

Comparaison des incertitudes sur la concentration moyenne d'un événement pluvieux induites par des stratégies de prélèvement proportionnellement au temps ou au volume

Uncertainties on a pollutant mean concentration for a rain event implied by sampling strategies controlled either by flowrate or by elapsed time

C. Joannis^a, A. Hannouche^b, G. Chebbo^{b,c}

^a LEE IFSTTAR – Centre de Nantes- Département géotechnique eau et risques - route de Bouaye CS4 – 44344 Bouguenais cedex (claudio.joannis@ifsttar.fr);

^b LEESU, UMR MA 102, Ecole des Ponts, AgroParisTech, UPEC, UPE, Champs-sur-Marne, France (chebbo@leesu.enpc.fr);

^c Faculty of Engineering III, Lebanese University, Hadath – Lebanon.

Résumé

La concentration moyenne d'un polluant donné lors d'un événement pluvieux est définie comme le rapport de la masse de polluant au volume d'eau correspondant. Elle est classiquement déterminée à partir de l'analyse d'un échantillon moyen, obtenu en cumulant plusieurs prélèvements répartis sur la durée de l'événement. La représentativité de cet échantillon moyen dépend de la stratégie de pilotage du préleveur. Cette étude a pour objectif d'évaluer à partir de données expérimentales les erreurs commises sur la concentration moyenne d'un événement associées à chacune des deux stratégies les plus courantes et d'en déduire des intervalles de confiance pour l'ensemble des événements.

Elle 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 au temps ou 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, tandis que l'asservissement au temps implique une erreur de pondération qui peut être considérable sur certains sites et certains événements. En revanche il est très délicat voire impossible d'évaluer a priori les incertitudes associées à chaque stratégie pour un site donné si on ne dispose pas de mesures en continu du débit et de pollution associés aux prélèvements.

Mots-clés : échantillonnage, prélèvement, eaux résiduaires, erreur, incertitude,

Abstract

The mean concentration of a given pollutant for a rain event can be defined as the ratio of the pollutant mass to the corresponding water volume. It is usually derived by analysing a compound sample cumulating several sub-samples dispatched along the duration of the event to provide a mean concentration value. The representativeness of the compound sample depends on the sampling strategy used to control the automatic sampler. This study uses experimental data to assess the errors on the mean concentration of a given event implied by

two usual sampling strategies. Then uncertainty ranges for all events are derived and compared.

It relies on discharge and turbidity recordings available on separate and combined drainage systems included in URBIS network Two sampling strategies controlled either by cumulated discharge or by elapsed time were simulated for several hundred rain events, with different parameter set-ups. Results clearly show that the control by cumulated volume provides a far more representative sample than the control by elapsed time. Indeed the latter implies a weighting error, which may be huge for some sites and some events. Unfortunately it is quite difficult if not impossible to assess the uncertainties implied by both strategies for a particular site unless flowrate and pollution are continuously monitored while sampling campaigns are performed

Key-words : sampling, wastewater, error, uncertainty

1 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. Les procédures correspondantes sont détaillées dans la norme AFNOR pour ce qui concerne les eaux résiduaires (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).

Ces opérations sont lourdes et les échantillons, en nombre nécessairement limité, doivent être prélevés à bon escient. De plus, pour limiter le nombre d'analyses tout en assurant une représentativité acceptable, un échantillon sera constitué de plusieurs prélèvements échelonnés sur la période à caractériser.

On s'intéresse le plus souvent à une concentration moyenne à l'échelle d'une journée ou d'un événement pluvieux, en général pour calculer une masse ou un flux en multipliant cette concentration moyenne par le volume correspondant. Pour que ce calcul de flux n'introduise pas de biais, il faut que la concentration moyenne intègre une *pondération par les débits*. On parle alors de concentration moyenne (d'un événement et d'un site).

En asservissant le préleveur à un débitmètre on peut constituer un échantillon unique, issu du mélange de plusieurs prélèvements, dont la concentration est représentative de la concentration moyenne. On peut aussi s'affranchir de la pondération par le débit et constituer un échantillon moyen à partir de prélèvements à pas de temps fixe : la concentration obtenue sera représentative de la moyenne arithmétique des concentrations.

La norme FD T90-523-2 évoque en une phrase la possibilité d'asservir les prélèvements au temps ou au volume écoulé et mentionne sans justification particulière que « dans le cas d'un prélèvement automatique, le prélèvement asservi à une mesure du débit est un mode d'échantillonnage satisfaisant » et que « 6 cycles de prélèvement par heure effective de rejet sont recommandés ».

Dans le cadre du programme OPUR 5 et du projet ANR Mentor il a semblé intéressant d'évaluer et de comparer les incertitudes sur les valeurs de concentration moyenne obtenues avec les deux stratégies de prélèvement les plus courantes.

2 Matériel et méthodes

2.1 Méthode

La méthode utilisée pour évaluer les erreurs d'échantillonnage est basée sur la simulation de prélèvements effectuées selon deux stratégies à partir de mesures en continu de débit et de turbidité selon une méthode voisine 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 chercher à la convertir en concentration, mais en considérant que la dynamique des variations de turbidité est représentative de celles des concentrations en MES et en polluants particulaires. La turbidité moyenne reconstituée \hat{T}_e à partir des prélèvements simulés a donc été comparée à une valeur de référence \tilde{T}_e obtenue à partir des mesures en continu

$$\tilde{T}_e = \frac{\sum_{1/2}^{n-1/2} (Q_i T_i + Q_{i-1} T_{i-1}) \Delta t_i}{\sum_{1/2}^{n-1/2} (Q_i + Q_{i-1}) \Delta t_i} \quad \text{Eq. 1}$$

Avec n : nombre de valeurs enregistrées pour un événement donné ;
 Q_i, T_i : valeurs de débit et de turbidité enregistrées à l'instant i ;
 Δt_i : intervalle de temps entre deux enregistrements.

\tilde{T}_e est une moyenne *pondérée par les débits* que l'on appellera turbidité moyenne de l'événement e . Transposée en concentration, la turbidité moyenne d'un événement permet de calculer une masse, en multipliant la concentration moyenne par le volume correspondant. La notion de concentration moyenne prend tout son intérêt si on parvient à la matérialiser dans un échantillon composite. On peut alors calculer une masse pour tous les polluants analysés sur cet échantillon sans passer par une mesure en continu.

La turbidité moyenne d'un événement diffère de la moyenne temporelle \overline{T}_e sur la durée de l'événement. Celle-ci s'exprime comme :

$$\overline{T}_e = \frac{\sum_{1/2}^{n-1/2} (T_i + T_{i-1}) \Delta t_i}{\sum_{1/2}^{n-1/2} \Delta t_i} \quad \text{Eq. 2}$$

L'écart entre \tilde{T}_e et \hat{T}_e représente l'erreur d'échantillonnage pour l'événement considéré.

En fonction du type de stratégie de prélèvement mis en œuvre, l'erreur d'échantillonnage peut inclure deux composantes :

- 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 *discrétisation* est due à la représentation simplifiée fournie par les prélèvements (cf. Figure 1). On peut exprimer analytiquement l'incertitude correspondante, mais les paramètres statistiques nécessaires à son calcul (variances et covariances à l'échelle du pas de temps ou du pas de volume) ne sont pas facilement accessibles (Joannis & Bertrand-Krajewski, 2009). Cependant l'erreur est clairement liée au nombre de prélèvements, et diminue lorsque le nombre de prélèvements augmente

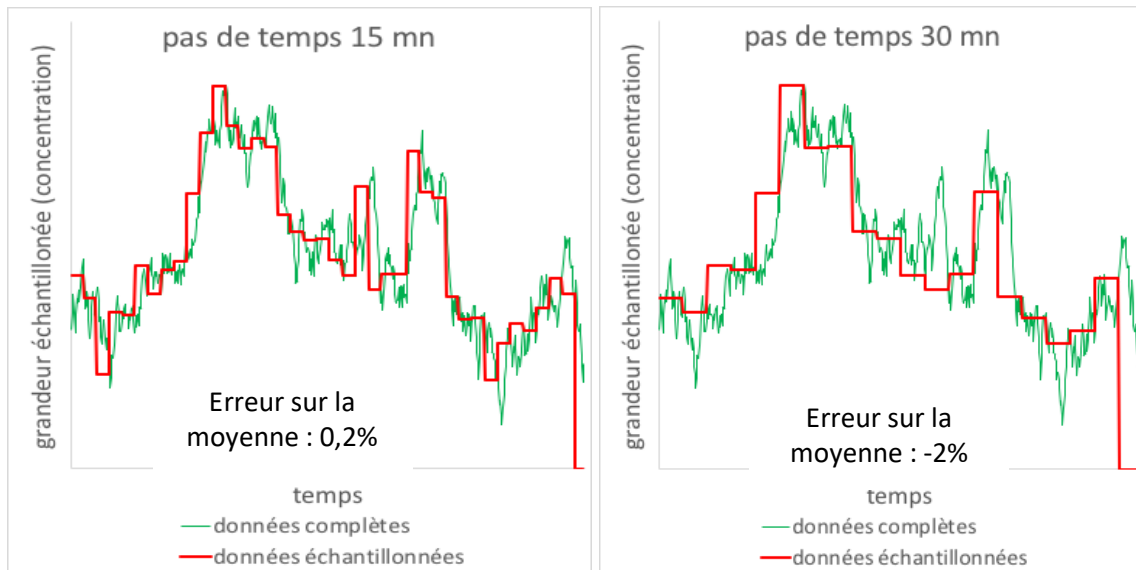


Figure 1 : Illustration de l'erreur de discrétisation pour un échantillon à pas de temps constant

L'erreur de *pondération* est due au fait que la turbidité moyenne d'un événement \tilde{T}_e diffère de la moyenne temporelle \overline{T}_e sur la durée de l'événement.

Celle-ci s'exprime comme :

$$\overline{T}_e = \frac{\sum_1^{n-1} (T_i + T_{i-1}) \Delta t_i}{\sum_1^n \Delta t_i} \quad \text{Eq. 3}$$

L'erreur de pondération est définie comme $\overline{T}_e - \tilde{T}_e$.

On peut démontrer cette erreur ne dépend pas du nombre d'échantillons prélevés, et n'est notable que si le débit et la turbidité sont corrélés (positivement ou négativement).

$$\frac{\overline{T}_e - \tilde{T}_e}{\overline{T}_e} = -\rho_{Q_e, C_e} \frac{\sigma(Q_e)}{\overline{Q}_e} \frac{\sigma(T_e)}{\overline{T}_e} \quad \text{Eq. 4}$$

\overline{Q}_e et \overline{T}_e sont les moyennes événementielles du débit et de la turbidité ;
 $\sigma(Q_e)$ et $\sigma(T_e)$ sont les écarts-types événementiels du débit et de la turbidité ;
 ρ_{Q_e, C_e} est le coefficient de corrélation entre le débit et la turbidité.

Bien qu'on dispose d'une expression analytique de l'erreur de pondération associée à la stratégie de prélèvement proportionnellement au temps, les simulations restent indispensables pour évaluer l'erreur globale d'échantillonnage, et pour aborder des aspects pratiques tels que les risques de saturation de la capacité de stockage du préleveur.

2.2 Données expérimentales

L'étude a été menée sur les données issues de 4 sites des observatoires OPUR et OTHU pour lesquels des enregistrements de débit et de turbidité étaient disponibles (Hannouche, 2012, Métadier, 2011). Les caractéristiques des sites et des jeux de données utilisées sont résumées dans le Tableau 1.

Tableau 1: Caractéristiques pluviométriques des événements identifiés sur les sites d'études (d_{10} - d_{90})

Type du réseau	Site	Superficie (ha)	Nombre d'événements	Lame d'eau (mm)	Intensité moyenne (mm.h^{-1})	Intensité max./5 mn (mm.h^{-1})
Unitaire	Clichy	942	88	1,4 - 12,2	1,1- 6,6	2,8 - 60,3
	Ecully	245	239	0,7 -- 14,2	0,3 - 3,8	1,0 - 8,8
Séparatif	Chassieu	185	263	1,0--14,7	0,3-2,8	1,5-17,5
	Sucy	228	119	1,4 -12,7	0,6-2,7	2,4-21,6

Le Tableau 2 détaille quelques caractéristiques moyennes de signaux enregistrés sur les différents sites et susceptibles d'influencer les erreurs d'échantillonnage. 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. Le site unitaire de Clichy est caractérisé par une faible variabilité du débit et de la turbidité. 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). Dans la suite, la plupart des résultats seront présentés uniquement pour les deux sites « extrêmes » Chassieu et Clichy.

Tableau 2: valeurs moyennes inter-événementielles des paramètres caractérisant les événements

	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 (prélèvements proportionnellement au temps)	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
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

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 pour éviter la saturation de la capacité du préleveur.

Simulations

L'application de deux stratégies d'échantillonnage a été simulée sur tous les événements de des 4 sites et les valeurs de turbidité moyennes ont été comparées à la valeur de référence calculée à partir de mesures en continu. Les deux stratégies réalisent respectivement un échantillonnage proportionnellement au temps (stratégie T) et au volume écoulé (stratégie V), et pour chacune plusieurs valeurs de paramètre (selon le cas pas de temps ou pas de volume) ont été testées.

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 V_{moy} , 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 de la capacité du préleveur, exprimée en nombre de prélèvements. Pour simuler l'impact de pertes de prélèvements due à la saturation du flacon récepteur, on a donc fait varier la capacité du préleveur pour un même paramètre d'échantillonnage.

Les valeurs retenues pour les paramètres sont indiquées dans le Tableau 3. 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 3: 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.

3 Résultats

Précisons tout d'abord que la plupart des résultats sont 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 turbidité moyenne 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*. Nous l'avons exprimée en % de la valeur mesurée.

La gamme de variation des erreurs sur l'ensemble des événements observés sur un site, pour une stratégie et un paramétrage donnés est représentée sous forme d'un intercentile contenant les valeurs observées pour X% des événements (ICX). X sera en général égal à 95 (IC95) sauf quand le nombre d'événements analysés est trop faible : on prendra alors X=90. On peut assimiler IC95 à un intervalle de confiance, autrement dit une incertitude, si on considère que la gamme des événements observés est représentative de la gamme visée pour évaluer la concentration caractéristique d'un site.

Les résultats présentés correspondent à une capacité du préleveur de 400 prélèvements, sauf pour l'étude de l'impact des prélèvements manqués.

On ne présentera en général que les résultats obtenus sur les deux cas extrêmes : le site séparatif de Chassieu et le site unitaire de Clichy.

3.1 Stratégie de prélèvement proportionnellement au volume

3.1.1 Nombre de prélèvements par événement

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.

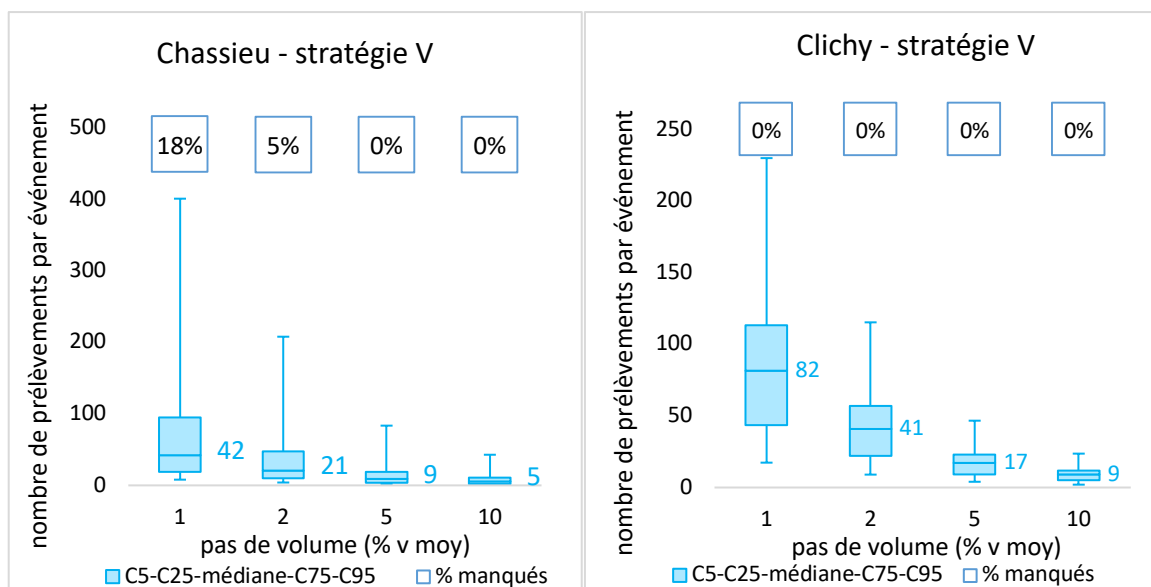


Figure 2 : nombre de prélèvements par événement en fonction du pas de volume (capacité du flacon 400 prélèvements)

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.

3.1.2 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 sous forme d'un intercentile contenant 95% des événements.

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

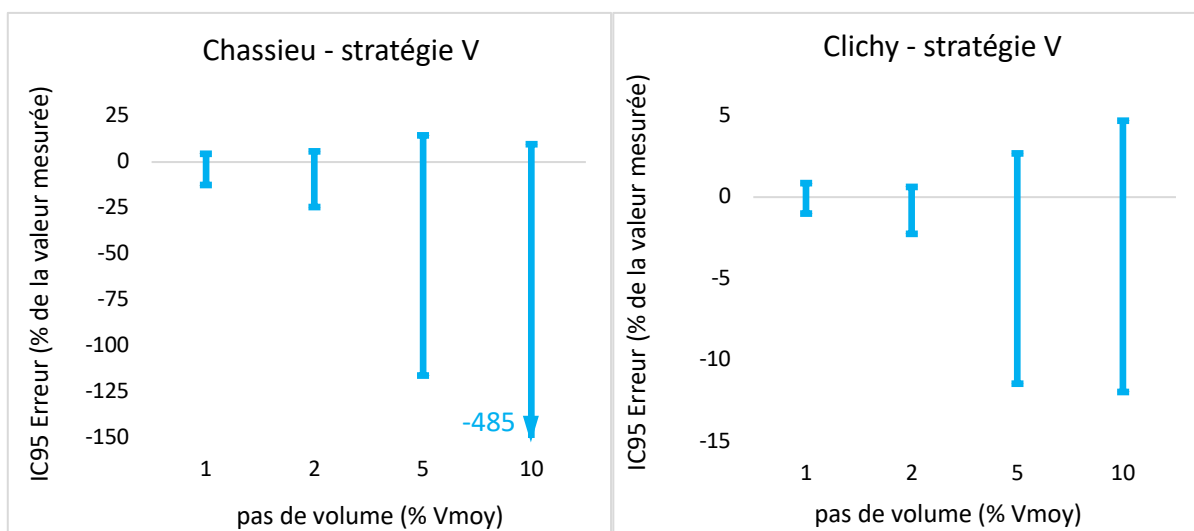


Figure 3 : IC95 des erreurs en fonction du pas de volume

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.

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.

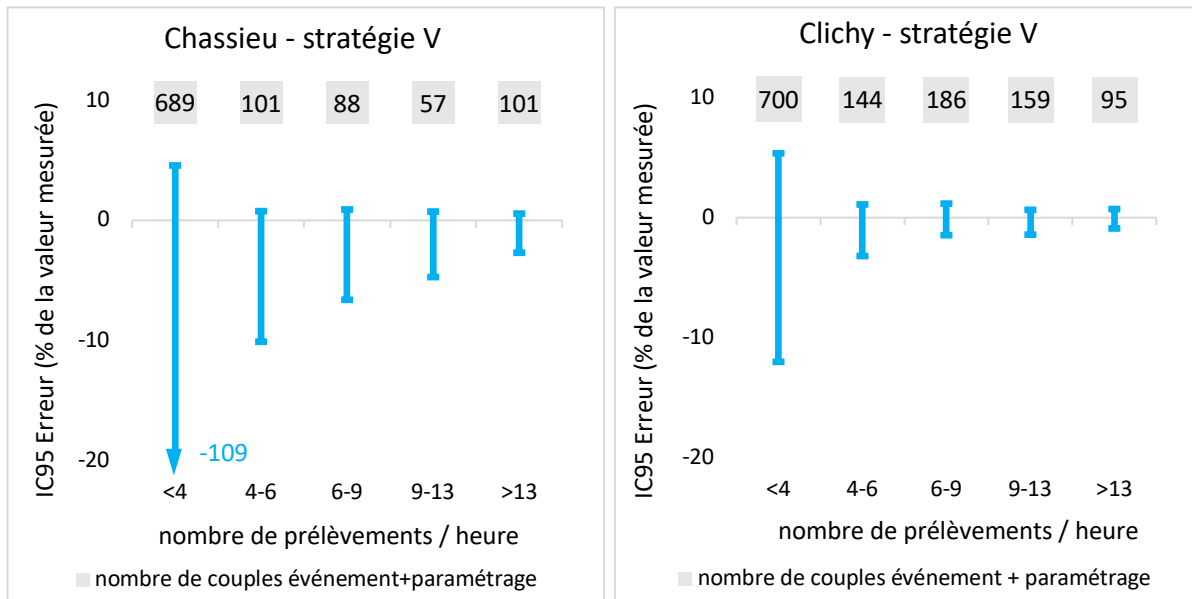


Figure 4 IC95 des erreurs en fonction du nombre de prélèvements par événement

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. C'est bien le cas, mais de manière moins prévisible, il apparaît que l'intervalle de confiance sur une moyenne de tirage aléatoire surévalue sensiblement (d'un facteur compris entre 3 et 5 pour les 4 sites étudiés et pour les pas de volume les plus représentatifs) l'intervalle de confiance des erreurs de discrétisation.

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 le même pour tous les événements, 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 manqués en fin d'événement. Pour ce faire nous avons diminué la capacité du flacon pour provoquer des saturations à pas de volume fixé. 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 une capacité de flacon provoquant des pertes de prélèvement avec l'erreur commise *sur le même événement* pour une capacité de flacon suffisante pour recevoir tous les prélèvements.

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 caractérisée par un intervalle de confiance à 90%. Cet intervalle de confiance est matérialisé par une couleur pastel pour le distinguer des intervalles de confiance à 95% présentés 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.

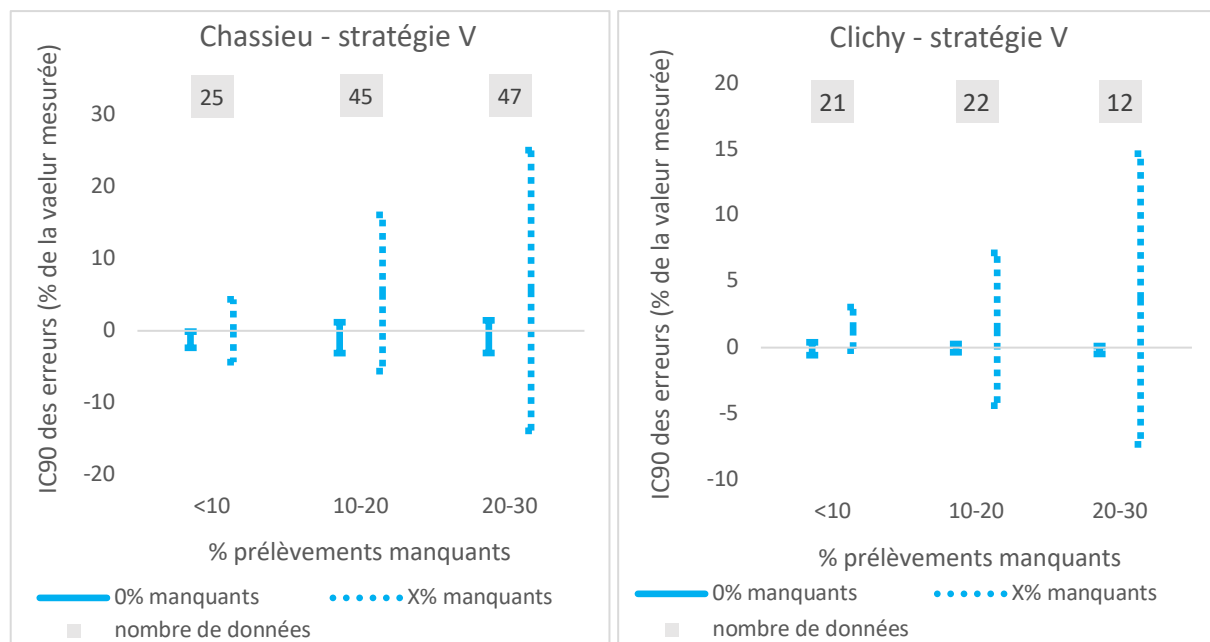


Figure 5 : IC90 des erreurs en fonction de la proportion de prélèvements manquants (pour un taux de perte donné - X%- on compare les erreurs observées pour ce taux de perte, et celles qui sont observées **pour les mêmes événements**, avec un paramétrage ne provoquant pas de perte de prélèvement – voir texte)

Le Tableau 4 synthétise les résultats pour le 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 4: 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

3.2 Stratégie de prélèvement proportionnellement au temps

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 paramètres choisis.

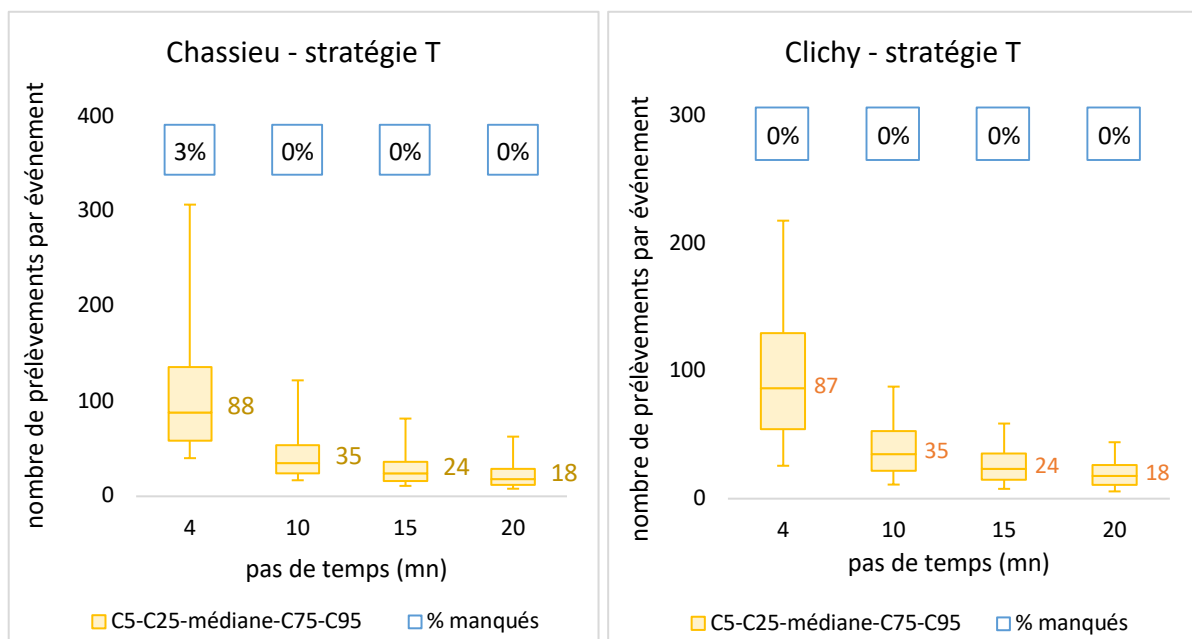


Figure 6 : nombre de prélèvements par événement en fonction du pas de temps

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.

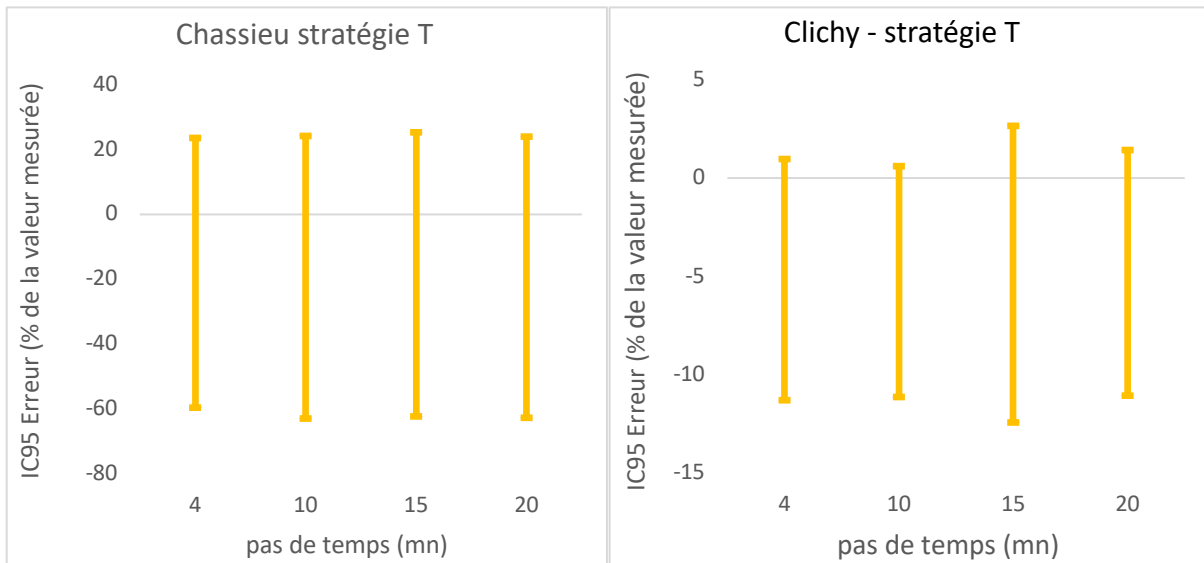


Figure 7 : IC95 des erreurs en fonction du pas de temps

3.2.3 Comparaison des erreurs d'échantillonnage avec les erreurs de pondération 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.

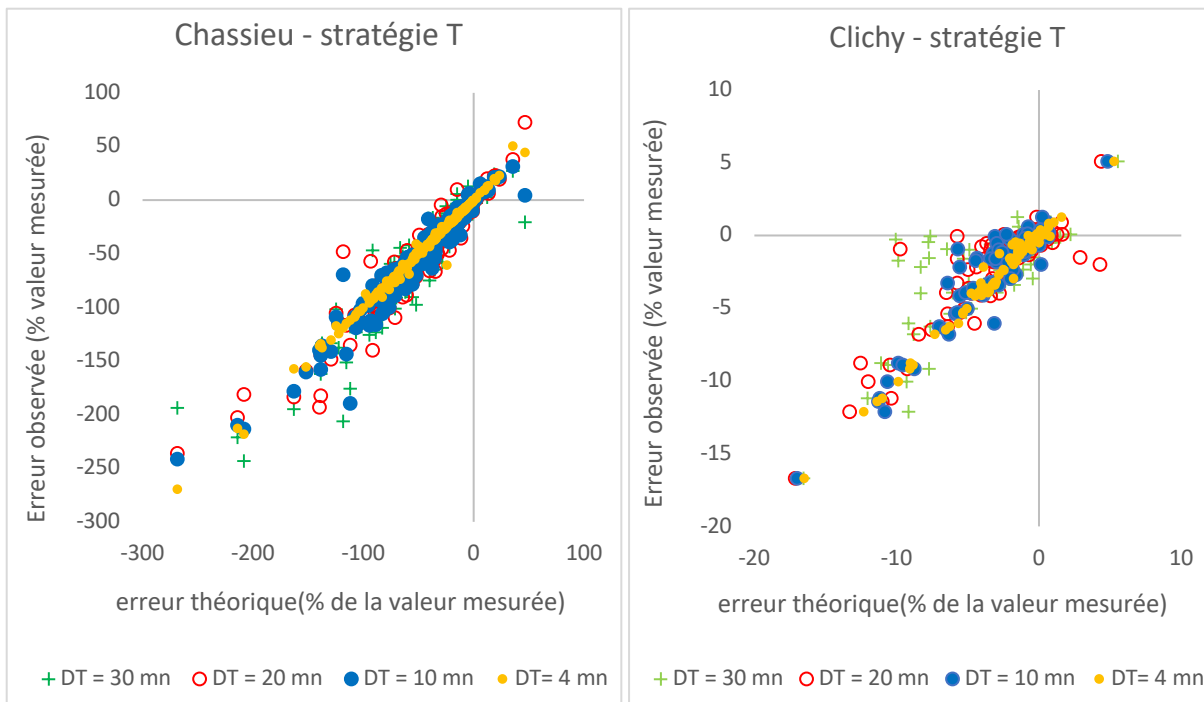


Figure 8 : comparaison des erreurs observées pour la stratégie T avec l'erreur théorique de pondération

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.

3.3 Comparaison des deux stratégies

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 a 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.

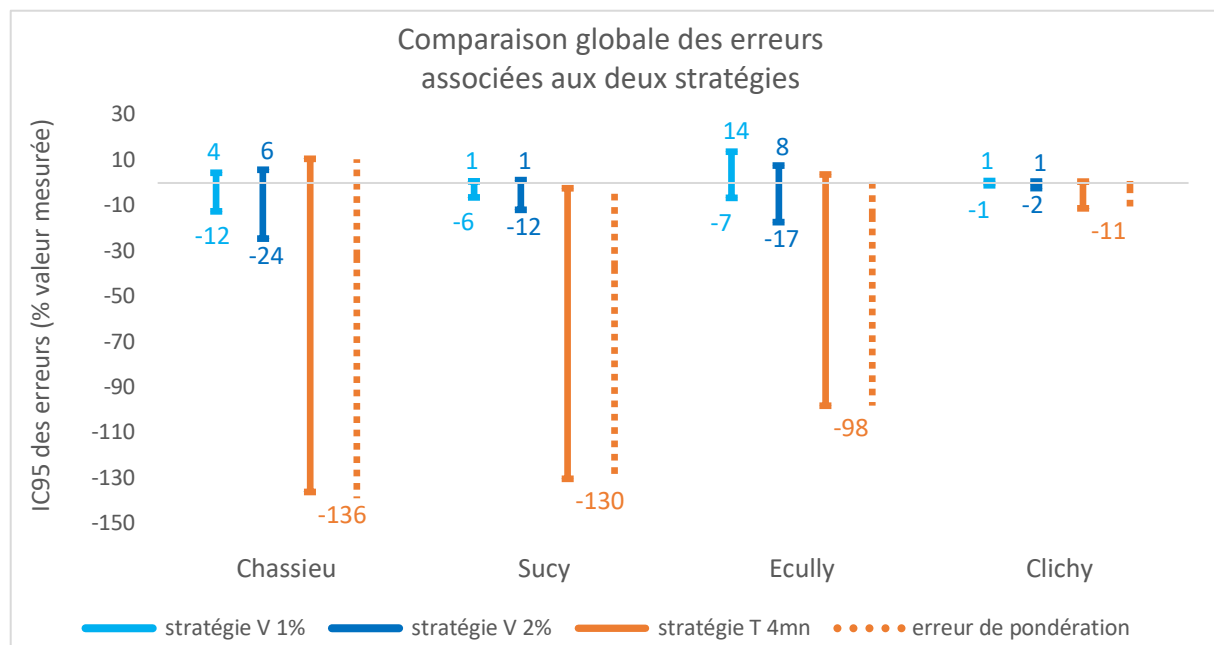


Figure 9 : Comparaison globale pour les mêmes événements des IC95 des erreurs associées aux deux stratégies

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 pour 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) On néglige l'impact des incertitudes de mesure de la turbidité et du débit sur les erreurs de discrétisation et de pondération.

De plus les valeurs numériques des erreurs fournies par les simulations correspondent à 4 sites, qui présentent certes de caractéristiques 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 ?

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.

4 Conclusions

- 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.

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.: 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).'

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

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, LGCIE, 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, 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