

OPUR : Observatoire des Polluants Urbains en Île de France

Thème 6 : Processus en réseau d'assainissement

Action 6.1 : Stratégies d'échantillonnage des flux polluants et incertitudes associées

**Comparaison de stratégies de prélèvement
proportionnellement au temps ou au volume
pour évaluer une masse
et une concentration moyenne de polluant
transportées lors d'un événement pluvieux**

*Rapport final
Date 22 juillet 2018*

C. JOANNIS (IFSTTAR), A. HANNOUCHE (LEESU), G. CHEBBO (LEESU)
Projet ANR ECOTECH MENTOR



Sommaire

A. Introduction	10
B. Étude théorique.....	11
B.1. Définitions.....	11
B.2. Méthodes approchées de calcul d'une concentration caractéristique à partir de mesures de concentration et erreurs associées	12
B.2.1. Cas hypothétique d'un prélèvement multi-flacon :.....	12
B.2.2. Cas d'un prélèvement mono flacon.....	12
B.2.2.1. Stratégie de prélèvement proportionnellement au temps.....	13
B.2.2.2. Stratégies de prélèvement pilotées par le débit.....	13
B.2.2.2.1. Prélèvement à pas de temps variable proportionnellement au volume	13
B.2.2.2.2. Prélèvement à pas de temps fixe proportionnellement au débit moyen	14
B.2.2.2.3. Prélèvement à pas de temps fixe proportionnellement au débit instantané	16
B.2.2.3. Erreur de pondération associée à la stratégie de prélèvement proportionnellement au temps	16
B.3. Conclusion.....	18
C. Étude expérimentale.....	20
C.1. Méthode	20
C.2. Données	21
C.2.1. Description et équipement des sites	21
C.2.2. Jeux de données	22
C.3. Résultats	24
C.3.1. Stratégie de prélèvement proportionnellement au volume.....	24
C.3.1.1. Nombre de prélèvements par événement	25
C.3.1.2. Erreurs en fonction du pas de volume	25
C.3.1.3. Erreurs en fonction du nombre de prélèvements/heure	26
C.3.1.4. Impact des prélèvements manqués.....	28
C.3.2. Stratégie de prélèvement proportionnellement au temps.....	30
C.3.2.1. Nombre de prélèvement par événement	30
C.3.2.2. Erreurs en fonction du pas de temps.....	31
C.3.2.3. Comparaison des erreurs de pondération avec leurs valeurs théoriques	31
C.3.3. Comparaison des deux stratégies	33
C.3.4. Discussion : ces résultats sont-ils transposables à d'autres sites ?.....	34
D. Conclusion.....	36

E. Références 37

Liste des figures

Figure 1 constitution d'un échantillon moyen proportionnel au débit par la méthode des aliquotes	15
Figure 2 : nombre de prélèvements par événement en fonction du pas de volume (capacité du flacon 400 prélèvements)	25
Figure 3 : IC95 des erreurs en fonction du pas de volume	25
Figure 4 IC95 des erreurs en fonction du nombre de prélèvements par événement.....	26
Figure 5 : Erreurs de discrétisation observées pour la stratégie V comparées aux incertitudes de référence (moyenne de tirages indépendants)	27
Figure 5 : IC90 des erreurs en fonction de la proportion de prélèvements manquants (pour un taux de perte donné - X%- on compare les erreurs observées pour ce taux de perte, et celles qui sont observées pour les mêmes événements , avec un paramétrage ne provoquant pas de perte de prélèvement – voir texte)	29
Figure 6 : nombre de prélèvements par événement en fonction du pas de temps	30
Figure 7 : IC95 des erreurs en fonction du pas de temps	31
Figure 8 : comparaison des erreurs observées pour la stratégie T avec l'erreur théorique de pondération	32
Figure 9 : Comparaison globale pour les mêmes événements des IC95 des erreurs associés aux deux stratégies	33

Liste de tableaux

Tableau 1: récapitulation des principales stratégies de prélèvement mono-flacon.....	18
Tableau 2: paramétrages utilisés dans les simulations pour des deux stratégies de prélèvements	20
Tableau 3: Description et caractéristiques des bassins versants étudiés et les données de mesure en continu utilisées sur chaque site.....	21
Tableau 4: Caractéristiques pluviométriques des événements identifiés sur les sites d'études (d_{10} - d_{90})	23
Tableau 5: valeurs moyennes inter-événementielles des paramètres caractérisant les événements et susceptibles d'influencer les erreurs de pondération et de discrétisation sur les sites d'études.....	23
Tableau 6: rapports entre l'intervalle de confiance d'une moyenne aléatoire et l'intervalle de confiance à 95% des erreurs de discrétisations observées.....	28
Tableau 7: bornes les plus pénalisantes des IC90 (i.e. correspondant à la plus grande erreur en valeur absolue), pour les différents sites et différents taux de pertes, comparées aux valeurs de référence observées pour les mêmes événements en absence de perte de prélèvement.....	30

Comparaison de stratégies de prélèvement proportionnellement au temps ou au volume pour évaluer une masse et une concentration moyenne de polluant transportées lors d'un événement pluvieux

Résumé :

La détermination d'une concentration moyenne d'un polluant donné lors d'un événement pluvieux est classiquement réalisée à partir de l'analyse d'un échantillon composite, obtenu en cumulant plusieurs prélèvements répartis sur la durée de l'événement. La représentativité de cet échantillon vis à vis de la masse de polluant associée à l'événement dépend de la stratégie de pilotage du préleveur.

Cette étude a pour objectif :

- de préciser les bases théoriques des principales stratégies de prélèvement d'un échantillon moyen
- de comparer à partir de données expérimentales les erreurs associées à chacune des deux stratégies les plus courantes et affectant l'estimation de la concentration moyennes et de la masse t associées à un événement pluvieux.

La partie expérimentale est basée sur des enregistrements de débit et de turbidité acquis sur des sites séparatifs et unitaires des observatoires du réseau URBIS. L'application de stratégies de prélèvement proportionnellement respectivement au temps et au volume écoulé a été simulée sur plusieurs centaines d'événements pluvieux pour plusieurs paramétrages.

L'analyse des résultats met en évidence la nette supériorité de la stratégie de prélèvement asservie au débit, car la stratégie de prélèvement asservie au temps implique une erreur de pondération qui peut être considérable sur certains sites et certains événements. En revanche il reste délicat d'évaluer les incertitudes de discrétisation et de pondération pour un site donné si on ne dispose pas d'une mesure en continu de pollution associée aux prélèvements.

A. Introduction

La surveillance des rejets polluants s'effectue classiquement à l'aide de préleveurs automatiques fournissant des échantillons d'eau. Ceux-ci sont ensuite collectés et transmis à un laboratoire pour analyser les concentrations en divers polluants. Ces opérations sont détaillées dans la norme AFNOR (FD T90-523-2 Qualité de l'eau - Guide de prélèvement pour le suivi de qualité des eaux dans l'environnement - Partie 2 : prélèvement d'eau résiduaire), dont une nouvelle version paraîtra en 2018.

Ces opérations sont lourdes et les prélèvements, en nombre nécessairement limité, doivent être déclenchés à bon escient. Dans le cas des rejets urbains de temps de pluie, les échantillons sont constitués lors des événements pluvieux, dont l'occurrence et l'importance sont aléatoires, ce qui rend difficile l'optimisation des prélèvements.

Pour améliorer sa représentativité un échantillon est en général constitué de plusieurs prélèvements. Dans certains cas on constitue plusieurs échantillons lors d'un événement pluvieux : on parle alors de prélèvement multi-flacons. Les prélèvements peuvent aussi être accumulés dans un flacon unique de grande capacité (par exemple 10 l). On parle alors de prélèvement mono-flacon.

Dans tous les cas la représentativité dépend de la stratégie de pilotage du préleveur, c'est à dire de la consigne provoquant le déclenchement de chaque prélèvement.

Nous nous proposons dans cette étude :

- de préciser les bases théoriques des principales stratégies de prélèvement d'un échantillon moyen
- de comparer à partir de données expérimentales les erreurs associées à chacune des deux stratégies les plus courantes lorsqu'il s'agit de constituer un échantillon moyen représentatif d'un événement pluvieux.

B. Étude théorique

B.1. Définitions

La masse M totale de polluant (en général appelée flux polluant) écoulee au cours d'une journée ou d'un événement pluvieux correspond à l'intégration du flux instantané $Q(t) \cdot C(t)$ par rapport au temps :

$$M = M(t_1, t_2) = \int_{t_1}^{t_2} Q(t) \cdot C(t) \cdot dt \quad \text{Eq. 1}$$

Où M désigne le flux d'un polluant donné cumulé sur un intervalle de temps (t_1, t_2)

$Q(t)$ le débit en fonction du temps

$C(t)$ la concentration du polluant considéré en fonction du temps

$T=(t_2-t_1)$: durée d'intégration

Si on rapporte cette masse au volume V écoulé sur la période (t_1, t_2) correspondant à la durée d'un événement on définit une concentration moyenne C_M , *pondérée par les débits*, que nous appellerons *concentration de masse* : elle permet de reconstituer une masse lorsqu'on dispose d'une mesure du volume écoulé.

$$C_M = C_M(t_1, t_2) = \frac{\int_{t_1}^{t_2} Q(t) \cdot C(t) \cdot dt}{\int_{t_1}^{t_2} Q(t) \cdot dt} = \frac{\overline{QC}}{V} \cdot (t_2 - t_1) \quad \text{Eq. 2}$$

Elle peut être généralisée pour définir la concentration *caractéristique d'un site*, en faisant une moyenne sur plusieurs événements, pondérée par les volumes des événements.

$$C_c = \frac{\sum_1^{ne} V_j C_{Mj}}{\sum_1^{ne} V_j} \quad \text{Eq. 3}$$

Avec ne : nombre d'événements ;

V_j, C_{Mj} : volume et concentration de masse mesurés pour l'événement j ;

La concentration caractéristique d'un site présente un intérêt pratique car elle permet certaines extrapolations. En effet le volume peut facilement être mesuré pour un site et un événement donné, alors que les mesures à mettre en œuvre pour obtenir un flux sont beaucoup plus lourdes. Pour estimer une masse on pourra donc choisir de combiner une mesure de débit avec une concentration caractéristique du site issue de mesures réalisées sur le site, de données bibliographiques ou d'un modèle.

Nous verrons plus loin que de nombreuses stratégies de prélèvement visent à matérialiser la concentration de masse d'un événement dans un échantillon composite.

La concentration de masse d'un événement diffère de la moyenne temporelle \bar{C} sur la même période. Celle-ci s'exprime comme :

$$\bar{C} = \bar{C}(t_1, t_2) = \frac{\int_{t_1}^{t_2} C(t) \cdot dt}{t_2 - t_1} = \frac{\bar{Q} \cdot \bar{C}}{V} \cdot (t_2 - t_1) \quad \text{Eq. 4}$$

et la différence entre les deux concentrations est :

$$\bar{C} - C_M = \frac{\bar{Q} \cdot \bar{C} - \overline{QC}}{V} \cdot (t_2 - t_1) = \frac{\text{cov}(Q, C)}{V} \cdot (t_2 - t_1) \quad \text{Eq. 5}$$

B.2. Méthodes approchées de calcul d'une concentration caractéristique à partir de mesures de concentration et erreurs associées

B.2.1. Cas hypothétique d'un prélèvement multi-flacon :

Supposons pour simplifier qu'on ne réalise qu'une prise par flacon. On dispose alors de différentes valeurs de concentration et des valeurs de débits associées et on peut approcher le flux M selon la formule (intégration par la méthode des trapèzes) par \widehat{M}_T :

$$\widehat{M}_T = \sum_1^n \frac{1}{2} (Q_i C_i + Q_{i-1} C_{i-1}) \Delta t_i \approx M \quad \text{Eq. 6}$$

où Q_{i-1} et Q_i représentent le débit instantané, C_{i-1} et C_i représentent la concentration du prélèvement respectivement au début et à la fin du pas de temps Δt_i : $Q_i = Q(t_i)$, $C_i = C(t_i)$
La différence entre M et \widehat{M}_T est l'erreur de discrétisation pour la méthode des trapèzes et la discrétisation temporelle choisie. Elle est d'autant plus grande que le pas de temps (moyen) de prélèvement est grand.

Une bonne précision sur l'évaluation du flux nécessiterait des pas de temps Δt_i de l'ordre de quelques minutes, et une analyse de chaque échantillon pour obtenir tous les C_i , ce qui impliquerait des moyens logistiques et des frais d'analyse très importants.

En pratique on peut réduire le nombre d'analyses en réalisant plusieurs prises par flacon, ce qui permet de diminuer le nombre de flacons par événement. En poussant la logique jusqu'au bout, on aboutit à un flacon unique contenant un mélange de divers prélèvements représentatif de l'ensemble de l'événement

B.2.2. Cas d'un prélèvement mono flacon

Le contenu du flacon correspond au mélange de n prises, de volume δv_i . Il s'agit ici du volume d'eau prélevée et transférée dans le flacon pour représenter un intervalle de temps Δt_i , à ne pas confondre avec le volume ΔV_i transité pendant le même temps dans l'écoulement à surveiller

La concentration dans le flacon vaudra

$$C_f = \frac{\sum_1^n C_i \delta v_i}{\sum_1^n \delta v_i} \quad \text{Eq. 7}$$

C_f est la moyenne des concentrations prélevées, *pondérée par le volume des prises*.

B.2.2.1. Stratégie de prélèvement proportionnellement au temps

C'est la stratégie la plus simple. On prélève un volume δv fixe à un pas de temps Δt fixe.

$$C_{ft} = \frac{\sum_1^n C_i \delta v}{\sum_1^n \delta v} = \frac{\sum_1^n C_i}{n} = \frac{\sum_1^n C_i \Delta t}{\sum_1^n \Delta t} \approx \bar{C} \quad \text{Eq. 8}$$

Dans ce cas la concentration de l'échantillon moyen contenu dans le flacon *ne correspond pas à une valeur moyenne pondérée par les volumes écoulés*. Si on la multiplie par le volume total écoulé pour calculer un flux, on introduira une nouvelle erreur que l'on appellera *erreur de pondération*.

D'un point de vue pratique, l'avantage de cette stratégie est d'être très simple à mettre en œuvre : elle ne nécessite pas de débitmètre et peut fonctionner sur l'horloge interne du préleveur (mais il faut néanmoins un signal extérieur pour démarrer et arrêter une série de prélèvements). Pour un volume de flacon et un volume de prélèvement donné, la capacité de prélèvement est déterminée par le pas de temps choisi et peut facilement atteindre quelques heures. On peut ainsi être assuré de couvrir une large gamme d'événement pluvieux, même si le plus courts ne font l'objet que d'un nombre limité de prélèvements.

B.2.2.2. Stratégies de prélèvement pilotées par le débit

B.2.2.2.1. Prélèvement à pas de temps variable proportionnellement au volume

On peut introduire une pondération par les débits en maintenant le volume de prise δv_i constant ($\delta v_i = \delta v$) et en faisant varier l'intervalle de temps entre deux prélèvements en raison inverse du débit moyen écoulé.

La pondération par les volumes écoulés est alors indirecte : les plus forts débits correspondent à des pas de temps plus courts, et donc à des prélèvements plus nombreux. Ceux-ci pèseront davantage dans la concentration moyenne obtenue dans le flacon que les prélèvements moins nombreux correspondant aux débits plus faibles.

On a $\Delta t_i = \frac{k}{\bar{Q}_i}$, avec \bar{Q}_i débit moyen sur l'intervalle de temps Δt_i compris entre l'instant t_i d'un prélèvement et l'instant t_{i-1} du prélèvement précédent.

$k = \bar{Q}_i \Delta t_i$ est une constante, égale au volume ΔV écoulé entre deux prélèvements.

Comme précédemment la concentration dans le flacon vaut

$$C_{fV} = \frac{\sum_1^n C_i \delta v}{\sum_1^n \delta v} = \frac{\sum_1^n C_i}{n} \quad \text{Eq. 9}$$

Mais puisque ΔV est constant et égal à $\bar{Q}_i \Delta t_i$

$$C_{fV} = \frac{\Delta V \sum_1^n C_i}{n \Delta V} = \frac{\sum_1^n \bar{Q}_i \cdot C_i \Delta t_i}{n \Delta V} \quad \text{Eq. 10}$$

Si on pose

$$\widehat{m}_V = \sum_1^n \bar{Q}_i \cdot C_i \Delta t_i \quad \text{Eq. 11}$$

\widehat{m}_V est une approximation de M obtenue par un schéma d'intégration de QC combinant la méthode des rectangles à pas de temps variable pour C et une intégration « juste » (ou basée sur un pas de discrétisation court) pour Q.

\widehat{m}_V ne prend pas en compte la masse écoulée pendant la durée séparant le dernier prélèvement de la fin de l'événement). On peut extrapoler cette masse pour calculer \widehat{M}_V :

$$\widehat{M}_V = \widehat{m}_V \frac{V}{n \Delta V} = V \cdot C_{fV} \quad \text{Eq. 12}$$

B.2.2.2.2. Prélèvement à pas de temps fixe proportionnellement au débit moyen

On peut obtenir un résultat analogue en maintenant Δt_i constant ($\Delta t_i = \Delta t$), et en faisant varier δv_i proportionnellement à \bar{Q}_i , débit moyen sur l'intervalle de temps Δt compris entre l'instant t_i d'un prélèvement et l'instant t_{i-1} du prélèvement précédent. Cela peut être fait automatiquement par un préleveur mono-flacon adapté.

On peut aussi utiliser un préleveur multi-flacon et reconstituer a posteriori un échantillon moyen à partir de prélèvements proportionnellement au temps et d'un enregistrement des débits. Dans ce cas on reconstitue un échantillon composite en prélevant dans chaque flacon un volume δv_i proportionnel au débit moyen \bar{Q}_i . C'est la méthode dite « des aliquotes ». Cette méthode était utilisée par le passé, lorsque les débitmètres et les préleveurs disposaient de moyens d'asservissement limités, mais son intérêt ici est surtout de permettre une visualisation graphique de la stratégie de prélèvement proportionnellement au débit moyen.

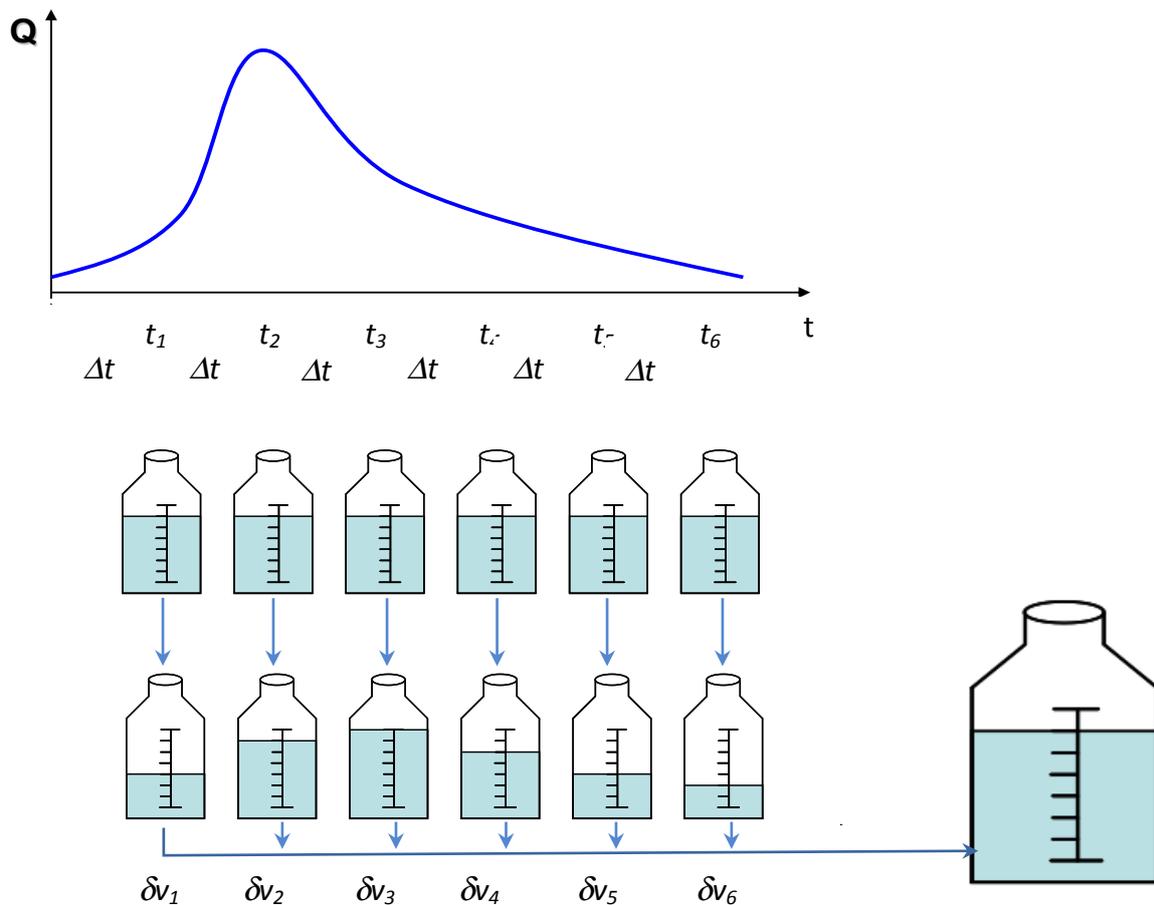


Figure 1 constitution d'un échantillon moyen proportionnel au débit par la méthode des aliquotes

$$C_{fv'} = \frac{\sum_1^n C_i \delta v}{\sum_1^n \delta v} = \frac{\sum_1^n k C_i \bar{Q}_i \Delta t}{\sum_1^n k \bar{Q}_i \Delta t} = \frac{\sum_1^n \bar{Q}_i \cdot C_i \Delta t}{\sum_1^n \Delta V_i} \quad \text{Eq. 13}$$

Comme précédemment on obtient une approximation \widehat{M}'_V de M en extrapolant un schéma d'intégration de QC combinant la méthode des rectangles pour C et une intégration « juste » (ou basée sur un pas de discrétisation court) pour Q .

$$\widehat{M}'_V = \frac{V}{\sum_1^n \Delta V_i} \sum_1^n \bar{Q}_i \cdot C_i \Delta t \quad \text{Eq. 14}$$

D'un point de vue pratique, ces deux stratégies ont pour l'essentiel les mêmes avantages et les mêmes défauts. Elles réalisent une moyenne des concentrations pondérées par les débits selon le même schéma d'intégration. Elles nécessitent un débitmètre et un système d'intégration pour obtenir des volumes cumulés. Ce système est en général implémenté dans le débitmètre qui fournit au préleveur des impulsions déclenchant les prélèvements.

Si les masses recherchées sont associées à des événements pluvieux, la gamme des volumes concernés peut être très large et il est difficile d'optimiser la relation entre les volumes

prélevés et les volumes écoulés pour couvrir toute la gamme : des prélèvements trop fréquents dépasseront la capacité du flacon pour les événements les plus importants en termes de volume ruisselé et la fin de ces événements ne sera pas échantillonnée. A l'inverse des prélèvements moins fréquents conduiront à sous-échantillonner des événements les plus faibles. Néanmoins les événements courts seront bien échantillonnés si les volumes écoulés sont assez importants.

B.2.2.2.3. Prélèvement à pas de temps fixe proportionnellement au débit instantané

Pour ce dernier type de stratégie on maintient encore $\Delta t_i = \Delta t$, mais on fait varier δv_i proportionnellement à Q_i , débit instantané au moment du prélèvement. Comme précédemment cela peut être fait automatiquement par un préleveur mono-flacon adapté, ou reconstitué a posteriori par la méthode des aliquotes

On aura alors

$$C_{fQ} = \frac{\sum_1^n C_i \delta v_i}{\sum_1^n \delta v_i} = \frac{\sum_1^n C_i k Q_i \Delta t}{\sum_1^n k Q_i \Delta t} = \frac{\sum_1^n C_i Q_i \Delta t}{\sum_1^n Q_i \Delta t} \quad \text{Eq. 15}$$

et comme $V \approx \sum_1^n Q_i \Delta t$ (l'approximation correspond au volume écoulé pendant la durée séparant le dernier prélèvement de la fin de l'événement)

$$C_{fQ} \approx \frac{\widehat{M}_Q}{V} = \widehat{C}_{cQ} \quad \text{Eq. 16}$$

\widehat{M}_Q est une approximation de M obtenue par un schéma d'intégration de QC par la méthode des rectangles.

$$\widehat{M}_Q = \sum_1^n Q_i \cdot C_i \Delta t_i \quad \text{Eq. 17}$$

D'un point de vue pratique, cette stratégie est comparable avec les deux précédentes, à cela près qu'elle ne nécessite pas de système de calcul des volumes cumulés. L'optimisation de la relation entre les volumes prélevés et les débits pour couvrir toute la gamme des événements pluvieux reste tout aussi délicate

B.2.2.3. Erreur de pondération associée à la stratégie de prélèvement proportionnellement au temps

Les 3 stratégies de prélèvement qui prennent en compte le débit fournissent des valeurs approchées \widehat{M} de la masse totale M . La différence entre \widehat{M} et M constitue l'erreur d'intégration associée à la discrétisation de la mesure de concentration.

La stratégie de prélèvement proportionnellement au temps implique également une erreur de discrétisation.

Mais elle génère une erreur supplémentaire en ne réalisant pas une intégration de QC mais une intégration de C. Lorsqu'on multiplie la concentration observée dans le flacon par le volume écoulé on commet une erreur que nous appellerons erreur de pondération. C'est la différence entre $V\bar{C}$ et VC_M :

$$E_{pond}(M) = T(\bar{Q} \cdot \bar{C} - \overline{QC}) = -Tcov(Q, C) = -T\rho_{Q,C}\sigma(Q)\sigma(C) \quad Eq. 18$$

avec

$\rho_{Q,C}$: coefficient de corrélation entre le débit et la concentration

$\sigma(Q)$: écart-type du débit

$\sigma(C)$: écart-type de la concentration

T : durée d'intégration

Transposée en concentration l'erreur de pondération s'exprime :

$$E_{pond}(C_{CE}) = -\frac{\rho_{Q,C}\sigma(Q)\sigma(C)}{\bar{Q}} \quad Eq. 19$$

Rapportées au *résultat de mesure* (et non à la valeur du mesurande) ces erreurs s'expriment sous forme relative comme

$$\frac{E_{pond}(M)}{V\bar{C}} = \frac{E_{pond}(C_M)}{\bar{C}} = -\rho_{Q,C} \frac{\sigma(Q)}{\bar{Q}} \frac{\sigma(C)}{\bar{C}} \quad Eq. 20$$

Si la corrélation entre Q et C est négligeable, ou si le coefficient de variation de C ou de Q est faible, $\Phi \approx V \cdot \bar{C}$ la concentration d'un prélèvement asservi au temps sera voisine de la concentration caractéristique obtenue par un prélèvement asservi au débit ou au volume. Sinon une corrélation positive entrainera une erreur par défaut, une corrélation négative une erreur par excès.

Des mesures en continu de Q, et pour certains paramètres de C, permettent une évaluation assez précise de $\rho_{Q,C}$, $CV(Q)$ et $CV(C)$ et donc de l'erreur de pondération.

Sous forme discrétisée les équations 17 et 18 deviennent

$$\Delta M = \widehat{M}_t - \widehat{M}_Q = -n\Delta T \sum_1^n (Q_i - \bar{Q}) \cdot (C_i - \bar{C}) = -n\Delta T \cdot \widehat{\rho}_{Q,C} \cdot \widehat{\sigma}(Q) \cdot \widehat{\sigma}(C) \quad Eq. 21$$

$$\Delta C_{CE} = \widehat{C}_{ct} - \widehat{C}_{cQ} = -\frac{\widehat{\rho}_{Q,C} \cdot \widehat{\sigma}(Q) \cdot \widehat{\sigma}(C)}{\widehat{Q}} \quad Eq. 22$$

$\widehat{\rho}_{Q,C}$, $\widehat{\sigma}(Q)$, $\widehat{\sigma}(C)$ et \widehat{Q} désignent respectivement les estimateurs de $\rho_{Q,C}$, $\sigma(Q)$, $\sigma(C)$ et \bar{Q} obtenus à partir des valeurs échantillonnées

Elles expriment l'écart entre la concentration obtenue avec un prélèvement proportionnel au temps et la concentration obtenue pour le même pas de temps avec un prélèvement proportionnel *au débit instantané*.

Exprimés sous forme relative ces écarts s'expriment

$$\frac{\Delta M}{\widehat{M}_t} = \frac{\Delta C_M}{\widehat{C}_{ct}} = -\widehat{\rho}_{Q,C} \cdot \widehat{CV}(Q) \cdot \widehat{CV}(C) \quad \text{Eq. 23}$$

où $\widehat{CV}(Q)$ et $\widehat{CV}(C)$ désignent les coefficients de variation empiriques, calculés à partir des mesures

L'erreur de pondération ne dépend pas directement de la fréquence de prélèvement. Celle-ci ne joue que sur l'erreur de discrétisation qui affecte chaque méthode

B.3. Conclusion

Les 4 stratégies de prélèvement mono-flacon récapitulées dans le tableau ci-après permettent d'obtenir des estimations de la concentration caractéristique sur une période de temps.

La stratégie de prélèvement proportionnellement au débit instantané correspond à une intégration du flux QC par la méthode des rectangles alors que les stratégies de prélèvement proportionnellement au volume ou au débit moyen, à pas de temps fixe ou variable, correspondent à un schéma d'intégration « mixte ». Ce schéma associe une méthode des rectangles pour les concentrations avec une intégration à pas de temps fin pour les débits.

Tableau 1: récapitulation des principales stratégies de prélèvement mono-flacon

Pas de temps	Volume de prise	Stratégie
Constant	Constant	Proportionnelle au temps
	Asservi au débit moyen (ou au volume écoulé)	Proportionnelle au débit moyen
	Asservi au débit instantané	Proportionnelle au débit instantané
Asservi au volume écoulé	Constant	Proportionnelle au volume

La représentativité d'un échantillon moyen vis à vis de la concentration caractéristique d'un événement dépend de la méthode d'intégration et de la fréquence de prélèvement. Elle est mesurée par l'erreur d'échantillonnage.

Pour les stratégies proportionnelles au débit moyen ou instantané, ou au volume, l'erreur d'échantillonnage se résume à l'erreur d'intégration, liée au nombre de prélèvements réalisés pour un événement donné.

La stratégie de prélèvement proportionnellement au temps est plus simple à mettre en œuvre que les trois précédentes. Mais puisqu'elle ne prend pas en compte les débits, son erreur d'échantillonnage intègre une erreur de pondération en sus de l'erreur d'intégration. Cette erreur de pondération est égale en valeur relative au produit du coefficient de corrélation entre concentration et débit par les coefficients de variation de la concentration et du débit. Cette erreur est indépendante de la fréquence d'échantillonnage.

En pratique les erreurs de discrétisation et de pondération sont inconnues. On peut évaluer l'erreur de pondération à partir des caractéristiques des signaux de débit et de turbidité, mais nous n'avons pas trouvé de méthode pour les erreurs de discrétisation. De plus les caractéristiques des signaux varient sensiblement d'un événement à l'autre et il serait sans doute délicat de les intégrer dans un calcul d'incertitude.

En revanche les erreurs de discrétisation et de pondération et pourraient être approchées par des mesures en continu menées parallèlement aux prélèvements pour chaque événement pluvieux.

C. Étude expérimentale

C.1. Méthode

La méthode utilisée pour évaluer les erreurs d'échantillonnage est basée sur la simulation de prélèvements effectués selon deux stratégies à partir de mesures « en continu » de débit et de turbidité selon une méthode voisines de celles qui ont été mises en œuvre par (Pouzol, 2010) pour des bilans 24h sur des réseaux d'eau usée et par (Sandoval, 2017) pour les événements pluvieux.

Nous avons étudié les erreurs d'échantillonnage de la turbidité, sans la convertir en concentration, en considérant que la dynamique des variations de turbidité est représentative de celles des concentrations en MES et en polluants particulaires. Les deux stratégies d'échantillonnage mono flacon testées sont les deux stratégies le plus couramment utilisées en pratique, à savoir

- prélèvement proportionnellement au temps (stratégie T)
- prélèvement à pas de temps variable asservi au volume écoulé (stratégie V)

Chaque stratégie est définie par un paramètre-clé, qui définit l'intensité de l'échantillonnage : pas de temps pour la stratégie T et pas de volume pour la stratégie V. On a donc fait varier la valeur de ces paramètres. Pour les pas de temps les valeurs étaient indépendantes des sites. Pour les pas de volume, les valeurs ont été normalisées par le volume moyen des événements observés sur chaque site.

Pour une valeur donnée du paramètre d'échantillonnage, la période de temps couverte par les prélèvements, ou autonomie hydraulique, dépend du volume des prises. Pour simuler l'impact de pertes de prélèvements due à la saturation du flacon récepteur, on a donc fait varier le volume des prises pour un même paramètre d'échantillonnage : un volume de prise plus grand aboutit à une saturation plus rapide et à une perte d'échantillons plus fréquente. Le volume de prise est exprimé en nombre maximum de prises admissible dans le flacon récepteur.

Les valeurs retenues pour les paramètres sont indiquées dans le Tableau 2. Les cases blanches correspondent aux paramétrages les plus représentatifs des bonnes pratiques. Les cases grisées correspondant à des paramétrages complémentaires, plus éloignés des conditions usuelles, mais utiles pour préciser le domaine d'application des conclusions et pour analyser l'impact des prélèvements manqués.

Tableau 2: paramétrages utilisés dans les simulations pour des deux stratégies de prélèvements

Pas de temps (mn)	4	10	15	20	Capacité du préleveur (nb. prélèvements)	50	100	200	400
Pas de volume (% V_{moy})	1	2	5	10					

On réalise une simulation par événement pour chaque combinaison stratégie + pas de temps ou pas de volume + capacité de préleveur.

C.2. Données

Nous avons utilisé des données acquises sur cinq bassins versants drainés par des réseaux d'assainissement unitaires et séparatifs pluviaux, à savoir : *Quais*, *Clichy* et *Sucy en Brie* à Paris, et *Ecully* et *Chassieu* à Lyon par deux observatoires français en hydrologie urbaine (OPUR-Paris et OTHU-Lyon)

C.2.1. Description et équipement des sites

Les principales caractéristiques et les données de mesure en continu utilisées dans le cadre de cette étude sur les six bassins versants sont synthétisées dans le *Tableau 3*.

Tableau 3: Description et caractéristiques des bassins versants étudiés et les données de mesure en continu utilisées sur chaque site

Type principal du réseau	Sites	Localisation	Surface (ha)	Coefficient d'imperméabilisation (%)	Pente moyenne (%)
Unitaire	Quais	Paris, centre-ville	402	64	0,14
	Clichy		942	68	0,1
	Ecully	Lyon	245	15	2,7
Séparatif pluvial	Chassieu	Paris, banlieue	185	30	0,40
	Sucy		228	23	0,18

Les points de mesure sur réseau unitaire sont situés en collecteur visitable de section ovoïde à double banquettes aux sites des *Quais* et *Clichy* (Hannouche, 2012). Les deux bassins versants correspondants sont emboîtés : *Quais* est inclus dans *Clichy*. Sur le site d'*Ecully* les mesures de la qualité sont réalisées dans un canal de mesure placé en dérivation dans un bungalow et alimenté en continu par une pompe péristaltique (Metadier, 2011). La pente moyenne du réseau d'*Ecully* (2,7%) est beaucoup plus importante que celle des autres bassins versants ($\leq 0,14\%$).

Les deux sites de *Sucy* et *Chassieu* sont desservis par un réseau séparatif pluvial. Le site de *Sucy* est représentatif d'un site pavillonnaire d'un centre urbain de la banlieue parisienne (Hannouche et al. 2013) et celui de *Chassieu* est à dominante industrielle (Metadier, 2011). Le point de mesure au site de *Sucy* se situe dans un collecteur circulaire visitable tandis que les mesures du site de *Chassieu* ont été réalisées en dérivation, comme pour le site d'*Ecully*.

Les sites de *Quais*, *Clichy* et *Sucy* ont été équipés de deux turbidimètres fournissant des mesures en atténuation (*Martec-Ponselle* dans l'infrarouge à 880 nm) étalonnés à l'aide d'un étalon de Formazine (FAU : Formazin Attenuation Unit), et d'une mesure de débit (*h*, *vitesse*

par doppler). Les sites d'Ecully et Chassieu sont équipés par deux turbidimètres fournissant des mesures en néphélogéométrie (Enders- Hauser dans l'infrarouge à 880 nm), d'une mesure du débit à l'aide des mesures de hauteur et vitesse par doppler effectuée directement dans le réseau.

Enfin, chacun de sites étudiés est équipé d'un ou deux pluviomètres couvrant le bassin versant et permettant de fournir les données pluviométriques des événements pluvieux à identifier à l'exutoire de chaque bassin versant.

C.2.2. Jeux de données

Le terme de « mesurage en continu » est un peu ambigu, car si les capteurs fournissent effectivement un signal continu, celui-ci est échantillonné au fil du temps et les valeurs successives sont enregistrées à un intervalle de temps défini par l'utilisateur.

Sur les sites des Quais et Clichy (2006), le pas du temps d'enregistrement du débit et de la turbidité choisi était de 1 minute (mesure instantanée). Sur les sites Lyonnais (2004→2008) et le site de Sucy à Paris (2011→2012), ce pas du temps était de 2 minutes pour le débit et la turbidité (mesures instantanées).

Sur chaque site, les données sont validées et découpées en événements pluvieux par (Hannouche et al. 2013 ; Hannouche, 2012 ; Métadier, 2011, Lacour 2009). La validation des données de turbidité a été effectuée en utilisant les mesures redondantes obtenues par les deux turbidimètres sur chaque site selon la méthode présentée dans (Joannis et al. 2010). Les précipitations enregistrées sur les pluviomètres ont été utilisées pour détecter les jours où il y a des événements pluvieux. Pour les sites unitaires (ou séparatif s'il y a du débit par temps sec : mauvais branchements ou infiltration), les jours où la précipitation est nulle permettent de construire une chronique moyenne du débit permettant d'identifier les événements pluvieux.

En effet, le début d'un événement pluvieux est caractérisé fréquemment par une augmentation importante du débit tandis que la fin de l'événement est délimitée par un retour au débit proche de celui de la chronique moyenne de temps sec. Sur la base de ces critères, on peut identifier : (a) respectivement 75 et 88 événements sur les sites de Quais et Clichy (cf Tableau 4) (sites emboîtés), (b) environ 250 événements aux sites Lyonnais et (c) environ 120 à Sucy

Le Tableau 4 indique les caractéristiques pluviométriques événements identifiés sur les sites d'études. Ces caractéristiques ne sont pas très différentes d'un site à un autre et les événements pluvieux ont des périodes de retour inférieur à 1 an sauf pour deux événements pluvieux communs aux Quais et Clichy ($I_{\max 5} > 70$ mm/h).

Tableau 4: Caractéristiques pluviométriques des événements identifiés sur les sites d'études (d_{10} - d_{90})

Type du réseau	Site	Nombre d'événements	Lame d'eau (mm)	Durée (hh:mn)	Intensité moyenne (mm.h^{-1})	Intensité max./5 mn (mm.h^{-1})
Unitaire	Quais	75	1,2 -- 11,7	00:30 – 09:10	1,0 - 6,4	2,3 - 61,6
	Clichy	88	1,4 - 12,2	00:50- – 10:30	1,1- 6,6	2,8 - 60,3
	Ecully	239	0,7 -- 14,2	00 :30 - 13 :30	0,3 - 3,8	1,0 - 8,8
Séparatif	Chassieu	263	1,0--14,7	00:40- 13:00	0,3-2,8	1,5-17,5
	Sucy	119	1,4 -12,7	01:00- 13:00	0,6-2,7	2,4-21,6

Le Tableau 5 caractérise globalement la dynamique des signaux observés sur chaque site ainsi que la variabilité des durées et des volumes d'événements, qui est à prendre en compte pour optimiser respectivement les stratégies T et V.

Tableau 5: valeurs moyennes inter-événementielles des paramètres caractérisant les événements et susceptibles d'influencer les erreurs de pondération et de discrétisation sur les sites d'études

	CVQ Coefficient de variation moyen du débit	CVT Coefficient de variation moyen de la turbidité	R Coefficient de corrélation moyen entre débit et turbidité	R*CV T*CV Q Erreur de pondération moyenne	CV D Coefficient de variation moyen de la durée des événements	CV V Coefficient de variation moyen du volume des événements
Sucy	0,83	0,78	0,66	0,44	0,88	1,07
Quais	0,21	0,26	0,43	0,03	0,58	0,71
Clichy	0,22	0,21	0,43	0,03	0,62	0,71
Chassieu	0,98	0,73	0,55	0,40	0,84	1,80
Ecully	0,57	0,64	0,57	0,21	1,05	1,73

Les sites séparatifs de Chassieu et Sucy présentent la plus grande variabilité pour tous les paramètres, ainsi que l'erreur moyenne de pondération la plus élevée.

Les sites unitaires emboîtés de Quais et Clichy ont des dynamiques très semblables, caractérisées par une faible variabilité du débit et de la turbidité. Les résultats obtenus par les simulations sont également très semblables, voire redondants, et dans la suite nous ferons abstractions des données des Quais pour nous concentrer les résultats de Clichy.

Le site unitaire d'Ecully présente une dynamique intermédiaire entre celles des deux groupes précédents pour le débit et la turbidité, mais se distingue par une grande variabilité des caractéristiques des événements (durée et volume).

Pour tous les sites, la variabilité des durées d'événements est plus faible que celle des volumes, ce qui doit (un peu) faciliter l'optimisation des stratégies de prélèvement proportionnelles au temps.

C.3. Résultats

Précisons tout d'abord que les résultats sont en général exprimés en termes d'erreur puisqu'on dispose de signaux « continus » de turbidité et de débit, les premiers étant supposés représentatifs de variations de concentration en polluants particulaires. On peut alors calculer une valeur de référence de la masse de polluant et de la concentration caractéristique de chaque événement. Comme on ne dispose que d'une réalisation du couple stratégie-paramétrage par événement, l'écart entre la concentration obtenue en simulant les prélèvements et la concentration de référence est parfaitement déterminé et représente une erreur. La gamme de variation des erreurs sur l'ensemble des événements observés sur un site, pour une stratégie et un paramétrage donnés, permet cependant de définir et d'évaluer une incertitude. Celle-ci sera en général évaluée par l'intervalle de confiance à 95% des erreurs (IC95) ou, lorsque les effectifs sont faibles par l'intervalle de confiance à 90% (IC90). Dans certains cas on agrègera les résultats issus de plusieurs paramétrages. La gamme de variation des erreurs ne peut plus être considérée comme une incertitude, mais on utilisera encore IC90 et IC95 pour caractériser les erreurs.

La quasi-totalité des résultats présentés ci-après correspondent au volume de prise minimum, qui assure une autonomie de 400 prélèvements et minimise le nombre de prélèvements manqués. Seule l'étude de l'impact des pertes a nécessité la prise en compte de volume de prise plus importants, précisément pour provoquer ces pertes.

C.3.1. Stratégie de prélèvement proportionnellement au volume

On ne présentera que deux cas extrêmes : le site séparatif de Chassieu et le site unitaire de Clichy.

Le nombre de prélèvements diminue bien sûr lorsqu'on augmente le pas de volume, et le risque de prélèvements manqués par saturation du flacon collecteur diminue aussi. Pour un pas de volume donné, le nombre de prélèvements réalisés est très variable, et reflète la variabilité du volume des événements. Les pas de volume les plus faibles permettent d'avoir un nombre de prélèvements conséquent pour la plupart des événements, même si les événements les plus modestes en volume restent peu échantillonnés.

Le nombre médian de prélèvements est deux fois plus faible à Chassieu qu'à Clichy, pourtant les centile 95% sont deux fois plus élevés, et les pertes de prélèvements, absentes à Clichy, sont notables pour le pas de volume le plus faible et présentes pour le pas de volume suivant. Elles correspondent à la variabilité des volumes des événements qui est beaucoup plus grande à Chassieu qu'à Clichy.

C.3.1.1. Nombre de prélèvements par événement

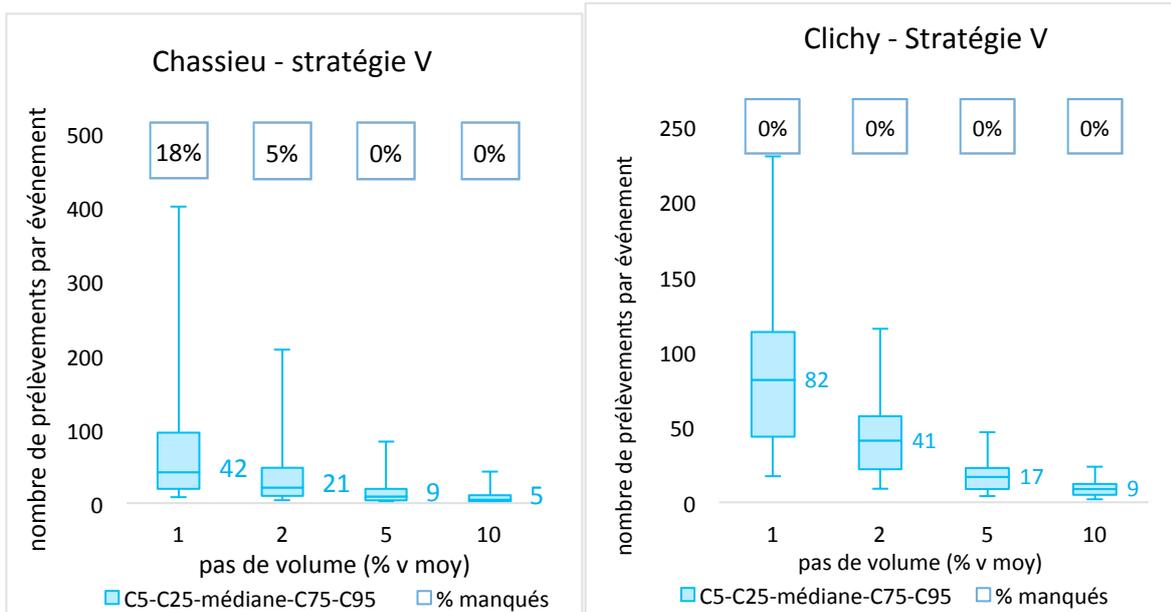


Figure 2 : nombre de prélèvements par événement en fonction du pas de volume (capacité du flacon 400 prélèvements)

C.3.1.2. Erreurs en fonction du pas de volume

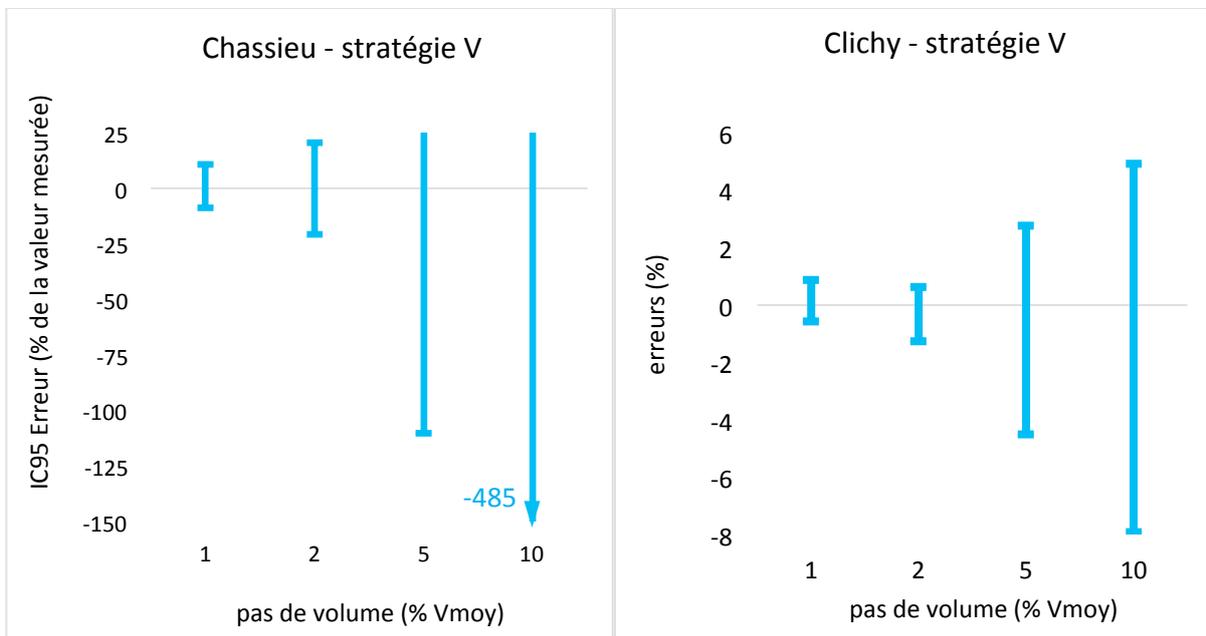


Figure 3 : IC95 des erreurs en fonction du pas de volume

La Figure 3 présente la gamme des erreurs de discrétisation selon les événements en fonction du pas de volume. Cette gamme est représentée par l'intervalle contenant, pour un pas de volume donné, 95% des erreurs observées sur 88 événements pour Clichy et 263 événements pour Chassieu. Cet « inter-centile » (IC) peut être assimilé à un intervalle de

confiance (IC aussi !), autrement dit une incertitude. Pour cela il faut considérer que la gamme des événements observés est représentative de celle qui est concernée par des évaluations de concentration caractéristique d'un site.

Sur les deux sites, les erreurs sont très variables en fonction des événements. Les erreurs beaucoup plus importantes observées à Chassieu sont à mettre en relation avec la grande variabilité des concentrations observées sur ce site séparatif. On constate une tendance à la sous-estimation, notamment sur Chassieu, sans qu'on en ait trouvé une explication. Sandoval (2017) ne l'a pas constaté sur les sites de Chassieu et Ecully, mais ses simulations incluaient d'autres sources d'incertitudes qui ont pu occulter ce biais. Si ce biais était confirmé, il impliquerait que les erreurs ont une composante systématique et ne se compensent pas totalement sur des moyennes inter-événementielles, réalisées pour déterminer une concentration caractéristique d'un site donné. Cette moyenne inter-événementielle sera donc sous-estimée elle aussi

Sur les deux sites l'incertitude augmente logiquement avec le pas de volume, puisqu'un pas de volume plus grand conduit à réaliser moins de prélèvements pour un événement donné ce qui conduit à un échantillon moins représentatif. Mais le nombre de prélèvements dépend aussi du volume de l'événement. Pour dissocier les effets du pas de volume et du volume de l'événement, nous avons étudié directement l'effet du nombre de prélèvements/heure sur l'erreur de discrétisation, tous pas de volumes confondus.

C.3.1.3. Erreurs en fonction du nombre de prélèvements/heure

L'incertitude diminue lorsque le nombre de prélèvements/heure augmente, et au-delà de 4 prélèvements/heure sur Clichy (et Sucy) et de 6 prélèvements/heure sur Chassieu (et Ecully) on obtient une erreur inférieure à 10% pour 95% des événements. Les préconisations de la norme T90-523-2 (au moins 6 prélèvements par heure) semblent donc convenir pour ce type de rejets.

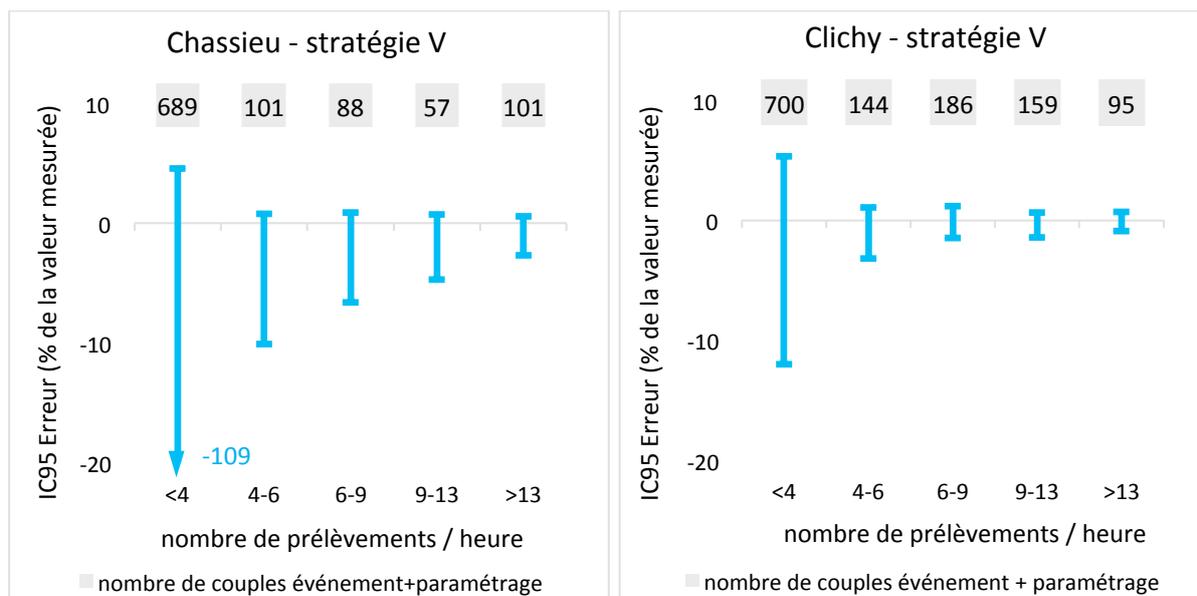


Figure 4 IC95 des erreurs en fonction du nombre de prélèvements par événement

Par curiosité nous avons comparé ces erreurs à l'incertitude affectant une moyenne arithmétique de tirages indépendants prise comme référence. Celle-ci est calculée pour chaque événement à partir de l'écart-type du signal de turbidité et du nombre de prélèvements réalisés. Comme les prélèvements réalisés avec la stratégie V ne sont pas aléatoires et leurs concentrations sont autocorrélées, on s'attend à trouver un résultat différent de l'incertitude sur un tirage aléatoire.

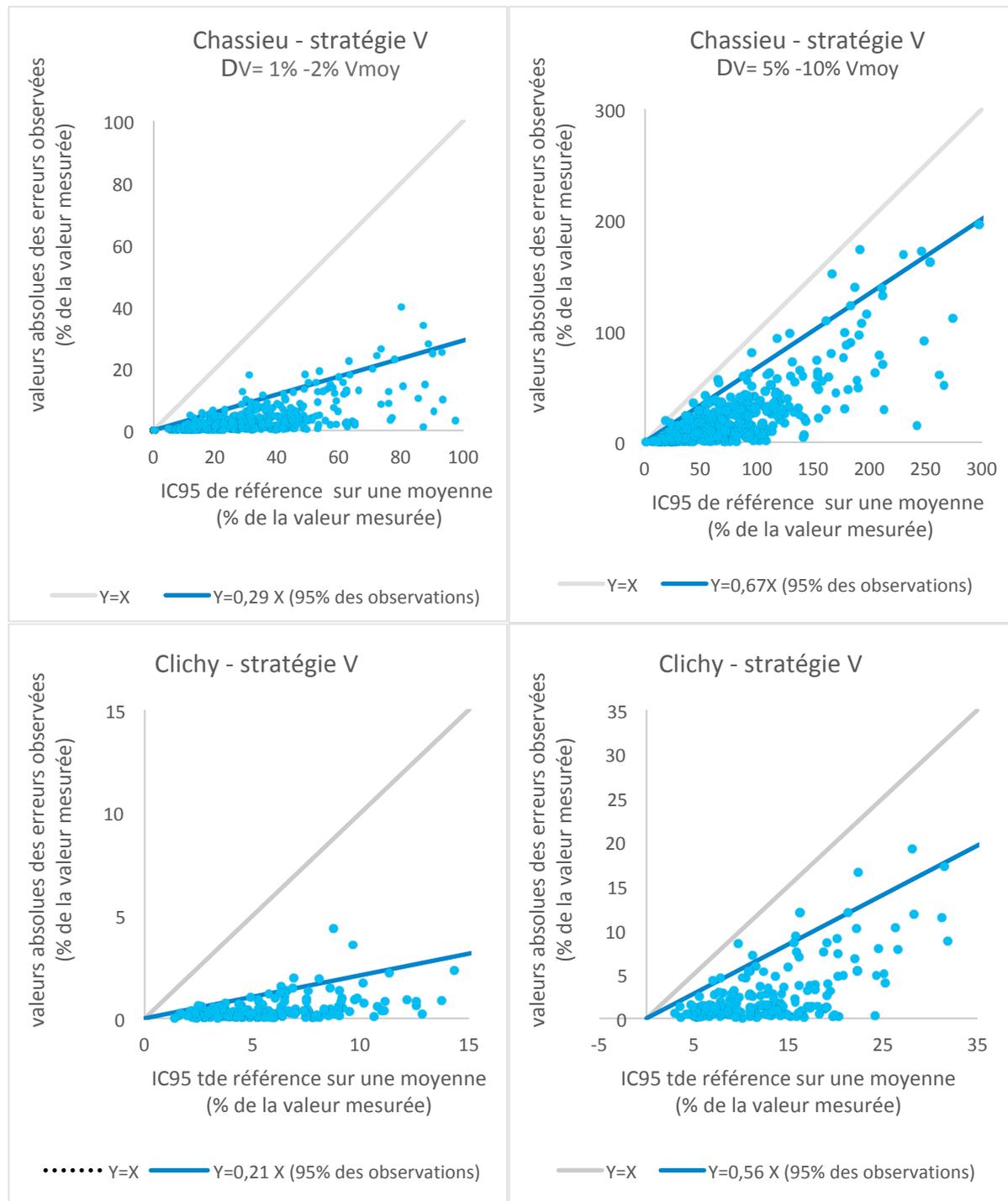


Figure 5 : Erreurs de discrétisation observées pour la stratégie V comparées aux incertitudes de référence (moyenne de tirages indépendants)

Il apparaît en effet (Figure 5) que les erreurs ont tendance à se regrouper bien à l'intérieur des intervalles de confiance à 95% pris comme référence, surtout pour les pas de volume les plus faibles. L'intervalle de confiance sur une moyenne aléatoire surévalue alors d'un facteur 3 à 5 l'intervalle de confiance des observations. Pour les pas de volume plus élevés, l'autocorrélation entre les prélèvements diminue et l'intervalle de confiance des observations se rapproche de l'intervalle de confiance sur une moyenne aléatoire, tout en restant inférieur.

Les facteurs de surévaluation sont relativement proches pour les 4 sites (Tableau 6). Ils n'en sont pas pour autant extrapolables, ne serait-ce que parce qu'ils dépendent fortement des pas de volume.

Tableau 6: rapports entre l'intervalle de confiance d'une moyenne aléatoire et l'intervalle de confiance à 95% des erreurs de discrétisations observées

	$\Delta V = 1\% - 2\%$	$\Delta V = 5\% - 10\%$
Chassieu	3,4	1,5
Clichy	4,8	1,8
Ecully	4	1,5
Sucy	4,5	1,8

Il existe une formulation théorique pour ce type d'erreur de discrétisation à pas de temps variable (Joannis & Bertrand-Krajewski, 2009). Elle fait intervenir les covariances non pas des concentrations des prélèvements successifs, mais des concentrations dans l'intervalle de temps séparant deux prélèvements : elle nécessite donc une connaissance fine du signal pour chaque événement et ne peut être calculée a posteriori que par simulation en décalant l'instant du prélèvement initial pour un événement donné. En principe on pourrait ainsi obtenir une incertitude (et non plus une erreur) de discrétisation à condition que la résolution temporelle du signal en continu soit suffisamment faible devant l'intervalle de temps entre deux prélèvements pour permettre de simuler de nombreuses valeurs de décalage. En pratique cette méthode est difficilement envisageable et l'incertitude de discrétisation n'est potentiellement accessible que pour un ensemble d'événements - c'est ce qui est réalisé dans ce travail – et non pour un événement particulier.

C.3.1.4. Impact des prélèvements manqués

Pour minimiser l'erreur de discrétisation, il faut chercher à réaliser un maximum de prélèvements pour chaque événement. Mais comme le paramétrage est fixe, cela induit un risque de dépassement du volume du flacon et donc la perte des derniers prélèvements pour les événements dont le volume est le plus important.

Nous avons donc cherché quel pouvait être l'impact sur l'erreur de prélèvements perdus en fin d'événement. Pour ce faire nous avons fait varier le pas de volume, mais aussi le volume de chaque prélèvement, afin de multiplier les cas de figures. Les résultats ont été analysés par classe de taux de perte, en se limitant à un maximum de 30%.

On a vu plus haut que l'erreur commise dépend beaucoup des événements. Nous avons donc comparé l'erreur commise sur un événement donné pour un paramétrage provoquant un taux de perte donné avec l'erreur commise sur le même événement pour un paramétrage ne provoquant pas de perte de prélèvement : ce deuxième paramétrage peut selon le cas correspondre à un pas de volume plus grand ou à un volume de prise plus faible que le premier.

Comme on ne maîtrise pas le nombre de prélèvements perdus, qui dépend pour chaque événement du paramétrage et dépasse souvent 30%, le nombre de couples événement+paramétrage avec et sans perte inclus dans chaque classe est finalement assez limité. Compte tenu de ces faibles effectifs, la gamme des erreurs a été caractérisé par un intervalle de confiance à 90%. Cet intervalle de confiance est matérialisé par une couleur pastel pour le distinguer de l'intervalle de confiance à 95% présenté précédemment.

Même si les résultats sont à prendre avec précaution du fait des faibles effectifs, on constate (Figure 5) que les pertes de prélèvement semblent avoir un impact sensible sur les erreurs, y compris pour un taux de pertes inférieur à 10%. Au-delà de 10% de pertes les erreurs deviennent relativement, voire très importantes. Si cet impact est confirmé le paramétrage d'une stratégie de prélèvement proportionnellement aux volumes serait donc à optimiser soigneusement pour éviter de manquer des prélèvements en fin d'événement.

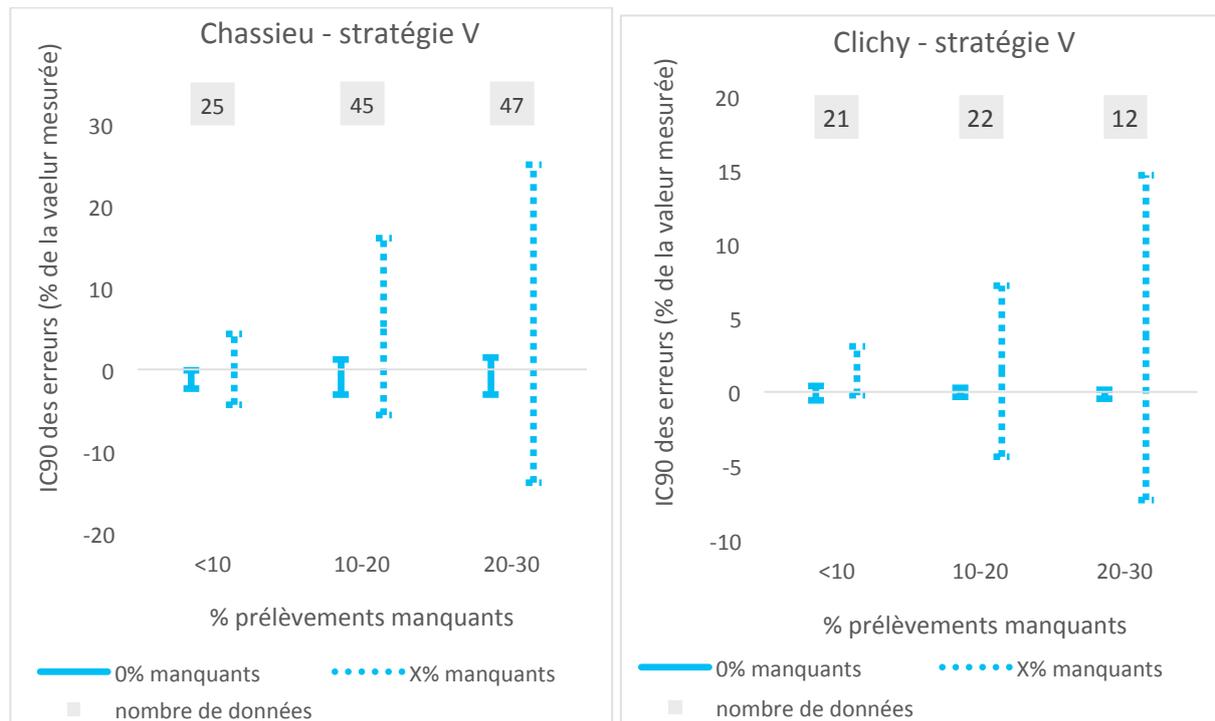


Figure 6 : IC90 des erreurs en fonction de la proportion de prélèvements manquants (pour un taux de perte donné - X%- on compare les erreurs observées pour ce taux de perte, et celles qui sont observées **pour les mêmes événements**, avec un paramétrage ne provoquant pas de perte de prélèvement – voir texte)

Le Tableau 7 synthétise les résultats pour les 4 sites et montre qu'au-delà de 10% de prélèvements manqués, les erreurs dépassent 15% pour les événements les plus pénalisants sur tous les sites sauf Clichy.

Tableau 7: bornes les plus pénalisantes des IC90 (i.e. correspondant à la plus grande erreur en valeur absolue), pour les différents sites et différents taux de pertes, comparées aux valeurs de référence observées pour les mêmes événements en absence de perte de prélèvement.

Taux de perte \ Site	<10%		10%-20%		20%-30%	
	référence	avec pertes	référence	avec pertes	référence	avec pertes
Chassieu	-2,4	4,4	-3,1	16	-3,1	25
Clichy	-1	3	-0,4	7	0,1	15
Ecully	-3	6	-3	15	-4	26
Sucy	-2,3	8	-1,6	20	-2	23

C.3.2. Stratégie de prélèvement proportionnellement au temps

C.3.2.1. Nombre de prélèvement par événement

Il est difficile de choisir a priori des gammes de valeurs permettant de faire correspondre les deux types de paramétrage et les nombres de prélèvements (Figure 6 et Figure 2). On constate par exemple à Chassieu un nombre de prélèvements deux fois plus important pour le pas de temps le plus court que pour le pas de volume le plus petit. Malgré ce nombre de prélèvements plus important, les pertes de prélèvements restent limitées du fait de la moindre variabilité des durées d'événements comparée à celle qui affecte les volumes la stratégie T, et on n'observe pas de perte de prélèvements. D'une manière plus générale, les pas de temps les plus longs fournissent davantage de prélèvements que les pas de volumes les plus importants pour les paramétrages choisis.

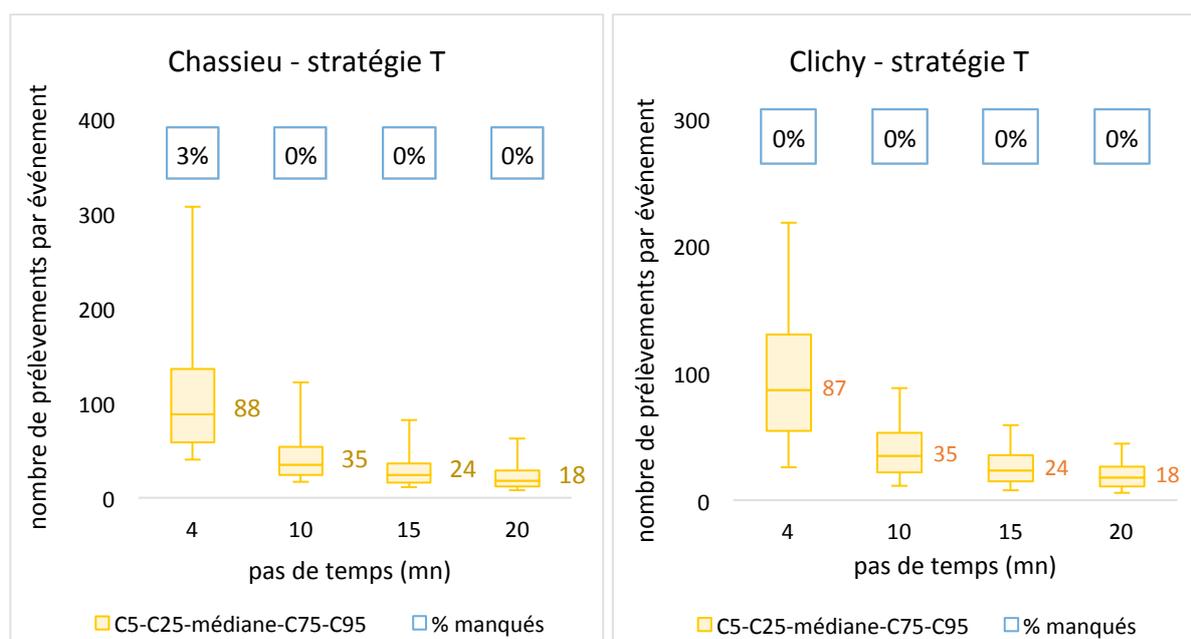


Figure 7 : nombre de prélèvements par événement en fonction du pas de temps

C.3.2.2. Erreurs en fonction du pas de temps

Les erreurs observées pour la stratégie T (Figure 7) incluent les erreurs de pondération et les erreurs de discrétisation. Comme pour la stratégie V elles sont toujours très variables en fonction des événements surtout à Chassieu, et sont beaucoup plus importantes (en valeurs absolues) pour Chassieu que pour Clichy. Elles ne dépendent pas du pas de temps et donc du nombre de prélèvements réalisé sur un événement donné. Ce comportement, caractéristique des erreurs de pondération, laisse supposer qu'elles sont prépondérantes par rapport aux erreurs de discrétisation pour tous les pas de temps. Dans ce cas les erreurs, en général négatives, correspondraient à des corrélations positives entre la turbidité et le débit pour la plupart des événements.

Comme pour la stratégie V on constate que les erreurs sont pour la plupart négatives. Mais dans ce cas le biais a également été observé par Sandoval (2017) et il est potentiellement explicable : il signifierait que la corrélation entre débit et concentration est positive pour la plupart des événements. La section suivante va permettre de vérifier cette hypothèse.

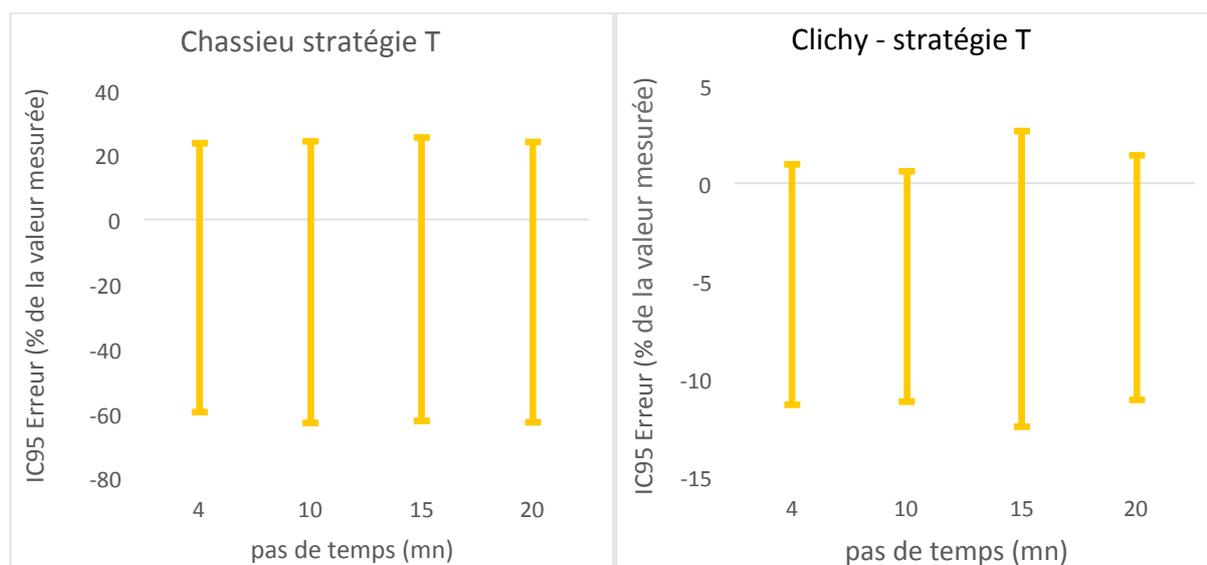


Figure 8 : IC95 des erreurs en fonction du pas de temps

C.3.2.3. Comparaison des erreurs de pondération avec leurs valeurs théoriques

L'erreur de pondération théorique est obtenue en multipliant pour chaque événement les coefficients de variation du débit et de la turbidité par le coefficient de corrélation entre débit et turbidité.

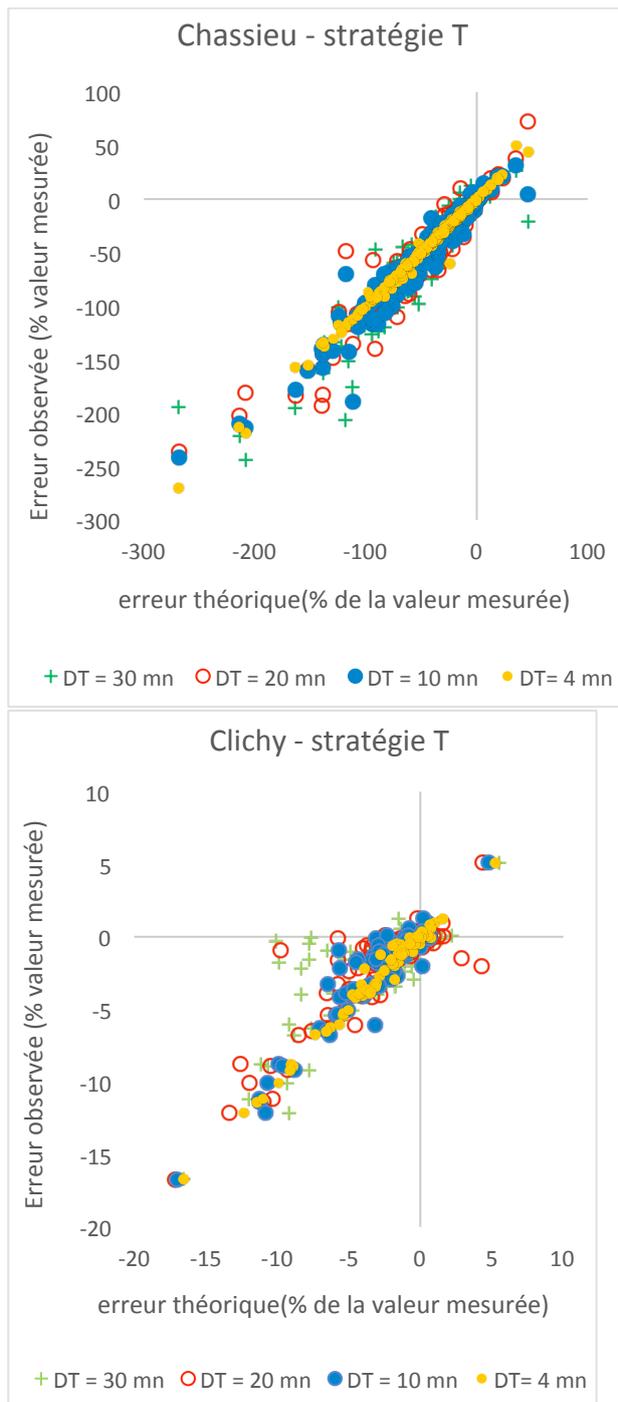


Figure 9 : comparaison des erreurs observées pour la stratégie T avec l'erreur théorique de pondération

L'erreur de pondération est effectivement presque toujours négative (Figure 8), et bien plus marquée sur Chassieu que sur Clichy comme en témoignent les échelles des ordonnées des deux graphiques. Les valeurs observées correspondent bien aux valeurs théoriques, surtout pour les pas de temps les plus courts. En effet l'erreur observée inclut non seulement l'erreur de pondération, mais aussi l'erreur de discrétisation, qui devient sensible lorsque le pas de temps augmente. A contrario, la bonne correspondance entre erreur de pondération théorique et erreur observée pour les pas de temps les plus courts confirme que l'erreur de discrétisation est négligeable devant l'erreur de pondération pour ces pas de temps.

C.3.3. Comparaison des deux stratégies

Il pourrait a priori être délicat de comparer les performances d'une stratégie à pas de temps fixe et d'une stratégie à pas de temps variable car il faudrait comparer les résultats obtenus sur les mêmes événements dans des conditions comparables. Les « conditions comparables » pourraient être caractérisées par l'effort d'échantillonnage, traduit par le nombre de prélèvements réalisés au court d'un événement. Mais comme les erreurs de pondération associées à la stratégie T sont insensibles à l'effort d'échantillonnage, et prépondérantes devant l'erreur de discrétisation dès lors que l'effort d'échantillonnage est raisonnable (pas de temps courts), nous avons choisi de comparer globalement les deux stratégies pour l'ensemble des événements.

Les valeurs globales (Figure 9) montrent que la stratégie V paramétrée avec des pas de volume suffisamment faibles produit toujours des erreurs sensiblement réduites (d'un facteur compris entre 5 et 20) par rapport la stratégie T. De plus le biais négatif des moyennes temporelles est confirmé pour tous les sites, et correspond à une corrélation positive entre la concentration et le débit.

Les résultats obtenus avec la stratégie V pour Chassieu et Ecully montrent que même avec un pas de volume réduit, les erreurs de discrétisation peuvent dépasser 10%. Ces valeurs élevées sont certainement dues aux taux importants de perte de prélèvement observé sur ces sites pour le pas de volume le plus faible (respectivement 18% et 17%). On observe encore des pertes, quoique beaucoup limitées pour le pas de volume de 2% du volume moyen : respectivement 5% et 4% de prélèvements manqués. Ces résultats confirment la sensibilité de la stratégie V à son paramétrage et la difficulté d'optimiser celui-ci.

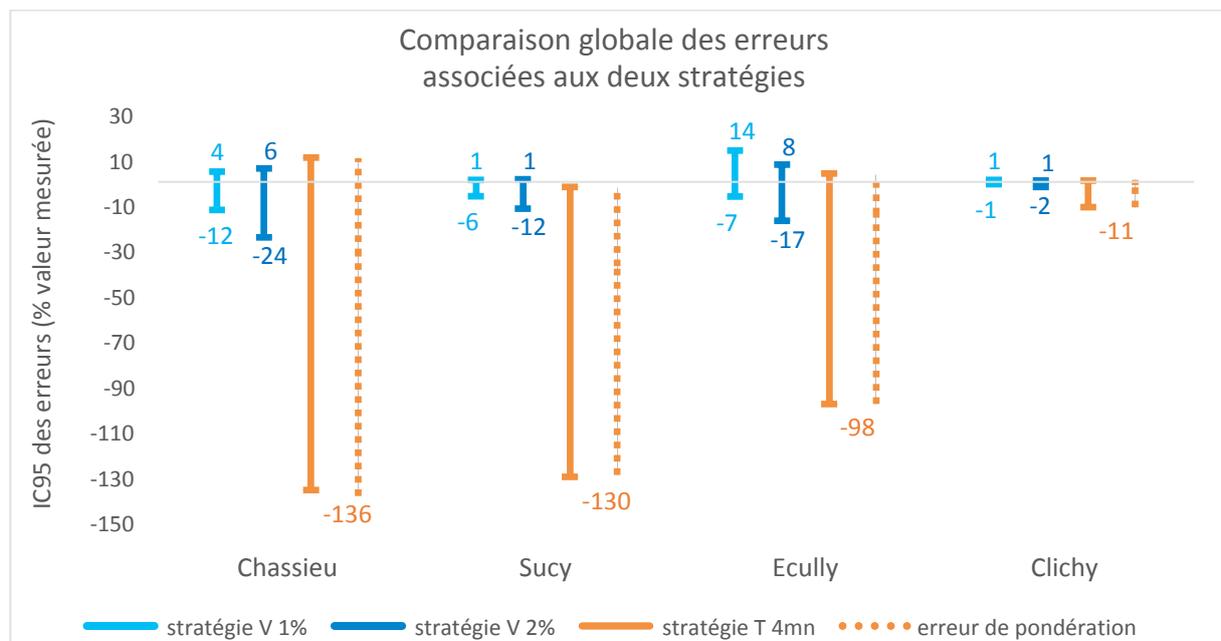


Figure 10 : Comparaison globale pour les mêmes événements des IC95 des erreurs associées aux deux stratégies

C.3.4. Discussion : ces résultats sont-ils transposables à d'autres sites ?

Cette étude repose sur une modélisation du processus de prélèvement basée sur trois hypothèses :

- 1) Les variations dans le temps des concentrations en polluants, au moins ceux qui sont sous forme particulaire, ont les mêmes caractéristiques (amplitude par rapport à la moyenne, autocorrélation, corrélation avec le débit...) que le signal de turbidité
- 2) On dispose pour chaque événement d'une mesure en quasi-continu de la turbidité qui permet de simuler les valeurs que l'on obtiendrait sur des échantillons issus d'un préleveur piloté selon différentes stratégies avec plusieurs paramétrages.
- 3) L'impact des incertitudes de mesure de la turbidité et du débit sur les erreurs de discrétisation et de pondération est a priori faible : il n'est donc pas nécessaire de les prendre en compte

De plus les valeurs numériques des erreurs fournies par les simulations correspondent à 5 sites, qui présentent certes de caractéristiques très différentes, mais ne peuvent pas prétendre pour autant couvrir toute la diversité des contextes possibles.

Les résultats obtenus montrent d'ailleurs une grande variabilité des incertitudes de pondération et de discrétisation entre les sites utilisés dans notre étude, ce qui incite à la prudence pour la transposition des résultats à un autre site : à quel site étudié pourrait-on l'assimiler ? faut-il systématiquement se placer dans le pire des cas ?

Rappelons que tout au long de ce rapport nous avons essentiellement parlé d'erreurs définies comme la différence entre la turbidité obtenue en simulant les prélèvements et la turbidité de référence calculées à partir des mesures en continu. Les variations observées de cette erreur en fonction des événements permettent définir une incertitude, caractérisant selon le cas l'erreur de discrétisation ou l'erreur de pondération, par exemple par son centile 95.

Pour l'incertitude de pondération, spécifique de la stratégie T, le résultat est assez clair : on dispose d'une gamme de valeurs possibles en fonction des sites et des événements qui montre que l'incertitude de pondération peut être très élevée pour certains sites ou certains événements. En l'absence d'information sur la dynamique des débits et des concentrations sur un site donné, il semble risqué d'utiliser une stratégie de prélèvement proportionnellement au temps pouvant conduire à des erreurs supérieures à 100 % de la valeur mesurée !

L'incertitude de discrétisation est quant à elle souvent négligeable devant l'incertitude de pondération pour la stratégie T. En revanche c'est la seule composante de l'incertitude d'échantillonnage pour la stratégie V. Elle est relativement limitée lorsqu'un échantillon est composé à partir d'un nombre de prélèvement suffisant (De 6 à 10 prélèvements/ heure avec la stratégie V pour les sites testés) mais elle semble assez sensible aux pertes de prélèvements. L'optimisation du paramétrage doit donc concilier un nombre maximum de prélèvements tout en assurant une couverture totale (i.e. ne pas manquer les derniers

prélèvements par saturation de la capacité du préleveur). Il semble néanmoins qu'il faille surtout viser une couverture complète des événements les plus importants en volume, qui sont le plus souvent importants également en masse de polluants, quitte à avoir une moins bonne estimation de la concentration moyenne des événements moins importants.

Si on disposait d'une formulation théorique permettant d'évaluer a priori l'incertitude d'échantillonnage pour un événement donné ou pour l'ensemble des événements observables sur un site donné, la question de la transposition ne se poserait pas. Il existe effectivement de telles formulations pour un événement donné, très simple pour l'erreur de pondération associée à la stratégie T, plus complexe pour l'incertitude de discrétisation associée aux stratégies T et V. Mais elles nécessitent d'avoir enregistré en continu avant les campagnes de prélèvement le débit et de turbidité (ou d'autres paramètres de qualité comme la conductivité) pour un nombre important d'événements et de les exploiter selon les méthodes présentées dans cet article.

A défaut, le caractère intégrateur du prélèvement mono flacon suggère de compléter les prélèvements avec une mesure en continu de débit et de turbidité (ou d'autres paramètres de qualité comme la conductivité). Celles-ci pourraient être utilisées en validation, en évaluant a posteriori les erreurs affectant chaque événement et en éliminant ceux pour lesquels l'erreur sur la turbidité moyenne serait trop importante. On pourrait aussi envisager de corriger les erreurs, mais cette correction ne serait que partielle car elle dépend de l'hypothèse de représentativité des mesures en continu. De plus l'incertitude après correction ne pourrait plus être évaluée.

D. Conclusion

- Les opérations d'échantillonnage des concentrations pour constituer un échantillon moyen représentatif du flux écoulé lors d'un événement pluvieux peuvent induire deux types d'erreurs :
 - une erreur de discrétisation, qui affecte toutes les stratégies de prélèvement et dépend du nombre d'échantillons prélevés ;
 - une erreur de pondération, présente uniquement si la stratégie de prélèvement n'est pas asservie au débit. L'erreur de pondération ne dépend pas du nombre d'échantillons prélevés, et n'est notable que si le débit et la concentration sont corrélés (positivement ou négativement).
- La stratégie de prélèvement proportionnellement au volume est recommandée pour évaluer la concentration moyenne et la masse de polluant associées à un événement pluvieux
- Cette stratégie doit être paramétrée pour éviter les pertes de prélèvements en fin d'événements, tout en assurant 6 à 10 prélèvements par heure. Un pas de volume de 1% à 2% du volume moyen événementiel a donné d'assez bons résultats pour les événements étudiés lors de nos simulations (95% des erreurs inférieures à 15% de la valeur mesurée pour un site donné).
- L'erreur de discrétisation de la stratégie de prélèvement proportionnellement au volume est très variable en fonction des sites et des événements, mais aussi du paramétrage de la méthode (pas de volume, capacité du flacon récepteur).
- L'incertitude de discrétisation est bien inférieure à l'incertitude d'une moyenne arithmétique de prélèvements aléatoires, mais elle est impossible à évaluer a priori.
- Des mesures en continu de qualité menées parallèlement aux prélèvements pourraient constituer un apport intéressant pour évaluer la représentativité des prélèvements et les erreurs associées à un événement donné, voire pour corriger en partie ces erreurs
- L'erreur de pondération associée à la stratégie de prélèvement proportionnellement au temps ne dépend pas du paramétrage (valeur du pas de temps) mais uniquement des caractéristiques des signaux de débit et de turbidité. D'après nos simulations, elle est prépondérante devant l'erreur de discrétisation pour les pas de temps les plus faibles (4 et 10 mn), et peut être très importante pour les événements présentant une corrélation notable entre le débit et la concentration. Le principal avantage de la stratégie de prélèvement proportionnellement au temps est de ne pas nécessiter la présence d'un débitmètre, mais cet avantage ne semble pas suffisant pour compenser le surcroît d'erreur dès lors que l'on veut convertir les concentrations en masse.

E. Références

.AFNOR (2008) FD T90-523-2 Qualité de l'eau - Guide de prélèvement pour le suivi de qualité des eaux dans l'environnement - Partie 2 : prélèvement d'eau résiduaire

Hannouche, A., Chebbo, G., Ruban, G., Tassin, B., Lemaire, B.J., Joannis, C. 2011. Relationship between turbidity and total suspended solids concentration within a combined sewer system. *Water Science and Technology*, 64(12), 2445-2452.

Hannouche, A.: 2012, 'Analysis of solid transport in combined sewer network during wet weather: exploitation of data acquired by the French observatories in urban hydrology, PhD Thesis, Université Paris-Est-France, 485 pp.(In French).'

Hannouche A., Chebbo G., Joannis C., J., Gasperi, Gromaire M.-C., Moilleron R., Barraud S., Ruban V. (2017), Stochastic evaluation of annual micropollutant loads and their uncertainties in separate storm sewers, *Environ Sci Pollut Res*, 24 (36) 28205-28219.
<https://doi.org/10.1007/s11356-017-0384-5>

Joannis, C., Ruban, G., Aumond, M., Bertrand-Krajewski, J. L., Battaglia, P., Lacour, C., Saad, M. & Chebbo, G.: 2010, 'Mise en œuvre de capteurs de turbidité en collecteurs d'assainissement', *Techniques Sciences Méthodes (TSM) numéros 1/2 - 2010*.

Joannis, C., Bertrand-Krajewski, J. L., 2009, Incertitudes sur un mesurande défini comme une valeur intégrée d'un signal continu discrétisé en fonction du temps - Application aux mesures hydrologiques enregistrées in situ, *La Houille Blanche*, 3, 82-91

Lacour, C.: 2009, 'Apport de la mesure en continu pour la gestion de la qualité des effluents de temps de pluie en réseau d'assainissement', thèse de doctorat, CERREVE, UNIVERSITE PARIS-EST, 310 pp.

Métadier, M.: 2011, 'Traitement et analyse de séries chronologiques continues de turbidité pour la formulation et le test de modèles des rejets urbains par temps de pluie. Thèse de doctorat au LGCI, INSA de Lyon. 408 pages'.

Sandoval S., 2017, Révision des modèles conceptuels de qualité des eaux pluviales sur un grand bassin versant urbain : mesures en continu, incertitudes sur les données et les modèles Thèse de doctorat au LDEEP, INSA de Lyon. 122 pages + annexes

Pouzol T., 2010, Optimisation des stratégies d'échantillonnage des eaux résiduaires industrielles et urbaines, Agence de l'eau Rhône Méditerranée et Corse- INSA de Lyon 26 pp