



HAL
open science

Suivi dynamique en Temps Réel de la qualité de l'air Intérieur dans un environnement de Bureaux

Olivier Ramalho, Rachid Ouaret, Anda Ionescu, Eline Le Ponner, Yves
Candau

► **To cite this version:**

Olivier Ramalho, Rachid Ouaret, Anda Ionescu, Eline Le Ponner, Yves Candau. Suivi dynamique en Temps Réel de la qualité de l'air Intérieur dans un environnement de Bureaux. DSC/OBS-QAI/2016-017, Centre scientifique et technique du bâtiment (CSTB), 84 avenue Jean Jaures BP2, 77447 Marne-la-Vallée cedex 2; Université Paris-Est Créteil. 2016, pp.151. hal-04911493

HAL Id: hal-04911493

<https://cstb.hal.science/hal-04911493v1>

Submitted on 24 Jan 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Open licence - etalab

CSTB
le futur en construction

TRIBU
**Suivi dynamique en Temps Réel de la
qualité de l'air Intérieur dans un
environnement de Bureaux -
**Contributions des sources et Modèle
prévisionnel****

Olivier RAMALHO, Rachid OUARET, Anda IONESCU,
Eline LE PONNER, Yves CANDAU

Rapport final n° DSC/OBS-QAI/2016-017

Mars 2016

Convention de subvention : 12-MRES-PRIMEQUAL-4-CVS-09

TRIBU

Suivi dynamique en Temps Réel de la qualité de l'air Intérieur dans un environnement de Bureaux - Contributions des sources et Modèle prévisionnel

RAPPORT Final n° DSC/OBS-QAI/2016-017

Olivier RAMALHO

Responsable scientifique (olivier.ramalho@cstb.fr)

Eline LE PONNER

Université Paris-Est, Centre Scientifique et Technique du Bâtiment, Direction Santé Confort, Marne-la-Vallée

Rachid OUARET

Université Paris-Est Créteil, Centre d'Études et de Recherche en Thermique, Environnement et Systèmes EA 3481

Anda IONESCU

Yves CANDAU

Avec la précieuse collaboration de Lucille LABAT, Bruno BERTHINEAU et Guillaume WYART

Et sur une idée initiale de Jacques RIBERON

Mars 2016

Convention de subvention : 12-MRES-PRIMEQUAL-4-CVS-09

Toute reproduction ou représentation intégrale ou partielle, par quelque procédé que ce soit, des pages publiées dans le présent document, faite sans l'autorisation du CSTB est illicite et constitue une contrefaçon. Seules sont autorisées, d'une part, les reproductions strictement réservées à l'usage du copiste et non destinées à une utilisation collective et, d'autre part, les analyses et courtes citations justifiées par le caractère scientifique ou d'information de l'œuvre dans laquelle elles sont incorporées (Loi du 1er juillet 1992 - art. L 122-4 et L 122-5 et Code Pénal art. 425).

Toute citation d'extraits ou reproduction doit obligatoirement faire apparaître la référence de ce document sous la forme : Ramalho O., Ionescu A., Ouaret R., Le Ponner E., Candau Y. *TRIBU. Suivi dynamique en temps réel de la qualité de l'air intérieur dans un environnement de bureaux - Contributions des sources et Modèle prévisionnel*, CSTB, Mars 2016, 151 pages. [DSC-OBS-QAI/2016-017].

© 2016 CSTB

Résumé

L'environnement intérieur réel est sujet aux fluctuations climatiques, mais aussi et surtout aux fluctuations pas toujours aléatoires imposées par l'homme, occupant les locaux, et ses activités. Alors que les modèles physico-chimiques parviennent assez bien à représenter l'évolution de différentes espèces chimiques dans un environnement contrôlé, la présence de l'homme rend particulièrement difficile l'application de ces modèles dans un environnement réel.

La mesure en continu des paramètres de qualité de l'air conjointement à des critères d'usages comme la présence ou l'ouverture des fenêtres peut permettre d'appréhender le rôle central de l'homme sur la qualité de l'air intérieur de son environnement.

Ce projet cherche ainsi à anticiper l'instrumentation active future d'un environnement intérieur pour le suivi d'un ou plusieurs polluants. Il vise notamment à fournir une méthodologie d'exploitation des données pour identifier leurs sources de fluctuations et développer un modèle de prévision de la qualité de l'air. Cette méthodologie devrait permettre au gestionnaire de contrôler et éventuellement prévenir tout dépassement de niveau de concentration au regard de valeurs limites instaurées par ailleurs.

La démarche adoptée est de suivre avec un pas de temps fin (entre 1 et 20 min) l'évolution de la concentration de polluants cibles (formaldéhyde et particules en nombre) dans un espace de bureau normalement occupé sur au moins une année. Les mesures sont également réalisées en parallèle à l'extérieur. L'espace de bureau est instrumenté pour suivre en temps réel l'occupation, la gestion des ouvrants, et les paramètres climatiques intérieurs et extérieurs, ainsi que durant certaines périodes des polluants d'origine extérieure (oxydes d'azote, ozone, CO). Ce jeu de données avec un suivi fin est utilisé pour développer des modèles de fluctuations des sources de polluants et de prévision des niveaux de polluants.

La variabilité des concentrations intérieures de formaldéhyde et de particules a été déterminée. Le formaldéhyde ne présente pas un profil journalier type tant les journées diffèrent les unes des autres en terme d'évolutions de la concentration. La présence et l'ouverture des ouvrants ont une influence variable sur les niveaux observés, mais qui reste globalement faible sur l'année. La saisonnalité du formaldéhyde associée notamment aux fluctuations de température et d'humidité spécifique est plus marquée fenêtres fermées en absence des occupants.

La variabilité temporelle des particules est fonction de leur taille : les plus fines étant très peu impactées par la présence des occupants ou l'ouverture des fenêtres, les plus grosses essentiellement influencées par la présence des occupants et en particulier le fonctionnement d'une imprimante dans l'espace de bureau, enfin les particules intermédiaires sont plus influencées par l'action des occupants sur les ouvrants que leur présence seule.

La variabilité du taux de renouvellement a été estimée à partir d'une méthode automatisée de détection et de calcul des décroissances de la concentration en CO₂. Bien que parcellaire, ces données montrent la forte variation de ce paramètre au fil du temps notamment durant les périodes de présence des occupants.

Les associations observées entre les concentrations et différents paramètres qui parfois évoluent de manière conjointes aux mêmes périodes soulignent la nécessité de déconvoluer leur contribution. A ce titre, quatre méthodes de séparation aveugle des sources ont été utilisées pour extraire les différentes sources de fluctuations des concentrations de particules et estimer leur contribution : l'ACI (analyse par composantes indépendantes), la NMF (non-negative matrix factorization), la PMF (positive matrix factorization) et l'ACP (analyse par composantes principales). Les composantes ou facteurs identifiées représentent autant de sources de fluctuations séparées. Mis à part l'ACP, les différentes méthodes montrent une similitude dans les sources extraites contributives du niveau de particules observé.

Un modèle de prédiction de l'ouverture des fenêtres par arbre de classification CART a été développé à partir des données calendaires, climatiques, et de présence. Malgré des performances optimales, il doit être retravaillé pour améliorer la prédiction avec efficacité de l'ouverture des fenêtres d'une année sur l'autre.

Pour prévoir l'évolution de la concentration intérieure des polluants, un modèle de prévision hybride a été développé en couplant un processus de décomposition spectrale de la concentration mesurée dans le temps par plusieurs FFT successives avec un modèle de prévision autorégressif à seuil (FFT/SETAR) ou un modèle basé sur la théorie des systèmes dynamiques (FFT/Chaos). L'horizon de prévision obtenu dans ces conditions est d'une dizaine d'heures.

Lorsque la saisonnalité de la concentration est suffisamment marquée, une décomposition saisonnière associée à un modèle autorégressif simple (STL/ARIMA) peut suffire et conduire selon le niveau de bruit de la saisonnalité à des horizons de prévision de plusieurs dizaines d'heures.

Cette étude a finalement cherché à montrer avec un minimum de paramètres mesurés comment la qualité de l'air dépend d'autres facteurs tous rattachés au temps et comment cette structure temporelle des données peut nous aider à mieux la comprendre pour mieux la prédire. L'homme y joue un rôle central en perturbant les relations existantes entre les différents paramètres. Mieux comprendre son comportement, c'est mieux comprendre notre environnement et la manière dont il évolue.

Mots clés : formaldéhyde, particules, bureau, variabilité temporelle, contribution des sources, prévision

Abstract

The indoor environment is subject to climatic fluctuations but also to other not always random fluctuations created by human, his presence and activities. Whereas physicochemical models can achieve successful dynamic simulations of the concentration of several chemical species in a controlled environments, the presence of man in real environments makes the use of such models particularly difficult.

Real-time monitoring of indoor air parameters combined with criteria of use like human presence and windows openings can help to assess the key role of human on indoor air quality in his environment.

The present works aims to anticipate active instrumentation of the indoor environment to measure one or several pollutants. In particular, methods to identify pollutant sources and their temporal variation along with a method to forecast the indoor air level of pollutants are expected. These tools could help a building manager to control and prevent exceedance of predefined reference values to occur.

The strategy adopted is to monitor with a short timestep (1 to 20 min) the indoor concentration of target pollutants (formaldehyde and particles) in an open space office during at least one year in real conditions. Outdoor levels are also monitored. The office is also equipped to monitor occupancy, the opening of windows and both indoor and outdoor climatic parameters, along with several pollutants from outdoors (nitrogen oxides, ozone, CO) during specific periods. These data are used to develop a model to extract pollutant source and a model to forecast indoor levels.

The variability of indoor concentration of both formaldehyde and particles was determined. Formaldehyde did not exhibit a particular daily pattern so different are day profiles from one another. The influence of human presence and windows openings on observed indoor levels is not always the same. But the overall effect remains weak throughout a year. The seasonality of indoor formaldehyde concentration is more apparent in the absence of occupants and windows closed.

The temporal variation of particle levels depends on their size. Fine particle levels are not affected by human presence or windows openings. Coarse particles are strongly impacted by the presence of occupants and in particular the activity of a laser printer in the open space office and not as much with windows openings. Intermediate particle levels rely more on the occupants opening windows than its mere presence.

The variability of air exchange rate was estimated with an automatic algorithm that detects decrease sequences of CO₂ levels and calculates the appropriate AER. Despite being fragmentary, the data show the high variation of AER with time in particular during human presence.

The observed relationships between the indoor concentration of pollutants and the other parameters which covariate most of the time, stress the need to deconvolute

their contribution. As such, four blind source separation methods were used to extract independent sources of fluctuations and estimate their relative contribution: ICA (independent component analysis), NMF (non-negative matrix factorization), PMF (positive matrix factorization) and PCA (principal component analysis). The factors extracted represent independent sources of fluctuations. Apart from PCA, the different methods bestow components that showed similar contribution profiles.

CART classification tree was used to develop a predictive model of windows opening state based on calendar, climatic and human presence data. Despite optimum performance, it needs improvements to achieve better predictions of the windows opening state from one year to the other.

To forecast indoor concentration of pollutants, a hybrid model was developed combining spectral decomposition of the monitored concentration by multistage FFT with a non-linear forecasting model based either on threshold autoregressive models (FFT/SETAR) or based on dynamical systems theory (FFT/Chaos). The forecast horizon for indoor formaldehyde concentration with this model is about 10 hours.

When the concentration dynamic bestow a significant seasonality, a seasonal decomposition coupled with a simple autoregressive model (STL/ARIMA) is sufficient to reach a forecast horizon of several tens of hours depending on the noise level around the seasonality.

This study has tried to show with a limited number of monitored parameters how indoor air quality depends on other factors all tied to time and how the temporal structure of data can help us to better understand it and better predict it.

Keywords: formaldehyde, particulate matter, office, temporal variability, source apportionment, forecasting

1 Introduction

Avec l'avènement des objets connectés, le foisonnement de capteurs et à l'aube du bâtiment numérique, l'information en temps réel sera bientôt disponible pour tout un chacun sur des domaines très variés et en particulier celui de la qualité de l'air intérieur. Cette pléthore de données, au-delà de fournir la simple prise d'information à un instant donné, renferme au sein de sa structure une grande partie des éléments permettant de comprendre les fluctuations observées et d'anticiper celles à venir.

L'environnement intérieur réel est sujet aux fluctuations climatiques, mais aussi et surtout aux fluctuations pas toujours aléatoires imposées par l'homme, occupant les locaux, et ses activités. Alors que les modèles physico-chimiques parviennent assez bien à représenter l'évolution de différentes espèces chimiques dans un environnement contrôlé, la présence de l'homme rend particulièrement difficile l'application de ces modèles dans un environnement réel.

La mesure en continu des paramètres de qualité de l'air conjointement à des critères d'usages comme la présence ou l'ouverture des fenêtres peut permettre d'appréhender le rôle central de l'homme sur la qualité de l'air intérieur de son environnement.

La qualité de l'air intérieur est aujourd'hui prise en compte comme objet de surveillance dans la loi Grenelle. Par ailleurs, les premières valeurs de référence pour certains polluants cibles commencent à apparaître et ont abouti à l'élaboration de textes réglementaires. L'élaboration des premiers protocoles, ayant comme objectif de définir la stratégie de mesurage en vue de se positionner par rapport à ces valeurs de référence, a montré d'importantes lacunes en terme de suivi temporel de polluants dans les environnements intérieurs. Dans un environnement réel, les sources de fluctuations des concentrations, leurs amplitudes sur une année, restent très mal connues. Il devient dès lors difficile de positionner une mesure effectuée à un moment précis par rapport à une distribution des données par exemple sur une année. Les conséquences font qu'il devient difficile voire impossible de comparer la mesure à une valeur de référence court ou long terme. Par ailleurs, une mesure ponctuelle seule ne permet pas de remonter aux sources ce qui limite le champ d'action en cas de dépassement de la valeur de référence.

C'est également le cas pour l'air extérieur. Néanmoins, dans ce domaine, des modèles statistiques inverses sont régulièrement utilisés pour déterminer les parts contributives des différentes sources à un polluant mesuré donné (Ionescu & Candau, 2007 ; Coman et al., 2008 ; Gros et al., 2011 ; Gaymoz et al., 2011).

A notre connaissance, ces modèles n'ont été que très peu employés dans les environnements intérieurs sur des mesures long-termes, du fait essentiellement du manque de données dynamiques. Un état de l'art plus complet des données disponibles pour les environnements intérieurs est présenté en annexe.

Le présent projet cherche à anticiper l'instrumentation active future d'un environnement intérieur (qui pourrait être par exemple un espace d'accueil de populations sensibles) pour le suivi d'un ou plusieurs polluants et de fournir une méthodologie d'exploitation de systèmes de données en temps réel à moindre coût aboutissant à un modèle de prévision de la qualité de l'air pour permettre au gestionnaire de contrôler et éventuellement prévenir tout dépassement de niveau de concentration au regard de valeurs limites instaurées par ailleurs.

La démarche adoptée dans ce projet est de suivre avec un pas de temps fin (entre 1 et 20 min) l'évolution de la concentration de polluants cibles (formaldéhyde et particules fines) dans un espace de bureau normalement occupé (entre 10 et 20 personnes) sur une année. Les mesures seront également réalisées en parallèle à l'extérieur. L'espace de bureau est également instrumenté pour suivre en temps réel l'occupation (par la mesure de CO₂ et détecteurs de mouvements), la gestion des ouvrants (contacteurs informant de l'état ouvert ou fermé), les paramètres climatiques intérieurs (température, humidité, irradiance) et extérieurs (température, humidité, irradiance, pluviométrie, pression, vent), ainsi que des mesures complémentaires d'autres paramètres d'intérêt tels que : ozone, oxydes d'azote et CO (proxy du trafic et des processus de combustion).

Ce jeu de données avec un suivi fin sert de base pour déterminer les fluctuations des sources de polluants (matériaux, équipements, activités, extérieur) qui pourront être transposés à d'autres environnements en adaptant le scénario. Ces mêmes modèles pourront également être intégrés aux modèles de prévision des niveaux de polluants permettant ainsi d'anticiper tout évènement de forte concentration. Les résultats obtenus ont également été comparés à ceux observés pour d'autres environnements tels qu'une maison expérimentale et un bureau individuel dans lesquels la concentration de particules ou du formaldéhyde ont pu être suivis.

2 Objectifs de recherche

A partir des mesures réalisées en continu, il est envisagé de répondre aux objectifs suivants :

- Documenter la qualité de l'air dans un environnement tertiaire de type bureau en continu sur le long terme avec un pas de temps fin. Ces données sont caractéristiques d'un niveau de fond intérieur dans un environnement de type périurbain.
- Développer un profil général d'occupation d'un environnement de bureau.
- Déterminer l'influence de la dynamique d'occupation sur la dynamique de la concentration de polluants cibles.
- Calculer l'exposition réelle des occupants au regard de l'exposition estimée à partir de mesures intégrées sur 8h, 24h ou 5 jours ouvrés (simulant des mesures par prélèvements passifs). Cela permettra de déterminer l'incertitude sur le calcul de l'exposition dans un espace de bureaux liée à des mesures intégrées.
- Déterminer l'influence de l'ouverture des ouvrants extérieurs sur la dynamique de la concentration de polluants cibles
- Développer un modèle général d'ouverture de fenêtres.
- Déterminer la contribution des sources intérieures à la concentration de polluants cibles (formaldéhyde, particules par compteur optique).
- Déterminer l'évolution sur le long terme des paramètres puits de la concentration de polluants :
 - o Taux de renouvellement d'air
 - o Dépôt particulaire par gamme de taille (analyse automatisée des décroissances de concentration de particules en lien avec l'analyse automatisée des renouvellements d'air).
- Développer un modèle de prévision de la concentration intérieure à court ou moyen terme du formaldéhyde et des particules.

3 Méthodologie

3.1 Suivi de la qualité de l'air dans un espace de bureaux

Un espace de bureau (open space) accueillant entre 10 et 20 personnes situé à Champs-sur-Marne (Seine-et-Marne), et équipé d'un système d'extraction d'air sans balayage, a été instrumenté pour suivre différents paramètres de façon continue:

- Suivi en temps réel de la concentration intérieure en formaldéhyde (mesure par analyseur AL1021 Aerolaser) dans le conduit d'extraction toutes les minutes (20 minutes après multiplexage). Certainement, la mesure la plus lourde à mettre en place car nécessitant un changement des solutions tous les 6 jours environ.
- Suivi en temps réel de la concentration intérieure et extérieure en particules fines [0,3 – 20 µm] (mesure optique Grimm Dust Monitor 1.108) toutes les minutes. En 2015, les particules ultrafines (de 0,02 à 0,7 µm) ont été mesurées en intérieur (mesure par charge de diffusion, DISCmini).
- Instrumentation de l'ouverture des fenêtres et de la porte principale par des contacteurs (changements d'état ouvert ou fermé enregistrés).
- Instrumentation de l'occupation par détecteurs de mouvements (quantité de mouvement toutes les 10 secondes) et niveau de dioxyde de carbone dans l'air (évaluation de la production métabolique) toutes les minutes.
- Suivi des paramètres climatiques intérieurs (température, humidité relative, irradiance) et extérieurs (température, humidité relative, irradiance, pluviométrie, pression, vent) toutes les minutes.
- Suivi en temps réel de la concentration intérieure en NO_x, Ozone, CO (mesure par Microstation Environnement SA) sous couvert de disponibilité.
- Utilisation possible des données de pollution extérieure à partir des données des réseaux de surveillance (PM, NO_x, O₃, HAM, HAP) et données trafic.

La plupart des capteurs mis en place disposent d'une technologie de communication sans fils permettant de minimiser les interventions.

Le suivi dynamique a été réalisé de janvier 2012 et juin 2015.

3.1.1 Description de l'espace de bureaux instrumenté

Le bâtiment cible est situé à Champs-sur-Marne (Seine et Marne) sur le site du Centre Scientifique et Technique du Bâtiment, 30 km à l'est de Paris. Il comporte des bureaux, une salle de réunion, des laboratoires et une halle expérimentale. Le bâtiment comporte deux niveaux, les bureaux étant situés au second niveau. Le bâtiment est situé en zone urbaine (densité modérée) à 2,2 km de l'autoroute A4 et à 5 km de

l'autoroute A104 (à vol d'oiseau au point le plus proche). Il est également proche de la voie rapide D199 (600 m) et de la rue passante distante de 130 mètres.

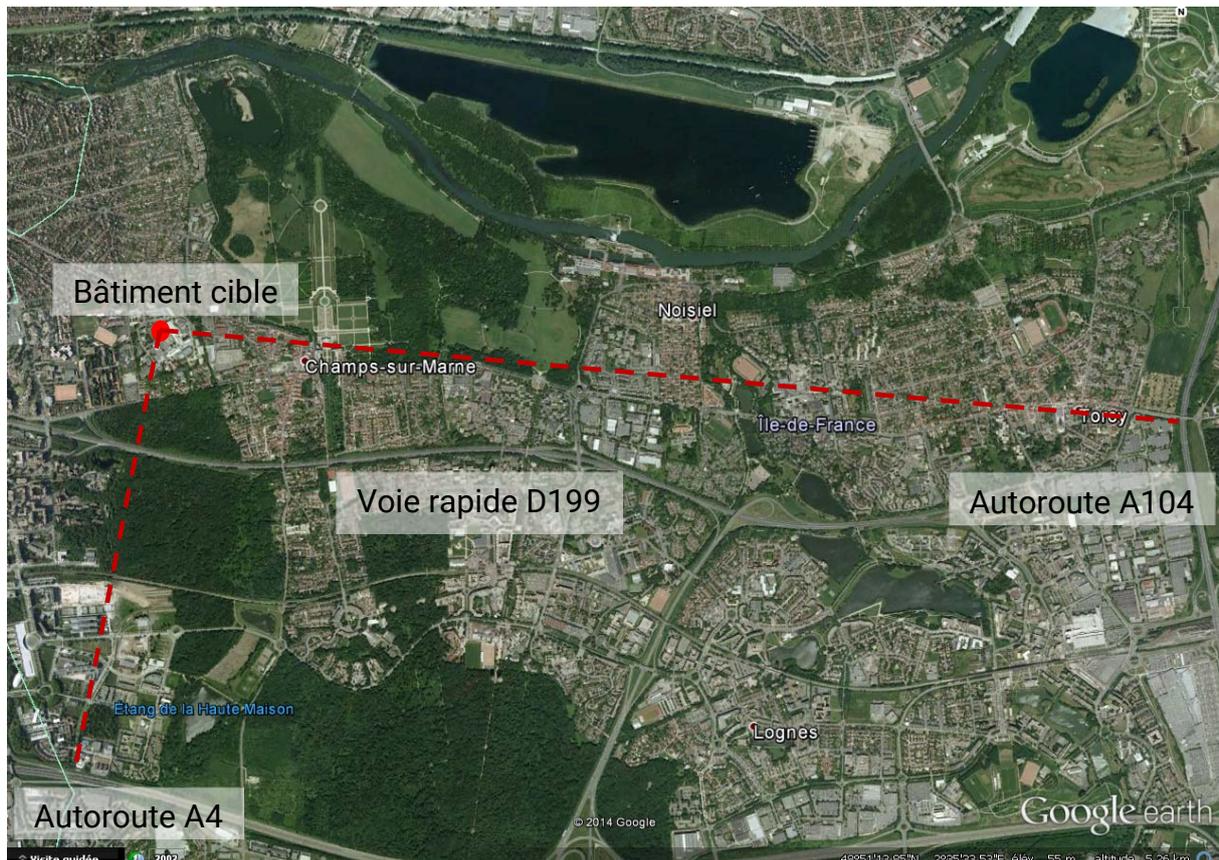


Figure 1 : Situation générale du bâtiment cible

L'espace instrumenté est un bureau paysager situé au 1^{er} étage et qui comprend également des bureaux individuels. Les cloisons de ces bureaux ont été montées après coup. Par conséquent, ces bureaux individuels font partie du même volume. Cet espace d'une surface de 132 m² et d'un volume total de 364 m³ accueille entre 6 et 12 personnes selon les jours et les périodes de l'année. Un système de ventilation simple flux assure un débit d'extraction d'air constant de 252 m³/h (mesuré en 2012 à $\pm 6\%$). En 2014, le débit d'extraction d'air était de 228 m³/h. Les 6 bouches d'extraction sont disposées dans les bureaux H1.03, H1.04 et H1.05 (voir plan). A l'opposé, 10 entrées d'air sont fixées sur les menuiseries des 5 fenêtres coulissantes. Les fenêtres sont exposées au sud-est. L'espace communique à travers le reste du bâtiment au travers d'une seule porte qui mène à une pièce de circulation menant soit à la halle expérimentale, soit à une petite cour extérieure permettant d'atteindre le toit, soit aux sanitaires, soit à l'aile gauche du bâtiment.

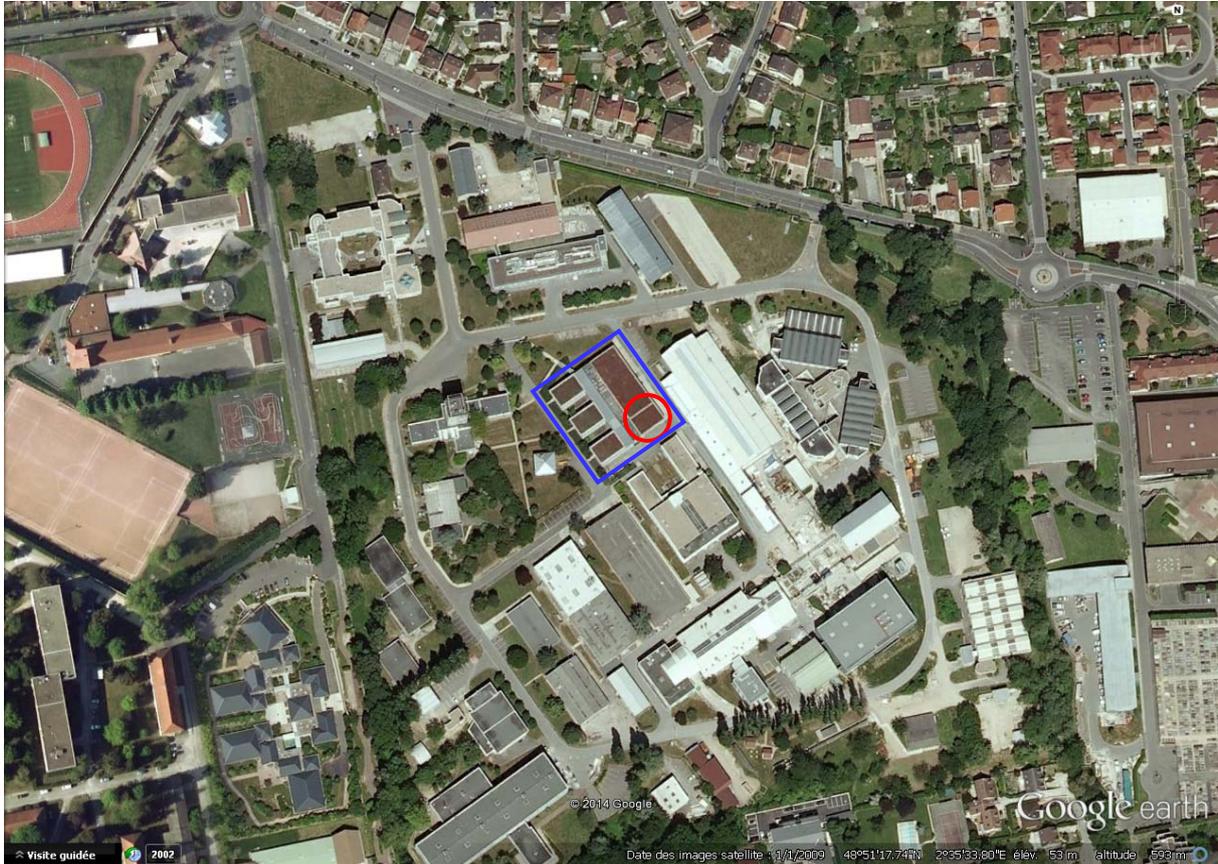


Figure 2 : Vue rapprochée du bâtiment cible avec en bleu, le bâtiment cible et en rouge la localisation de l'espace instrumenté.

Les murs sont peints et une moquette couvre le sol de l'espace. Le faux plafond en dalles est situé à 2,75 mètres. Le mobilier est composé de bureaux en bois compact mélaminé, d'armoires en aluminium et de chaises en polypropylène. Le nombre de postes informatiques actifs varie avec les occupants et une imprimante-copieur multifonction est située à proximité du bureau H1.05.

L'entretien du local est limité au passage de l'aspirateur généralement en fin de semaine le vendredi en fin de journée. La moquette a subi un nettoyage par voie humide poussé avec une shampoineuse durant les vacances de fin d'année en 2013.



Figure 3 : Vue de l'espace de bureau depuis la porte d'entrée. La configuration des tables varie avec le nombre d'occupants de cet espace.

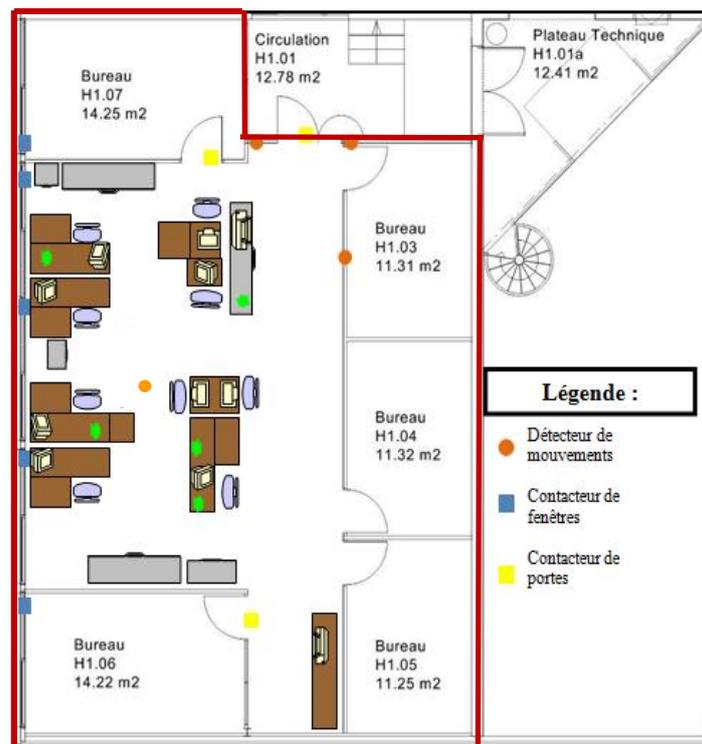


Figure 4 : Plan de l'espace de bureau (132 m², 364 m³).

Plusieurs paramètres sont mesurés en continu dans cet espace. Les instruments et détecteurs sont soit disposés dans le volume central de l'espace ou bien mesurent directement dans le conduit d'extraction. Les paramètres mesurés sont listés ci-après :

- La concentration de formaldéhyde (HCHO) par réaction de Hantzsch en phase liquide
- La concentration en nombre de particules fines (0,3 à 20 μm) par un compteur optique.
- La concentration en particules ultrafines (0,02 à 0,7 μm) par un compteur à diffusion de charge (en 2015)
- L'ouverture des fenêtres par détecteur d'ouverture (CSTBox)
- L'ouverture des portes par détecteur d'ouverture (CSTBox)
- La détection de mouvement par capteur passif infrarouge (CSTBox)
- Un compteur d'impulsion électrique branché sur l'imprimante multifonctions
- L'irradiance par un solarimètre
- La température
- L'humidité relative
- La concentration en dioxyde de carbone (CO_2) par capteur NDIR
- La concentration en oxydes d'azote (NO et NO_2) (2013 seulement)
- La concentration en ozone (O_3)
- La concentration en monoxyde de carbone (CO)

Par ailleurs, plusieurs paramètres sont également suivis à l'extérieur soit par le biais de la station météorologique permanente du site placée sur le toit du bâtiment cible, soit par l'intermédiaire d'instruments protégés dans un boîtier environnemental. La liste des paramètres mesurés à l'extérieur est détaillée ci-après :

- La concentration de formaldéhyde (via une ligne de prélèvement)
- La concentration en particules fines par compteur optique dans un boîtier environnemental dédié
- Les paramètres climatiques (température, humidité relative, vitesse du vent, direction du vent, pression atmosphérique, irradiance, détecteur de pluie)
- La concentration de CO_2
- Les concentrations de NO, NO_2 et O_3 sont récupérées à partir de la station de surveillance AIRPARIF de Lognes située à 3,4 km au sud-est du bâtiment cible.

3.1.2 Occupation et ouvrants

Pour recueillir des informations complémentaires en continu sur l'environnement occupé, les ouvertures de portes et fenêtres ont été instrumentées ainsi que l'occupation de l'espace par l'intermédiaire d'un dispositif spécifique, la CSTBox (CSTB

sensing and Telemonitoring Box). Il est ainsi possible d'appréhender certaines variables relatives aux usages pour déterminer leur impact sur les concentrations de polluants dans l'air intérieur.

La CSTBox est un dispositif extensible pour la collecte, l'analyse et le contrôle des données du bâtiment associée à un réseau de capteurs, d'actionneurs ou d'interacteurs. La CSTBox se matérialise par un mini PC industriel et un logiciel développé en licence libre par le CSTB. Le réseau, extensible à la demande, se compose de capteurs, d'actionneurs et de dispositifs connectables à la CSTBox. Elle a été développée par le CSTB de Sophia Antipolis dans un cadre gérontologique (programme Gerhome¹ ; Zouba et al., 2009 ; Zarli et al., 2010) pour la surveillance des activités afin notamment de détecter des dérives anormales des habitudes de vie, pouvant être des symptômes de pathologies ou de dégénérescence globale, ou la prévention des accidents domestiques, par la détection de situation à risque.

Des modules de détection d'ouverture d'ouvrants et des détecteurs de mouvement sans fil ont été associés à la CSTbox.

3.1.2.1 Détection d'ouverture des fenêtres et portes (CSTBox-DIN4)

Ces modules sont dotés de 4 entrées logiques pouvant accepter des contacts secs (interrupteurs à lames souples ou ILS, relais,...) ou tout dispositif fournissant un signal logique TTL 3V. Dans le cas présent, des ILS sont montés au niveau des ouvrants pour détecter la position des portes et fenêtres. Les modules fonctionnent sur alimentation autonome (3 accus 1,2V ou 2 piles alcalines 1,5V) et sur une durée théorique estimée à 2 ans pour une stimulation correspondant à l'utilisation habituelle de portes ou fenêtres. Ils communiquent par ondes radio leur état à la CSTBox (Module radio Xbee série 1, bande ISM (2.4GHz), protocole 802.15.4).



Figure 5 : Détecteur d'ouverture des fenêtres : à gauche, position fermée, au centre, position ouverte et à droite la CSTBox qui recueille l'ensemble des informations.

¹ <http://gerhome.cstb.fr>

Avec ce type de capteur, il est difficile d'appréhender le degré d'ouverture de la fenêtre (% de surface ouverte). Le module de détection ne fournit qu'une donnée binaire ouvert / fermé pour chaque pas de temps. Chaque module est programmé pour renvoyer une information dès lors qu'un changement d'état survient, mais les informations sont recodées pour un pas de temps de 1 minute. Pour une fenêtre donnée, plusieurs changements d'état peuvent parfois survenir dans la même minute. Dans ce cas, c'est le dernier état enregistré qui l'emporte.

A l'échelle de l'espace instrumenté, plusieurs possibilités d'agrégation de cette information existent :

- Déclarer les fenêtres ouvertes si au moins une fenêtre sur les cinq est dans un état ouvert pour chaque minute
- Compter le nombre (ou le ratio) de fenêtres ouvertes à chaque minute
- Introduire ou non les conditions d'ouverture des portes des bureaux H1.06 et H1.07 pour prendre en compte ou non l'état d'ouverture de ces deux fenêtres. Si les portes sont fermées, il est possible de considérer que l'état d'ouverture de ces fenêtres ne vienne pas impacter de manière fondamentale sur le reste du volume.

3.1.2.2 Détection de mouvement (CSTBox-PIR)

Ces modules sont équipés d'un capteur passif infrarouge (PIR) capable de détecter un mouvement dans une demi-sphère presque complète. Les modèles sélectionnés sont en outre à très basse consommation (quelque μA en veille) ce qui permet de les utiliser sans problème avec une alimentation autonome (3 accus 1,2 V ou 2 piles alcalines 1,5 V). Ils communiquent par ondes radio leur état à la CSTBox (Module radio Xbee série 1, bande ISM (2.4 GHz), protocole 802.15.4). Lorsqu'aucun mouvement n'est détecté, aucune trame n'est envoyée. Dès qu'un mouvement est détecté, la quantité est enregistrée pendant 10 secondes. Une trame contenant un tableau de 10 échantillons est envoyée. Le capteur se met ensuite en veille profond pendant 1 seconde avant qu'une nouvelle détection soit enregistrée.



Figure 6 : Détecteur de mouvement disposé au niveau du faux plafond ou sur les cloisons des bureaux individuels.

A partir des enregistrements de variations de quantité de mouvement (donnée brute), les données sont recodées au pas de temps de 1 minute en sommant les informations.

La quantité de mouvement ne représente toutefois pas le nombre d'occupants, une personne pouvant produire autant de quantité de mouvement que deux ou plus. Aussi, cette donnée ne peut être interprétée que de façon binaire : "0" pour absence d'occupation et "1" pour présence d'occupants dans le volume. Sur les 3 détecteurs de mouvement utilisés, l'occupation est déclarée effective pour chaque minute où au moins un des 3 détecteurs a signalé une quantité de mouvement différente de 0.

3.1.3 Mesure du formaldéhyde en continu

3.1.3.1 Principe de fonctionnement

La mesure du formaldéhyde a été réalisée en continu par l'analyseur AL4021 (Aerolaser GmbH). Le principe de détection de cet appareil est basé sur la réaction de Hantzsch correspondant à la cyclisation de la 2,4-pentanedione avec de l'ammoniac et le formaldéhyde transféré de façon continue de la phase gazeuse à la phase aqueuse pour former la 3,5-diacetyl-1,4-dihydrolutidine (ou DDL). Ce dernier est coloré et fluorescent. Excité à 400 nm, il émet un rayonnement fluorescent à 510 nm.

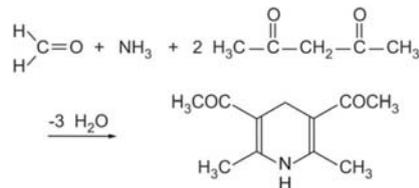


Figure 7 : Réaction de Hantzsch pour former la DDL

L'air est prélevé à un débit de 1 L/min et mis en contact avec la solution de stripping (eau distillée tamponnée) prélevée à un débit de 500 ml/min au sein d'une boucle d'échange (stripper) maintenue à 10 °C (pour une efficacité proche de 100%). La très forte solubilité du formaldéhyde permet de transférer l'intégralité du formaldéhyde contenu dans l'air vers la solution aqueuse. Cette solution est ensuite mise en contact avec le réactif de Hantzsch (prélevé en continu) au sein d'un réacteur maintenu à 68 °C. Le produit de réaction, la DDL est détectée au moyen d'un fluorimètre qui présente l'avantage par rapport à une méthode colorimétrique d'être plus sensible, plus rapide et du fait de temps de réaction courts d'être moins sensible aux interférences d'autres aldéhydes ou cétones. Il comprend une lampe UV-LED pour l'excitation de la molécule à 410 nm et un photomultiplicateur pour la détection qui convertit le signal en concentration.

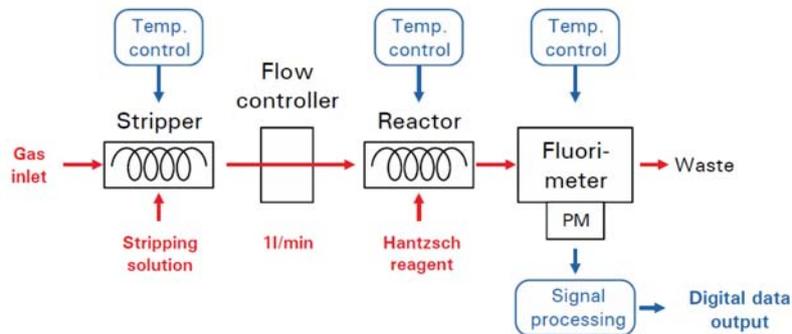


Figure 8 : Schéma simplifié du processus de mesure.

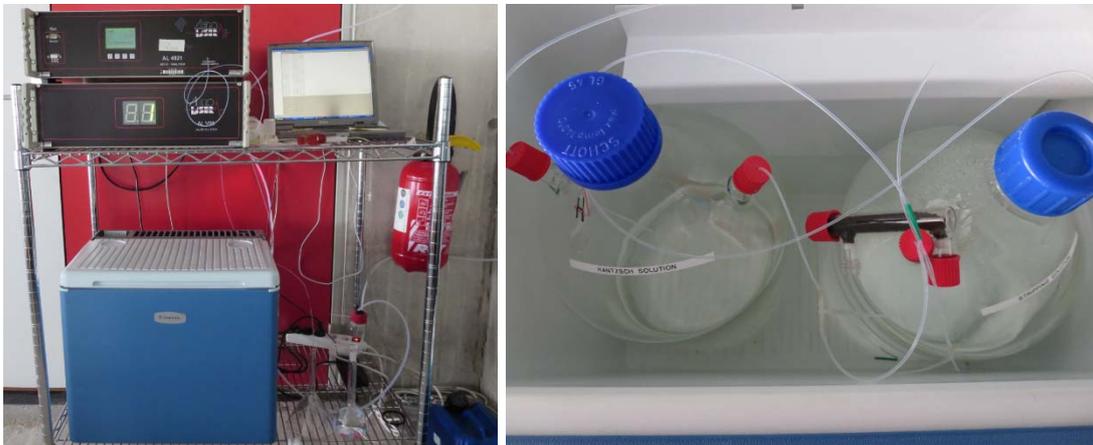


Figure 9 : Vue globale de l'analyseur AL4021 avec le module de séquençage des voies de mesure et les solutions de stripping et réactif dans la glacière.

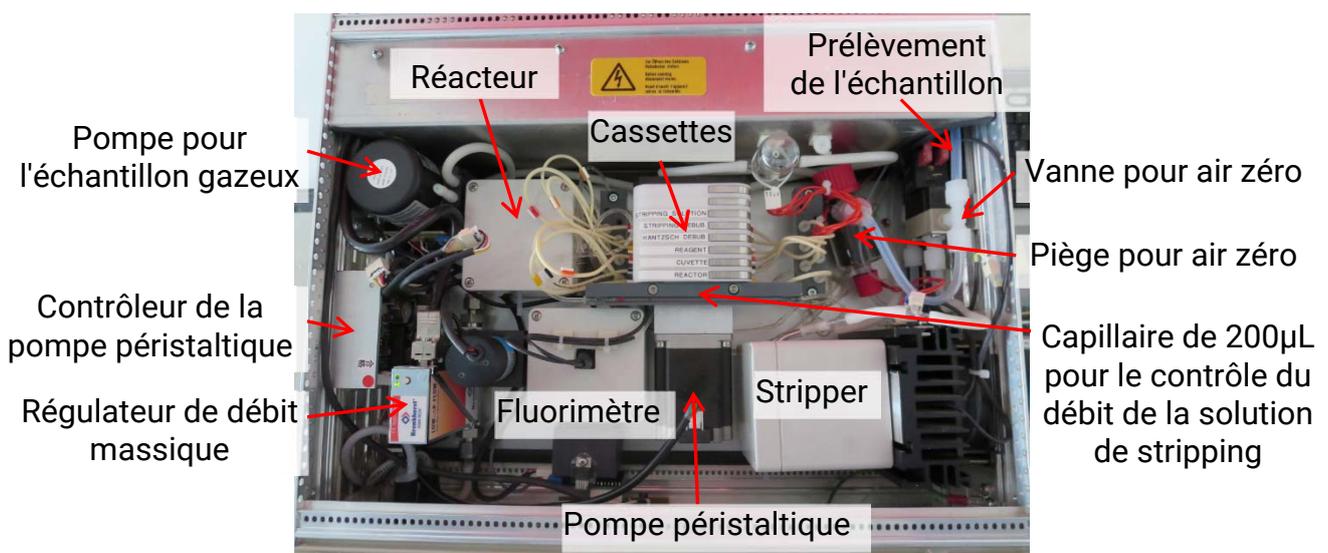


Figure 10 : Vue détaillée de l'analyseur de formaldéhyde AL4021. Les cassettes et les tubes ismaprène® de différents diamètres intérieurs permettent de réguler le débit des liquides.

L'appareil fonctionne selon 2 modes opératoires possibles pour la mesure du formaldéhyde dans l'air :

- en mode manuel en étalonnant l'appareil à partir d'une solution aqueuse à concentration connue de formaldéhyde qui va définir la plage de mesure. C'est la méthode de mesure la plus précise, mais elle nécessite de vérifier périodiquement la réponse de l'appareil de façon manuelle. C'est ce mode de mesure qui est retenu dans cette étude. Les débits des solutions de solubilisation (stripping) et de prélèvement d'air fixent le coefficient de conversion de la concentration en phase aqueuse à la concentration en phase gazeuse.
- en mode automatique, en étalonnant l'appareil à partir d'une concentration de formaldéhyde en phase gazeuse générée par un four à perméation interne. Dans cette configuration, l'appareil peut s'autocalibrer à une fréquence prédéfinie (ici toutes les 24 heures) et vérifier le zéro toutes les 4 heures. Durant ces phases, l'appareil ne mesure pas le formaldéhyde ce qui génère des trous dans le suivi continu.

Le signal zéro augmente au cours du temps en mode automatique. L'appareil ajuste alors de lui-même la sensibilité, ce qui réduit la plage de mesure effective. Le signal reste lui toujours compris entre 0 et 5 V. Il est alors nécessaire de refaire manuellement un étalonnage liquide pour reparamétrer l'appareil, généralement après plus de 3 semaines. Les solutions de piégeage et de réactif sont quant à elles changées tous les 5 jours. Les solutions peuvent être préparées à l'avance et conservées au froid.

Les spécifications de la méthode de mesure sont données dans le tableau ci-après.

Tableau 1 : Spécifications de l'analyseur de formaldéhyde AL4021 (Aérolaser)

Spécification	générale	gamme choisie
Domaine de mesure	100 ppt – 3 ppm (gaz) 150 ng/L – 5 mg/L	0,1 – 50 ppb 0,15 – 100 µg/L
Limite de détection	100 ppt (gaz) 150 ng/L (liquide)	
Signal de sortie	0 à 5 V	
Bruit	2 % pleine échelle	
Temps de réponse	Délai de réponse : 5 min Temps de réponse : 1min30 (10 % - 90 %)	
Interférences	H ₂ O ₂ < 1 : 100 Glyoxal < 1 : 123 O ₃ < 1 : 500 Acétaldéhyde < 1 : 10 000 Acroléine < 1 : 10 000	NH ₃ < 1 : 10 000 Benzaldéhyde < 1 : 20 000 Propanal < 1 : 20 000 Méthanol < 1 : 50 000
Absence d'interférences	Acétone SO ₂ NO ₂	NO Isobutane

3.1.3.2 Etalonnage

L'instrument est étalonné au moyen d'une solution liquide de formaldéhyde autour de 100 µg/L préparée à partir d'une solution mère de formaldéhyde à 37 %. Cet étalonnage liquide sert ensuite à définir le taux de perméation de la source interne de formaldéhyde (maintenue à 45 °C) utilisé pour l'étalonnage en phase gazeuse.

La concentration de formaldéhyde en phase gazeuse dépend directement de la concentration en phase liquide et des débits de prélèvement d'air et de la solution de stripping selon la formule suivante :

$$C_{gaz} = \frac{[HCHO]_{liq} \times q_{stripper} \times V_M}{q_{air} \times M}$$

Avec C_{gaz} la concentration de formaldéhyde en phase gazeuse exprimée en ppb, $[HCHO]_{liq}$ la concentration de formaldéhyde en phase liquide exprimée en µg/L, $q_{stripper}$ le débit de la solution de stripping en mL/min, q_{air} le débit de prélèvement d'air en L/min, V_M le volume molaire en L/mol et M la masse molaire en g/mol.

Dans ces conditions, une concentration liquide de formaldéhyde de 100 µg/L équivaut à une concentration dans l'air de 40 ppb pour un débit de solution de stripping de 0,5 mL/min, un débit de prélèvement d'air de 1 L/min et une pression atmosphérique de 1 atm à 20 °C.

Une comparaison de la mesure de formaldéhyde par l'analyseur Aérolaser avec un prélèvement actif sur cartouches Sep-Pak DNPH a été réalisé à partir d'émissions de différents matériaux (bois aggloméré, contreplaqué, planche OSB, stratifié, lingettes nettoyantes, et colles cellulose) et produits émettant du formaldéhyde à des niveaux très variables et testés en chambre expérimentale. La mesure donne des résultats très satisfaisants sur une grande gamme de concentration.

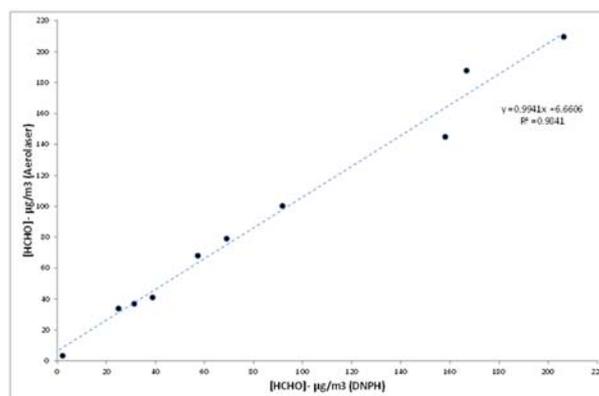


Figure 11 : Comparaison de la mesure de formaldéhyde par l'analyseur Aerolaser et de la mesure active sur cartouches Sep-Pak DNPH pour des données d'émission de différents matériaux et produits testés en enceinte.

3.1.3.3 Multiplexage

De manière à suivre l'évolution de la concentration de formaldéhyde en différents points, l'instrument a été associé à un module de multiplexage (AL V08, Aerolaser GmbH) permettant de suivre jusqu'à 8 voies de façon séquentielle. Dans notre cas, seules 2 voies sont utilisées l'une vers le conduit d'extraction de l'espace de bureaux et l'autre vers l'extérieur. Compte tenu du temps de réponse de l'instrument et des lignes de prélèvement (en PTFE® longueur 8 m et id 6,35 mm, soit 250 ml de volume mort), et après quelques tests, le changement de voie est effectué toutes les 10 minutes. La dernière valeur de chaque période correspondant à une moyenne sur une minute est conservée. Ainsi, en un point de prélèvement donné, la concentration de formaldéhyde est renseignée toutes les 20 minutes.

Le module de multiplexage n'a été acquis qu'en mars 2014. Au préalable pour 2013, l'instrument a été utilisé de manière continue en enregistrant une moyenne toutes les minutes mais pour un seul point de mesure. Les données mesurées à l'extérieur pour 2014 ont été toutefois invalidées car les débits de prélèvement étaient un peu faibles et les mesures perturbées par la voie intérieure. Seules les données extérieures de formaldéhyde de janvier à juin 2015 sont disponibles et valides.

3.1.3.4 Incertitude de mesure

L'incertitude de la mesure dépend de la mesure de la concentration de formaldéhyde en phase liquide, du débit de prélèvement d'air, du débit de la solution de stripping, ainsi que de l'évolution de la température qui affecte le calcul du volume molaire.

L'erreur associée à la mesure de la concentration de formaldéhyde en phase liquide dépend du bruit de l'instrument (au plus 2 % du signal) et de la sensibilité fonction de la solution d'étalonnage utilisée (pour 100 µg/L, la sensibilité est de l'ordre de 0,04 V/µg/L). Cette erreur représente dans ces conditions ± 2 µg/L pour un niveau de 100 µg/L.

Le débit de prélèvement d'air est enregistrée par l'instrument et ne varie que très peu ($\pm 0,1$ %). Le débit de la solution de stripping entre deux périodes d'étalonnage varie de $\pm 1,5$ %. Il est ajusté à chaque étalonnage.

Au sein d'une même période où l'instrument est étalonné, l'incertitude absolue est de $\pm 1,7$ ppb pour un niveau de 40 ppb en intérieur pour des faibles variations de température comprises entre 18 et 22 °C. L'incertitude est due pour 47 % au bruit du signal, pour 35 % à l'erreur sur le débit de la solution de stripping, et pour 16 % associés aux fluctuations de température. L'erreur sur le débit de prélèvement d'air ne représente que 2 %.

Pour une mesure extérieure, l'incertitude absolue est plus importante, $\pm 4,5$ ppb pour un niveau de 40 ppb pour des variations de température comprises entre -10 et 35 °C. L'incertitude est due pour 68 % aux fluctuations de température, pour 18 %

au bruit de l'instrument et pour 13 % à l'erreur sur le débit de la solution de stripping, et 1 % associé au débit de prélèvement d'air.

L'incertitude relative est donc estimée à $\pm 4,3$ % pour une mesure intérieure et à $\pm 11,3$ % pour une mesure extérieure entre deux étalonnages de l'appareil.

3.1.4 Mesure des particules en continu par un compteur optique

3.1.4.1 Principe

Le compteur optique Dust Monitor modèle 1.108 (GRIMM) mesure en continu le nombre de particules par tranche de taille (diamètre) et par litre d'air prélevé. L'appareil utilise la diffraction lumineuse pour compter les particules présentes dans l'échantillon d'air prélevé qui traversent le faisceau lumineux (diode laser monochromatique). La distribution en taille (15 classes de tailles entre 0,3 et 20 μm) des particules prélevées est obtenue en quantifiant la dispersion angulaire causée par le passage de particules de taille variable au travers d'un rayon lumineux produit par une diode laser monochromatique sur un angle de 60-120°. Le débit de l'appareil est de 1,2 L/min. L'instrument permet de compter les particules jusqu'à 2000 particules/cm³ sans effets de coïncidence avec une sensibilité de 0,001 particule/cm³ et une reproductibilité de 2%.

Le compteur optique utilise 2 puissances laser pour effectuer ses mesures. Entre 0,3 et 2 μm , c'est la forte puissance laser qui est utilisée. Entre 2 et 20 μm , la plus petite puissance laser est utilisée. La mesure à 2 μm est réalisée 2 fois à forte et petite puissance laser et le résultat est la moyenne des 2 valeurs obtenues. Lors de l'étalonnage de chaque appareil, l'erreur maximale tolérée est de 10 % entre 0,3 et 2 μm et de 20 % entre 2 et 20 μm .

Une mesure est réalisée toutes les minutes en 3 points à l'aide de 3 instruments : deux points dans l'espace de bureaux (l'un proche de l'imprimante multifonctions et l'autre situé dans la diagonale opposée) et un point à l'extérieur sur le toit du bâtiment. L'instrument situé à l'extérieur est abrité dans un boîtier environnemental dédié.

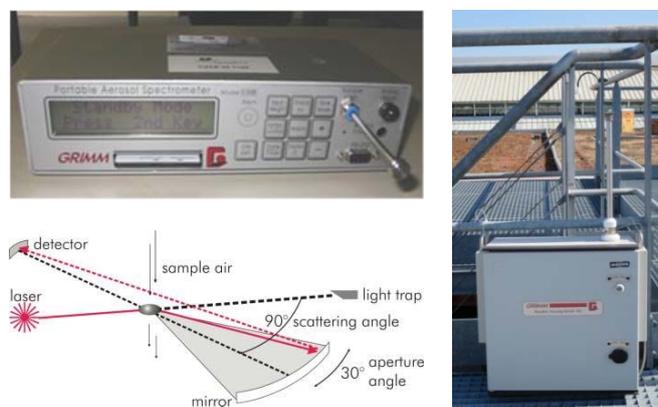


Figure 12 : Le compteur optique Dust Monitor 1.108, son principe et le boîtier extérieur.

Le prélèvement extérieur a été mis en place fin août 2013. L'instrument positionné au niveau de l'imprimante multi-fonctions est défectueux et en cours de remplacement.

Les concentrations de particules sont enregistrées toutes les minutes dans une carte mémoire d'une autonomie de plus de 45 jours. Les données sont récupérées environ tous les 15 jours.

3.1.4.2 Comparaisons entre les instruments

Les compteurs utilisés ont été comparés dans l'espace de bureau en 2013 et en 2014. Les résultats montrent une bonne corrélation des données pour une gamme de taille allant jusqu'à 7,5 μm . Pour les particules de diamètre supérieur, la corrélation des résultats entre les deux instruments est moins marquée, du fait de valeurs mesurées très faibles par ailleurs. Entre 0,3 et 7,5 μm , la dispersion des mesures entre les deux instruments, exprimée par l'écart relatif moyen, varie entre 1,4 % pour les particules entre 1 et 1,6 μm et 15 % pour les particules entre 2 et 3 μm .

Compte tenu des différences observées entre les instruments, les concentrations en nombre de particules supérieures à 7,5 μm sont difficilement comparables. En effet la comparaison des instruments dans une ambiance n'est pas adéquate pour les plus grosses particules sujettes à des variations locales importantes du fait des faibles mouvements d'air et ce malgré la présence de ventilateurs d'homogénéisation.

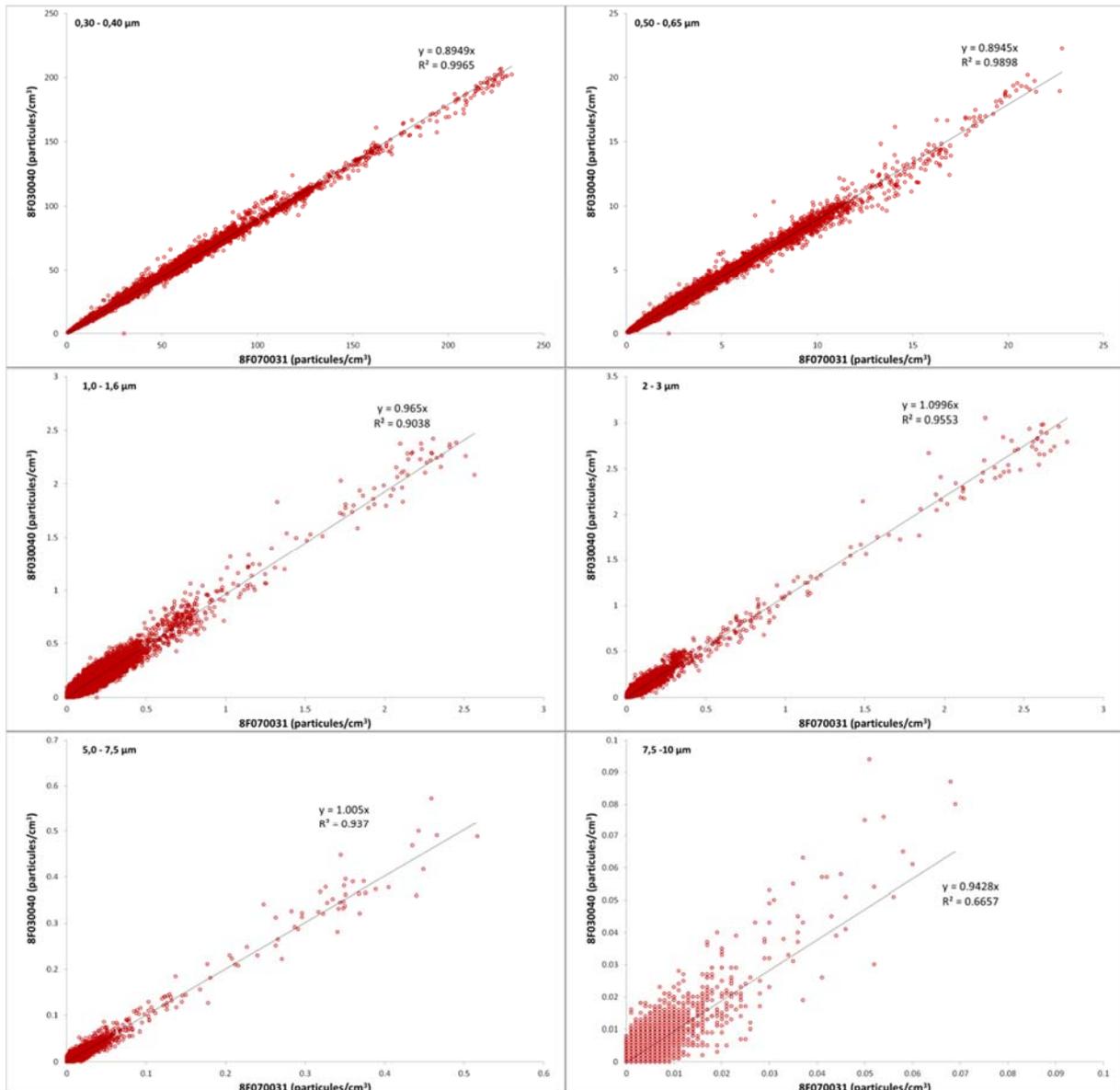


Figure 13 : Comparaison de deux compteurs optiques pour différentes tailles de particules (2013).

En 2014, l'intercomparaison a montré des écarts plus conséquents entre les instruments jusqu'à un facteur 2 pour certains canaux. Les mesures ont été ajustées en conséquence de façon rétrospective. Aucune correction n'a par contre été apportée pour les grosses particules compte tenu de la difficulté de comparer les instruments pour cette gamme de taille.

3.1.5 Mesure des particules ultrafines par diffusion de charge

En 2015, pour pallier à l'absence d'information sur les particules ultrafines, nous avons pu utiliser deux compteurs à diffusion de charge (DISCmini, Matter aerosol) permettant de mesurer la concentration en nombre de particules de 0,02 à 0,7 μm . L'un était disposé à proximité de l'imprimante multifonctions et l'autre à l'opposé à proximité d'une fenêtre. Ces appareils ne peuvent par contre pas être utilisés en extérieur.

L'instrument charge l'aérosol dans un chargeur à diffusion d'ions unipolaires. Les particules chargées passent à travers un étage de diffusion et les particules restantes collectées à l'étage de filtration. Le courant électrique mesuré en fA à ces deux étages par un électromètre sensible est proportionnel au nombre de particules déposées. Un impacteur en entrée élimine les grosses particules ($> 0,7 \mu\text{m}$) pour éviter l'encrassement excessif des étages de collecte. Le courant mesuré peut également être converti en surface alvéolaire de dépôt dans les poumons (A-LDSA ou alveolar lung deposit surface area). Le diamètre moyen de particules (dans la gamme 0,02 – 0,7 μm) est également renseigné en continu. L'intercomparaison entre les deux instruments placés côte à côte dans une ambiance intérieure est excellente avec des écarts inférieurs à 10 % sur la période et une pente d'ajustement de 0,997. La comparaison du nombre de particules observées avec un compteur à noyaux de condensation CPC-3007 (gamme 0,1- 1 μm et corrigé des effets de coïncidence) en présence de fumée d'encens montre également de très bons résultats avec des niveaux légèrement supérieurs d'environ 20 %. Malgré la correction des effets de coïncidence, le niveau réel de particules peut être sous-estimé par le CPC-3007.

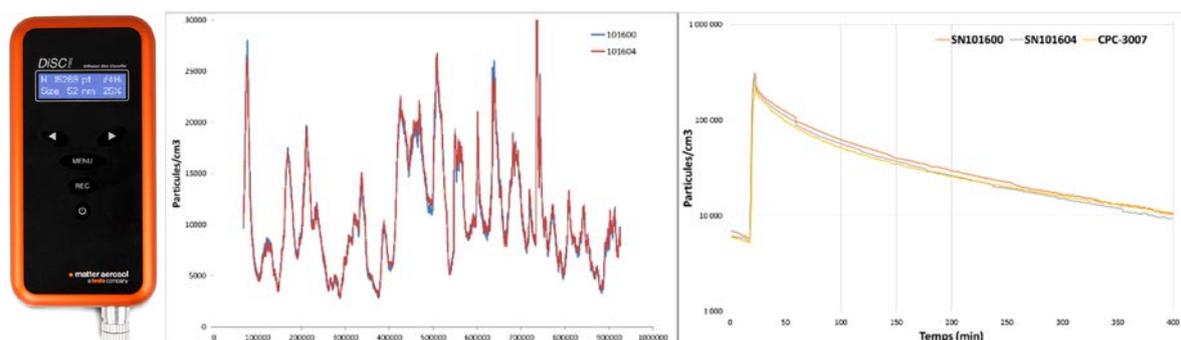


Figure 14 : Le compteur de particules à diffusion de charge DISCmini (à gauche), l'intercomparaison des deux instruments dans l'ambiance (centre) et avec le CPC-3007 (à droite).

3.1.6 Mesure complémentaire des oxydes d'azote et de l'ozone

3.1.6.1 Principe

Dans le cadre du projet, une micro station de mesure (MMS, Environnement SA) a été acquise pour mesurer la concentration à l'intérieur des oxydes d'azote (NO et NO₂), de l'ozone et du monoxyde de carbone. Ces mesures viennent en complément des mesures des deux paramètres principaux que sont le formaldéhyde et les particules. L'instrument combine trois modules de mesures : le module AC32M pour la mesure des oxydes d'azote par chimiluminescence (norme ISO 7996), le module O3 42M pour la mesure de l'ozone par photométrie dans l'ultraviolet (norme ISO 13964) et le module CO 12M pour la mesure du monoxyde de carbone par spectrométrie dans l'infrarouge en non dispersif (norme ISO 4224).

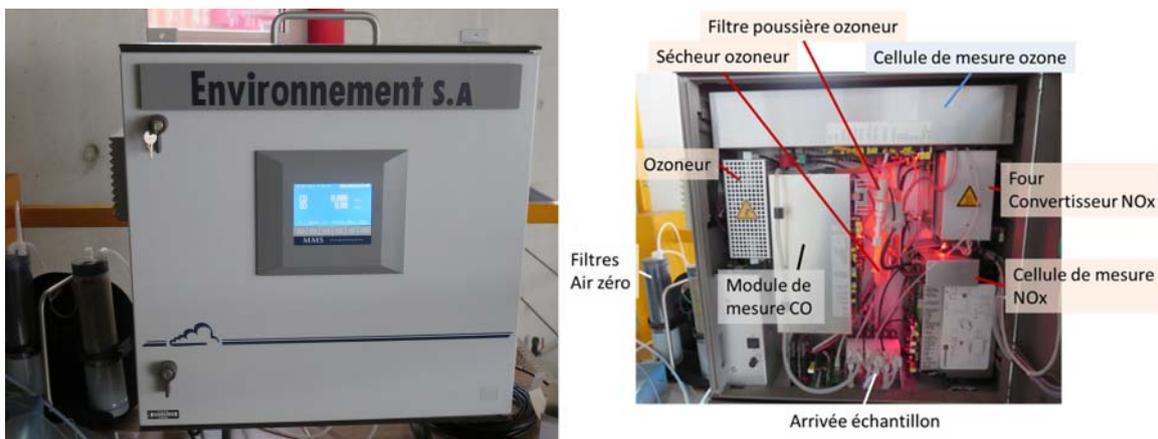


Figure 15 : Vue extérieure et intérieure de la micro station de mesure.

Les spécifications de l'instrument sont données dans le tableau ci-après. Le pas de temps d'acquisition est fixé à une valeur moyenne toutes les minutes. Une correction automatique du zéro pour la mesure du CO est réalisée quotidiennement et l'étalonnage de l'instrument est réalisé manuellement tous les mois.

Le point de mesure de ces paramètres est situé dans le conduit d'extraction de l'espace de bureaux. Les données sont enregistrées toutes les minutes dans la mémoire interne de l'appareil.

Aucun point de mesure n'est pris à l'extérieur, car l'instrument ne possède pas de voie de multiplexage et que les données de station du réseau de surveillance extérieure sont disponibles auprès d'Airparif. Les données de la station la plus proche (station urbaine de Lognes située à 3,4 km au sud-est de l'espace de bureaux) sont prises en compte comme données représentatives de l'air extérieur. Elles sont données en moyennes horaires et converties en ppb à 20 °C (température de référence).

Tableau 2 : Spécifications de la micro station de mesure MMS

Module	AC32	CO12	O342
Gammes de mesure	: programmables (50 ppm maximum)	: programmables (200 ppm maximum)	: programmables (10.00 ppm maximum)
Unités	: ppb, ppm ou µg/m ³ (programmable)	: ppm ou mg/m ³ (programmable)	: ppb, ppm ou µg/m ³ (programmable)
Bruit (σ):	: 0.2 ppb (temps de réponse automatique TR11)	: 0.025 ppm (temps de réponse automatique TR11)	: 0.2 ppb (temps de réponse automatique TR11)
Minimum détectable (2 σ)	: 0.4 ppb (temps de réponse automatique TR11)	: 0.05 ppm (temps de réponse automatique TR11)	: 0.4 ppb (temps de réponse automatique TR11)
Temps de réponse (0-90 %)	: 40 sec (programmable)	: 40 sec (programmable)	: 10-90 secondes (programmable)
Linéarité	: ± 1 %	: ± 1 %	: ± 1 %
Débit échantillon	: 44 Nl / h	: 55 Nl / h.	: 55 Nl / h
Débit Ozone (Module AC32M)	: 4 Nl / h		
Pression de la chambre	250 mbars absolu	1 bar absolu	1 bar absolu
Convertisseur NO ₂ → NO	Molybdène à : 340°C		
Température PM	Régulé à 10°C		
Température de la chambre de mesure	60°C	55°C	Interne

Les données d'oxydes d'azote n'ont pu être mesurées à l'intérieur qu'en 2013. L'instrument n'a pas pu fournir des mesures valides d'oxydes d'azote depuis ce moment, malgré plusieurs interventions en 2014 et en 2015. Seules les données CO et ozone sont restés valides en 2014 et 2015.

3.1.6.2 Comparaisons avec l'extérieur

En septembre 2013 durant 15 jours, les moyennes horaires données par l'instrument ont été comparées aux données de la station de Lognes (ou des stations urbaines de Paris Centre et d'Aubervilliers pour la comparaison du CO, les plus proches du site). Les résultats sont représentés à la Figure 16.

Une bonne adéquation des données d'ozone est obtenue entre l'instrument et la station de Lognes avec un coefficient de corrélation de Pearson de 0,93. L'instrument a tendance à donner des résultats plus faibles que ceux de la station (pente de 0,89). Pour NO₂, la corrélation est satisfaisante avec un coefficient de 0,83 mais l'instrument fournit des résultats plus faibles que ceux de la station (pente 0,68). Pour NO, le coefficient de corrélation est de 0,81 et l'instrument tend à sous-estimer également les valeurs par rapport à la station (pente de 0,59). Les données du CO restent en général très faibles au niveau de l'instrument et des stations de mesure, inférieures la plupart du temps à 0,5 ppm. Dans l'ensemble des situations, les données relatives au CO sont pratiquement à la limite de détection des instruments. Il est par conséquent délicat de comparer ces résultats pour des niveaux si faibles.

L'écart entre l'instrument et la station de Lognes peut être dû à la fois aux spécificités locales ou à celles de l'instrument. L'écart est jugé acceptable dans ces conditions.

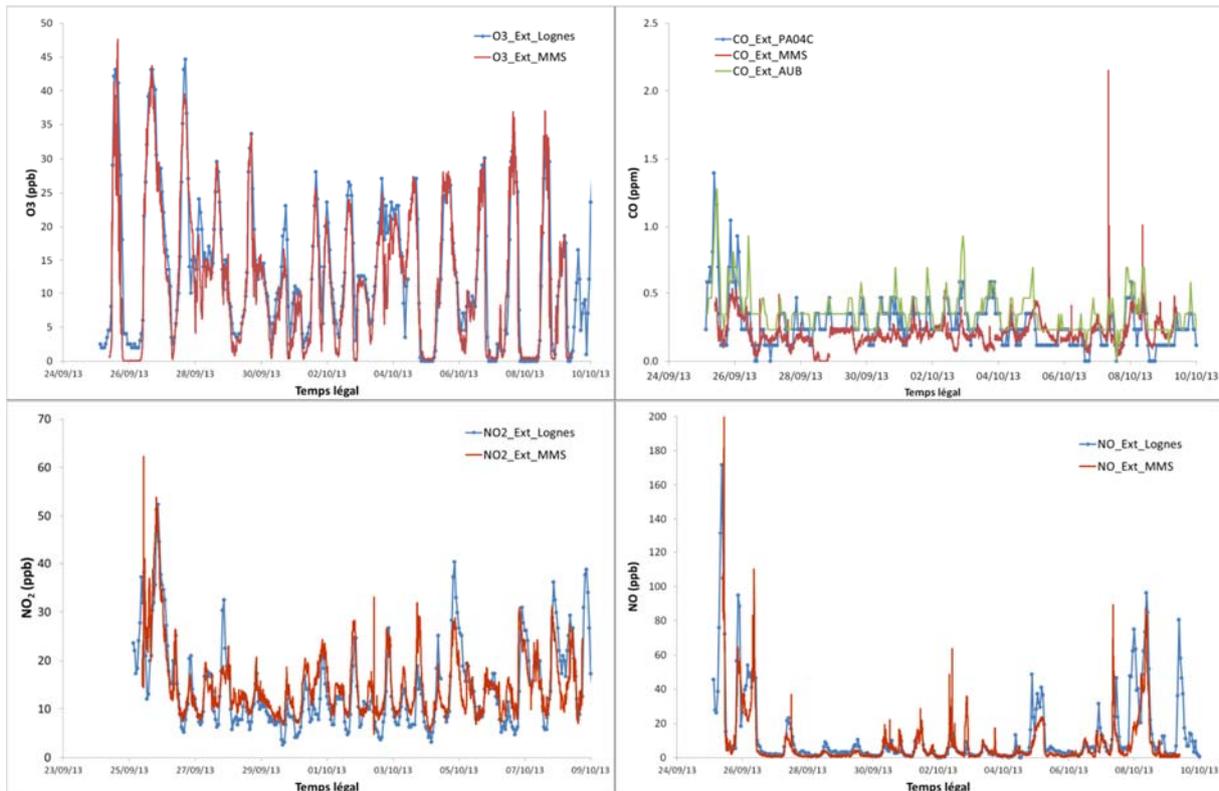


Figure 16 : Comparaisons des mesures MMS avec les données de station du réseau de surveillance extérieure en 2013 (station de Lognes ou de Paris et Aubervilliers pour le CO).

Fin 2014, une nouvelle comparaison des niveaux extérieurs a été menée. L'instrument fournit des résultats en ozone environ 30 % plus faibles que ceux de la station (sauf pour les minimales) mais l'évolution est pratiquement identique (pente de 0,70). La concentration de CO mesurée par le MMS montre une très bonne adéquation avec les mesures des stations de Paris centre et d'Aubervilliers avec notamment un pic de 1,5 ppm observé le 1^{er} janvier 2015.

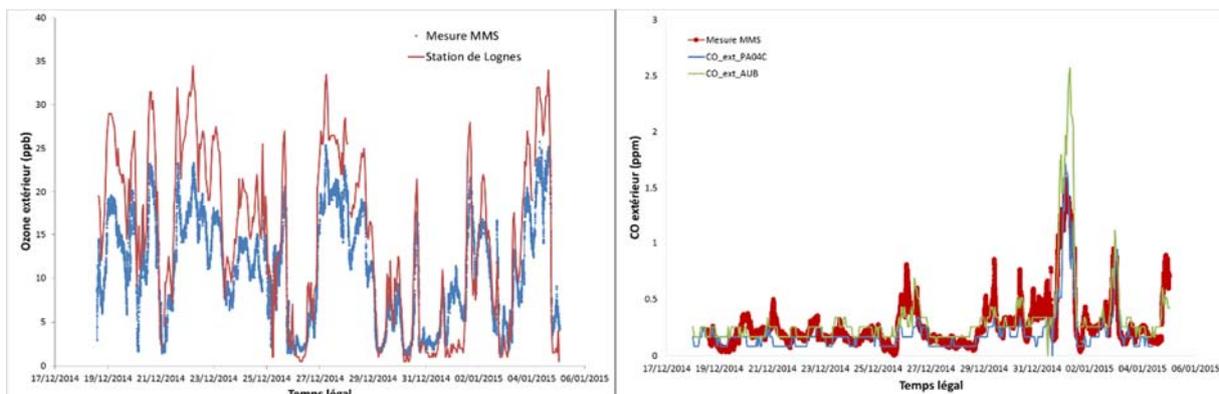


Figure 17 : Comparaisons des mesures MMS en ozone et CO avec les données de station du réseau de surveillance extérieure en 2014 (station de Lognes ou de Paris et Aubervilliers pour le CO).

3.1.7 Mesure complémentaire du dioxyde de carbone (CO₂)

Le dioxyde de carbone est un traceur de l'activité métabolique mais aussi des activités de combustion. Dans l'espace de bureau, en l'absence de sources de combustion, la seule source de CO₂ est l'activité métabolique associée à la présence humaine. La mesure de CO₂ est donc liée à aux périodes d'occupation, périodes durant lesquelles le niveau de CO₂ tend à augmenter. Cette mesure est donc indicative de la présence des occupants dans le local comme pour les détecteurs de présence, mais elle va également renseigner sur la présence d'autres sources (d'autres occupants) aux alentours du local, ce que les détecteurs ne permettent pas. Par ailleurs, le niveau de CO₂ dépend également du renouvellement de l'air du local. La mesure n'est donc pas totalement corrélée avec la présence des occupants. Mais, de ce fait l'analyse des fluctuations de la concentration de CO₂ en connaissant l'occupation des locaux permet de renseigner cette variable difficile à appréhender que représente le renouvellement de l'air. Grâce à cette mesure en continu, il est possible d'estimer ce renouvellement d'air à différents moments de la journée. De manière générale, le renouvellement d'air est déterminé ponctuellement de façon expérimentale en injectant un gaz traceur. Mais la variabilité du renouvellement de l'air n'est pratiquement jamais renseignée. Or l'usage des occupants, notamment au travers de l'ouverture des fenêtres et portes peut modifier considérablement cette variable au cours de la journée. Sa modulation va avoir un effet direct sur les fluctuations de la concentration de CO₂ et des autres polluants. Caractériser la variabilité du renouvellement d'air est donc essentiel.

La mesure du CO₂ dans l'espace de bureau est assurée par absorption infrarouge non dispersif (NDIR) avec une sonde Q-Track modèle 8550 (TSI inc.) qui enregistre toutes les minutes les valeurs moyennes de CO₂, de température et d'humidité relative. A l'extérieur, la mesure du CO₂ est mesurée par un instrument développé au CSTB sur le même principe (capteur NDIR) et utilisé dans le cadre de la campagne pilote de surveillance de la qualité de l'air dans les écoles et les crèches pilotée par la Direction Générale de la Prévention des Risques. L'instrument utilisé (prototype Lum'Air) enregistre la valeur moyenne de CO₂ toutes les 10 minutes. Il est disposé à l'intérieur du boîtier environnemental qui abrite le compteur de particules. Le boîtier est ventilé ce qui permet la mesure du CO₂ et l'abri de l'appareil par rapport aux conditions extérieures.

Les instruments préalablement étalonnés ont été comparés entre eux à 3 niveaux de concentration de CO₂. Le coefficient de variation entre les instruments est de 6 % à 400 ppm et autour de 1 % à 1000 et 2500 ppm.

Les instruments ont une plage de 0 à 5000 ppm (jusqu'à 6000 ppm pour la sonde Q-Track) et partagent la même incertitude de $\pm (50 \text{ ppm} + 3 \% \text{ de la lecture})$. Les instruments sont équipés d'un système de compensation de la dérive, mais sur le long terme, en l'absence d'interventions, il est tout de même nécessaire de corriger une dérive du zéro de l'ordre de +3 ppm par mois pour les prototypes Lum'Air.

3.1.8 Mesure de l'activité de l'imprimante multifonction

Un compteur d'impulsion électrique a été branché sur l'imprimante multifonction. Il permet de distinguer les périodes d'utilisation de l'imprimante des périodes de veille et des périodes d'arrêt. Les données sont enregistrées toutes les minutes en mémoire et récupérées par le logiciel MadgeTech data recorder v2.00.74 toutes les semaines.

3.1.9 Mesure de l'irradiance solaire

La mesure de l'irradiance solaire (entre 400 et 1100 nm de longueur d'onde en W/m^2) est réalisée à l'intérieur et à l'extérieur à l'aide d'un solarimètre SL 200 (Kimo). A l'intérieur, un seul point de mesure est réalisé à proximité des fenêtres mais à l'abri des rayonnements directs. Toutefois, ce seul point de mesure ne peut en aucun cas représenter l'irradiance globale dans l'espace de bureau. La mesure n'est qu'indicative, elle permet juste de donner une indication sur l'abattement de l'irradiance solaire à l'intérieur par rapport à l'extérieur. Cette indication peut sous réserves fournir une première information quant à l'activité photochimique dans l'espace de bureau au regard de l'extérieur, même si la fraction de rayonnement ultraviolet la plus énergétique n'est pas prise en compte.

Le solarimètre à l'extérieur est disposé en permanence sur le toit du bâtiment cible et fait partie intégrante de la station de mesure de paramètres climatiques qui y est fixée. La précision de mesure des deux instruments est de $\pm 5\%$. Une mesure moyenne intérieure comme extérieure est enregistrée toutes les minutes.

Une intercomparaison des 2 instruments a été réalisée durant 2,5 jours. Une corrélation linéaire de 0,91 est observée avec une pente de réponse de 0,93. L'écart des moyennes est de l'ordre de $1 W/m^2$ et autour de $70 W/m^2$ pour les valeurs les plus élevées.

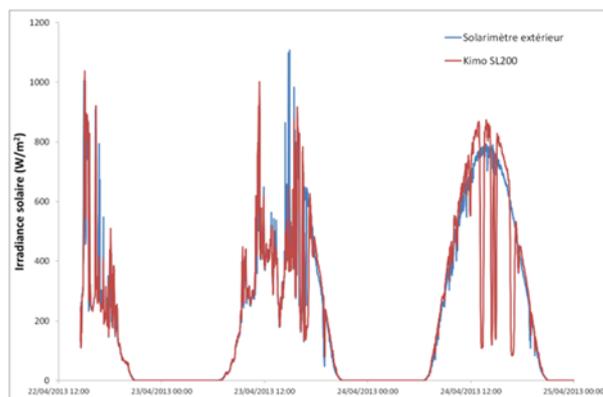


Figure 18 : Comparaisons des deux instruments de mesure de l'irradiance solaire. Les chutes d'irradiance observées le 3^{ème} jour sur le Kimo SL200 sont dues à des ombres parasites spécifiques.

3.1.10 Données climatiques

A l'extérieur, une station de mesure permanente située sur le toit du bâtiment cible permet d'enregistrer les valeurs moyennes de température, d'humidité relative, de vitesse et de direction du vent, de pression atmosphérique, et d'irradiance solaire de façon automatique. Un détecteur de pluie permet de renseigner l'activité pluvieuse au cours du temps de manière qualitative. En l'absence de pluie, le détecteur renvoie une valeur de 0 et à l'inverse renvoie la valeur de 5 dès qu'il reçoit les premières gouttes.

A l'intérieur, la température et l'humidité relative sont mesurées par la sonde Q-Track (TSI Inc.) toutes les minutes. A partir de la mesure de l'humidité relative, de la température et de la pression, l'humidité spécifique est calculée en utilisant la formule de Rankine pour approximer la pression de vapeur saturante nécessaire au calcul.

$$H_{abs} \left(\frac{g}{kg} \text{ air humide} \right) = \frac{HR}{100} \times \frac{M_{eau}}{M_{air}} \times e^{\left(13.7 - \frac{5120}{T+273.15} \right)} \times 1000$$

$$H_{spec} \left(\frac{g}{kg} \text{ air sec} \right) = \frac{H_{abs}}{(1000 - H_{abs})} \times 1000$$

Des pannes de courant et des problèmes de communication entre les instruments et la centrale d'acquisition ont été observés. Lors de ces épisodes, les données d'une station météo privée² située à Lagny-sur-Marne (8,5 km du site) ont été récupérées dans la mesure du possible (base d'acquisition au pas de temps de 5 min) pour compléter les valeurs manquantes. Ces données présentent une bonne adéquation avec les nôtres.



Figure 19 : Station permanente de mesure des paramètres climatiques sur le toit du bâtiment cible.

² <http://www.toutletempslaniaque.net> (avec l'aimable autorisation de Jean-Michel Guillemain)

3.1.11 Outils statistiques

De nombreux outils statistiques et logiciels ont été amenés à être utilisés dans ce projet :

- Le logiciel R avec ses nombreux packages, notamment :
 - l'excellent package « Openair » développé par David Carslaw (Carslaw et Ropkins, 2012 ; Carslaw, 2015), pour les variations selon les différentes échelles temporelles
 - package RTisean (Nonlinear time series analysis, Hegger et al., 1999) pour toute la modélisation autour des séries temporelles
 - tsDyn (Nonlinear time series models with regime switching)), l'analyse par composantes indépendantes (FastICA)
- Le logiciel SPAD version 7.4 (Coheris) pour les analyses factorielles notamment
- Le logiciel SAS version 9.1.3 pour les régressions
- Le logiciel CART (Salford systems) pour les arbres de classification
- Le logiciel XLSTAT (Addinsoft) pour différents tests de comparaison
- Le logiciel Matlab pour d'autres programmations

4 Résultats

4.1 Caractéristiques des mesures réalisées

4.1.1 Global

Le suivi dans l'espace de bureau s'est déroulé de fin janvier 2012 à début juillet 2015, avec un nombre de paramètres mesurés variables. Le bilan des paramètres mesurés par année est présenté au tableau ci-dessous.

Tableau 3 : Bilan des paramètres mesurés dans l'espace de bureaux de 2012 à 2015 en % de minutes de l'année couvert par chaque mesure. (En grisé : données aux pas de temps de 10 ou 20 minutes).

Paramètre mesuré	Unité	Environnement	2012	2013	2014	2015*	Cumul (jours)
Vitesse du vent	m/s	extérieur	13%	79%	100%	94%	1219
Direction du vent	degré	extérieur	13%	79%	100%	71%	1170
Température	°C	extérieur	13%	83%	100%	94%	1233
Humidité relative	%	extérieur	13%	83%	100%	94%	1233
Humidité spécifique	g/kg sec	extérieur	13%	83%	100%	94%	1233
Irradiance	W/m ²	extérieur	13%	83%	100%	94%	1233
Pression	hPa	extérieur	13%	83%	100%	94%	1233
Pluie	0 / 1	extérieur	13%	83%	100%	94%	1233
CO ₂	ppm	extérieur	--	--	8%	10%	504
Particules [0.35-20 µm]	part. / L	extérieur	--	36%	95%	34%	550
Formaldéhyde	ppb	extérieur	--	--	--	3%	127
Présence	0 / 1	intérieur	78%	78%	99%	100%	1143
Ouverture des fenêtres	0 / 1	intérieur	41%	78%	99%	100%	1008
CO ₂	ppm	intérieur	79%	65%	97%	96%	1084
Température	°C	intérieur	79%	65%	96%	83%	1053
Humidité relative	%	intérieur	79%	65%	96%	83%	1053
Humidité spécifique	g/kg sec	intérieur	79%	65%	96%	83%	1053
Irradiance	W/m ²	intérieur	63%	68%	91%	84%	989
Impulsions imprimante		intérieur	6%	90%	80%	61%	772
CO	ppm	intérieur	--	41%	37%	28%	344
NO	ppb	intérieur	--	42%	--	--	152
NO ₂	ppb	intérieur	--	33%	--	--	119
Ozone	ppb	intérieur	--	42%	47%	92%	518
Particules [0.35-20 µm] - imprimante	part. / L	intérieur	--	22%	--	--	80
Particules [0.35-20 µm] - fenêtre	part. / L	intérieur	77%	43%	96%	86%	971
Particules [0.02-0.7 µm] - imprimante	part./cm ³	intérieur	--	--	--	75%	159
Particules [0.02-0.7 µm] - fenêtre	part./cm ³	intérieur	--	--	--	76%	161
Formaldéhyde	ppb	intérieur	--	25%	3%	3%	437

*Pour l'année 2015, seuls les mois de janvier à juillet sont pris en compte.

Le pourcentage indique le nombre de minutes de l'année renseigné pour le paramètre en question. Certains paramètres n'ont pas été renseignés à l'échelle de la minute. C'est le cas des données météorologiques en 2012 (la plupart du temps au pas de temps de 10 minutes), de la mesure du CO₂ en extérieur (pas de temps de 10 minutes) et du formaldéhyde en 2014 et 2015 (pas de temps de 20 minutes). Dès lors, les 3 % de minutes renseignées pour le formaldéhyde en 2015 se répartissent sur 60 % de l'année.

La colonne cumul représente le nombre de jours cumulés renseignés au total pour chaque paramètre.

Le suivi sur le long terme d'un aussi grand nombre de paramètres est sujet à des incidents matériels et nécessite malgré tout une surveillance régulière du matériel par du personnel qualifié pas toujours disponible tout au long de l'année. Cela explique la variation selon les années des mesures valides disponibles. Certains paramètres n'ont été mis en place que tardivement comme le comptage des impulsions électriques.

L'exploitation de ces données a été réalisée au pas de temps d'acquisition des mesures sans recourir à un pas de temps plus important qui diminuerait le nombre de données manquantes.

L'ensemble de ces données est destiné à être mis à disposition de la communauté scientifique internationale, une fois les principaux résultats de ces travaux publiés.

4.1.2 Formaldéhyde

4.1.2.1 Evolution de la concentration

L'analyseur de formaldéhyde a été utilisé sans multiplexage d'avril à septembre 2013 (avec une interruption en août) dans l'espace de bureau. Durant cette période, la concentration de formaldéhyde était mesurée toutes les minutes au niveau de la gaine d'extraction de l'espace de bureau. La concentration de formaldéhyde mesurée durant cette période varie de 1,5 à 25 ppb (Figure 20). La concentration est donc globalement faible (en moyenne 9 ppb) mais montre des variations très importantes au cours d'une journée pouvant atteindre une amplitude de 10 ppb.

Pour l'année 2014, la mesure de formaldéhyde intérieure est disponible de fin mars à octobre sur un pas de temps de 20 minutes (Figure 21). En 2015, la mesure de formaldéhyde couvre les 6 premiers mois de l'année de janvier à juin sur un pas de temps de 20 minutes et renseigne aussi bien l'intérieur de l'espace de bureaux que l'extérieur (Figure 22).

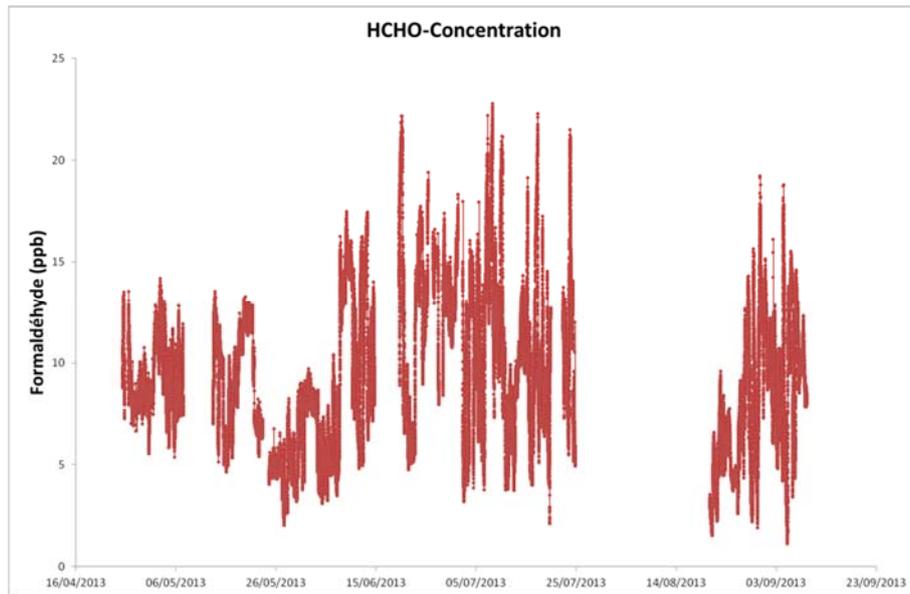


Figure 20 : Concentration de formaldéhyde dans l'espace de bureau d'avril à septembre 2013.

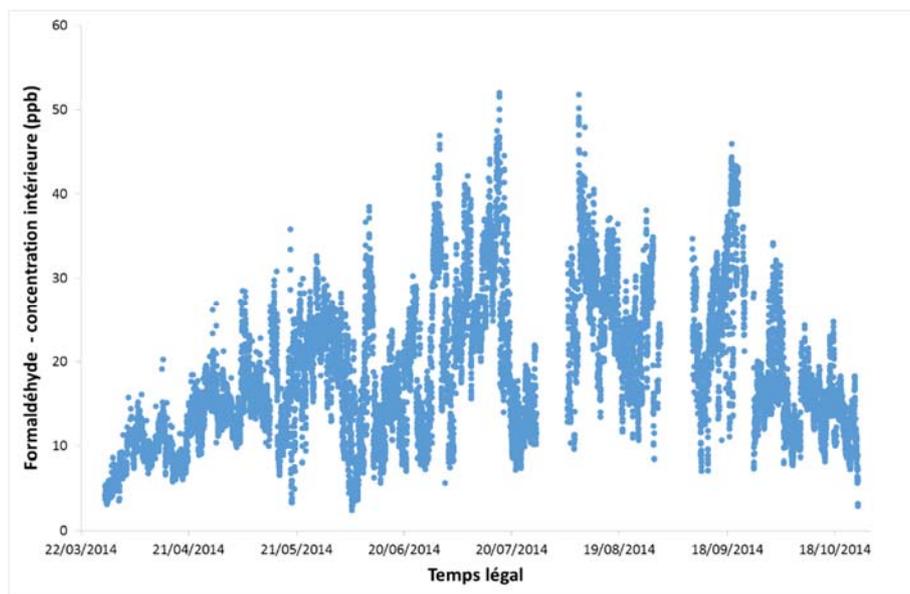


Figure 21 : Concentration de formaldéhyde dans l'espace de bureau de mars à octobre 2014.

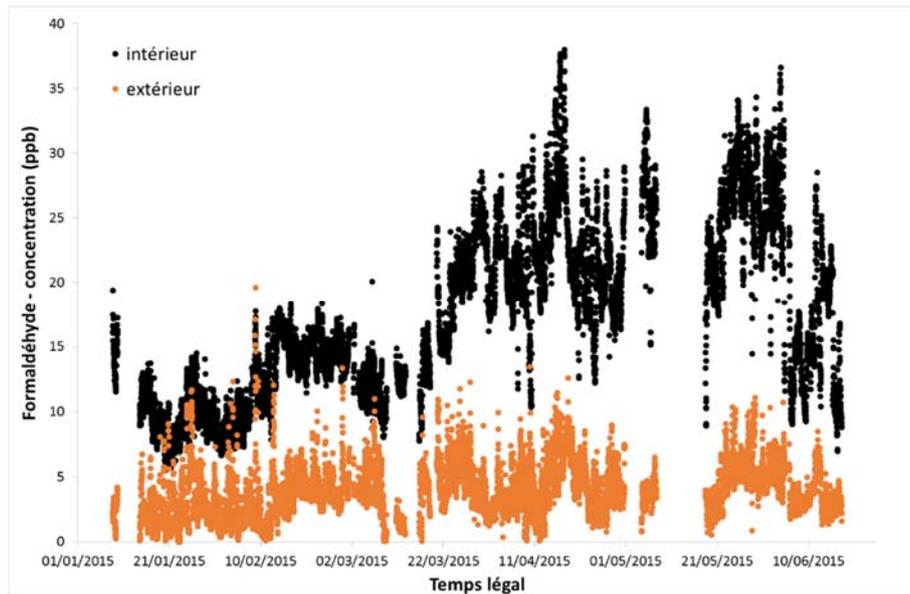


Figure 22 : Concentration de formaldéhyde dans l'espace de bureau de janvier à juin 2015.

Les niveaux intérieurs de formaldéhyde sont globalement plus élevés en 2014 et 2015 (médiane de 16 ppb) par rapport à 2013 (9 ppb). La concentration extérieure se situe en moyenne à 4 ppb. Le ratio intérieur/extérieur calculé sur les moyennes journalières en 2015 se situe en valeur médiane à 4,4.

Tableau 4 : Statistiques des mesures de formaldéhyde (rapports de mélanges en ppb)

Variable	Année	N	Min	P10	P25	Médiane	P75	P90	Moyenne	Ecart-type	Max
Intérieur	2013	130370	1.1	4.7	6.4	8.7	11.8	14.0	9.2	3.6	22.8
Intérieur	2014	13410	2.4	8.6	11.9	16.3	23.4	30.4	18.1	8.4	52.0
Intérieur	2015	9350	5.5	9.3	11.9	16.3	22.3	26.7	17.4	6.6	38.0
Extérieur	2015	9322	< 0.1	1.5	2.5	3.7	5.1	6.5	3.9	2.0	19.7
Ratio I/E (j)	2015	138	2.2	2.9	3.5	4.4	5.7	7.9	4.9	2.2	17

4.1.2.2 Variabilité temporelle de la concentration en formaldéhyde

En 2013, les variations de la concentration en formaldéhyde exprimées en termes de coefficient de variation (rapport de l'écart-type à la moyenne) sont de l'ordre de 40 % à l'échelle de la minute ou de l'heure et de 33 % pour une variabilité journalière. L'année 2014 montre la variabilité la plus élevée de la concentration intérieure en formaldéhyde avec des variations comprises entre 46 % (toutes les 20 minutes) et 40 % (variation mensuelle). En 2015, la variabilité de la concentration intérieure reste élevée, à près de 38 % à l'échelle de l'heure et 36 % à l'échelle de la journée. La concentration extérieure

présente des variations plus importantes à hauteur de 50 % à l'échelle de l'heure et 39% à l'échelle de la journée. Les variations mensuelles sont par contre plus importantes en intérieur (33 %) qu'en extérieur (19 %).

La variabilité temporelle de la concentration de formaldéhyde est représentée à l'échelle de la semaine, du jour et du mois sous forme de profil temporel moyen de 2013 à 2015 de Figure 23 à la Figure 25. Les profils sont déterminés en temps légal. La bande de couleur représente l'intervalle de confiance de la moyenne temporelle à partir des données horaires. Pour 2015, les profils sont normalisés par rapport à la moyenne annuelle pour permettre de mieux comparer la variation des concentrations intérieure et extérieure.

En 2013, les fluctuations de la concentration intérieure de formaldéhyde restent très faibles avec un profil moyen journalier assez peu marqué et des variations irrégulières entre les différents jours de la semaine. Un profil assez semblable est observé en 2014 mais avec une amplitude de variation plus marquée. La concentration de formaldéhyde est un peu plus élevée la nuit et plutôt faible en fin d'après-midi et en début de matinée.

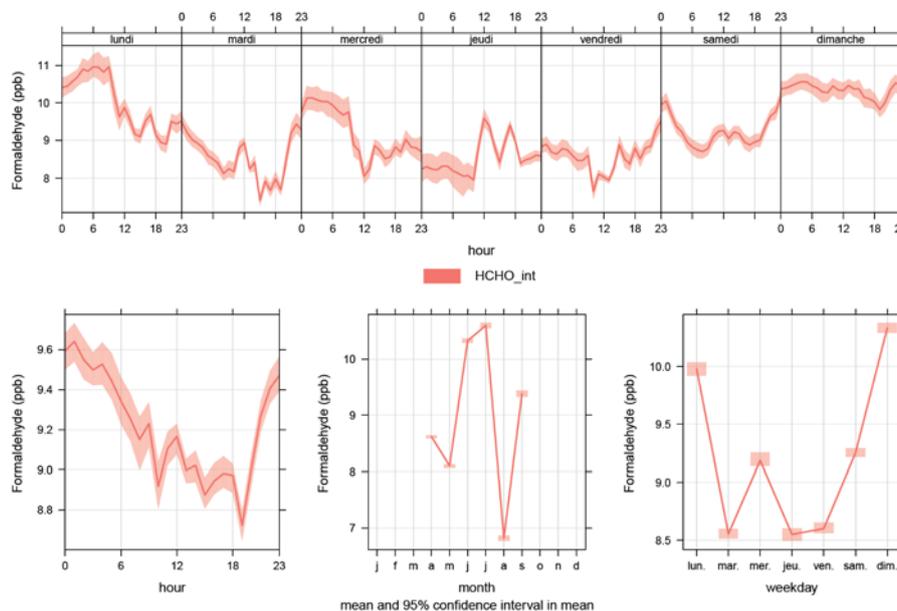


Figure 23 : Variations journalière, hebdomadaire et mensuelle de la concentration intérieure de formaldéhyde en 2013.

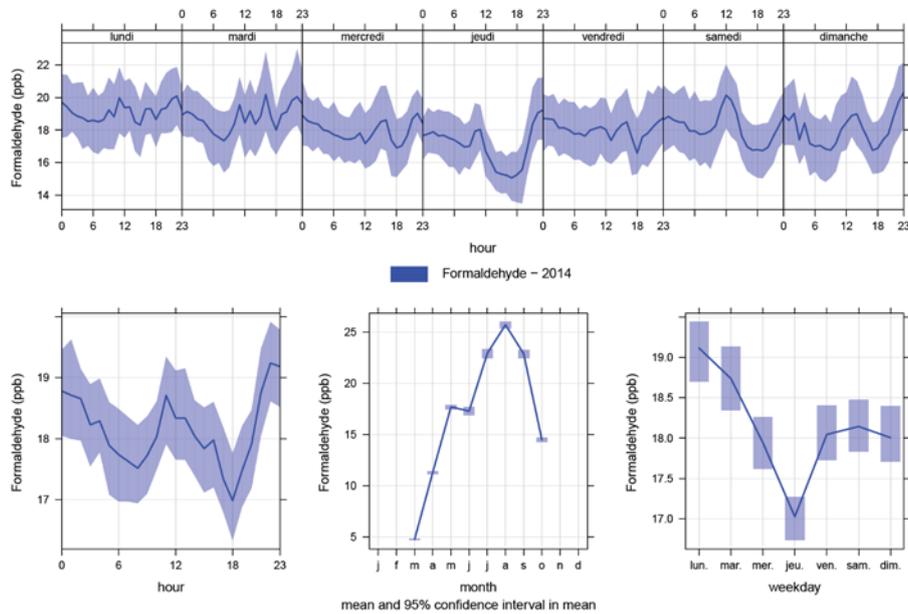


Figure 24 : Variations journalière, hebdomadaire et mensuelle de la concentration intérieure de formaldéhyde en 2014.

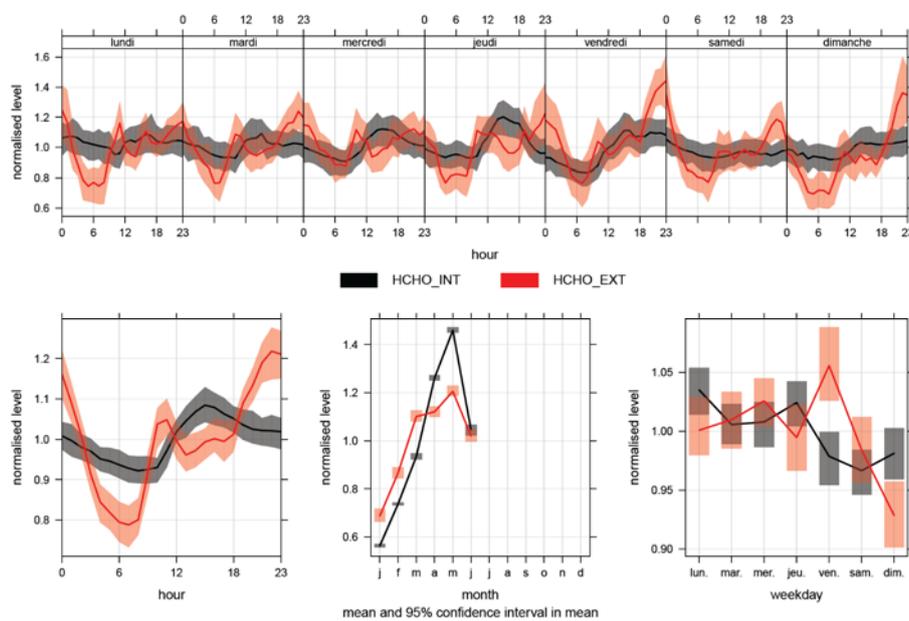


Figure 25 : Variations journalière, hebdomadaire et mensuelle de la concentration normalisée en formaldéhyde à l'intérieur et à l'extérieur en 2015.

En 2015, le profil moyen journalier de la concentration intérieure en formaldéhyde montre un maximum en milieu d'après-midi et un minimum en fin de matinée. Ce profil est également plus marqué les jours de la semaine que le week-end. La concentration

extérieure est généralement maximale en début de nuit pour décroître vers son minimum en début de matinée. Dans la journée, la concentration extérieure augmente jusqu'en fin de matinée et stagne dans le courant de l'après-midi pour remonter ensuite en soirée. La concentration extérieure ne présente pas de variations selon les jours de la semaine. Ce profil extérieur semble finalement assez proche du profil intérieur observé en 2013 et 2014.

Cette différence entre les profils de 2013, 2014 et 2015 pourrait être liée au comportement à l'ouverture des fenêtres qui diffère d'un jour à l'autre et d'une année sur l'autre. Cela conduit à des journées de mesures qui ne se ressemblent pas. Pour illustrer cela, différentes journées ont été sélectionnées et comparées aux autres. Les journées sélectionnées représentent différents jours de semaine. L'évolution de la concentration de formaldéhyde lors de ces journées est représentée à la Figure 26.

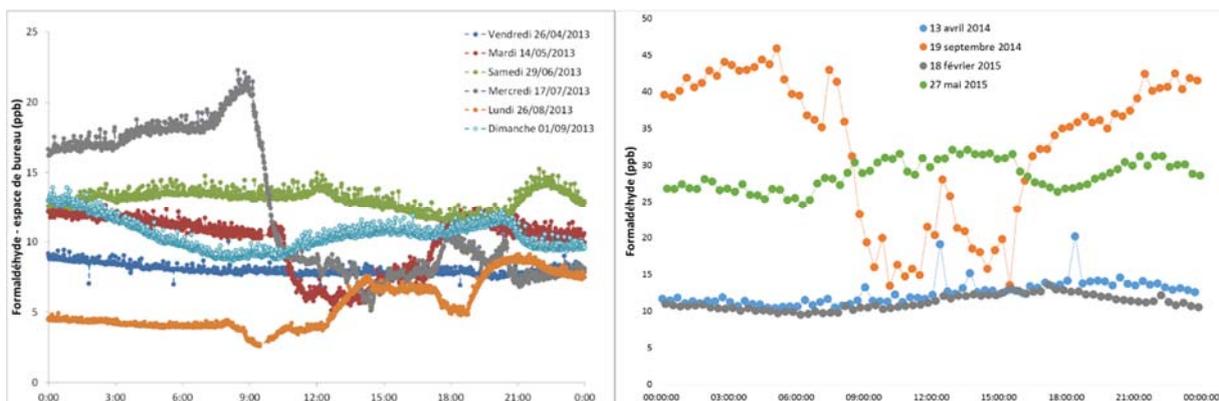


Figure 26 : Profils journaliers types de concentration intérieure de formaldéhyde dans l'espace de bureau au cours de l'année 2013 (à gauche) et 2014/2015 (à droite).

Les différentes journées sélectionnées montrent des évolutions très différentes de la concentration de formaldéhyde : soit très peu de variations (26 avril 2013), des variations brusques (14 mai et 17 juillet 2013, 19 septembre 2014) ou des variations moins marquées (26 août et 1^{er} septembre 2013). On peut également observer que durant la nuit, le niveau de formaldéhyde est généralement stable, mais très variable d'un jour à l'autre. La chute de concentration du formaldéhyde du 14 mai 2013 de 10h30 à 16h30 est associée à l'ouverture d'une des fenêtres de l'espace de bureau. Cette ouverture s'accompagne d'une brusque élévation de la concentration d'ozone à l'intérieur et d'une diminution de la température intérieure (ainsi que de l'humidité spécifique non représentée). L'ouverture de plusieurs fenêtres ne conduit pas forcément à une nouvelle diminution de la concentration de formaldéhyde (17 juillet 2013). Passer de une à cinq fenêtres ouvertes diminue la concentration intérieure de formaldéhyde le 11 juin 2014, mais cette même situation n'entraîne pas de baisse de concentration le lendemain. Des variations de la concentration intérieure sont

également observées lorsque les fenêtres sont fermées, par exemple du 9 au 18 février 2015. Durant ces journées, les variations du formaldéhyde sont étroitement liées aux variations d'humidité spécifique aussi bien intérieure qu'extérieure. Pour autant, les fluctuations de la concentration extérieure de formaldéhyde ne semblent pas dépendre de l'humidité spécifique.

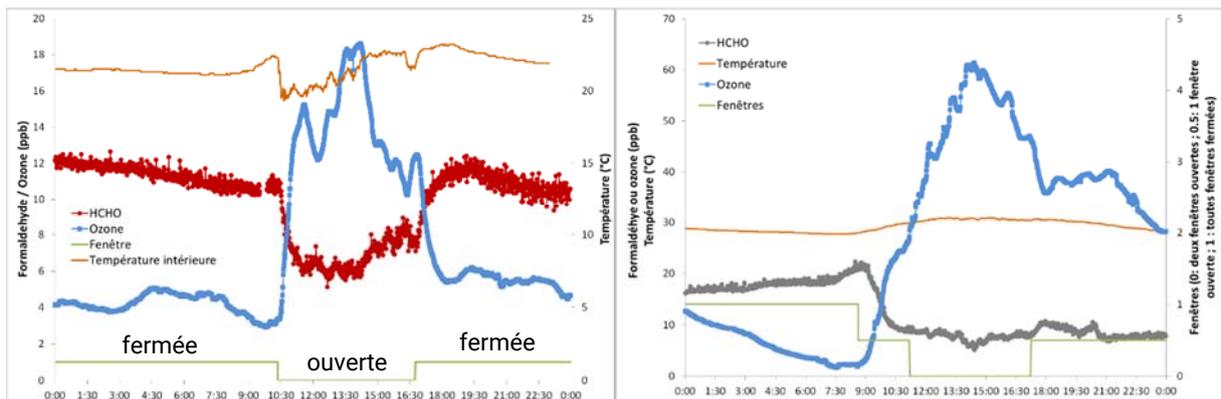


Figure 27 : Variation brusque de la concentration de formaldéhyde liée à l'ouverture d'une fenêtre lors de la journée du 14 mai 2013 (à gauche) et de une à deux fenêtres le 17 juillet 2013 (à droite).

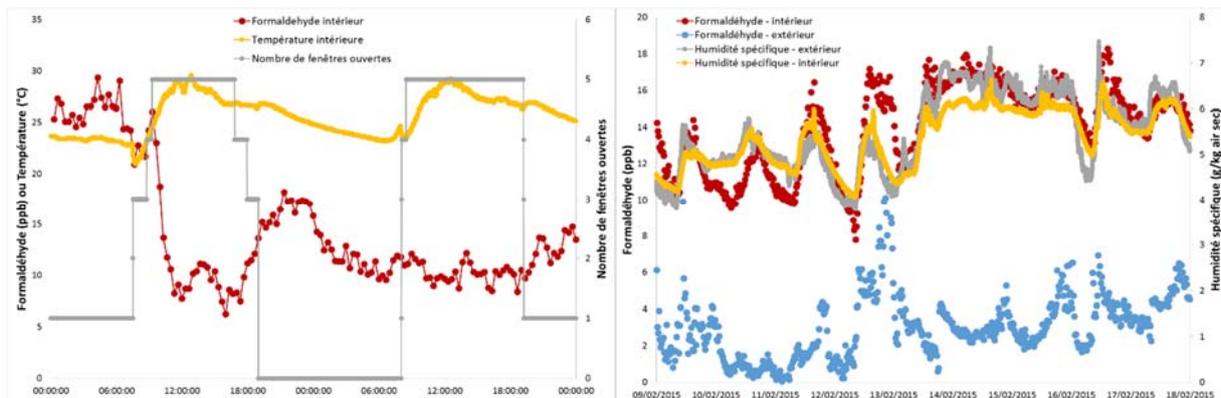


Figure 28 : Variation de la concentration de formaldéhyde liée à l'ouverture de fenêtres les 11 et 12 juin 2014 (à gauche) et à l'humidité spécifique en l'absence d'ouverture de fenêtres du 9 au 18 février 2015 (à droite).

L'influence d'un paramètre tel que l'ouverture des fenêtres n'est pas systématique et dépend d'autres conditions. Cette influence sera étudiée en détail par la suite.

4.1.3 Particules

4.1.3.1 Evolution des concentrations

La concentration de particules entre 0,3 et 20 μm a été mesurée par compteur optique dans l'espace de bureau de fin janvier à décembre 2012 (un seul point intérieur) et depuis juillet 2013 jusqu'en juin 2015 (mesures intérieures et extérieures). En 2013, un troisième compteur optique a été placé pendant quelques mois à proximité de l'imprimante multifonctions. En 2015, deux compteurs à diffusion électrique ont été déployés dans l'espace de bureau, l'un au niveau de la position principale (à proximité d'une fenêtre) et le second à l'opposé au niveau de l'imprimante.

Entre 2012 et 2015, la concentration moyenne de particules dans l'espace de bureau reste relativement stable entre 0,03 et 0,05 particules/ cm^3 pour les particules de taille moyenne ($d_i^3 = 0,25 \mu\text{m}$) et entre 11 et 16 particules/ cm^3 pour les particules plus fines (0,35 μm). Les statistiques des concentrations en nombre pour deux tailles de particules (0,35 et 2,5 μm) sont présentées au Tableau 5. Les ratios I/E journaliers représentent le rapport de la concentration moyenne journalière intérieure par rapport à l'extérieur. Les évolutions observées pour ces deux gammes de taille sont représentées à la Figure 29.

³ d_i représente le diamètre moyen de la gamme de taille comprise entre le diamètre inférieur d_{p1} et le diamètre supérieur d_{p2} .

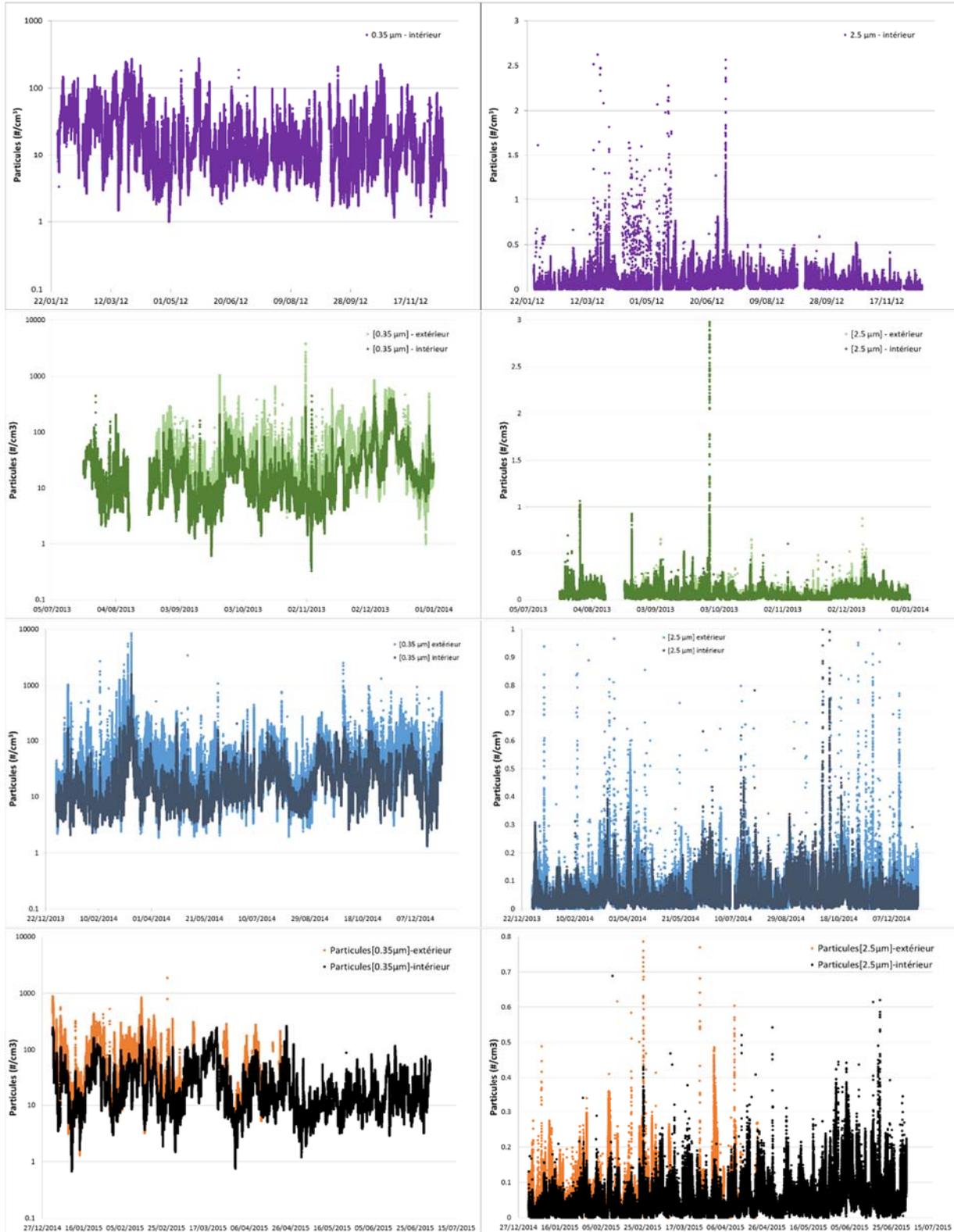


Figure 29 : Evolution des concentrations en nombre de particules de diamètre 0,35 µm (gauche) et 2,5 µm (droite) dans l'espace de bureau de 2012 à 2015.

Tableau 5 : Extrait des statistiques des mesures en nombre de particules (particules par cm³)

Intérieur	Année	N	Min	P10	P25	Médiane	P75	P90	Moyenne	Ecart-type	Max
[0,3-0,4 µm]	2012	406404	1.0	4.7	7.2	13	28	57	23	28	274
[0,3-0,4 µm]	2013	224236	0.3	4.7	7.7	16	30	56	28	42	444
[0,3-0,4 µm]-prin	2013	114864	0.4	4.2	6.1	11	24	40	18	20	382
[0,3-0,4 µm]	2014	505571	1.3	5.8	8.6	15	28	50	24	31	1589
[0,3-0,4 µm]	2015	262084	0.68	5.9	8.9	16	34	61	27	31	263
[2-3 µm]	2012	406404	<LD	0.01	0.02	0.05	0.09	0.14	0.06	0.07	3.3
[2-3 µm]	2013	224236	<LD	0.01	0.03	0.05	0.08	0.12	0.06	0.06	3.1
[2-3 µm]-prin	2013	114864	<LD	0.01	0.01	0.03	0.05	0.08	0.04	0.06	2.8
[2-3 µm]	2014	505571	<LD	0.01	0.02	0.04	0.06	0.09	0.04	0.04	2.9
[2-3 µm]	2015	262083	<LD	0.01	0.02	0.03	0.06	0.10	0.05	0.04	0.92
[0,02-0,7 µm]-fen	2015	230819	1027	3744	5197	7336	10263	14667	8386	4596	61898
[0,02-0,7 µm]-prin	2015	230162	1114	3824	5362	7644	10620	14823	8611	4627	82872
Extérieur	Année	N	Min	P10	P25	Médiane	P75	P90	Moyenne	Ecart-type	Max
[0,3-0,4 µm]	2013	188713	0.4	7.6	13	29	63	132	53	71	3849
[0,3-0,4 µm]	2014	497586	1.4	8	13	26	60	112	50	84	8486
[0,3-0,4 µm]	2015	105114	0.87	9.7	16	43	107	172	73	82	1863
[2-3 µm]	2013	188713	<LD	0.02	0.03	0.05	0.08	0.11	0.06	0.04	0.9
[2-3 µm]	2014	497586	<LD	0.02	0.03	0.05	0.08	0.11	0.06	0.05	2.8
[2-3 µm]	2015	105114	<LD	0.01	0.03	0.05	0.08	0.14	0.06	0.06	1.7
Ratio I/E (j)	Année	N(j)	Min	P10	P25	Médiane	P75	P90	Moyenne	Ecart-type	Max
[0,3-0,4 µm]	2013	132	0.16	0.28	0.38	0.52	0.74	0.96	0.59	0.30	1.6
[0,3-0,4 µm]	2014	354	0.14	0.31	0.39	0.54	0.73	0.87	0.57	0.22	1.2
[0,3-0,4 µm]	2015	85	0.17	0.30	0.37	0.44	0.55	0.65	0.47	0.18	1.4
[2-3 µm]	2013	132	0.008	0.49	0.62	0.86	1.09	1.38	0.90	0.40	2.4
[2-3 µm]	2014	354	0.08	0.43	0.56	0.72	0.94	1.2	0.77	0.31	2.1
[2-3 µm]	2015	85	0.21	0.33	0.43	0.55	0.66	0.74	0.54	0.16	0.85

Le ratio I/E journalier moyen est représenté par année selon le diamètre des particules à la Figure 30. Ce ratio moyen est systématiquement inférieur à 1 et voisin de 0,5 pour les particules de diamètre inférieur à 5 µm. Pour des diamètres supérieurs, le ratio moyen très variable varie entre 1 et 6, le maximum étant observé pour les particules comprises entre 10 et 15 µm. Pour les particules de diamètre > 20 µm (d_i = 22,5 µm), le ratio redevient inférieur à 1. Mais, pour cette gamme de taille, le compteur optique utilisé n'est pas l'instrumentation la plus appropriée. Son efficacité de collecte est relativement faible comparativement à des cyclones. Les résultats relatifs à cette gamme de taille sont donc à considérer avec précaution.

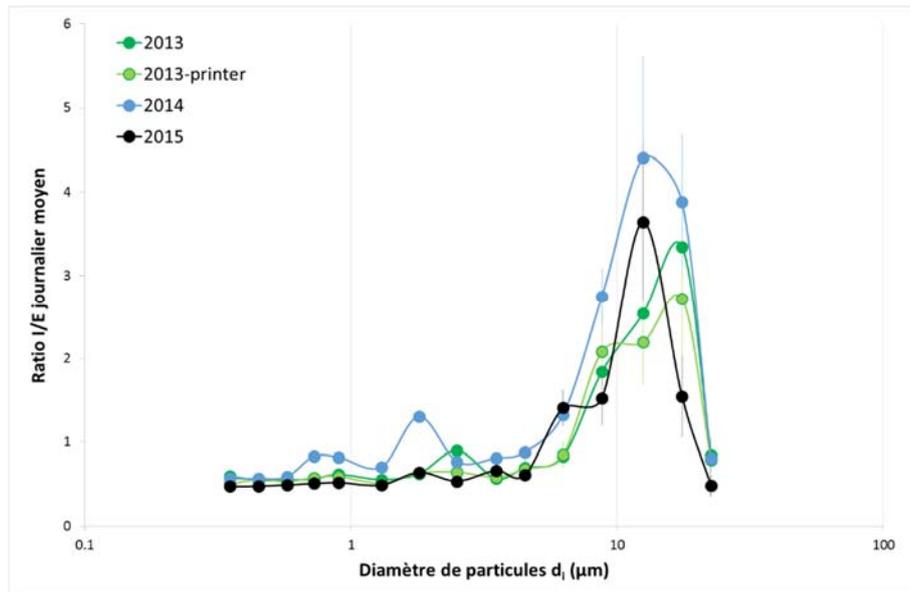


Figure 30 : Ratio I/E journalier moyen selon le diamètre des particules et par année. Les barres d'erreur représentent l'intervalle de confiance du ratio I/E moyen par taille. L'année 2013 couvre les mois de septembre à décembre et l'année 2015 de janvier à avril.

Cette variation du ratio I/E donne déjà une indication quant aux origines possibles de ces particules. Les particules inférieures à 5 µm avec un ratio I/E < 1 semblent majoritairement issues de l'extérieur et subissent des pertes liées à la fois au transfert de l'extérieur vers l'intérieur et aux dépôts par diffusion et sédimentation sur les surfaces. L'absence de sources de combustion dans l'environnement de bureaux limite les sources intérieures de ces particules.

Pour les plus grosses particules, le ratio supérieur à 1 indique la présence de processus d'émission à l'intérieur de l'espace de bureaux suffisamment important pour compenser les pertes par sédimentation.

La ventilation simple flux sans traitement d'air ne joue ici qu'un rôle de transport de particules et ne représente pas un puits proprement dit.

En 2013, le ratio I/E moyen entre les deux compteurs disposés dans l'espace de bureaux est sensiblement identique pour des particules inférieures à 6,5 µm. Il diffère pour des particules de diamètre supérieur. Cela indique que ces particules plus grosses ne sont pas distribuées de manière homogène dans l'espace de bureaux et dépendent par conséquent de conditions locales thermo-aérauliques, d'émissions ponctuelles localisées et de pertes par dépôt spatialement hétérogènes.

4.1.3.2 Variabilité temporelle des concentrations de particules

Quelle que soit l'année, la variabilité des données mesurées à l'intérieur (moyennes à la minute ou moyennes horaires), exprimée en coefficient de variation varie entre de 70 % (particules entre 0,8 et 2 μm) à plus de 200 % pour les particules supérieures à 3 μm . La variabilité est du même ordre à l'extérieur, mais avec une variabilité plus prononcée pour les plus grosses particules. Cette variabilité diminue pour des moyennes journalières (de 50 % à 150 %) ou mensuelles (de 15 % à 80 %). Dans ces deux derniers cas, la variabilité la plus faible est généralement observée pour la fraction comprise entre 0,8 et 2 μm . Les particules ultrafines mesurées en 2015 présentent la variabilité la plus faible autour de 50 % (horaire), 30 % (journalière) et 15 % (mensuelle).

Cette variabilité temporelle de la concentration de particules est représentée à l'échelle de la semaine, du jour et du mois pour quelques tailles de particules (0,35 μm ; 2,5 μm ; 12,5 μm et les particules ultrafines) sous forme de profil temporel moyen de 2012 à 2015 de la Figure 31 à la Figure 40. Les profils sont normalisés par rapport à la moyenne annuelle pour permettre de superposer différents profils de niveaux différents. Les profils sont déterminés en temps légal. La bande de couleur représente l'intervalle de confiance de la moyenne temporelle à partir des données horaires.

Le profil de variation temporelle est très différent selon la taille des particules. Les particules de 0,35 μm présentent des concentrations intérieures et extérieures maximales en soirée autour de minuit avec un pic dans la matinée et un minimum observé en milieu d'après-midi. Les variations intérieures sont étroitement associées aux variations extérieures. La concentration de ces particules est également généralement plus élevée durant le mois de mars (2012 et 2014 notamment). La différence entre les jours de la semaine est peu marquée. Les particules ultrafines suivent un profil de variation similaire avec un pic plus marqué dans la matinée et des niveaux plus faibles en fin de semaine.

Les particules de 2,5 μm montrent des variations plus marquées à l'intérieur au cours de la journée avec un maximum observé en fin de matinée et une lente décroissance dans l'après-midi. Les niveaux sont minimaux durant la nuit. Le profil de variation de la concentration extérieure est proche de celui observé à l'intérieur mais avec une amplitude de variation plus faible. En fin de semaine, les niveaux intérieurs comme extérieurs sont généralement plus faibles. Les mois de juin et de juillet présentent souvent des niveaux élevés en intérieur comme à l'extérieur pour cette taille de particules.

Les particules plus grosses de 12,5 μm présentent les variations les plus marquées à l'intérieur au cours de la journée avec un maximum entre 8h et 17h correspondant à l'occupation des locaux et des niveaux très faibles la nuit et les week-ends. L'amplitude des variations extérieures est beaucoup plus faible. Les concentrations intérieure et extérieure sont généralement plus élevées durant le printemps et l'été et plus faibles en hiver.

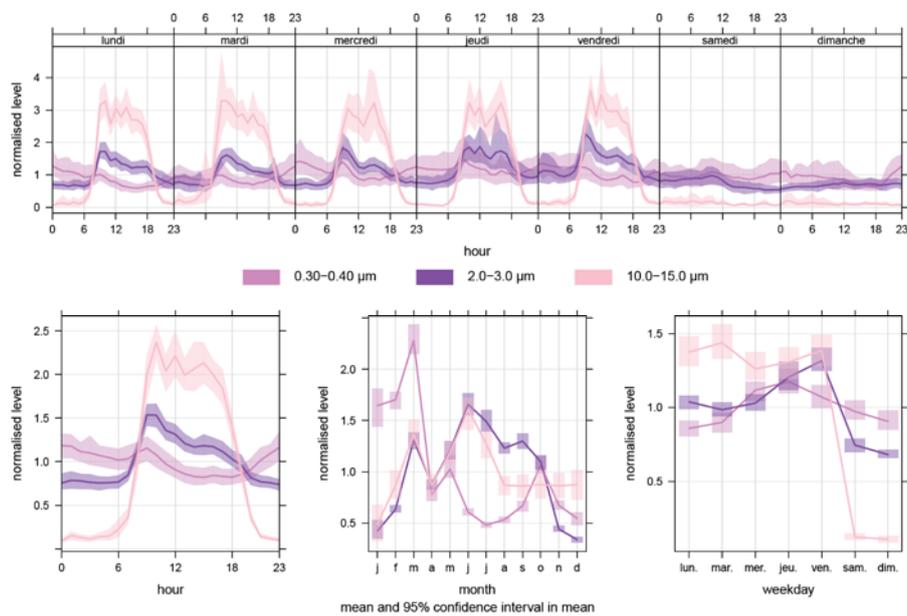


Figure 31 : Variations journalière, hebdomadaire et mensuelle de la concentration intérieure en particules sur l'année 2012 (trois tailles de particules représentées).

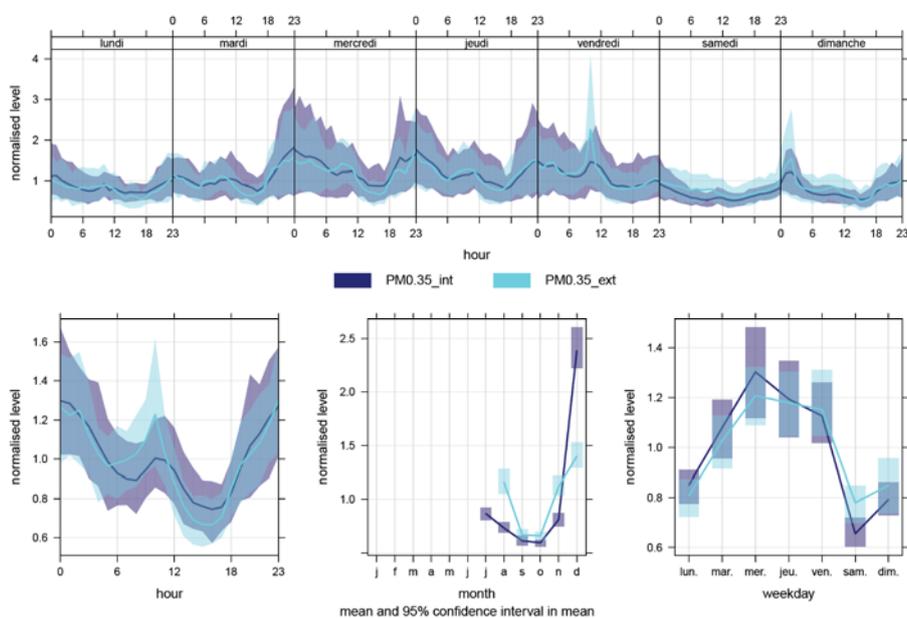


Figure 32 : Variations journalière, hebdomadaire et mensuelle de la concentration intérieure et extérieure en particules de 0,35 µm sur le second semestre de l'année 2013.

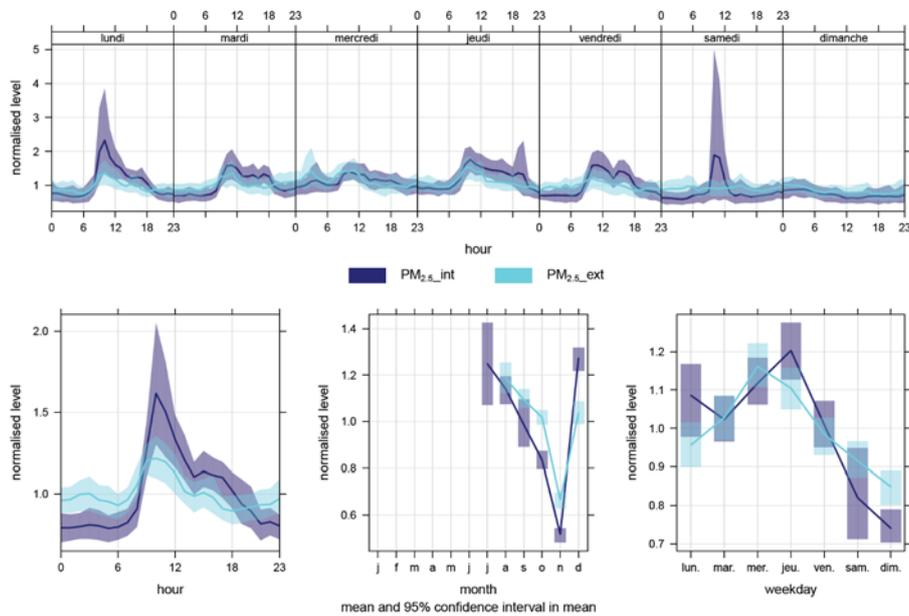


Figure 33 : Variations journalière, hebdomadaire et mensuelle de la concentration intérieure et extérieure en particules de 2,5 µm sur le second semestre de l'année 2013.

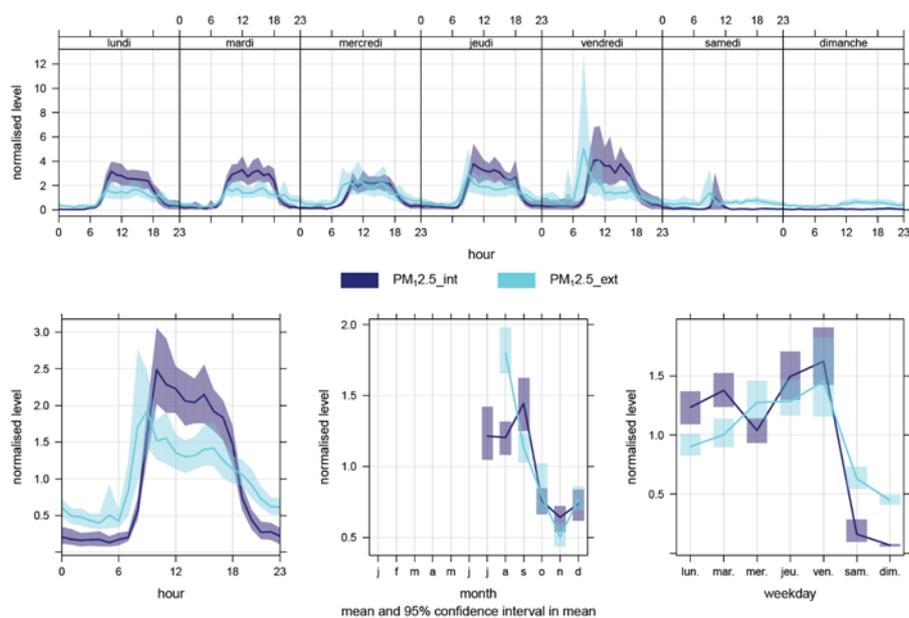


Figure 34 : Variations journalière, hebdomadaire et mensuelle de la concentration intérieure et extérieure en particules de 12,5 µm sur le second semestre de l'année 2013.

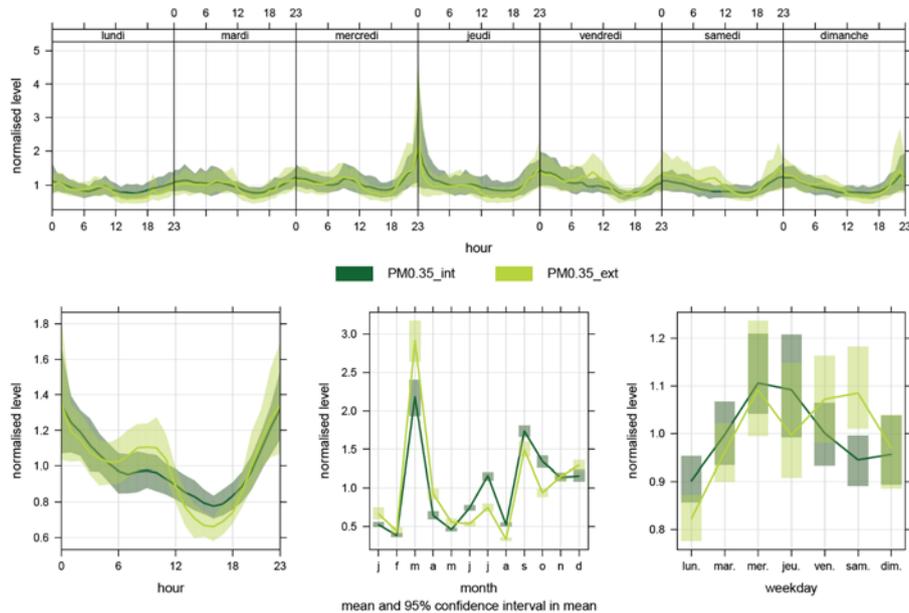


Figure 35 : Variations journalière, hebdomadaire et mensuelle de la concentration intérieure et extérieure en particules de 0,35 µm sur l'année 2014.

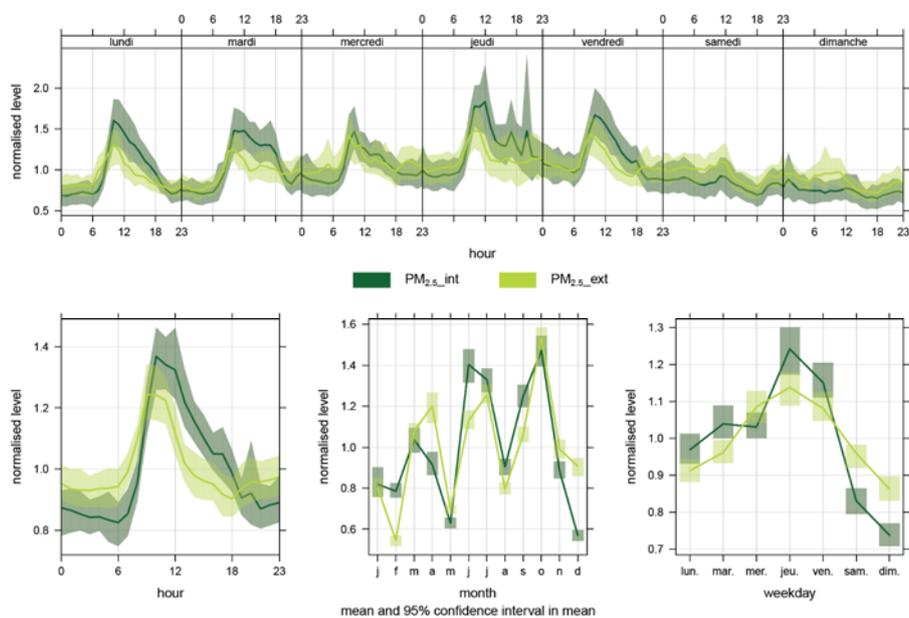


Figure 36 : Variations journalière, hebdomadaire et mensuelle de la concentration intérieure et extérieure en particules de 2,5 µm sur l'année 2014.

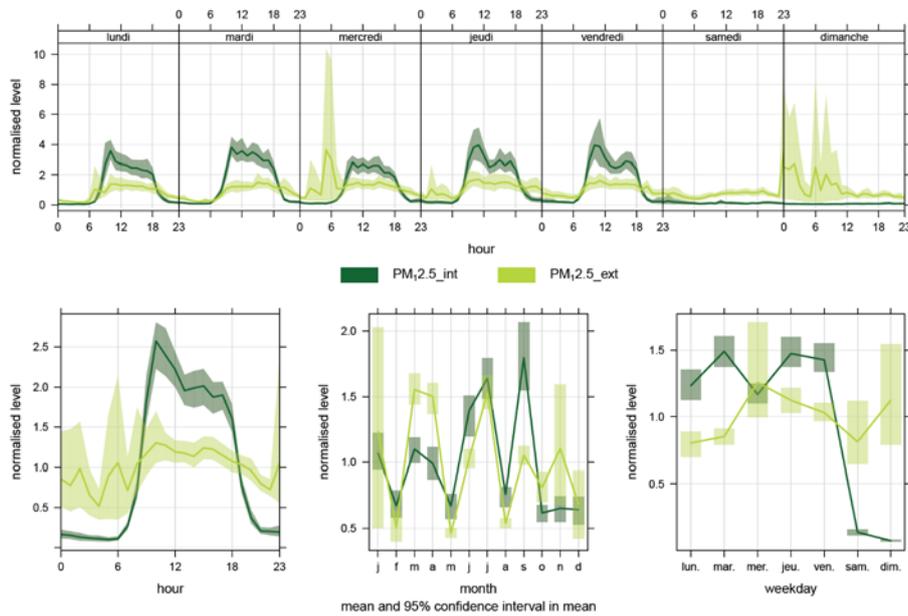


Figure 37 : Variations journalière, hebdomadaire et mensuelle de la concentration intérieure et extérieure en particules de 12,5 µm sur l'année 2014.

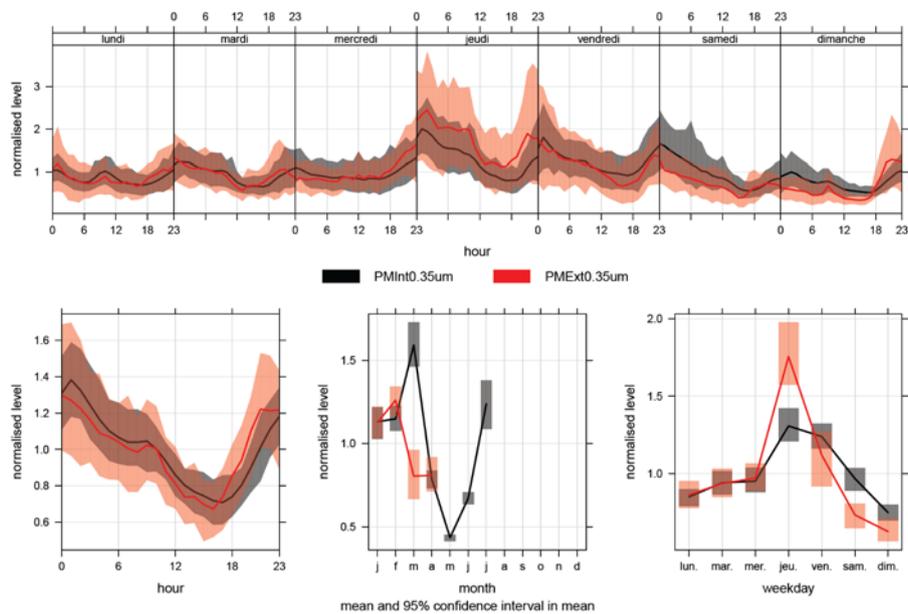


Figure 38 : Variations journalière, hebdomadaire et mensuelle de la concentration intérieure et extérieure en particules de 0,35 µm sur le premier semestre de l'année 2015.

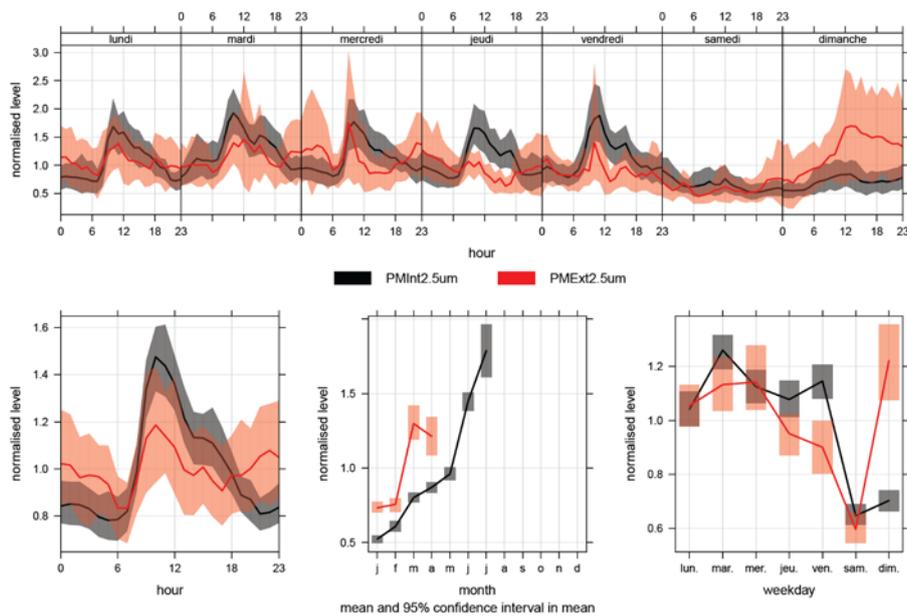


Figure 39 : Variations journalière, hebdomadaire et mensuelle de la concentration intérieure et extérieure en particules de 2,5 µm sur le premier semestre de l'année 2015.

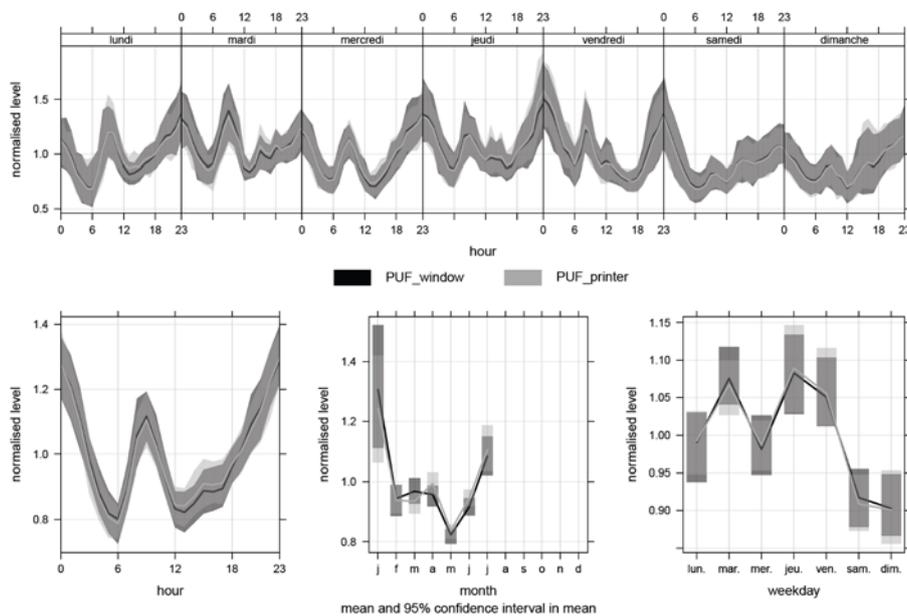


Figure 40 : Variations journalière, hebdomadaire et mensuelle de la concentration intérieure en particules ultrafines [0,02 – 0,7 µm] sur le premier semestre de l'année 2015.

4.1.4 Corrélations entre les variables

Pour illustrer les corrélations existantes entre les différentes variables, une analyse par composantes principales normée (ACP) a été réalisée sur l'ensemble des variables continues de 2013 à 2015. Les données sont fortement multidimensionnelles avec les deux premières composantes qui ne représentent que 29 % à 34 % de la variance globale selon les ans. Le premier axe est porté par les variations induites par les concentrations de particules intérieures et extérieures. Le second axe se caractérise par les paramètres climatiques intérieurs et extérieurs saisonniers (température, humidité, irradiance), mais également l'ozone intérieur. Cet axe discrimine les particules intérieures les plus grosses et celles intermédiaires, dont la concentration évolue conjointement aux paramètres saisonniers (elle tend à augmenter en saison estivale), des particules intérieures plus fines qui évolue de la même façon que les particules fines extérieures et dont la concentration est plutôt plus élevée en hiver. Les particules extérieures de plus gros diamètre varient selon l'axe 3 de manière pratiquement indépendante des autres paramètres.

Le formaldéhyde ne présente pas d'association avec une composante en particulier, ce qui indique que les facteurs qui contribuent à l'évolution de sa concentration sont multiples et/ou non pris en compte dans l'étude.

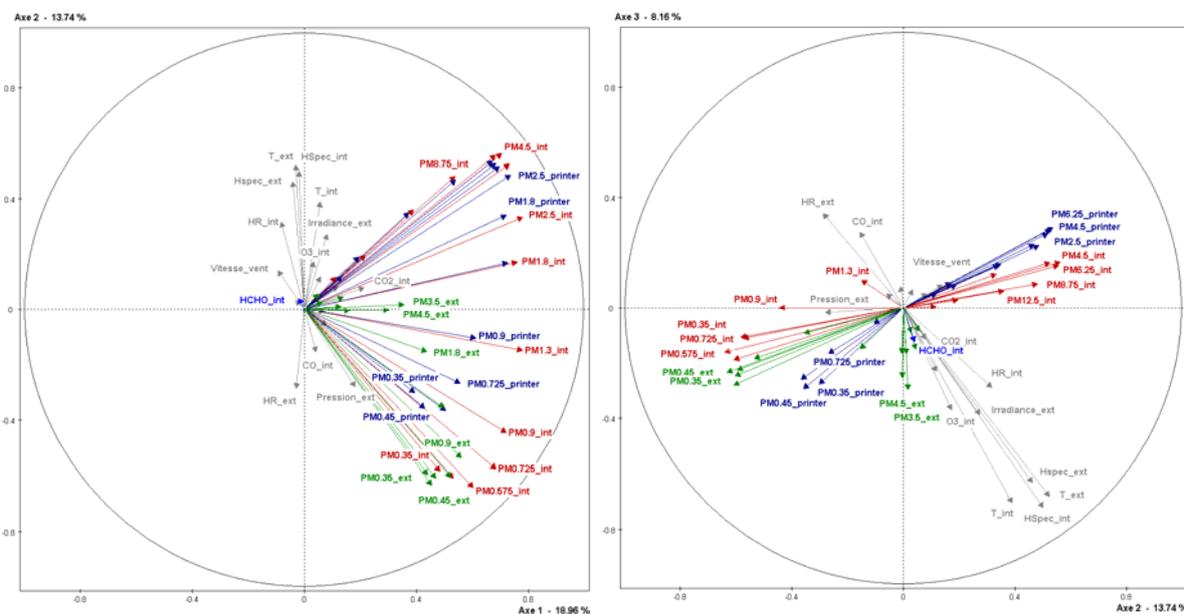


Figure 41 : ACP normée des données 2013

4.2 Exposition réelle des occupants

A l'aide du suivi fin des concentrations de formaldéhyde et de particules, il est possible de déterminer la concentration d'exposition selon plusieurs critères :

- Intégration sur 5 jours du lundi matin 9h au vendredi soir 17h (simulation de mesures par prélèvement passif)
- Intégration sur 24 heures (journée complète)
- Intégration sur 8 heures (journée de travail 9h – 17h)
- Intégration selon la présence des occupants (capteur de mouvement)

Ces différents calculs de la concentration d'exposition sont ensuite comparés à l'exposition de référence définie ici par la présence des occupants selon les capteurs de mouvement. L'absence de mouvement peut dans certains cas être associée à une présence immobile ne déclenchant pas de signal des occupants. Par contre, le moindre mouvement est forcément associé à une occupation. Les capteurs de mouvements représentent donc un indicateur a minima de la présence. Cette notion sera développée plus en détail dans le chapitre traitant de l'occupation.

Pour se comparer aux mesures de formaldéhyde réalisées au pas de temps de 20 min en 2014 et 2015, la variable présence caractérisée toutes les minutes a été agrégée au pas de temps de 20 minutes en y associant une occupation dès lors qu'un mouvement était détecté durant ce laps de temps.

4.2.1 Exposition au formaldéhyde

Pour les années 2013 et 2014, les écarts entre les 4 modes de calculs sont très faibles et non significatifs. L'écart maximal observé est de +4 % à +5 % entre la moyenne calculée sur des données intégrées sur 24 heures et la moyenne calculée sur des données intégrées selon la présence des occupants. Cet écart est de +2 % à +3 % en utilisant la moyenne des données intégrées sur 5 jours. Il est inférieur à 1 % en utilisant des données intégrées sur 8 heures. Le fait de ne pas tenir compte de la présence surestime donc que très légèrement l'exposition réelle des occupants.

En 2015, les écarts entre les 4 modes de calcul sont un peu plus importants mais restent non significatifs. Par contre, ces écarts vont dans le sens opposé à ceux observés en 2013 et 2014. L'écart maximal observé est de -8 % pour une moyenne déterminée à partir de données intégrées sur 5 jours par rapport à l'exposition calculée sur la présence. Ici, la non-prise en compte de la présence tend à légèrement sous-estimer l'exposition réelle. L'exposition calculée sur les données intégrées sur 8 heures reste très voisine de l'exposition de référence déterminée sur la présence effective des occupants.

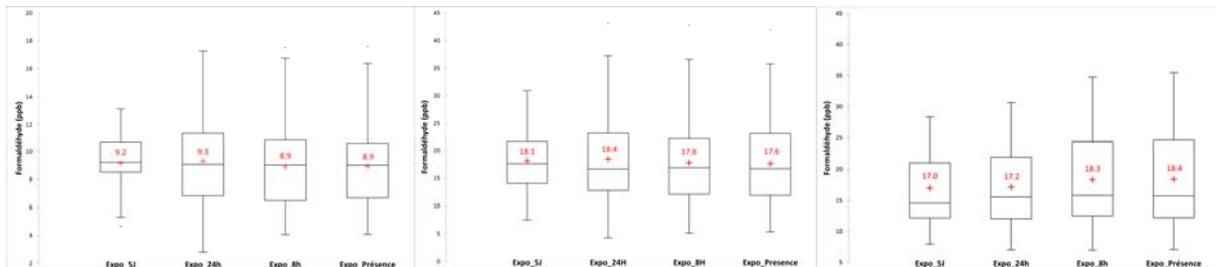


Figure 44 : Comparaison de la concentration d'exposition en formaldéhyde selon le mode de calcul sur les années 2013, 2014 et 2015 (de gauche à droite).

Ces résultats indiquent que dans des situations d'occupation similaires à celles observées dans les bureaux, calculer l'exposition sur 8 heures reste le meilleur compromis lorsque l'information relative à la présence des occupants n'est pas disponible. Malgré tout, les autres modes d'intégration fournissent des résultats très proches de ceux attendus. Au-delà de la situation d'occupation, les usages doivent également être pris en considération pour pouvoir étendre ces observations à d'autres environnements. Dans l'espace de bureau, les activités des occupants tournent autour du matériel bureautique et informatique. Ce n'est pas forcément le cas dans d'autres environnements. L'entretien du local réalisé par un prestataire extérieur reste sporadique et ne contribue pas aux variations observées.

4.2.2 Exposition aux particules

Trois gammes de taille ont été sélectionnées pour comparer les différents modes de calcul de l'exposition : la gamme 0,3 – 0,4 μm ; 2 – 3 μm ; 10 – 15 μm . Ces trois gammes présentaient des profils de variation particulièrement différents. Ces comparaisons ont été réalisées sur les données de 2013, 2014 et 2015. Fin 2013, les capteurs de mouvement n'étaient pas en fonctionnement lors d'un épisode de pollution aux particules fines, si bien que l'exposition calculée par leur intermédiaire sous-estime finalement l'exposition réelle (visible pour la fraction 0,3 – 0,4 μm).

Pour la fraction [0,3 – 0,4 μm], des écarts de l'ordre de +20 % sont observés en calculant l'exposition à partir de données intégrées sur 5 jours. L'écart relatif n'est plus que de +10 % en utilisant les données intégrées sur 24 heures et de -1 % avec des données sur 8 heures. Ces écarts sont néanmoins faiblement significatifs (test de Kruskal-Wallis, $p < 0.1$), mais ils restent du même ordre d'une année sur l'autre.

Un comportement similaire est observé pour les particules ultrafines [0,02 – 0,7 μm] en 2015 avec une surestimation de +8 % (intégration sur 5 jours), de +5 % (intégration 24 heures) ou -3 % (intégration 8 heures). Compte tenu des variations importantes d'un jour sur l'autre, ces écarts restent également peu significatifs ($p < 0.1$). Les écarts sont pratiquement identiques entre les deux instruments situés près d'une fenêtre ou près de l'imprimante.

La fraction intermédiaire [2 – 3 µm] montre des écarts systématiquement significatifs d'une année sur l'autre ($p < 0.0001$) pour des données intégrées sur 5 jours ($-22 \pm 2 \%$), pour des données intégrées sur 24 heures ($-31 \pm 2 \%$). La détermination de la concentration d'exposition à partir des données sur 8 heures n'est pas significative et s'écarte peu de l'exposition de référence ($-3 \pm 0,2 \%$).

Pour la fraction correspondant aux grosses particules [10 – 15 µm], les écarts significatifs sont encore plus importants et fidèles d'une année sur l'autre avec $-55 \pm 2 \%$ (intégration 5 jours) et $-70 \pm 2 \%$ (intégration 24 heures). L'intégration sur 8 heures présente également des écarts plus importants, mais variables d'une année sur l'autre et non significatifs ($-11 \pm 5 \%$).

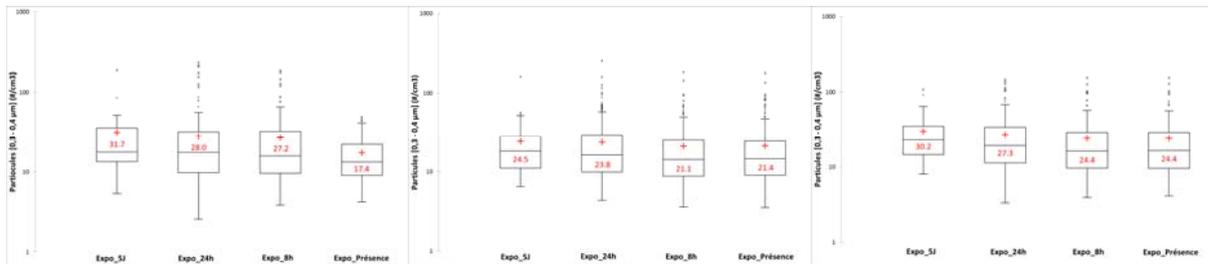


Figure 45 : Comparaison de la concentration d'exposition en particules de 0,3 à 0,4 µm de diamètre selon le mode de calcul sur les années 2013, 2014 et 2015 (de gauche à droite).

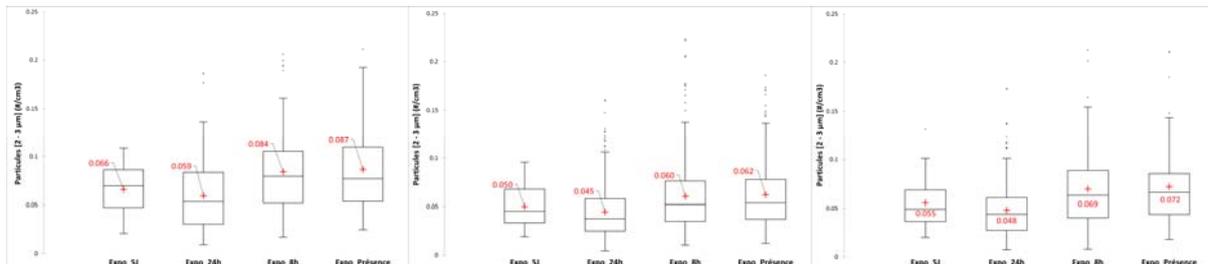


Figure 46 : Comparaison de la concentration d'exposition en particules de 2 à 3 µm de diamètre selon le mode de calcul sur les années 2013, 2014 et 2015 (de gauche à droite).

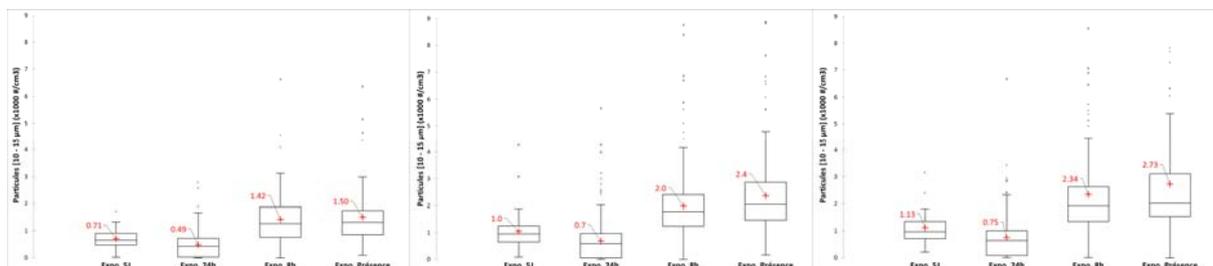


Figure 47 : Comparaison de la concentration d'exposition en particules de 10 à 15 µm de diamètre selon le mode de calcul sur les années 2013, 2014 et 2015 (de gauche à droite).

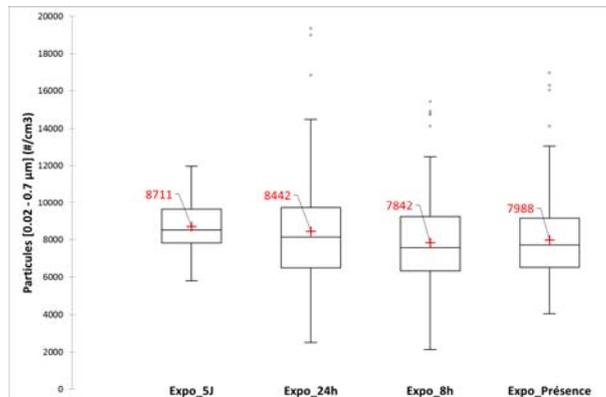


Figure 48 : Comparaison de la concentration d'exposition en particules ultrafines [0,02 – 0,7 µm] selon le mode de calcul pour 2015.

Selon le diamètre des particules considéré, le mode de calcul de l'exposition va sous-estimer ou surestimer l'exposition. Les particules de gros diamètre ou intermédiaire plus lourdes sont fortement sous-estimées si la période de présence n'est pas prise en compte alors que les plus fines sont surestimées. En termes de collecte de particules, ne pas considérer la présence de l'occupant devrait conduire en général à une masse plus faible de particules prélevées. Mais surtout, cela conduit à une déformation du profil granulométrique collecté qui n'est plus représentatif de ce à quoi l'occupant est réellement exposé.

L'espace de bureau étudié présente très peu de sources caractéristiques d'émission de particules, en particulier aucune activité de combustion. Ces activités en grande partie associées aux usages dans les logements pourraient venir modifier nos conclusions sur les particules fines en termes de comparaison de moyenne mais peu en termes de comparaison de médiane. Par contre, une grande partie du temps de présence dans le logement se fait la nuit dans la chambre avec une activité quasi-inexistante. Par ailleurs, le temps de présence dans le logement est de 16 heures en moyenne par jour, soit 2/3 du temps. Dans ces conditions, les différences observées entre les modes de calcul de l'exposition devraient être plus faibles que celles observées dans l'espace de bureau.

Finalement, l'exposition calculée sur 8 heures de présence théoriques constitue une bonne alternative pour déterminer l'exposition réelle des occupants avec une marge d'erreur très faible.

4.3 Influence de l'occupation

L'information sur l'occupation de l'espace de bureau est fournie à la fois par les détecteurs de mouvement et la mesure de la concentration de CO₂ dans le volume. Par contre, le nombre d'occupants à un instant donné n'est pas renseigné.

La variation du niveau de CO₂ peut permettre d'avoir une idée de l'occupation d'un local si aucune source de combustion n'est présente. Dès lors, toute augmentation ou stagnation de la concentration (au-delà du niveau de fond) correspond à une présence dans le local. A l'inverse, une diminution du niveau de concentration serait associée à l'absence d'occupants. Mais un doute subsiste pour des locaux bien ventilés où une diminution de la concentration peut être associée à une augmentation du taux de renouvellement d'air ou une variation du nombre d'occupants sans qu'il soit nul. Par ailleurs, en cas de recirculation d'air (ce qui n'est pas le cas de l'espace de bureau), les variations du niveau de CO₂ sont plus difficiles à interpréter. Les données recueillies par les détecteurs de mouvement sont donc préférables pour renseigner l'occupation. L'évolution de la concentration de CO₂ reste utile comme traceur d'une pollution d'origine métabolique.

4.3.1 Profil d'occupation

De 2012 et 2014, l'occupation du local a représenté de 555 à 685 heures (soit entre 6 % et 8 % de l'année) mais durant 22 % du temps en 2012 et 2013, l'information n'était pas disponible. En 2015, l'occupation sur les 7 premiers mois représentait 274 heures (soit 5 % du temps).

Durant la semaine, l'espace de bureau est occupé du lundi au vendredi entre 9 % et 15 % de la durée journalière (cumulée à l'échelle de l'année) de 2012 à 2014, avec le minimum observé le mercredi (en moyenne 10 % du temps). En 2015, la fréquence d'occupation journalière est globalement plus faible de 6 % (lundi) à 8 % (du mardi au jeudi). En fin de semaine, l'occupation est inexistante hormis durant très brèves et rares périodes (moins de 0,1 %). Les mardis et jeudis sont les jours où les occupants sont le plus présents.

L'occupation est également variable selon le mois de l'année avec un minimum en avril pour 2012, en décembre pour 2013 et 2014 et en juillet pour 2015. A l'échelle de la journée, la durée cumulée d'occupation par heure varie entre 0 % entre 21h00 et 05h00 à plus de 15 % entre 09h00 et 16h00, l'occupation du local étant en général plus fréquente le matin que l'après-midi (entre 4 % et 5 % de différence).

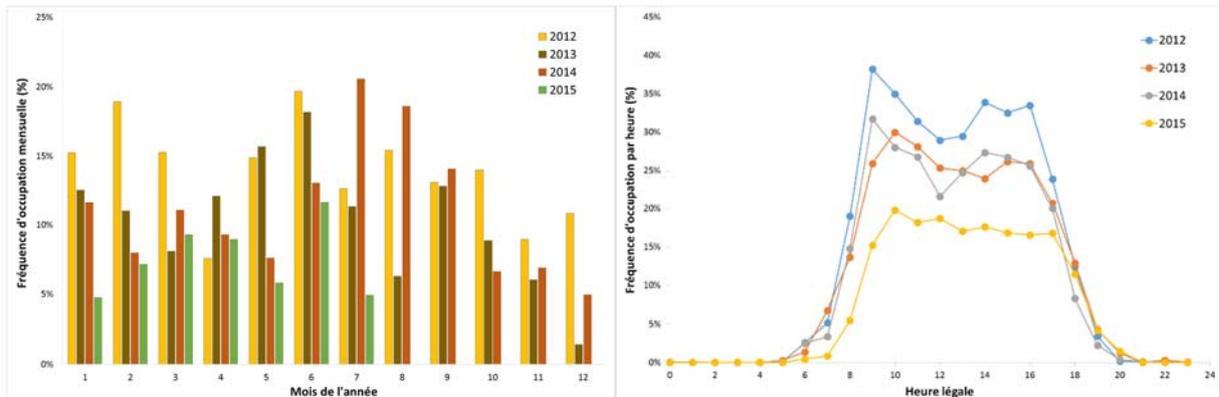


Figure 49 : Variation de la fréquence d'occupation en fonction du mois et de l'heure de la journée hors week-ends (% par rapport aux données renseignées).

Les profils d'occupation sont corrélés entre eux à l'échelle de la journée ou de la semaine, mais pas du tout à l'échelle de l'année où l'occupation varie d'un mois sur l'autre (de 26 % à 44 % selon les ans) et pour un même mois entre les années (de 5 % à 83 % selon les mois).

4.3.2 Influence de l'occupation sur la concentration en formaldéhyde

En 2014 et 2015, la mesure du formaldéhyde au pas de temps de 20 minutes implique de reconsidérer la variable présence renseignée elle au pas de temps de 1 minute. La variable présence a donc été ramenée au pas de temps de 20 minutes en considérant que le local était occupé sur ce laps de temps lorsqu'une présence était enregistrée sur au moins une minute.

Des écarts significatifs sont observés entre la concentration intérieure de formaldéhyde en présence par rapport à la concentration en absence d'occupants (test de Mann-Whitney, $p = 0.05$ en 2013, <0.01 en 2014 et <0.0001 en 2015). Ces écarts restent en médiane et en valeur moyenne très faibles : < 2 ppb en valeur absolue soit moins de 10 % d'écart. Aucune différence significative n'a été observée pour le ratio I/E de la concentration de formaldéhyde (calculé toutes les 20 minutes) en fonction de la présence ou de l'absence d'occupants et ce malgré une valeur moyenne du ratio plus élevée en absence d'occupants (valeurs extrêmes plus nombreuses).

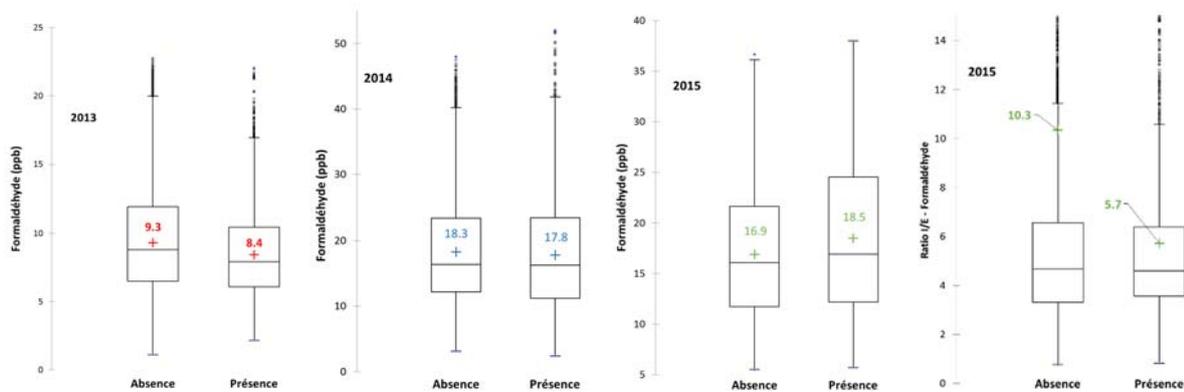


Figure 50 : Influence de la présence des occupants sur la concentration intérieure de formaldéhyde et sur le ratio intérieur / extérieur.

L'influence de l'occupation sur la variation de la concentration au cours du temps est illustrée sur les 4 figures suivantes. En l'absence d'occupants, la concentration intérieure de formaldéhyde tend à diminuer à l'échelle de la semaine entre le lundi et le vendredi, mais pas particulièrement à l'échelle de la journée. Le ratio I/E est maximum en fin de nuit et minimum en début de soirée. En présence d'occupants, l'évolution de la concentration intérieure en formaldéhyde varie selon l'année. A l'échelle de la semaine, comme à l'échelle de la journée, la concentration de formaldéhyde tend à baisser légèrement en 2013 et 2014, alors qu'elle tend plutôt à augmenter en 2015. La concentration extérieure de formaldéhyde suit également cette tendance en 2015. Au cours de la journée, l'augmentation est notamment marquée le matin et ne varie que très peu dans l'après-midi. Le ratio I/E élevé durant la nuit décroît pour atteindre un minimum en début de matinée puis augmente jusqu'en début d'après-midi après quoi il décroît à nouveau lentement. C'est cette évolution rapide du matin qui semble caractéristique de la présence des occupants.

Le fait que les concentrations intérieure et extérieure évoluent de manière conjointe signifie que d'autres phénomènes associés à l'occupation sont en partie responsables des évolutions observées, comme la saison par exemple. En dissociant selon la saison, l'augmentation matinale de la concentration intérieure en formaldéhyde reste marquée durant les deux saisons instrumentées en 2015 (janvier-avril et mai-juin). Durant la saison estivale, la concentration extérieure varie de façon similaire quel que soit le statut d'occupation de l'espace de bureau. Ce n'est pas tout à fait le cas entre janvier et avril, où la concentration extérieure varie de façon plus monotone durant les périodes de présence. Le ratio intérieur/extérieur moyen est pratiquement identique entre les deux saisons en présence d'occupants avec une diminution marquée en début de matinée. Par contre, en absence d'occupants, le ratio I/E varie beaucoup moins au cours de la journée de mai à juin que de janvier à avril.

Même si la présence des occupants ne peut normalement pas affecter la concentration extérieure de formaldéhyde, elle correspond à des périodes spécifiques

durant lesquelles des activités extérieures surviennent qui peuvent par contre affecter cette concentration. Les périodes d'absence des occupants correspondent à la nuit, aux week-ends, aux jours fériés essentiellement, périodes durant lesquelles les activités extérieures sont généralement différentes du reste de la semaine.

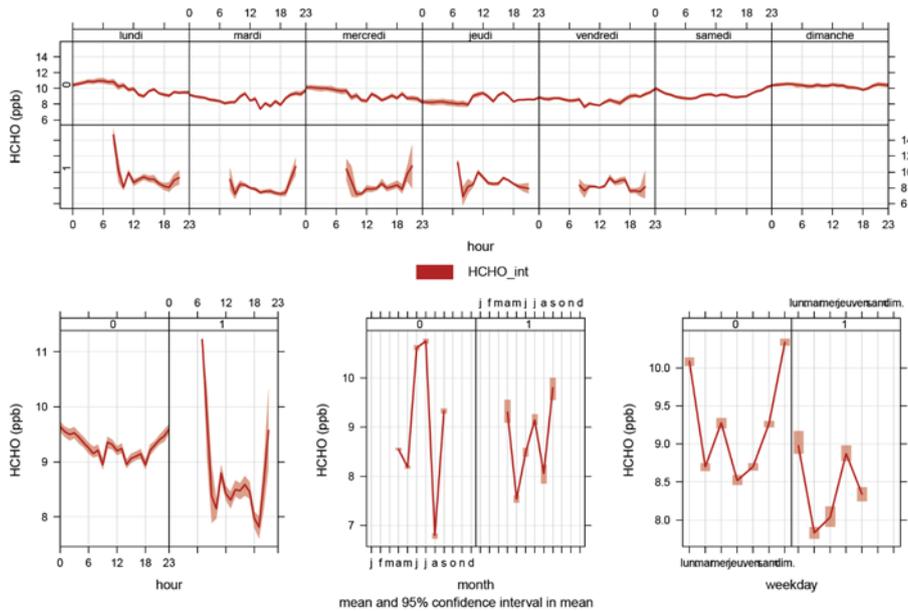


Figure 51 : Variation temporelle de la concentration en formaldéhyde selon la présence (1) ou l'absence d'occupants (0) en 2013.

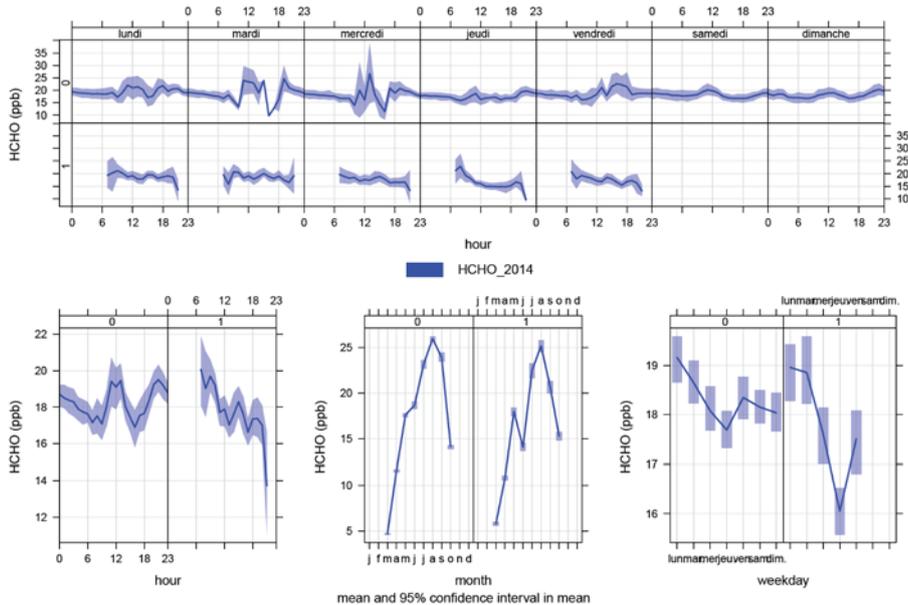


Figure 52 : Variation temporelle de la concentration en formaldéhyde selon la présence (1) ou l'absence d'occupants (0) en 2014.

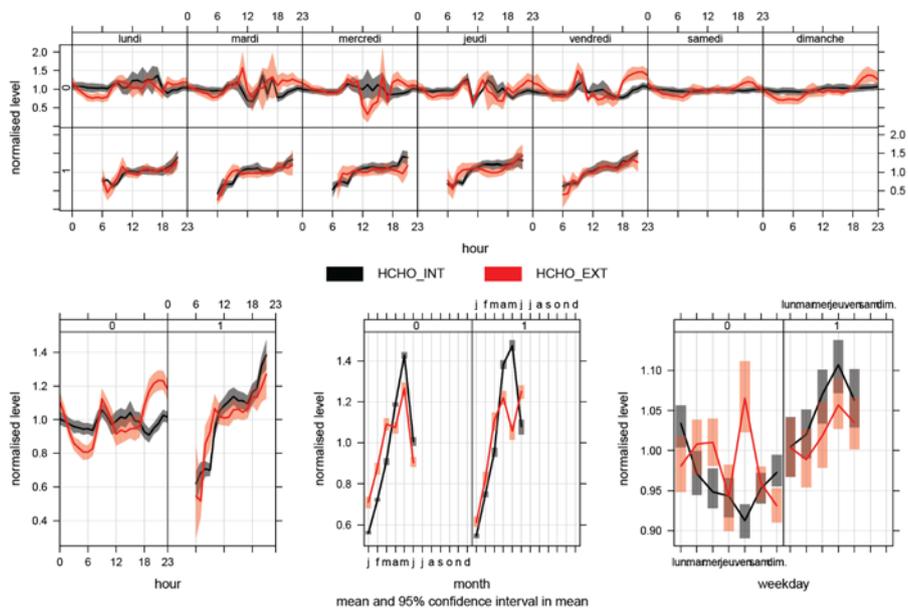


Figure 53 : Variation temporelle de la concentration en formaldéhyde selon la présence (1) ou l'absence d'occupants (0) en 2015.

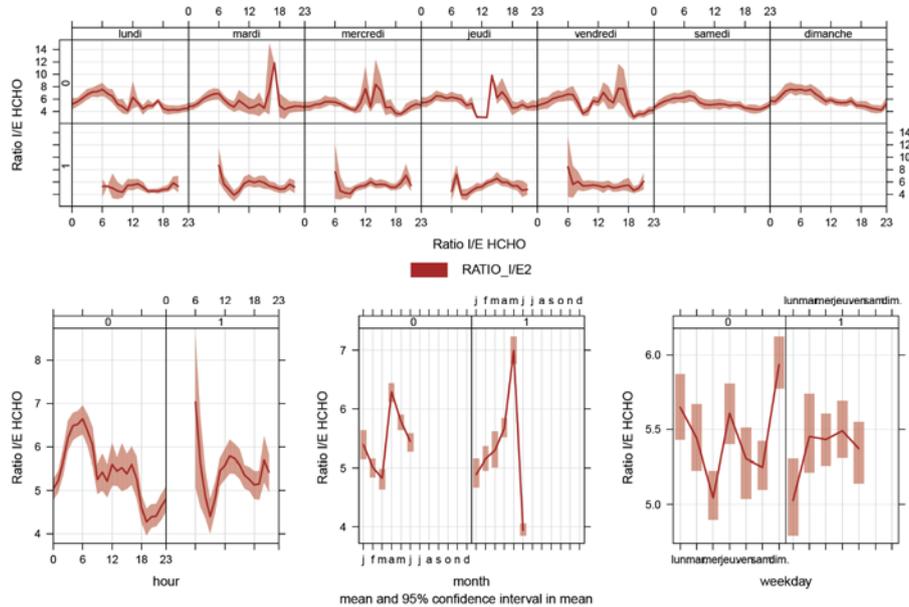


Figure 54 : Variation temporelle du ratio I/E en formaldéhyde selon la présence (1) ou l'absence d'occupants (0) en 2015.

Au final, malgré le faible écart observé en général dans la concentration moyenne intérieure de formaldéhyde selon l'occupation des locaux, cette dernière influence comment la concentration de formaldéhyde va varier au cours de la journée et au cours de la semaine. Cette influence n'est pas toujours la même. D'autres facteurs, comme

les activités exercées par les occupants ou leur action sur les ouvrants sont à même de venir modifier l'évolution de la concentration en formaldéhyde.

4.3.3 Influence de l'occupation sur la concentration en particules

La concentration de particules dans l'espace de bureau dépend également de la présence des occupants, mais de façon différentielle selon le diamètre des particules. Pour illustrer ce point, le ratio des concentrations de particules en période d'occupation (O) par rapport à la période d'inoccupation (U) a été calculé chaque jour. En 2015, le ratio O/U des particules ultrafines a été déterminé en leur affectant le diamètre médian en nombre (CMD) fourni par l'instrument pour chaque période. Le CMD moyen était de 0,05 μm sans différence significative en absence ou en présence d'occupants.

Le ratio O/U moyen journalier est significativement supérieur à 1 pour les particules de diamètre supérieur à 1 μm . Il augmente avec le diamètre des particules et devient supérieur à 3 à partir de 10 μm . Pour ces tailles de particules, des variations importantes sont observées en fonction des jours. Le profil d'évolution du rapport O/U en fonction du diamètre des particules reste similaire selon les différentes années.

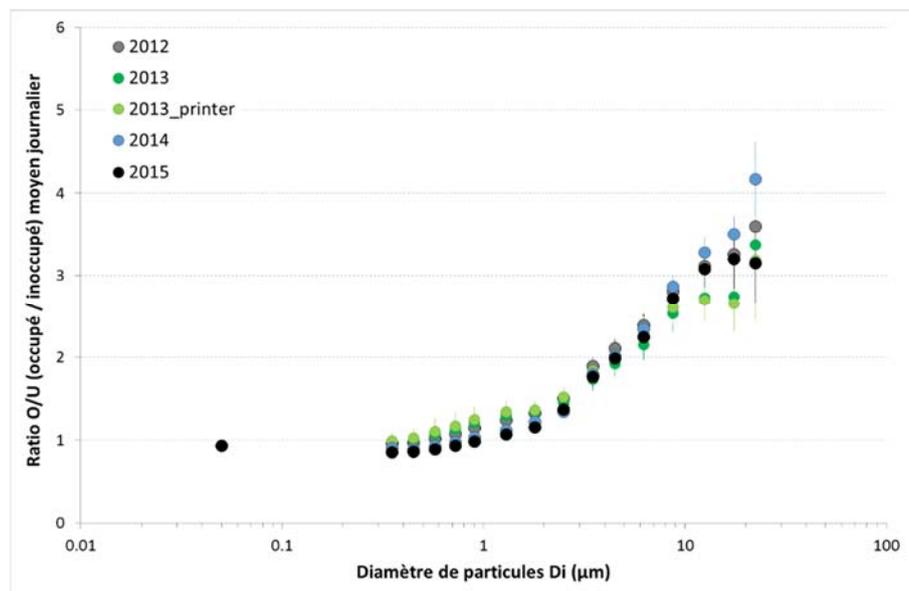


Figure 55 : Evolution du ratio O/U moyen journalier de la concentration en particules selon le diamètre et l'année (barres d'erreur : intervalle de confiance de la moyenne à 95 %)

Le ratio I/E peut également évoluer selon la présence des occupants. Ce ratio a été calculé à partir des moyennes journalières en considérant uniquement les périodes d'absence ou de présence des occupants. Il est ensuite représenté en fonction du

diamètre des particules. Pour les particules inférieures à 2 μm , le ratio I/E reste inférieur à 1 que ce soit en présence ou en absence d'occupants. En absence d'occupants, le ratio I/E reste inférieur à 1 pour les particules de moins de 5 μm . En présence d'occupants, le ratio I/E devient supérieur à 1 pour les particules de plus de 3 μm . Il est maximal pour les grosses particules comprises entre 10 et 20 μm avec des valeurs de ratio pouvant dépasser 10. En l'absence d'occupants, le ratio I/E de ces grosses particules est également élevé, mais il ne dépasse que rarement la valeur de 4. Cette valeur élevée peut être liée au fait que ces particules remises en suspension ou générées par les activités des occupants mettent un certain temps à se déposer une fois les occupants partis.

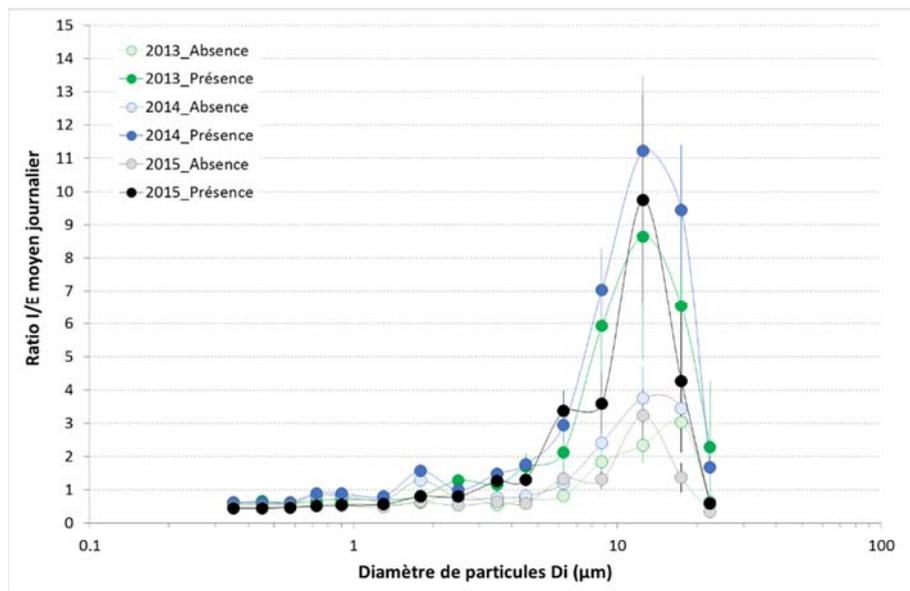


Figure 56 : Evolution du ratio I/E moyen journalier de la concentration en particules selon leur diamètre et l'année (barres d'erreur : intervalles de confiance de la moyenne à 95 %)

L'impact de la présence des occupants sur la variation temporelle de la concentration en particules est illustré dans les 3 figures suivantes en prenant 2014 comme année de référence (la différence selon les années étant faible) et les trois tailles D_i de particules suivantes 0,35 ; 2,5 et 12,5 μm . Les concentrations intérieures et extérieures sont normalisées selon leur moyenne annuelle.

La variation des particules les plus fines à l'intérieur suit celle observée à l'extérieur avec des variations moins prononcées quelle que soit la présence des occupants. Elle reste inférieure à la concentration extérieure à tout moment.

Pour les particules intermédiaires de 2,5 μm , la concentration intérieure suit les variations de la concentration extérieure tout en restant systématiquement inférieure

en absence d'occupants. Alors qu'en présence d'occupants, l'évolution de la concentration intérieure épouse quasiment celle de la concentration extérieure.

Les grosses particules sont systématiquement supérieures en concentration à l'intérieur par rapport à l'extérieur et ce quel que soit la présence d'occupants à l'exception du week-end. Par contre, leur concentration intérieure est toujours plus élevée en présence d'occupants.

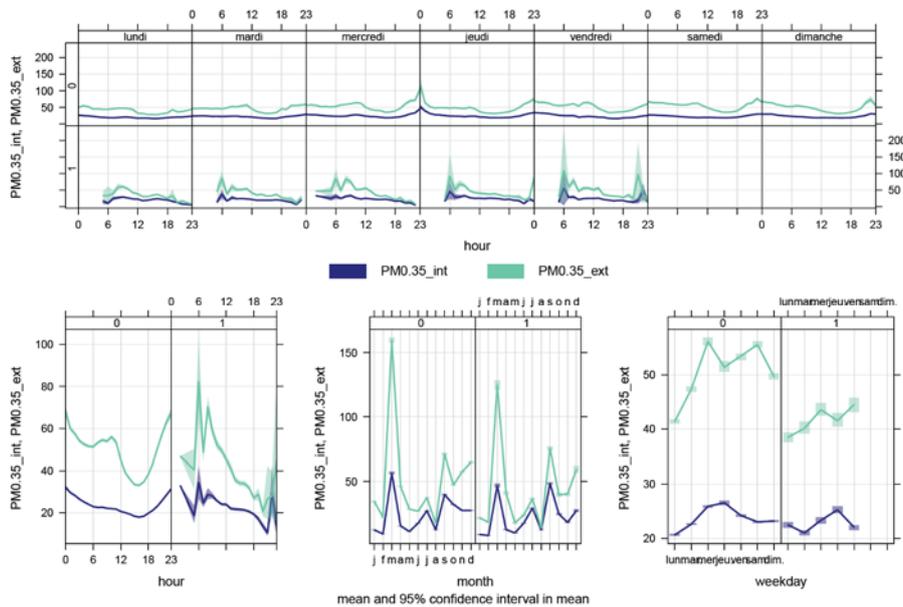


Figure 57 : Variation temporelle de la concentration moyenne intérieure en particules [0.3-0.4 µm] selon la présence des occupants en 2014 (en particules/cm³)

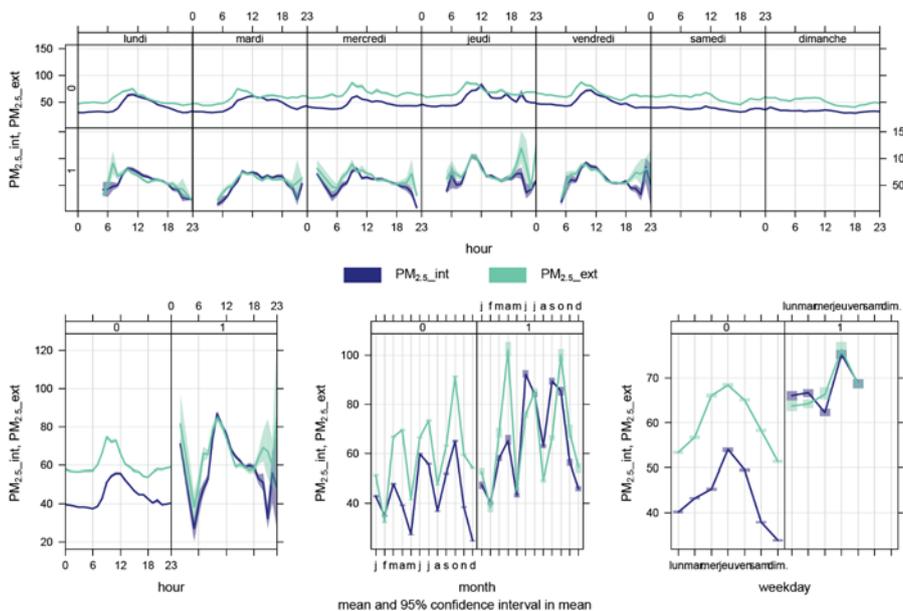


Figure 58 : Variation temporelle de la concentration intérieure moyenne en particules [2-3 µm] selon la présence des occupants en 2014 (x1000 particules/cm³)

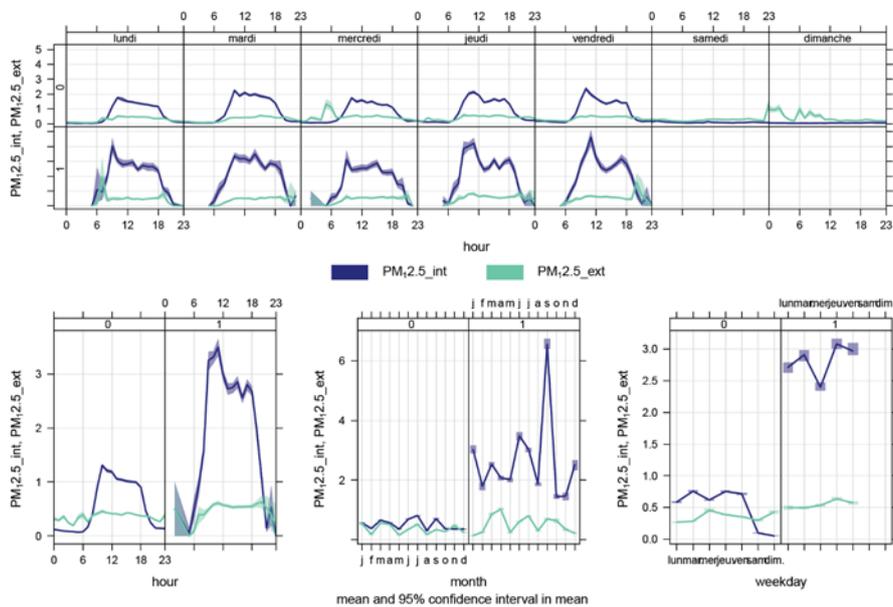


Figure 59 : Variation temporelle de la concentration intérieure moyenne en particules [10-15 μm] selon la présence des occupants en 2014 ($\times 1000$ particules/ cm^3)

L'influence de l'occupation sur la concentration en particules est plus marquée que pour le formaldéhyde. Seule l'influence de la simple présence d'occupants a pu être évaluée. L'influence de leur nombre n'a pas pu être examinée, cette information n'étant pas disponible. La présence d'occupants augmente la surface d'échange disponible en fournissant une surface corporelle d'environ $1,7 \text{ m}^2$ par occupant, sans compter la surface développée des vêtements. Par ailleurs, les occupants respirent à un débit d'environ $0,8 \text{ m}^3/\text{heure}$ chacun pour une activité sédentaire. La respiration peut filtrer en partie les polluants entrant dans le corps humain, notamment certaines tailles de particules, mais cela reste faible au regard du taux renouvellement d'air induit par le système spécifique de ventilation ou l'ouverture des ouvrants. Cette dernière se manifeste principalement en présence des occupants. Il est nécessaire de la prendre en compte tout en considérant l'occupation pour bien évaluer l'influence propre à chacun.

4.4 Influence de l'ouverture des ouvrants

4.4.1 Description de l'ouverture des ouvrants

L'ouverture des fenêtres a pu être documentée grâce aux contacteurs installés depuis le 20 janvier 2012. Ces contacteurs ne permettent toutefois pas de déterminer l'amplitude d'ouverture de chaque fenêtre. Les fenêtres de l'espace de bureau sont des fenêtres coulissantes qui n'entraînent pas un encombrement du volume intérieur comme des fenêtres battantes par exemple. Le type de fenêtre et sa manœuvrabilité peut possiblement influencer son ouverture et/ou l'amplitude de son ouverture.

En 2012 et 2013, 20 % des données sont manquantes du fait de perte de connexion entre les contacteurs et la CSTBox ou de pannes de courant. Ce manque d'informations n'est pas réparti de la même façon entre les mois de mesure. Ainsi, les mois de janvier, février, octobre et décembre 2012 présentent plus de 40 % d'informations manquantes. En 2013, ce sont les mois de mars, novembre et décembre qui présentent plus de 40 % d'informations manquantes. A l'inverse, les mois de mai à juillet sont les moins affectés par le manque de données.

L'ouverture d'au moins une fenêtre a été plus fréquente en 2012 et 2014 (respectivement 55 % et 45 % du temps) par rapport à 2013 (26 % du temps). Il n'y a toutefois pas de différences entre 2012 et 2013 sur la fréquence d'ouverture de toutes les fenêtres (respectivement 6 % et 7 %). En 2014, cette fréquence est plus faible autour de 2,6 %. Les périodes durant lesquelles toutes les fenêtres étaient fermées ne représentaient que 21 % en 2012 contre respectivement 53 % et 55 % en 2013 et 2014. En 2015 durant le premier semestre, au moins une fenêtre était ouverte 24 % du temps et 76 % du temps elles sont restées fermées.

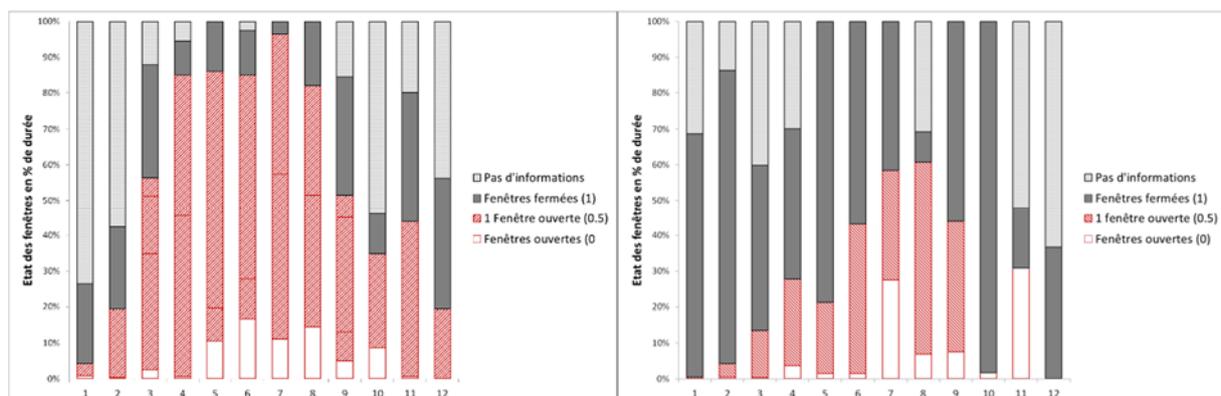


Figure 60 : Variation de l'ouverture des fenêtres au cours de l'année 2012 (gauche) et 2013 (droite).

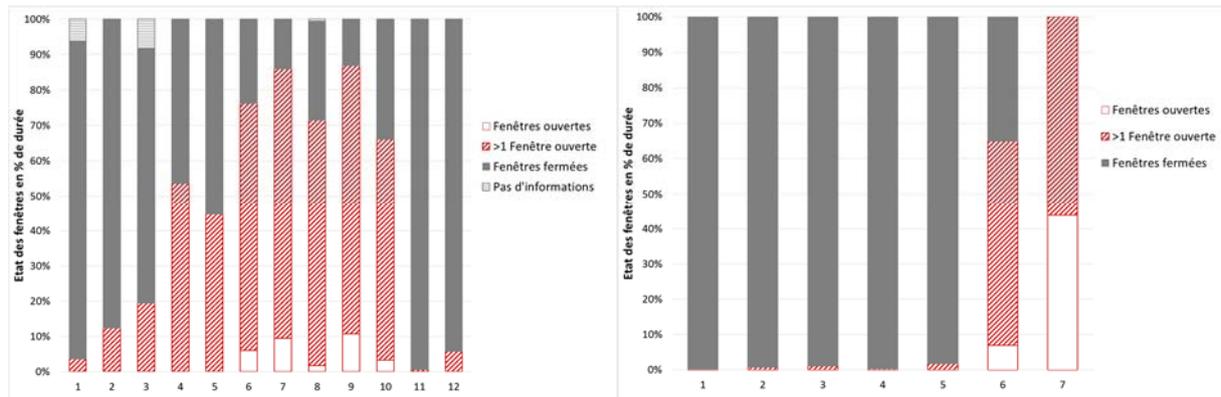


Figure 61 : Variation de l'ouverture des fenêtres en 2014 (gauche) et en 2015 (droite).

L'ouverture des fenêtres est variable dans le temps selon les mois de l'année, l'heure de la journée et de manière moins marquée selon le jour de la semaine. Ainsi, les fenêtres sont plus souvent ouvertes en période estivale (96 % du temps en juillet 2012 avec au moins une fenêtre ouverte). Cette fréquence mensuelle varie entre 4 % et 96 % du temps en 2012, entre 0,5 % et 61 % en 2013, entre 0,6 % et 87 % en 2014 et entre 0 % et 100 % de janvier à juillet 2015.

A l'échelle de la semaine, la fréquence d'ouverture d'au moins une fenêtre ne varie pas beaucoup d'un jour sur l'autre. Elle se situe entre 50 % et 60 % en 2012 et entre 20 % et 30 % en 2013 et 2015, et autour de 40 % - 50 % en 2014. Par contre, l'ensemble des fenêtres restent ouvertes moins souvent durant le week-end et rarement toutes ensembles. Enfin à l'échelle d'une journée, si l'ouverture des fenêtres survient pendant les heures d'occupation, elles ne sont pas toujours refermées la nuit. Si bien que le temps durant lequel une ou toutes les fenêtres sont ouvertes varie assez peu entre le jour et la nuit en ce qui concerne cet espace de bureau en particulier. Le matin est en général un peu plus propice à l'ouverture des fenêtres que l'après-midi.

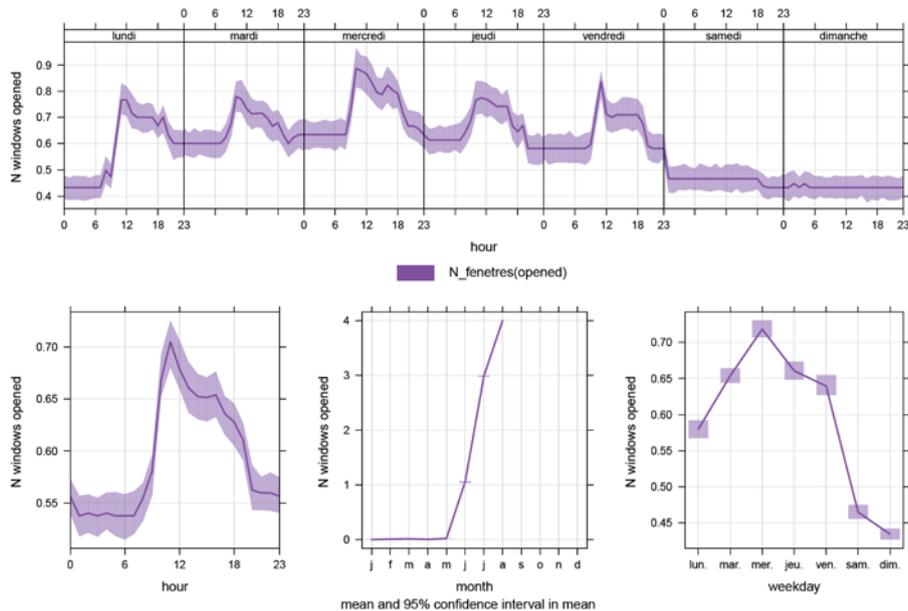


Figure 62 : Variation du nombre moyen de fenêtres ouvertes durant le premier semestre 2015.

4.4.2 Influence de l'ouverture des fenêtres

4.4.2.1 Influence de l'ouverture des fenêtres sur la concentration en formaldéhyde

Pour déterminer cette influence et prendre également en compte les périodes d'occupation, les concentrations intérieures ont été moyennées à l'échelle de la semaine pour permettre de déterminer un ratio de concentration (fenêtres ouvertes)/(fenêtres fermées) en absence ou en présence des occupants. L'échelle de la journée était trop courte car elle ne couvre pas forcément l'ensemble de ces situations. Ainsi, la moyenne des concentrations de formaldéhyde a été calculée sur chaque période d'absence (P0) ou de présence des occupants (P1) définie également par le nombre de fenêtres ouvertes (F0 à F3+). Les périodes qui représentaient moins de 10 minutes en cumulé sur la semaine ont été écartées.

En 2013 et 2015, l'ouverture d'au moins une fenêtre diminue en moyenne de façon modérée la concentration intérieure du formaldéhyde entre 19 % et 26 %. L'impact est surtout visible lors de l'ouverture d'une seule fenêtre. Ouvrir d'autres fenêtres ne permet pas de diminuer à nouveau la concentration de formaldéhyde. Cette influence se traduit également au niveau du ratio I/E moyen hebdomadaire de l'ordre de 5 lorsque les fenêtres sont fermées en présence ou absence des occupants, et qui baisse à 3 ou 4 lorsqu'elles sont ouvertes. Le ratio O/F moyen est de l'ordre de 0,8 pour une fenêtre ouverte et descend à un peu plus de 0,6 pour au moins trois fenêtres ouvertes avec peu d'écart selon la présence de l'occupant.

L'influence des fenêtres sur la concentration intérieure en formaldéhyde n'est pas perceptible en moyenne sur l'année 2014. Ceci étant lié aux fortes variations observées au cours de l'année aussi bien lorsque les fenêtres sont fermées que

lorsqu'elles sont ouvertes. Néanmoins, le ratio O/F tend à diminuer avec le nombre de fenêtres ouvertes de 1,1 à 0,9. Dans certains cas, la concentration de formaldéhyde mesurée en périodes de fenêtres ouvertes dépasse celle observée toutes fenêtres fermées et ce même en ramenant la comparaison à l'échelle de la semaine. L'influence de l'ouverture des fenêtres sur la concentration intérieure en formaldéhyde n'est donc pas systématique et fait intervenir d'autres facteurs qui varient dans le temps. La présence des occupants ne semble pas jouer un rôle prépondérant.

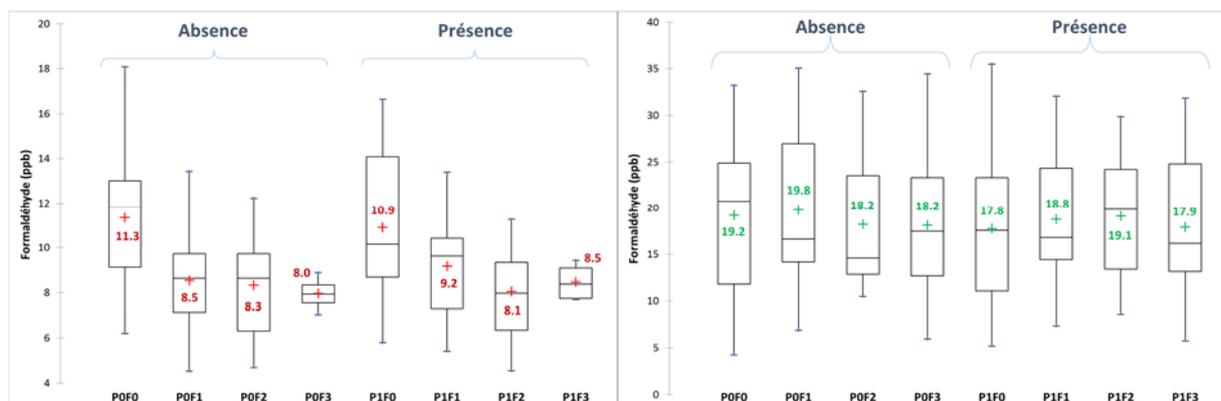


Figure 63 : Variation de la concentration hebdomadaire en formaldéhyde selon le nombre de fenêtres ouvertes (0 à 3+) en l'absence (P0) ou en présence (P1) des occupants en 2013 (gauche) et en 2014 (droite)

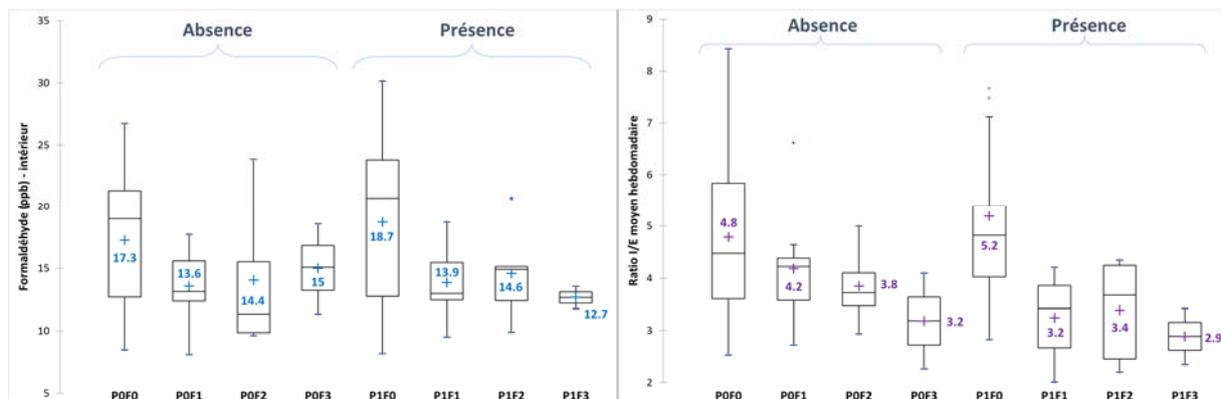


Figure 64 : Variation de la concentration hebdomadaire et du ratio I/E du formaldéhyde selon le nombre de fenêtres ouvertes (0 à 3+) en l'absence (P0) ou en présence (P1) des occupants en 2015.

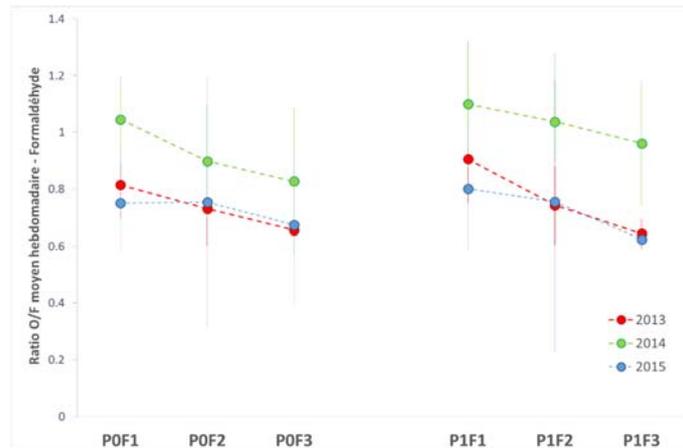


Figure 65 : Variation du ratio O/F du formaldéhyde selon le nombre de fenêtres ouvertes (0 à 3+) en l'absence (P0) ou en présence (P1) des occupants.

L'influence de l'ouverture des fenêtres peut être variable dans le temps. La Figure 66 montre les différences observées dans les profils temporels du formaldéhyde en 2015 selon l'état des ouvrants. Fenêtres ouvertes, le profil moyen journalier de la concentration intérieure suit les mêmes variations que la concentration extérieure du formaldéhyde avec des niveaux qui restent toutefois plus élevés. Ce n'est pas du tout le cas en situation de fenêtres fermées, où l'évolution du formaldéhyde suit sans surprise l'allure générale observée préalablement (les périodes où les fenêtres restent fermées étant plus nombreuses).

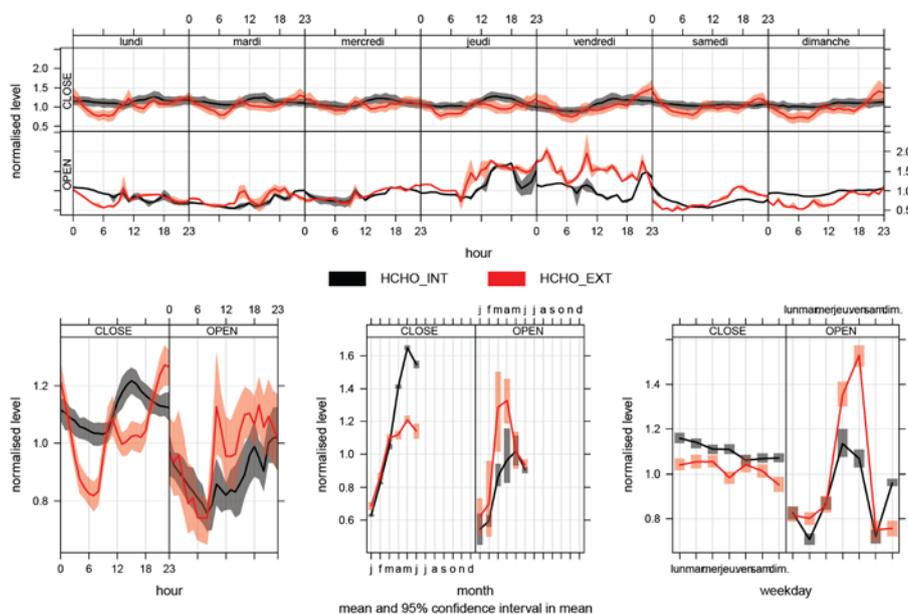


Figure 66 : Variation de la concentration en formaldéhyde intérieure et extérieure selon l'ouverture des fenêtres sans tenir compte de la présence des occupants en 2015.

En considérant la présence des occupants, l'influence de l'ouverture des fenêtres sur le profil moyen journalier reste marquée en 2013 et 2015. En l'absence d'occupants, le profil intérieur suit à peu près les variations extérieures lorsque les fenêtres sont ouvertes et la différence avec le niveau observé fenêtres fermées est assez nette. En présence d'occupants, cette influence est moins prononcée. Pour l'année 2014, aucune différence n'est observée sur le profil moyen journalier selon l'ouverture de fenêtres en absence d'occupants. Le niveau par contre augmente en ouvrant les fenêtres en présence d'occupants, niveau qui tend à décroître au fil de la journée.

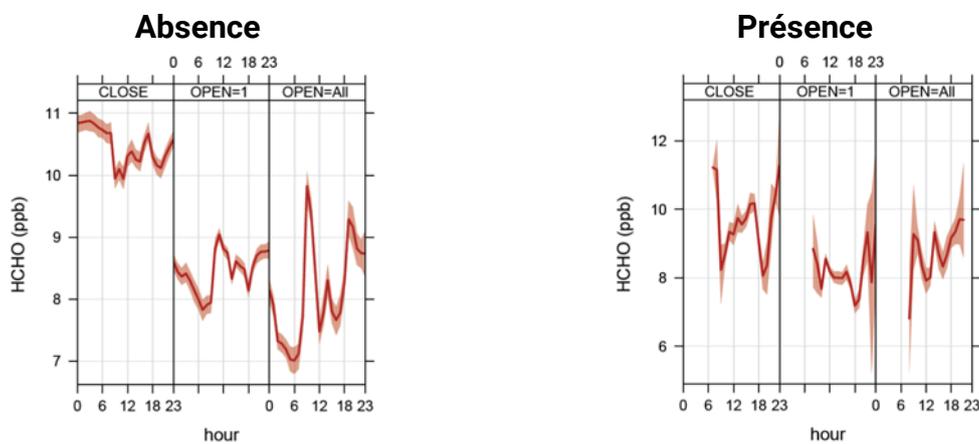


Figure 67 : Variation de la concentration en formaldéhyde selon l'ouverture des fenêtres en absence ou en présence d'occupants en 2013.

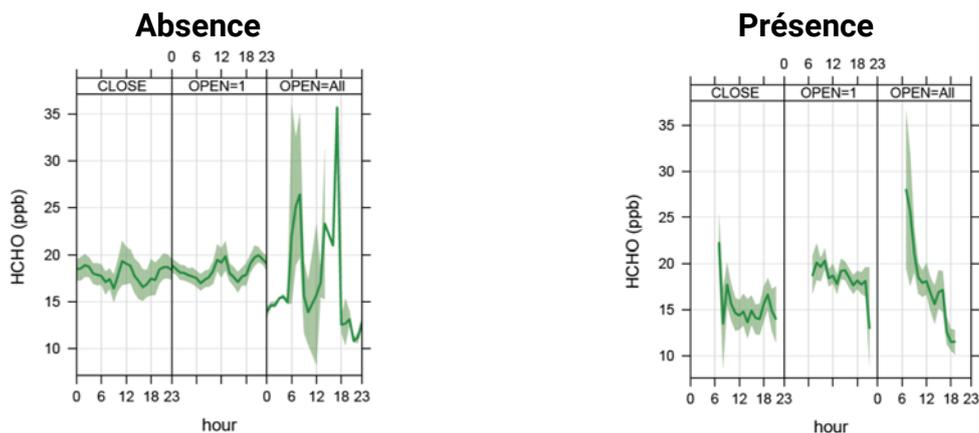


Figure 68 : Variation de la concentration en formaldéhyde selon l'ouverture des fenêtres en absence ou en présence d'occupants en 2014.

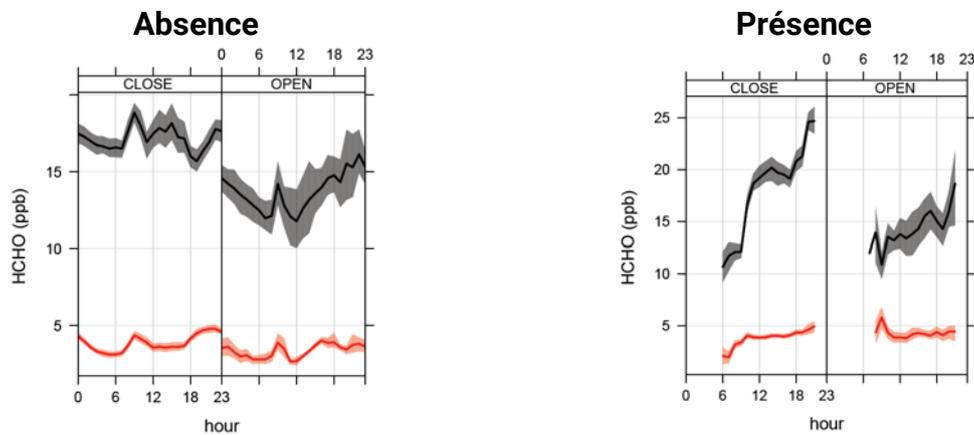


Figure 69 : Variation de la concentration en formaldéhyde selon l'ouverture des fenêtres en absence ou en présence d'occupants en 2015.

Tant la présence que l'ouverture des fenêtres présente une influence modérée sur la concentration intérieure du formaldéhyde, influence qui varie selon les périodes considérées. Cette influence variable souligne le rôle d'autres facteurs qui varient également dans le temps : les paramètres climatiques, la réactivité chimique ou la mise en œuvre de sources spécifiques dans le bâtiment au voisinage de l'espace de bureau.

4.4.2.2 Influence de l'ouverture des ouvrants sur la concentration en particules

L'ouverture des fenêtres joue également un rôle sur la concentration des particules. Elle fait intervenir plusieurs phénomènes : l'apport de particules d'origine extérieure à l'intérieur, l'évacuation de particules de l'intérieur vers l'extérieur, une modification des conditions thermoaérauliques pouvant affecter le dépôt et la resuspension de particules. L'ouverture d'une seule fenêtre n'est pas toujours suffisante pour générer un courant d'air et accentuer le renouvellement d'air. La contribution relative de ces différents phénomènes va jouer sur le sens et l'intensité de variation de la concentration de particules dans l'espace de bureau.

Le rapport de la concentration de particules en période de fenêtre ouverte sur la concentration observée fenêtres fermées (ratio O/F) a été utilisé pour mettre en évidence l'influence de l'ouverture des fenêtres en fonction de la présence des occupants. Les périodes ont été définies à l'échelle de la semaine pour le calcul des ratios O/F. Les périodes qui selon les cas représentaient moins de 10 minutes cumulées n'ont pas été prises en compte.

En l'absence d'occupants, le ratio O/F moyen hebdomadaire varie entre 1,2 et 4 selon le diamètre des particules pour une fenêtre ouverte, de 1,4 à plus de 10 pour deux fenêtres ouvertes et de 1,2 à plus de 20 pour 3 fenêtres ouvertes ou plus. Pour les plus

grosses particules, le ratio O/F varie énormément selon les semaines si bien qu'il devient impossible de différencier l'influence de l'ouverture de 3 fenêtres ou plus de celle de 2 fenêtres. Par contre, ouvrir une seule fenêtre a significativement moins d'impact en l'absence d'occupants qu'en ouvrir plusieurs.

La présence d'occupants vient bouleverser ces observations. En effet, l'augmentation du ratio O/F avec le diamètre des particules n'est plus apparente. Le ratio O/F évolue de manière général entre 1 et 2 sauf en 2015 où les fenêtres étaient ouvertes lors d'un épisode de pollution particulaire. L'ouverture de plus d'une fenêtre n'augmente pas systématiquement le ratio O/F. Cela peut s'expliquer par le fort impact de la présence des occupants sur la concentration intérieure des plus grosses particules et l'apport extérieur de ces grosses particules apparaît dès lors négligeable. Par contre, en absence d'occupants, les niveaux intérieurs de grosses particules sont redevenus très faibles et la contribution extérieure devient prépondérante.

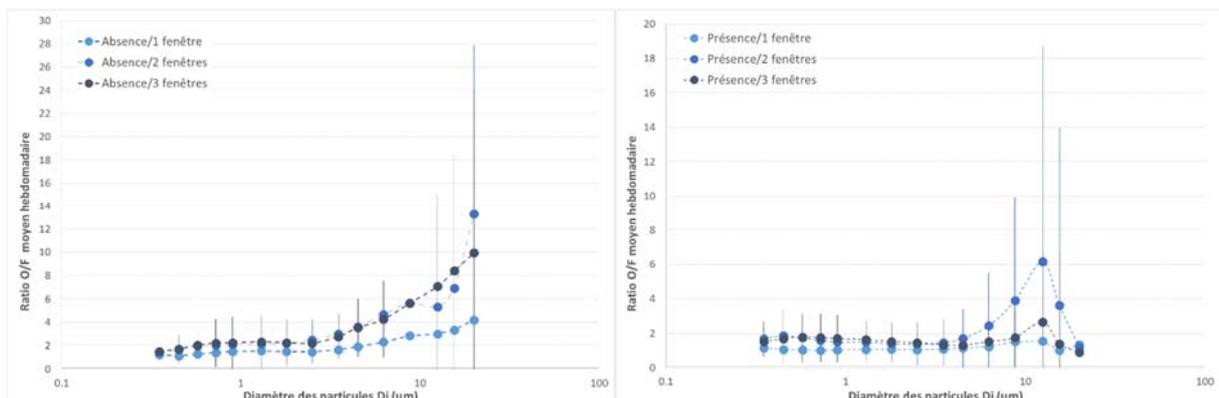


Figure 70 : Evolution du ratio O/F moyen selon la présence d'occupants, le nombre de fenêtres ouvertes et le diamètre des particules sur l'année 2013.

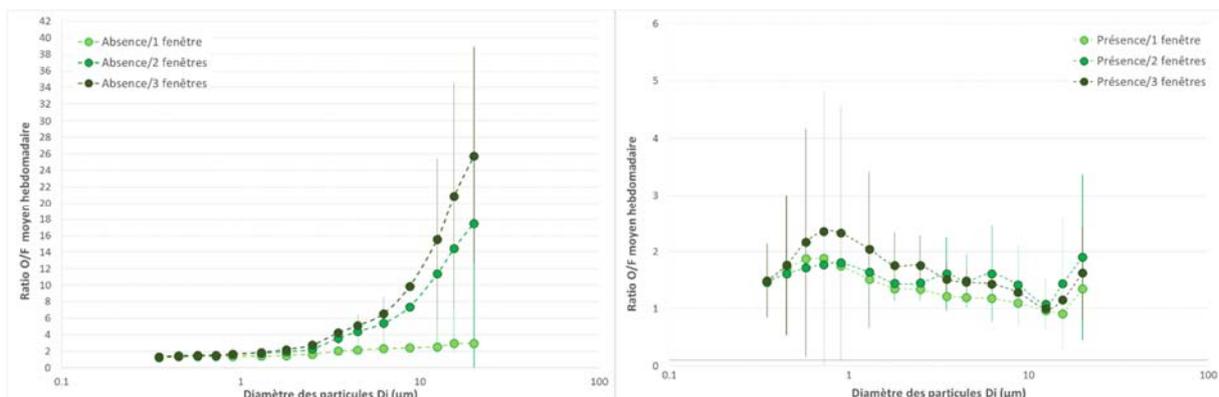


Figure 71 : Evolution du ratio O/F moyen selon la présence d'occupants, le nombre de fenêtres ouvertes et le diamètre des particules sur l'année 2014.

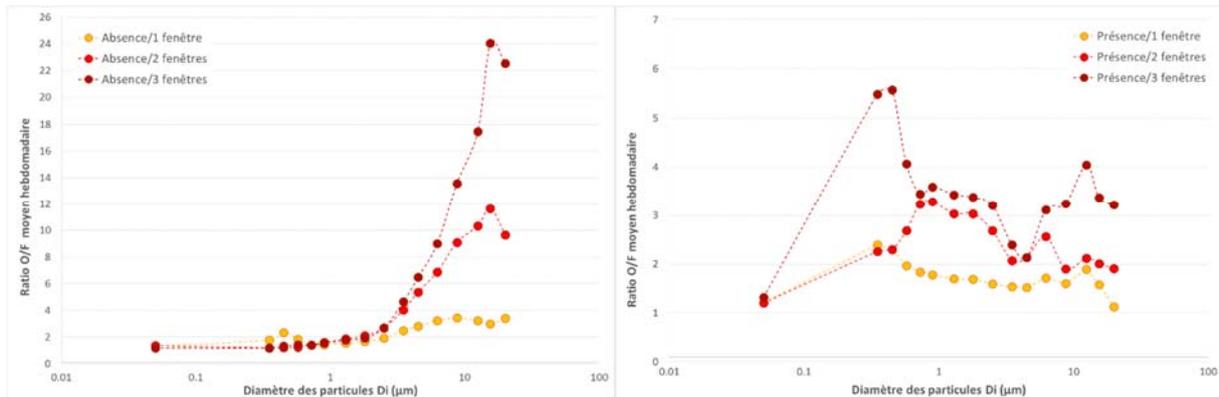


Figure 72 : Evolution du ratio O/F moyen selon la présence d'occupants, le nombre de fenêtres ouvertes et le diamètre des particules sur l'année 2015.

Outre les ratios O/F, il est important de considérer également comment varie la concentration extérieure des particules aux moments où les fenêtres sont ouvertes. Pour en tenir compte, le meilleur moyen est de regarder l'influence de l'ouverture des fenêtres sur le ratio I/E.

En l'absence d'occupants, le ratio I/E varie de 0,5 (particules fines) à 2 (grosses particules) toutes fenêtres fermées. Un ratio I/E plus faible autour de 0,2 est observé pour les particules > 20 µm, mais leur efficacité de collecte n'est pas optimale pour ces compteurs. L'ouverture d'une fenêtre entraîne une légère augmentation du ratio I/E pour les plus grosses particules fines et intermédiaires (entre 0,2 et 0,3 fois le nombre de fenêtres ouvertes en moyenne). Le ratio I/E dépasse 1 lorsque 3 fenêtres ou plus sont ouvertes. L'augmentation est encore plus importante pour les plus grosses particules jusqu'à 400% mais elle n'est plus linéaire selon le nombre de fenêtres ouvertes.

En présence d'occupants, le ratio I/E fenêtres fermées s'étend de 0,5 à plus de 10 pour les grosses particules. L'augmentation du ratio I/E est moins importante pour les plus fines particules et le ratio I/E dépasse rarement 1 pour les particules < 0,65 µm même en ouvrant 3 fenêtres ou plus, contrairement à ce qui est observé en absence des occupants. Pour ces particules, l'apport extérieur est sans doute compensé par des phénomènes de pertes plus importantes du fait de la présence des occupants ou d'évènements survenant lorsque les occupants sont présents. Dans le premier cas, il pourrait s'agir d'une augmentation du dépôt sur les surfaces liée soit à l'apport de surfaces disponibles (les occupants et leurs vêtements), soit à une augmentation de la vitesse de dépôt du fait d'une plus grande homogénéisation du volume assurée par le mouvement des occupants. Dans le second cas, l'augmentation de la concentration extérieure en particules fines durant les heures de présence des occupants pourrait également venir diminuer le ratio I/E.

Le ratio I/E des plus grosses particules supérieures à 3 µm se situe déjà au-dessus de 1 avec toutes les fenêtres fermées. Pour ces particules, l'ouverture des fenêtres conduit à une modification du ratio I/E plus aléatoire pouvant tout aussi bien

l'augmenter que le baisser (mais rarement en deçà de 1) et ce indépendamment du nombre de fenêtres ouvertes. Dans ces conditions, l'influence de l'ouverture des fenêtres sur les plus grosses particules reste très difficile à déterminer. L'apport extérieur, la remise en suspension de particules déposées, l'évacuation des particules en suspension, la modification des champs thermo-aérauliques et leur impact à la fois sur le dépôt de particules et l'homogénéité de la concentration dans le volume sont autant de facteurs qui vont définir l'impact de l'ouverture d'une ou plusieurs fenêtres. A cela, il est nécessaire de rajouter le nombre d'occupants présent dans le volume au moment où la ou les fenêtres sont ouvertes, car la présence des occupants reste le principal facteur responsable d'un ratio I/E élevé pour les grosses particules.

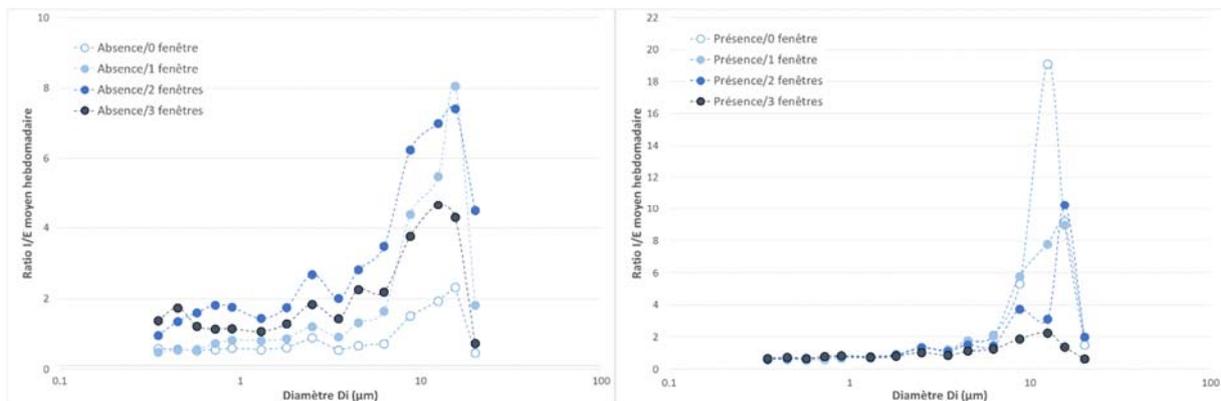


Figure 73 : Evolution du ratio I/E moyen selon la présence d'occupants, le nombre de fenêtres ouvertes et le diamètre des particules sur l'année 2013.

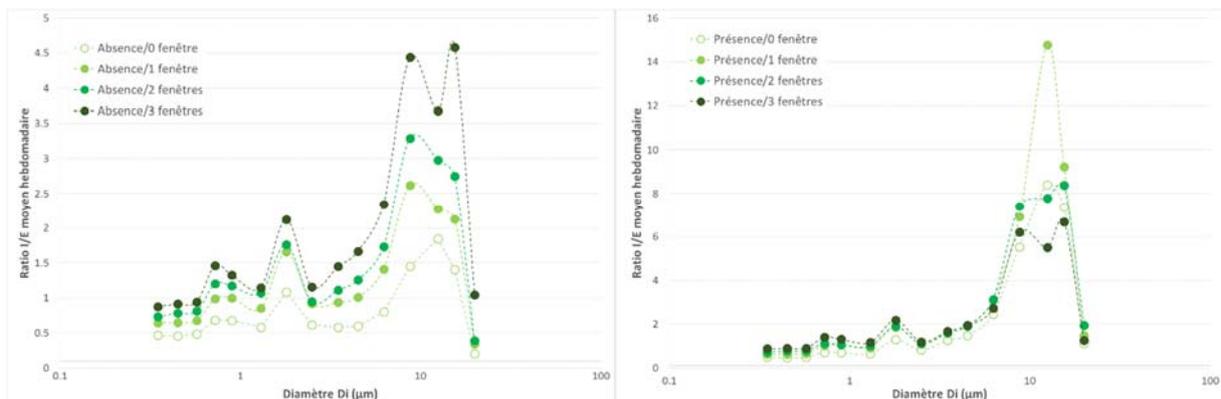


Figure 74 : Evolution du ratio I/E moyen selon la présence d'occupants, le nombre de fenêtres ouvertes et le diamètre des particules sur l'année 2014.

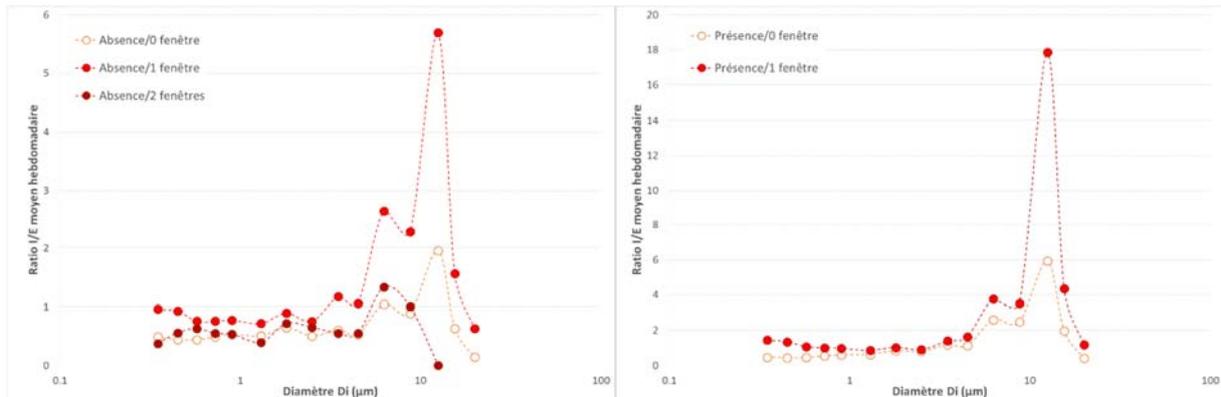


Figure 75 : Evolution du ratio I/E moyen selon la présence d'occupants, le nombre de fenêtres ouvertes et le diamètre des particules sur l'année 2015.

En termes de variations temporelles, la concentration intérieure de particules de 0,35 µm reste systématiquement inférieure à la concentration extérieure et en moyenne du même niveau quelle que soit la présence des occupants ou l'ouverture des fenêtres. Ces deux facteurs viennent finalement très peu perturbés les variations de cette gamme de particules.

Le même phénomène est observé pour les particules de 2,5 µm à l'exception du fait que l'ouverture des fenêtres augmente la concentration intérieure de ces particules qui se retrouve au même niveau que la concentration extérieure. Pour un même état d'ouverture des fenêtres, la concentration des particules de 2,5 µm est en moyenne du même ordre en absence ou en présence des occupants. De ce fait, l'influence de l'occupant sur la concentration de cette gamme de taille intervient surtout au travers de son action sur les ouvrants et très peu du fait de sa présence.

A l'inverse, les plus grosses particules de 12,5 µm sont plus nombreuses dans l'air de l'espace de bureau en présence des occupants et systématiquement plus concentrées qu'à l'extérieur. L'impact de l'ouverture des fenêtres est moins prononcé, sauf lorsque l'ensemble des fenêtres sont ouvertes. L'ouverture des fenêtres ne vient pas diminuer la concentration intérieure de ces particules, mais au contraire l'augmenter. Pour cette taille de particules, l'occupant influence plus par sa présence que par son action sur les ouvrants.

Pour les particules ultrafines mesurées en 2015, l'ouverture des ouvrants (au moins une fenêtre ouverte) modifie le profil moyen journalier observé en absence d'occupants. En effet, la concentration intérieure de particules ultrafines est plus élevée en journée lors de l'ouverture des fenêtres et cette différence est encore plus marquée l'après-midi. C'est surtout durant le week-end que cette augmentation de l'après-midi est la plus évidente. Cette situation n'est pas observée par contre en présence des occupants durant laquelle l'ouverture des fenêtres ne modifie pratiquement pas le profil moyen journalier. L'absence de données extérieures pour

cette gamme de taille ne permet pas d'avancer d'hypothèses pour expliquer ce phénomène.

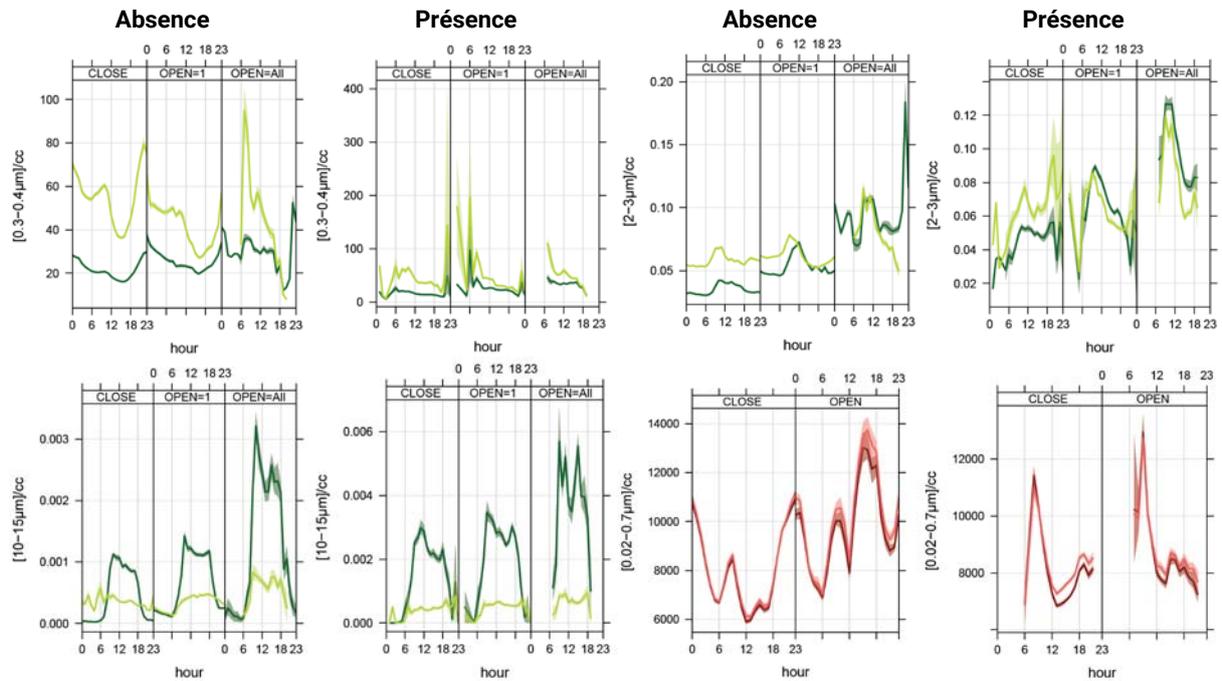


Figure 76 : Evolution du profil moyen journalier des particules selon la présence d'occupants et l'ouverture des fenêtres en 2014 (vert foncé : intérieur et vert clair : extérieur) et pour les particules ultrafines intérieures en 2015 (rouge foncé : position fenêtre et rouge clair : position imprimante).

4.5 Taux de renouvellement d'air

4.5.1 Estimation du taux de renouvellement d'air

Le dioxyde de carbone est un traceur des activités métaboliques (et donc de la présence humaine) et des activités de combustion. Dans l'espace de bureau, aucune source intérieure de combustion n'est présente. Dès lors, la mesure du CO₂ va tracer la présence des occupants. La production métabolique de CO₂ va dépendre du nombre d'occupants, de leur corpulence et de leur état physiologique (activité physique intense, modérée ou sommeil). La concentration de CO₂ qui en résulte va dépendre à la fois de la production métabolique, mais également du renouvellement d'air qui vient moduler les échanges avec le niveau de CO₂ extérieur. De fait, la concentration de CO₂ est un outil par lequel il est possible d'estimer le taux de renouvellement d'air. Et c'est principalement pour cette raison, que la mesure du CO₂ est particulièrement utile.

L'estimation du taux de renouvellement d'air à partir de la concentration de CO₂ dans l'air peut être réalisée à partir de 3 périodes distinctes :

- La phase de production : pour estimer le taux de renouvellement d'air durant cette phase où la concentration de CO₂ augmente, il est nécessaire d'estimer au préalable la production métabolique et donc d'émettre des hypothèses sur le nombre réel d'occupants et la production individuelle de CO₂. Le taux de renouvellement d'air est ensuite estimé en simulant l'augmentation de la concentration de CO₂ avec ces paramètres et en les confrontant avec les données mesurées. Dans l'espace de bureau, le nombre réel d'occupants n'est pas connu. Mais surtout les hypothèses sur l'état physiologique et la masse corporelle de chaque individu qui vont définir leur métabolisme sont difficilement vérifiables.
- La phase stationnaire : durant cette phase, la concentration de CO₂ ne varie plus et il devient alors possible de déterminer très simplement le taux de renouvellement d'air en ayant connaissance de la production métabolique. Bien que plus simple à mettre en œuvre, nous sommes confrontés aux mêmes réserves quant aux hypothèses à prendre en compte sur la production métabolique.
- La phase de décroissance. Ici, l'activité métabolique s'est arrêtée au moment du départ des occupants de la pièce. La concentration de CO₂ va dépendre du taux de renouvellement d'air, de la concentration de CO₂ initiale et du niveau extérieur. Le taux de renouvellement d'air durant cette période peut être facilement estimé sans avoir recours à des hypothèses sur la production métabolique. L'inconvénient est de ne pas pouvoir disposer d'informations durant les périodes d'occupation.

Pour estimer le taux de renouvellement d'air par décroissance, un algorithme spécifique d'exploitation des données, développé au préalable pour les écoles (Ramalho et al., 2013) a été utilisé en modifiant les paramètres. Il permet, à partir des

mesures de CO₂ dans une pièce, de calculer, par un processus automatique, le taux de renouvellement d'air (TRA) pour chaque séquence de décroissance de la concentration en CO₂. Chaque séquence de décroissance est supposée correspondre à une absence de toute occupation dans la pièce. Cette hypothèse n'est pas forcément vérifiée en cas d'ouverture de fenêtres en cours d'occupation. Ce qui pourrait conduire malgré l'occupation de la salle à une diminution de la concentration en CO₂ en fonction du rapport entre le taux de production et le taux de disparition par renouvellement d'air. Dans ce cas, le renouvellement d'air effectif serait sous-estimé.

Dans ces conditions, le taux de renouvellement d'air est proportionnel à la concentration en CO₂ dans l'air, selon la formule suivante :

$$\ln \frac{(C_{ext} - C_{int})}{(C_{ext} - C_{int}^0)} = -\lambda(t - t_0)$$

Avec C_{ext} la concentration extérieure en CO₂ prise comme constante et égale au minimum enregistré par l'appareil pour tenir compte de ses fluctuations autour du niveau de fond,

C_{int} la concentration intérieure en CO₂ en ppm, supposée homogène dans la pièce

C_{int}⁰ la concentration initiale en CO₂ en ppm de la séquence de décroissance,

λ, le taux de renouvellement d'air exprimé en volume/heure et,

t, le temps en heures (t₀ : temps initial de la séquence de décroissance).

La concentration extérieure de CO₂ n'a pas été mesurée dans un premier temps. Ce niveau se situe autour de 400 ppm, mais peut monter jusqu'à 600 ppm pendant de brèves périodes en atmosphère urbaine de fond. La mesure de CO₂ a été mise en place à l'extérieur en janvier 2014. Au préalable, le minimum observé par l'instrument (305 ppm) a été pris par défaut comme valeur représentative de l'extérieur pour prendre en compte les défauts d'ajustement de ces instruments à ces niveaux de fond.

L'hypothèse d'homogénéité de la concentration de CO₂ au sein de la pièce n'est pas vérifiée. Si bien que le taux de renouvellement d'air estimé peut ne pas être représentatif du volume global de l'espace de bureau. Néanmoins, les occupants ne restent pas statiques et leur déplacement au sein de cet espace contribue à l'homogénéité des concentrations dans le volume.

Chaque séquence de décroissance identifiée dans l'enregistrement des mesures déclenche automatiquement le calcul d'un taux de renouvellement d'air par régression linéaire. Un seuil minimum de CO₂ (400 ppm) fixe le nombre d'enregistrements de mesure pris en compte dans la régression. L'algorithme de calcul considère la fin de la phase de décroissance lorsque la concentration de CO₂ atteint ce seuil. Il en est de même également dès lors que la concentration de CO₂ augmente.

On obtient ainsi une série de taux de renouvellement d'air calculés, ainsi que les paramètres permettant de juger de la pertinence du calcul : coefficient de détermination de la régression r^2 , somme des moindres carrés résiduels RMSE, le nombre de points pris en compte dans la régression (durée de l'évènement) et la concentration initiale en CO₂ au moment de la décroissance.

Les critères suivants sont fixés pour sélectionner les TRA valides :

- Coefficient de détermination de la régression $r^2 > 0,9$
- Nombre d'enregistrements de mesure pris en compte dans la régression ≥ 3 (soit 3 min de décroissance minimum, de manière également à essayer de prendre en compte les brusques variations associées à l'ouverture des fenêtres)
- Concentration initiale en CO₂ au moment de la décroissance ≥ 500 ppm (seuil permettant d'éviter de prendre en compte les faibles variations de CO₂ proches de l'incertitude de mesure).

Par ailleurs, l'état de la porte de l'espace de bureau (ouverte ou fermée) n'est pas distingué lors du calcul. Les deux cas peuvent conduire à une décroissance du niveau de CO₂ parfois différente. En effet, si la porte est ouverte, le taux de renouvellement d'air observé tend à représenter un renouvellement d'air global au niveau du bloc bâtementaire car les échanges entre compartiments peuvent être prédominants. En revanche, si la porte est maintenue fermée, le taux de renouvellement d'air calculé représente celui de l'espace de bureau seul.

La méthode de calcul a été implémentée dans une procédure du logiciel statistique SAS (version 9.1.3) permettant d'identifier et d'extraire les séquences de décroissance des niveaux de CO₂. Les incertitudes de la méthode ont été déterminées dans le cadre des écoles. Elles étaient estimées à 22 % dans les cas les plus défavorables.

4.5.2 Variation du taux de renouvellement d'air

Les données de CO₂ sont mesurées depuis janvier 2012 dans l'espace de bureau (avec une interruption de février à mars 2013) et depuis janvier 2014 à l'extérieur. A l'intérieur, la concentration médiane est de l'ordre de 480 ppm avec des pics pouvant atteindre dans de rares occasions 2200 ppm. A l'extérieur, la médiane se situe autour de 415 ppm avec des valeurs maximales pouvant atteindre les 600 ppm. Les fluctuations intérieures sont dépendantes de l'occupation et de l'ouverture des ouvrants. A l'extérieur, la concentration oscille avec un maximum en début de matinée et un minimum en fin d'après-midi.

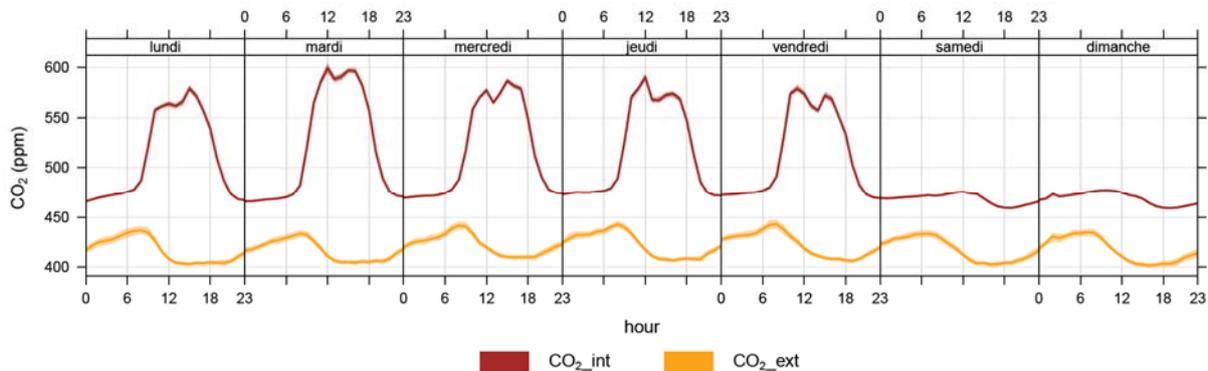


Figure 77 : Variation moyenne hebdomadaire de la concentration de CO₂ à l'intérieur et à l'extérieur.

Entre janvier 2012 et juillet 2015, plus de 100 000 séquences de décroissance de la concentration de CO₂ ont été identifiées. Seules 40 000 séquences ont été jugées valides selon les critères mentionnés préalablement. Pour chacune de ces séquences, une valeur de taux de renouvellement d'air a été déterminée et est associée à une durée comprise entre 3 et 240 minutes. La durée cumulée pour laquelle le taux de renouvellement d'air est renseignée représente environ 150 jours soit environ 12 % du temps. Il reste donc une grande partie du temps pour laquelle le taux de renouvellement d'air dans l'espace de bureau n'est pas connu.

L'hypothèse d'absence d'occupation durant la décroissance peut ne pas s'avérer exacte durant les heures habituelles de présence des occupants. Les données de présence recueillies par ailleurs permettraient de valider ou d'invalidier les séquences observées durant cette période. Ce n'est pas chose aisée, car la présence éventuelle dans les locaux voisins peut également mettre à mal cette hypothèse. Néanmoins, la présence d'occupants durant une décroissance de la concentration de CO₂ impliquerait un taux de renouvellement d'air plus important que celui qui serait calculé. En d'autres termes, les valeurs calculées par la méthode automatique sont sous-estimées durant les heures habituelles de présence.

Les taux de renouvellement d'air (TRA) obtenus se situent autour de 1,2 vol/h en moyenne horaire avec un coefficient de variation de 80 %. En moyenne journalière, cette variabilité est de 47 % et tombe à 26 % lorsque l'on compare des moyennes mensuelles. Ce taux de renouvellement d'air moyen est supérieur au renouvellement d'air associé à la seule extraction qui est de l'ordre de 0,6 vol/h. La perméabilité de la pièce et l'ouverture des fenêtres tendent à apporter un renouvellement d'air complémentaire par rapport à la ventilation mécanique par extraction.

La Figure 40 montre l'évolution du taux de renouvellement d'air journalier de janvier 2012 à mars 2014 (sélection des jours avec un minimum de 30 minutes de données). Au cours d'une même journée, l'amplitude des variations peut être supérieure à 5. Le taux de renouvellement d'air annuel moyen a baissé de manière significative de 1,4 en 2012 à 1,1 en 2013, 1,2 en 2014 et 0,8 au premier semestre 2015. Une variation est

également observée au cours de l'année, avec un mois de juin qui présente le niveau moyen le plus élevé à 1,7 vol/h et les mois de février, août et octobre associés aux niveaux les plus faibles autour de 1 vol/h en moyenne. La variation journalière est encore plus significative. Elle montre que le TRA entre 20h00 et 07h00 est stable autour de 0,4 vol/h avec peu de variation (40 %). Par contre entre 08h00 et 20h00, le TRA moyen horaire augmente de façon significative jusqu'à 1,5 vol/h avec d'importantes variations au sein d'une même heure (jusqu'à 156 %). Ce qui tend à montrer que durant les heures d'occupation, le TRA est plus important et que par conséquent les usages des occupants jouent un rôle prépondérant dans le renouvellement d'air du local.

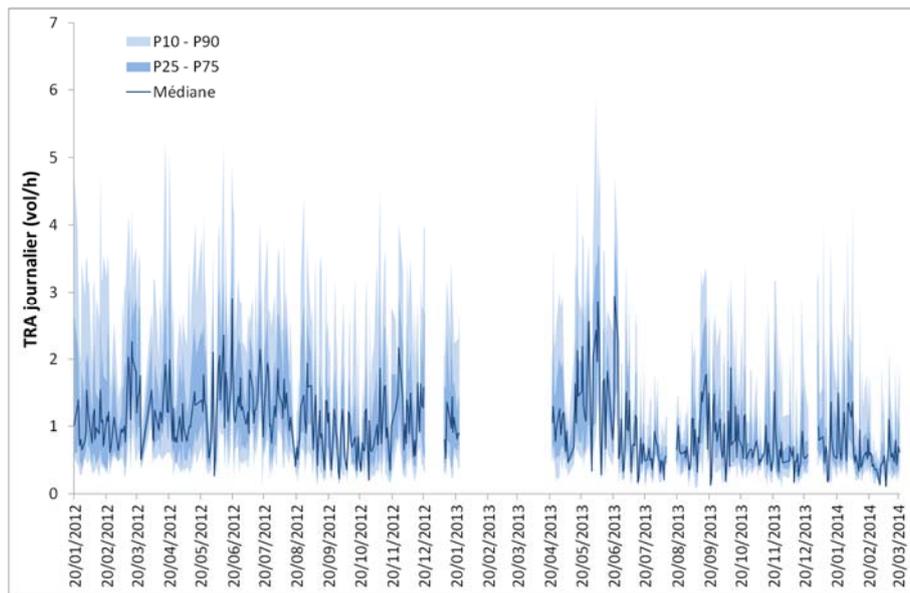


Figure 78 : Variation du taux de renouvellement d'air (TRA) journalier médian (hors occupation) de 2012 au premier trimestre 2014.

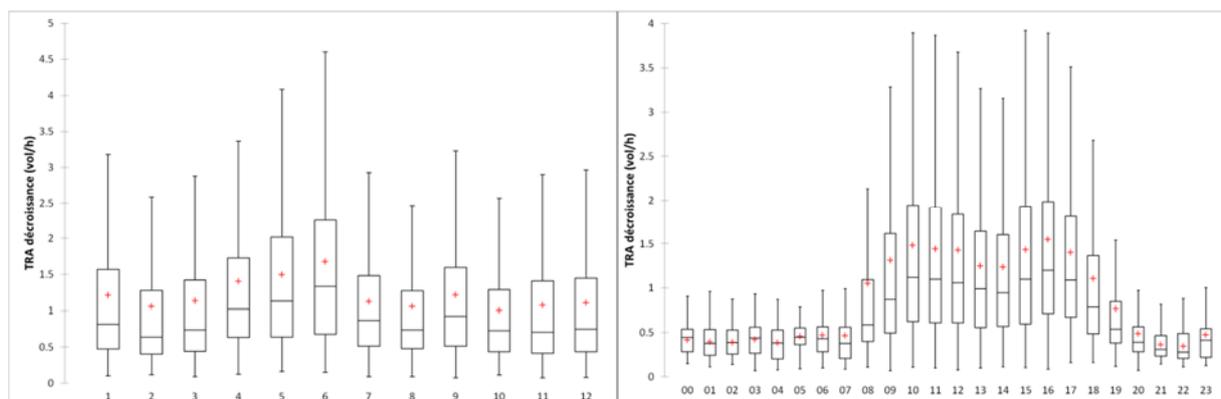


Figure 79 : Variation du taux de renouvellement d'air (TRA) journalier médian (hors occupation) de janvier 2012 à mars 2014.

La connaissance de la variation du taux de renouvellement d'air dans le temps est une donnée très importante qui affecte directement la variation de la concentration des polluants dans l'air intérieur. Pouvoir estimer les fluctuations de cette variable d'échange entre extérieur et intérieur, c'est permettre d'améliorer sensiblement les modèles de prévision de l'évolution des concentrations et d'estimation de l'exposition des occupants. La méthode utilisée pour estimer le taux de renouvellement d'air est loin d'être parfaite, mais à notre connaissance elle est la seule à nous permettre d'avoir une idée de sa variabilité temporelle en tenant compte des usages des occupants (bien qu'il reste encore à développer une estimation du taux de renouvellement durant l'occupation en minimisant les risques d'erreurs sur les hypothèses).

Il reste encore à croiser les données de TRA avec les ouvertures de fenêtres pour identifier la relation et tenter de quantifier l'effet. Pour le croisement avec les autres données, cela sera sans doute plus difficile tant les données de TRA restent parcellaires, à moins de travailler sur des moyennes journalières. Mais ces journées ne représentent que 62 % de la période complète.

4.5.3 Influence des ouvrants sur le taux de renouvellement d'air

L'influence des ouvrants sur le taux de renouvellement d'air a été évaluée sur l'année 2014. Chaque valeur de TRA calculée est associée à la date exacte du début de décroissance. Cette date est la clé qui permet de croiser la variable TRA avec les autres variables d'intérêt comme l'ouverture des ouvrants. En réalité, le TRA calculé est associé à une durée (celle de la séquence) et non à un instant précis. Il serait nécessaire de reconstruire cette variable en tenant compte de cet aspect. Mais, dans un premier temps, nous l'avons considéré comme associé à un instant précis pour faciliter nos croisements. Ce qui donne plus de poids aux valeurs élevées de TRA souvent associées à des durées plus courtes que les valeurs plus faibles.

L'influence de l'état des fenêtres (fermées / au moins une fenêtre ouverte) et celle de la porte d'accès à l'espace de bureau sur les niveaux de TRA ont été évaluées. Les deux ouvrants ont une influence significative (test de Mann-Whitney, $p < 0.0001$) sur le niveau de TRA. La porte d'accès ouverte conduit globalement à des niveaux de TRA plus élevées (médiane de $1,05 \text{ h}^{-1}$ par rapport à $0,85 \text{ h}^{-1}$ porte fermée). L'effet est plus important lorsqu'au moins une des fenêtres est ouverte (médiane de $1,25 \text{ h}^{-1}$ par rapport à $0,80 \text{ h}^{-1}$). Il est nécessaire de rappeler que le TRA estimé représente avant tout un renouvellement d'air par dilution et pas forcément un renouvellement d'air neuf provenant exclusivement de l'extérieur.

La porte d'accès donne depuis l'espace de bureaux vers une zone de passage qui communique par trois portes différentes avec un grand hall, l'extérieur et une halle technique. Une porte d'accès ouverte a par conséquent de fortes chances de contribuer à l'augmentation du renouvellement d'air.

L'influence de la porte d'accès n'est pas aussi prononcée lorsque l'on observe son effet sur le profil moyen journalier du TRA. L'influence des fenêtres se fait sentir principalement durant les heures d'occupation des locaux. Leur effet est globalement moins marqué la nuit.

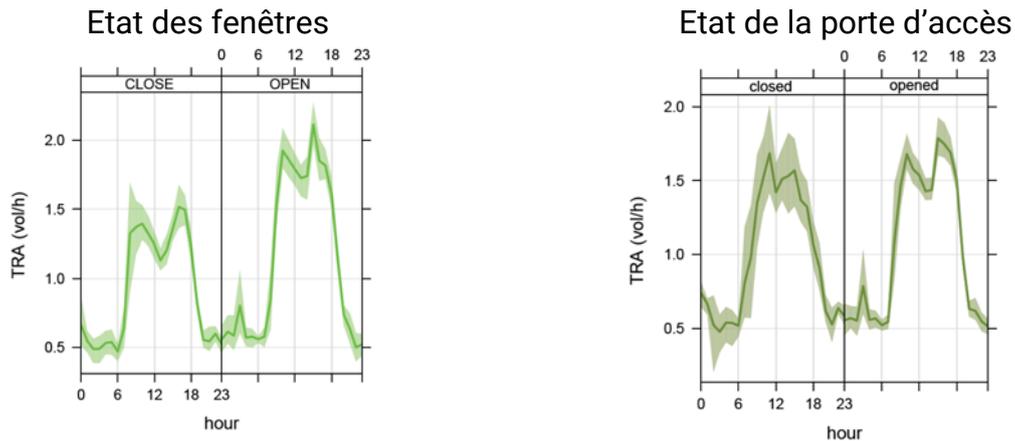


Figure 80 : Profil moyen journalier du TRA selon l'état d'ouverture des fenêtres ou de la porte

En regardant de manière conjointe l'influence de l'ouverture de la porte et des fenêtres, on se rend compte que l'effet de l'ouverture de la porte disparaît lorsque les fenêtres sont ouvertes. Il n'est significatif qu'à partir du moment où les fenêtres sont fermées (test MW, $p < 0,0001$). L'influence de l'ouverture des fenêtres est par contre systématique quel que soit l'état de la porte.

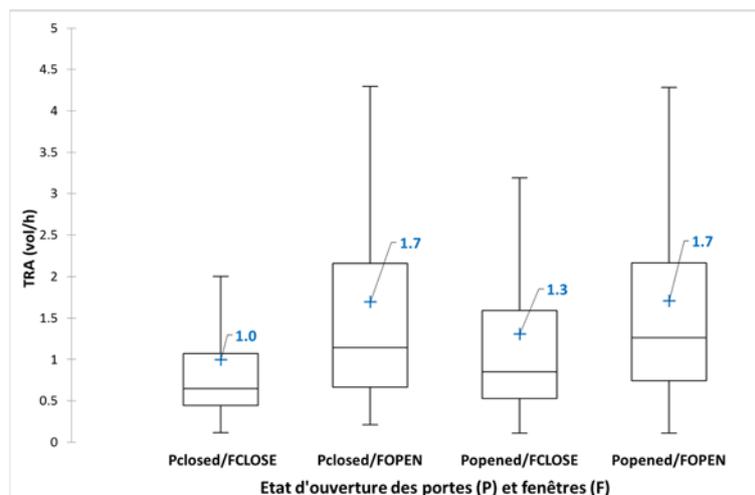


Figure 81 : Influence conjointe de la porte et des fenêtres sur le taux de renouvellement d'air (valeurs extrêmes non représentées)

Plus le nombre de fenêtres ouvertes augmente, plus le TRA a tendance à augmenter. Cette augmentation est significative à toute ouverture d'une nouvelle fenêtre à l'exception du passage de 2 à 3 fenêtres ouvertes et de 4 à 5 fenêtres ouvertes.

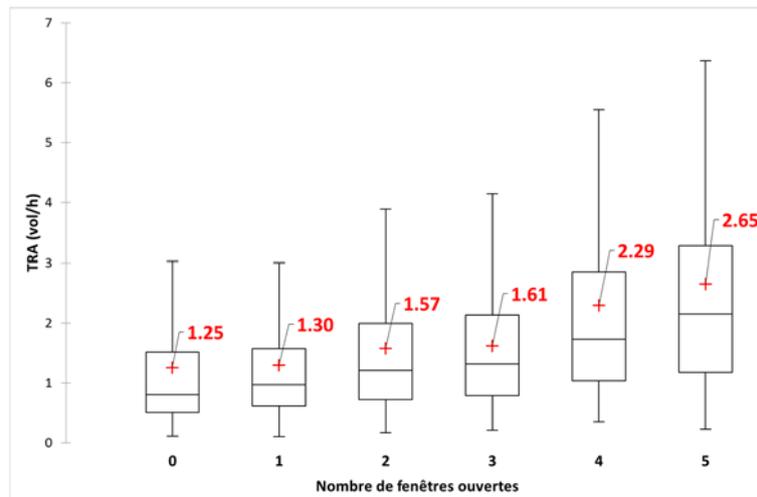


Figure 82 : Variation du TRA en fonction de nombre de fenêtres ouvertes en 2014 (valeurs extrêmes non représentées)

L'estimation du TRA parvient à montrer des différences qui semblent cohérentes dans l'influence des ouvrants et ce malgré les nombreuses hypothèses qui l'entoure. La difficulté réside surtout dans son utilisation avec les autres données tant les informations disponibles qui le concernent restent parcellaires.

4.6 Dépôt des particules

Au travers de la connaissance du taux de renouvellement d'air sur certaines périodes, il est possible de déterminer d'autres caractéristiques, notamment celles associées au dépôt des particules in situ. Des premiers travaux ont été engagés dans ce sens sur la base des données du premier semestre 2014. Ils ne sont pas jugés satisfaisants. Ils sont présentés ici à titre illustratif mais ils nécessitent d'être revus en termes de méthodes employées et de l'ensemble des données disponibles.

4.6.1 Estimation de la constante de dépôt des particules

Le dépôt des particules sur les surfaces peut être modélisé de différentes façons. Dans Halios et Helmis (2007), il est modélisé par un paramètre supposé constant, la vitesse de dépôt, qui peut être déterminé expérimentalement, sous des conditions contrôlées. La vitesse de dépôt v_d intervient dans le bilan de masse, qui peut être simplifié si on fait l'hypothèse qu'il n'y a pas de source de particules à l'intérieur, qu'il n'y a pas de processus physique ou chimique générateur ou consommateur de particules, et que le coefficient de pénétration (c'est-à-dire l'efficacité de transfert) des particules vaut 1.

$$\frac{dN_i}{dt} = \lambda_v(N_e - N_i) - \frac{v_d A N_i}{V}$$

Où N_i et N_e sont respectivement les concentrations en nombre de particules de la gamme de taille considérée, λ_v le taux de renouvellement d'air constant sur la période en h^{-1} , A la surface disponible de l'espace de bureau pour le dépôt des particules (m^2) et V son volume en m^3 .

Sur les périodes où on connaît le taux de renouvellement d'air λ_v , on peut calculer le ratio $\frac{v_d A N_i}{V}$ en approximant la dérivée de la concentration intérieure, selon l'équation suivante.

$$\frac{v_d A}{V} N_i(t) = \lambda_v(N_e(t) - N_i(t)) - \frac{N_i(t + \Delta t) - N_i(t)}{\Delta t}$$

Où le facteur $\frac{v_d A}{V} = \lambda_d$ est la constante de dépôt (h^{-1}). La constante de dépôt moyenne est calculée avec cette expression pour les 50 journées où le taux de renouvellement d'air a pu être déterminé durant le premier semestre 2014 et en ne tenant compte que des données disponibles après la présence des occupants pour vérifier l'hypothèse d'absence de sources intérieures. Les résultats indiquent de nombreuses valeurs négatives. C'est essentiellement dû au fait que la constante de dépôt a été calculé

toutes les 10 minutes et que la variabilité conjuguée des concentrations intérieures et extérieures augmente le risque d'avoir une différence $N_e - N_i$ négative. La méthode est donc à revoir de ce point de vue en intégrant les concentrations sur des périodes plus longues. Par ailleurs, les plus grosses particules remises en suspension par la présence des occupants sont encore très élevées en début de soirée et très supérieures en nombre à leur niveau à l'extérieur.

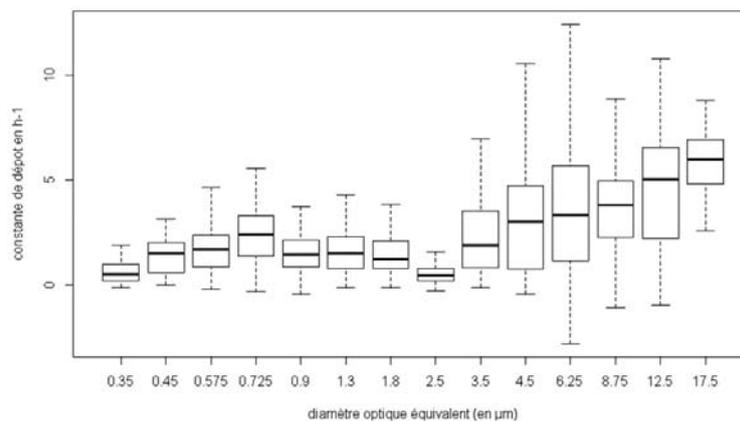


Figure 83 : Estimation de la constante de dépôt sur le premier semestre 2014

La vitesse de dépôt v_d peut être estimée à partir de ces données. Pour le local étudié, A vaut 572 m^2 (toutes surfaces verticales et horizontales). La variabilité des valeurs de vitesse de dépôt est très élevée. Elle est en grande partie liée à la méthode de calcul, mais elle peut souligner des variations réelles liées à différentes configurations thermo-aérauliques notamment et des températures de surfaces des parois variables au cours de l'année (en particulier au niveau des vitrages).

Tout comme le taux de renouvellement d'air, il est difficile de considérer que la vitesse de dépôt est constante. Si on souhaite utiliser la vitesse de dépôt pour modéliser l'interaction des particules avec les parois, il serait plus juste d'utiliser un paramètre variable dans le temps. Il existe d'autres façons de modéliser l'interaction des particules avec les parois. Un modèle semi-empirique est défini par Lai et Nazaroff (2000). Ce modèle de dépôt en atmosphère intérieure sur les surfaces lisses prend en compte les turbulences au voisinage des surfaces, et simule la diffusion brownienne, la diffusion turbulente et l'attraction gravitationnelle. Ainsi le dépôt des particules est modélisé en fonction de la taille et de la masse volumique des particules. Ce modèle est utilisé par Gao et Niu (2007), couplé à une simulation de mécanique des fluides CFD, pour estimer le risque de contamination par un bioaérosol ou un agent chimique particulaire dans un local ventilé.

Tableau 6 : Estimation de la vitesse de dépôt moyenne des particules

Diamètre (µm)	0.35	0.45	0.575	0.725	0.9	1.3	1.8	2.5	3.5	4.5	6.25	8.75	12.5	17.5
Moyenne (m/h)	0.49	1.11	1.17	1.53	1.00	1.05	1.00	0.41	1.50	2.20	2.48	2.72	3.29	3.84
Ecart type (m/h)	0.56	1.24	1.07	1.13	0.70	0.73	0.68	0.44	1.06	1.88	2.25	2.19	2.55	4.35
CV (%)	115	111	91	74	70	70	68	109	70	86	91	81	77	113

4.6.2 Estimation de la constante de disparition des particules

La constante de disparition globale λ des particules peut être estimée en additionnant le taux de renouvellement d'air et la constante de dépôt calculé précédemment sur les mêmes périodes.

Chen et Zhao (2011) propose une autre façon de calculer λ . En régime stationnaire, le ratio $\frac{N_i}{N_e}$ peut s'écrire :

$$\frac{N_i}{N_e} = \frac{\lambda_v * P_0}{\lambda} + \frac{S}{\lambda * V * N_e}$$

Où P_0 est le facteur de pénétration, supposé ici égal à 1. En réalité, ce facteur de pénétration varie avec la taille des particules (Chen et Zhao, 2011) et avec l'ouverture des fenêtres. Ce n'est donc pas une constante également. La méthode de calcul tout comme la stratégie de mesure ne permet pas de distinguer le facteur de pénétration dans le calcul de la constante de disparition. La variation du premier se répercute sur le calcul du second.

Sur les périodes d'absence des occupants, on peut calculer λ en considérant les concentrations en particules moyennes mesurées sur chaque plage de données où le taux de renouvellement d'air est connu. Le Tableau suivant présente les valeurs numériques de λ calculées par les deux méthodes décrites précédemment.

Les deux méthodes fournissent des résultats très proches jusqu'à 2,5 µm. Pour les diamètres plus élevés, l'écart entre les deux moyennes augmente (0,2 à 3,5 h⁻¹ de différence). Dans la deuxième méthode, l'hypothèse d'un régime stationnaire n'est pas forcément vérifiée pour les particules de taille supérieure à 2,5 µm, ce qui peut expliquer l'écart avec la première méthode. Par ailleurs, l'écart-type sur les valeurs calculées par les deux méthodes diffère seulement pour les particules supérieures à 8,75 µm.

Tableau 7 : Estimation de la constante de disparition des particules selon deux méthodes de calcul

Diamètre (µm)	0.35	0.45	0.575	0.725	0.9	1.3	1.8	2.5	3.5	4.5	6.25	8.75	12.5	17.5
	$\lambda_{TRA} + \lambda_d$													
Moyenne (h⁻¹)	1.18	2.17	2.27	2.84	1.99	2.04	1.95	0.98	2.70	3.79	4.18	4.48	5.29	5.82
Ecart type (h⁻¹)	1.03	2.09	1.83	1.93	1.24	1.28	1.18	0.63	1.76	3.08	3.60	3.29	3.52	5.79
	ratio Ni/Ne													
Moyenne (h⁻¹)	1.21	2.22	2.33	2.89	2.04	2.04	1.91	0.90	2.51	3.60	3.93	3.58	3.99	9.33
Ecart type (h⁻¹)	1.07	2.11	1.83	1.91	1.21	1.24	1.18	0.65	1.84	3.07	3.57	3.11	3.73	11.88

Comme indiqué au début de ce chapitre, les méthodes et hypothèses utilisées pour déterminer ces premiers résultats ne sont pas complètement satisfaisantes. Les résultats présentés restent indicatifs. Ils doivent être confirmés par un travail d'exploitation plus prononcé et étendus sur l'ensemble des données.

4.7 Contribution des sources

4.7.1 Influence de l'imprimante multifonction sur les concentrations de particules

Le compteur d'impulsions électriques a permis de mettre en évidence que l'imprimante multifonction n'est active (en cours d'impression) que moins de 10 % du temps dans l'espace de bureau en 2013. Le reste du temps, l'imprimante est en état de veille soit active soit en veille prolongée (mode économie d'énergie). L'imprimante précédente ne disposait pas d'un mode de veille prolongée. Elle a été changée au début avril 2013. En 2014 et 2015, l'imprimante peut être active durant près de 36 % de son temps de fonctionnement certains mois.

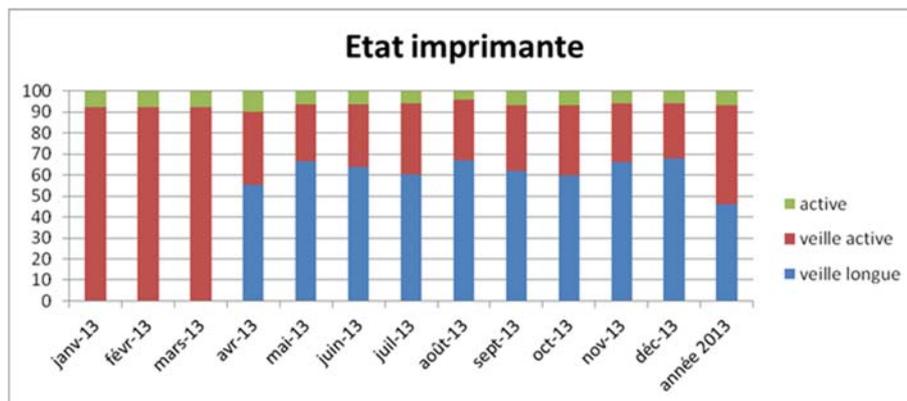
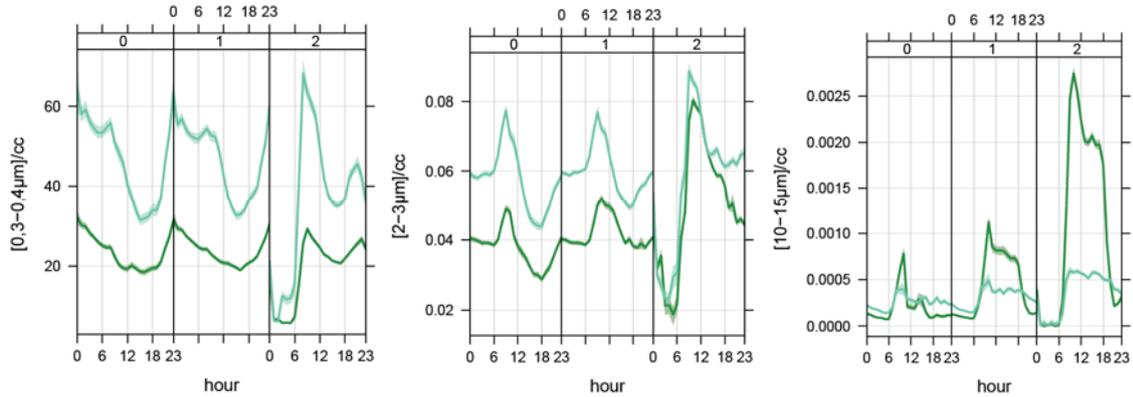


Figure 84 : Variation de l'état de l'imprimante multifonction sur l'année 2013 (changement d'imprimante entre mars et avril 2013).

L'effet du fonctionnement de l'imprimante a été évalué sur la dynamique de concentration des particules en 2014 en considérant trois gammes de taille 0,35 µm, 2,5 µm et 12,5 µm. Les résultats présentés sur la figure ci-dessous montrent l'évolution de la concentration de ces 3 tailles de particules selon l'état de fonctionnement de l'imprimante : en veille prolongée (0), en veille active (1) et active (2).

En considérant l'ensemble des périodes d'occupation et d'inoccupation, l'activité de l'imprimante affecte surtout la dynamique des plus grosses particules. Mais, les périodes où l'imprimante est active coïncident avec les périodes de présence des occupants. Il est donc nécessaire de ne considérer l'influence de l'imprimante que durant ces périodes. Dès lors, l'influence de l'imprimante est beaucoup moins marquée mais reste significative pour l'ensemble des tailles de particules, en particulier entre le mode veille prolongée et les autres états. Cette influence se confirme plus encore au regard des ratios I/E des concentrations moyennes déterminées sur ces mêmes périodes notamment pour les plus grosses particules.

En période d'occupation et d'inoccupation



En période d'occupation exclusivement

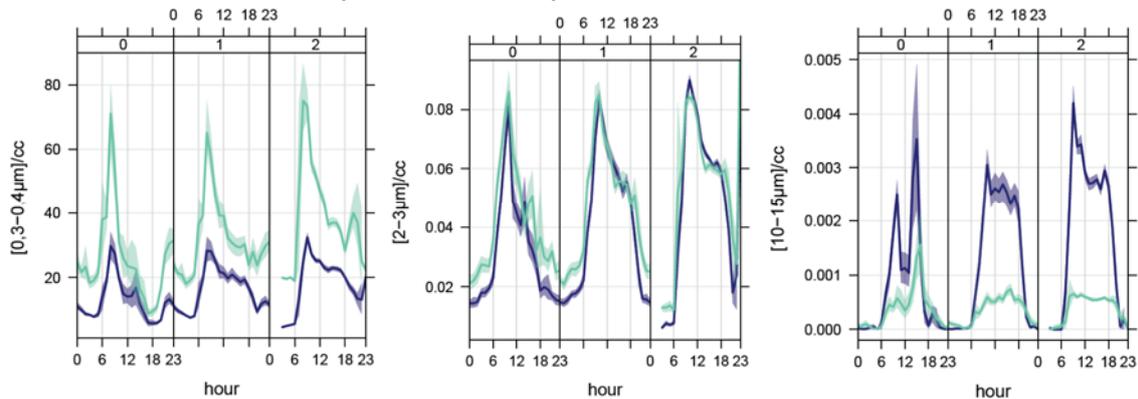


Figure 85 : Evolution journalière moyenne de la concentration de particules selon l'état de fonctionnement de l'imprimante (année 2014).

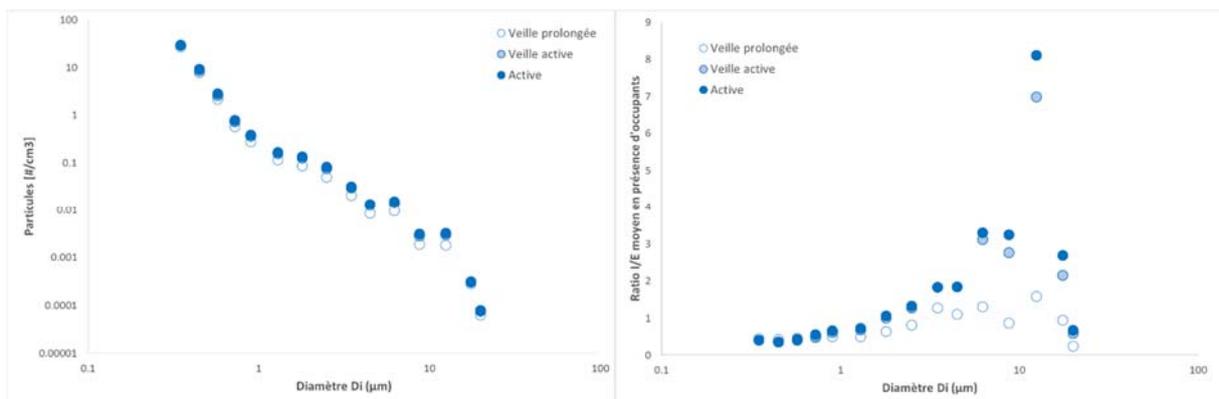


Figure 86 : Concentration moyenne de particules et ratio I/E selon leur diamètre et l'état de l'imprimante durant les périodes de présence des occupants (année 2015).

L'influence sur les particules ultrafines est différente. En considérant toutes les périodes d'occupation et d'inoccupation, aucun effet n'est perceptible. Durant les

seules périodes de présence, les différences observées entre la veille prolongée et les autres états sont significatives. Elles indiquent toutefois que les concentrations de particules ultrafines sont plus élevées durant ces phases de veille prolongée que durant le fonctionnement de l'imprimante. Le positionnement différent des deux instruments ne montre pas de différences significatives.

Le fonctionnement de l'imprimante (mise en chauffe, vibrations) pourrait émettre ou conduire à des modifications thermoaérauliques qui favoriseraient la remise en suspension des plus grosses particules mais également le dépôt des particules ultrafines. Cela ne reste que des hypothèses que cette étude ne permet pas de confirmer.

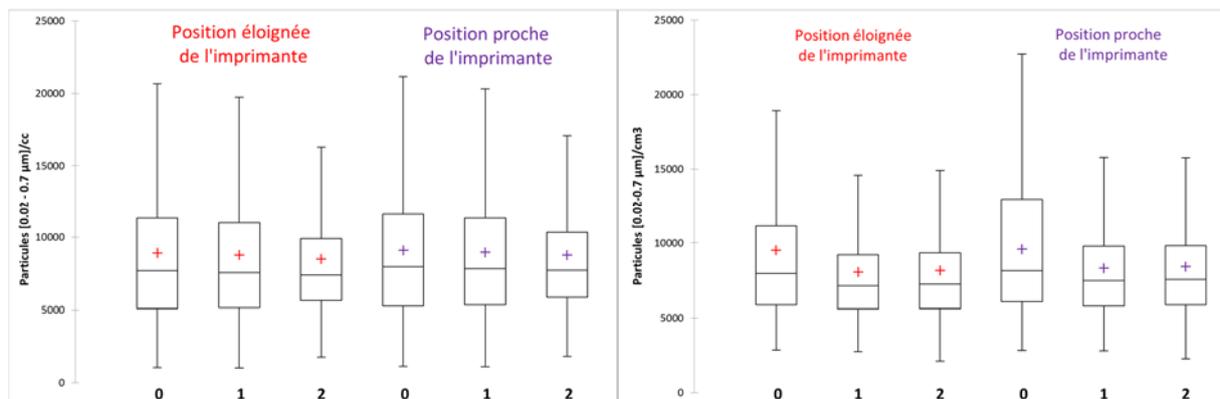


Figure 87 : Influence du fonctionnement de l'imprimante sur les niveaux de particules ultrafines toutes périodes confondues (gauche) et en présence d'occupants uniquement (droite)

Le fonctionnement de l'imprimante n'a par contre aucune influence sur les concentrations de formaldéhyde. Une légère augmentation significative de la concentration intérieure en CO est observée lorsque l'imprimante est active (0,2 ppm par rapport à 0,1 ppm en moyenne, $p = 0,0002$). A l'inverse, l'imprimante en phase active diminue en moyenne la concentration intérieure d'ozone de 13 ppb en veille prolongée à 9 ppb en fonctionnement actif ($p < 0,0001$). Cette diminution pourrait aussi bien être attribuée à la modification des conditions thermo-aérauliques locales qu'à l'émission de composés organiques volatils réagissant avec l'ozone. Le fonctionnement de l'imprimante contribuerait alors à la réactivité chimique globale de l'espace de bureau.

4.7.2 Influence de la vitesse et de la direction du vent sur les concentrations intérieures de polluants

Un des aspects assez peu étudié de l'influence des conditions climatiques sur la concentration intérieures des polluants est lié au vent. Sa vitesse et sa direction modifient les concentrations extérieures, mais le vent peut intervenir sur d'autres aspects du bâti qui peuvent in fine influencer les niveaux intérieurs des paramètres de la qualité de l'air. Associer cette variable aux interactions entre différents paramètres intérieurs permet d'apporter un autre regard sur les relations observées.

La fonction Polarplot du package Openair permet d'observer ces relations en fonction de la direction du vent. La surface représentée couvre l'ensemble des valeurs d'une variable étudiée (par exemple la vitesse du vent) selon la direction du vent associée. Les valeurs prises par la variable d'intérêt sont ensuite représentées sous forme d'une échelle de couleur sur la surface définie par les deux variables précédentes.

La Figure 88 montre que la concentration intérieure de formaldéhyde selon la présence des occupants et l'état d'ouverture des fenêtres n'est pas associée aux mêmes conditions de vent. En période d'inoccupation, fenêtres fermées (0), les concentrations de formaldéhyde les plus élevées sont associées à des vents de secteur ouest / sud-ouest. Toutes fenêtres ouvertes (1), les concentrations les plus élevées sont plutôt associées à des vents de secteur nord-est. En comparaison, la concentration d'ozone est plus élevée toutes fenêtres ouvertes et plutôt associée à des vents provenant de l'est sur la période d'avril à juillet 2013.

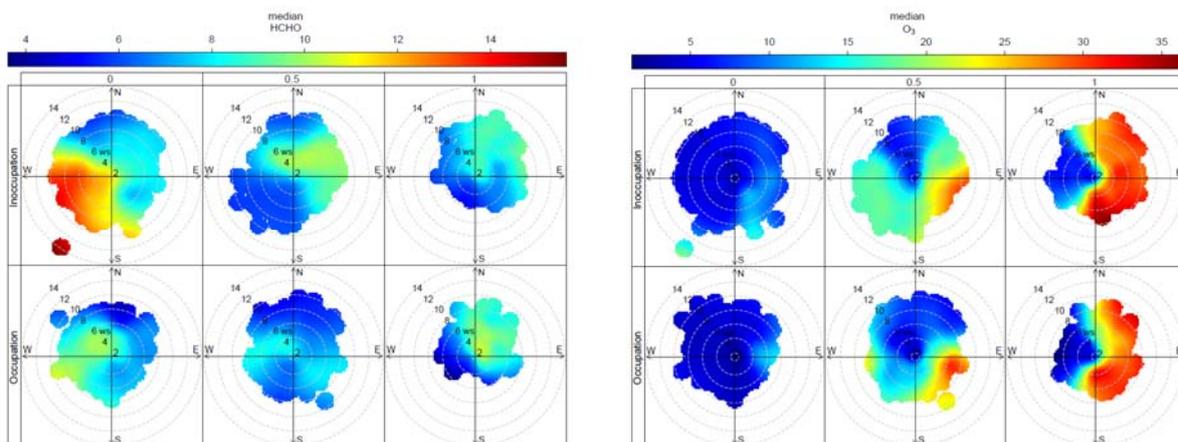


Figure 88 : Distribution polaire de la concentration en formaldéhyde et ozone selon la vitesse et la direction du vent et en fonction de la présence des occupants et de l'ouverture des fenêtres (période d'avril à juillet 2013)

Ce même exercice a été réalisé en remplaçant la vitesse du vent par la température ou l'humidité spécifique intérieure. La Figure 89 montre que la relation entre le formaldéhyde et la température intérieure n'est pas associée à une direction de vent

particulière que ce soit en période d'occupation ou d'inoccupation. Cette relation semble plus intense lorsque les fenêtres sont fermées.

La relation entre l'humidité spécifique et la concentration en formaldéhyde est également plus marquée lorsque les fenêtres sont fermées, en particulier durant les périodes d'inoccupation. Elle ne semble pas dépendre de la direction du vent, bien que les concentrations de formaldéhyde soient plus faibles lorsque le vent est de secteur nord-est.

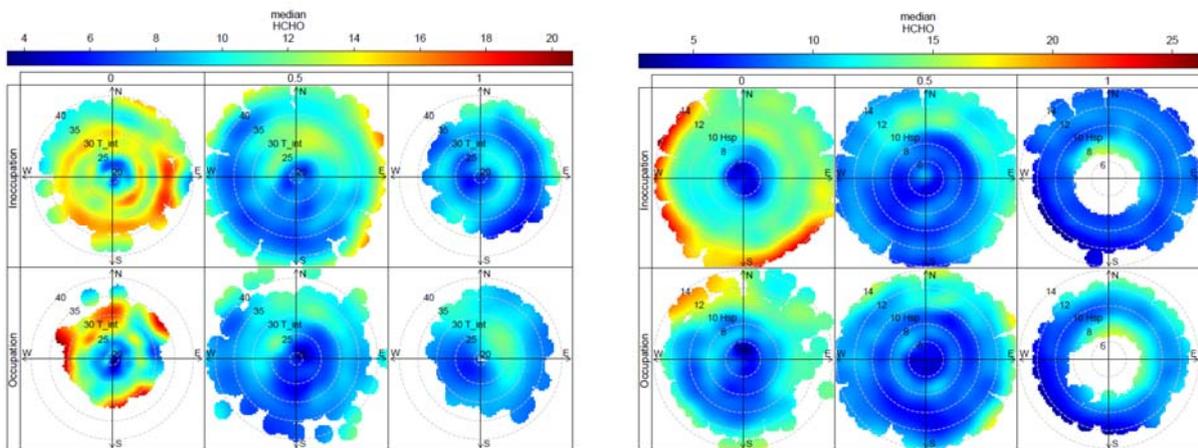


Figure 89 : Distribution polaire de la concentration en formaldéhyde selon la direction du vent et la température ou l'humidité spécifique intérieure en fonction de la présence des occupants et de l'ouverture des fenêtres (avril à juillet 2013).

Un exercice similaire a été réalisé sur les concentrations de particules toutes périodes confondues en les associant à une vitesse et une direction de vent. La figure suivante illustre les résultats obtenus pour 3 tailles de particules. Pour les particules fines et de taille intermédiaire, les concentrations les plus élevées coïncident avec celles observées à l'extérieur à la fois en vitesse et en direction du vent. Par contre, selon la gamme de taille des particules considérées, la vitesse et la direction du vent correspondant aux valeurs les plus élevées n'est pas la même. Pour les particules les plus fines, il s'agit de vents forts provenant de l'est ou des vents modérés de secteur nord. Tandis que des vents modérés de secteur nord-ouest caractérise les concentrations les plus élevées de particules de 2,5 μm . Enfin, les plus grosses particules de 12,5 μm sont plutôt associées à des vents modérés de secteur nord-est.

Il reste encore de nombreuses possibilités d'associations des concentrations des polluants cibles que sont le formaldéhyde et les particules au regard des autres paramètres (concentrations intérieures d'ozone et d'oxydes d'azote notamment, concentrations extérieures issues de la station de surveillance de Lognes) et des paramètres climatiques plus spécifiquement que nous n'avons pas eu le temps d'explorer dans le cadre de ce projet. Tous ces paramètres sont liés et il est alors

difficile de faire la part des différentes contributions de chacun. L'utilisation de méthodes de déconvolution ou de séparation des sources de fluctuations pourrait permettre de démêler ces observations environnementales et d'en extraire les caractéristiques particulières.

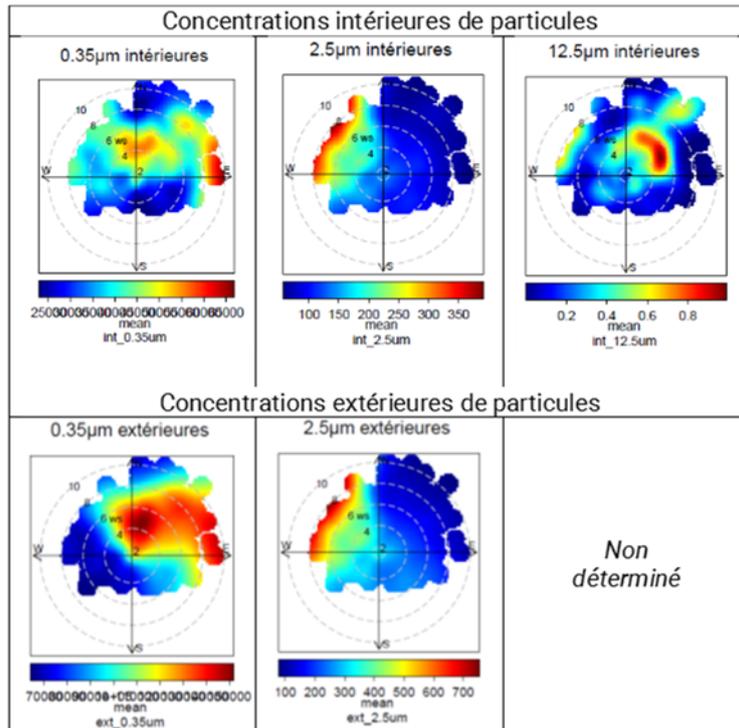


Figure 90 : Distribution polaire de la concentration en particules par litre selon la vitesse et la direction du vent en 2013.

4.7.3 Séparation des sources de fluctuations

Différentes techniques de séparation aveugle des sources ont été utilisées sur les concentrations de particules mesurées dans l'espace de bureau de 2012 à 2015, mais également dans la maison expérimentale sans occupation en 2010 ou encore dans un bureau individuel en 2011.

Le principe général de ces méthodes (Figure 91) consiste à extraire les vecteurs sources à partir d'une matrice d'observations en isolant le bruit du signal et la matrice de mélange. La manière dont cette extraction est réalisée diffère selon les méthodes. Ces méthodes sont détaillées en annexes.

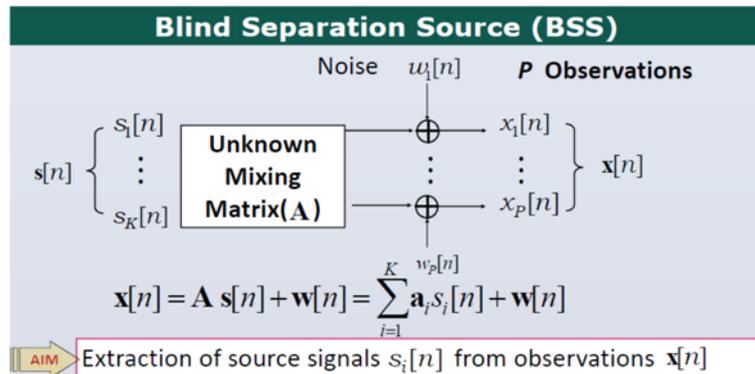


Figure 91 : Principe général des méthodes d'extraction des sources

4.7.3.1 Séparation des sources de fluctuations dans un bureau individuel

La caractérisation des sources de particules mesurées dans un bureau individuel a été entreprise. Plus précisément, il s'agit de caractériser la structure des profils temporels des sources de particules. Pour cela, nous avons d'abord entrepris une analyse statistique des concentrations en nombre de particules dans un environnement de type bureau, étape préalable au développement de modèles permettant de quantifier la contribution des sources. Ensuite, nous avons fait appel aux méthodes appliquées en statistique exploratoire des données et en traitement du signal pour la séparation des sources, notamment l'analyse en composantes principales (ACP) et l'analyse en composantes indépendantes (ACI) avec la méthode FastICA. Cette étude a fait l'objet d'une communication au 29^{ème} Congrès Français sur les Aérosols en janvier 2014 à Paris (Ouaret et al., 2014a).

Les données prises en compte dans cette étude concernent 45 jours consécutifs, du 19 février 2011 à minuit jusqu'au 4 avril 2011 à 23h50. Les résultats de la statistique descriptive montrent qu'en termes de variabilité, plus les particules sont fines, plus leur écart-type est grand et plus elles sont corrélées entre elles. En effet, les coefficients de corrélation de Pearson pour les particules entre $0,3 \mu\text{m}$ et $0,80 \mu\text{m}$ varient entre 0,75 et 0,98. Les coefficients de corrélation entre les particules $> 20 \mu\text{m}$ et $< 4,5 \mu\text{m}$ ne dépassent pas 0,35. Les coefficients d'asymétrie des paramètres mesurés sont positifs, leurs queues de distribution sont étalées vers la droite ; ceci laisse présager des distributions log-normales.

Outre la mesure d'aplatissement et d'asymétrie, les distributions de probabilités ont été estimées par une méthode non paramétrique (à noyau gaussien). Une distribution log-normale a été également ajustée aux données. De façon générale, on remarque que pour les particules fines, les distributions en nombre sont caractérisées par une densité de probabilité de type log-normale ; ce n'est pas le cas pour les grosses particules, en particulier du fait que leur nombre est très souvent voisin de zéro.

L'approche statistique développée ici consiste à exploiter l'ensemble des données mesurées et à construire un jeu de variables permettant d'identifier les éventuelles structures temporelles des sources.

En utilisant l'analyse de la fonction d'autocorrélation et l'estimation de l'exposant de Hurst, nous étudions les propriétés de la structure temporelle des sources de particules. La Figure 92 présente la méthodologie générale suivie en vue de caractériser la typologie des émissions de source. Dans le but de comprendre la structure de ces fluctuations, on s'intéresse, d'une part aux propriétés intrinsèques des séries (la forme de la fonction d'autocorrélation), et d'autre part aux spécificités de la variabilité, en faisant une classification pour trouver des journées-types.

L'analyse des profils diurnes des séries temporelles dans un bureau individuel met en évidence la présence des niveaux de concentration élevés durant les heures de travail, notamment, pour les grosses particules et le CO₂, tandis que pour les particules comprises entre 0,30 et 0,5 µm de diamètre, les profils diurnes exhibent une faible variabilité.

Les caractéristiques de la variabilité des phénomènes qui ont généré les séries temporelles ne sont pas forcément connues, la seule information dont nous disposons est l'historique des séries mesurées. Justement, la fonction d'autocorrélation (ACF) nous permet d'explorer en partie les structures statistiques des séries.

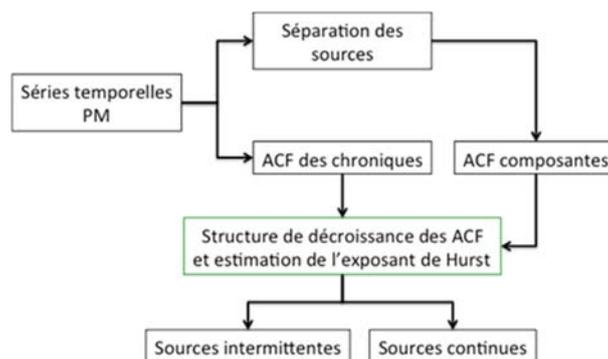


Figure 92 : Caractérisation de la structure de profil des sources (ACF : fonction d'autocorrélation).

La Figure 93 présente les courbes de la fonction d'autocorrélation (ACF) (le corrélogramme) calculées à partir des séries temporelles de la concentration en nombre de particules et du niveau du CO₂. On remarque que les valeurs de l'ACF de la série des concentrations de CO₂ sont significativement non nulles jusqu'à ce que le corrélogramme soit tronqué à partir de 8h (480 minutes). Le même type de comportement est observé pour les fractions de taille supérieure à 1,6 µm. La plupart des fonctions d'autocorrélation se comportent comme un mélange de fonctions exponentielles/sinusoidales amorties, en particulier pour les particules de taille

moyenne et le CO₂. En effet, l'ACF de ces fractions traduit l'aspect de saisonnalité de leur concentration.

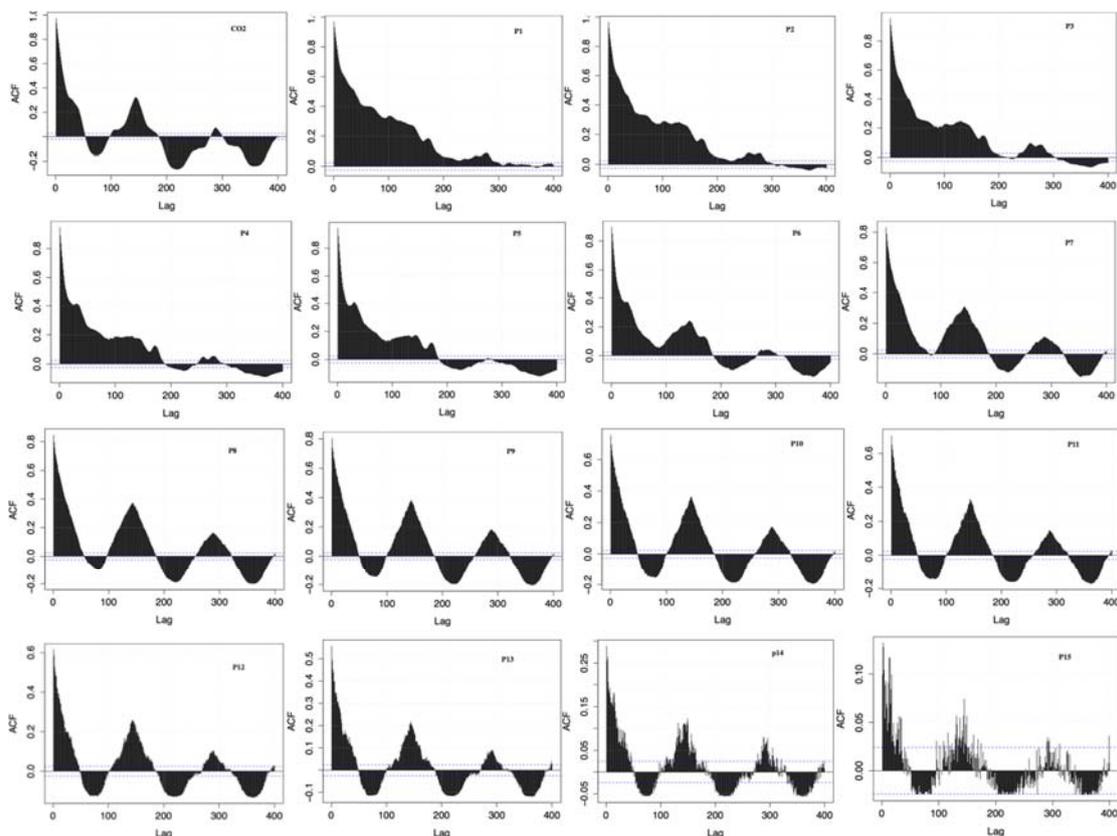


Figure 93 : Fonctions d'autocorrélation des séries temporelles particulières des 15 fractions granulométriques et du CO₂; le retard (lag) est en $\times 10$ minutes. Les deux lignes horizontales pointillées représentent l'intervalle de confiance à 95% pour les autocorrélations.

En revanche, les corrélogrammes correspondants aux tailles de particules fines présentent des autocorrélations relativement élevées avec une décroissance très lente, le premier passage à zéro se produisant après 2 jours seulement. Les fortes valeurs de l'ACF expriment ici, à la fois la force et l'étalement de la persistance. En fait, cette persistance matérialisée par une décroissance lente des autocorrélations (corrélation à long terme) représente un mécanisme complexe associées aux sources de particules fines. Ces sources présentent des fluctuations multi-fréquentielles, c'est-à-dire selon différentes échelles temporelles.

L'Analyse en Composantes Principales a été appliquée aux séries temporelles des polluants (toutes les fractions de particules), chaque série étant considérée comme une variable. Les variables ont été standardisées (centrées-réduites) pour neutraliser l'effet d'échelle provoqué essentiellement par le nombre des particules fines. Chaque variable reçoit ainsi la même importance, quelle que soit sa moyenne et son écart-type.

Les six premières composantes principales expliquent 89% de la variance totale (Figure 94A), avec 55% de variabilité expliquée par la première composante principale. Le corrélogramme associé à la première composante principale (Figure 94B) possède une forme sinusoïdale et alterne de signe. On retrouve la même forme de corrélogramme que celle associée à la concentration de CO₂ et des particules de tailles moyennes (1,6 µm à 7,5 µm). Cette structure semblable d'autocorrélation laisse présager l'existence d'une source commune, vraisemblablement liée à l'occupation, dont le principal indicateur est le CO₂. Les trois autres composantes traduisent la variabilité associée aux particules fines comme on peut en juger avec la forme des fonctions d'autocorrélations associées. Elles peuvent être associées à trois sources distinctes de variation de la concentration des particules fines. Les profils temporels des composantes indiquent le schéma de variation de ces sources dans le temps.

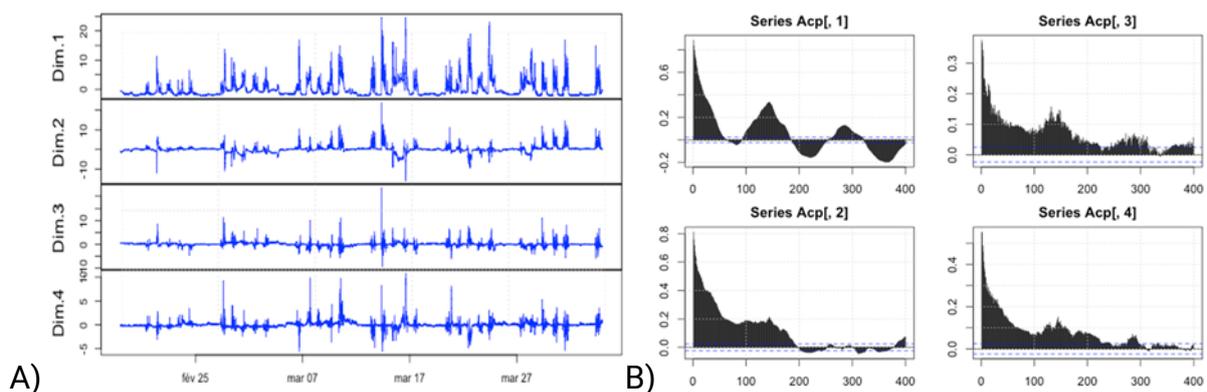


Figure 94 : A) Profils temporels des quatre premières composantes principales. B) : corrélogramme associés aux composantes principales.

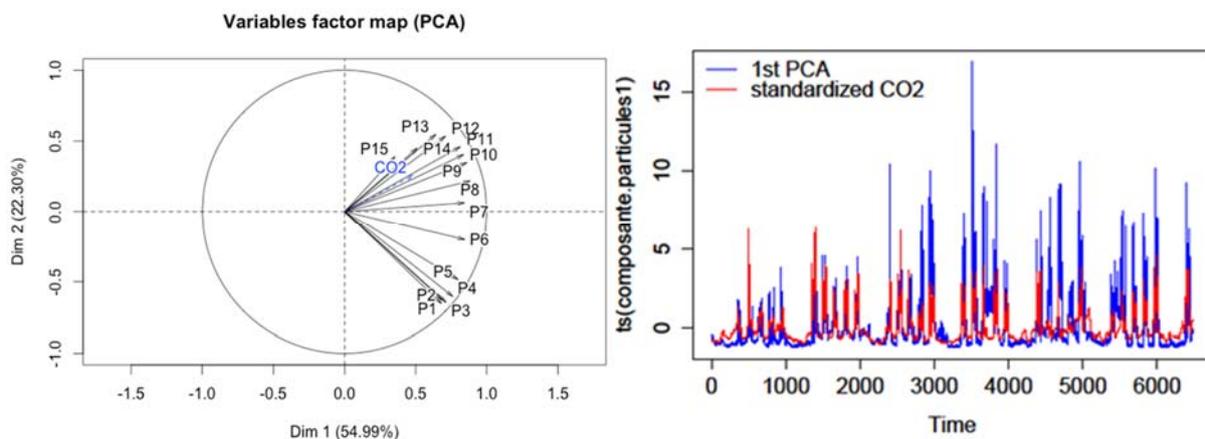


Figure 95 : Plan principal de l'ACP et représentation de la 1ère composante principale et de la concentration de CO₂ standardisée.

La trajectoire des concentrations du CO₂ et de la première composante principale produisent le même type de variation (Figure 95). Elles sont caractérisées par des phases d'alternances (expansion et contraction) des fluctuations. Ceci montre l'importance de l'occupation et l'activité des occupants au regard de la variabilité de la pollution particulaire.

L'analyse en composantes indépendantes (ACI) requiert de fixer en entrée, le nombre de composantes (ou sources de variations). Plusieurs simulations ont été faites afin d'évaluer le critère RMSE (racine carrée de l'erreur quadratique moyenne) en fonction du nombre de composante indépendantes (Figure 96). Cette erreur représente l'écart entre les séries temporelles initiales et les variables reconstituées à partir des composantes indépendantes.

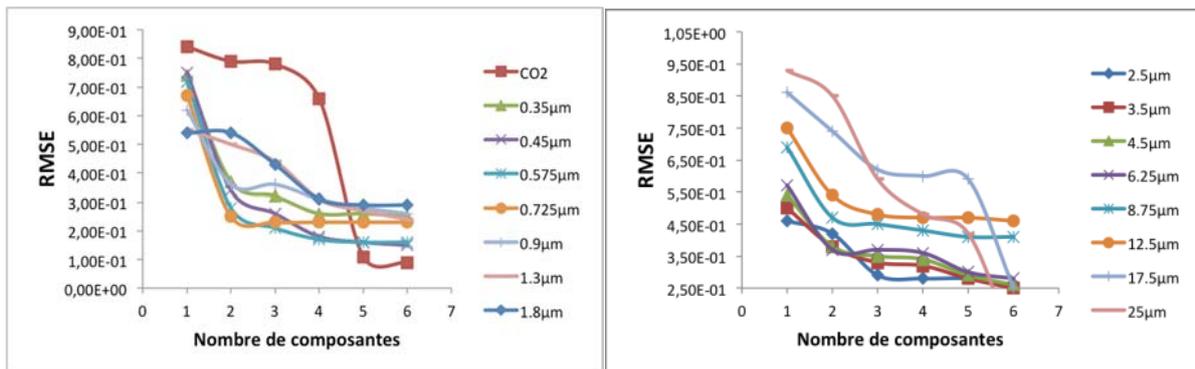


Figure 96 : Evolution du RMSE en fonction du nombre de composantes indépendantes. A gauche : fractions de tailles médianes 0,35-1,8 µm et du CO₂. A droite : fractions de tailles médianes 2,5-25 µm.

Le choix du nombre de composantes retenues est basé sur la recherche du premier minimum local du RMSE. On remarque qu'en se basant sur ce critère, le nombre de composantes à retenir pour chaque fraction varie en fonction de la taille médiane des particules. Pour les séries 0,72 µm, 0,9 µm, 6,25 µm et 8,75 µm, seules 2 composantes suffisent pour atteindre le minimum local. Sinon pour la plupart des fractions, 4 composantes sont nécessaires, et par conséquent, nous retiendrons 4 composantes indépendantes pour l'ACI.

La Figure 97 montre les fluctuations temporelles des quatre composantes indépendantes obtenues sur les 15 fractions, ainsi que les graphiques de la fonction d'autocorrélation de toutes les composantes indépendantes (c1...c4) obtenues sur les 15 canaux.

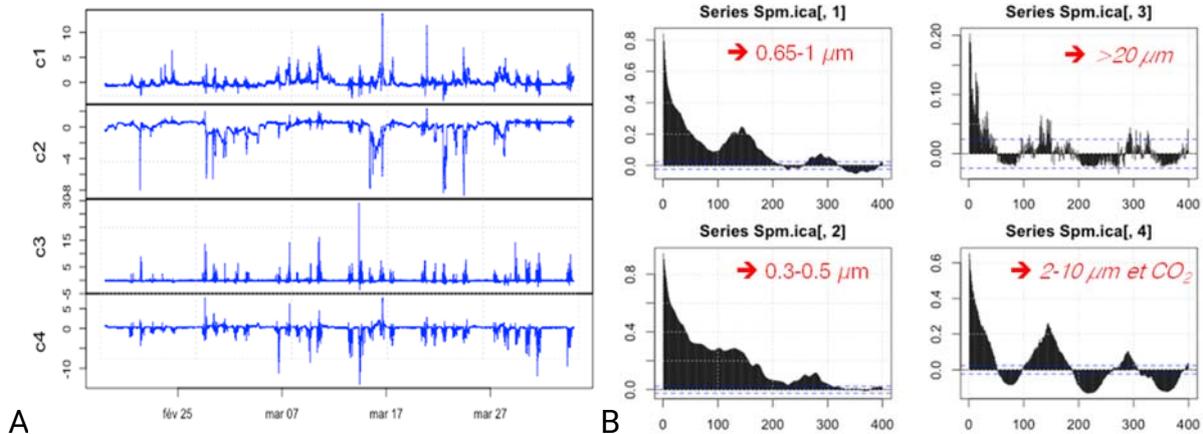


Figure 97 : A) : Profils temporels des quatre premières composantes indépendantes. B) : corrélogrammes associés aux composantes indépendantes.

La deuxième et la quatrième composante indépendante obtenues sont inversées en signe, mais ceci n'est qu'un artefact, car les solutions sont déterminées à un facteur près, donc la multiplication par le facteur (-1) est envisageable. Les autocorrélations de la 1^{ère} et de la 2^{ème} composante sont significatives pour des retards importants (2 jours), montrant ainsi l'influence des observations très éloignées sur les plus récentes, d'où l'idée d'une mémoire longue de la série. La lente décroissance hyperbolique de la fonction d'autocorrélation indique que la variabilité des concentrations des particules fines admet les caractéristiques des séries à mémoire longue. De façon générale, une série présentant des corrélations à long terme est définie comme une série présentant une fonction d'autocorrélation de la forme $\zeta(\tau) = \tau^{-\lambda}$ avec $0 < \lambda < 1$. En d'autres termes, la fonction d'autocorrélation suit une loi de puissance. Ce type de comportement est lié aux émissions des sources chroniques.

Tout comme la première composante principale, la structure de décroissance du corrélogramme de la 4^{ème} composante indépendante indique que celle-ci présente des fluctuations du niveau de particules similaire aux fluctuations du niveau de CO₂. Avec quatre composantes, l'ACI permet de mettre en évidence l'existence d'une source responsable des fluctuations des grosses particules (la 3^{ème} composante indépendante), tandis qu'il faut retenir six dimensions pour l'ACP (95% de variance expliquée) pour pouvoir faire apparaître ce type de variation au niveau des autocorrélations. Les valeurs de l'autocorrélation de la 3^{ème} composante indépendante se présentent comme des pics discrets avec des très faibles corrélations. Cette discontinuité est souvent liée à des phénomènes de faibles probabilités d'occurrence. Les fonctions d'autocorrélations des composantes indépendantes 1 et 2 semblent proches de celles observées pour les composantes principales 2, 3 et 4. Pourtant, les profils temporels de ces composantes ne se ressemblent pas.

La contribution des sources estimée par l'ACI pour chaque composante indépendante sur chaque fraction de particules est présentée à la Figure 98. Les quatre composantes indépendantes contribuent de manière distributive en fonction de la taille des particules. Pour les fractions les plus fines de particules (0,35 μm , 0,45 μm , 0,57 μm), la contribution de la composante 2 avoisine les 80% de l'ensemble des contributions des autres composantes. Elle représente en plus, la source la moins contributive pour les grosses particules (>20 μm). La composante 3 représente sur cette gamme de particules plus de 95% de l'ensemble des autres contributions de sources. Notons par ailleurs la particularité de discontinuité de la fonction d'autocorrélation pour la composante 3 : elle caractérise les émissions dues aux phénomènes aléatoires, tandis que pour la composante 2, les autocorrélations de type fonction de puissance caractérise les émissions des sources déterministes.

Pour la fraction de taille inférieure à 0,9 μm , les composantes indépendantes 2 et 1 partagent à part égale plus de 93% de l'ensemble des contributions. La composante 4 caractérisé par une décroissance sinusoïdale des autocorrélations correspond à la source majoritaire pour la gamme de taille 2-10 μm . Elle est de plus associée aux fluctuations de concentration du CO_2 , elle est par conséquent reliée à la présence d'occupants dans la pièce. Il s'agit vraisemblablement de particules remises en suspension ou générées par les activités de l'occupant dans le bureau.

Au niveau de l'ACP, les coordonnées d'un variable avec chaque axe représentent le coefficient de corrélation linéaire entre cette variable et l'axe en question. Elles sont donc directement associées à la contribution des composantes principales. En prenant en compte la valeur propre de chaque composante, il est alors possible de représenter ces contributions qui peuvent être positives ou négatives selon les variables. Ainsi, la première composante montre une contribution quasiment identique aux différentes tailles de particules, résultat lié à la forte corrélation qui existe entre les différentes fractions. La seconde composante contribue de manière positive aux particules > 2 μm et négativement pour les autres. De même pour les composantes 3 et 4 dont les contributions négatives et positives varient selon la taille des particules.

Les contributions des composantes principales sont très différentes de celles déterminées à partir des composantes indépendantes. Au regard de l'association dans chaque cas d'une des composantes avec la variation du niveau de CO_2 , la contribution de la composante indépendante 4 semble mieux correspondre avec ce qui a été observé par ailleurs : à savoir que les particules les plus fines ne sont pas corrélées avec le CO_2 . Au contraire, la première composante principale, très corrélée avec le CO_2 fournit une contribution pratiquement identique quelle que soit la taille des particules. Dès lors, les composantes indépendantes semblent plus facilement interprétables en termes de sources de fluctuations de particules que les composantes principales.

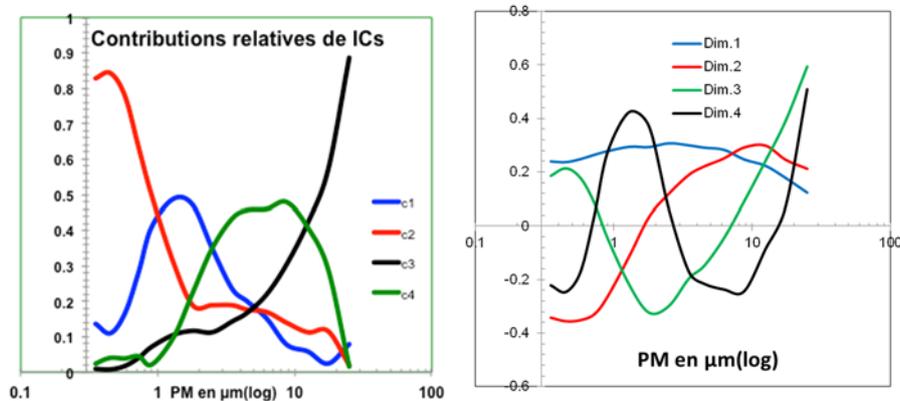


Figure 98 : Contribution relative des composantes indépendantes (gauche) et vecteurs propres de l'ACP selon le diamètre des particules (droite).

4.7.3.2 Séparation de sources du formaldéhyde dans la maison expérimentale

Le jeu de données utilisé pour ce volet comporte la mesure de particules en nombre (0,3 à 20 µm) et du formaldéhyde parmi d'autres paramètres intérieurs et extérieurs sur une période de 4 jours (du 19/04/2010 au 23/04/2010).

Dans ce volet de l'étude, nous avons exploré le modèle de mélange gaussien pour rechercher les sources de fluctuations de la concentration en formaldéhyde dans la maison expérimentale inoccupée. Ce travail a fait l'objet d'une communication aux Journées Interdisciplinaires de la Qualité de l'Air en 2014 à Villeneuve d'Ascq (Ouaret et al., 2014c).

La concentration moyenne de formaldéhyde sur la période est de 14,3 ppb avec un écart-type de 2,6 ppb. L'histogramme de répartition des concentrations suggère une structure polymodale et donc l'existence d'un mélange de distributions. Par conséquent, cette distribution correspond à un mélange de composantes en nombre fini, qui sont modélisées par autant de fonctions de densité de probabilités.

Par ailleurs, l'évolution de la concentration de formaldéhyde dans le temps montre une structure cyclique journalière associée aux variations de température intérieure et d'humidité spécifique. Le profil journalier type de la concentration de formaldéhyde dans la maison expérimentale est corrélé avec celui de la température et de l'humidité spécifique intérieure. Pour rappel dans l'espace de bureau en 2013, la concentration de formaldéhyde ne présentait pas de variations importantes au cours de la journée. Cette différence est attribuée à d'une part l'absence d'ouverture de fenêtres dans la maison expérimentale durant la campagne de mesure et à d'autre part, l'existence d'une source (ou d'un ensemble de sources) caractéristique dont l'émission de formaldéhyde varie avec la température et/ou l'humidité. En l'occurrence dans la maison expérimentale, la source principale de formaldéhyde est l'isolant thermique situé dans les combles avec des émissions de formaldéhyde qui diffusent vers le reste de la maison au travers de la trappe d'accès et du plafond du premier étage.

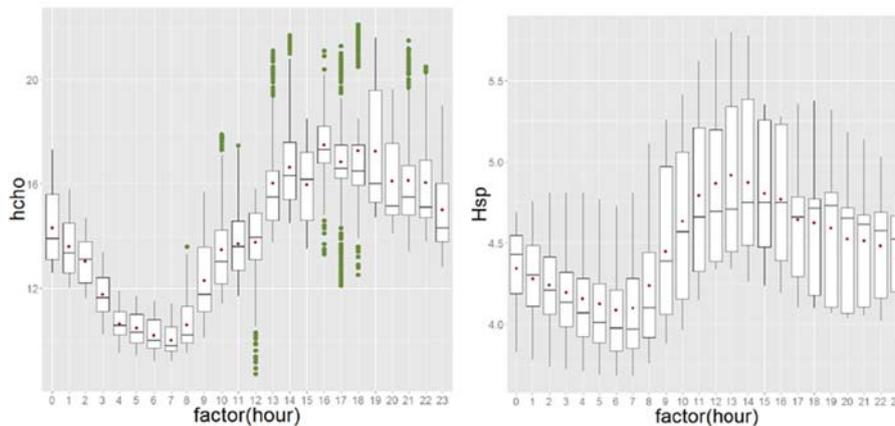


Figure 99 : Profil journalier type de la concentration de formaldéhyde et de l'humidité spécifique (en g/kg) dans la maison expérimentale.

L'estimation d'un mélange gaussien de données préalablement filtrées par une transformée de Fourier rapide (FFT) a été appliqué aux mesures de formaldéhyde (HCHO). En effet, il est possible de considérer que la densité de probabilité de la variabilité temporelle du HCHO totale est la résultante d'un mélange des processus sources. De plus, chaque source est supposée être caractérisée par une distribution de probabilité propre. Cette hypothèse est présentée par Pearson en 1896 : il justifie l'asymétrie d'un histogramme en supposant que l'échantillon qu'il étudie provient de deux populations gaussiennes distinctes et non d'une seule.

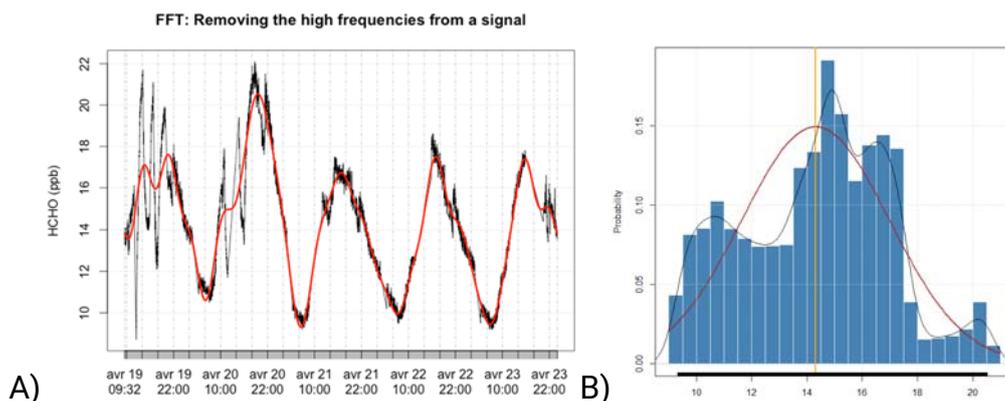


Figure 100 : Extraction des composantes de FFT appliquées aux variations de HCHO (A) et distribution de la concentration de HCHO (B).

Du fait de la forme quasi-sinusoïdale des concentrations observées, nous souhaitons à partir d'une transformée de Fourier rapide (FFT) ne garder que la structure « déterministe » de la dynamique des fluctuations, ensuite estimer les différentes densités de probabilités qui la constituent et qui sont mélangées. La méthode statistique utilisée pour ajuster le modèle aux données est l'estimation par maximum de vraisemblance. Les algorithmes de calcul sont basés sur une méthode de type Newton et l'algorithme est celui de "Expectation Maximization" (EM) (Dempster, et al.

1977). Dans un premier temps, un ajustement a été réalisé en faisant une hypothèse de normalité des distributions de la composante de Fourier. Un autre ajustement a été réalisé avec le logarithme népérien de la composante de Fourier. L'estimation du nombre de modes est basée sur le critère de BIC (*Bayesian Information Criterion*) (Schwarz, 1978). Les différentes simulations pour l'optimisation du critère BIC ont été faites sans contrainte sur les proportions, moyennes et écarts types de différentes composantes que peuvent constituer la densité de probabilité totale.

La Figure 101 montre l'ajustement de mélanges de lois normales appliqué au jeu de données de fréquence de la composante de Fourier du HCHO. Les deux lois de distributions estimées dans le cas d'une transformation logarithmique (Ln) de la composante de Fourier indiquent une première distribution centrée sur 10,6 ppb et une seconde sur 15,2 ppb, correspondant finalement à une moyenne des valeurs faibles (périodes creuses) et à une moyenne des valeurs fortes (crêtes) respectivement. L'ajustement de la distribution de concentration par 5 lois normales est certes plus fin mais également plus discutable en termes d'interprétation de processus sous-jacents permettant d'expliquer ces variations.

L'approche par modèle de mélange gaussien présente dans tous les cas le désavantage de ne pouvoir appréhender la fluctuation temporelle des processus mis en jeu.

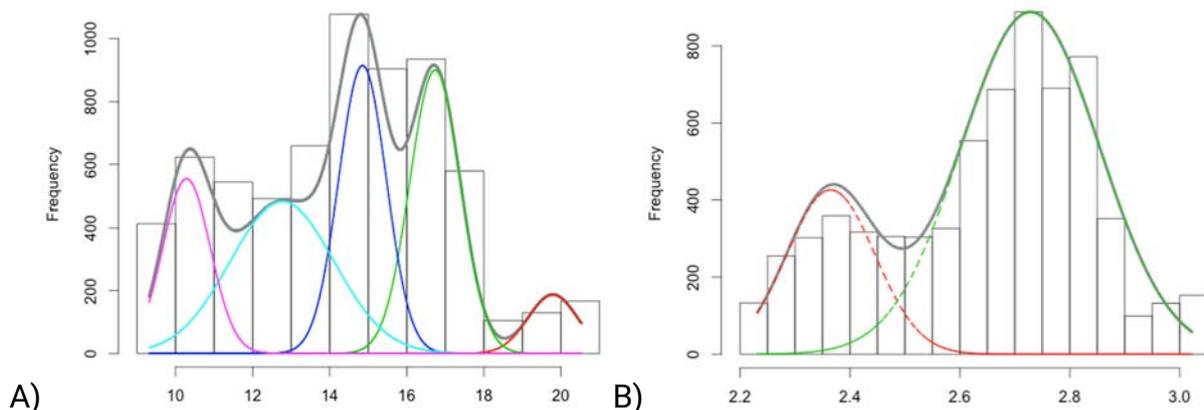


Figure 101 : A) Distribution des concentrations filtrées de HCHO et ajustement de 5 lois normales estimées. B) L'estimation après transformation en logarithme népérien.

4.7.3.3 Analyse par composantes indépendantes des fluctuations de particules dans la maison expérimentale

La campagne de mesure dans la maison expérimentale a été réalisée en absence d'occupation et d'ouverture de fenêtres. Si bien que durant la période couverte dans ce volet de l'étude (4 journées en avril 2010), la contribution majoritaire sur les variations

de concentration intérieure pour les particules fines est d'origine extérieure. Ce travail est intégré à la communication aux Journées Interdisciplinaires de la Qualité de l'Air en 2014 à Villeneuve d'Ascq (Oualet et al., 2014c).

Sur l'ensemble des 8 gammes de tailles de particules comparées entre l'intérieur et l'extérieur, le ratio I/E reste inférieur à 1, voire inférieur à 0,5 pour les particules inférieures à 0,6 μm de diamètre (Figure 102). Pour ces dernières, le ratio atteint un minimum au cours de la journée, qui est lié au fait que le niveau extérieur augmente alors que le niveau intérieur reste stable (peu d'occupation ou d'activités à l'intérieur durant cette période). Pour les particules comprises entre 0,8 et 3 μm , le ratio I/E est relativement constant quel que soit l'heure de la journée. Par ailleurs, des valeurs extrêmes se sont manifestées plutôt dans la nuit (22 h-00 h) pour les très fines particules (0,35 μm et 0,57 μm). Elles sont liées à un pic de particules extérieures dans la nuit du 19 au 20 avril 2010.

La Figure 107 montre les fluctuations temporelles des quatre composantes indépendantes obtenues sur l'ensemble des fractions de particules mesurées à l'intérieur et à l'extérieur. La seconde composante est liée à l'important pic extérieur nocturne enregistré durant cette période. Les autres traduisent des événements d'origine extérieure et notamment le changement de direction du vent de secteur nord-ouest les deux premiers jours accompagnés de niveaux de particules plus élevées vers un vent de secteur nord-est qui apporte moins de particules.

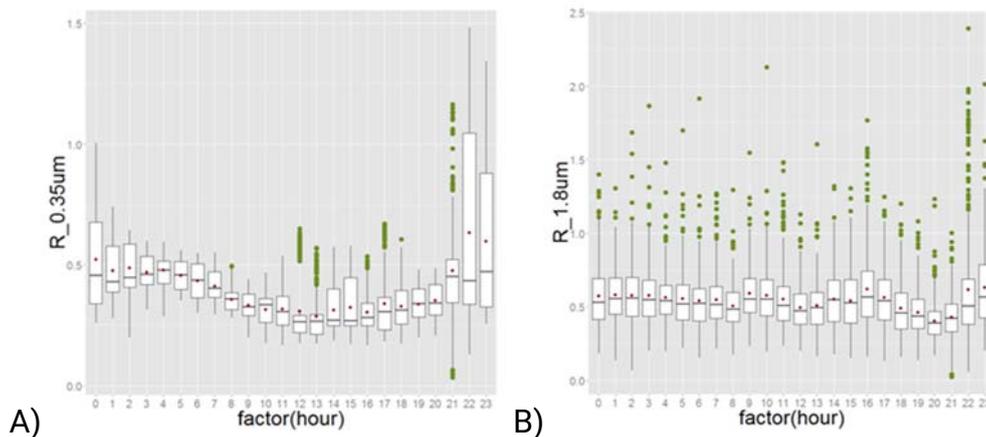


Figure 102 : Profil journalier moyen du ratio I/E des particules de diamètre 0,35 μm (A) et 1,8 μm (B) dans la maison expérimentale

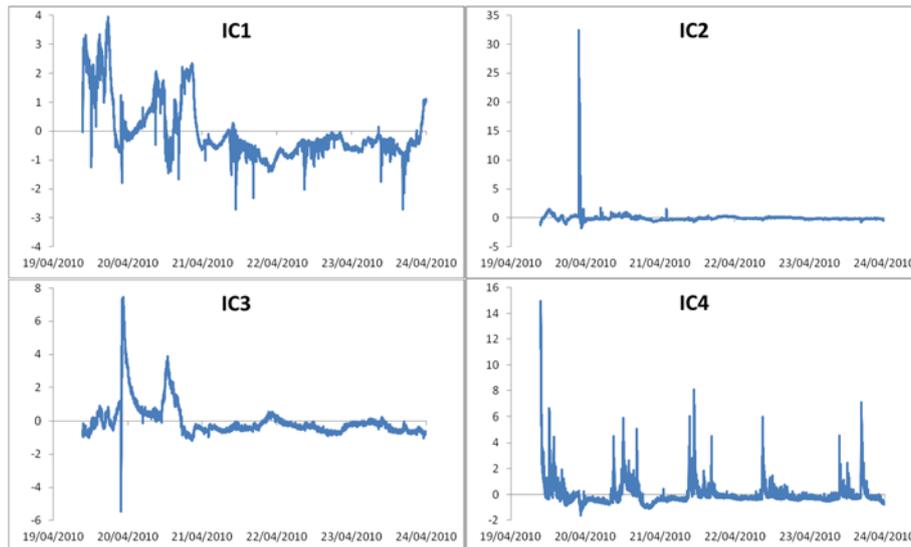


Figure 103 : Profils temporels des quatre composantes indépendantes extraites

La contribution des sources estimée par l'ACI pour chaque composante indépendante sur les fractions de particules intérieures et extérieures sont représentées à la Figure 104. Les quatre composantes contribuent de façon distributive en fonction de la taille des particules.

Pour les fractions les plus fines ($0.35 \mu\text{m}$, $0.45 \mu\text{m}$), la contribution de la première composante (IC1) se situe entre 50 % et 60 % à l'intérieur comme à l'extérieur. Le profil de cette composante est similaire entre l'intérieur et l'extérieur sur la gamme de taille $0,35 - 2,5 \mu\text{m}$. Son profil temporel montre des événements qui contribuent à l'augmentation de la concentration de particules les deux premiers jours puis qui à l'inverse y contribue de façon négative. Cette composante marque vraisemblablement le bruit de fond associé au trafic et le changement de direction du vent.

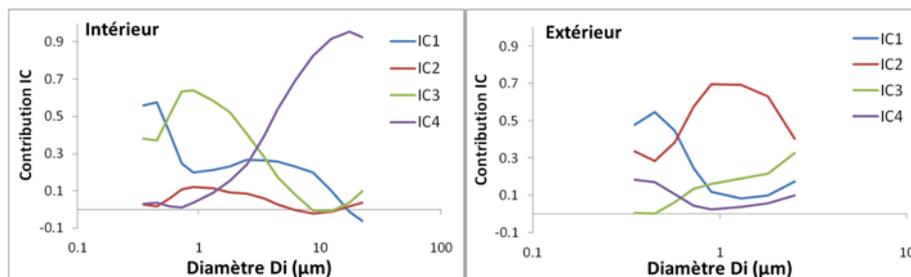


Figure 104 : Contribution relative des composantes indépendantes selon le diamètre des particules à l'intérieur jusqu'à $20 \mu\text{m}$ (gauche) et à l'extérieur jusqu'à $2,5 \mu\text{m}$ (droite).

La seconde composante (IC2) ne contribue pratiquement pas dans l'air intérieur (moins de 10 % quel que soit le diamètre des particules). Par contre, elle est majoritaire dans l'air extérieur pour les fractions comprises entre $0,8$ et $2 \mu\text{m}$, fractions

correspondantes au pic nocturne enregistré à l'extérieur. Son impact ayant été beaucoup moins ressenti à l'intérieur (rapport du maximum/moyenne durant le pic de 8 à l'intérieur contre 50 à l'extérieur), il est normal que la contribution de cette composante soit plus faible à l'intérieur.

La troisième composante est plutôt associée aux particules comprises entre 0,6 et 2 µm à l'intérieur (50 % à 60 % de contributions) et se caractérise par deux événements importants l'un correspondant au pic nocturne de particules et l'autre qui survient dans l'après-midi du 20 avril. La contribution de cette composante est pratiquement nulle par la suite. A l'extérieur, cette composante est plutôt associée avec les particules > 0,8 avec une contribution maximale de 30 % pour les particules de 2,5 µm.

La quatrième composante est marquée par une série de pics irréguliers qui reviennent chaque jour. La contribution de cette composante aux particules extérieures reste très faible (< 20 %). Elle est par contre très importante à l'intérieur (> 50 %) pour les particules de plus de 5 µm. Ces pics sont en grande partie associés à des pics de concentration de CO₂ qui marquent les rares et brèves interventions des expérimentateurs dans la maison. L'émission de ces grosses particules est vraisemblablement associée à une remise en suspension par les mouvements des expérimentateurs.

En comparant les composantes indépendantes identifiées dans la maison expérimentale avec celles identifiées dans le bureau individuel, quelques similitudes sont observées dans leur profil de contribution selon le diamètre des particules. Ainsi, la composante IC1 (maison) présente une forme voisine de la composante IC2 (bureau) avec une forte contribution des particules les plus fines (< 0,6 µm). La composante IC3 (maison) est proche de la composante IC1 (bureau) avec une prédominance des particules entre 1 et 2 µm. La composante IC4 (maison) est de même similaire à la composante IC3 (bureau) avec une forte contribution des grosses particules > 10 µm. La composante IC4 (bureau) associée à l'occupation ne se retrouve pas dans la maison expérimentale où l'occupation est plus épisodique. De même, la composante IC2 (maison) associée à un événement particulier d'origine extérieure n'apparaît pas au niveau du bureau individuel.

4.7.3.4 Comparaison de 4 méthodes de séparation des sources ACI, PMF, NNMF et ACP

Une comparaison de 4 méthodes de séparation des sources : l'analyse par composantes indépendantes (ACI), la factorisation positive (Positive Matrix Factorization), la factorisation non-négative (NNMF) et l'analyse par composantes principales (ACP) a été réalisée sur des données mesurées dans le bureau individuel en 2011. Cette comparaison a fait l'objet d'un poster présenté à l'International Aerosol Conference à Busan (Corée du Sud) en août 2014 (Ouaret et al., 2014b).

Chacune de ces méthodes utilise une métrique particulière pour extraire les vecteurs sources à partir des observations.

Tableau 8 : Métrique utilisée par les 4 méthodes de séparation des sources

Matrix factorization conditions				
Method	Initialisation	Algorithm/Procedure	Metric	Iteration
ICA		Fast ICA	Neg-Entropy	200
NMF	random	Brunet	Kullback-Leibler	20
PMF		Multilinear Engine		bootstrap analysis
PCA	-	Orthogonal linear transformation	Euclidian	-

Un des paramètres critiques de ces méthodes est l'estimation du nombre de sources/facteurs en jeu. Dans la plupart des études, il est imposé considérant que le processus global est connu et seules les contributions respectives sont recherchées. Ce n'est pas le cas, lorsque ces méthodes sont utilisées en aveugle. Ce nombre de sources peut toutefois être estimé en comparant les erreurs résiduelles de plusieurs runs. Par exemple, dans cette étude le nombre optimal de facteurs à prendre en compte est de 4 dans le bureau individuel et de 3 dans l'espace de bureau, correspondant au point d'inflexion.

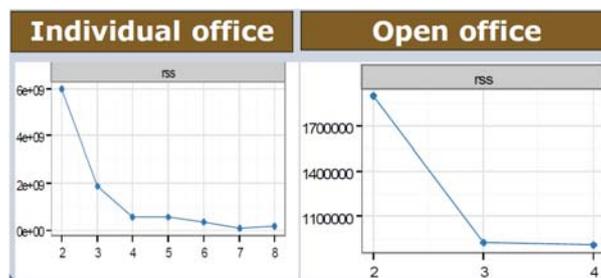
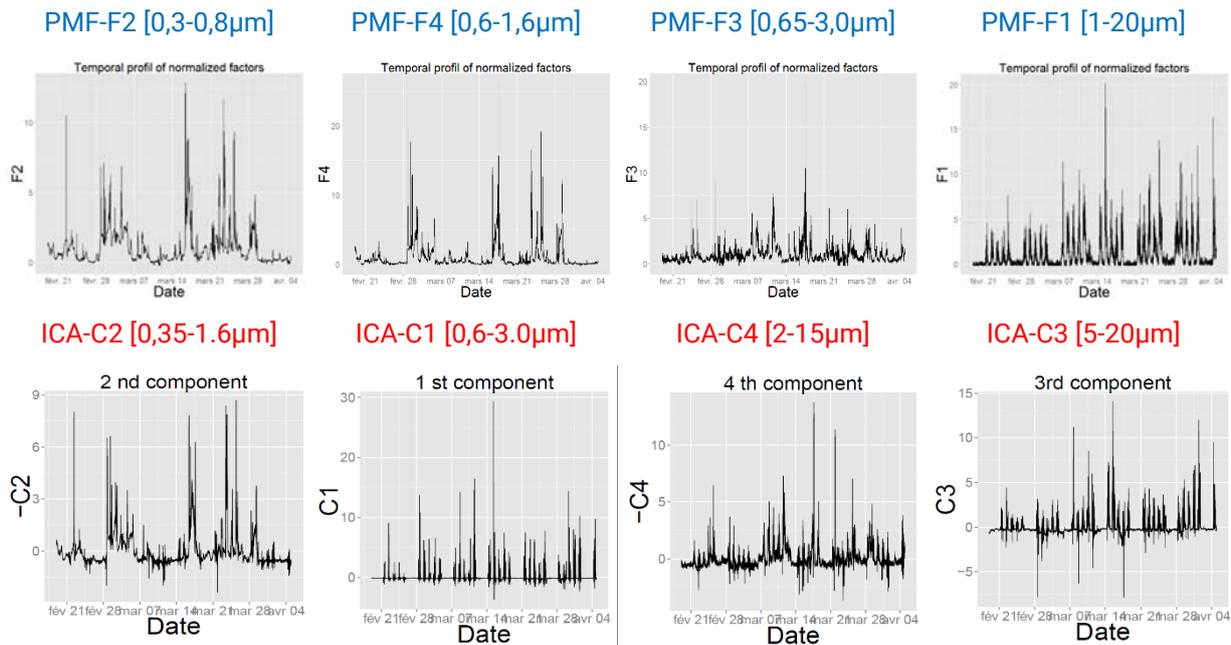


Tableau 9 : Erreur résiduelle obtenue après plusieurs runs (rss : residual sum of squares)

Les profils temporels normalisés des 4 facteurs extraits par la PMF et par l'ACI pour le bureau individuel sont représentés ci-dessous. Chaque facteur est associé à la gamme de taille de particules à laquelle il contribue le plus, à hauteur : de plus de 10 % pour la PMF et à hauteur de plus de 30 % pour l'ACI. Le facteur le plus à droite F1 pour la PMF et C3 pour l'ACI présente des fluctuations temporelles qui rappellent celles du CO₂. Les composantes de l'ACI apparaissent plus bruitées avec des valeurs négatives. Pourtant, les profils des composantes de l'ACI présentent des allures similaires aux facteurs de la PMF. Ainsi la composante C2 évoque les facteurs F1 et F2 par ailleurs très

ressemblants. Les composantes C3 et C4 rappellent respectivement les facteurs F3 et F1.

Figure 105 : Profils temporels des facteurs/composantes extraites par PMF ou ACI de la variation des concentrations de particules dans le bureau individuel



Les contributions de chaque facteurs aux différentes tailles de particules selon les différentes méthodes sont représentées sur la figure suivante. La numérotation des composantes variant selon la méthode, elle n'est pas à prendre en compte. Les composantes obtenues par ICA et NMF présentent des profils de contributions relativement similaires. La composante C2 de l'ACI et H3 de la NMF pourraient correspondre en termes de profils de contributions à l'association des facteurs F2 et F4 de la PMF. A l'inverse, le facteur 1 de la PMF semble plutôt combiner les composantes C3 et C4 de l'ACI ou H2 et H4 de la NMF. Les trois méthodes fournissent des résultats assez proches en termes d'interprétation possible. La PMF se différencie des deux autres par le fait que ses facteurs affichent un contraste marqué dans les contributions des différentes tailles de particules, alors que pour l'ACI et la NMF les contributions sont rarement nulles.

Les composantes extraites par l'ACI représentent des sources (ou groupes de sources « indépendantes » au sens statistique. Quel que soit le critère statistique utilisé (indépendance) ou mathématique (positivité des matrices), les 3 méthodes de séparation des sources fournissent des facteurs robustes interprétables sur le plan temporel et sur la contribution moyenne des particules.

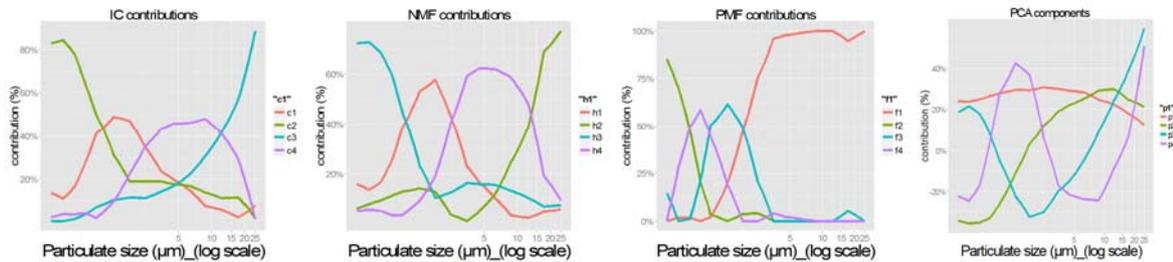


Figure 106 : Profils de contribution moyenne des facteurs aux différentes tailles de particules selon la méthode d'extraction

Pour l'ACP, chaque composante est une combinaison linéaire des différentes tailles de particules. Pour être homogène avec les autres procédures, seules les 4 premières composantes sont prises en compte. La contribution est alors calculée en comparant le coefficient de la variable aux coefficients de cette même variable sur les quatre composantes. Ce coefficient correspond à la contribution de la variable au facteur. Les contributions négatives sont prises en compte et conduisent à la difficulté d'interpréter les profils de contributions. Ces contributions calculées sont plus associées à la corrélation linéaire de la variable avec la composante que des contributions au sens strictes.

Le choix des quatre premières composantes est arbitraire et d'autres composantes moins bien représentées pouvaient avoir également du sens. Toutes les tailles de particules contribuent à la première composante du fait de leur corrélation significative entre elles. L'information qu'elle fournit est donc assez pauvre. De manière générale, il apparaît nécessaire d'adapter l'ACP ou d'utiliser des variantes de l'ACP pour permettre l'interprétation des composantes de la même façon que les autres méthodes.

4.7.3.5 Discussion sur la séparation des sources de fluctuations

Ces premiers travaux montrent la similitude des différentes méthodes pour séparer à l'aveugle les sources contributives au niveau de particules observé dans un environnement intérieur. L'ACP classique n'offre pas le même confort d'interprétation que les autres méthodes.

Pour autant, séparer n'est pas identifier. Sans informations extérieures, les facteurs extraits restent difficilement identifiables autrement que par l'observation de fluctuations caractéristiques d'un processus particulier. Après leur extraction, une seconde phase d'exploitation doit donc être menée pour rechercher les associations avec d'autres phénomènes observés.

Ces travaux ont pris appui sur des mesures de particules en nombre relativement corrélées entre elles dont l'ACI, la NMF et la PMF ont su extraire les composantes indépendantes ou caractéristiques. Il est donc nécessaire de disposer de plusieurs

enregistrements d'un même polluant pour que l'ACI soit utilisable. Néanmoins, nous nous posons la question de son utilisation sur un jeu de données hétérogènes, mais pour lequel nous suspectons des processus ou sources communes permettant d'expliquer les différentes variations observées. Ainsi, des variables exogènes comme la température, l'humidité, l'occupation et l'ouverture des fenêtres seraient autant de paramètres qui permettraient de se prononcer sur l'interprétation des composantes extraites. Pour l'instant, ces données sont utilisées a posteriori de l'analyse. Mais rien n'interdit de les introduire dans l'analyse elle-même à condition de les exprimer toutes sous une même métrique.

4.8 Prédiction des concentrations intérieures de polluants

L'ouverture des fenêtres influence en partie les variations de la concentration des polluants gazeux et particulaires des environnements intérieurs, et en particulier le formaldéhyde et les particules de taille intermédiaire. Prévoir la variation de leur concentration intérieure serait sans doute facilitée par la prédiction préalable de l'ouverture des fenêtres.

4.8.1 Prédiction de l'ouverture des fenêtres

Plusieurs approches statistiques ont été appliquées afin de prédire l'ouverture des fenêtres. Des régressions logistiques ont été réalisées pour expliquer les 3 variables associées à l'ouverture des fenêtres (le nombre de fenêtres ouvertes, l'état global ouvert ou fermé de toutes les fenêtres (0/1), l'état global d'ouverture en 3 catégories (fenêtres fermées / 1 ou plusieurs fenêtres ouvertes / toutes fenêtres ouvertes). Neuf variables quantitatives explicatives (heure de la journée, vitesse du vent, température extérieure, irradiance extérieure, pression, pluie, CO₂ intérieur, température intérieure et l'humidité relative intérieure) et 5 variables qualitatives explicatives (présence, direction du vent, jour de la semaine) ont été introduites. La variable « état de la porte » n'a pas été introduite car elle était significativement dépendante de la présence. La direction du vent n'a pas été prise en compte non plus. La procédure PROC LOGISTIC du logiciel SAS a été utilisée. Cette procédure utilise une approximation d'une étape de validation croisée pour obtenir des probabilités prédites. Les observations sont ensuite classées selon que les probabilités d'événements prévus sont au-dessus ou au-dessous une valeur de césure z dans la gamme de 0 à 1. Une observation est prédite comme un événement si la probabilité de l'événement prédit est supérieure ou égale à z . En outre, les faux positifs et faux négatifs peuvent être calculés comme des probabilités a posteriori en utilisant le théorème de Bayes. Les modèles présentés dans ce rapport convergent.

De plus, deux régressions logistiques polytomiques additives ont été appliquées pour rechercher les déterminants des ouvertures des fenêtres composées de plus de 2 modalités. La régression logistique polytomique additive permet de comparer une catégorie de la variable dépendante (aucune fenêtre ouverte) avec toutes les autres catégories qui lui sont supérieures (au moins une fenêtre ouverte/toutes les fenêtres ouvertes et 1, 2, 3 ou 4 fenêtres ouvertes). Ce type de régression a été préféré à la régression logistique polytomique cumulative car l'effet d'une variable est significativement différent selon sa modalité. Les modèles présentés dans ce rapport vérifient cette hypothèse. La technique d'optimisation utilisée est l'algorithme de Newton-Raphson.

Par ailleurs, une régression linéaire généralisée a été effectuée sur le nombre de fenêtres prise comme variable continue. Une k-validation croisée a été effectuée (avec

$k = 10$). De plus, une recherche du meilleur modèle a été effectuée par la méthode STEPWISE.

Pour finir, les prévisions ont également été effectuées par arbre de classification à l'aide de la méthode CART (*Classification And Regression Tree*). Les méthodes de segmentation qui regroupent les arbres de classification et de régression cherchent à résoudre les problèmes de discrimination ou de régression en segmentant de façon progressive l'échantillon pour obtenir un arbre de décision binaire. Elle vise à prédire une variable cible, qualitative (arbre de classification) ou quantitative (arbre de régression), à partir de variables « prédictives » qualitatives ou quantitatives. C'est une méthode non paramétrique et non linéaire : aucune hypothèse implicite n'est faite sur la relation entre la variable expliquée et les variables indépendantes.

Le principe est de chercher la dichotomie (induite à chaque pas par l'une des variables) la plus "liée" à la variable à expliquer. La méthode de segmentation consiste à rechercher d'abord la variable qui explique le mieux l'ouverture des fenêtres. Cette variable définit une première division de l'échantillon en deux sous-ensembles appelés segments. La procédure est réitérée à l'intérieur de chacun de ces deux segments en recherchant la deuxième meilleure variable, et ainsi de suite.

Lorsque la variable à expliquer est continue, le critère de sélection de la meilleure division d'un nœud est basé sur la variance de cette variable dans les segments descendants. Celle-ci doit être plus faible dans le segment descendant par rapport au nœud parent.

Le résultat final de la segmentation par arbre binaire est une règle de décision qui conduit à classer les individus de la population étudiée dans un des nœuds terminaux de l'arbre de classification ou de régression retenu.

Le défaut de la méthode tient à son caractère séquentiel car les covariations qui servent à sélectionner les variables ne mesurent pas un lien causal et une variable peut en cacher une autre, beaucoup plus fondamentale, qui n'a plus aucune chance d'apparaître dans la suite du processus (Lebart et al., 2006).

Différents algorithmes sont disponibles pour construire un arbre de décision binaire. La méthode CART a été utilisée. La méthode CART (*Classification and Regression Tree*) n'impose aucune règle (ou seuil) d'arrêt de division des segments (Breiman et al., 1984). Elle fournit à partir de l'arbre binaire complet, la séquence des sous-arbres obtenue en utilisant une procédure d'élagage. Celle-ci est basée sur la suppression successive des branches les moins informatives en terme de discrimination, permettant d'obtenir un sous-arbre optimal en se fondant sur l'estimation de l'erreur théorique d'affectation obtenue par un échantillon test ou par validation croisée.

La détermination de la bonne taille de l'arbre s'effectue par post-élagage. L'arbre est dans un premier temps complètement développé avec le critère de pureté sur un premier échantillon, puis, dans un second temps, il est réduit de manière à optimiser le taux de mauvais classement calculé sur un second échantillon dit d'élagage. La

variable cible ayant deux modalités, le critère de division (critère d'impureté, c'est-à-dire l'homogénéité du nœud) est le critère de Gini : l'indice d'impureté représente la probabilité qu'une observation tirée au hasard parmi les individus du nœud soit mal classée.

Le logiciel utilisé (CART, Salford Systems) ne permet pas d'utiliser plus de 50 000 observations pour réaliser l'arbre de classification. Une sélection aléatoire des données a été réalisée en conservant 50 000 observations parmi les 525 600 disponibles sur une année. Une k-validation croisée a été réalisée avec k=10 pour l'élagage. Pour chaque modèle, l'arbre optimal choisi correspond à l'arbre avec le moins de nœuds possibles apportant une sensibilité de 98 %. En effet, 2% d'erreur sur une année, c'est plus de 10 000 minutes d'erreur soit plus de 150 heures d'erreur. Cet arbre peut être différent de l'arbre optimal déterminé par le logiciel. La variable MOIS a été prise en compte dans les arbres de classification. Néanmoins, lors de la prédiction pour les autres années cette variable a été remplacée par la variable concurrente car pour l'année 2015 les données sont collectées entre janvier et juillet.

4.8.1.1 Résultats de la régression logistique binaire sur l'année 2015

La variable à expliquer est une variable binaire (fenêtre ouverte/fenêtre fermée). 244 379 observations ont été incluses dans le modèle (sur 305 220) dont 34 634 avec une fenêtre ouverte (soit 14 %).

Les résultats du modèle sont présentés au Tableau 10 et Tableau 11. La présence des occupants, la vitesse du vent, la température intérieure et extérieure contribuent le plus à l'ouverture des fenêtres. A l'inverse, la pluie et l'heure de la journée contribue plus à leur fermeture.

Le modèle prédit correctement les situations de fenêtres fermées dans 96,6 % des cas (spécificité) et les situations de fenêtres ouvertes dans seulement 54,5 % des cas (sensibilité). Un état de fenêtres prédit comme fermées l'est effectivement dans 93 % des cas (valeur prédictive négative). Un état de fenêtres prédit comme ouvertes l'est effectivement dans 73 % des cas (valeur prédictive positive ou VPP). Le gain apporté par le modèle dans la prévision des situations de fenêtres ouvertes, mesuré comme l'écart entre la VPP et la prévalence initiale de 14 %, est de 58 %.

Des résultats du même ordre ont été observés en utilisant 2013 ou 2014 comme données d'apprentissage, avec des sensibilités parfois meilleures mais au détriment d'une plus mauvaise spécificité.

Tableau 10 : Résultats de la régression logistique binaire (modalité de la probabilité d'ouvrir les fenêtres)

	Odd-Ratio	95% Intervalle de confiance de Wald	
Présence (Réf : aucun mouvement)	1.611	1.515	1.714
Température_ext	1.290	1.283	1.297
VitesseVent	1.167	1.160	1.175
Température_int	1.140	1.126	1.155
Humidité relative_int	1.081	1.079	1.084
Pression atmosphérique	1.065	1.062	1.067
Irradiance_ext	0.999	0.999	1.000
CO₂_int	0.995	0.995	0.995
Heure de la journée	0.913	0.910	0.915
Pluie	0.756	0.732	0.781
p-value du test du rapport de vraisemblance : <0,0001 p-value du test de Wald : <0,0001 AIC : 107607 -2LL : 107586 D de Somers : 0,855 Gamma : 0,855 Tau-a : 0,208 c = 0,927 AUC(ROC) = 0.9273			

Tableau 11 : Tableau de classification : ouverture observée / prédite

		Ouverture prédite		
	Effectif % % en ligne % en colonne	CLOSE	OPEN	Total
Ouverture observée	CLOSE	202616 82.91 96.60 92.79	7129 2.92 3.40 27.41	209745 85.83
	OPEN	15754 6.45 45.49 7.21	18880 7.73 54.51 72.59	34634 14.17
Total		218370 89.36	26009 10.64	244379 100.00

4.8.1.2 Résultats de la prévision de l'ouverture de plus d'une fenêtre par régression

Un modèle de régression logistique polytomique cumulatif a été utilisé pour expliquer la variable d'ouverture des fenêtres en 3 modalités (toutes les fenêtres sont ouvertes – au moins une fenêtre est ouverte – les fenêtres sont fermées) sur l'année 2015 notamment. 244 379 observations ont été incluses dans le modèle (sur 305 220) dont 3 758 où toutes les fenêtres sont ouvertes, 30 876 où au moins une fenêtre est ouverte et 209 745 où toutes les fenêtres sont fermées.

Le modèle est moins sensible pour prédire les situations avec au moins une fenêtre ouverte (44,7 %) et surtout les situations où toutes les fenêtres sont ouvertes sont considérées à 85 % comme les cas à une seule fenêtre ouverte.

Le même constat est observé en utilisant comme variable à expliquer le nombre de fenêtres ouvertes en 4 modalités (1, 2, 3 ou 4+ fenêtres ouvertes). 244 379 observations ont été incluses dans le modèle (sur 305 220) dont 3 079 où 4 fenêtres ou plus sont ouvertes, 4 209 observations où 3 fenêtres sont ouvertes, 6 930 observations où 2 fenêtres sont ouvertes, 20 416 observations où 1 fenêtre est ouverte et 209 745 observations où aucune fenêtre n'est ouverte.

Le modèle présente la meilleure sensibilité pour la prédiction des situations avec 2 fenêtres ouvertes (60 %), mais affiche une sensibilité inférieure à 18 % dans les autres cas. Ces modèles apparaissent donc moins performants que la régression logistique binaire.

Des régressions linéaires multiples (RLM) ont également été utilisées pour prédire le nombre de fenêtres ouvertes selon la présence et les paramètres climatiques. La performance de ces modèles n'est pas exceptionnelle notamment pour la prévision des situations où les fenêtres sont fermées. Par contre, une situation où une seule fenêtre est ouverte est rarement considérée comme une situation avec fenêtres fermées. Le modèle RLM ne permet pas de distinguer entre 1 et plusieurs fenêtres ouvertes.

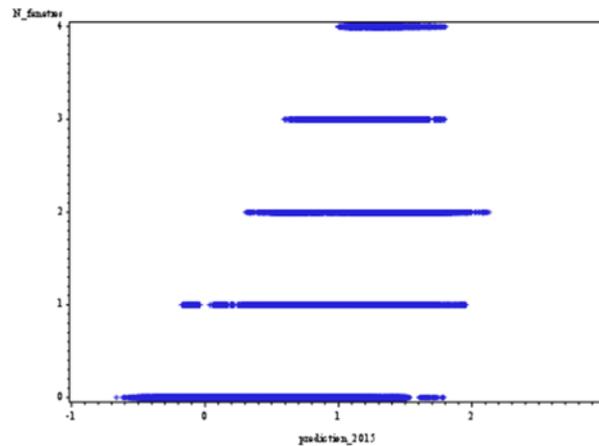


Figure 107 : Prédiction du nombre de fenêtres ouvertes en 2015 selon le modèle RLM de 2013.

4.8.1.3 Prédiction de l'ouverture des fenêtres par arbre de décision (CART)

Sans conteste les meilleures performances ont été observées avec la méthode CART pour la prédiction de l'état d'ouverture général des fenêtres (variable binaire : ouvert / fermé). La méthode CART a été utilisée sur les données des années 2013, 2014 et 2015 prises séparément. Les données d'apprentissage étaient composées à chaque fois de 50 000 observations (données sur 1 minute) en intégrant comme variables explicatives les données calendaires, de présence et les paramètres climatiques.

Pour l'année 2013, la variable à expliquer à travers l'arbre de classification est l'ouverture des fenêtres (en binaire). L'arbre optimal a un coût relatif de 0,087 et possède 704 nœuds terminaux. Pour l'année 2014, l'arbre optimal obtenu a un coût relatif de 0,030 et possède 598 nœuds terminaux. Et sur les données de 2015, l'arbre optimal a un coût relatif de 0,009 et possède 132 nœuds terminaux.

Le Tableau 12 regroupe les principales variables qui ont contribué à la construction de l'arbre optimal de classification de l'état d'ouverture des fenêtres selon les années. La variable principale est le mois de l'année suivie par les conditions de température et d'humidité extérieures et intérieures.

Tableau 12 : Importance relative de l'ensemble des variables (séparatrices, concurrentes et suppléantes) dans l'arbre de classification selon les années

	2013	2014	2015
Variable	Score	Score	Score
Mois	100,00	100,00	100,00
Température_Ext	96,89	68,23	52,86
Humidité Relative_Int	95,58	44,71	2,28
Température_Int	93,17	36,12	43,22
Humidité Spécifique_Int	93,04	64,98	41,63
Humidité Spécifique_Ext	87,76	67,46	37,36
CO ₂ _Int	85,53	19,85	3,82
Pression atmosphérique	85,26	29,22	7,81
Humidité Relative_Ext	56,70	16,81	2,78
Irradiance_Ext	49,04	8,10	1,09
Jour_Semaine	46,30	26,98	2,55
Vitesse_vent	36,61	6,29	0,83
Irradiance_Int	27,15	5,07	5,01
Heure de la journée	17,91	8,47	1,92
Etat de la porte	9,38	5,66	Non utilisé
Direction du vent	8,37	7,16	0,56
Pluie	6,13	0,68	0,04
Présence	1,75	0,31	0,00

La performance des trois modèles CART obtenus sur les données d'apprentissage est présentée au Tableau 13. La sensibilité et la spécificité dépassent systématiquement 95 %. Il reste toutefois à vérifier les performances des modèles sur des observations qui n'ont pas servi à la construction de l'arbre.

Pour vérifier ces performances, les règles qui définissent l'arbre de classification ont été exportées et appliquées à l'ensemble des autres observations de la même année. Cela permet de vérifier le pouvoir prédictif de l'ouverture des fenêtres la même année. Les résultats sont présentés au Tableau 14. La prédiction de l'ouverture des fenêtres reste correcte avec près de 78 % pour l'année 2013 avec l'arbre CART 2013. Ces données souffrent d'importantes données manquantes relatives à l'état d'ouverture des fenêtres. Par contre, la prédiction fournie par les arbres CART de 2013 et 2014 montre d'excellentes performances avec respectivement une sensibilité de 98,5 % et 99,3 % sur les observations de 2013 et 2014, tout en conservant une spécificité (prédiction des situations de fermeture des fenêtres de plus de 98 %).

Le modèle CART construit sur quelques observations d'une année parvient donc à prédire l'état d'ouverture et de fermeture des fenêtres de l'espace de bureau sur la même année par l'intermédiaire de données calendaires, de données climatiques et de présence des occupants avec un taux d'erreur inférieur à 2 %.

Tableau 13 : Performance de classification des modèles CART sur données d'apprentissage

2013			
OBSERVE	% bonne prédiction (en ligne)	PREDIT	
		Fermée	Ouverte
Fermée	95,56 %	23 778	1 104
Ouverte	95,76 %	609	13 763
2014			
OBSERVE	% bonne prédiction (en ligne)	PREDIT	
		Fermée	Ouverte
Fermée	98,67 %	26 997	365
Ouverte	98,35 %	363	21 637
2015			
OBSERVE	% bonne prédiction (en ligne)	PREDIT	
		Fermée	Ouverte
Fermée	99,65 %	37 895	134
Ouverte	99,49 %	61	11 900

Tableau 14 : Performance de prédiction de l'ouverture des fenêtres la même année

Prédiction 2013 | CART 2013

	Fréquence % % en ligne % en col.	Prédit		
		CLOSE	OPEN	Total
Observée	CLOSE	208 258	17 342	225 600
		58.87	4.9	63.77
		92.31	7.69	
	87.85	14.86		
OPEN	28 794	99 353	128 147	
	8.14	28.09	36.23	
	22.47	77.53		
	12.15	85.14		
Total		237 052	116 695	353 747
		67.01	32.99	100
Fréquence manquante = 121 854				

Prédiction 2014 | CART 2014

	Fréquence % % en ligne % en col.	Prédit		
		CLOSE	OPEN	Total
Observée	CLOSE	256 139	3 905	260 044
		54.55	0.83	55.39
		98.5	1.5	
	98.77	1.86		
OPEN	3 181	206 290	209 471	
	0.68	43.94	44.61	
	1.52	98.48		
	1.23	98.14		
Total		259 320	210 195	469 515
		55.23	44.77	100
Fréquence manquante = 6 085				

Prédiction 2015 | CART 2015

	Fréquence % % en ligne % en col.	Prédit		
		CLOSE	OPEN	Total
Observée	CLOSE	188 577	413	188 990
		79.79	0.17	79.96
		99.78	0.22	
	99.82	0.87		
OPEN	336	47 025	47 361	
	0.14	19.9	20.04	
	0.71	99.29		
	0.18	99.13		
Total		188 913	47 438	236 351
		79.93	20.07	100
Fréquence manquante = 18 869				

La question se pose de savoir si un modèle construit sur une année permet de prédire l'état d'ouverture des fenêtres sur d'autres années. Le modèle est-il performant au-delà de l'année en cours ? Pour répondre à cette question, les règles de l'arbre de classification testées ont été appliquées aux données observées les autres années. Les résultats obtenus sont présentés du Tableau 15 au Tableau 17. Les performances ne sont pas à la hauteur de celles atteintes pour prédire l'ouverture des fenêtres la même année. Le modèle CART 2013 est le moins performant lorsqu'il est appliqué sur les données de 2015. En effet, il affiche la sensibilité la plus faible de 39 %, et surtout ses prédictions de fenêtres ouvertes sont fausses dans près de 73 % des cas (VPP de 27 %). Le modèle CART 2014 n'est pas non plus très performant lorsqu'il est appliqué sur les données de 2013 malgré une sensibilité de 61 %, mais sa prédiction de fenêtres

ouvertes est fautive dans 60 % des cas et il affiche la plus mauvaise spécificité de 48 %. C'est pourtant le modèle CART 2014 qui fournit les meilleures performances lorsqu'il est appliqué sur les données 2015 avec une sensibilité de 82 % et surtout une valeur prédictive négative de 93 % (prédiction de fenêtres fermées correctes dans 93 % des cas). C'est finalement le modèle CART 2015 qui apparaît comme le plus robuste avec des résultats de sensibilité (47 %) et de spécificité (82 %) très voisins sur 2013 et 2014. Il affiche également les VPP les plus élevées, supérieures à 57 %.

Tableau 15 : Performance de prédiction de l'ouverture des fenêtres avec le modèle CART 2013 sur d'autres années

Prédiction 2014 | CART 2013

	Fréquence % % en ligne % en col.	Prédit		
		CLOSE	OPEN	Total
Observée	CLOSE	191 334	84 165	275 499
		38.43	16.9	55.33
		69.45	30.55	
	62.01	44.44		
OPEN	117 198	105 223	222 421	
	23.54	21.13	44.67	
	52.69	47.31		
	37.99	55.56		
Total		308 532	189 388	497 920
		61.96	38.04	100
Fréquence manquante = 27 680				

Prédiction 2015 | CART 2013

	Fréquence % % en ligne % en col.	Prédit		
		CLOSE	OPEN	Total
Observée	CLOSE	140 967	69 338	210 305
		50.69	24.93	75.63
		67.03	32.97	
	77.24	72.55		
OPEN	41 546	26 231	67 777	
	14.94	9.43	24.37	
	61.3	38.7		
	22.76	27.45		
Total		182 513	95 569	278 082
		65.63	34.37	100
Fréquence manquante = 27 138				

Tableau 16 : Performance de prédiction de l'ouverture des fenêtres avec le modèle CART 2014 sur d'autres années

Prédiction 2013 | CART 2014

	Fréquence % % en ligne % en col.	Prédit		
		CLOSE	OPEN	Total
Observée	CLOSE	115 484	122 948	238 432
		31.05	33.06	64.11
		48.43	51.57	
	69.05	60.07		
OPEN	51 751	81 716	133 467	
	13.92	21.97	35.89	
	38.77	61.23		
	30.95	39.93		
Total		167 235	204 664	371 899
		44.97	55.03	100
Fréquence manquante = 103 702				

Prédiction 2015 | CART 2014

	Fréquence % % en ligne % en col.	Prédit		
		CLOSE	OPEN	Total
Observée	CLOSE	175 006	54 722	229 728
		57.87	18.1	75.97
		76.18	23.82	
	93.02	47.89		
OPEN	13 124	59 535	72 659	
	4.34	19.69	24.03	
	18.06	81.94		
	6.98	52.11		
Total		188 130	114 257	302 387
		62.21	37.79	100
Fréquence manquante = 2 833				

Tableau 17 : Performance de prédiction de l'ouverture des fenêtres avec le modèle CART 2015 sur d'autres années

Prédiction 2013 | CART 2015

	Fréquence % % en ligne % en col.	Prédit		
		CLOSE	OPEN	Total
Observée	CLOSE	126 738	26 235	152 973
		54.91	11.37	66.28
		82.85	17.15	
		75.08	42.31	
Observée	OPEN	42 070	35 767	77 837
		18.23	15.5	33.72
		54.05	45.95	
		24.92	57.69	
Total	Total	168 808	62 002	230 810
		73.14	26.86	100
Fréquence manquante = 244 791				

Prédiction 2014 | CART 2015

	Fréquence % % en ligne % en col.	Prédit		
		CLOSE	OPEN	Total
Observée	CLOSE	230 718	49 195	279 913
		45.48	9.7	55.18
		82.42	17.58	
		65.56	31.67	
Observée	OPEN	121 226	106 121	227 347
		23.9	20.92	44.82
		53.32	46.68	
		34.44	68.33	
Total	Total	351 944	155 316	507 260
		69.38	30.62	100
Fréquence manquante = 2 833				

La différence de performance du modèle CART lorsqu'il est appliqué la même année en comparaison à d'autres années pourrait s'expliquer par le fait que les conditions combinées de calendrier, de paramètres climatiques et de présence qui ont conduit à l'ouverture des fenêtres une année donnée n'entraîne plus forcément l'ouverture des fenêtres pour d'autres années. Les occupants n'étant pas toujours les mêmes d'une année sur l'autre peuvent avoir des critères de décision d'ouverture des fenêtres différents entre eux. De la même manière, un occupant donné peut modifier son critère de décision d'une année sur l'autre. Néanmoins, les performances du modèle comparées à la fréquence initiale d'ouverture des fenêtres sur une année montrent que le modèle CART améliore la prédiction par rapport à une répartition des événements d'ouverture dues au hasard. Le modèle CART 2015 fournit ainsi un gain de +24 % pour prédire l'ouverture des fenêtres et de +10 % pour prédire la fermeture des fenêtres.

Le modèle CART ne prend pas en compte la dimension temporelle ni les changements d'état. Les observations sont utilisées comme des états indépendants. Les fluctuations des variables explicatives peuvent conduire à basculer plusieurs fois d'un état ouvert à un état fermé dans un court laps de temps, ce qui n'est pas toujours le cas en réalité. Les excellentes performances des modèles observées sur une même année ne vont pas dans ce sens. Mais ces modèles pourraient sans doute être améliorés en intégrant cette dimension temporelle.

Un modèle CART a également été utilisé pour prédire le nombre de fenêtres ouvertes (variable qualitative). L'arbre optimal construit sur les données de 2014 a un coût relatif de 0,087 et possède 1250 nœuds terminaux. Sa performance sur les données d'apprentissage varie de 97,7 % (0 fenêtre ouverte) à 89,1 % (5 fenêtres ouvertes). Ses performances de prédiction sur l'année 2015 restent médiocres avec un taux de bien classés au global de 60 % concentrés sur les deux modalités principales (0 et 1 fenêtre ouverte).

4.8.2 Prédiction de la concentration intérieure en formaldéhyde par régression

Un premier moyen de prédire la concentration intérieure en formaldéhyde est d'utiliser les autres données disponibles. Une régression linéaire de la concentration intérieure de formaldéhyde a été réalisée en intégrant comme variables du modèle les paramètres climatiques intérieurs (température, humidité spécifique, irradiance) et extérieurs (vitesse et direction du vent, pression atmosphérique, pluie, température, humidité spécifique et irradiance), les occupants (présence, concentration de CO₂, activité de l'imprimante mesurée par impulsions électriques) et l'état des ouvrants (porte et fenêtres codées en binaire ouvert/fermé). Etant donné que la régression vise essentiellement la prédiction et non l'explication de la concentration du formaldéhyde, la multicollinéarité entre variables n'est pas prise en compte.

Un exemple de prédiction par régression linéaire a été réalisé sur les données de 2015. Il a ensuite été appliqué aux données observées en 2013 et 2014. Les résultats de ce modèle sont présentés au Tableau 18, Tableau 19 et Figure 108. Les concentrations de CO et d'ozone intérieures étaient initialement introduites dans le modèle mais elles n'amélioreraient pas sensiblement les performances et ont donc été retirées.

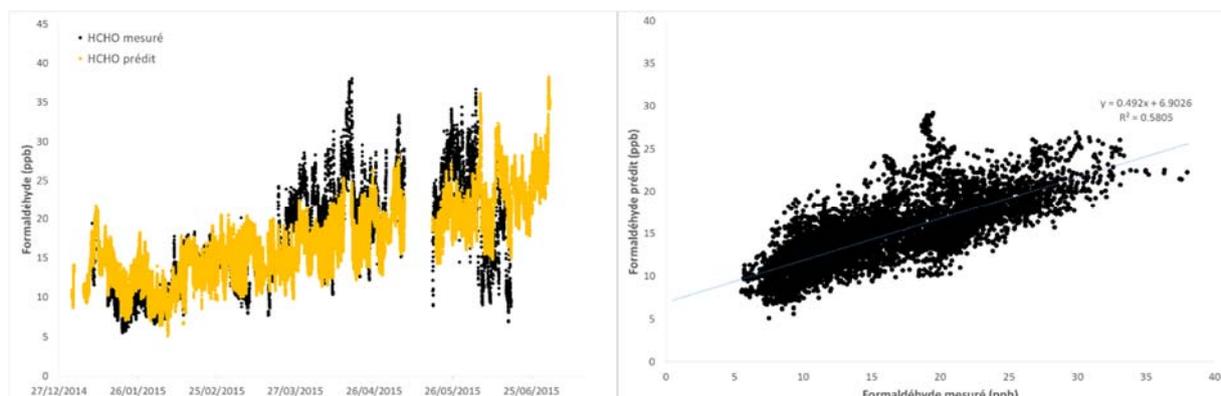


Figure 108 : Prédiction de la concentration intérieure en formaldéhyde par régression en 2015

Tableau 18 : Paramètres d'ajustement de la régression linéaire du formaldéhyde (2015)

Test d'ajustement global

Libellé de la statistique	Valeur
Somme des carrés des écarts (SCE)	46767
Coefficient de corrélation multiple (R)	0.7324
Coefficient de corrélation multiple (R ²)	0.5365
Variance estimée des résidus (S ²)	9.60
Ecart-type estimé des résidus (S)	3.10
Test de nullité simultanée des coefficients	
Fisher	352.5
Nb de degrés de liberté au numérateur	16
Nb de degrés de liberté au dénominateur	4873
Probabilité critique	< 0.0001
Valeur-Test	60.4428

Tableau 19 : Coefficients de la régression linéaire de la concentration intérieure en formaldéhyde (2015)

Ajustement des moindres carrés (avec terme constant)

4890 individus, 17 paramètres (constante en queue).

Paramètre	Coefficient	Ecart-type	Student - 4873	Probabilité
Porte – fermée ouverte	-1.17036	0.077	15.292	<0.0001
Humidité spécifique_int	2.12439	0.161	13.161	<0.0001
Température extérieure	0.25500	0.023	11.260	<0.0001
CO ₂ _intérieur	-0.00728	0.001	8.020	<0.0001
Pluie – non oui	-0.68189	0.092	7.416	<0.0001
Température_intérieure	0.39008	0.054	7.262	<0.0001
Vitesse du vent	-0.18517	0.029	6.346	<0.0001
Humidité spécifique_ext	-0.37040	0.113	3.264	0.001
Direction du vent	-0.00124	0.000	2.845	0.004
Impulsions imprimante	0.07953	0.037	2.156	0.031
Pression	0.00963	0.005	1.945	0.052
Irradiance_extérieure	0.00071	0.000	1.444	0.149
Présence – non oui	-0.10393	0.094	1.109	0.267
Fenêtres – fermées ouvertes	0.19150	0.268	0.714	0.475
Irradiance_int	-0.00070	0.001	0.498	0.618
CONSTANTE	-12.2737	4.887	2.511	0.012

La régression parvient à s'ajuster assez bien à la concentration de formaldéhyde observée tant que les données des variables d'entrée sont disponibles. La présence

de valeurs manquantes dans ces paramètres ne permet pas de fournir des prédictions sur certaines périodes. Le modèle se basant sur 15 variables d'entrée, le risque n'en est que plus grand. L'irradiance, la présence (mais pas la concentration en CO₂) et l'état des fenêtres contribuent peu à la prédiction du formaldéhyde. L'état de la porte d'accès à l'espace de bureau et l'humidité spécifique intérieure sont par contre des paramètres importants dans la prédiction.

Lorsque ce modèle est appliqué aux données de 2013, les prédictions sont systématiquement supérieures aux observations. Cela indique qu'il manque un ou plusieurs paramètres au modèle responsable des variations observées entre 2013 et 2015. Sur l'année 2014, le modèle de régression ne rend pas compte des importantes amplitudes de variation observées au cours de l'année.

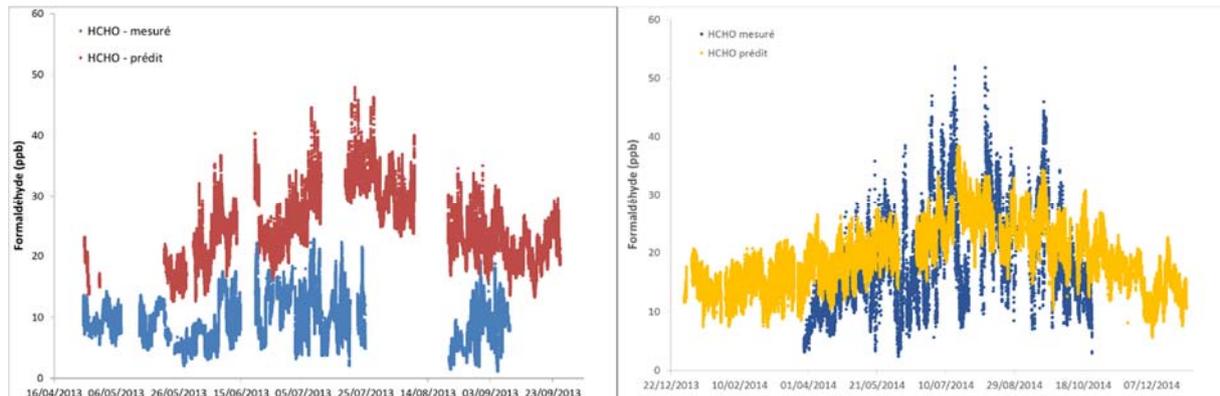


Figure 109 : Application du modèle de régression 2015 pour prédire la concentration de formaldéhyde en 2013 (gauche) et 2014 (droite).

Pour faire de la prévision, ce type de modèle n'est pas optimal car il nécessite de disposer des mesures et surtout des prévisions de chacun des paramètres entrant dans le modèle pour permettre de prédire à chaque instant futur la concentration du formaldéhyde. Les prévisions des paramètres climatiques extérieures existent et pourraient être utilisées. Ceux de l'environnement intérieur pourraient également être estimés à partir des prévisions extérieures. Mais la présence, les activités et l'état des ouvrants nécessitent le développement de modèles de prévision spécifiques.

A des fins de prévision de l'évolution de la concentration, il est dès lors plus intéressant de se pencher sur des modèles ne requérant qu'un minimum de connaissances, à savoir un historique des mesures de concentration enregistrées au préalable sans introduire d'informations complémentaires (par exemple des données climatiques ou d'autres mesures de polluants). Ce type de prévision va chercher à exploiter les propriétés intrinsèques à la dynamique des séries temporelles.

4.8.3 Structure temporelle des mesures en continu et prévisions

Les paramètres mesurés dans l'environnement fournissent généralement des suites numériques, et les valeurs prises par un paramètre donné sont des réalisations d'un ensemble de variables aléatoires et forment, avec l'indice temps ($t \in T$) une "image de la trajectoire". Pour des applications empiriques, on dispose que d'un nombre fini de données, donc de réalisation d'un nombre fini de variables ($X(t_1), X(t_2), \dots, X(t_n)$) pour une suite finie d'instant $\{t_1 < t_2 < \dots < t_n\}$. Ici, X_{t_i} et $X(t_i)$ sont interchangeables du moment qu'on ne spécifie pas la nature discrète ou continue de T . $X(t)$ étant réservée généralement aux processus à temps continu, nous ne distinguerons pas le processus X_t d'une de ses réalisations $X_t(\omega)$.

La suite d'observations possède au moins deux caractéristiques :

- (a) en dehors de l'évolution à peu près polynômiale d'une tendance, les observations présentent des aspects locaux très irréguliers, nécessitant un très grand nombre de liberté pour pouvoir modéliser la courbe associée aux observations,
- (b) le processus "physique" est irréversible : impossible de reproduire la suite de mesures dans leurs conditions réelles.

Qu'elle soit exprimée par une concentration ou par un rapport de quantités, la série d'observations associée aux réalisations des variables aléatoires fournit des informations sur l'état du système qui l'a générée. Les informations portées par ces variables sont plus au moins "bruitées", elles dépendent du type du paramètre étudié et les causes qui l'ont généré. En résumé, l'aspect irrégulier traduit la variabilité des observations, elle est d'autant plus importante avec une résolution temporelle fine.

Lorsqu'on travaille sur un nombre fini de variables aléatoires, nous cherchons souvent les dépendances ou des structures de corrélations entre les différents retards des observations, en d'autres termes leur niveau d'autocorrélation. Les processus à mémoire courte se manifestent par une décroissance rapide (exponentielle) de la fonction d'autocorrélation vers zéro. Au contraire, une série comporte une mémoire longue si l'autocorrélation décroît très lentement (fonction puissance) vers zéro. Puisque les fonctions d'autocorrélation et la densité spectrale forment une paire de transformée de Fourier, la connaissance de l'une implique la connaissance de l'autre.

La transformée de Fourier est un outil de base de l'analyse spectrale pour représenter les séries temporelles, et qui va permettre de décomposer les séries en fonction de leurs périodicités. Ainsi, les basses fréquences vont représenter les dynamiques de long terme et de faible variabilité tandis que les hautes fréquences vont décrire les dynamiques de court terme.

Une autre approche à l'analyse des séries temporelles est la décomposition en tendance, cycles (ou/et saisonnalité) et en composantes aléatoires. Cette décomposition repose sur l'idée selon laquelle la structure inhérente de la série

temporelle est un composite de plusieurs éléments déterministes et stochastiques qu'il faut étudier séparément.

Le problème de prévision se résume à une question fondamentale, à savoir comment à partir des observations x_1, \dots, x_T , il est possible d'obtenir des valeurs futures $x_{T+1}; x_{T+2}; \dots$ jugées "utiles" compte tenu du phénomène étudié. Pour des observations univariées, il s'agit de modéliser la dynamique à travers laquelle les valeurs passées influencent la valeur présente. Dans le cadre des processus stochastiques, les valeurs observées sont supposées être des réalisations des variables aléatoires X_1, \dots, X_T dont il faudra spécifier la dynamique. La prévision à la date $T+1$ consiste, d'un point de vue mathématique, à projeter X_{T+1} sur un ensemble de fonctionnelles de X_1, \dots, X_T (projection linéaire, espérance conditionnelle, etc.), et à estimer l'intervalle de précision associé à cette projection.

La modélisation univariée et la démarche générale de prévision par la procédure de Box-Jenkins (1970) sont fondées sur les processus ARIMA(p; d; q) : *Autoregressive Integrated Moving Average* avec p l'ordre du modèle autoregressif, d le degré de différenciation et q l'ordre de la moyenne mobile. La procédure proposée est en quatre étapes :

- Identification : on utilise généralement l'estimation de la fonction d'autocorrélation (ACF) et de la fonction d'autocorrélation partielle (PACF). En revanche, il existe d'autres méthodes d'identification développées à partir des critères d'entropie. Citons, entre autre le critère AIC d'Akaike (1974) (*Akaike's Information Criterion*), le critère BIC de Schwarz et al. (1978) (*Bayesian Information Criterion*) et le critère CAT de Parzen (1975) (*Criterion Autoregressive Transfer Function*) ;
- Estimation : différentes procédures itératives peuvent être appliquées pour l'estimation des paramètres du processus ARMA. Les plus utilisées pour le critère des moindres carrés sont les méthodes de type Gauss-Newton ou le "compromis de Marquardt" ;
- Tests de diagnostic : il s'agit d'examiner si le modèle estimé est ou non compatible avec les hypothèses sous-jacentes aux modèles. Il en existe plusieurs : des tests graphiques de l'autocorrélation des résidus et des tests basés sur la statistique de 2 tel que le test de Ljung-Box ou celui de Box-pièce ;
- Prévisions : consiste à respecter le principe d'extrapolation en utilisant les formules de la prévision optimale pour les processus ARMA (voir ci-après).

Dans le cadre de l'analyse des structures de variabilité de la concentration en polluants dans l'air intérieur, l'hypothèse de stationnarité n'est pas tenable. Donc, la prévision par un modèle de type ARMA sur ces séries, conduit dans la plupart des cas à un échec. En revanche, si l'on considère une série de transformation des données, par exemple, une désaisonnalisation du logarithme de la série puis différenciation, l'hypothèse de

stationnarité devient plus vraisemblable. Le résultat de la désaisonnalisation (la série désaisonnalisée) est ensuite modélisé par, une approche type ARMA, lissage exponentiel, ou autre. Les différents types de transformations dépendent énormément de la nature du polluant, de la présence des variations abruptes, ainsi que de la résolution temporelle (pas de temps).

4.8.4 Prédiction de la concentration intérieure en formaldéhyde par régression non-paramétrique STL

La méthode "Seasonal Trend Loess" (STL) proposée par Cleveland (1979) est une méthode de régression non-paramétrique par une pondération localement lisse (locally weighted scatterplot smoother (lowess)). Pour une application aux concentrations de CO₂ extérieures, Cleveland et al. (1990) montrent l'effet de la tendance au cours des dernières années.

Soit $(x_i)_{i=1,\dots,n}$ la série temporelle à décomposer, avec le temps comme variable indépendante et x_t variable dépendante. L'idée de la régression linéaire locale réside dans l'analyse des moindres carrés localement polynomiale (type ajustement k-NN). De manière plus précise, elle consiste à ajuster une fonction polynomiale au voisinage du point t_0 de la variable d'intérêt. On note le voisinage de t par $N(t_0)$. Il s'agit de trouver le paramètre β :

$$\arg \min_{\beta} \sum_{t_i \in N(t_0)} w\left(\frac{t_i - t_0}{\lambda}\right) \left[X_i - \sum_{j=0}^p \beta_j t^j \right]^2$$

avec le paramètre de lissage λ et la pondération par une fonction kernel w . Le noyau proposé par Cleveland (1979) et qui a été dès lors, le plus utilisé est le noyau cubique :

$$w(t_i - t_0; \lambda) = \max\left(\left[1 - \left|\frac{t_i - t_0}{\lambda}\right|^3\right]^3, 0\right)$$

L'approche par décomposition STL requiert de fixer la fréquence. La procédure STL est conçue principalement pour des séries de basses fréquences, donc pour des données journalières, voire même mensuelles. Ceci constitue un inconvénient majeur pour l'analyse de nos données, qui sont enregistrées avec un pas de temps fin. Néanmoins, cette méthode peut être appliquée pour des séries qui exhibent un comportement saisonnier.

Les performances de cette méthode ont été évaluées sur un enregistrement en continu du formaldéhyde au pas de temps de 1 minute réalisé en 2010 dans le séjour de la maison expérimentale MARIA du CSTB en absence d'occupants. Ces données montraient en effet un caractère saisonnier marqué, idéal pour tester la méthode. Elle a également été testée sur des données avec un caractère saisonnier moins

prononcé : la concentration de particules également mesurées dans la maison expérimentale sur la même période.

Les décompositions obtenues sont présentées sur les deux figures suivantes. Chaque figure indique les données de départ, sa composante saisonnière, sa tendance et le reste de la décomposition.

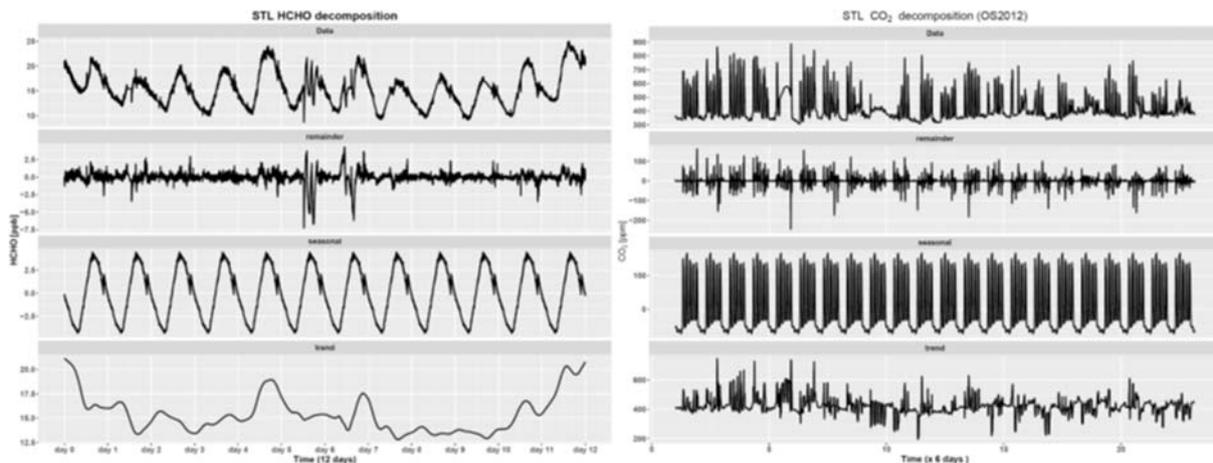


Figure 110 : Décomposition par STL de la variabilité de la concentration en formaldéhyde dans la maison expérimentale (gauche) et comparaison avec le CO₂ mesuré dans l'espace de bureau en 2012 (droite).

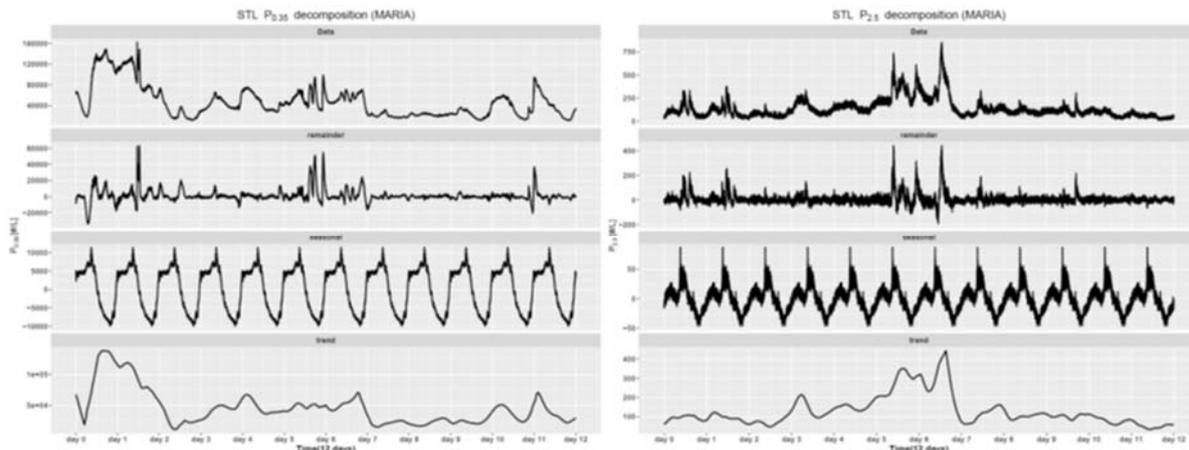


Figure 111 : Décomposition par STL de la variabilité de la concentration en particules de 0,35 µm et 2,5 µm dans la maison expérimentale.

Les fréquences fondamentales de la décomposition saisonnière sont d'un jour⁻¹ pour le formaldéhyde et à titre de comparaison de 7 jours⁻¹ pour le CO₂ mesuré dans l'espace de bureau en 2012.

Cette méthode permet également de donner une indication sur la contribution de chaque composante dans les fluctuations de la série originale. Elle se calcule en termes d'écart-type relatif par rapport à l'écart-type de la série originale et aux écarts-types des autres composantes. Après décomposition de la série, la somme des écarts-types des trois composantes est généralement supérieure à l'écart-type estimé de la série temporelle initiale (σ_p). La contribution relative (Ct_r) de variabilité de chaque composante peut être estimée comme suit :

$$\widehat{Ct}_r = \frac{\widehat{\sigma}_C - \frac{\widehat{\sigma}_{diff}}{3}}{\widehat{\sigma}_p}$$

où $\widehat{\sigma}_C$ est l'écart-type de chaque composante : la saisonnalité S , la tendance T et le reste e ($\widehat{\sigma}_S, \widehat{\sigma}_T, \widehat{\sigma}_e$), et $\widehat{\sigma}_{diff} = (\widehat{\sigma}_S + \widehat{\sigma}_T + \widehat{\sigma}_e) - \widehat{\sigma}_p$.

Le Tableau 20 donne les contributions relatives de chaque composante de la décomposition STL de quelques séries temporelles de polluants et paramètres mesurés dans la maison expérimentale et dans l'espace de bureau.

Tableau 20 : Contribution relative en variabilité des composantes de différents paramètres mesurés dans la maison expérimentale (MARIA) ou dans l'espace de bureau en 2012 (OS2012). La colonne $t:span$ (trend span) correspond à la taille de la fenêtre retard (délai) pour l'extraction de la tendance.

Environnement	Variable	$t.span$ (h)	Contribution (Ct_r %)		
			Saisonnalité	Tendance	Reste
MARIA	HCHO	8	60.34	35.94	3.72
	P _{0.35} -P _{0.45}	8	7.36-6.21	79.26-75.60	13.38-18.19
	P _{0.575} -P _{0.725}	8	5.48-9.35	64.01-44.92	30.51-45.72
	P _{0.9} -P _{1.3}	8	8.26-7.52	46.56-54.62	45.18-37.86
	P _{1.8} -P _{2.5}	8	7.52-6.78	54.62-64.47	37.86-28.75
	P _{3.5} -P _{4.5}	8	7.27-14.46	42.48-20.12	50.25-65.41
	P _{6.25} -P _{8.75}	8	20.20-20.26	6.12-3.27	73.67-76.47
	P _{12.5} -P _{17.5}	1.2	19.46-23.38	27.57-18.32	52.96-58.28
OS2012	CO ₂	6	57.25	39.19	3.56
	P _{0.35} -P _{0.9}	4	9.16-12.96	87.12-80.37	3.72-6.67
	P _{1.8} -P _{4.5}	4	17.05-36.61	75.65-58.05	7.29-5.34
	P _{8.75}	5	46.69	42.24	11.07
	T _{int} -HS _{int}	8-9	34.43-11.04	62.01-88.37	3.56-0.59
	T _{ext} -HS _{ext}	12-6	19.23-9.37	80.46-89.21	0.31-1.42

Bien que la série du formaldéhyde soit au pas de temps d'une minute, la décomposition semble relativement facile et une prévision par un modèle basée sur cette décomposition est envisageable car la composante saisonnière (ici diurne) est majoritaire, de l'ordre de 60 %, contre 36 % pour la tendance et 4 % pour le reste de la décomposition.

La contribution de la tendance polynomiale reste majoritaire pour les particules de taille entre $0:35 \mu\text{m}$ et $2:5 \mu\text{m}$ et diminue avec l'augmentation de leur diamètre. La contribution relative de la saisonnalité augmente avec la taille des particules. Les plus grosses particules sont plutôt associées à une contribution importante des résidus de la décomposition.

Il s'avère que la méthode STL est très sensible au caractère discret de la variabilité des grosses particules, l'estimation de chaque composante nécessite le contrôle de la taille de la fenêtre nécessaire pour l'extraction de la tendance, sinon elle est biaisée. Cette méthode est alors inadaptée pour la prévision des fluctuations des séries des grosses particules.

Un modèle autoregressif ARIMA combiné à une décomposition STL a été utilisé pour prédire les concentrations de formaldéhyde et de particules mesurées dans la maison expérimentale, mais également dans l'espace de bureau où la présence des occupants et leurs activités rendent la saisonnalité moins marquée. Les données d'apprentissages correspondent aux premiers huit jours et la partie test correspond aux quatre derniers jours, sur un total de douze jours. La partie ombrée des prévisions sur les figures est associée à l'intervalle de prévision obtenue par la modélisation ARIMA.

Pour le formaldéhyde, le modèle affiche de très bonnes performances avec un horizon de 2 jours de prévision sur les mesures réalisées dans la maison expérimentale. Il est clairement moins performant sur les mesures réalisées dans l'espace de bureau en février 2015 qui montrent une saisonnalité plus bruitée.

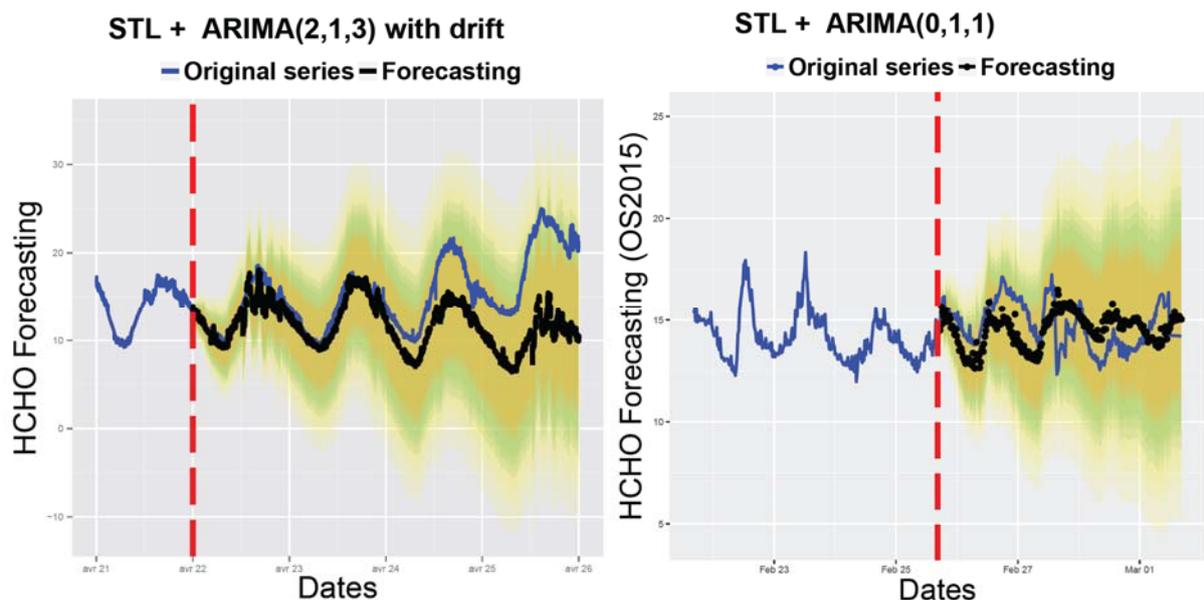


Figure 112 : Prédiction des concentrations du formaldéhyde dans la maison expérimentale en 2010 (à gauche) et dans l'espace paysager en 2015 (à droite) par un modèle STL/ARIMA.

Pour les particules, la performance du modèle varie avec leur taille. Pour les particules fines, la prévision est basée sur une transformation en log de la concentration en nombre $\ln[N]$. L'évolution de la concentration des particules fines avec une saisonnalité très peu marquée n'est pas restituée par le modèle malgré des écarts de prévisions faibles. La saisonnalité plus marquée des plus grosses particules est mieux restituée par le modèle mais les autres variations avec un caractère plus aléatoire ne le sont pas.

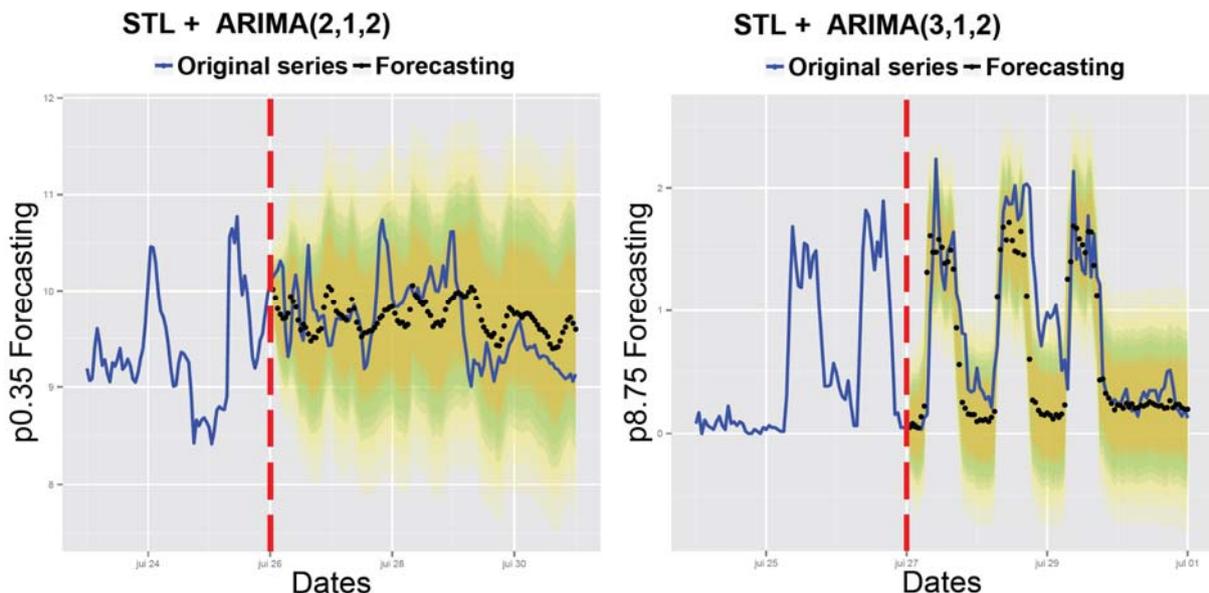


Figure 113 : Prédiction des concentrations horaires de particules de 0,35 μm et 8,75 μm dans l'espace paysager en 2012 par un modèle STL/ARIMA.

4.8.5 Prédiction de la concentration intérieure en formaldéhyde après décomposition spectrale

Un modèle combinant la décomposition par bande spectrale au moyen d'une transformée de Fourier (FFT) avec un modèle non-linéaire des composantes FFT a été testé pour prévoir la concentration du formaldéhyde dans l'espace de bureaux. Une série continue de 14 jours au pas de temps de 1 minute du 24 mai au 7 juin 2013 a été retenue pour tester cette configuration de modèle. Deux modèles non-linéaires ont été comparés : un modèle autoregressif à seuil (TAR) et un autre basé sur la théorie des systèmes dynamiques (chaos). Les données (20 000 minutes) ont été divisées en 3 groupes. 80 % de ces données représentent le jeu d'apprentissage avec lequel les paramètres des modèles sont déterminés. Une première validation du modèle est

réalisée sur 15 % des données permettant de l'optimiser au besoin. Et les 5 % restants servent à déterminer les performances effectives du modèle.

La présence de points de variations abruptes dans l'évolution des concentrations justifie le choix de modèles de prévision non-linéaires et en particulier ceux gérant les changements de régime d'une série temporelle, comme par exemple un modèle SETAR (*Self-Exciting Threshold Auto-Regressive*) développé par Tong et al. (1980 ; 2011). La transition entre les régimes dans les modèles à seuil de type TAR est gouvernée par un mécanisme qui s'effectue à l'aide d'une variable de transition observable, d'un seuil et d'une fonction de transition. Typiquement, deux mécanismes de transition sont envisageables suivant la forme de la fonction de transition. Les modèles à seuils à transition de phase abrupte et les modèles à transition lisse où le passage d'un régime à l'autre se fait moins brutalement. Lorsque la variable de transition sélectionnée est une variable endogène retardée X_{t-d} d'un délai d , le modèle TAR devient un modèle SETAR. A cause du caractère discret de la variable indicatrice, les transitions dans ces modèles sont abruptes et le passage d'un régime à l'autre se fait instantanément au pas de temps près.

Ce modèle a été appliqué dans un premier temps à des données préalablement filtrées par transformée de Fourier (FFT) pour ne conserver que la tendance globale de variation du formaldéhyde. Les résidus de ce premier lissage ont été filtrés à leur tour par FFT et ainsi de suite. Les n composantes FFT ainsi obtenues sont ensuite modélisées chacune par SETAR ou par dynamique du chaos avant que les résultats soit additionnés pour reconstituer la concentration finale prévue.

Dans la dynamique du chaos ou théorie des systèmes dynamiques, disposer de toute l'information du système dynamique déterministe permet de prédire complètement l'évolution de ce système. Ce pouvoir prédictif remarquable peut être préservé pour les variables observées de ce système sous certaines conditions. L'étude d'un système complexe fait entrer en ligne de compte l'ensemble des positions de trajectoire, des vitesses de variations que le système est a priori susceptible d'adopter, qui est appelé espace d'état. Tout système dont les états varient dans le temps est un système dynamique ; pour un tel système, un état est une fonction $s(t)$ du temps t . Dans un système dynamique déterministe, l'état initial $s(0)$ est le seul qui détermine les états futurs $s(t)$; $t > 0$: Dans le cas d'un système dynamique stochastique, une application unique entre les états (futur, présent et passé) peut ne pas exister. Il est donc possible qu'il n'y ait aucune corrélation entre les états ; d'ailleurs un bruit aléatoire pur comme un bruit blanc, l'est par définition.

L'observation directe de tous les états de l'environnement intérieur comme états d'un système dynamique $s(t)$ n'est pas directement accessible. L'attention va donc se focaliser sur les propriétés évolutives de la variable mesurée, tant pour sa prévision que dans sa relation avec son passé. La reconstitution de la dynamique consiste en la détermination de la relation empirique entre les états $s(t)$ et la seule connaissance des variables observées $x(t)$. A partir d'une seule variable, la seule manière d'obtenir de l'information additionnelle est de revenir aux observations précédentes. Cette

reconstitution nécessite le choix approprié de deux paramètres clés : le délai et la dimension de plongement. Le délai est généralement estimé à partir de la fonction d'autocorrélation et la dimension de plongement par la méthode des « Faux Plus Proches Voisins » (FNN, Kennel et al., 1992).

La procédure de prévision à partir de la théorie des systèmes dynamiques suppose que si le système a été observé durant un temps suffisamment long, il existe alors probablement des états du passé aux caractéristiques proches de l'état présent. La prévision consiste alors à appliquer l'évolution observée dans le passé dans des conditions quasi-identiques. L'algorithme de prévision permettant de rechercher les plus proches voisins de l'état présent est généralement la prédiction par constante locale (Local Constant Prediction, Kantz et Schreiber, 2004).

La décomposition spectrale permet d'améliorer les prévisions comparées à l'utilisation des données brutes. Elle permet d'éliminer du bruit nuisible au modèle mais écarte également la possibilité de reproduire des fluctuations très courtes de la concentration selon le niveau de lissage effectué. Elle permet également d'améliorer le temps de calcul et du coup fournit de meilleures performances. En règle générale, la performance du modèle hybride est un compromis entre le niveau de lissage des données réalisé par la décomposition spectrale et la performance du modèle de prévision sur les données lissées.

En effet, le prétraitement des données joue sur la qualité des prévisions, mais surtout elle permet de s'interroger sur l'importance que l'on veut donner aux variations observées et quelles variabilités sont utiles à la prévision. Les données prises en compte avec un pas de temps de 1 min présentes un bruit en partie liée aux fluctuations du signal de l'instrument et donc à l'incertitude de mesure qu'il est inutile de chercher à reproduire. Dans notre cas, nous avons pris en compte des variations de période supérieures à 4,5 heures. Il serait certainement souhaitable de descendre à l'échelle de 1 heure, pour des prévisions de court terme.

Les deux modèles utilisés reproduisent l'aspect non-linéaire de l'évolution de la concentration du formaldéhyde avec des performances légèrement supérieures pour le modèle basé sur la dynamique du chaos. Ces modèles hybrides permettent de prévoir la concentration du formaldéhyde avec de bonnes performances à un horizon de 12 heures.

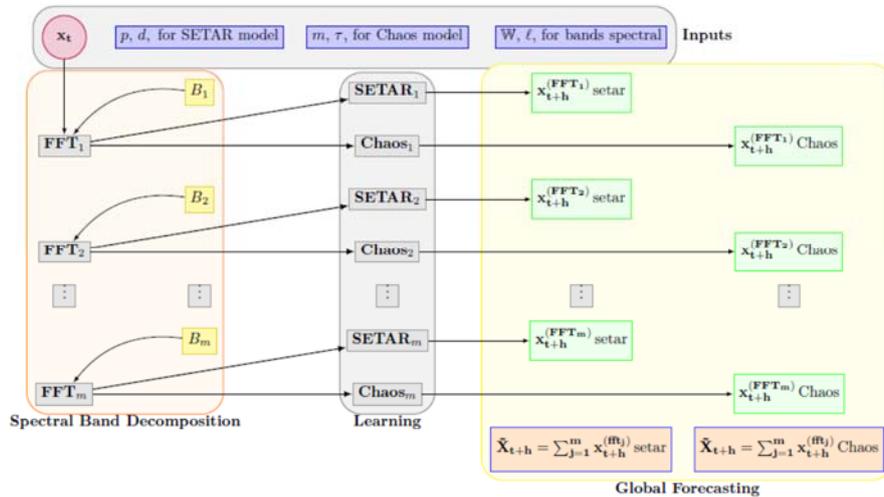


Figure 114 : Processus de prévision par un modèle hybride FFT/TAR ou FFT/Chaos.

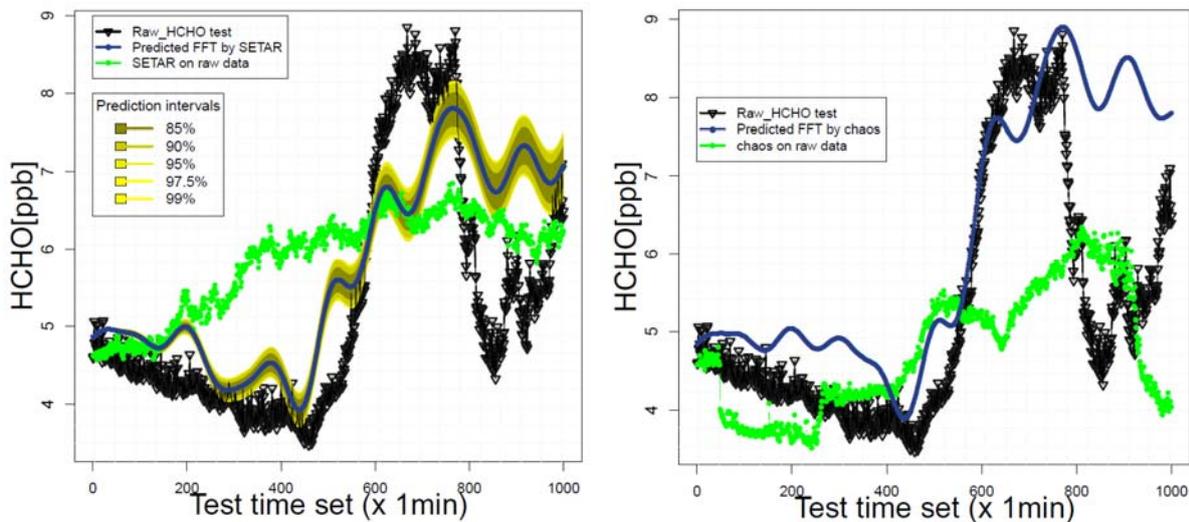


Figure 115 : Prévisions obtenues sur les données tests par FFT/SETAR et FFT/Chaos

En général, la prévision de la première composante FFT1, celle qui extrait la tendance générale de variation de la concentration détermine en grande partie la qualité de la prévision finale. Lorsqu'elle n'est pas performante, les prévisions sur les autres composantes FFT ne permettent pas d'améliorer la prévision finale.

La décomposition spectrale n'est pas toujours nécessaire si les paramètres étudiés présentent une saisonnalité marquée. C'est le cas des particules les plus grosses, pour lesquelles une décomposition spectrale par FFT introduit des oscillations parasites lors des périodes monotones comme les week-ends. A titre de comparaison, la figure suivante montre les résultats de la prévision sur des concentrations de particules

mesurées dans l'espace de bureaux en 2012 par chaos. Les particules fines de $0,9 \mu\text{m}$ nécessitent une décomposition spectrale préalable de leur concentration alors qu'un résultat correct est observé pour les particules de $8,75 \mu\text{m}$ en appliquant directement le modèle sur les données brutes.

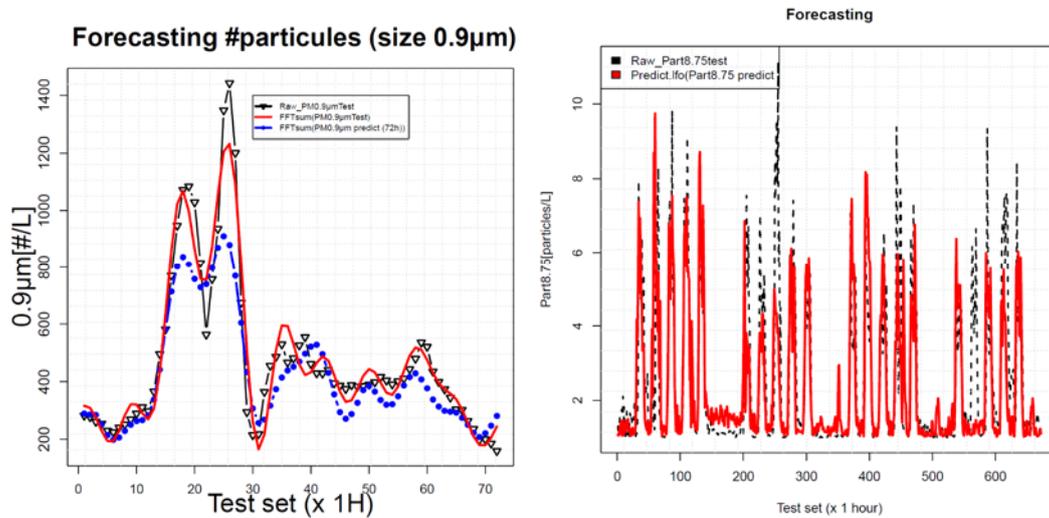


Figure 116 : Prévisions obtenues sur les données tests par FFT/Chaos pour les particules de $0,9 \mu\text{m}$ et directement par chaos pour les particules de $8,75 \mu\text{m}$

Plus de détails sur ces modèles hybrides sont fournies dans une première version d'article qui a été soumis à la revue *Environmetrics* et disponible en annexe. Les travaux préalables à cet article notamment sur le développement du modèle FFT/SETAR ont fait l'objet de deux communications dans des conférences internationales également disponibles en annexe. Enfin, la thèse de Rachid Ouaret (2016) reprend en détail les différents aspects et contraintes de la prévision des paramètres de la qualité de l'air intérieur.

4.8.6 Discussion sur la prévision des concentrations intérieures de polluants

La prévision de l'évolution des paramètres mesurés dans les environnements intérieurs basée sur la dynamique des séries temporelles visait en premier lieu l'identification des spécificités des environnements intérieurs. Dans ces environnements, la présence des occupants et leurs actions sur les ouvrants se traduit par des variations irrégulières plus brutales comparées à ce qui peut être observé dans l'environnement extérieur. Ramener les données mesurées à un pas de temps long permettrait de lisser ces variations brutales et on serait alors amené à traiter des données qui seraient assez voisines de celles enregistrées à l'extérieur. Mais pour

autant, la part de variation relative à la présence des occupants et de leurs activités serait en grande partie gommée. La prise en compte de ces phénomènes se fait avec un pas de temps fin et en fonction de la nature du polluant, la variabilité est généralement plus fluctuante. Des composantes aléatoires viennent alors se greffer au sein de la série en modifiant les structures de régularité. Elles nécessitent donc une modélisation plus complexe par rapport à ce qui est réalisé pour l'air extérieur avec une résolution temporelle de l'ordre de l'heure ou de la journée.

D'autre part, ces modèles de prévisions sont très peu étudiés et appliqués aux environnements intérieurs. Il s'agissait donc dans un premier temps d'examiner les différentes approches possibles pour fournir des prévisions bonnes et utiles aux gestionnaires-décideurs. Pour ces raisons, notre travail s'est appuyé sur les résultats des modèles de prévisions obtenus dans différentes disciplines, tant au niveau théorique que pratique. Par exemple, les prévisions délivrées par un modèle type ARMA convergent rapidement vers la moyenne parce que l'on perd très vite de l'information sur les perturbations aléatoires quand l'horizon de prévision augmente. C'est précisément ce que nous voulions éviter étant donnée la résolution temporelle utilisée dans la plupart de nos séries.

Les travaux menés dans le cadre de ce projet ont permis de poser quelques "jalons", mais les éléments de réponses pour d'autres paramètres de la QAI demanderont des études supplémentaires pour être précisés.

Les modèles d'ajustement appliqués ici pour la QAI ne reposent pas forcément ni sur une construction théorique unique ni sur la théorie physique de l'environnement intérieur. L'extraction des composantes déterministes via une méthode statistique consiste à identifier l'équilibre de long terme (tendance) de la variabilité et la fluctuation oscillatoire (saisonnalité) clairement reconnaissable. Cette distinction ne suffit pas pour l'extraction des différentes composantes, elle nécessite donc des hypothèses supplémentaires.

La manifestation de quelques patterns statistiques dans la concentration de certains polluants relève le fait que la vraie variable dépend fortement des facteurs endogènes et/ou exogènes imprévisibles dont l'influence varie selon des conditions partiellement identifiables. Par conséquent, le vrai modèle doit contenir ces facteurs de manière à refléter au mieux le comportement des fluctuations. Donc, un bon modèle serait en mesure de prévoir cette dynamique sur un horizon donné.

Les modèles à changements de régimes décrivent un ensemble d'états dans lesquelles chaque portion de variabilité se révèle explicative du phénomène environnemental étudié. Si la succession de ces phases de variabilité se manifestent de façon régulière, celles-ci peuvent être attribuées à une forme de saisonnalité. En effet, la succession des états occupations-inoccupations dans un environnement réel serait la cause directe et indirecte d'au moins deux types de variation : un comportement quasi-saisonnier des concentrations du CO₂ et un comportement plus erratique des concentrations du formaldéhyde. Pour ce dernier, la dynamique du

formaldéhyde est liée d'une certaine mesure aux variations de l'état des ouvrants qui sont déterminées par la succession des phases d'occupations et d'inoccupations.

L'aspect délicat de la prévision en règle générale est qu'il est d'une part nécessaire d'avoir pour chaque environnement et paramètre étudié un temps d'observation suffisamment long et que d'autre part la paramétrisation du modèle de prévision va autant dépendre des données observées qui ont servi à l'apprentissage que de l'horizon de prévision souhaité. Un modèle bien paramétré peut être performant à un horizon donné, et ne pas l'être pour des horizons plus courts ou plus longs pour lesquels une autre paramétrisation apporterait de meilleurs résultats. C'est un aspect à prendre en compte pour toute volonté d'automatisation du processus de prévision.

Enfin, la qualité de prévision dépend du "niveau de prédictibilité" des observations. En d'autres termes, la prévision des concentrations de certains polluants est plus "facile" que d'autres. Par exemple, une variabilité qui présente un profil régulier s'avère plus facile à estimer. Par contre, une forte irrégularité de la variabilité nécessite une exploitation plus approfondie dans la recherche des périodicités. En outre, la prévision sur certaines périodes est plus facile que pendant d'autres périodes, car la complexité d'un phénomène est conditionnée par la dynamique de ses différentes composantes. Cette paramétrisation gagnerait en précision avec une meilleure connaissance de la structure des séries temporelles étudiées comme par exemple ses dimensions fractales.

D'autres modèles de type lissage exponentiel et notamment le modèle de Holt-Winters ont été appliqués sur certaines données avec des résultats parfois satisfaisants (bonne performance après 41 heures pour prévoir le formaldéhyde dans la maison expérimentale).

Des modèles introduisant des variables exogènes observées ou prédites comme la température, l'humidité spécifique ou relative ou encore l'ouverture des fenêtres, sont en cours d'évaluation. Il s'agit notamment de modèles exogènes autorégressifs non-linéaires (NARX). Leur fonction non-linéaire est ici assurée par un réseau de neurones.

5 Conclusion

La concentration de particules et de formaldéhyde a été suivie avec un pas de temps fin (entre 1 et 20 minutes) dans un espace de bureau entre 2012 et 2015. Ces mesures ont été systématiquement accompagnées d'enregistrement des paramètres climatiques intérieurs et extérieurs, ainsi que d'un suivi expérimental de la présence des occupants et de l'état des ouvrants.

Ce jeu de données conséquent malgré la présence inéluctable de valeurs manquantes constitue une formidable base de travail à partir de laquelle il est possible d'étudier et de développer des outils permettant de séparer et d'identifier les sources ou encore de prévoir l'évolution des concentrations à venir.

Ces données seront mises à disposition de la communauté scientifique et du public une fois les principaux résultats de cette étude publiés dans un journal à comité de lecture que les utilisateurs pourront citer. La mise à disposition de ces données permettra, nous l'espérons, de voir de nouveaux outils se développer pour la caractérisation de la structure temporelle des paramètres relatifs à la qualité de l'air intérieur et de nouveaux modèles de prévision.

La variabilité de la concentration intérieure en formaldéhyde et en particules a été caractérisée à différentes échelles temporelles. Le formaldéhyde montre des variations importantes au jour le jour liées en partie à des conditions d'ouverture des ouvrants et d'occupation variables aux conséquences également variables, mais également aux fluctuations d'humidité spécifique et de température. La concentration intérieure de particules montre des fluctuations variables selon leur taille. Les plus fines de 0,35 µm suivent les fluctuations observées à l'extérieur pour la même gamme de taille alors que les plus grosses reproduisent le schéma d'occupation de la pièce.

Le profil d'occupation de l'espace de bureaux a été caractérisé. Il est relativement stable d'un jour sur l'autre et d'une semaine sur l'autre. L'occupation varie beaucoup plus d'un mois sur l'autre et d'une année sur l'autre.

L'influence de l'occupation est peu marquée sur les concentrations de formaldéhyde avec des variations en moyenne annuelle de l'ordre de 10 %. Elle modifie par contre la façon dont le formaldéhyde varie au cours de la journée. Son influence est plus importante sur les particules et caractérisée par un ratio O/U moyen (rapport de la concentration en présence sur la concentration en absence d'occupants) variant de 1 pour les particules fines et ultrafines à 4 pour les plus grosses particules. L'occupation modifie également le ratio I/E moyen (rapport de la concentration en présence sur la concentration en absence d'occupants) pour les plus grosses particules d'environ 3 en absence d'occupants à 10 en leur présence. Le fonctionnement de l'imprimante dans l'espace de bureau pendant les périodes d'occupation contribue en grande partie à cet impact.

Les pratiques d'ouvertures des fenêtres varient essentiellement selon le mois de l'année et d'une année sur l'autre. L'ouverture des fenêtres permet de diminuer de façon modérée la concentration intérieure en formaldéhyde en moyenne d'environ 20 % pour une fenêtre ouverte jusqu'à 40 % pour plus de 3 fenêtres ouvertes. Le ratio I/E diminue de 5 (fenêtres fermées) à 3-4 (fenêtres ouvertes). Néanmoins, en 2014 l'ouverture d'une fenêtre est associée en moyenne à une augmentation de la concentration de 10 %. L'influence de l'ouverture des fenêtres n'est donc pas systématique et dépend d'autres facteurs. Elle est en général plus marquée en absence d'occupants que durant les périodes de présence.

L'impact de l'ouverture des fenêtres sur la concentration de particules est également fonction de la présence des occupants. Il peut se décliner de trois façons selon la taille de particules :

- Pour les plus fines particules (0,35 μm), la concentration intérieure reste systématiquement inférieure à la concentration extérieure et en moyenne du même niveau quelle que soit la présence des occupants ou l'ouverture des fenêtres. Le ratio O/F moyen est de l'ordre de 1,2 et le ratio I/E se maintient à 0,5 quelle que soit la présence des occupants. La présence comme l'ouverture des fenêtres viennent finalement très peu perturbés les variations de cette gamme de particules.
- Pour les particules de diamètre intermédiaire (2,5 μm), la concentration intérieure est en moyenne du même ordre en absence ou en présence des occupants pour un même état d'ouverture des fenêtres. De ce fait, l'influence de l'occupant sur la concentration de cette gamme de taille intervient surtout au travers de son action sur les ouvrants et très peu du fait de sa présence.
- L'impact de l'ouverture des fenêtres est moins prononcé pour les plus grosses particules (12,5 μm notamment), sauf lorsque l'ensemble des fenêtres sont ouvertes. L'ouverture des fenêtres ne vient pas diminuer la concentration intérieure de ces particules, mais au contraire l'augmenter. Le ratio O/F atteint en moyenne 30 en absence d'occupants mais reste autour de 2 en leur présence. Pour cette taille de particules, l'occupant influence plus par sa présence que par son action sur les ouvrants.

Un modèle par arbre de classification CART a cherché à prédire l'ouverture des fenêtres. Il permet de prédire à plus de 98 % sur une année les situations où au moins une fenêtre est ouverte à partir des données calendaires et climatiques intérieures et extérieures, et de la présence des occupants et de l'état de la porte d'accès. Les performances du modèle se dégradent lorsqu'il est utilisé pour prédire les situations de fenêtre ouverte d'une autre année. L'association d'évènements qui ont conduit à l'ouverture d'une fenêtre une année n'est pas complètement transposable l'année suivante. Il nécessite encore quelques améliorations pour pouvoir être couplé à un modèle de prévision des concentrations.

Le calcul de l'exposition moyenne sur l'année des occupants de l'espace de bureau a été déterminé par différentes méthodes soit basé sur le capteur de mouvement, soit intégré sur une durée de 8 h, 24 h ou 5 jours. Il montre que l'erreur associée à la non-prise en compte des périodes d'occupation est négligeable dans le calcul de l'exposition annuelle au formaldéhyde. Pour les particules, elle conduit toutefois à une surestimation des particules les plus fines de 10 % à 20 % et par contre une forte sous-estimation des plus grosses particules jusqu'à 70 % et donc par conséquent à une modification notable du profil granulométrique. Cela peut induire des erreurs dans l'interprétation des expositions à certaines substances particulières comme les composés organiques semi-volatils, issues d'analyses de fractions particulières collectées sans tenir compte des périodes d'occupation et des activités qui en dépendent.

A partir des données recueillies, il a été possible d'estimer le taux de renouvellement d'air durant chaque épisode de décroissance de la concentration de CO₂ par une méthode automatisée de calcul. Bien que parcellaire, car ne prenant pas en compte les périodes de présence des occupants correspondant à une production de CO₂, cette démarche a permis d'appréhender la variabilité de ce renouvellement d'air au cours du temps. Alors qu'il est stable autour de 0,4 h⁻¹ avec 40 % de variation en période nocturne, il atteint en moyenne 1,5 h⁻¹ entre 08h00 et 20h00 mais avec des variations plus importantes avec une amplitude pouvant être supérieure à 5. Le renouvellement d'air est en général plus faible lorsque la porte d'accès est ouverte. L'effet de l'ouverture de la porte sur le renouvellement d'air disparaît lorsque les fenêtres sont ouvertes. Il n'est significatif qu'à partir du moment où les fenêtres sont fermées. L'influence de l'ouverture des fenêtres reste par contre systématique quel que soit l'état de la porte.

La variabilité du renouvellement d'air associée à celle de l'occupation et à celle de l'ouverture des fenêtres nous montre que les pratiques des occupants sont au cœur de l'évolution de la concentration des polluants dans l'air intérieur. Pour déterminer cette évolution en situation réelle d'occupation, il est primordial de renseigner la variabilité du renouvellement d'air qui représente la fonction de transfert de matière entre les deux compartiments intérieur et extérieur, et de coupler cette information avec les variations des paramètres climatiques.

Un exercice similaire a été entrepris pour estimer le dépôt des particules. Mais, la méthodologie employée pour ces premiers travaux est à revoir.

L'influence des paramètres climatiques a été abordée au travers du croisement des données observées selon la présence des occupants et l'ouverture des ouvrants au regard de la vitesse et de la direction du vent. Cette dernière est très peu prise en compte dans l'exploitation des données de paramètres intérieurs alors qu'elle peut permettre d'expliquer certaines variations observées. Les concentrations élevées en formaldéhyde sont par exemple associées à des directions du vent particulières qui varient selon l'ouverture des fenêtres. Pour les particules, les directions de vent associées aux concentrations les plus élevées varient en fonction de la taille des

particules. Ces travaux soulignent la nécessité de déconvoluer la contribution des différents paramètres qui parfois évoluent de manière conjointes aux mêmes périodes.

Différentes méthodes de séparation des sources ont donc été utilisées pour extraire les différentes sources de fluctuations et estimer leur contribution. Ces méthodes ont principalement été utilisées sur les données de particules mesurées dans l'espace de bureau, mais également dans d'autres environnements comme la maison expérimentale ou un bureau individuel situés dans le même cadre. Quatre différentes méthodes ont été comparées : l'ACI (analyse par composantes indépendantes), la NMF (*non-negative matrix factorization*), la PMF (*positive matrix factorization*) et l'ACP (analyse par composantes principales). Les composantes ou facteurs identifiées représentent autant de sources de fluctuations séparées. Les composantes peuvent être interprétées en termes de contributions variables selon le diamètre des particules, mais également en termes de fluctuations ou d'occurrence au fil du temps.

Ces premiers travaux montrent la similitude des différentes méthodes pour séparer à l'aveugle les sources contributives au niveau de particules observé dans un environnement intérieur. L'ACP classique n'offre pas le même confort d'interprétation que les autres méthodes.

Pour autant, séparer n'est pas identifier. Sans informations extérieures, les facteurs extraits restent difficilement identifiables autrement que par l'observation de fluctuations caractéristiques d'un processus particulier. Après leur extraction, une seconde phase d'exploitation doit donc être menée pour rechercher les associations avec d'autres phénomènes observés.

Un modèle de mélange gaussien a été appliqué à la distribution des concentrations de formaldéhyde mesurées dans la maison expérimentale pour séparer les sources. La méthode permet de distinguer les différentes densités de distributions qui contribuent à la distribution globale observée. Elle distingue deux processus, l'un centré sur un mode à 11 ppb et l'autre centré sur 16 ppb. Les hypothèses sur la forme des distributions et surtout le manque d'informations sur les fluctuations temporelles rendent difficiles l'interprétation des processus identifiés.

Pour prédire les concentrations observées de formaldéhyde, une première approche a été utilisée un modèle de régression linéaire simple basées sur les autres données observées en particulier les données calendaires, climatiques intérieures et extérieures et celles relatives à l'occupation et aux ouvrants. Tout comme le modèle prédictif de l'ouverture, ce modèle restitue des prédictions convenables pour l'année en cours, mais insuffisants pour restituer les variations observées les autres années.

Pour faire de la prévision, ce type de modèle n'est pas optimal car il nécessite de disposer des mesures et surtout des prévisions de chacun des paramètres entrant dans le modèle pour permettre de prédire à chaque instant futur la concentration du formaldéhyde. Les prévisions des paramètres climatiques extérieures existent et pourraient être utilisées. Ceux de l'environnement intérieur pourraient également être

estimés à partir des prévisions extérieures. Mais la présence, les activités et l'état des ouvrants nécessitent le développement de modèles de prévision spécifiques.

A des fins de prévision de l'évolution de la concentration, il est dès lors plus intéressant de se pencher sur des modèles ne requérant qu'un minimum de connaissances, à savoir un historique des mesures de concentration enregistrées au préalable sans introduire d'informations complémentaires (par exemple des données climatiques ou d'autres mesures de polluants). Ce type de prévision va chercher à exploiter les propriétés intrinsèques à la dynamique des séries temporelles.

Un modèle de prévision hybride a été développé en couplant un processus de décomposition spectrale de la concentration mesurée dans le temps par plusieurs FFT successives avec un modèle de prévision autorégressif à seuil (FFT/SETAR) ou un modèle basé sur la théorie des systèmes dynamiques (FFT/Chaos). L'horizon de prévision obtenu dans ces conditions est d'une dizaine d'heures.

Lorsque la saisonnalité de la concentration est suffisamment marquée, une décomposition saisonnière associée à un modèle autorégressif simple (STL/ARIMA) peut suffire et conduire selon le niveau de bruit de la saisonnalité à des horizons de prévision de plusieurs dizaines d'heures.

Une plus grande connaissance de la structure temporelle de l'évolution des paramètres dans les environnements intérieurs est nécessaire pour optimiser la paramétrisation de ces modèles.

L'utilisation de variable exogène observée ou prédite comme l'ouverture des fenêtres dans des modèles autorégressifs non-linéaires de type NARX est actuellement à l'étude.

Bien que ce projet utilise des technologies parfois délicates à mettre en œuvre, ce qui se ressent notamment lors de l'étape fatidique de validation des données, il montre l'opportunité offerte par l'enregistrement conjoint de données de diverses origines pour mieux comprendre la dynamique de l'évolution des polluants dans des environnements intérieurs réels en présence des occupants.

Il serait nécessaire de poursuivre cette même approche d'observation sur le long terme dans d'autres environnements intérieurs. Chacun ayant sa spécificité, les résultats obtenus dans l'un ne sont pas complètement transposables dans un autre où l'occupation et les pratiques d'ouverture des fenêtres diffèrent.

Un aspect nécessiterait d'être développé pour approfondir ces analyses. Il s'agit de la possibilité de suivre en temps réel les propriétés de surfaces des matériaux notamment les variations des émissions de composés organiques volatils dont le formaldéhyde. Cela permettrait de mettre plus en évidence les interactions matériaux-environnement très peu caractérisées dans les environnements réels.

Le bâtiment numérique, la multiplication des objets connectés et surtout les nouvelles technologies de capteurs vont permettre par l'observation de croiser des données d'origines variées dans leur dimension temporelle et ainsi de mieux comprendre le processus à l'origine des fluctuations observées, de prévoir et d'anticiper les situations à risques, d'utiliser cette connaissance pour optimiser leur gestion en adaptant les pratiques d'ouverture des ouvrants extérieurs et intérieurs, en modifiant les scénarios de ventilation selon ces pratiques et la présence des occupants ou encore en temporisant certaines activités en fonction des conditions extérieures.

Bien sûr, cela nécessitera d'avoir foi dans les données observées et donc qu'elles soient un minimum validées pour permettre leur exploitation. Le gestionnaire du bâtiment le plus à même d'être récepteur de ces données, si ce n'est l'occupant lui-même, aura-t-il l'expertise suffisante pour valider ces données ? Devra-t-il passer par un tiers ? La validation est sans doute le principal frein aujourd'hui, en dehors de la disponibilité des technologies à grande échelle, et certainement la protection des libertés (car toute mesure reste intrusive), au déploiement de ces réseaux d'observations dans le bâtiment.

Notre étude a finalement cherché à montrer avec un minimum de paramètres mesurés comment traiter cette masse d'informations collectées au fil du temps et avec quels outils disponibles, comment la qualité de l'air dépend d'autres facteurs tous rattachés au temps et comment cette structure temporelle des données peut nous aider à mieux comprendre la QAI pour mieux la prédire. L'homme y joue un rôle central en perturbant les relations existantes entre les différents paramètres. Mieux comprendre son comportement, c'est mieux comprendre notre environnement et la manière dont il évolue. Nous en avons dévoilé finalement qu'une petite partie dans le cadre de ce rapport tant les possibilités d'exploitation de ces données restent nombreuses en les croisant ou non avec d'autres données disponibles.

Nous pensons également avoir montré l'intérêt de suivre de manière expérimentale les événements propres au comportement des occupants de leur simple présence à l'ouverture des ouvrants ou à certaines de leurs activités comme l'utilisation d'une imprimante par la mesure de ses impulsions électriques.

Pour finir, il existe encore trop peu d'études qui ont exploré la dimension temporelle de la qualité de l'air dans les environnements intérieurs. Nous en sommes encore à la photographie ou au court-métrage, mais aucun film ne nous permet d'appréhender vraiment l'environnement dans lequel nous passons le plus de temps et comment il interagit au fil du temps avec l'extérieur et comment nous interagissons au jour le jour avec lui.

6 Valorisations

6.1 Articles :

IONESCU A. (2013), *Prévision statistique des concentrations de particules dans l'air : tour d'horizon des principaux outils mathématiques*, **Pollution atmosphérique**, n° 217, [En ligne], mis à jour le : 16/04/2013, URL : <http://odel.irevues.inist.fr/pollution-atmospherique/index.php?id=873>

OUARET R., IONESCU A., PETREHUS V., CANDAU Y., RAMALHO O. (2016), *Spectral Band Decomposition Coupled with Nonlinear Models for Indoor Formaldehyde Concentration Forecasting*, **Environmetrics**, soumis.

6.2 Communications :

OUARET R., IONESCU A., RAMALHO O., PETREHUS V., CANDAU Y. (2014), *Caractérisation et identification de la variabilité temporelle des sources de particules dans un environnement intérieur : approche statistique*, **29^{ème} Congrès Français sur les Aérosols**, Paris, 22-23 janvier 2014, 6 p.

OUARET R., IONESCU A., RAMALHO O., PETREHUS V., CANDAU Y. (2014), *Identification des sources de variabilité des niveaux de polluants par analyse en composantes indépendantes et par estimation de mélange des données de l'air intérieur*, **Journées Interdisciplinaires de la Qualité de l'Air**, Villeneuve d'Ascq, 10-11 février 2014, Paris, 18 p.

OUARET R., IONESCU A., RAMALHO O., CANDAU Y., PETREHUS V. (2014), *Forecasting indoor pollutants concentrations using Fast Fourier Transform (FFT) and Regime Switching Models*, **ITISE 2014**, International work-conference on Time-Series analysis, June 25-27 2014, Granada (Spain), Vol. 1, 52-63.

OUARET R., IONESCU A., RAMALHO O., CANDAU Y., PETREHUS V., LABAT L. (2014), *Modelling the time fluctuation of indoor air formaldehyde concentrations: variability structure identification and forecasting using non-linear models*, Proc. **Indoor Air 2014**, July 7-12 2014, Hong Kong, Vol. V, 321-328.

OUARET R., IONESCU A., RAMALHO O., CANDAU Y., GEHIN E., PETREHUS V. (2014), *Analysis of the temporal variability of indoor particulate matter concentrations using Blind Source Separation methods: a comparative study*, **International Aerosol Conference 2014**, Aug. 28 – Sep. 2, 2014, Busan, South Korea, poster.

OUARET R., IONESCU A., PETREHUS V., CANDAU Y., RAMALHO O. (2016), *Forecasting particulate matter concentrations in an indoor environment*, **22nd European Aerosol Conference**, Tours (France), September 4th - 9th 2016, Poster. (accepted)

OUARET R., IONESCU A., PETREHUS V., CANDAU Y., RAMALHO O. (2016), *Particulate matter variability sources in an open-plan office: comparison of two monitoring campaigns*, **22nd European Aerosol Conference**, Tours (France), September 4th - 9th 2016, Poster. (accepted)

7 Références

Akaike H (1974), A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6):716–723.

Box G, Jenkins GM (1970), Time series analysis: forecasting and control. Holden-Day.

Breiman L, Friedman JH, Olshen RA, Stone CI (1984). Classification and regression trees. Belmont, Calif.: Wadsworth.

Carslaw D (2015), The openair manual – open-source tools for analysing air pollution data. Manual for version 1-1.4, King's College London.

Carslaw DC, Ropkins K (2012), openair – an R package for air quality data analysis, *Environmental Modelling and Software*, 27-28, 52-61.

Chen C, Zhao B (2011). Review of relationship between indoor and outdoor particles: I/O ratio, infiltration factor and penetration factor. *Atmospheric Environment* 45(2): 275-288

Cleveland WS (1979), Robust locally weighted regression and smoothing scatterplots. *Journal of the American statistical association*, 74(368):829–836.

Cleveland RB, Cleveland WS, McRae JE, Terpenning I (1990), STL: A seasonal-trend decomposition procedure based on loess. *Journal of Official Statistics*, 6(1):3–73, 1990.

Coman A., Ionescu A., Candau Y. (2008), Hourly ozone prediction for a 24 hour horizon using neural networks, *Env. Model. Soft.* 23 (2008), 1407-1421.

Gao NP, Niu JL (2007). Modeling particle dispersion and deposition in indoor environments. *Atmospheric Environment* 41(18): 3862-3876.

Halios CH, Helmis CG (2007). On the estimation of characteristic indoor air quality parameters using analytical and numerical methods. *Science of The Total Environment*, 381(1–3): 222-232.

Hegger R, Kantz H, Schreiber T (1999). Practical implementation of nonlinear time series methods: The tisean package. *Chaos: An Interdisciplinary Journal of Nonlinear Science* 9(2): 413–435.

Kantz H, Schreiber T (2004), Non-linear time series analysis, Cambridge University Press.

Kennel M, Brown R, Abarbanel H (1992). Determining embedding dimension for phase-space reconstruction using a geometrical construction. *Physical Review A*, 45 (6): 3403–3411.

Lai K, Nazaroff WW (2000). Modeling indoor particle deposition from turbulent flow onto smooth surfaces. *Journal of aerosol science* 31(4): 463-476

Lebart L, Piron M, Morineau A (2006), *Statistique Exploratoire Multidimensionnelle*, Dunod, 2006 (4ème édition, refondue), 480p.

Ouaret R. (2016), *Déterminants et prévision des fluctuations de la concentration en polluants dans un environnement intérieur*, Thèse de l'Université Paris-Est Créteil.

Parzen E (1975), Multiple time series: Determining the order of approximating autoregressive schemes. Rapport technique, DTIC Document.

Schwarz G (1978), Estimating the dimension of a model. *The annals of statistics*, 6(2):461–464, 1978.

Tong H. & Lim K S (1980) Threshold Autoregression, Limit Cycles and Cyclical Data, *Journal of the Royal Statistical Society, Series B*, 42, 245-292.

Tong H (2011). Threshold models in time series analysis –30 years on, *Statistics & its Interface*, 4, 107-136.

Zarli A., Pascual E., Cheung D. (2010), Information and communication technology for intelligent and integrated control in buildings: current developments and future research, *Proc. of the CIB W78 2010: 27th International Conference –Cairo, Egypt*, 16-18 November, 9 p.

Zouba N., Bremond F., Thonnat M., Anfosso A., Pascual E., Mallea P., Mailland V., Guérin O. (2009), A computer system to monitor older adults at home: Preliminary results, *Gerontechnology*, 8(3), 129-139.



Le futur en construction

Établissement public au service de l'innovation dans le bâtiment, le CSTB, Centre Scientifique et Technique du Bâtiment, exerce quatre activités clés : la recherche et expertise, l'évaluation, la certification et la diffusion des connaissances, organisées pour répondre aux enjeux de la transition énergétique dans le monde de la construction. Son champ de compétence couvre les produits de construction, les bâtiments et leur intégration dans les quartiers et les villes.

Avec plus de 900 collaborateurs, ses filiales et ses réseaux de partenaires nationaux, européens et internationaux, le groupe CSTB est au service de l'ensemble des parties prenantes de la construction pour faire progresser la qualité et la sécurité des bâtiments.

Centre Scientifique et Technique du Bâtiment

Siège social > 84 avenue Jean Jaurès – Champs-sur-Marne – 77447 Marne-la-Vallée cedex 2

Tél. : +33 (0)1 64 68 82 82 – www.cstb.fr

MARNE-LA-VALLÉE / PARIS / GRENOBLE / NANTES / SOPHIA ANTIPOLIS