

# Retour d'expérience sur l'essai inter-laboratoire Norman de 2019 couplant l'échantillonnage passif à l'analyse non ciblée

APPLICATION A DES ECHANTILLONS D'EAUX DE RIVIERE ET D'EAUX D'UNE FILIERE DE POTABILISATION

S. Merel, T. Ducrocq, C. Miège

Janvier 2023

Document final

Avec le soutien de



## Contexte de programmation et de réalisation

---

Ce rapport a été réalisé dans le cadre du programme scientifique et technique AQUAREF pour l'année 2022, au titre de l'action FG1.1a « Analyse non ciblée (NTS) : Inventorier et fiabiliser les méthodes d'acquisition et de traitement des données pour la caractérisation chimique non-ciblée de matrices intégratrices (EIP, sédiment, biote) des milieux aquatiques » du thème FG.

Auteur (s) :

*Sylvain Merel*  
INRAE  
[sylvain.merel@inrae.fr](mailto:sylvain.merel@inrae.fr)

*Tom Ducrocq*  
INRAE  
[tom.ducrocq@inrae.fr](mailto:tom.ducrocq@inrae.fr)

*Cécile Miège*  
INRAE  
[Cecile.miège@inrae.fr](mailto:Cecile.miège@inrae.fr)

---

Vérification du document :

*Béatrice Lalère*  
LNE  
[beatrice.lalere@lne.fr](mailto:beatrice.lalere@lne.fr)

*Anne Togola*  
BRGM  
[a.togola@brgm.fr](mailto:a.togola@brgm.fr)

## Les correspondants

---

OFB : P-F Staub, ONEMA, [pierre-francois.staub@ofb.gouv.fr](mailto:pierre-francois.staub@ofb.gouv.fr)

INRAE : Marina Coquery, [marina.coquery@inrae.fr](mailto:marina.coquery@inrae.fr)

Référence du document : S. Merel, T. Ducrocq, C. Miège - Retour d'expérience sur l'essai inter-laboratoire Norman de 2019 couplant l'échantillonnage passif à l'analyse non ciblée - Rapport AQUAREF 2022 - 37 p.

Droits d'usage :	<i>Accès libre</i>
Couverture géographique :	<i>International</i>
Niveau géographique :	<i>National</i>
Niveau de lecture :	<i>Professionnels, experts</i>
Nature de la ressource :	<i>Document</i>

<b>1. INTRODUCTION.....</b>	<b>6</b>
<b>2. DESCRIPTION DE L'ESSAI INTER-LABORATOIRE.....</b>	<b>7</b>
<b>3. DEFINITION DES OBJECTIFS.....</b>	<b>8</b>
<b>4. RESULTATS DE L'ESSAI-INTERLABORATOIRE .....</b>	<b>8</b>
4.1 Comparaison qualitative des résultats HRMS de l'EIL.....	8
4.2 Comparaison des logiciels libres MZmine et workflow4metabolomics (W4M) .....	12
4.3 Comparaison des empreintes chimiques .....	14
4.4 Effet de la durée d'exposition des échantillonneurs passifs .....	16
<b>5. RECOMMANDATIONS POUR L'ORGANISATION DE FUTURS EIL ET PISTES     POUR DE FUTURS EXPLOITATIONS DES DONNEES EXISTANTES.....</b>	<b>19</b>
<b>6. REFERENCES BIBLIOGRAPHIQUES .....</b>	<b>21</b>

Liste des annexes :

**Annexe 1 : Description détaillée de l'essai inter-laboratoire**

**Annexe 2 : Poster présenté lors de la conférence internationale sur le NTS du 4 au 7 octobre 2021 à Erding (Allemagne).**

**Annexe 3 : Détails des méthodes analytiques utilisées par les quatre laboratoires considérés dans ce rapport**

## RESUME

Afin de caractériser de manière optimale la contamination du milieu aquatique, il faut réunir à la fois une stratégie d'échantillonnage et une approche analytique permettant la détection des contaminants la plus large possible. Des prélèvements ponctuels ne permettent qu'un état des lieux à un moment précis, mais le choix d'un échantillonnage passif reposant sur l'adsorption des micropolluants par un support déterminé permet d'estimer la contamination moyenne durant la période d'exposition du support. Après désorption des contaminants, leur détection est généralement assurée par chromatographie couplée à la spectrométrie de masse. L'approche traditionnelle ciblée n'autorise que la détection de contaminants présélectionnés. Pour détecter le plus grand nombre de contaminant possible il faut alors recourir à une approche non-ciblée basée sur la spectrométrie de masse à haute résolution (HRMS). Cependant, le couplage de l'échantillonnage passif et de l'analyse HRMS n'est pas une approche commune et un essai inter-laboratoire (EIL) a donc été organisé et réalisé en 2019 dans le cadre du réseau Norman. Dans le cadre de cet EIL, 21 laboratoires devaient utiliser une méthode LC-HRMS prédéfinie pour analyser les mêmes extraits de POCIS exposés pendant 24 et 48 heures en entrée et en sortie d'une usine de potabilisation d'eau. Ce rapport présente donc un premier retour d'expérience sur les résultats de cet EIL en examinant plus particulièrement les performances des laboratoires pour la recherche de molécules spécifiques dopées dans les échantillons, les logiciels libres pour le traitement des données HRMS, la comparabilité des empreintes chimiques générées par les différents laboratoires, puis l'effet de la durée d'exposition de l'échantillonneur passif sur la nature des contaminants détectés. En premier lieu, on remarque que la méthode analytique imposée par les organisateurs de l'EIL n'a pas été rigoureusement suivie par les 21 laboratoires participants. Les variations observées peuvent concerner le volume injecté sur la colonne tout comme le gradient chromatographique (composition, débit). Une telle disparité des conditions analytiques crée une première source de difficulté pour la comparaison des résultats des différents laboratoires. Néanmoins, lorsque l'on ne considère plus un ensemble de molécules inconnues (caractérisé par une masse et un temps de rétention) mais des molécules connues (caractérisée par une masse, des fragments et éventuellement un temps de rétention), la comparaison reste tout à fait possible. On observe alors que les 6 molécules dopées dans les échantillons sont effectivement bien détectées par les différents laboratoires avec une fréquence de détection comprise entre 90 et 95%. De plus on remarque que l'erreur sur la masse est généralement comprise entre 0 et 3 ppm, conformément au niveau de tolérance préconisée dans la littérature pour l'identification d'un composé lors d'une analyse par HRMS. Cependant, on remarque que, pour certains cas plus minoritaires, la déviation en masse peut atteindre 10 ppm. Une analyse suspectée avec une tolérance usuelle de 5 ppm aboutirait alors à un certain nombre de « faux négatifs » tandis qu'adapter la tolérance à 10 ppm aboutirait à augmenter le nombre de « faux positifs ». Ensuite, afin d'examiner les données HRMS avec une approche totalement non-ciblée, ce rapport compare deux logiciels en accès libre : MZmine et workflow4metabolomics (W4M). D'une manière générale, MZmine révèle un plus grand nombre de « features » (signal associée à une molécule) que W4M. Cependant, cela n'est vraisemblablement pas lié à la performance du logiciel mais à la différence des critères utilisés pour le « peak picking » (recherche automatisée des signaux). En raison des différences dans les méthodes d'acquisition évoquées auparavant, la comparaison directe des empreintes chimiques générées par les laboratoires participants s'est avérée très complexe et nous conduit à des différences. Néanmoins il en ressort une observation assez inattendue sur la durée d'exposition des échantillonneurs passifs. En effet, sur l'ensemble des composés détectés, environ un tiers est détecté uniquement dans les échantillonneurs exposés 2 jours à l'eau de rivière, un tiers est détecté uniquement dans les échantillonneurs exposés 4 jours, et un tiers est détectés dans les deux échantillonneurs. Un examen plus détaillé des « features » suggère qu'une augmentation du temps d'exposition des échantillonneurs favoriserait la détection des composés de masse plus importante et plus lipophiles.

## Mots clés (thématique et géographique) :

Echantillonnage passif, spectrométrie de masse à haute résolution (HRMS), Analyse non-ciblée, EIL Norman, POCIS

*FEEDBACK ON THE 2019 NORMAN COLLABORATIVE TRIAL ON PASSIVE SAMPLING COUPLED TO NON-TARGETED ANALYSIS*

Sylvain Merel, Tom Ducrocq, Cécile Miège

**ABSTRACT**

The optimal characterization of water contamination requires the combination of sampling strategy with an analytical approach that allow the broadest possible detection of contaminants. While grab samples allow a survey of water contamination at a specific time point, passive samplers that rely on the sorption of contaminants on a selected substrate allow a survey of the average water contamination during a defined exposure period. After extracting the contaminants from the substrate, their detection is usually achieved by coupling chromatography to mass spectrometry. The common targeted approach only detects a limited number of preselected contaminants. In order to obtain a comprehensive detection of contaminants it is necessary to apply a non-targeted analysis through high resolution mass spectrometry (HRMS). As coupling passive sampling to non-targeted analysis is not a common approach, a collaborative trial was organized in 2019 through the Norman network. The current report provides a first feedback on this collaborative trial considering the performance of the laboratories when searching for spiked molecules, examining two open source software to process HRMS data, and evaluating the impact of passive samplers' exposure time (2 days vs. 4 days) on the contaminants detected.

The analytical method imposed by the organizers of the collaborative trial was rather poorly followed by the 21 participating laboratories. The differences observed were related to the sample volume injected on the analytical apparatus as well as the chromatographic gradient itself (composition, flow...). Such differences in analytical conditions are a major issue when comparing results. Nonetheless, comparing results remains perfectly possible when considering known molecules. In fact, all the 6 molecules spiked into the samples were properly detected by the laboratories involved with a detection frequency ranging from 90% to 95%. We also observe that the mass accuracy usually remains in the range 0-3 ppm, which is consistent with the values recommended in the literature. However, in some cases, mass accuracy was as high as 10 ppm. Therefore, performing a suspect screening with a usual 5 ppm tolerance might result in "false negative" while increasing the tolerance to 10 ppm might increase the possibility of "false positive" to occur. Afterwards, in order to examine HRMS data with a non-target approach, two open access software were compared: MZmine and workflow4metabolomics (W4M). Overall, MZmine tends to reveal a greater number of features than W4M. However, this is not likely due to the intrinsic performance of the software but rather to the different criteria used to perform the "peak picking". In practice, due to the previously mentioned differences in analytical conditions, comparing the chemical fingerprints generated by different laboratories was rather complex and unsuccessful. Nonetheless, an unexpected observation was made concerning the impact of the exposure time of the passive sampler and the contaminants detected. In fact, when considering all the compounds detected in passive samplers exposed to river water, approximately a third of them were detected only in samplers exposed for 2 days, another third were detected only in samplers exposed for 4 days and another third were detected in both. A deeper examination of the features suggests that increasing the exposure time favors the detection of larger and more lipophilic compounds.

**Key words** (thematic and geographical area):

Passive sampling, High resolution mass spectrometry (HRMS), Non-targeted analysis, Norman collaborative trial, POCIS

## 1. Introduction

Le nombre de substances synthétisées et mises sur le marché pour satisfaire aux besoins de notre société n'a cessé d'augmenter aux cours des dernières décennies. Aux cours de leur cycle de vie, ces substances peuvent atteindre différents compartiments environnementaux (eau, air, sol, biote...) et devenir des contaminants.

Afin de connaître l'impact de ces contaminants sur l'environnement et la santé publique, il est nécessaire de pouvoir les détecter et de les quantifier de manière spécifique et fiable. Pour ce faire, l'outil le plus couramment employé est la spectrométrie de masse couplée à une technique de séparation chromatographique préalable. Dans de nombreux laboratoires, l'approche classique par spectrométrie de masse en tandem selon une approche ciblée est appliquée en routine. Cependant, cette approche n'autorise que la détection d'un nombre limité de contaminants organiques sélectionnés au préalable et pour lesquels un étalon analytique est disponible au laboratoire.

Afin de garantir une détection plus exhaustive des contaminants organiques, il faut recourir à une approche d'analyse non-ciblée par spectrométrie de masse à haute résolution (HRMS). Cette approche permet en effet de détecter l'ensemble des composés organiques présents dans un échantillon à la condition qu'ils puissent être ionisés par le spectromètre de masse et que leur concentration soit suffisante pour que le signal associé soit observable. L'analyse HRMS génère donc un large volume de données où chaque composé détecté est alors considéré comme un composé inconnu caractérisé par un ratio masse/charge ( $m/z$ ), une abondance, un temps de rétention (si couplage avec une séparation chromatographique préalable) et des ions fragments (si leur acquisition est permise par la méthode HRMS appliquée). L'identification des composés se fait a posteriori selon diverses approches possibles. Alors que l'application de la HRMS devient de plus en plus courante en raison de la sensibilité croissante des instruments pour un investissement financier plus abordable, de nombreux laboratoires développent leur propre méthode pour le traitement des données. A l'heure actuelle il n'y a pas de consensus au sujet des différentes étapes (workflow) à appliquer pour le traitement des données ni au sujet des outils disponibles (quelques logiciels fournisseurs et plusieurs autres en libre accès). Dans ce contexte, et afin que l'analyse non-ciblée démontre pleinement sa complémentarité avec l'analyse ciblée traditionnelle, le défi actuel est de définir comment permettre que les résultats obtenus en HRMS soient comparables d'un laboratoire à un autre, malgré les différences dans l'acquisition (type d'instrument et de méthode analytique) et le traitement (type de workflow et software utilisé) de la donnée.

Ensuite, l'analyse des contaminants organiques dans l'environnement repose également sur un échantillonnage in situ adapté. Alors que l'échantillonnage ponctuel permet d'établir un état de contamination à un instant précis, il ne permet pas de déterminer des concentrations représentatives de périodes plus longues des contaminants organiques dans un cours d'eau par exemple. Alors qu'une stratégie pour pallier cette lacune consiste à faire des prélèvements multiples ou en continu, et asservi au débit ou au temps pendant une période allant généralement jusqu'à 24 heures, des dispositifs d'échantillonnage passif sont également développés. Ces échantillonneurs passifs sont exposés dans le milieu et permettent de capter les contaminants sur une période plus longue,

pouvant aller de plusieurs jours à plusieurs semaines. Ils permettent de concentrer les contaminants et d'abaisser leurs limites de détection. Ils permettent aussi de recalculer des concentrations dans l'eau moyennées sur leur durée d'exposition. Cependant, des interrogations existent encore, en particulier sur la durée d'exposition optimale pour capter et détecter le plus grand nombre et la plus grande diversité de molécules par analyse non ciblée. Dans ce contexte, un essai inter-laboratoire (EIL) a été réalisé en 2019 afin d'examiner la possibilité de comparer les résultats d'analyses HRMS générés par différents laboratoires, en prenant pour objet d'analyse, des extraits d'échantillonneurs passifs. L'objectif de ce rapport est donc de proposer un retour d'expérience sur cette EIL, incluant les premiers résultats portant sur la détection de contaminants spécifiques, une étude additionnelle examinant l'empreinte chimique HRMS dans sa globalité, ainsi que des recommandations pour l'organisation de futurs EIL.

## 2. Description de l'essai inter-laboratoire

Dans le cadre du réseau européen Norman (<https://www.norman-network.net/>), un essai inter-laboratoire (EIL) a été préparé et organisé en 2019 (description détaillée en annexe 1) par les organismes INRAE (Institut national français de recherche pour l'agriculture, l'alimentation et l'environnement), NIVA (Institut Norvégien pour la recherche sur l'eau, Norsk institutt for vannforskning) et RECETOX (centre de recherche tchèque sur les composés toxiques dans l'environnement). Cet EIL, auquel 21 laboratoires de plusieurs pays européens ont participé, couple d'une part la collecte d'échantillons d'eau par un dispositif d'échantillonnage passif et d'autre part leur analyse par spectrométrie de masse à haute résolution.

Des échantillonneurs passifs (disques Horizon Atlantic HLB) ont été déployés en entrée (rivière) et en sortie d'une usine de production d'eau potable située en république tchèque. Un groupe d'échantillonneurs a été exposé durant 2 jours tandis qu'un second groupe d'échantillonneurs a été exposé durant 4 jours. Afin de maximiser l'accumulation des contaminants, un système d'échantillonnage passif dynamique a été utilisé. Ce système consiste à faire passer, par pompage, un large volume d'eau dans la cellule d'exposition des échantillonneurs passifs. De plus, un groupe d'échantillonneurs non-exposés a également été utilisé afin de générer un échantillon blanc. Pour chaque groupe d'échantillonneurs passifs, les extraits ont été mélangés afin de constituer une solution homogène, par la suite aliquotée et envoyée aux laboratoires participants. Chaque laboratoire a donc reçu un total de 6 vials : un extrait d'échantillonneur passif non exposés, deux extraits d'échantillonneurs passif exposés respectivement 2 et 4 jours dans la rivière, deux extraits d'échantillonneurs passif exposés respectivement 2 et 4 jours dans l'eau potable, et un vial contenant un mélange d'étalons analytiques destiné à établir un index de temps de rétention (RTI). Chaque vial était également dopé avec un mélange comprenant des isotopes stables de 6 composés (caféine-<sup>13</sup>C<sub>3</sub>, carbamazépine-D<sub>10</sub>, cotinine-D<sub>3</sub>, diuron-D<sub>6</sub>, nicotine-D<sub>4</sub>, et simazine-D<sub>10</sub>) afin de vérifier la performance des différents laboratoires participants. Les participants à l'EIL ont alors été invités à analyser chaque vial



par spectrométrie de masse à haute résolution avec une méthode analytique (chromatographie liquide et type ionisation) prédéfinie par les organisateurs de l'EIL (annexe 1), puis avec leur propre méthode d'analyse non-ciblée. Les laboratoires avaient ensuite libre choix de la méthode pour le traitement des données et l'identification des composés par comparaison avec des bases de données spectrales (analyse suspectée).

### 3. Définition des objectifs

En se basant sur l'EIL décrit auparavant, dans ce rapport nous traiterons quatre objectifs particuliers :

- 1) Comparer de manière qualitative les résultats HRMS fournis par les 21 laboratoires participant à l'EIL en examinant deux paramètres : la fréquence de détection des 6 composés dopés et l'erreur sur la masse qui y est associée
- 2) Comparer différents logiciels en accès libre pour le traitement des données HRMS afin de déterminer s'ils influencent les résultats.
- 3) Comparer les empreintes chimiques générées par des laboratoires différents pour un même échantillon afin d'étudier la similitude des données HRMS qui pourrait être influencée par plusieurs facteurs (type d'instruments, performance de l'instrument, contamination potentielle par l'analyste ou l'environnement du laboratoire...).
- 4) Examiner l'impact de la durée d'exposition d'un échantillonneur passif sur le nombre et la nature des composés observés (en considérant néanmoins qu'il est question d'un dispositif dynamique), afin de déterminer la durée optimum d'exposition pour l'évaluation des contaminations chimiques non pré-ciblées.

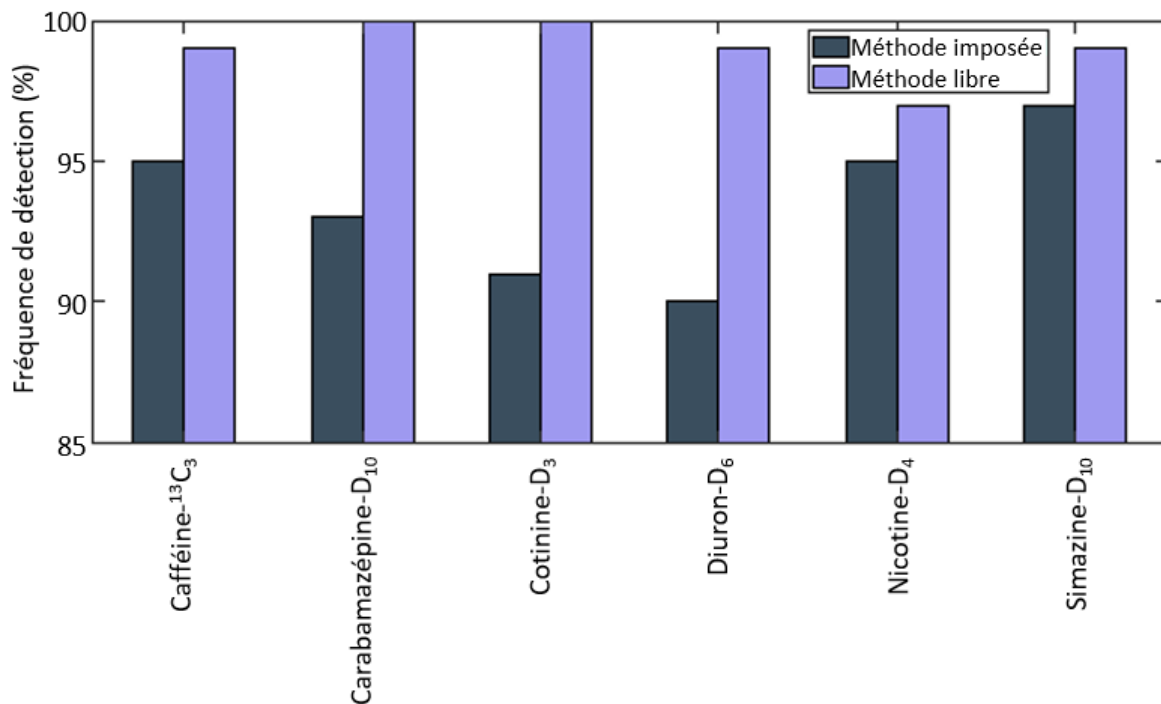
## 4. Résultats de l'essai inter-laboratoire

### 4.1. Comparaison qualitative des résultats HRMS de l'EIL

D'une manière générale, 20 laboratoires sur les 21 participants ont apporté des modifications à la méthode analytique « imposée » par les organisateurs. Celles-ci concernent aussi bien la composition de l'extrait (reconstitution dans des solutions différentes), que le volume injecté lors de l'analyse (jusqu'à 200 µL injectés contre 10 µL demandés) et le gradient chromatographique (débit, composition...). De plus, le choix de la colonne analytique étant libre, tous les laboratoires ont analysé les échantillons à l'aide d'une colonne analytique en phase inverse C18, mais de nombreuses divergences existent dans les caractéristiques plus spécifiques de la colonne (diamètre interne, longueur, taille des particules et fournisseurs).

Pour tous les échantillons (entrée ou sortie d'usine de potabilisation d'eau) et bien, qu'il était précisé que l'acquisition des données devait se faire sans l'utilisation de « lock mass », le suivi des isotopes stables des 6 composés

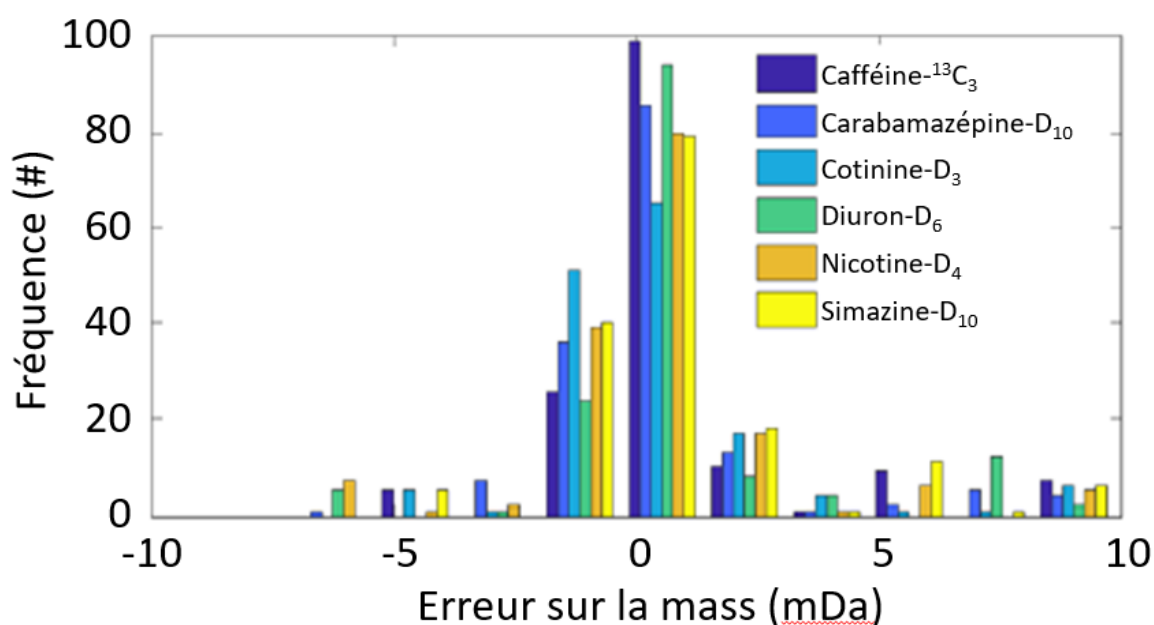
ajoutés aux échantillons révèle une bonne performance de l'ensemble des laboratoires. En effet, la fréquence de détection de ces composés n'est jamais inférieure à 90% avec la méthode imposée par les organisateurs de l'EIL (Figure 1). On remarque que la fréquence de détection augmente lorsque les laboratoires utilisent leur propre méthode ; leur expertise propre permet d'améliorer les résultats. Cela représente donc une première indication selon laquelle une méthode analytique imposée (et non optimisée par le laboratoire participant) peut engendrer un effet négatif sur la qualité des résultats, dans ce cas l'augmentation des faux négatifs (composés présents mais non détectés) lors de l'application d'une approche de type « analyse suspectée ». Une telle observation peut avoir plusieurs explications potentielles. Une première hypothèse peut envisager un effet de l'instrument. En effet, en raison des multiples brevets industriels, chaque instrument possède une géométrie de source différente selon le fournisseur. Chaque géométrie de source a donc des conditions d'ionisation optimale (température, débit de gaz, voltages...). L'ionisation peut également être influencée par la nature des solvants utilisés pour la séparation chromatographique. On comprend donc qu'une méthode imposée puisse aboutir à un nombre de détection plus faible qu'une méthode optimisée par chaque laboratoire. Une seconde hypothèse envisage alors l'impact d'un changement de méthode d'acquisition sur le traitement des données. En effet, chaque laboratoire optimise son traitement de données HRMS en fonction de la méthode d'acquisition. Par exemple, l'étape de détection des pics (peak picking) repose sur un algorithme considérant divers facteurs influencés par la méthode chromatographique. Le changement de la méthode chromatographique pour un débit de phase mobile plus élevé aboutira à des pics plus étroits qui pourraient ne plus être détectés sans une adaptation de l'algorithme. Enfin une troisième hypothèse considère tout simplement la sensibilité de la méthode. En effet, si le volume d'échantillon injecté sur la colonne chromatographique est plus faible avec la méthode d'acquisition imposée par les organisateurs de l'EIL (5  $\mu$ L) qu'avec la méthode d'acquisition propre à chaque laboratoire, alors (sans considérer les éventuels effets matrices) la méthode imposée sera moins sensible et moins de composés seront détectés.



**Figure 1.** Fréquence de détection des isotopes stables de 6 composés ajoutés aux échantillons (Schulze et al. 2021).

Au-delà de la fréquence de détection des 6 composés ajoutés aux échantillons, le biais sur la masse a également été examinée car c'est le premier critère qui intervient dans l'identification d'un composé. On observe alors que pour la majorité des échantillons analysés, l'erreur entre la masse mesurée et la masse théorique des 6 composés est inférieure à 3 ppm (figure 2). Ces résultats sont cohérents avec la littérature indiquant que la tolérance relative à la masse est généralement fixée à un maximum de 5 ppm mais qu'elle pourrait être restreinte à 3 ppm pour les instruments acquis au cours de la dernière décennie (Celma et al. 2020). Cependant, on remarque également sur la figure 2 que les 6 composés sont dans certains cas détectés avec une erreur sur la masse pouvant aller jusqu'à 10 ppm. Bien que la fréquence de telles erreurs soit moindre, cela conduit néanmoins à certaines interrogations. En effet, lorsqu'un laboratoire obtient ces valeurs extrêmes d'erreur sur la masse lors de la recherche de composés dopés servant de contrôle qualité, cela implique qu'une tolérance plus élevée doit être appliquée lors du traitement des données par analyse suspectée. En effet, appliquer la tolérance recommandée à 5 ppm alors que l'erreur sur la masse peut atteindre 10 ppm augmentera nécessairement le nombre de « faux négatifs » lors de l'analyse suspectée. En revanche, autoriser une tolérance en masse allant jusqu'à 10 ppm augmentera alors le nombre de « faux positifs ». Les erreurs importantes sur la masse posent donc la question de la qualité de l'acquisition des données. Pour corriger en continue la déviation en masse de l'instrument, certains vendeurs (Agilent, Waters...) proposent en général un système permettant d'infuser un ou plusieurs étalons analytiques (« lock mass ») au cours de l'analyse. L'opérateur de l'instrument peut donc utiliser ou non cette option. De plus, si l'opérateur décide d'acquérir le signal d'une « lock mass », il peut généralement choisir d'appliquer la correction nécessaire directement pendant l'acquisition de la donnée, ou d'appliquer cette correction a posteriori lors du traitement de la donnée. En effet, considérant

que l'acquisition des données devait se faire sans l'utilisation de « lock mass », cela explique donc que certaines valeurs d'erreur sur la masse puissent atteindre 10 ppm. Il est alors probable qu'une méthode libre utilisant la correction par la « lock mass » permette de restaurer une déviation maximale de 5 ppm. Néanmoins, cela soulève une interrogation relative aux critères à prendre en considération pour l'identification des composés par analyse suspectée. Par exemple, pour limiter le nombre de « faux négatifs », ne serait-il pas judicieux d'appliquer une tolérance en masse allant jusqu'à 10 ppm avant d'éliminer les éventuels « faux positifs » en requérant la détection d'au moins un fragment spécifique parmi ceux présents dans la bibliothèque de spectres ?



**Figure 2.** Distribution du biais sur la masse lors de l'identification des isotopes stables de 6 composés ajoutés aux échantillons (Schulze et al. 2021).

Ces premiers résultats de l'EIL ont fait l'objet d'une valorisation par le biais d'un « data paper » (Schulze et al. 2021) publié dans la revue « scientifique data » en 2021. Cet article explicite l'organisation de l'EIL et met à disposition de la communauté scientifique la liste des données acquises par les différents laboratoires, après leur conversion au format mzXML. Sous ce format, les données peuvent être lues et exploitées par la majorité des logiciels disponibles pour les utilisateurs de HRMS. Cette valorisation de l'EIL est un point très positif car, un an après sa publication, l'article a déjà été cité à 4 reprises, principalement en introduction afin d'illustrer les essais collaboratifs réalisés par le réseau NORMAN (Mohammed Taha et al. 2022) ainsi que le potentiel de la HRMS (Nanusha et al. 2022) et de son couplage à l'échantillonnage passif (Hohrenk-Danzouma et al. 2022) pour réaliser un large screening des contaminant dans l'environnement.

De plus, les méthodes d'acquisition et de traitement des données sont peu ou pas harmonisées en HRMS et la comparaison des résultats est donc généralement compliquée, limitée ou tout simplement impossible. Dans ce contexte, l'acquisition de données HRMS sur des échantillons identiques par 21 laboratoires

représente également une opportunité rare de comparer les résultats obtenus. Au-delà des résultats présentés dans cette partie 3, un travail additionnel a donc été entrepris basé sur une analyse strictement non-ciblée des données HRMS disponibles.

#### 4.2. Comparaison des logiciels libres MZmine et Workflow4metabolomics (W4M)

Les deux logiciels libres MZmine et Workflow4metabolomics (W4M) sont fréquemment utilisés pour le traitement des données HRMS. Chaque logiciel offre des options différentes. Par exemple, W4M permet de restreindre la quantité de données à examiner à l'aide de différents filtres basés, par exemple, sur la répétabilité du signal (considération du coefficient de variation) ou sur la fréquence de détection dans des répliquas...). Ces options avancées ne sont pas prises en considération dans ce rapport car leur utilité et leur conditions d'application dépendent de la question scientifique à laquelle l'utilisateur des données HRMS souhaite répondre. Par exemple, le coefficient de variation jugé acceptable est à l'appréciation de l'analyste mais il peut aussi varier suivant l'intensité du signal (plus le signal est proche du bruit et plus le coefficient de variation sera grand). De même, alors qu'un analyste jugera fiable un composé détecté dans 2 répliquas sur 3, un autre analyste ne jugera fiable qu'un composé détecté dans tous les répliquas. La comparaison des deux logiciels se fera donc uniquement sur la bases des premières étapes de traitement des données : « peak picking » et « alignement ».

Brièvement, le « peak picking » est la première étape du traitement de données HRMS. Elle consiste à examiner l'ensemble des spectres de masse acquis lors de l'analyse afin d'identifier les signaux pouvant refléter la présence d'un composé en particulier dans l'échantillon. Le logiciel isole donc l'ensemble des  $m/z$  détectés lors de l'analyse HRMS et recherche ceux formant des pics chromatographiques. Pour chaque échantillon analysé, le « peak picking » établit donc une liste de signaux caractérisés par 3 paramètres : un rapport masse/charge ( $m/z$ ), une abondance et un temps de rétention. D'un échantillon à un autre, la variation sur la mesure se traduit par le fait qu'un même composé sera associé à un signal qui diffère légèrement en masse et en temps de rétention. Cette variation rend donc impossible la comparaison directe des résultats du « peak picking ». Il faut alors introduire une seconde étape appelée « alignement ». Cette étape examine l'ensemble des listes issues du « peak picking » et identifie les signaux identiques en considérant une tolérance (définie par l'utilisateur) sur la valeur de  $m/z$  (généralement 5 à 10 ppm) et sur le temps de rétention (généralement 0,5 à 1 minute). Il en résulte une nouvelle liste de signaux comprenant la valeur moyenne de  $m/z$ , de temps de rétention et les échantillons dans lesquels ils ont été détectés.

Afin de comparer efficacement les deux logiciels, W4M et MZmine, seulement les données HRMS (méthodes d'acquisition en annexe 3) provenant de deux laboratoires (laboratoire 16 et 13) ont été considérées. Ces deux laboratoires ont été sélectionnés arbitrairement parmi les quatre laboratoires ayant participé à

l'EIL avec des instruments identiques (QTOF de type G2 du fournisseur Waters). Les deux étapes de « peak picking » et « alignement » ont été paramétrées comme indiqué sur le tableau 1. Le peak picking est réalisé par les deux logiciels, mais on remarque que les paramètres utilisés sont différents. Par exemple, alors que W4M utilise un ratio d'intensité signal/bruit (S/N) pour identifier les signaux d'intérêt ainsi que l'intensité minimale qu'un signal doit atteindre, MZmine ne considère pas le ratio S/N. Pour ce qui concerne l'étape d'alignement, elle est définie de manière identique mais elle repose sur un algorithme différent pour chaque logiciel. Ces différences dans les paramètres et les algorithmes utilisés par chaque logiciel laisse supposer que les résultats obtenus pourraient varier de manière significative selon les logiciels.

**Tableau 1.** Paramètres utilisés pour « peak picking » et « alignement » dans Workflow4metabolomics (W4M) et MZmine (Ducrocq, 2021).

	<b>W4M</b>	<b>MZmine</b>
<b>Peak picking</b>	<b>Centwave</b> S/N = 10 Larger de pic = 3-120 secondes Filtre relatif au bruit = 3000	<b>Centroïde</b> Seuil de bruit= 3000 Lissage avec 7 points Minimum de point par pic = 5
<b>Alignement</b>	<b>Join Aligner</b> Tolérance en m/z = 10 ppm Tolérance en temps de rétention = 2 min	<b>Ransac aligner</b> Tolérance en m/z = 10 ppm Tolérance en temps de rétention = 2 min

D'une manière générale, ces différences attendues ont bien été observées, comme le montre le tableau 2. En effet, on remarque que MZmine rapporte un nombre de signaux plus élevé que W4M. Cela s'explique par le fait que W4M requière un niveau minimum de S/N à atteindre et par conséquent est plus restrictif. Cette hypothèse se confirme car on remarque que la différence entre les deux logiciels s'accroît lorsque l'échantillon devient plus complexe. En effet, pour le laboratoire n°16, par exemple, alors que les deux logiciels rapportent environ 2000 signaux dans le blanc analysé, W4M rapporte environ 3500 signaux dans l'eau de rivière tandis que MZmine en rapporte plus de 6000. Une tendance similaire se dégage après l'examen des données du laboratoire n°13. Néanmoins, puisque l'étape de « peak picking » repose sur une intensité minimale à atteindre pour chaque signal, il faut également attendre une forte différence entre plusieurs laboratoires avec des instruments différents. En effet, l'état de l'instrument au moment de l'analyse peut avoir des conséquences importantes. Par exemple, des données acquises sur un instrument ayant eu récemment une opération de maintenance préventive révéleront probablement un nombre plus élevé de signaux que des données acquises sur un instrument n'ayant pas eu de maintenance et donc moins sensible.

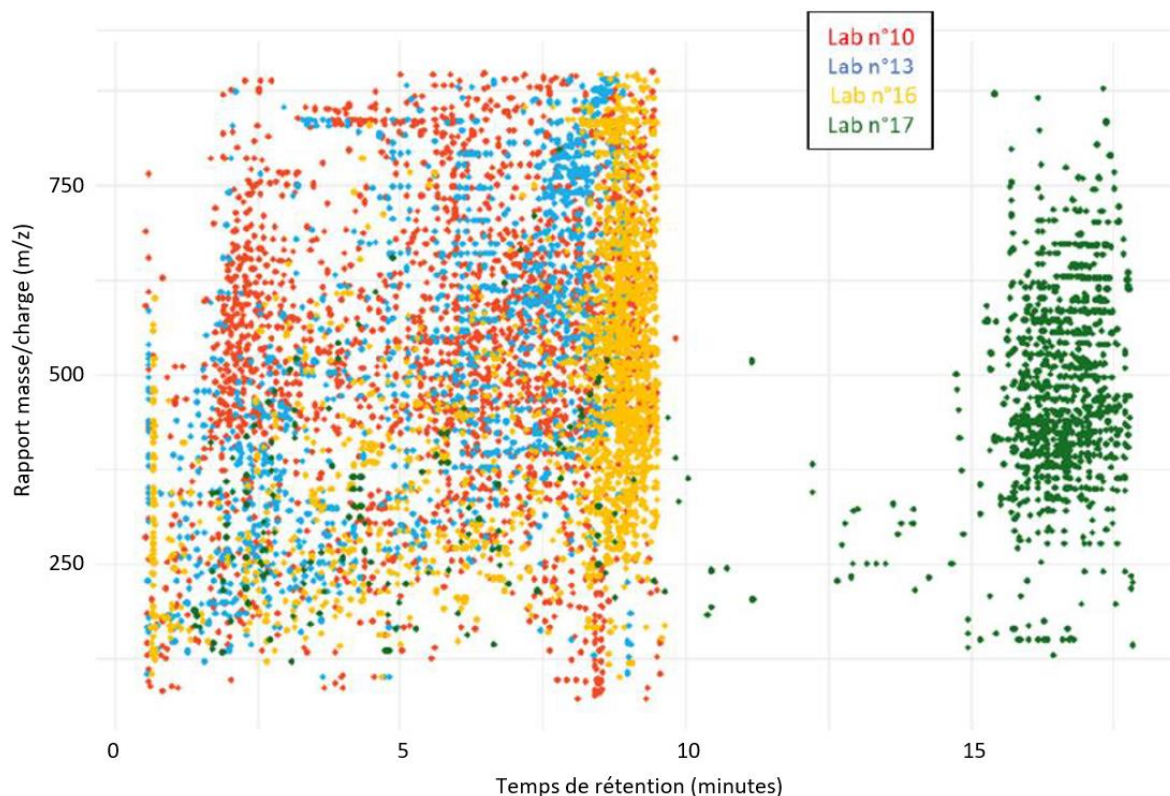
**Tableau 2.** Comparaison du nombre de signaux rapportés par les logiciels workflow4metabolomics (W4M) et MZmine dans différentes matrices analysées en ESI+ par deux laboratoires (Ducrocq, 2021).

Laboratoire	Logiciel	Blanc	Eau de rivière		Eau potable	
			(2 jours)	(4 jours)	(2 jours)	(4 jours)
Laboratoire N°16	W4M	1949	3779	3472	2678	2365
	MZmine	2158	6233	6179	2936	2171
Laboratoire N°13	W4M	1002	2931	2620	1215	1206
	MZmine	1723	5541	5411	2127	2236

Du fait des différences dans les paramètres utilisés par chaque logiciel, le seul examen des données ne permet pas d'identifier clairement une meilleure approche entre W4M et MZmine. Le choix sera donc majoritairement basé sur la préférence de l'utilisateur. D'une manière générale, MZmine est simple d'utilisation et les chromatogrammes des signaux sont accessibles à chaque étape du traitement des données. MZmine offre également un plus grand nombre de paramètres pouvant être considérés ainsi qu'une interface plus visuelle et intuitive. Cependant, W4M fonctionne à l'aide d'une plateforme et ne requiert donc pas d'outils informatiques puissants pour le traitement des données. De plus, W4M est également plus simple à mettre en œuvre pour un utilisateur peu d'expérience en HRMS car il est possible de réaliser les différentes étapes du traitement des données avec des valeurs « par défaut » pour chaque paramètre. Pour la suite de ce rapport, les données ont été traitées avec MZmine.

### 4.3. Comparaison des empreintes chimiques

Afin de mieux appréhender la diversité ou la similarité des empreintes chimiques, les données des quatre laboratoires ayant participé à l'EIL avec des instruments identiques (QTOF de type G2 du fournisseur Waters) ont été examinées (détails des méthodes analytiques en annexe 3). Une manière simple, et rapide d'examiner les données HRMS dans leur ensemble est de placer les signaux sur un graphique représentant le rapport m/z en fonction du temps de rétention. Sur la figure 3, on remarque alors que la distribution des signaux générés par le laboratoire 17 diffère très largement de celle des autres laboratoires. Cela est directement attribuable au fait que ce laboratoire utilise une colonne plus longue (150 mm) et un débit de phase mobile plus faible (0,3 mL/min) que les autres laboratoires (100 mm pour la taille de la colonne et 0,4 mL/min pour le débit de phase mobile). Par conséquent, les temps de rétention sont très fortement augmentés pour le laboratoire 17 et la comparaison avec les autres laboratoires est exclue.



**Figure 3.** Signaux obtenus après traitement MZmine des données HRMS générées par les laboratoires 10, 13, 16 et 17 pour un extrait d'échantillonneur passif exposé pendant 48 heures dans de l'eau de rivière.

Les signaux des trois autres laboratoires sont répartis entre 0 et 10 minutes, ce qui est cohérent avec l'utilisation d'une colonne (type et dimensions) et d'une phase mobile (solvants et gradient) similaires. Cependant, on remarque encore peu de superposition des résultats, ce qui traduit une empreinte chimique différente. Par exemple, on note que la majorité des signaux pour le laboratoire 16 se concentrent entre 8 et 10 minutes. Des différences significatives de temps de rétention malgré des méthodes chromatographiques similaires peuvent provenir du type d'instrument utilisé. Par exemple, les mélangeurs utilisés pour les solvants de la phase mobile peuvent avoir des volumes différents, ce qui se répercute sur le temps de rétention. En effet, avec un débit de phase mobile de 0,4 mL/min, des composés détectés par deux instruments ayant respectivement un mélangeur de 100  $\mu$ L et 2mL auront des temps de rétention qui diffèrent de plus de 4 minutes. De plus, des colonnes chromatographiques phase inverse greffées C18 légèrement différentes (inclusion d'un groupement amide ou sulfonamide) peuvent même conduire à inverser l'ordre d'élution de certaines molécules. Enfin, deux colonnes chromatographiques rigoureusement identiques peuvent également donner des temps de rétention différents en fonction de leur stade de vieillissement et du type de matrice généralement analysée (eaux résiduelles, extraits de matrices biologiques).

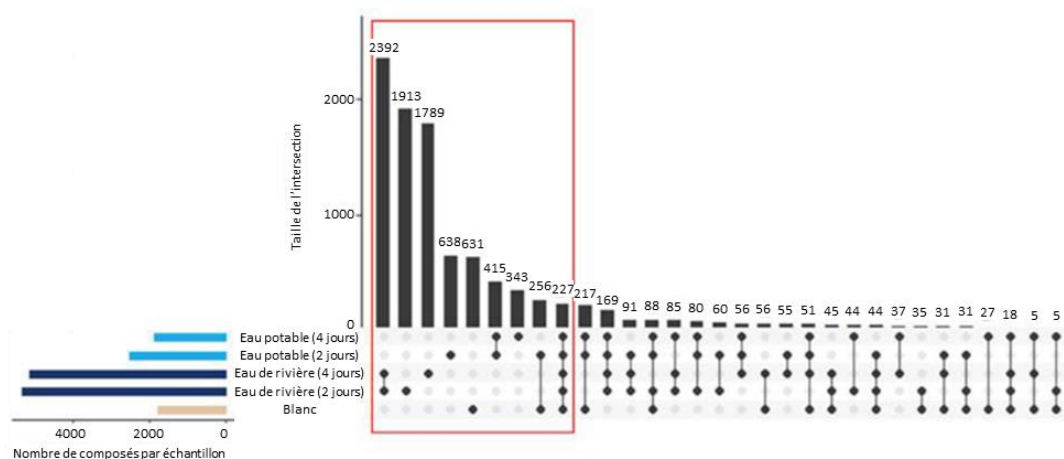


Enfin, de manière moins évidente sur le graphique, des différences sur les masses mesurées existent également. Par exemple, pour le laboratoire 10, l'erreur sur la masse pour les 6 composés (caféine-13C3, carbamazépine-D10, cotinine-D3, diuron-D6, nicotine-D4, et simazine-D10) ajoutés aux échantillons varie entre 14 et 30 ppm, ce qui est bien au-delà de la tolérance de 5 ppm généralement appliquée. Un tel écart entre la masse mesurée et la masse théorique s'explique probablement par un problème de calibration de l'instrument, l'absence de ou la mauvaise correction (sélection d'une masse erronée) par la « lock mass ». Une erreur sur la masse, probablement due à la non utilisation de la « lock mass » (comme demandé par les organisateurs de l'EIL) rend impossible la comparaison de l'empreinte chimique du laboratoire 10 avec celles des autres laboratoires.

D'une manière générale, la comparaison d'empreinte chimique d'un laboratoire à un autre s'avère particulièrement complexe, même si des méthodes chromatographiques similaires ont été appliquées. Une correction partielle des temps de rétention peut être appliquée à l'aide de l'indice de rétention mais la correction de la masse est strictement impossible pour des composés inconnus. Afin de mettre en évidence les incohérences sur les mesures de la masse, il est indispensable d'ajouter à chaque échantillon au moins un certain nombre de composés connus (isotopes stables), comme cela a été le cas dans cet EIL.

#### 4.4. Effet de la durée d'exposition des échantillonneurs passifs

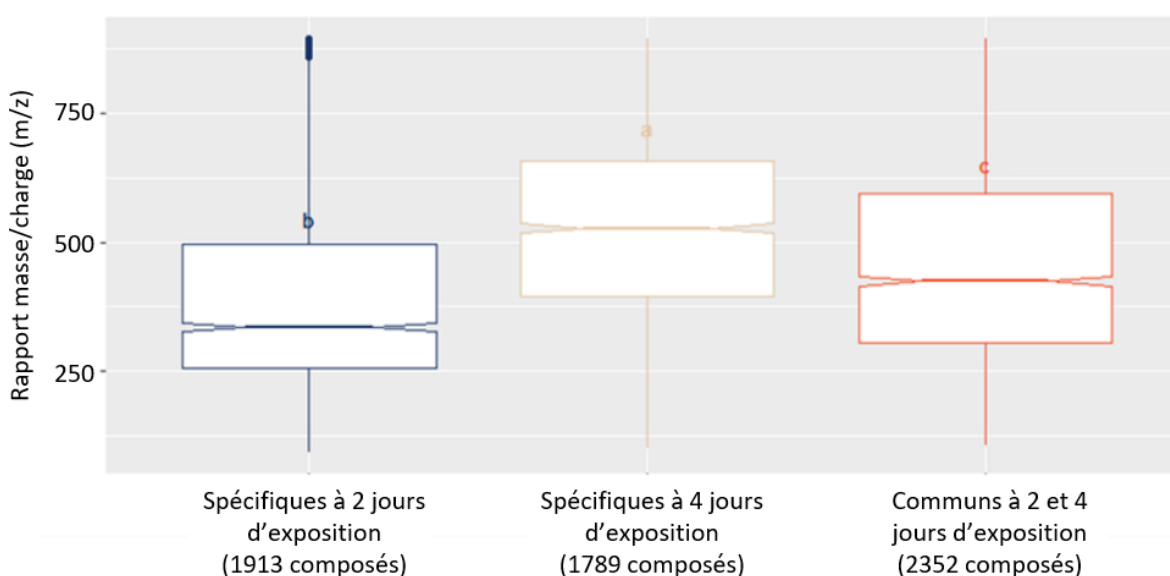
L'effet de la durée d'exposition des échantillonneurs passifs n'a pas pu être étudié de manière simultanée sur l'ensemble des 4 laboratoires sélectionnés en raison de l'hétérogénéité des résultats (cf. 4.1 et 4.2). Dans cette partie, l'effet de la durée d'exposition a été examinée grâce aux données du laboratoire 13 présentées sur la figure 4.



**Figure 4.** Effet de la durée d'exposition des échantillonneurs passifs (2 ou 4 j) et du traitement de potabilisation de l'eau (avant ou après la station) sur la composition de l'échantillon en terme de nombre de signaux.

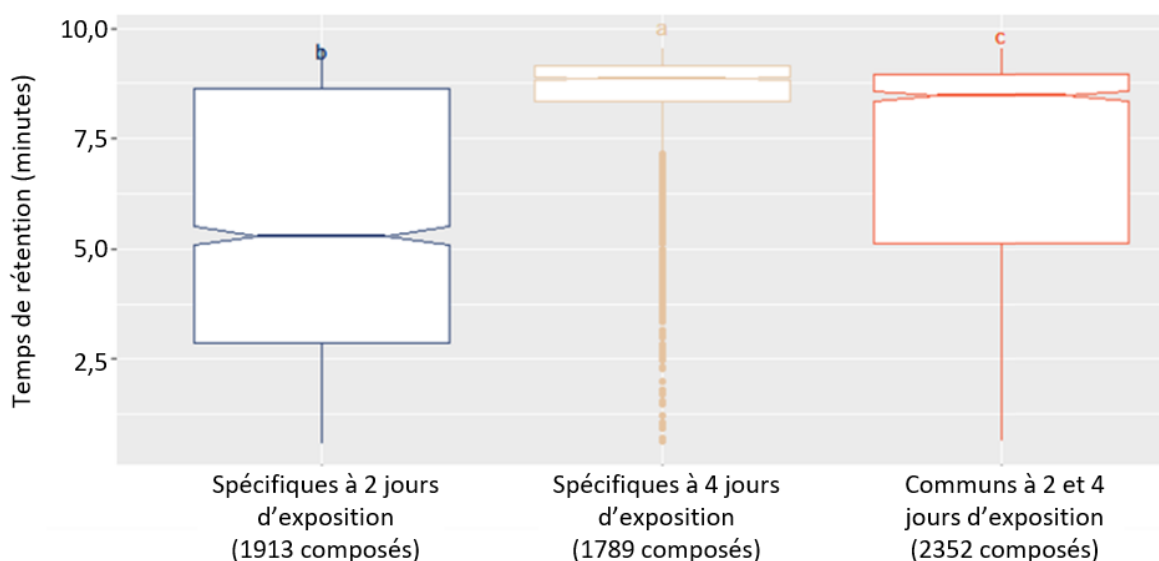
Alors que la variation de la durée d'exposition de 2 et 4 jours se traduit par un nombre de signaux similaire dans l'eau de rivière (environ 5000) ou dans l'eau potable (environ 2000), on remarque que ces signaux ne sont pas totalement identiques. Par exemple, tandis que 2392 signaux sont identiques dans l'échantillon de rivière après 2 et 4 jours, 1913 signaux supplémentaires sont observés uniquement après deux jours d'exposition et 1789 autres composés sont observés uniquement après 4 jours d'exposition. Ce résultat est assez inattendu car, on s'attend à ce que les composés adsorbés durant les 2 premiers jours soient également adsorbés sur les POCIS exposés 4 jours. En effet, en augmentant la durée d'exposition d'un POCIS, l'hypothèse la plus commune est de considérer que les composés présents dans l'échantillons s'absorberont de la même manière mais en quantité plus importante, ce qui facilitera leur détection.

Afin de comprendre ces observations, les rapports m/z des composés détectés après 2 et 4 jours d'exposition ont été examinés plus en détail. Les composés détectés uniquement après 2 jours d'exposition ont des rapports m/z inférieurs à ceux détectés uniquement après 4 jours d'exposition, comme le montre la figure 5. En effet la valeur médiane des m/z se situe aux environs de 320 après 2 jours d'exposition, tandis qu'elle se situe aux alentours de 530 après 4 jours d'exposition. Enfin, la médiane des m/z se situe aux alentours de 420 pour les composés détectés à la fois après 2 et 4 jours d'exposition.



**Figure 5.** Distribution des rapports m/z détectés sur les extraits d'échantillonneurs passifs après 2 et 4 jours d'exposition dans l'eau de rivière.

De la même manière, pour chaque temps d'exposition des échantillonneurs passifs, le temps de rétention des composés détectés a également été examiné. Sur la figure 6, on observe alors que le temps de rétention des composés détectés uniquement après 2 jours d'exposition (valeur médiane aux environs de 5 minutes) est plus faible que pour les composés détectés uniquement après 4 jours d'exposition (valeur médiane aux environs de 8,8 minutes). En comparaison, les composés détectés après 2 et 4 jours d'exposition ont des temps de rétention intermédiaires.



**Figure 6.** Distribution des temps de rétention des composés détectés sur les extraits d'échantillonneurs passifs après 2 et 4 jours d'exposition dans l'eau de rivière.

Sur la base de ces observations, la durée d'exposition des échantillonneurs passifs a une influence sur les composés détectés. Un temps d'exposition court favoriserait la détection des plus petites molécules ( $m/z$  plus faible) hydrophiles (temps de rétention plus faible) alors qu'un temps d'exposition plus long favoriserait la détection de composés de taille plus grande ( $m/z$  plus élevé) et hydrophobes (temps de rétention plus long). Ce constat pourrait alors être expliqué par les différentes étapes de diffusion, dans l'eau, la membrane et l'adsorbant, qui sont plus ou moins limitantes selon la nature des molécules. Ceci peut aussi s'expliquer par des différences d'affinité pour l'adsorbant selon les molécules. Les petites molécules hydrophiles vont diffuser plus vite dans l'eau mais ont une moins forte affinité pour la phase adsorbante ; au contraire des grosses molécules plus hydrophobes. En augmentant le temps d'exposition, on permet à plus de molécules hydrophobes d'atteindre l'adsorbant et éventuellement de prendre la place de molécules plus hydrophiles déjà adsorbées. Un examen plus approfondi des données serait nécessaire pour creuser et préciser ces interprétations.

Ce type d'étude n'a pas été rapporté dans la littérature et pourrait permettre de mieux comprendre et d'améliorer l'utilisation des échantillonneurs passifs pour le suivi de la qualité de l'eau. Généralement, les extraits d'échantillonneurs passifs sont analysés par une approche ciblée focalisée sur un nombre restreint de composés connus, avec des durées d'exposition calibrées et optimisées. A noter cependant que l'échantillonnage mis en œuvre pour l'EIL a été réalisé à l'aide d'une pompe faisant circuler l'eau à travers l'échantillonneur. Par conséquent, les échantillons générés ne sont pas complètement représentatifs du fonctionnement d'un échantillonneur passif dans des conditions réelles de déploiement. Ces observations portant sur le temps d'exposition doivent donc être confirmées à l'aide de dispositifs non-dynamiques.

## 5. Recommandations pour l'organisation de futurs EIL et pistes pour de futures exploitations des données existantes

Les premiers résultats de cet EIL combinant échantillonnage pseudo-passif et analyse HRMS mettent en évidence plusieurs points à considérer pour l'organisation de futurs essais. En premier lieu, il apparaît évident que les résultats doivent être valorisés sous forme de « data paper » car la mise à disposition des données de l'EIL permet d'optimiser leur valorisation et de rentabiliser l'effort fourni pour son organisation. Un an après sa publication, le « data paper » relatif à l'EIL bénéficie déjà de 4 citations, ce qui atteste de sa visibilité et de l'intérêt qu'il suscite pour la communauté scientifique.

D'autre part, concernant l'organisation de l'EIL du point de vue de l'analyse des échantillons, le fait d'imposer une méthode analytique spécifique ne s'avère pas utile voire même contre-productive. En effet, les laboratoires participant à l'étude n'appliquent que partiellement ou pas du tout les paramètres méthodologiques recommandés par les organisateurs. Les nombreuses variations dans la méthode analytique imposée deviennent même un facteur d'incertitude supplémentaire. De plus, cet EIL a montré que le fait d'imposer la méthode d'analyse HRMS à un laboratoire conduit généralement celui-ci à produire des résultats de qualité moindre. Typiquement, cela se traduit par une fréquence de détection des composés dopés plus faible avec la méthode imposée.

Ensuite, les différences de méthodes analytiques, d'instruments et d'outils pour le traitement des données HRMS entre les différents laboratoires confirme l'importance des contrôles qualité (CQ). En effet, certains résultats ont pu être exclus grâce à la mise en évidence d'une déviation en masse des composés dopés non acceptable. Cependant, d'autres CQ pourraient être envisagés. Par exemple, dans le domaine de la métabolomique, il est fréquent de préparer un « CQ pool ». Celui-ci est simplement un mélange dans des proportions égales de tous les échantillons analysés dans une série d'analyse. L'objectif est de rassembler et de détecter dans un seul échantillon l'ensemble des composés présents dans l'ensemble des échantillons analysés. Pour un laboratoire, l'intérêt de ce type de CQ est de pouvoir suivre la variation de la sensibilité de l'instrument au cours d'une série entière d'analyse. Dans le cadre d'un EIL, ce type de CQ permettra de comparer les sensibilités instrumentales des différents laboratoires de manière plus efficace qu'avec un nombre limité de composés dopés.

Pour de futurs EIL, la réalisation de l'échantillonnage passif avec un dispositif de pompage est discutable. En effet, bien que ce dispositif permette l'obtention des échantillons de manière plus rapide qu'une exposition des échantillonneurs en condition réelle, les résultats obtenus sont plus difficilement interprétables

et extrapolables à un échantillonneur passif classique. Pour de futurs EIL, utiliser un tel dispositif de pompage reste possible si l'objectif est de se focaliser sur l'analyse HRMS. En revanche, si l'objectif est de mieux comprendre le fonctionnement des échantillonneurs passifs et leur utilité pour le suivi de la qualité de l'eau, le déploiement d'échantillonneurs passifs classiques semble indispensable.

Enfin, les premiers résultats présentés dans ce rapport montrent que la comparaison directe d'empreintes chimiques obtenues par différents laboratoires s'avère particulièrement complexe et demande un investissement humain et statistique important. D'une manière générale, en raison des différences de méthode analytique et de sensibilité des instruments, même une comparaison des résultats non plus sur la base des couples masses vs. temps de rétention mais sur la base des couples masses vs. RTI (indice de temps de rétention) semble difficilement envisageable. A notre connaissance, à l'heure actuelle il n'existe pas de norme ni de guide portant sur la comparaison de résultat d'analyse non-ciblée entre différents laboratoires. Pour une comparaison efficace et rapide des données HRMS de plusieurs laboratoires différents, l'approche plus facilement réalisable est une analyse suspectée utilisant une large base de données. Dans un premier temps il sera nécessaire de construire une bibliothèque de spectres HRMS pour un large nombre (idéalement 3000-5000) de composés potentiellement présents dans les milieux aquatiques. Cette bibliothèque de spectre peut être construite à l'aide des listes de substances du réseau Norman (<https://www.norman-network.com/nds/SLE/>). Pour certains composés, les fragments spécifiques ne sont pas disponibles sur le réseau Norman. Dans ce cas, la fragmentation pourra éventuellement être prédite à l'aide d'outils tel que CFM-ID (<https://cfmid.wishartlab.com>) et disponible en ligne. Ainsi, pour chaque laboratoire il est possible de quantifier rapidement le nombre de composés identifiés et d'établir des listes avec l'indice de confiance associé (Schymanski et al. 2014). D'autres approches de valorisation des données peuvent être mises en œuvre. Par exemple, il est possible d'appliquer une stratégie de traitement des données, telle que l'analyse du défaut de masse (Merel 2023), pour rechercher un groupe de composés particuliers comme les produits de transformation d'un composé fréquemment détecté ou encore une famille spécifique de molécules comme les composés poly- ou perfluorés (PFASs). Enfin, la recherche de produits de transformation pour les composés les plus fréquemment détectés pourra également être complétée par une approche complémentaire : la recherche de fragments communs. En effet, les produits de transformations ont généralement une structure relativement proche du composé parent et il est donc probable qu'il se fragmentent de la même manière lors d'une analyse par spectrométrie de masse en tandem.

## 6. Références bibliographiques

- Celma, A., Sancho, J.V., Schymanski, E.L., Fabregat-Safont, D., Ibáñez, M., Goshawk, J., Barknowitz, G., Hernández, F. and Bijlsma, L. (2020) Improving Target and Suspect Screening High-Resolution Mass Spectrometry Workflows in Environmental Analysis by Ion Mobility Separation. *Environmental Science and Technology* 54(23), 15120-15131.
- Ducrocq, T. (2021) Treatment of High Resolution Mass Spectrometry data set from NORMAN network inter-laboratory trial using Non-Target Screening strategies, Ecole des Métiers de l'Environnement EME – Unilasalle Rennes.
- Hohrenk-Danzouma, L.L., Vosough, M., Merkus, V.I., Drees, F. and Schmidt, T.C. (2022) Non-target Analysis and Chemometric Evaluation of a Passive Sampler Monitoring of Small Streams. *Environmental Science and Technology* 56(9), 5466-5477.
- Merel, S. (2023) Critical assessment of the Kendrick mass defect analysis as an innovative approach to process high resolution mass spectrometry data for environmental applications. *Chemosphere* 313.
- Mohammed Taha, H., Aalizadeh, R., Alygizakis, N., Antignac, J.P., Arp, H.P.H., Bade, R., Baker, N., Belova, L., Bijlsma, L., Bolton, E.E., Brack, W., Celma, A., Chen, W.L., Cheng, T., Chirsir, P., Čirka, L., D'Agostino, L.A., Djoumbou Feunang, Y., Dulio, V., Fischer, S., Gago-Ferrero, P., Galani, A., Geueke, B., Głowacka, N., Glüge, J., Groh, K., Grosse, S., Haglund, P., Hakkinen, P.J., Hale, S.E., Hernandez, F., Janssen, E.M.L., Jonkers, T., Kiefer, K., Kirchner, M., Koschorreck, J., Krauss, M., Krier, J., Lamoree, M.H., Letzel, M., Letzel, T., Li, Q., Little, J., Liu, Y., Lunderberg, D.M., Martin, J.W., McEachran, A.D., McLean, J.A., Meier, C., Meijer, J., Menger, F., Merino, C., Muncke, J., Muschket, M., Neumann, M., Neveu, V., Ng, K., Oberacher, H., O'Brien, J., Oswald, P., Oswaldova, M., Picache, J.A., Postigo, C., Ramirez, N., Reemtsma, T., Renaud, J., Rostkowski, P., Rüdell, H., Salek, R.M., Samanipour, S., Scheringer, M., Schliebner, I., Schulz, W., Schulze, T., Sengl, M., Shoemaker, B.A., Sims, K., Singer, H., Singh, R.R., Sumarah, M., Thiessen, P.A., Thomas, K.V., Torres, S., Trier, X., van Wezel, A.P., Vermeulen, R.C.H., Vlaanderen, J.J., von der Ohe, P.C., Wang, Z., Williams, A.J., Willighagen, E.L., Wishart, D.S., Zhang, J., Thomaidis, N.S., Hollender, J., Slobodnik, J. and Schymanski, E.L. (2022) The NORMAN Suspect List Exchange (NORMAN-SLE): facilitating European and worldwide collaboration on suspect screening in high resolution mass spectrometry. *Environmental Sciences Europe* 34(1).
- Nanusha, M.Y., Frøkjær, E.E., Liigand, J., Christensen, M.R., Hansen, H.R. and Hansen, M. (2022) Unravelling the occurrence of trace contaminants in surface waters using semi-quantitative suspected non-target screening analyses. *Environmental Pollution* 315.
- Schulze, B., van Herwerden, D., Allan, I., Bijlsma, L., Etxebarria, N., Hansen, M., Merel, S., Vrana, B., Aalizadeh, R., Bajema, B., Dubocq, F., Coppola, G., Fildier, A., Fialová, P., Frøkjær, E., Grabic, R., Gago-Ferrero, P., Gravert, T., Hollender, J., Huynh, N., Jacobs, G., Jonkers, T., Kaserzon, S., Lamoree, M., Le Roux, J., Mairinger, T., Margoum, C., Mascolo, G., Mebold, E., Menger, F., Miège, C., Meijer, J., Moilleron, R., Murgolo, S., Peruzzo, M., Pijnappels, M., Reid, M., Roscioli, C., Soulier, C., Valsecchi, S., Thomaidis, N., Vulliet, E., Young, R. and Samanipour, S. (2021) Inter-laboratory mass spectrometry dataset based on passive sampling of drinking water for non-target analysis. *Scientific Data* 8(1).
- Schymanski, E.L., Jeon, J., Gulde, R., Fenner, K., Ruff, M., Singer, H.P. and Hollender, J. (2014) Identifying Small Molecules via High Resolution Mass Spectrometry: Communicating Confidence. *Environmental Science & Technology* 48(4), 2097-2098.

**Annexe 1 :**  
**Description détaillée de l'essai inter laboratoire**

# NORMAN Collaborative Trial on Passive Sampling and Non-target Screening (NTS)

This collaborative trial is based on passive sampling undertaken at a drinking water treatment plant in the Czech Republic. Passive samplers were exposed to source water (river water) and treated drinking water.

Through this collaborative trial we will attempt to answer the following scientific questions:

1. Which chemicals are (i) present and removed from source water (river) and (ii) present in drinking water and generated during drinking water treatment?
2. What is the most effective extended suspect screening workflow for the detection of these chemicals in the passive sampling extracts?
3. Can passive sampling combined with NTS be an effective strategy for the water monitoring?
  - a. Can the use of longer or multiple passive sampler exposure times increase the number of identified chemicals and improve the level of confidence in IDs?
  - b. Can passive sampling increase or decrease the matrix effects?

In order to answer these questions, we need to assess the adequacy of extended suspect screening workflows for these type of analyses via identifying the weak points of this procedure. To do so, we have set up an experimental design where passive samplers are deployed in the intake (i.e. river water) and output of a drinking water treatment plant for two different exposure times (i.e. 2 and 4 days), Figure1. Each participant is asked to follow the instructions in the task list. The tasks that are marked as compulsory, must be completed for reporting while the optional tasks are participant dependent.



Figure 1 Sampling site in the Czech Republic. Yellow dots represent river and drinking water sampling sites.





Figure 2, depicts the design of experiment used during the sampling campaign.

### Samples:

The samples are extracts of the passive sampling conducted using Horizon Atlantic® HLB disks pooled together to generate a homogenous sample to be analyzed by different participants. In addition to the samples, we included a procedural blank which is the extract of unexposed passive samplers processed in the same way as the exposed ones. For the extraction procedure, we utilized a generic method used during “joint Danube Survey 3” (i.e. three consecutive extractions with acetone and solvent exchange to methanol) in order to cover a wide range of chemical and physical properties. At present, we do not know accurately the volume of water sampled by the disks and therefore the equivalent volumes of water provided to the participants. For the moment we can assume that the sampling rates for samplers deployed in the river water and in drinking will be the same.

In total each participant receives five extracts (i.e. four passive sampler extracts and one procedural blank) and one retention indices (RTI) mixture, thus 6 vials. The list of RTI standards is included in the RTI reporting template which will be sent alongside with this document.

Vial 1	PS extract – River water – 2-day exposure	S2 2
Vial 2	PS extract – River water – 4-day exposure	S2 4
Vial 3	PS extract – drinking water – 2-day exposure	S1 2
Vial 4	PS extract – drinking water – 4-day exposure	S1 4
Vial 5	Procedural blank (field blank)	B
Vial 6	RTI mixture standard	RTI

### Task List:

1. **(Compulsory)** Run each sample using the predefined LC-MS method **ESI+**.
2. **(Compulsory)** Run each sample using their own NTS LC-MS method **ESI+**.
3. **(Compulsory)** Perform feature detection on both datasets.
4. **(Compulsory)** Perform feature identification (of 50 unique features) in each source water sample (i.e. before drinking water treatment plant and for the 2 PS duration exposure).
5. **(Optional)** Perform feature identification of 50 unique features in each treated water sample (i.e. after drinking water treatment plant and for the 2 PS duration exposure).
6. **(Compulsory)** Report the results

**Task 1**, predefined method: the reason behind this is to minimize the number of variables that may affect the results of the identification procedure. In addition to that by using this method we

can assess the results of the identifications via our target screening method. We did not want to drastically modify the composition of the samples by spiking them with a large number of internal standards. All the samples, except the RTI mixture, are spiked with only five labeled internal standards (i.e. Caffeine-13C3, Nicotine-D4, Cotinine-D3, Simazine-D10, Carbamazepine-D10, and Diuron-d6) as quality assurance means. *Please check your samples and make sure that you are able to detect these chemicals in your chromatograms.* The details of the instrumental conditions for the pre-defined method are provided in the Appendix 1.

**Task 2**, own method: please use your own favorite LC-MS method for the analysis of the same samples. This dataset will enable us to evaluate the effect of the chromatographic and MS conditions on the performance of the identification procedure.

**Task 3**, feature detection: we do not have any preferences on the feature detection algorithms used by the participants. So, feel free to use the one that you are the most comfortable with. However, please make sure that the feature detection is performed in a non-targeted way. In other words, please generate a feature list for each run chromatogram. Also, please make sure that the adducts, isotopes *are not removed* from the feature lists. These raw feature lists will be part of the report to the coordinators. Using these raw feature lists, we will try to identify the parameter affecting the feature detection the most, given that we are looking at the same sample run using similar LC conditions. Please make sure that you report all the parameters set during the feature detection. This information is essential to our evaluation and data interpretation, during the QA/QC stage by the coordinators and the publication preparation.

**Task 4 and 5**, feature identification: in this case also we do not have any preferences in terms of the software and database used for the identification. We have limited the number of features to be identified to 50, which would make it manageable for the participants. For the selection of the features, please use your own prioritization routine and specify that in your report. However, we encourage the participants to take the full advantage of the experimental design in order to prioritize the features that are unique to the sample sets (i.e. by comparing the different exposure times and the sampling site to isolated the features removed from the source water and added to the treated water). When reporting the IDs, please make sure to report the top three candidates, the matched fragments and the level of confidence, based on the Schymanski scale. This task is crucial to successful completion of our trial. We will use this information to generate a comprehensive guideline for performing extended suspect screening, which will further help the harmonization as well as the transparency in the results.

**Task 6**, reporting: along with this document, you are provided with an excel file for reporting the results. Please generate one file for each chromatographic run (i.e. 12 excel files). Additionally, please submit the raw datasets as well as the converted data in an OpenMS format (i.e. netCDF or mzXML). If your instrument provider has its own data convertor, please use that. Otherwise please use the instructions provided in Appendix 2 for the data conversion. Once you have all the files ready, use the below link to upload the data. Please make sure to first create a folder with the name of your institute and upload all your data in that folder (i.e. raw chromatograms, converted ones, and excel files).

Finally, if you use any scripting language for your data processing, we would appreciate if you could submit your scripts during the reporting process. All this information will further help us to perform QA/QC of the final results.

Please note that uploading the data into this folder implies that you agreed with the terms and conditions of participating in this trial and its data stewardship (i.e. open access and public). The terms and conditions are provided in the appendix 3.

[https://niva365-my.sharepoint.com/:f/g/personal/nivalab\\_niva\\_no/EjL80BclLH5A|gtiVtzPcKwBATfUiGb8PBgZwYmZiUSREg?e=IONcWy](https://niva365-my.sharepoint.com/:f/g/personal/nivalab_niva_no/EjL80BclLH5A|gtiVtzPcKwBATfUiGb8PBgZwYmZiUSREg?e=IONcWy)

**Please make sure all the reports and data are submitted before October 15<sup>th</sup>, 2019. However, if you need more time to complete the tasks, please get in touch with us.**

#### **Outcomes:**

We are planning to use what we have learnt from this trial to create a comprehensive and detailed guideline for extended suspect screening. We expect to publish the following scientific articles including the participants as coauthors:

1. A guideline for extended suspect screening based on an international collaborative trial.
2. Detection and identification of compounds according to the 2 PS exposure durations, and discussion/conclusion on this aspect.
3. Detection and identification of the transformation products of the chemicals of emerging concern in drinking water.
4. A dataset based on an international collaborative trial for suspect screening.

Additionally, we envisage to organize two workshops with the active participants for further discussion of the results and workflows prior to the final publication of the results. Finally, this publicly available dataset will be a valuable source, particularly for Norman member, for future developments of NTS.

**Finally, as coordinators, we would like to thank you for your participation in this collaborative trial and if you have any questions regarding the tasks or procedures listed here please do not hesitate to contact us.**

#### **Acknowledgements**

We would like to extend our thanks to Pavla Fialova and Branislav Vrana at Recetox (Masaryk University, Czech Republic), Helena Švecová and Roman Grabic at the University of South Bohemia (Czech Republic), the drinking water treatment plant and Biotage for their contribution to the Horizon disks.

## **Appendix 1: Instrumental conditions**

### **LC Method:**

#### **Allowable column chemistries:**

- Agilent Eclipse Plus C18
- Agilent Eclipse XDB C18
- Agilent Extend C18
- Kromasil C18
- AMT Halo C18
- Dionex Acclaim 120 C18
- Epic C18
- Intersil ODS-2
- Phenomenex Luca C18
- Phenomenex Prodigy C18
- Supelco Ascentis Express C18
- Supelco Discovery HS C18
- Thermo Hypersil Elite C18
- Waters Acquity HSS C18
- Waters Acquity HSS C18 SB
- Waters XSelect HSS C18
- Waters XSelect HSS C18 SB

If you do not have any of these columns, please open the below link on Chrome and activate the flash-player, Figure 2. Select “Acquity UPLC BEH C18” column (Figure 3) and find the column that you own, which is the closest to that, Figure 4.

[https://www.waters.com/waters/promotionDetail.htm?id=10048475&alias=Alias\\_selectivitychart\\_CHEMISTRY&locale=en\\_NO](https://www.waters.com/waters/promotionDetail.htm?id=10048475&alias=Alias_selectivitychart_CHEMISTRY&locale=en_NO)



Figure 1, opening the applet.

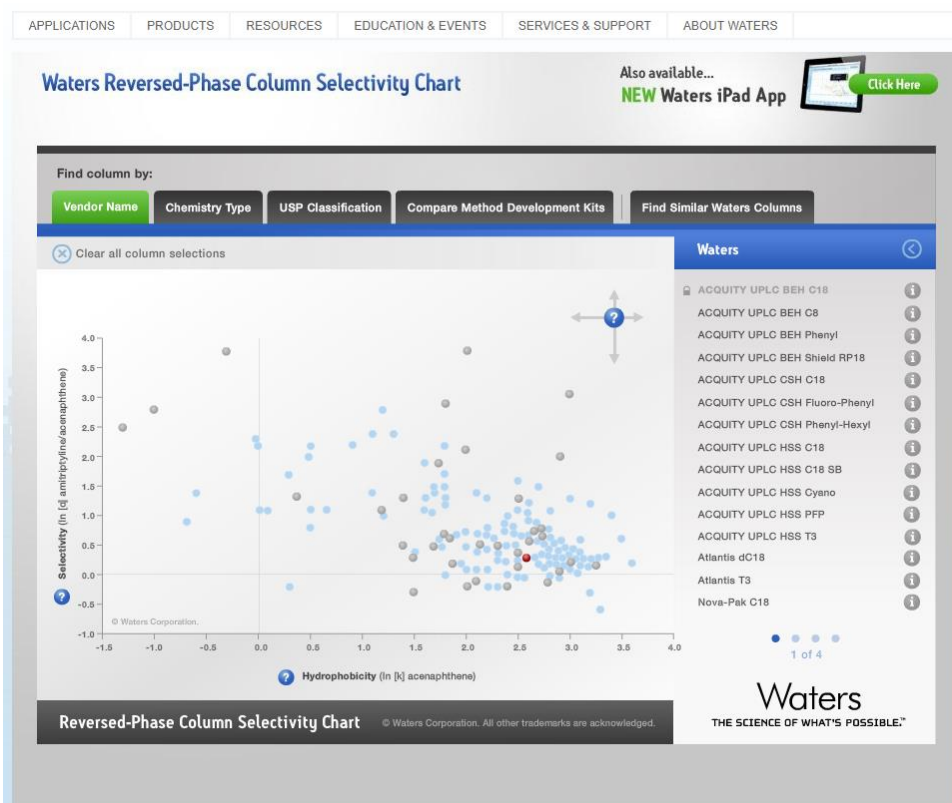


Figure 2, the applet with the Acquity UPLC BEH C18 column selected (red dot).

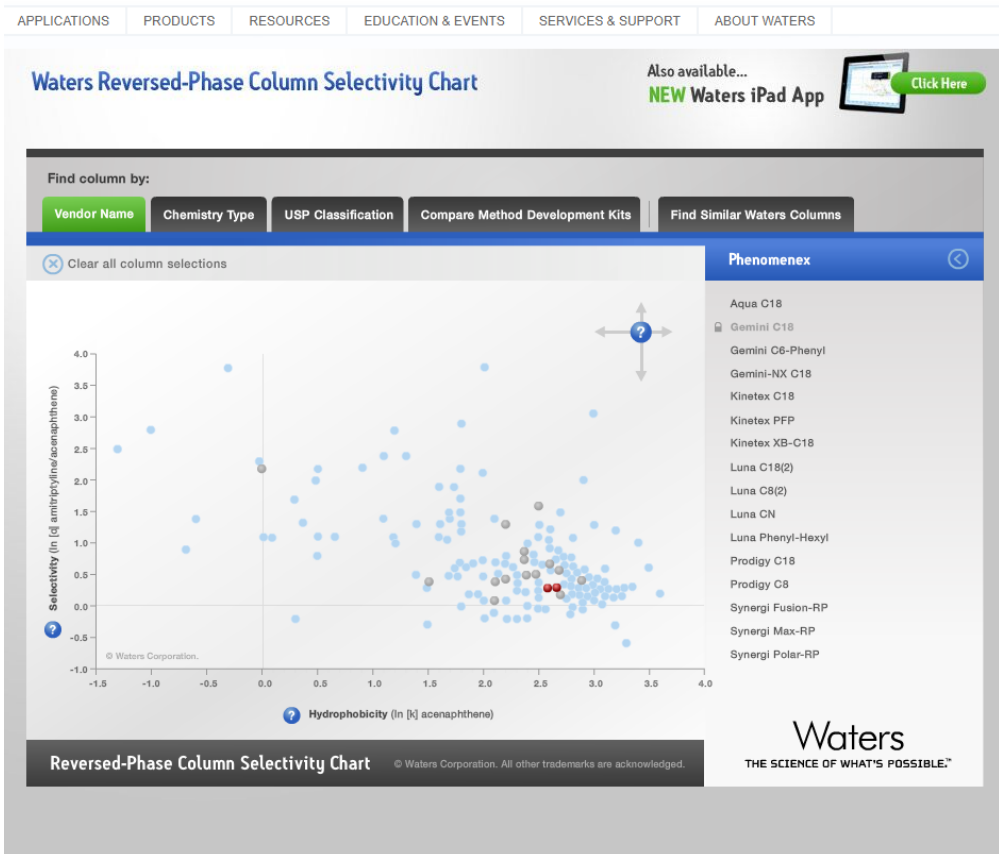


Figure 3, the column Gemini C18 selected as an example (second red dot) for a column with similar chemistry to the reference one.

**Allowable column dimensions:**

Please use the longest column with smallest particle size possible. Ideally 150 mm x 2.1, 1.8 µm.

**Mobile phases:**

Mobile Phase A: 5 mM aqueous ammonium formate, adjusted to pH 3.0.

Mobile Phase B: Acetonitrile containing 0.1% formic acid

**Gradient:**

This will depend on the column dimensions used.

The following gradient is based on a 150 mm x 2.1 mm column with 1.8 µm particle size.

Flow rate = 0.4 mL/min

Time	%A	%B
0.00	87	13
0.50	87	13
10.00	50	50
10.75	5	95
12.25	5	95

12.50 87 13

15.00 87 13

Use the following calculator to scale the above gradient to your particular column/system dimensions:

<https://www.sigmaaldrich.com/analytical-chromatography/hplc/method-transfer-calculator.html#gradient>

**Injection volume** = 5  $\mu$ L

**Mass Spectrometer Conditions:**

TOF:

Acquisition: data independent and profile mode. In case the SWATH mode is used, please make sure that the total number of SWATH windows are not more than 10. Also please submit your SWATH window information as a csv file.

Ionization mode: Positive

Ionization source: Electron spray (ESI+)

Collision Energy (CE) for MS2: either use a ramp from 10-45 eV or use an average CE of 30-40 eV.

Sampling rate: please make sure that your sampling rate is larger than 2 Hz and smaller 5 Hz.

Mass range: full scan between 60-900 Da. This applies to both MS1 and MS2 channels.

Mass calibration (**only for Waters instrument**): please make sure that you are not applying the lock-mass correction on the fly. Please the lock-mass channel and apply it later on during your identification. This is due to the fact that the lock-mass applied data when converted to OpenMS **corrupt the data**, especially in MS2.

Orbi:

Acquisition: all ion fragmentation (not variable data-independent acquisition) and profile.

Ionization mode: Positive

Ionization source: Electron spray (ESI+)

Collision Energy (CE) for MS2: either use a ramp from 10-45 eV or use an average CE of 30-40 eV.

Resolution: 60,000.

Mass range: full scan between 60-900 Da. This applies to both MS1 and MS2 channels.

**If you are not able to setup the method using these tools, please contact us and we will do it together.**

**Appendix 2:** data conversion,

If you do not have the software package "Proteowizard" already installed on your computer, please use this link to download the package (<http://proteowizard.sourceforge.net/download.html>). For installation please follow the

instructions. Once installed, please open the software MSConvert installed with Proteowizard, Figure 2.

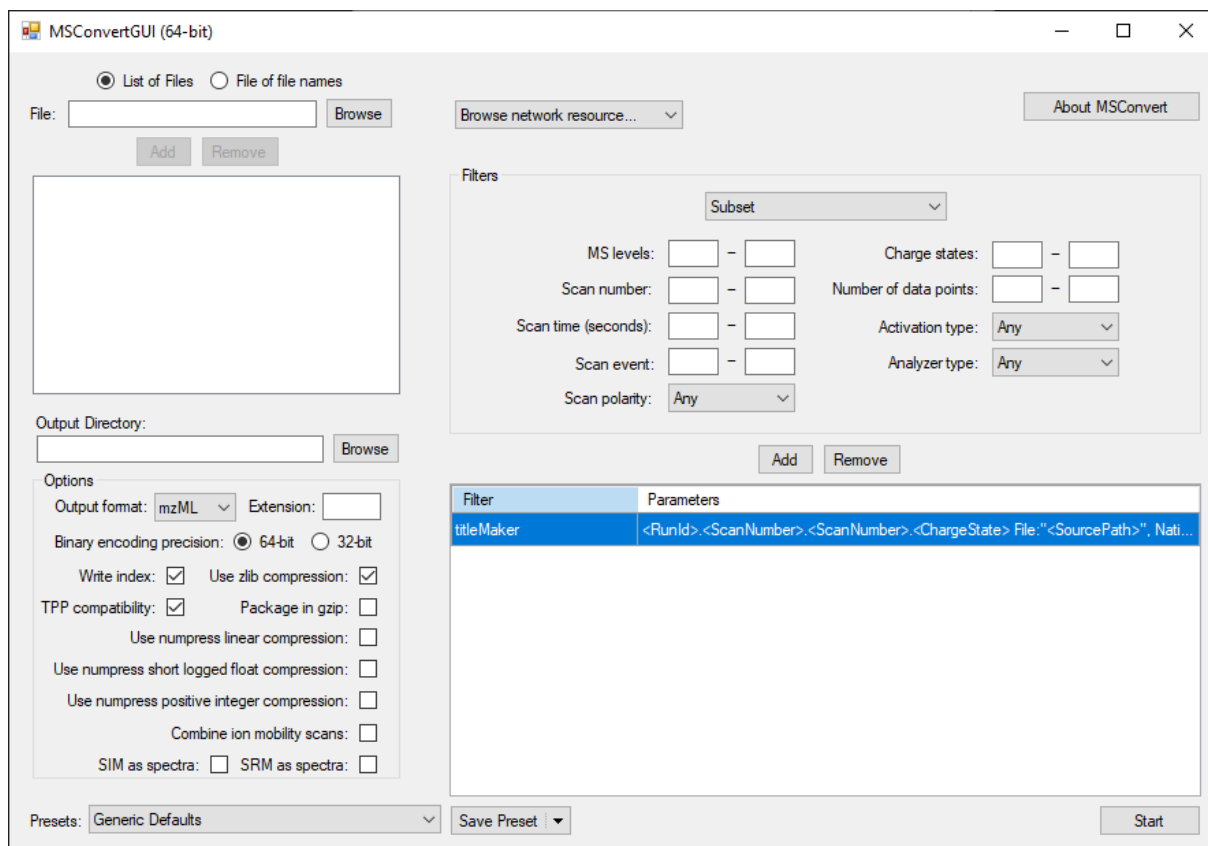


Figure 4, the MSConvert window.

Once you see this window, use the Browse button to navigate to the folder of your raw data. In the next step, select the files that you would like to convert. Please set the conversion parameters as follows:

Output Directory: path to the folder where you want to save the converted files.

Output format: mzXML

Extension: mzXML

Binary encoding precision: 32-bit

All other options need to be unchecked except Write index, Figure 3. At this point you can hit the Start button in order to convert your data. For the Waters users, please use the software package "DBridge or DataBridge" shipped with MassLynx for the data conversion.



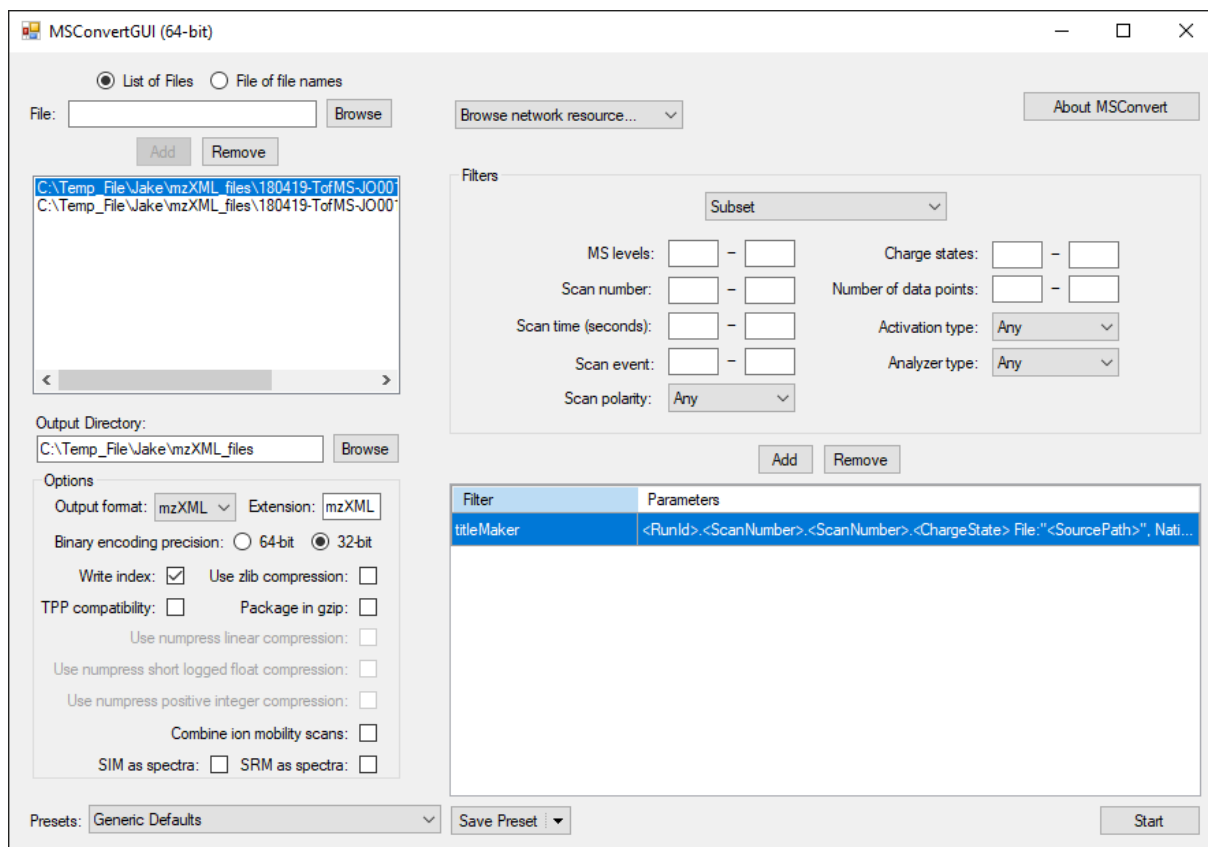


Figure 5, MSConvert settings.

**If you are not able to convert your data to OpenMS format, please contact us and we will do it together.**

### Appendix 3, terms and conditions:

Each Data Provider relinquishes all rights and ownership to its data provided in this “NORMAN Collaborative Trial on Passive Sampling and Non-target Screening (NTS)”.

Each Data Provider warrants that the Data Provider has exercised reasonable care, to assure that the Data it relinquishes to NIVA was created or generated by it or was obtained from others with the right to provide the data under this statement, and that publication of such data does not violate any privacy or confidentiality obligation undertaken by the Data Provider.

NIVA shall, in any case be exempted/free of any liability for any direct, indirect, incidental, special, exemplary, or consequential damages, arising out of the use or distribution of the collected data.

Data Provider(s), collectively and individually, waive and/or agree not to assert, to the extent permitted by law, any moral rights they hold in the collected data.

The Data Provider(s) collectively and individually hereby agree that NIVA, as the main receiver of the collected data can make the data available as open source.

**Annexe 2 :**

**Poster présenté lors de la conférence internationale sur le NTS du 4 au 7  
octobre 2021 à Erding (Allemagne)**

**Référence HAL-INRAE : hal-03967265v1**

Ducrocq, T.<sup>1</sup>, Merel, S.<sup>1</sup>, Rocco, K.<sup>1</sup>, Samanipour S.<sup>2,3</sup>, Bijlsma, L.<sup>4</sup>, Guillemain, C.<sup>1</sup>, Dubocq, F.<sup>3</sup>, Menger, F.<sup>6</sup>, Vrana, B.<sup>7</sup>, Allan, I.<sup>2</sup>, Miège, C.<sup>1</sup>

<sup>1</sup>INRAE (France), <sup>2</sup>NIVA (Norway), <sup>3</sup>University of Amsterdam (Netherlands), <sup>4</sup>University Jaume I (Spain), <sup>5</sup>Örebro University (Sweden), <sup>6</sup>University of Agricultural Sciences (Sweden), <sup>7</sup>Masaryk University (Czech Republic)

**OBJECTIVES**

An Inter-laboratory assay has been conducted by the NORMAN network, in order to compare similarity and repeatability of Non Target Analysis (NTA) by High Resolution Mass Spectrometry (HRMS) coupled to Liquid Chromatography (LC). A total of 21 laboratories in Europe had to analyse by HRMS four samples from passive samplers (PS) placed at the input and the output of drinking water treatment plant after 2 days and 4 days exposure.

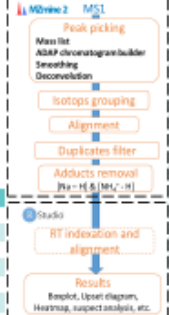
- Objective: Take advantage of the data generated by NORMAN to:
- Test different open access workflows for data treatment
  - Explore data focusing on passive sampling, particularly how it contributes to assess substances attenuated or generated during drinking water treatment, and what is the influence of the exposure time.

**METHODOLOGY**

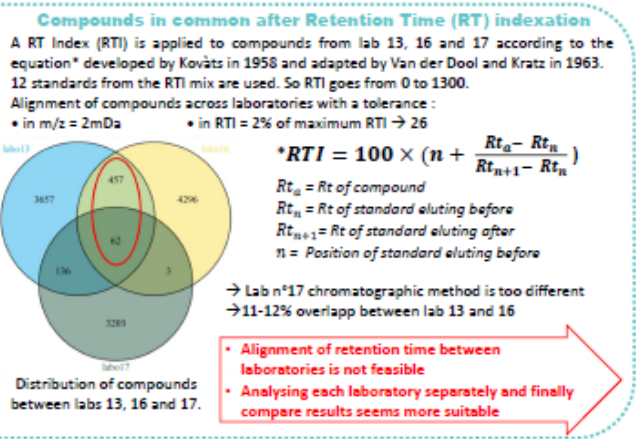
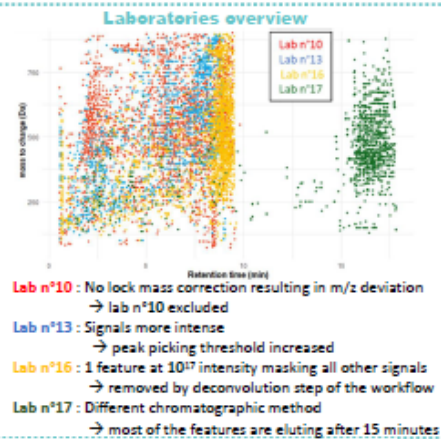
For this preliminary exploration, we selected 4 laboratories using same brand of HRMS instruments. A workflow using Mzmine and Rstudio was applied.

PS : Horizon Atlantic® HLB-L disks  
 LC : Reverse phase using a C18 column.  
 HRMS : ESI+ mode with a m/z range of 60-900Da.  
 Only data from MS1 was treated.

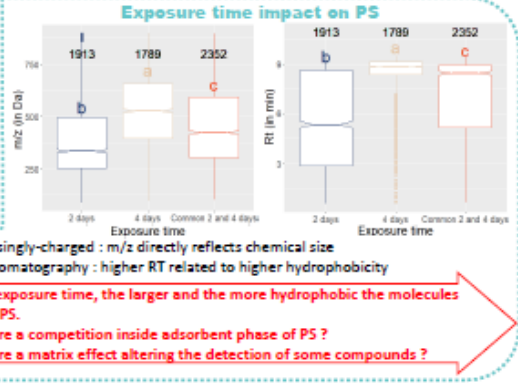
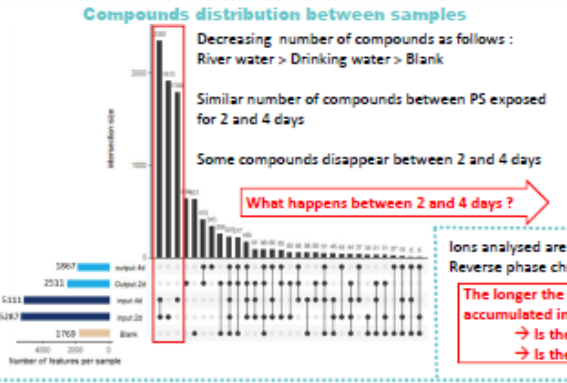
Sample	Matrix	Exposure time	Equivalent volume
001	Blank		
121	River water	2 days	4,8 L
141	River water	4 days	8,7 L
221	Drinking water	2 days	4,0 L
241	Drinking water	4 days	7,4 L
RTmix	15 standards		



**ANALYTICAL**



**ENVIRONMENTAL**



### Nature of compounds through water treatment

Type	Laboratoire n°13		
	Time	2 days	4 days
Hydrophobic and large compounds	Only in 2 days exposed PS	4486	3418
	Intensity increasing between 2 and 4 days exposed PS	1189	257
	Intensity decreasing between 2 and 4 days exposed PS	454	226
	Intensity increasing between 2 and 4 days exposed PS	335	144
Hydrophilic and small compounds	Only in 4 days exposed PS	1295	462
	Only in 2 days exposed PS	1213	329
	Only in 4 days exposed PS	911	314

Log(SM/Intensity)

Number of compounds

Deployed specific to one water / Predicted Common to both water / Detected Specific to drinking water

**PERSPECTIVES**

Even if lists of features are not comparable between laboratories, trends on the influence of PS exposure time or on the effect of water treatment are similar.

Further experiments are necessary to better understand what happen in passive sampling (test multiple exposure time in controlled media).

A suspect screening using NORMAN databases could be achieved. Identification confidence can be improved using RTI and fragmentation spectra.

### **Annexe 3 :**

**Détails des méthodes analytiques utilisées par les quatre laboratoires  
considérés dans ce rapport**

Lab	QTOF	Colonne	Injection	Phase mobile	Gradient	Débit
10	Xevo G2S	HSS T3 (100; 2.1; 1.8)	5µL	A= Eau + 5mM formate d'ammonium + 0,1% acide formique; B=acetonitrile + 0,1% acide formique	0-0,1 min isocratique (13%B), 0,1-2 min gradient (13%-50%B); 2-5,2 min gradient (50%-95%B); 5,2-8,2 min isocratique (95%B)	0,4 µL/min
13	Xevo G2-XS	BEH C18 (100; 2.1; 1.7)	5-10µL	A= Eau,+ 5mM formate d'ammonium; B=acetonitrile + 0,1% acide formique	0-0,33 min isocratique (13%B); 0.33-2.65 min gradient (13%-50%B); 2.65-7.17 min gradient (50%-95%B); 7.17-8.17min isocratique 95%B	0,4 µL/min
16	Xevo G2S	HSS T3 (100; 2.1; 1.8)	5µL	A=Eau,+ 5mM formate d'ammonium (pH=3); B=acetonitrile + 0,1% acide formique	0-0,3 min isocratique (13%B); 0.3-6.7 min gradient (13%-50%B); 6,7-7,2 min gradient (50%-95%B); 7,2-8,2 min isocratique (95%B)	0,4 µL/min
17	Xevo G2	Cortecs C18 (150; 2.1; 2.7)	5µL	A=Eau,+ 5mM formate d'ammonium (pH=3); B=acetonitrile + 0,1% acide formique	0-0.67 min isocratique (13%B); 0,67-13,33 Gradient (13%-50%B); 13,33-14,33 Gradient (50%-95%B); 14,33-16,33 min isocratique (95%B)	0,3 µL/min