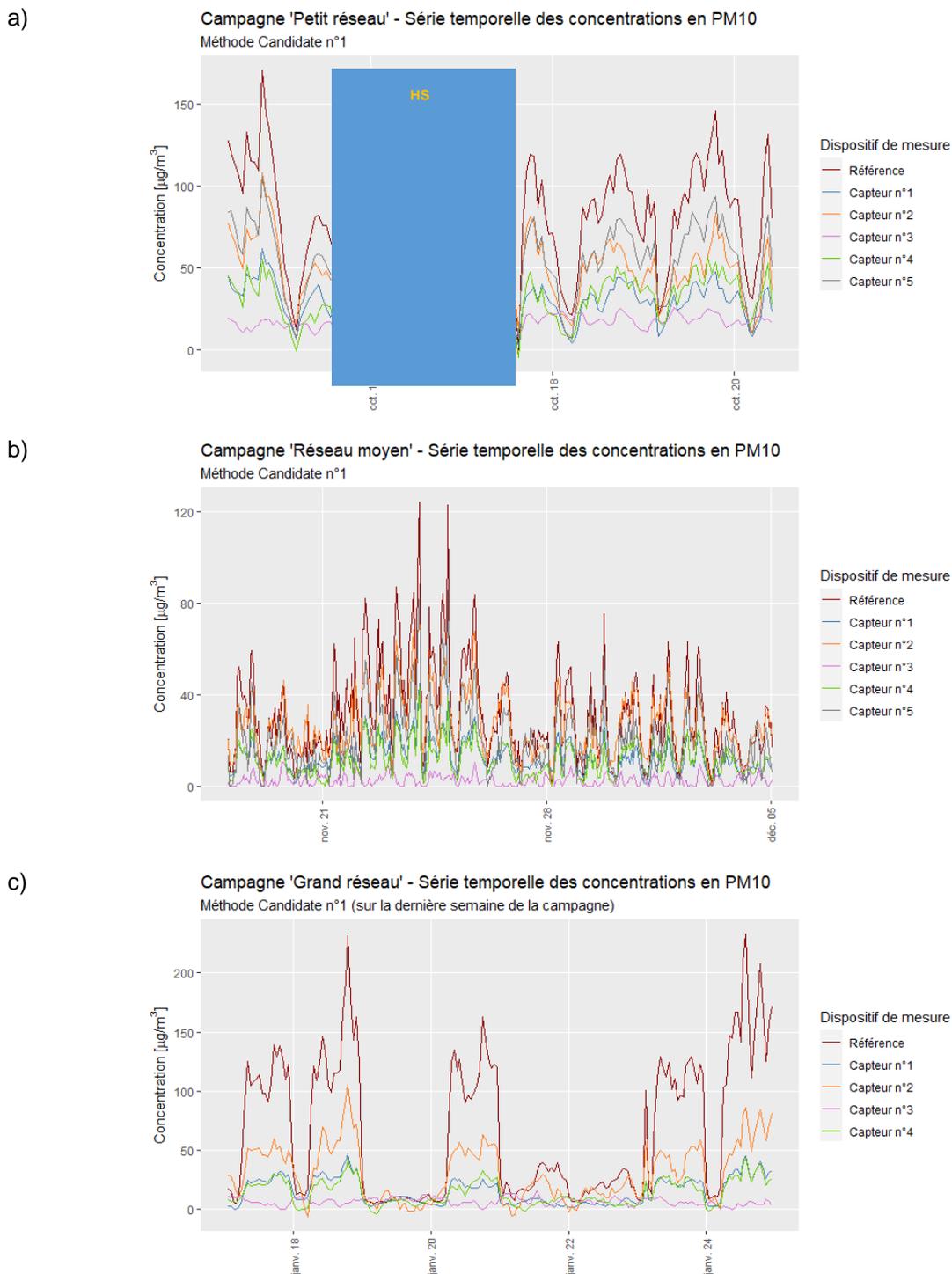
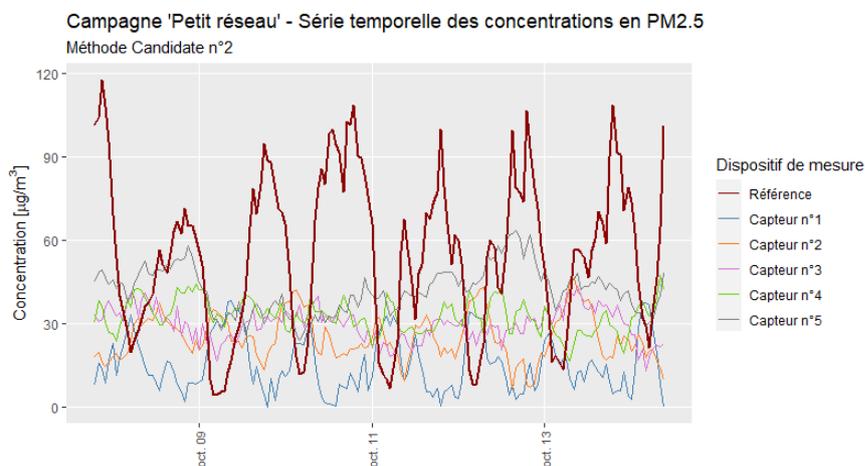


Figure 11. Suivi temporel des concentrations horaires en PM₁₀ par la méthode candidate n°1.

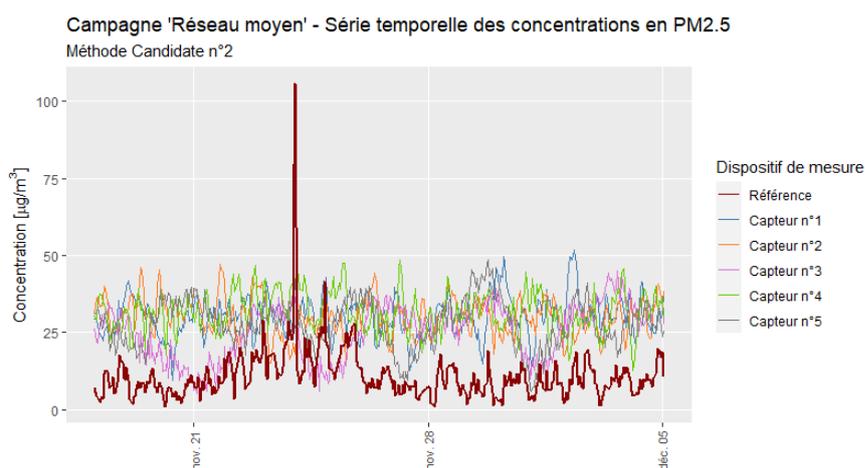
3.2.1.3 Commentaires qualitatifs pour des PM_{2.5} par la méthode candidate n°2

Contrairement à la méthode candidate n°1 qui semblait restituer les dynamiques d'évolution des concentrations à la fois pour les PM_{2.5} et les PM₁₀, il n'est pas possible de reconnaître les schémas d'évolution proposés par les méthode prise pour référence à partir de la méthode candidate n°2. Lors de la campagne n°1 (type 'Petit réseau'), certaines évolutions sont même en opposition à celles de la méthode prise pour référence. Lors des deux autres campagnes, les comportements dynamiques semblent être mal voire pas du tout restitués.

a)



b)



c)

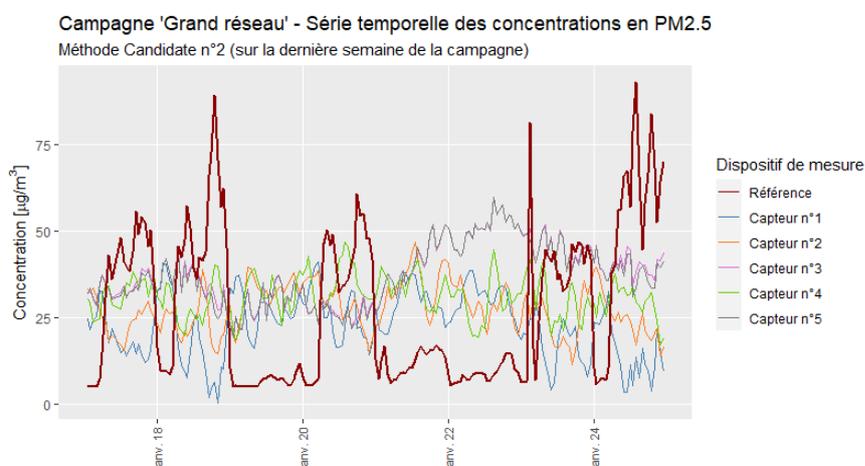


Figure 12. Suivi temporel des concentrations horaires en PM_{2.5} par la méthode candidate n°2.

De manière générale, la méthode n°2 surestime très fortement les concentrations notamment pour les campagnes n°2 (type 'Réseau moyen') et n°3 (type 'Grand réseau'). Les dynamiques étant mal perçues, il semblerait qu'il y ait une forte erreur systématique fournie par l'ensemble des dispositifs. Une problématique de limite de détection/quantification haute semble également apparaître dans la mesure où la surestimation des teneurs est plus présente pour des teneurs prises pour référence plus faibles.

3.2.1.4 Commentaires qualitatifs pour des PM₁₀ par la méthode candidate n°2

A l'instar de ce qui a pu être constaté pour les PM_{2.5}, la méthode n°2 ne restitue pas le comportement dynamique de l'évolution des PM₁₀. Etant donné l'écart (tant en valeurs absolues qu'en relatif) entre les teneurs retournées par les différents dispositifs, il semblerait que la méthode ne soit ni fidèle, ni juste.

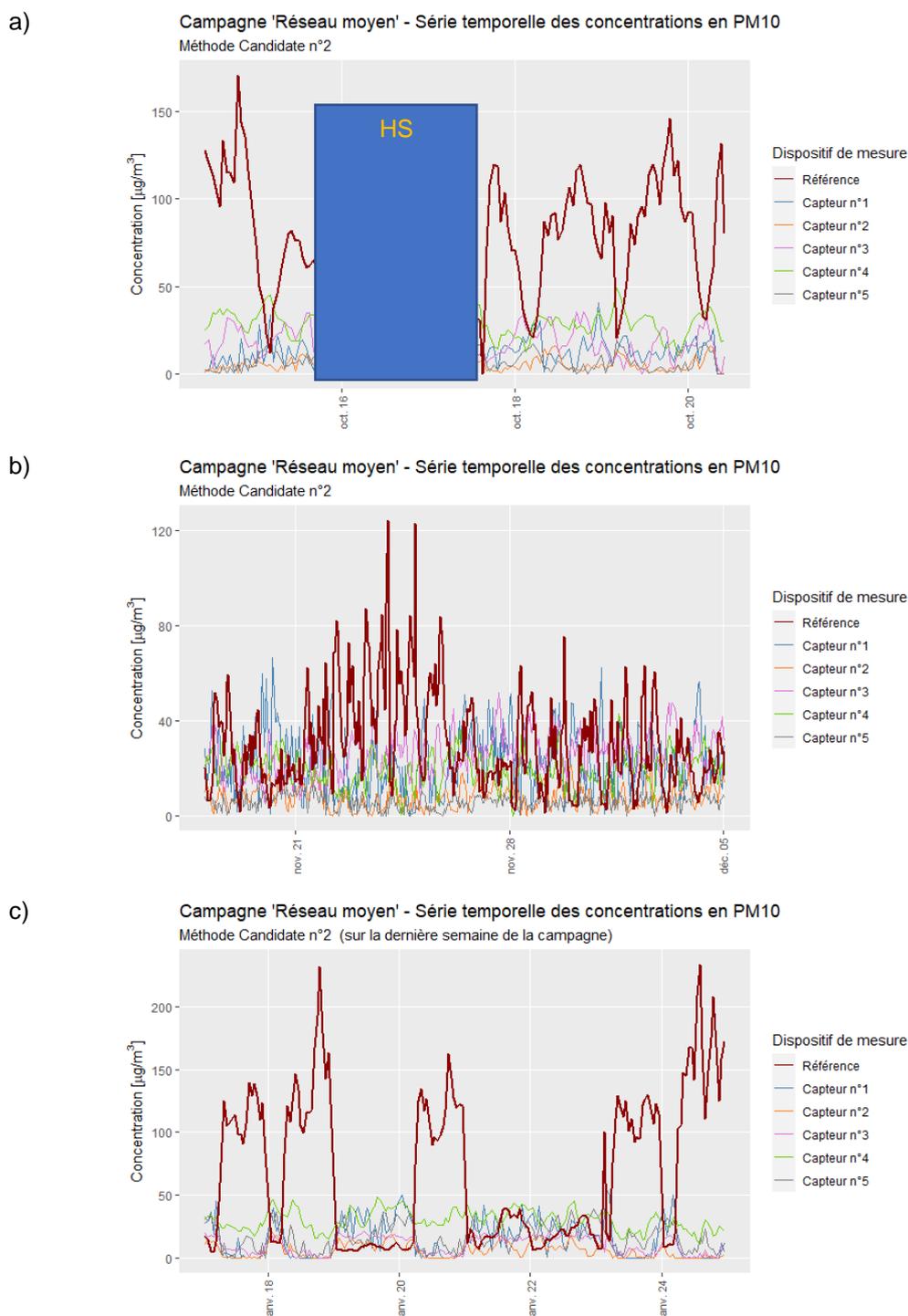


Figure 13. Suivi temporel des concentrations horaires en PM₁₀ par la méthode candidate n°2.

La méthode semble surestimer les teneurs de particules inférieures à $\sim 20\mu\text{g}/\text{m}^3$. Au-dessus de ces valeurs, la méthode semble sous-estimer les concentrations.

La surestimation des données étant plus présente quand les valeurs prises pour référence des teneurs en PM sont basses, il semblerait que la méthode n°2 possède (i) une limite de détection/quantification trop haute pour les ambiances investiguées, (ii) une erreur systématique importante.

3.2.2 Estimation des métriques de performance

Les métriques évoquées dans la section « Matériel & Méthodes » (sous-section 2.5.1) sont évaluées pour l'ensemble des dispositifs et méthodes sur les trois campagnes de mesure effectuées dans le cadre de cette étude. Les tableaux Tableau 4, Tableau 5 et Tableau 6 résument les valeurs obtenues. En vert dans ces tableaux, sont mises en relief les valeurs répondant aux critères évoqués en sous-section 2.5.2.

Tableau 4. Estimation des métriques pour la campagne de mesure n°1 (type « Petit réseau »).

| | Méthode candidate n°1 | | | | Méthode candidate n°2 | | | |
|-------------------|-----------------------|-------------------|------------------|---------------------|-----------------------|--------------|------------------|--------------|
| | PM _{2,5} | | PM ₁₀ | | PM _{2,5} | | PM ₁₀ | |
| Précision | SD | CV | SD | CV | SD | CV | SD | CV |
| Tous les capteurs | 4,1 | 32% | 3,71 | 29% | 13 | 45,00% | 16,44 | 114% |
| Biais | a | b | a | b | a | b | a | b |
| Capteur n° 1 | 0,34 ± 0,01 | 2,1 ± 0,69 | 0,33 ± 0,01 | 3,09 ± 1,09 | -0,29 ± 0,02 | 31,77 ± 1,12 | -0,11 ± 0,02 | 22,19 ± 2 |
| Capteur n° 2 | 0,56 ± 0,02 | 3,15 ± 1,4 | 0,56 ± 0,02 | 3,59 ± 2,04 | -0,17 ± 0,02 | 34,95 ± 1,3 | -0,05 ± 0,01 | 9,86 ± 0,93 |
| Capteur n° 3 | 0,37 ± 0,02 | 6,39 ± 1,27 | 0,02 ± 0,01 | 15,6 ± 1,02 | 0,06 ± 0,02 | 26,68 ± 1,04 | -0,01 ± 0,03 | 19,22 ± 2,57 |
| Capteur n° 4 | 0,04 ± 0,02 | 13,1 ± 0,96 | 0,37 ± 0,02 | 0,95 ± 1,51 | -0,01 ± 0,02 | 33,23 ± 1,05 | -0,1 ± 0,02 | 36,44 ± 1,91 |
| Capteur n° 5 | 0,43 ± 0,03 | 4,9 ± 0,71 | 0,67 ± 0,01 | -0,04 ± 1,21 | 0,09 ± 0,02 | 36,98 ± 1,51 | -0,13 ± 0,01 | 17,07 ± 0,76 |
| Linéarité | R ² | | R ² | | R ² | | R ² | |
| Capteur n° 1 | 0,86 | | 0,88 | | 0,62 | | 0,20 | |
| Capteur n° 2 | 0,80 | | 0,86 | | 0,31 | | 0,21 | |
| Capteur n° 3 | 0,81 | | 0,83 | | 0,06 | | 0,07 | |
| Capteur n° 4 | 0,04 | | 0,04 | | 0,00 | | 0,17 | |
| Capteur n° 5 | 0,90 | | 0,92 | | 0,09 | | 0,69 | |
| Erreur | RMSE | NMSE | RMSE | NMSE | RMSE | NMSE | RMSE | NMSE |
| Tous les capteurs | 42,04 | 196% | 82,85 | 144% | 41,75 | 143% | 77 | 544% |

Tableau 5. Estimation des métriques pour la campagne de mesure n°2 (type « Réseau moyen »).

| | Méthode candidate n°1 | | | | Méthode candidate n°2 | | | |
|-------------------|-----------------------|-------------------|--------------------|---------------------|-----------------------|-----------------|------------------|--------------|
| | PM _{2,5} | | PM ₁₀ | | PM _{2,5} | | PM ₁₀ | |
| Précision | SD | CV | SD | CV | SD | CV | SD | CV |
| Tous les capteurs | 3,23 | 24% | 4,34 | 30% | 7,4 | 26,00% | 11,61 | 69,00% |
| Biais | a | b | a | b | a | b | a | b |
| Capteur n° 1 | 0,33 ± 0,01 | 1,44 ± 0,1 | 0,34 ± 0 | 1,89 ± 0,11 | 0,2 ± 0,1 | 28,02 ± 1,42 | -0,4 ± 0,04 | 35,62 ± 1,55 |
| Capteur n° 2 | 0,61 ± 0,01 | 1,7 ± 0,1 | 0,67 ± 0,01 | 6,2 ± 0,52 | -0,22 ± 0,08 | 32,39 ± 1,04 | -0,09 ± 0,01 | 11,26 ± 0,44 |
| Capteur n° 3 | 0,39 ± 0,01 | 10,52 ± 0,16 | 0,33 ± 0,01 | 1,32 ± 0,25 | 4,77 ± 7,51 | -26,46 ± 102,46 | -0,04 ± 0,03 | 27,35 ± 0,98 |
| Capteur n° 4 | 0,02 ± 0,02 | 15,15 ± 0,3 | 0,02 ± 0,01 | 1,97 ± 0,22 | 0,44 ± 0,07 | 26,97 ± 0,95 | -0,09 ± 0,02 | 22,54 ± 0,73 |
| Capteur n° 5 | 0,68 ± 0,01 | -0,9 ± 0,2 | 0,68 ± 0,01 | -0,54 ± 0,49 | 0,74 ± 0,21 | 19,77 ± 2,91 | -0,1 ± 0,01 | 9,43 ± 0,2 |
| Linéarité | R ² | | R ² | | R ² | | R ² | |
| Capteur n° 1 | 0,84 | | 0,97 | | 0,01 | | 0,18 | |
| Capteur n° 2 | 0,95 | | 0,85 | | 0,12 | | 0,13 | |
| Capteur n° 3 | 0,74 | | 0,85 | | 0,00 | | 0,00 | |
| Capteur n° 4 | 0,00 | | 0,03 | | 0,09 | | 0,00 | |
| Capteur n° 5 | 0,84 | | 0,87 | | 0,03 | | 0,46 | |
| Erreur | RMSE | NMSE | RMSE | NMSE | RMSE | NMSE | RMSE | NMSE |
| Tous les capteurs | 7,18 | 72,00% | 22,77 | 154,00% | 21,1 | 73,00% | 29,02 | 174,00% |

Tableau 6. Estimation des métriques pour la campagne de mesure n°3 (type « Grand réseau »).

| | Méthode candidate n°1 | | | | Méthode candidate n°2 | | | |
|-------------------|-----------------------|--------------------|------------------|--------------------|-----------------------|--------------|------------------|--------------|
| | PM _{2,5} | | PM ₁₀ | | PM _{2,5} | | PM ₁₀ | |
| Précision | SD | CV | SD | CV | SD | CV | SD | CV |
| Tous les capteurs | 3,92 | 33% | 4,57 | 34% | 9,6 | 31,00% | 25,8 | 75 |
| Biais | a | b | a | b | a | b | a | b |
| Capteur n° 1 | 0,35 ± 0 | 2,17 ± 0,18 | 0,18 ± 0 | 2,93 ± 0,22 | -0,32 ± 0,01 | 34,07 ± 0,37 | -0,2 ± 0,02 | 27,75 ± 0,49 |
| Capteur n° 2 | 0,61 ± 0,01 | 2,72 ± 0,37 | 0,38 ± 0 | 6,25 ± 0,44 | -0,13 ± 0,01 | 32,78 ± 0,38 | -0,07 ± 0,03 | 10,59 ± 0,23 |
| Capteur n° 3 | -0,03 ± 0,02 | 15,1 ± 0,17 | -0,01 ± 0 | 7,52 ± 0,16 | 0 ± 0,02 | 28,63 ± 0,8 | -0,1 ± 0,01 | 18,2 ± 0,14 |
| Capteur n° 4 | 0,35 ± 0,1 | 11,22 ± 0,16 | 0,17 ± 0 | 3 ± 0,18 | 0,02 ± 0,01 | 31,87 ± 0,4 | -0,1 ± 0,02 | 37,38 ± 0,35 |
| Capteur n° 5 | - | - | - | - | 0,74 ± 0,21 | 19,77 ± 2,91 | -0,12 ± 0,01 | 26,09 ± 0,42 |
| Linéarité | R ² | | R ² | | R ² | | R ² | |
| Capteur n° 1 | 0,83 | | 0,86 | | 0,50 | | 0,60 | |
| Capteur n° 2 | 0,78 | | 0,87 | | 0,13 | | 0,48 | |
| Capteur n° 3 | 0,03 | | 0,07 | | 0,00 | | 0,00 | |
| Capteur n° 4 | 0,86 | | 0,89 | | 0,00 | | 0,43 | |
| Capteur n° 5 | - | | - | | 0,03 | | 0,43 | |
| Erreur | RMSE | NMSE | RMSE | NMSE | RMSE | NMSE | RMSE | NMSE |
| Tous les capteurs | 21,03 | 115% | 76 | 333% | 26,85 | 86% | 71,47 | 209,00% |

3.2.2.1 Pour la mesure des PM_{2,5} et PM₁₀ par la méthode candidate n°1

Dans l'ordre des critères listés dans les tableaux, nous pouvons constater que la méthode n°1 semble satisfaire aux exigences de précision (critères sur SD et CV) qui pourraient être attendues d'une mesure indicative. La méthode n°1 sous-estime les valeurs vraies avec des coefficients directeurs (a, issu de la régression linéaire orthogonale) systématiquement inférieurs à 1. Sur les 5 dispositifs, 4 (capteurs n°1, n°2, n°3 et n°5) présente des R² satisfaisant indiquant une bonne capacité à suivre les comportements dynamiques sur l'heure et la journée. Pour ce qui est de l'ordonnée à l'origine (coefficient b obtenu lors de la régression linéaire orthogonale), les valeurs obtenues satisfont un nombre conséquent de fois le critère fixé en sous-section 2.5.2. Ceci pourrait indiquer des limites de détection et quantification des appareils compatibles avec les niveaux de teneurs en PM des EFS. Si cela peut supposer que des améliorations sont possibles, en l'état actuel cette méthode ne permet pas de réaliser des mesures en EFS avec un quelconque objectif métrologique (NMSE >> 30%).

3.2.2.2 Pour la mesure des PM_{2,5} et PM₁₀ par la méthode candidate n°2

Aucun dispositif de la méthode n°2 ne semble remplir les critères de performances présentés en sous-section 2.5.2. Les performances évaluées semblent cohérentes avec les constatations faites sur les séries temporelles : (i) la méthode n'est pas fidèle (SD >> 5 µg/m³ et CV >> 30%), (ii) la méthode n'est pas juste (NMSE >> 30%), (iii) la méthode ne restitue pas ou difficilement les évolutions dynamiques des teneurs (a < 1, R² < 0,75) et une relation linéaire ne semble pas exister⁸ entre les deux séries temporelles (prise pour référence et candidate), (iv) présente une erreur systématique importante (RMSE >> 7 µg/m³ et valeurs de b en moyenne > 20 µg/m³). En conclusion, cette méthode ne permet pas du tout de réaliser des mesures en EFS avec un quelconque objectif métrologique et les indications sur le suivi des évolutions dynamiques des teneurs en PM semblent empreintes de nombreux biais.

3.2.1 Bilan par méthode candidate

3.2.1.1 Pour la méthode candidate n°1

Au final, la méthode candidate n°1 semble capable de suivre l'évolution dynamique des teneurs massiques en PM pour les trois campagnes de mesure. Les cinq dispositifs ont par ailleurs montré qu'ils pouvaient fournir des informations cohérentes entre elles avec des niveaux de reproductibilité satisfaisants. Si cet aspect est intéressant pour comparer de manière relative et qualitative des niveaux de concentration auxquels les individus peuvent être exposés dans les différents microenvironnements ciblés, les sous-estimations des teneurs fournies par les équipements sont trop importantes pour que les mesures soient considérées comme acceptables à des fins de quantification. Une utilisation de cette

⁸ Une analyse des données plus poussée semble indiquer que la relation est plutôt quadratique entre les deux séries.

méthode à des fins indicatives (*i.e.* une méthode respectant des objectifs de qualité des données moins stricts que ceux requis pour les mesures fixes dans la Directive 2008/50/CE) serait plus adéquat.

3.2.1.2 Pour la méthode candidate n°2

En comparaison avec la méthode candidate n°1, la méthode candidate n°2 a proposé des performances de mesure plus variables. Entre eux, les dispositifs ne semblent pas atteindre les mêmes niveaux de performance et ce, pour toutes les campagnes de mesure. Elle présente de fortes erreurs systématiques et un biais sur les dynamiques d'évolution des teneurs. Il semblerait qu'il ne soit pas possible d'obtenir une mesure fidèle, juste ou bien encore indicative et qualitative à partir des équipements de cette marque.

3.2.1.3 Pour la méthode candidate n°3

Sur les trois méthodes candidates, la méthode candidate n°3 n'a pas permis de fournir les mesures attendues sur chaque campagne (moins de 10%). Alors même qu'elle possède le même élément sensible que la méthode candidate n°1 ; ce constat illustre les difficultés auxquelles peuvent être confrontées les utilisateurs. Entre problématiques d'approvisionnement des composants électroniques, problématiques de communication, d'application... voire de la qualité de la donnée, les dysfonctionnements ont diverses origines et soulignent les difficultés de fabrication ou d'interprétation qui peuvent également surgir dans un marché en plein essor.

4 Conclusion et Perspectives

L'essor des systèmes capteurs notamment personnels, leurs développements technologiques et leur accès facilité contribuent à un marché qui se développe en dehors d'un cadre normatif. L'utilisation de systèmes capteurs par un nombre croissant d'utilisateurs et d'experts conduit à une multiplication des données qui peuvent notamment être acquises en EFS. Compte-tenu des retours d'expérience disponibles quant à l'emploi de tels dispositifs ainsi que des sources multiples d'erreur d'interprétation et d'utilisation, la question de la qualité des données acquises par de tels équipements se pose.

Pour répondre à ce besoin de clarification des performances actuelles des dispositifs individuels disponibles sur le marché, l'Ineris a réalisé, dans le cadre de ses missions d'appui technique au ministère chargé de l'environnement, une étude d'évaluation des performances métrologiques avec des méthodes prises pour référence en EFS. En cohérence avec les données disponibles dans la littérature portant sur la qualité de l'air dans les enceintes ferroviaires souterraines, cette étude s'est plus particulièrement intéressée aux performances de mesure des systèmes capteurs individuels fournissant des informations relatives aux teneurs massiques en particules dans l'air ($PM_{2.5}$ et PM_{10}), polluants d'intérêt dit prioritaire pour ce type d'environnement.

Pour tenir compte des niveaux de concentrations variables entre les différentes typologies de stations en France, trois stations de trois réseaux ferroviaires ont été sélectionnées pour accueillir une quinzaine de dispositifs de mesure parmi lesquels des instruments pris pour référence et des équipements de méthodes candidates. Pour ces dernières, trois marques disponibles sur le marché français ont été sélectionnées et cinq dispositifs de chaque marque ont été acquis afin d'appréhender indirectement les problématiques de reproductibilité de fabrication de ces équipements. Les campagnes de mesure toutes initiées en 2022 ont duré *a minima* deux semaines sur chaque site de mesure.

Cette étude renforce les constatations déjà émises dans la littérature au sujet des limites des systèmes capteurs pour l'exposition individuelle. La qualité métrologique comme la disparité des performances de mesure des dispositifs disponibles sur le marché sont autant de difficultés auxquelles le grand public peut se heurter. Ainsi à l'heure actuelle, la comparaison des données générées par ce type de systèmes à des valeurs sanitaires n'est pas envisageable. L'emploi de ces données qualitatives pour une hiérarchisation et/ou un suivi des évolutions temporelles des teneurs particulières sur quais pourrait être envisagé pour certains systèmes candidats à condition que des précautions métrologiques puissent être mises en place. Pour atténuer les biais métrologiques, les dispositions à prendre doivent permettre d'identifier, évaluer et maîtriser au mieux les sources d'incertitude qui peuvent peser sur la donnée par le dispositif employé. Il revient ainsi aux opérateurs et utilisateurs de ce type de dispositif de s'assurer d'une démarche qualité en adéquation avec le contexte et les objectifs pour lesquels ce type de dispositif peuvent être déployés avec par exemple la mise en place de phases de calibration ou de vérification des données acquises en regard de valeurs prises pour référence.

Si à l'heure actuelle, les performances de ces systèmes capteurs sont limitées, il existe cependant un fort potentiel de progrès au niveau des principes de détection (avec des solutions issues de la recherche) et des améliorations notables autour de la mise en œuvre des éléments sensibles (optimisation des aspects fluidiques, l'intégration de filtres fonctionnalisés, limitation des interférences chimiques ou physique...). Les avancées récentes également en matière de traitement des données avec les changements de paradigmes liés à l'intelligence artificielle sont également d'autres voies d'amélioration de la qualité de l'information délivrée par les systèmes capteurs. L'ensemble de ces perspectives constituent autant de raisons pour poursuivre les efforts pour :

- (i) l'amélioration des dispositifs par les développeurs/fabricants/intégrateurs de capteurs ;
- (ii) la poursuite des études d'évaluation des performances des systèmes capteurs pour l'usage individuel dans les EFS, l'accompagnement/sensibilisation des utilisateurs aux bons usages, à la compréhension et l'interprétation des informations générées par les systèmes capteurs du côté des acteurs publiques ;
- (iii) proposer, sur la base de ce premier retour d'expérience avec ce type de systèmes en EFS, une étude ouverte à l'ensemble des fabricants/intégrateurs/distributeurs volontaires de systèmes capteur compatibles avec les objectifs de hiérarchisation des atmosphères rencontrées en EFS.

5 Références

- [1] PNSE4, (2021) Un environnement, une santé : 4e Plan National Santé Environnement. 80 pages, <https://sante.gouv.fr/IMG/pdf/pnse4-2.pdf>
- [2] Ineris (2020). "Recommandations pour la réalisation de mesures harmonisées de la qualité de l'air dans les enceintes ferroviaires souterraines." Rapport Institut national de l'environnement industriel et des risques; Contract No.: DRC-19-152419-04847A. 39 p. ([lien](#))
- [3] OpinionWay pour ADEME, (2022). Attitude des Français à l'égard de la qualité de l'air et de l'énergie. 9ème vague de l'enquête annuelle sur la qualité de l'air, les énergies renouvelables et les économies d'énergie dans le logement. 50 pages.
- [4] Lewis, AC and Edwards, P. (2016). Validate personal air-pollution sensors, *Nature*, 535, 29-31.
- [5] N Castell et al. (2016) Can Commercial Low-Cost Sensor Platforms Contribute to Air Quality Monitoring and Exposure Estimates? *Environment International*. 99, 293-302.
- [6] Karagulian F., Barbieri M., Kotsev A., Spinelle L., Gerboles M., Lagler F., Redon N., Crunaire S., Borowiak A. (2019). Review of the Performance of Low-Cost Sensors for Air Quality Monitoring. *Atmosphere*, vol. 10: p. 1-41.
- [7] AC Rai et al. (2017) End-User Perspective of Low-Cost Sensors for Outdoor Air Pollution Monitoring *Science of the Total Environment*, 607-608, 691-705.
- [8] Anses (2022) Avis de l'Anses relatif à l'utilisation de micro-capteurs pour le suivi de la qualité de l'air intérieur et extérieur. Rapports d'expertise collective (Saisine n° « 2018-SA-0271 »). Maisons-Alfort : Anses, 318 p.
- [9] Redon N., Allard L., Debert C., Gabet S., Hanoune B., et al. (2022) Micro-capteurs pour le suivi de la qualité de l'air intérieur et extérieur - Volume 1 De l'usage de micro-capteurs à des fins d'évaluation de l'exposition individuelle. : Avis de l'Anses Rapports d'expertise collective. Saisine n° 2018-SA-0271, Anses.318 p. ([hal-03875171](#)).
- [10] Fishbain, B., Lerner, U., Castell, N., Cole-Hunter, T., Popoola, O., Broday, D. M., ... & Bartonova, A. (2017). An evaluation tool kit of air quality micro-sensing units. *Science of the Total Environment*, 575, 639-648.
- [11] Schneider, P., Bartonova, A., Castell, N., Dauge, F.R., Gerboles, M., Hagler, G.S.W., Hügline, C., Jones, R.L., Khan, S., Lewis, A.C., et al., (2019). Toward a unified terminology of processing levels for low-cost air-quality sensors. *Environmental Science & Technology*, 53, 8485-8487, <http://dx.doi.org/10.1021/acs.est.9b03950>.
- [12] Williams, R., Duvall, R., Kilaru, V., Hagler, G., Hassinger, L., Benedict, K., Rice, J., Kaufman, A., Judge, R., Pierce, G., et al., (2019). Deliberating performance targets workshop: Potential paths for emerging PM2.5 and O3 air sensor progress. *Atmospheric Environment: X*, 2, 100031, [doi:10.1016/j.aeaoa.2019.100031](https://doi.org/10.1016/j.aeaoa.2019.100031).
- [13] Crunaire, S., Redon, N. et Spinelle, L. (2018). 1er Essai national d'Aptitude des micro-Capteurs (EAμC) pour la surveillance de la qualité de l'air : Synthèse des résultats. *Rapport LCSQA*.
- [14] Certification micro-capteurs : « AIR QUALITY SENSOR ». <https://www.ineris.fr/fr/risques/dossiers-thematiques/qualite-air/qualite-air-ambiant/certification-micro-capteurs-air>
- [15] Spinelle, L., Aleixandre, M., & Gerboles, M. (2013). Protocol of evaluation and calibration of low-cost gas sensors for the monitoring of air pollution. *Publications Office of the European Union: Luxembourg*.
- [16] Spinelle, L., Gerboles, M., & Aleixandre, M. (2015). Performance evaluation of amperometric sensors for the monitoring of O3 and NO2 in ambient air at ppb level. *Procedia engineering*, 120, 480-483.
- [117] Spinelle, L., Gerboles, M., Villani, M. G., Aleixandre, M., & Bonavitacola, F. (2017). Field calibration of a cluster of low-cost commercially available sensors for air quality monitoring. Part B: NO, CO and CO2. *Sensors and Actuators B: Chemical*, 238, 706-715.

[18] Borrego, C., Costa, A. M., Ginja, J., Amorim, M., Coutinho, M., Karatzas, K., ... & Penza, M. (2016). Assessment of air quality microsensors versus reference methods: The EuNetAir joint exercise. *Atmospheric Environment*, 147, 246-263.

[19] World Meteorological Organization (2021). An update on low-cost sensors for the measurement of atmospheric composition. WMO. Geneva, CH.

[20] South Coast AQMD, Air Quality Sensor Performance Evaluation Center (AQ-SPEC). <http://www.aqmd.gov/aq-spec>

[21] Duvall, R., Clements, A., Hagler, G., Kamal, A., Kilaru, V., Goodman, L., ... et Dye, T. (2021). Performance Testing Protocols, Metrics, and Target Values for Fine Particulate Matter Air Sensors: Use in Ambient, Outdoor, Fixed Sites, Non-Regulatory Supplemental and Informational Monitoring Applications. *US EPA Office of Research and Development*.

6 Liste des annexes

Liste des annexes :

Annexe A : Formule d'estimation des métriques

Annexe A : Formule d'estimation des métriques

1.1 Moyennes horaires et/ou journalières

La moyenne sur la période considérée (horaire ou journalière) est évaluée en imposant une exigence de complétude des données (encore appelée taux de recouvrement des données) d'a minima 70% sur chaque intervalle selon la formule :

$$x_{ik}^h = \frac{1}{n} \cdot \sum_{m=1}^n c_{mk} \quad (1)$$

avec :

- x_{ik}^h , valeur moyennée sur l'intervalle d'intérêt (horaire h ou journalier j) pour la i^{ème} valeur de depuis le début des essais et le k^{ème} instrument en $\mu\text{g}/\text{m}^3$.
- n, nombre de mesures à considérer pour l'évaluation de la moyenne sur l'intervalle.
- c_{mk} , m^{ème} valeur de l'intervalle considéré pour le k^{ème} instrument en $\mu\text{g}/\text{m}^3$.

1.2 Moyennes (horaires et/ou journalières) par microcapteur candidat

Pour chaque intervalle et chaque microcapteur, la valeur moyenne des cinq dispositifs déployés est évaluée à partir de :

$$\bar{x} = \frac{1}{M} \cdot \sum_{k=1}^M \left[\frac{1}{N} \cdot \sum_{i=1}^N x_{ik} \right] \quad (2)$$

avec :

- M, le nombre de dispositifs déployés (ici, M = 5).
- N, le nombre de mesures pour l'évaluation sur un intervalle donné.
- x_{ik} , i^{ème} valeur considérée sur l'intervalle pour le k^{ème} dispositif.

1.3 Précision

La précision entre capteurs identiques peut être appréhendée en considérant deux métriques : l'écart-type de mesure pour une méthode candidate σ ou encore le coefficient de variation (CV). Ces paramètres peuvent être évalués en utilisant les formules suivantes :

$$\sigma = \sqrt{\frac{1}{(N \times M) - 1} \cdot \sum_{j=1}^M \left[\sum_{t=1}^N (x_{tj} - \bar{x}_t)^2 \right]} \quad (3)$$

avec :

- σ , l'écart-type des mesures fournies par l'ensemble des capteurs pour des conditions d'exposition identiques pour l'ensemble des capteurs en $\mu\text{g}/\text{m}^3$.
- M, le nombre de dispositifs déployés (ici, M = 5)
- N, le nombre de mesures pour l'évaluation sur un intervalle donné
- x_{tk} , valeur moyenne des concentrations pour le t^{ème} intervalle et le k^{ème} dispositif en $\mu\text{g}/\text{m}^3$.
- \bar{x}_t , valeur moyenne sur l'ensemble des dispositifs pour le t^{ème} intervalle en $\mu\text{g}/\text{m}^3$.

$$CV = \frac{\sigma}{\bar{x}} \cdot 100 \quad (4)$$

avec :

- CV, le coefficient de variation en %.
- σ , l'écart-type des mesures fournies par l'ensemble des capteurs pour des conditions d'exposition identiques pour l'ensemble des capteurs en $\mu\text{g}/\text{m}^3$.
- \bar{x} , la moyenne des concentrations fournies par les microcapteurs lors de la campagne de mesure

1.4 Biais et linéarité

A ce stade, un modèle de régression reliant les valeurs fournies par les méthodes prises pour référence et candidates est requis. Le modèle de régression linéaire est communément utilisé dans ce type de situations, il permet d'obtenir une relation de la forme :

$$y = a \cdot x + b \quad (5)$$

avec :

- y , la variable dépendante et constituée par les données issues des méthodes candidates, en $\mu\text{g}/\text{m}^3$;
- x , la variable indépendante, constituée par l'ensemble des données fournies par les méthodes prises pour référence, en $\mu\text{g}/\text{m}^3$;
- (a, b) , la paramètres du modèle exprimés respectivement sans unité et en $\mu\text{g}/\text{m}^3$.

Pour inférer les paramètres a et b , deux types de régression seront retenues dans le cadre de cette étude : la régression linéaire ordinaire et la régression linéaire orthogonale. La première vise à optimiser l'erreur quadratique entre les valeurs observées et le modèle en jouant sur les valeurs a et b . S'il s'agit de la méthode la plus communément utilisée et implémentée, le résultat obtenu est sensible aux valeurs extrêmes et aberrantes. Se faisant, il a été choisi de travailler également avec la régression orthogonale qui optimise la distance des points de mesure à la courbe de régression.

A titre d'information, la Figure 14 présente visuellement la différence entre les deux types de régression considérés notamment les distances considérées entre les valeurs observées et les modèles.

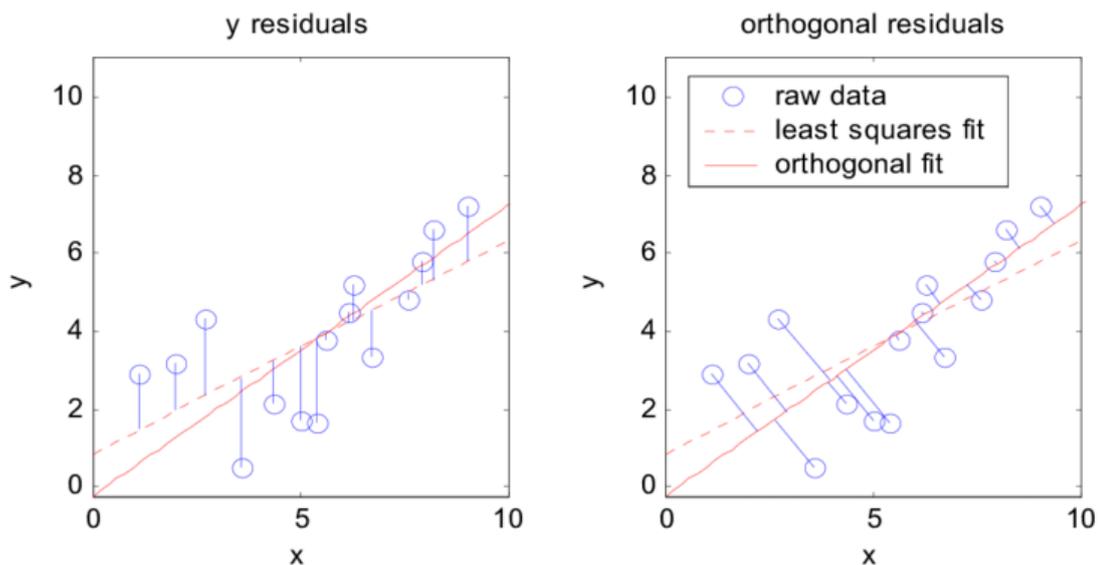


Figure 14. Illustration des distances considérées dans les deux types de régression linéaire considérés dans l'étude.

Outre les paramètres a et b , le coefficient de détermination (R^2) est également évalué lors des régressions. Ce paramètre permet pour des modèles de régression similaires de différencier les capteurs ou les méthodes candidates. En effet, un R^2 plus élevé indique en règle générale des données plus précises.

1.5 Erreur

Pour qualifier l'erreur entre sur les résidus (différences entre les valeurs prédites par le modèle et les valeurs observées), deux critères ont été sélectionnés :

- la racine de l'erreur quadratique moyenne (en anglais, *root-mean-square error* ou *RMSE*) a été sélectionné :

$$RMSE = \sqrt{\frac{1}{N \times M} \cdot \sum_{i=1}^M \left[\sum_{t=1}^N (x_{tj} - r_t)^2 \right]} \quad (6)$$

avec :

- M, le nombre de dispositifs déployés (ici, M = 5)
 - N, le nombre de mesures pour l'évaluation sur un intervalle donné
 - x_{tj} , valeur fournie par le j^{ème} microcapteur pour le t^{ème} intervalle en $\mu\text{g}/\text{m}^3$.
 - r_t , valeur fournie par la méthode prise pour référence pour le t^{ème} intervalle en $\mu\text{g}/\text{m}^3$.
- la racine de l'erreur quadratique moyenne normalisée (en anglais, *normalised root-mean-square error* ou *NMSE*) :

$$NMSE = \frac{RMSE}{\bar{r}} \cdot 100 \quad (7)$$

avec :

- \bar{r} , la valeur moyennée des mesures fournies par la méthode prise pour référence sur l'ensemble de la durée des essais.

