
Evaluation des flux à partir de suivis discrets

De manière théorique, le flux de matière cumulé (dissoute ou en suspension) transporté pendant une période de temps (entre t_1 et t_2), se détermine par intégration dans le temps des flux instantanés. Ce dernier est évalué par intégration sur la section transversale (S) des produits " vitesses $v(x,y,t)$ et concentrations $c(x,y,t)$ mesurées en tout point de la transversale " :

$$Flux_{t_1}^{t_2} = \int_{t_1}^{t_2} Flux(t) dt \text{ et } Flux(t) = \int_S v(x,y,t) * c(x,y,t) dS$$

Différentes approximations sont réalisées afin de calculer ce flux. Une première approximation est réalisée sachant que l'hypothèse d'une connaissance parfaite des fonctions $v(x,y,t)$ et $c(x,y,t)$ en tout point de la section est irréaliste. Cependant des moyens technologiques plus performants essayent de combler ces lacunes. Ainsi pour quantifier le flux de matière en suspension, plusieurs auteurs évaluent actuellement la technique de l'Acoustic Doppler Current Profile (ADCP) qui enregistre l'intensité de rétrodiffusion acoustique renvoyé par les particules (Jugari Tiron et al., 2009). Cette intensité est un indicateur acceptable des concentrations de MES (Gartner, 2004 ; Dinehart et Burau, 2005 ; Kostaschuk et al., 2005). Cependant les méthodes de calibration pour convertir cette intensité de rétrodiffusion en concentration de MES requiert d'importante campagne d'échantillonnage (Rotaru et al., 2006). Même si une calibration entre les données de rétrodiffusion et les concentrations est réalisée, des écarts allant jusqu'à 30% peuvent apparaître dus à différents facteurs comme la température de l'eau, le mode de mesure, du dispositif acoustique lui-même ainsi que des propriétés des sédiments (granulométrie, concentration). Malgré tout, en raison de limitations techniques, le nombre d'échantillons n'est pas assez élevé pour bien comprendre les raisons de ces écarts (Jugaru Tiron et al., 2009).

Dans ces conditions, on pose donc:

$$\int_S v(x,y,t) * c(x,y,t) dS = \overline{C(t)} * Q(t)$$

où $\overline{C(t)}$ représente la concentration moyenne sur la section au moment t et $Q(t)$ est le débit enregistré au même moment, généralement par l'intermédiaire d'un capteur de niveau et d'une courbe de tarage.

Une deuxième approximation est liée à la concentration moyenne sur la section $\overline{C(t)}$ qui est le plus souvent la concentration en un point quelconque de la section. On formule l'hypothèse d'uniformité spatiale des concentrations et de représentativité spatiale de la concentration mesurée en un point. De fortes incertitudes peuvent être liées à la nature de la mesure des concentrations (directe ou indirecte), notamment pour les matières en suspension où l'on peut utiliser une mesure de turbidité

(Birgand et al., 2004). Et enfin la troisième approximation, est liée à la connaissance des grandeurs C et Q seulement pour des mesures à intervalles discrets, que nous supposons réguliers (Δt). Dans cette thèse, on évalue les incertitudes liées à cette dernière approximation :

$$Flux_{t_1}^{t_2} = \int_{t_1}^{t_2} C(t) Q(t) dt \approx \Delta t * \sum (Q_i C_i)$$

L'intervalle Δt doit être choisi assez petit pour que les quantités Q_i et C_i puissent être raisonnablement considérées comme constantes durant cet intervalle. Dans le cas idéal, on aurait donc des débits et des concentrations en continu nous permettant de calculer également les flux en continu (figure 1- 1a).

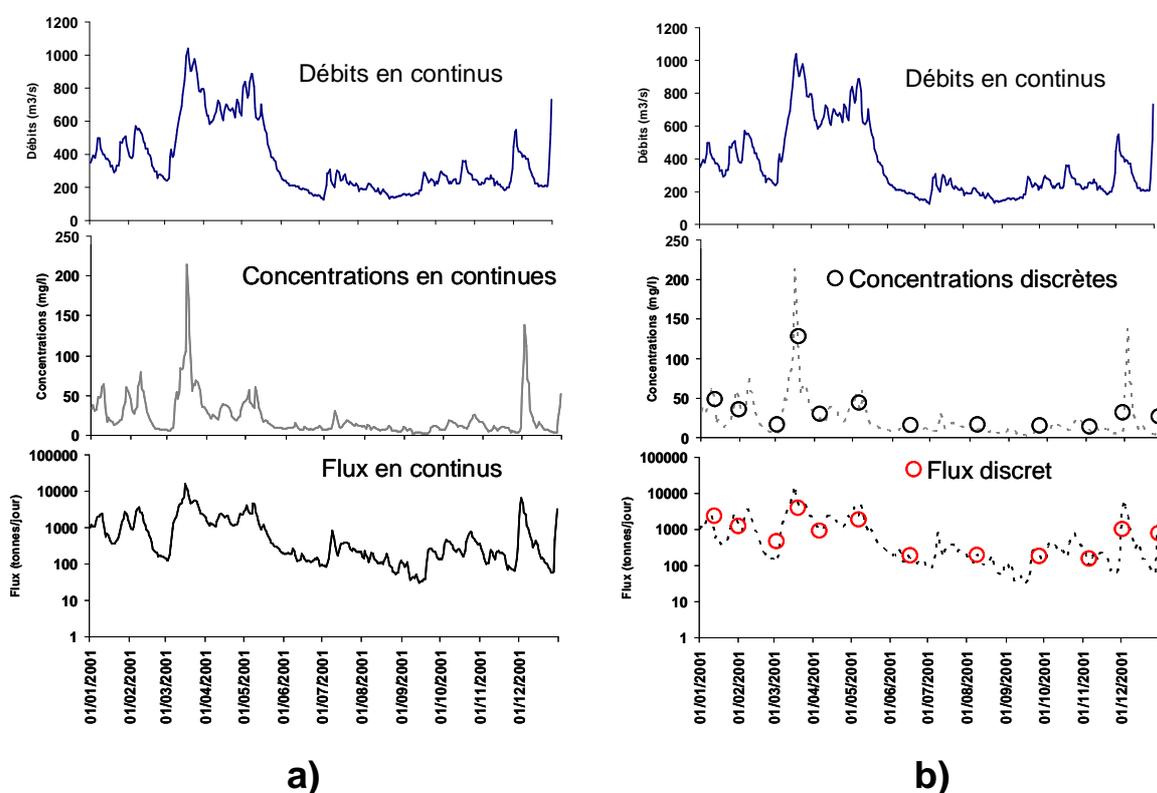


Figure 1-1: Exemple de chronique des concentrations, débits et flux pour une année : a) cas idéal avec données à haute résolution temporelle (journalière) ; b) cas réel avec données discrètes à un intervalle mensuel.

Cependant, pour des raisons techniques et économiques, les mesures de concentration sont toujours moins denses que celles de débit. Les flux sont donc estimés à partir de mesures discrètes de concentrations et des débits en continu (figure 1- 1b). Le choix d'une méthode de calcul performante devient donc primordial pour estimer au mieux les flux annuels.

Ces approximations ainsi que les modalités de réalisation des mesures de concentrations et débits représentent des sources potentielles d'erreurs dans le calcul des flux. Les flux de matières présentent des variabilités liées aux concentrations et aux débits, qui n'ont pas vraiment été exploré par manque de chroniques à haute résolution temporelle. Dans la plupart des cas, les suivis sont au

mieux saisonniers, l'échelle spatiale et le nombre de traceurs sont privilégiés plutôt que les variations temporelles fines.

Afin de réduire les erreurs liées à ce manque de données, différentes stratégies d'échantillonnage ont été mis en place.

1.1 Stratégies d'échantillonnage

Pour acquérir les mesures de concentrations et de débits afin d'estimer au mieux les flux de matériaux, diverses stratégies d'échantillonnage sont mises en œuvre en fonction de la variabilité hydrologique, d'autant plus grande que la taille du bassin est petite et des constituants. Ainsi on peut distinguer un échantillonnage régulier, d'un échantillonnage stratifié (Meybeck et al., 1992, 1993) et un échantillonnage composite (Shih et al., 1994 ; Birgand et al., 2010)

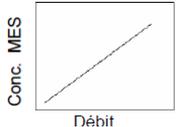
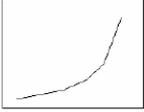
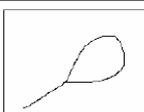
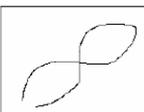
L'échantillonnage régulier consiste à collecter manuellement des échantillons à des intervalles fixes. Plusieurs études, notamment sur le transport des matières en suspension, ont utilisé une large gamme de fréquence d'échantillonnage, avec des intervalles mensuels (Martins et Probst, 1990, sur les grands fleuves africains : Niger, Congo ; Brunet et Astin, 1999, sur le bassin de l'Adour : 16 773km² ; Holmes et al., 2002, pour 19 rivières arctique ; les Agences de l'Eau), hebdomadaires (Tappin et al., 2003, dans l'estuaire Humber en Angleterre : 24 000km²) et journaliers (Audry et al., 2004, sur le bassin de la Garonne : 51 500km²). Cependant, dans beaucoup de rivières, une grande part de la charge annuelle peut être charriée en seulement 5 à 10 jours surtout pour les matières en suspension (Meade et Parker, 1985 ; Meybeck, 2001). Une estimation correcte du flux de MES est donc difficile à réaliser (Walling, 1977 ; Syvitski et Morehead, 1999 ; Serrat et al., 2001). Bien que beaucoup d'études soulignent le problème de la qualité de l'estimation des flux de MES (Meybeck et al., 1993 ; Guyot et al, 1999 ; Ludwig et Probst, 1998), la quantification des erreurs liée à un échantillonnage inadapté a été très peu documentée (Walling et Webb, 1981 ; De Vries et Klavers, 1994 ; Robertson et Roerish, 1999 ; Horowitz, 2003 ; Coynel et al., 2004).

L'échantillonnage stratifié consiste à prélever à des moments privilégiés. Il peut ainsi se dérouler comme une collecte manuelle incluant des périodes déterminées (échantillon mensuel ou bimensuel), complétée par de multiples échantillons prélevés lors de plus forts débits. C'est ce qui a notamment été fait lors du *Nonpoint Programm* conduit par l'United State Geological Surveys (USGS) dans le district du Wisconsin (Graczyk et al., 1993). Typiquement lors de ce programme, une collecte de 100 à 200 échantillons par an est réalisée pour des bassins de taille inférieure à 100 km².

Une alternative à ces deux stratégies consiste à réaliser un échantillonnage à une fréquence fixe, complété par seulement quelques prélèvements pendant les crues (Robertson et Roerish, 1999). C'est le principe utilisé par le *National Water Quality Assessment Program* (NAWQA ; Hirsch et al., 1991). L'échantillonnage mené à l'échelle des Etats-Unis par l'USGS permet de collecter à une période fixe (généralement mensuel) et de compléter par quatre à huit échantillons par an, pendant les fortes crues (Gilliom et al., 1995). L'échantillonnage stratifié en crue est devenu assez répandu surtout dans les petits bassins très réactifs, suite aux nombreuses études sur le comportement des concentrations (C) et des débits en crue (Q). Des modèles de relation C-Q ont été mis en évidence

dans la littérature pour des événements de crue spécifiques, comme par exemple l'hystérésis des concentrations de matières en suspension (Williams, 1989) (tableau 1-1) et peuvent être basées sur des variations du second ordre de type concave ou convexe (Asselman, 2000 ; Crowder et al., 2007).

Tableau 1-1: Différentes classes d'hystérésis dans les relations MES-débit et leurs interprétations

Type de relation	Illustrations	Descriptions
Rectiligne		<ul style="list-style-type: none"> -Crue de transport des matériaux (Jansson, 2002) -Matériaux désolidarisés des berges et déjà érodés (Lenzi et Lorenzo, 2000) -Après des crues où les particules fines sont parties (Hudson, 2003) -Arrivée contenue de matériaux selon débit et granulométrie (Hudson, 2003)
Curviligne convexe		
Curviligne concave		
Boucle horaire: orthograde		<ul style="list-style-type: none"> -Remise en suspension puis érosion et dilution (Jansson, 2002, Picouet et al, 2003) -Epuisement du stock de sédiment (Lenzi et Lorenzo, 2000, Jansson, 2002) -Particules des zones proches du corridor (Williams, 1989, Seeger et al, 2004)
Boucle anti-horaire: retrograde		<ul style="list-style-type: none"> -Sources lointaines (Orwin et smart, 2004) -Après un grand événement pluvieux qui a tout enlevé (Hudson, 2003) -Différence de vitesse entre onde de particule et onde d'eau (Hudson, 2003)
Rectiligne puis boucle		-Dépend de la disponibilité, du stock et du transport de MES (Williams, 1989)
En huit		-Dépend de la disponibilité en particules, et des débits (Williams, 1989)

Dans le cas d'un échantillonnage composite, plusieurs échantillons d'eau sont prélevés à des intervalles de temps régulier et sont emmagasinés dans une bouteille. A partir de ces échantillons, les flux, correspondant au temps qu'il faut pour remplir la bouteille, sont calculés en multipliant les concentrations de chaque échantillon par le volume d'eau cumulé pendant cette période. Cette méthode devrait théoriquement être robuste car l'eau est échantillonnée bien plus souvent que pour un échantillonnage discret. Cependant les concentrations utilisées pour calculer les flux de chaque bouteille sont des moyennes arithmétiques des concentrations alors que théoriquement des concentrations moyennes pondérées par les débits devraient être utilisées. Pour cette raison, les résultats peuvent être biaisés. Cette méthode d'échantillonnage composite nécessite des équipements automatisés très onéreux (Stone et al., 2000).

Bien que les stratégies d'échantillonnage varient entre les différents programmes de surveillance, ils ont tous comme point commun d'essayer d'estimer, avec la plus faible incertitude, le flux de matériaux transporté par la rivière. Cette estimation dépend aussi des méthodes de calculs utilisées à partir de ces données collectées.

1.2 Calcul des flux et de leurs incertitudes à partir de suivis discrets

Les méthodes de calcul des flux sont un domaine en plein développement grâce aux moyens de calculs toujours plus performants, aux séries de plus en plus longues, donc plus informatives et à l'inventivité des chercheurs. Certaines études de synthèse ou comparatives marquent de réelles avancées telles que : Walling et Webb, 1981, Walling, 1984, Ferguson, 1986, Phillips et al., 1999, Moatar et Meybeck, 2005.

1.2.1 Catégories de méthodes de calcul

Walling et Webb (1977) sont parmi les premiers à s'intéresser aux incertitudes sur l'évaluation des flux de matières, notamment MES, et a avoir proposé une terminologie suite à des tirages Monte-Carlo.

Les méthodes de calcul ont été classées différemment selon les auteurs : Walling et Webb (Méthode d'interpolation ou de moyenne et d'extrapolation ou de régression), Tim Cohn, 1995, dans le rapport de l'USGS classe ces méthodes en deux catégories principales : i) les méthodes permettant d'estimer les concentrations en continu, qui consiste à estimer les concentrations manquantes par des relations empiriques en faisant intervenir le débit ; ii) les méthodes permettant une estimation directe notamment en calculant les flux à partir d'un échantillonnage stratifié (Verhoff et al., 1980 ; Thomas et Lewis, 1995).

Cependant c'est en 1999 que Phillips et al., comparent un panel de 22 méthodes de calculs pour évaluer le flux de MES sur deux rivières britanniques : la rivière Ouse à Skelton (3 315 km²) et la rivière Swale à Catterick Bridge (499.3 km²) dans le cadre du programme River and Coastal Research in the Land Ocean Interaction Study (LOIS) ; (Wilkinson et al., 1997). C'est à l'heure actuelle, l'étude comparant le plus grand nombre de méthodes. Dans ce travail, Phillips et al. ont divisé les 22 méthodes en deux catégories (tableau 1-2): i) les méthodes d'extrapolations basées sur la reconstitution des concentrations par des régressions logarithmiques concentrations-débits (M1 à M12), ii) les méthodes d'interpolations basées sur des valeurs moyennes de concentration et de débit (M13 à M22). Le détail de ces méthodes est décrit en Annexe1.

Tableau 1-2: Récapitulatif des 22 méthodes employées par Phillips et al. (1999)

Nom	Formulation	Auteur
M1	Extrapolation des concentrations avec une relation $C=a \cdot Q^b$	
M2	Méthode M1 avec correction de Ferguson	Ferguson (1986)
M3	Méthode M1 avec correction de Duan	Duan (1983)
M4	Méthode M1 segmentée par saison (6 mois)	
M5	Méthode M1 segmentée par trimestre (3 mois)	
M6	Méthode M1 segmentée par saison et en fonction de l'instant où est réalisé le prélèvement sur l'hydrogramme (phase de montée ou de descente des eaux)	
M7	Méthode M1 segmentée en fonction de l'instant où est réalisé le prélèvement sur l'hydrogramme (phase de montée ou de descente des eaux)	
M8	Méthode M1 avec correction de Walling	Walling (1984)
M9	Méthode M1 avec correction de Phillips	Phillips et al.. (1999)
M10	Méthode M1 appliquée par classe de débits	Walling (1984)
M11	Méthode M10 avec correction de Ferguson	Ferguson (1986)
M12	Méthode M10 avec correction de Duan	Duan (1983)
M13	Moyenne des concentrations par classes de débits	Walling (1984)
M14	Produit des moyennes arithmétiques des concentrations et des débits	
M15	Moyenne arithmétique des flux instantanés	
M16	Produit des concentrations (jours de prélèvements) et des débits moyens entre les jours de prélèvements	
M17	Produit des moyennes des concentrations mesurées et du débit de l'intervalle d'étude	
M18	Produit des concentrations moyennes pondérées par les débits et le débit moyen	
M19	Méthode M18 corrigé par Beale (1962)	Littlewood (1995)
M20	Méthode M19 en séparant la base de données en deux selon le débit, si celui-ci est supérieur ou inférieur à la moyenne	Dolan et al.. (1981)
M21	Méthode M19 en séparant la base de données en deux selon le débit, si celui-ci est supérieur ou inférieur à la médiane	Dolan et al.. (1981)
M22	Séparation des flux d'eau de surface et de sub-surface sur la base de la séparation d'hydrogramme d'après Walling et Webb (1982)	Littlewood (1995)

Les incertitudes des 22 méthodes sont présentées à la figure 1-2. On peut remarquer des différences importantes entre les méthodes suivant les fréquences mensuelles ou bimensuelles.

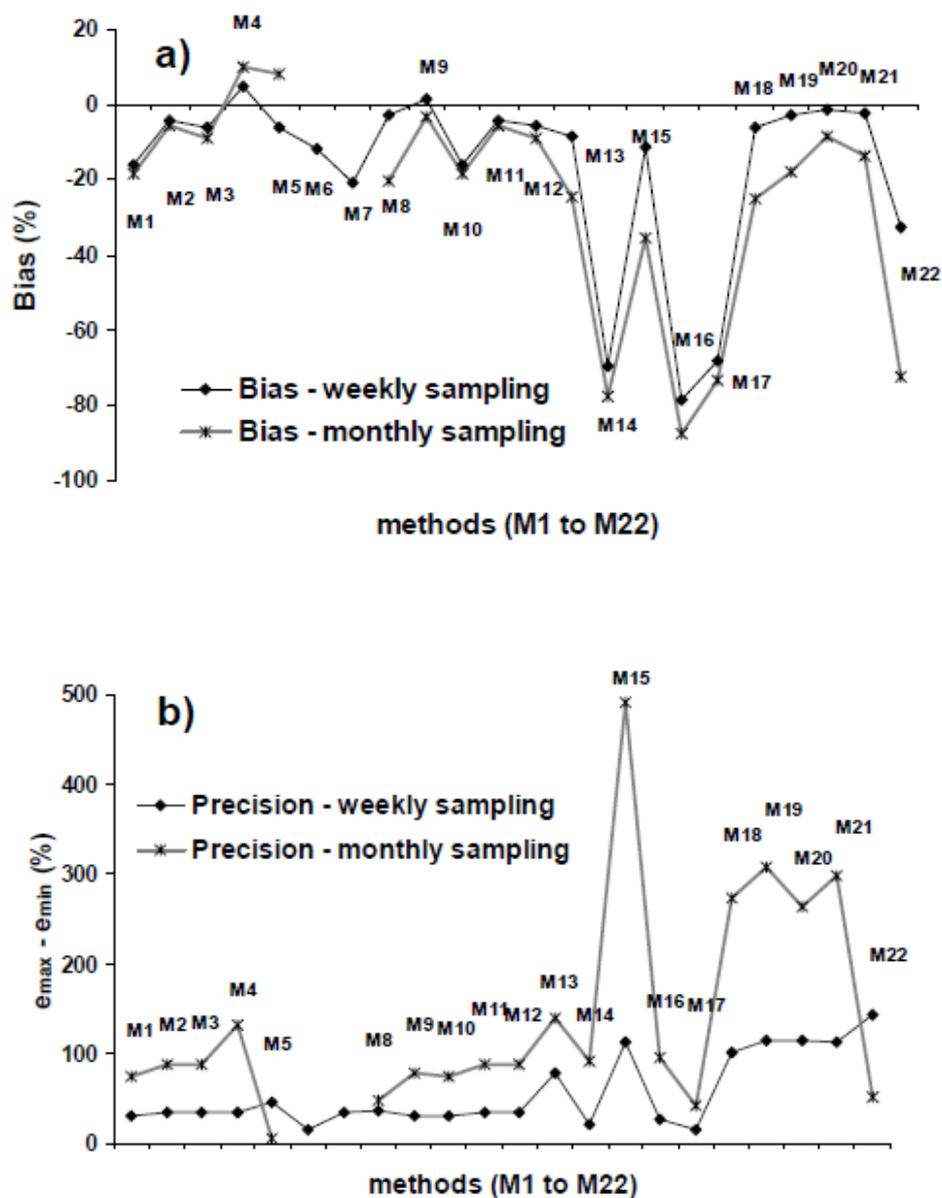


Figure 1-2 : Incertitudes des 22 méthodes de calcul des flux annuels de MES testée par Phillips et al. (1999) pour la rivière Ouse (3 315 km², Angleterre) en considérant des suivis hebdomadaires et mensuels : a) biais, défini comme étant la médiane (e50) des erreurs relatives ; b) l'imprécision définie ici comme étant la différence des erreurs maximales et minimales. M1 à M12 : Méthodes de régression ; M13 à M21 : méthodes basées sur des estimateurs de moyennes de concentrations ou des flux. (d'après les résultats de Phillips et al., 1999, dans Moatar et al., 2007).

L'analyse de ces graphiques montre une amélioration de l'imprécision lorsque l'intervalle de mesure diminue (passage d'une fréquence mensuelle à hebdomadaire) (figure 1-2b). Cependant le gain de biais est moindre pour ce cours d'eau sauf pour les méthodes dites « d'interpolation », M14 à M21. Les régressions linéaires logarithmiques entre les concentrations et les débits (Figure 1-3) sont les modèles les plus communs (Walling, 1974).

$$\text{Log } C = b_0 + b_1 \log Q + \varepsilon$$

où b_0 et b_1 sont les paramètres de régression et ε est en théorie le résidu de distribution normale de moyenne 0 et de variance constante. Cette méthode, également appelée « rating curve », permet une estimation du flux de sédiments sur une période de temps où les débits sont mesurés en continu. Cependant le résultat peut être fortement biaisé. En effet, Walling et Webb (1981, 1988) montrent que de telles méthodes peuvent produire une sous-estimation de l'ordre de -83% à -23%. Ce biais est dû à l'application d'une transformation logarithmique sans considérer la variance non expliquée par la régression. Ce problème est maintenant connu et de nombreux facteurs de corrections ont été proposés (Jansson, 1985 ; Ferguson, 1986, 1987 ; Cohn et al., 1989 ; Singh et Durgunoglu, 1989). Gilroy et al., (1990) ont comparé les performances de trois facteurs de corrections fréquemment utilisés. Ils ont montré que ces facteurs de correction réduisent le biais. Ces méthodes classiques sont encore largement utilisées pour l'évaluation des flux bien qu'elles ne soient pas entièrement satisfaisantes et qu'elles produisent des estimations imprécises, surtout dans le cas de données de concentrations éparses ou lorsque le jeu de données ne représente pas une gamme représentative de concentration et de débit (Cooper et Watts, 2002 ; Moatar et Meybeck, 2005 ; Salles et al., 2008) ou encore lorsque la relation n'est pas linéaire (exemple des nutriments) ; (Cohn et al., 1992 ; Horowitz, 2003).

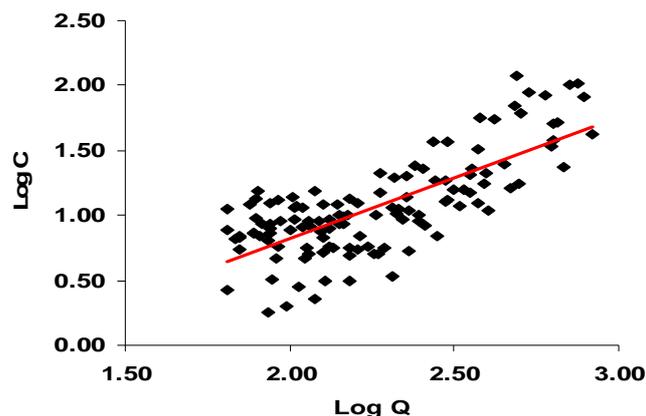


Figure 1-3: Régression logarithmique concentrations- débits

Une alternative à la régression simple est proposée par Cohn et al., (1992) qui suggèrent l'utilisation d'un modèle multivarié. Les concentrations sont estimées à partir d'une régression basée sur sept paramètres : constante, ajustement quadratique au logarithme des débits (deux paramètres), ajustement quadratique avec le temps (deux paramètres), une fonction sinusoidale pour enlever les effets de la saisonnalité (deux paramètres).

$$\ln (C) = \beta_0 + \beta_1 \ln (Q / \bar{Q}) + \beta_2 [\ln (Q / \bar{Q})]^2 + \beta_3 \ln (T / \bar{T}) + \beta_4 [\ln (T / \bar{T})]^2 + \beta_5 \sin (2 \pi T) + \beta_6 \cos (2 \pi T) + \varepsilon$$

où Q est le débit, T est le temps en année, \bar{Q} et \bar{T} sont les valeurs moyennes des séries, les β sont les paramètres du modèle à estimer à partir des données et ε est le résidu de la régression supposé normalement distribué de moyenne 0 et de variance constante. Dans le cas où les relations C-Q ne sont pas linéaires, cette méthode peut être intéressante.

Les méthodes basées sur les moyennes, aussi appelé méthodes d'interpolation ou d'intégration, relient les moyennes des concentrations disponibles et du débit sur un intervalle de temps. Plusieurs procédures de ce type existent (Walling et Webb, 1981, 1988), en utilisant soit les débits journaliers (Q_j) soit les débits mesurés les jours d'échantillonnage des concentrations (Q_i) (figure 1-4).

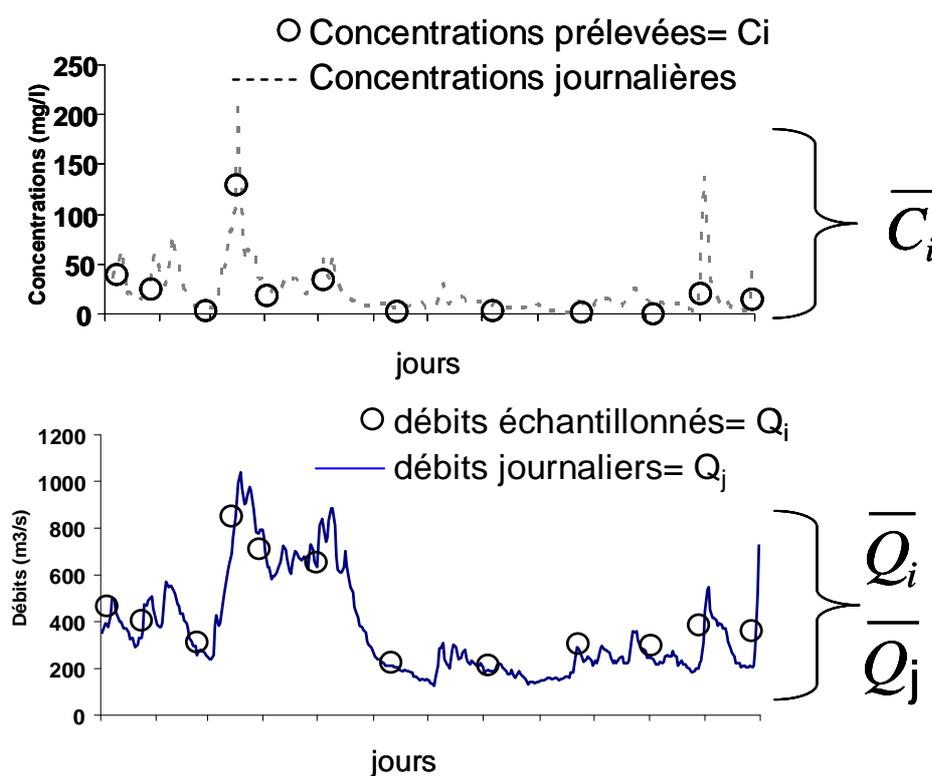


Figure 1-4: Représentation des données servant à calculer les flux à partir des méthodes de moyenne, C_i , Q_i , Q_j , avec ici $i=1$ à 12 ; valeurs mensuelles.

Plus récemment Quilbé et al., (2006) propose une classification plus complète des méthodes standard utilisées pour estimer la charge de sédiments en ajoutant deux catégories à celles définies par Phillips et al. (1999) : les estimateurs de planification et les méthodes de rapport d'estimation.

Les estimateurs par planification sont les méthodes les plus simples. Le flux d'une période donnée est estimé comme le produit d'un volume d'écoulement estimé sur la période et d'une valeur représentative de concentrations sur cette même période. Comme il n'y a pas besoin de données spécifiques au site, soient pour les écoulements ou les concentrations, il est souvent utilisé comme une première estimation de la charge de polluants (Schwartz et Naiman, 1999).

En raison de leur simplicité, ces estimateurs sont largement utilisés, spécialement quand la corrélation concentration-débit (C-Q) est faible. Néanmoins plusieurs auteurs (Walling et Webb, 1981, Ferguson, 1987) ont montré que ces estimateurs peuvent être biaisés et imprécis.

Les méthodes basées sur les moyennes génèrent un biais sur l'estimation du flux si l'échantillonnage ne couvre pas une gamme représentative de la variabilité des débits et des concentrations (Ferguson, 1987 ; Dolan et al., 1981) La méthode empirique de rapport d'estimation, dérivée des méthodes basées sur les moyennes, a été développée par Beale (1962) et plus tard modifiée par Preston et al., (1989) pour réduire ce biais. Elle sera détaillée par la suite dans la section 2.1.4.1. Ces méthodes pondèrent les concentrations avec le débit correspondant à l'échantillonnage. Beale (1962) développe ainsi un rapport d'estimation qui réduit le biais en prenant en compte la covariance entre les flux et les débits. Cet estimateur réduit le biais pour les grands bassins versants, notamment pour les MES. Il donne des estimations précises sur les flux de nitrate et de phosphore pour des bassins inférieurs à 20 km² en Angleterre (Littlewood, 1995) et en Finlande (Rekolainen et al., 1991) et est adapté au cas où peu de données de concentrations sont disponibles par rapport aux données de débits. Dolan et al., (1981) ainsi que Rekolainen et al., (1991) ont trouvé que ces estimateurs sont plus précis que les méthodes basées sur les moyennes. Il est à noter que toutes les méthodes décrites ci-dessus peuvent être utilisées sur les débits stratifiés quand un grand nombre de données est disponible.

Lemke (1991) a utilisé les modèles de fonction de transfert pour estimer les concentrations de sédiments en suspension sur deux bassins versants dans l'Iowa (USA). Les concentrations dépendent dans ce cas des débits du jour même mais également des jours précédents. Les résultats du modèle indiquent que les formes et l'estimation des paramètres peuvent être reliées à la taille du bassin versant, au type d'occupation du sol ainsi qu'aux caractéristiques physiques du bassin. Cette méthode génère un modèle complexe avec de nombreux paramètres à estimer même s'il n'emploie qu'une seule variable. Littlewood (1995) a appliqué ces modèles de fonctions de transfert aux flux de nitrates pour deux petits bassins versants anglais. Clarke (1990) a proposé une autre approche qui suppose que les concentrations en sédiments et les débits suivent une loi de distribution log-normale. Des approches plus sophistiquées ont été développées ou sont en cours de développement comme les approches déterministes non linéaires (Sivakumar et Wallender, 2004) ou les modèles de réseaux de neurones (Kisi, 2005). Quelques auteurs ont commencé à utiliser les réseaux de neurones pour tenter d'estimer les flux sédimentaires à l'exutoire d'un bassin versant (Cigizoglu, 2004 ; Alp et Cigizoglu, 2006). Ces méthodes ont déjà fait leur preuve pour la représentation d'autres paramètres comme la température ou le pH (Moatar, 1997 ; Moatar et al., 1999). En utilisant des données hydrométéorologiques (essentiellement pluie et débit), Alp et Cigizoglu (2006) sont parvenus à des résultats tout à fait encourageants, en reconstituant avec une bonne précision les flux sédimentaires journaliers. Plus récemment, Mailhot et al. (2008) ont développé une nouvelle méthode pour estimer le flux annuel de sédiments à partir de la distribution statistique des concentrations. La concentration est considérée comme une variable aléatoire qui peut être représenté par une distribution statistique dont les paramètres sont fonctions des débits et du jour calendaire. Le modèle est donc défini à travers un modèle explicatif utilisant une loi de distribution (par exemple Log-normale, Gamma, Weibull), un modèle de moyenne et un modèle de coefficient de variation des concentrations et dont la moyenne et la variance s'exprime en fonction de covariables (par exemple le débit et le jour julien). Cette méthode sera utilisée dans cette thèse et sera décrite au chapitre 2.

1.2.2 Incertitudes associées aux méthodes de calculs

Dans la littérature, l'évaluation des flux porte souvent sur un seul constituant : Le phosphore par Jordan et al., (2005,2007), Moosman et al. (2005), le nitrate (Birgand et al., 2009), Le constituant le plus étudié est la matière en suspension (MES). Walling et Webb (1981) ont travaillé sur l'estimation des MES sur la rivière Creedy au Devon (Royaume-Uni), Walling (1977) a étudié les limites de la méthode de régression sur les MES pour trois rivières du Devon (Rivière Creedy, Dart et Exe mais également Morehead et al. (2003), de Vries et Klavers (1994), Kao et al. (2005), Holtschlag (2001) Tramblay et al. (2007). Chacune de ces études portent sur un nombre de bassins versants limité (1 à 9) et souvent de petites tailles. Horowitz (2001) commençait à étudier l'estimation des flux de plusieurs paramètres (SPM et éléments traces, phosphore) sur des grands bassins Nord américains (Mississippi, Colorado, Columbia, Rio Grande).

Le tableau 1- 3 fait un inventaire de la bibliographie et donne un ordre de grandeur des incertitudes signalées par les différents auteurs. Les gammes de variations des incertitudes diffèrent selon les paramètres. Elles peuvent aller, selon les études, de -80 à 360% pour les MES, de -2 à 20% pour les nitrates et sont très faibles pour les chlorures.

Différents auteurs ont également évalué les incertitudes des flux déterminés pour des durées de plusieurs années. D'après Littlewood (1995), les flux de nitrate déterminés sur 2 et 14 ans avec la méthode des concentrations moyennes pondérées par les débits sont plus précis que les flux annuels (déterminé sur un an). Même constatation pour Horowitz (2003), qui sur la base de la méthode de régression C-Q, observe une diminution des erreurs de -10% à 3% sur le flux de MES pour des durée de 1 et 5 ans. Cet auteur observe également une diminution de la dispersion des erreurs en augmentant la fréquence d'échantillonnage quelle que soit la durée du calcul des flux.

Tableau 1-3: Inventaire des travaux scientifiques sur les incertitudes associées aux méthodes de calculs des flux en fonction du nombre de stations et du nombre de matériaux étudiés.

Auteurs	Type de méthodes	Variables	Nombre de stations testées	Incertitudes
Walling (1977)	Méthode de régression (4 types)	MES	1	-10% à 30%
De Vries et Klaver (1994)	6 méthodes : Moyenne, interpolation, Méthode de régression	MES, NH4 et chlorures	4	-80% à 350%
Walling et Webb (1985)	Méthode de moyenne pondérée par les débits	MES	1	-10% à 300%
Webb et al. (1997)	7 méthodes : 5 d'interpolation et 2 méthodes de régression	MES	1	-50% à 320%
Phillips et al. (1999)	22 méthodes : 14 d'interpolations et 8 de régressions	MES	2	Jusqu'à 110%
Asselman (2000)	Méthode de régression	MES	12	-40% à 120%
Holtschlag (2001)	Méthode de régression, interpolation linéaire, interpolation par spline	MES	10	-50% à 160%
Horowitz (2001)	Méthode de régression	MES, Ptot, C organique	29	-50% à 13%
Guo (2002)	Méthode de régression, Moyenne pondérée par les débits+ Facteur de correction	Nitrate	1	-2% à 20%
Horowitz (2003)	Méthode de régression	MES	5	-60% à 146%
Moatar et Meybeck (2005)	7 méthodes	Nutriments	1	13% à 34%
Moatar et al. (2006)	Moyenne pondérée par les débits	MES	36	-55%
Aulenbach et Hooper (2006)	Méthode de régression	Alcalinité, chlorures	1	-0.1% à -2%
Moatar et Meybeck (2007)	Moyenne pondérée par les débits	MES, TDS, Nutriments	6	-55% à 5%
Horowitz (2008)	Méthode de régression: linéaire ou polynomiale	MES+métaux lourds, Nutriments	12	-64% pour les SPM

1.3 Les avancées du projet VARIFLUX

Cette thèse s'est déroulée à la suite du projet VARIFLUX (2005-2008) soutenue par le programme Ecosphère continentale CNRS/INSU et qui associe plusieurs laboratoires du CNRS, des universités et du Cemagref ; l'objectif de ce projet était d'étudier les variations temporelles des concentrations et flux de matière pour des bassins de nature contrastée. Nous présentons ici les principaux résultats sur l'estimation des flux et de leurs incertitudes qui ont permis de déboucher sur ce travail.

A partir des séries existantes à haute résolution temporelle (journalières), les erreurs, générées par la méthode des concentrations moyennes pondérées par les débits, ont été déterminées pour des sites spécifiques par les participants au projet: rivières du sud-ouest et alpines pour les MES, rivières bretonnes et de la Loire moyenne pour les nutriments. Ainsi Coynel (2005), montre que pour une même région (Sud-Ouest de la France), les erreurs dépendent de la taille des bassins et de la variabilité des flux (Figure 1-5).

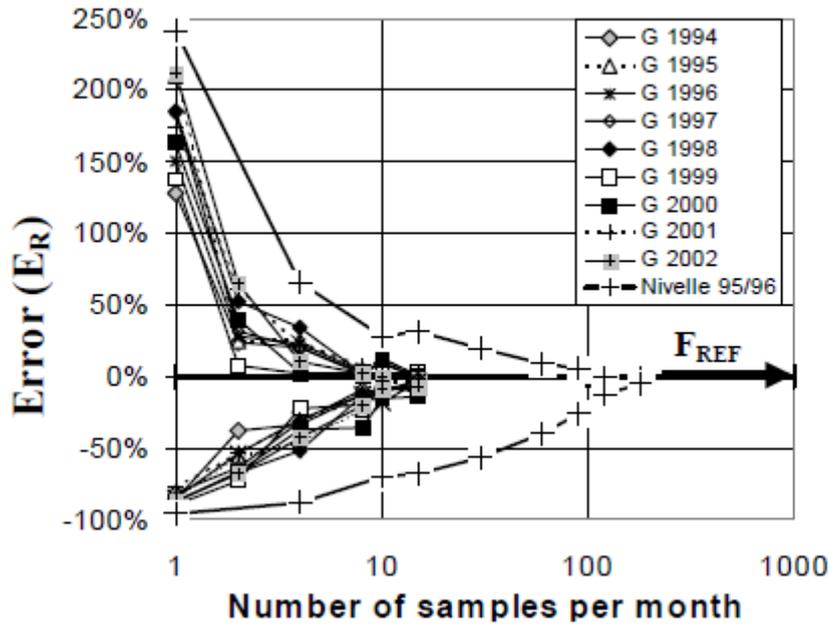


Figure 1-5: Comparaison des erreurs sur les flux annuels de MES simulés pour la Garonne et la Nivelles en fonction de la fréquence d'échantillonnage (nombre d'échantillons par mois). Les nombres d'échantillons sont sur une échelle logarithmique (Coynel, 2005)

Ainsi pour la Garonne à la Réole (~50 000 km²) la fréquence optimale pour avoir une erreur sur le flux de MES comprise entre $\pm 20\%$ est de 3 jours, de 7 à 10 jours pour des bassins de 7 000 à 10 000 km² (Lot, Adour, Isle, Charente) et de 7 heures pour des bassins de 165 km² (Nivelles). Une étude similaire a été réalisée par Mano (2008) sur des rivières Alpines de 100 à 1000 km². Pour des petits bassins versant torrentiels de ce type, la variabilité des flux de matière est très grande. Avec un intervalle d'échantillonnage mensuel, le biais sur les flux de MES est de l'ordre de -100% et l'imprécision de 300%. Pour réduire ces erreurs à $\pm 20\%$, il faut passer à des suivis sub-journaliers. Pour la Bléone (900 km²) dans le bassin de la Durance, la fréquence optimale pour un flux de MES compris entre $\pm 20\%$ est de 8 heures (Figure 1-6) ; de 6 heures pour des bassins de 230 km² (Romanche) et de 4h pour des bassins de 80 km² (Ferrand dans le bassin de l'Isère).

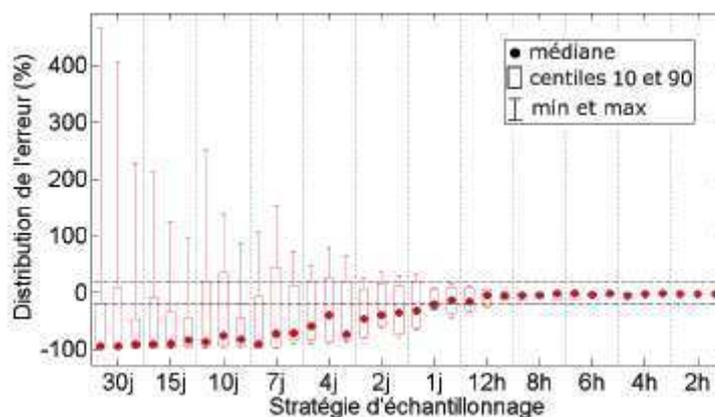


Figure 1-6 : Distributions statistiques de l'erreur sur les flux annuels estimés par des simulations de sous échantillonnage de la base de données sur la Bléone (2001-2003). Pour chaque stratégie d'échantillonnage, la

distribution des erreurs est présentée pour les trois années de mesures. Les pointillés représentent la limite d'erreur à $\pm 20\%$ (Mano, 2008)

Moatar et Meybeck (2005) ont comparé les performances de sept méthodes de calcul sur l'estimation des flux de nutriments à un intervalle d'échantillonnage mensuel pour la Loire à Orléans : 36 970 km² (Figure 1-7).

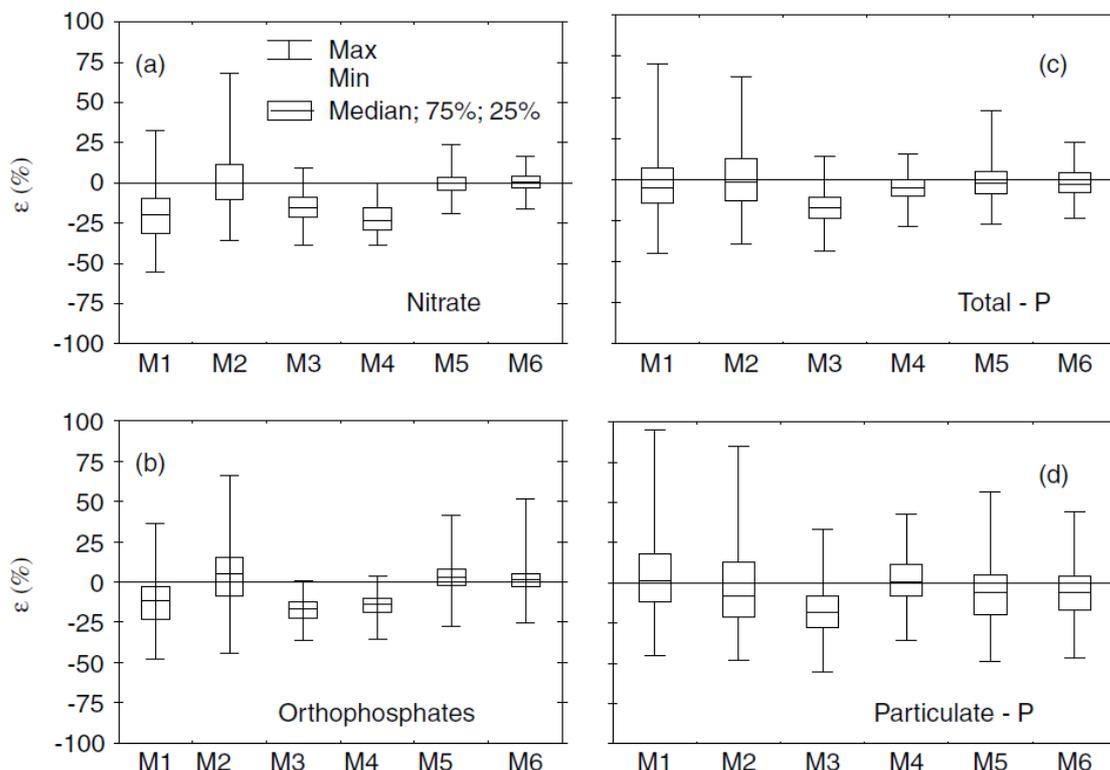


Figure 1-7: Comparaison des gammes d'erreurs sur les flux annuels de nutriments générées par six méthodes de calcul à un intervalle d'échantillonnage discret pour la Loire à Orléans : a) nitrate-N; b) orthophosphate-P; c) phosphore total ; d) phosphore particulaire. (Moatar et Meybeck, 2005)

Les méthodes M1 à M4 sont des méthodes simples de moyenne des concentrations et des débits, la méthode M5 est la méthode des concentrations moyennes pondérées par les débits (M18 selon Philipps et al., 1999) et M6 est une méthode d'interpolation linéaire. La méthode M5 apparaît ici comme la méthode la plus performante pour évaluer les flux de nitrate et d'orthophosphate avec une estimation non biaisée et une dispersion des erreurs très faible. La fréquence d'échantillonnage optimale proposée pour estimer les flux de nutriments avec une précision de $\pm 10\%$, est de 15 jours pour les nitrates, de 10 jours pour les orthophosphates et le phosphore total et de 5 jours pour le phosphore particulaire.

Les performances de différentes méthodes de calcul en fonction de la fréquence d'échantillonnage a aussi été évaluée sur des petits bassins bretons (Birgand et al., 2009). Les incertitudes sur les flux ont été calculées pour l'Elorn pour des fréquences allant de 6/an à 52/an. Sur la Figure 1-8 sont reportés les pourcentages d'erreurs par rapport aux flux réels – ici déterminés sur des échantillons toutes les 12 h – pour chaque fréquence testée. On distingue ensuite les médianes (e_{50}) des erreurs qui vont correspondre aux biais et les percentiles (e_{10} et e_{90}) des erreurs qui définiront les imprécisions, qui

augmentent avec la diminution de la fréquence. La Figure 1-8 représente une sélection des performances pour quatre méthodes de calcul.

En choisissant la meilleure méthode, M5 (méthode des concentrations moyennes pondérées par les débits) pour NO₃- et pour des erreurs e90 inférieures à 10%, la fréquence d'échantillonnage requise est inférieure à 6 échantillons par an (e90 < 7%) tant pour le flux annuel que pour le flux interannuel de NO₃-.

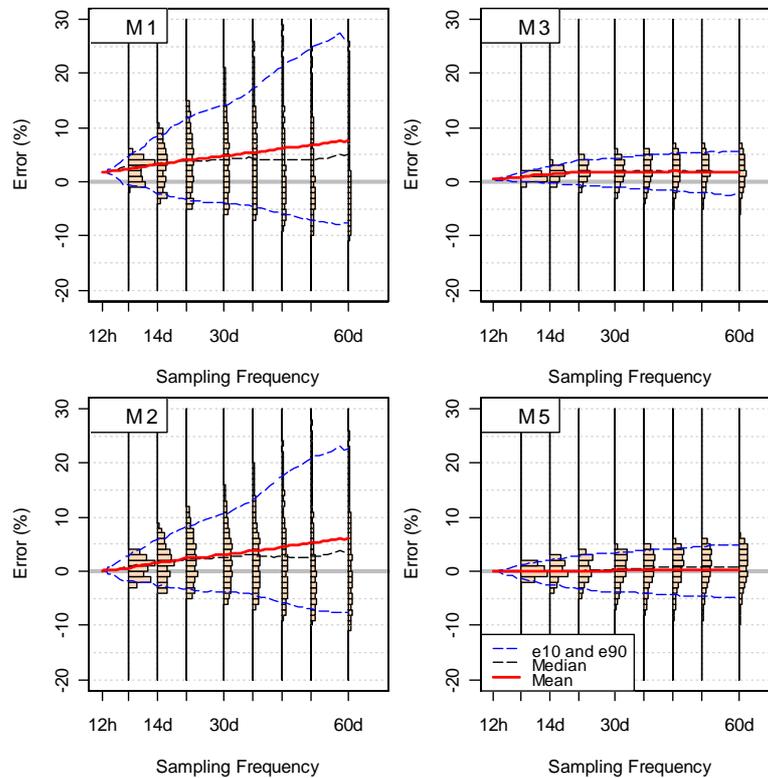


Figure 1-8. Histogrammes des erreurs des méthodes de calcul des flux, M1, M2, M3 et M5 pour les flux de nitrates de l'Elorn (1992/1993) (d'après Moatar et al., 2007)

Des études ont ensuite reliées les erreurs générées par la méthode M5 (Moatar et Meybeck, 2005 ; Birgand et al., 2009) à des descripteurs de variabilité hydrologique et/ou hydrochimique. L'un de ces indicateurs est le $M_{2\%}$, défini comme le pourcentage du flux de matière transporté en 2% du temps (Figure 1-9).

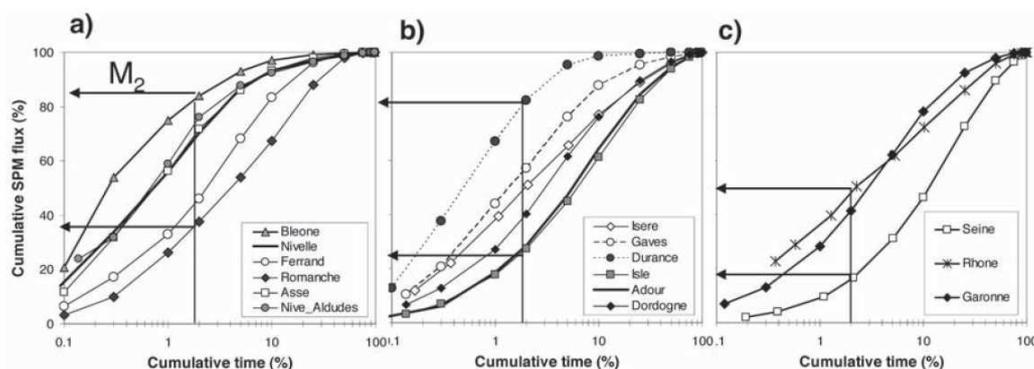


Figure 1-9 : Courbes de durée des flux de MES pour quelques rivières françaises : a) 100 à 1 000 km² ; b) 1 000 à 10 000 km² ; c) 10 000 à 100 000 km² (tiré de Mano et al., 2006)

Les flux journaliers sont classés du plus fort au plus faible sur la période considérée, puis cumulés en partant du plus élevé. On peut ainsi déterminer les flux maximaux d'eau et de matière transportés en 1%, 2%, 10% du temps ; les courbes flux – temps ainsi obtenues sont appelées courbes de durée (Walling, 1974). Elles ont généralement une allure sigmoïdale. L'exemple de la Figure 1-9 présente un faisceau de courbes pour les MES pour trois groupes de stations françaises. Le pourcentage du flux maximal transporté en 2% du temps, noté $M_{2\%}$, a été choisi pour décrire la variabilité des flux pendant la période des crues ; il varie dans les rivières françaises documentées entre 82% (Bléone) et moins de 8% (Seine). Cette dispersion est d'abord observée en fonction de la taille des bassins, les plus petits ayant des $M_{2\%}$ plus élevés. La plupart des courbes de durée peuvent être linéarisées (Figure 1-10) lorsqu'elles sont reportées dans une double échelle de probabilité d'une loi normale, ce qui permet plus facilement leur comparaison et la détermination des descripteurs tels que $M_{2\%}$ (Moatar et al., 2006).

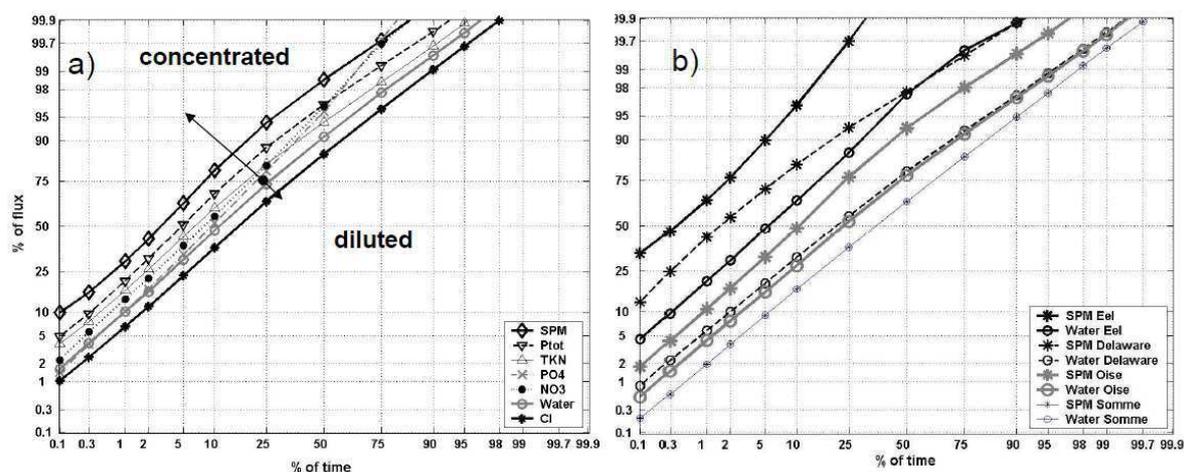


Figure 1-10: Exemple de flux d'eau et de matière sur une double échelle de probabilités : a) rivière Maumee à Waterville (Ohio, USA) pour les chlorures, nitrate, phosphate, azote Kjeldahl, phosphore total et MES ; b) flux d'eau et de MES pour la rivière Eel (Californie, USA), la rivière Delaware (New Jersey, USA), l'Oise et la Somme (France) (d'après Moatar et al., 2010)

Les courbes de durée des flux d'eau et de matière sont quasi linéaires dans cette double échelle de probabilité. Pour une même rivière, les matériaux dilués comme les chlorures dans cet exemple, sont représentés par des lignes parallèles situées en dessous de la courbe de durée des flux d'eau, alors que les matériaux concentrés (TKN, Ptot et MES) sont situés au dessus (Figure 1-10a). Les courbes de durée linéarisées ont aussi l'avantage de pouvoir comparer les stations entre elles pour un même type de matériaux (Figure 1-10b). La linéarité des courbes de flux d'eau et de matière ou la présence de « cassure » peut être reliée à la distribution unimodale vs bimodale des flux d'eau et de matière. En effet, une distorsion des courbes d'eau est observée pour les rivières Delaware et Eel qui peut être reliée à la distribution bimodale des débits observée par Nash (1994) et Syvitski et Morehead (1999). Pour ces rivières, la distance entre les courbes de durée d'eau et de MES est élevée pour Eel et le Delaware mais n'est pas visible pour la Somme dont l'aquifère est dominé par les craies et dont les concentrations en MES sont très basses même pendant les crues à cause d'une érosion mécanique très limitée dans ce bassin.

Moatar et al., (2006) ont ensuite établi un abaque qui relie les biais et les incertitudes sur les flux annuels ou interannuels estimés (Figure 1-11), par la méthode des moyennes des concentrations pondérées par les débits et le débit moyen, au descripteur $M_{2\%}$.

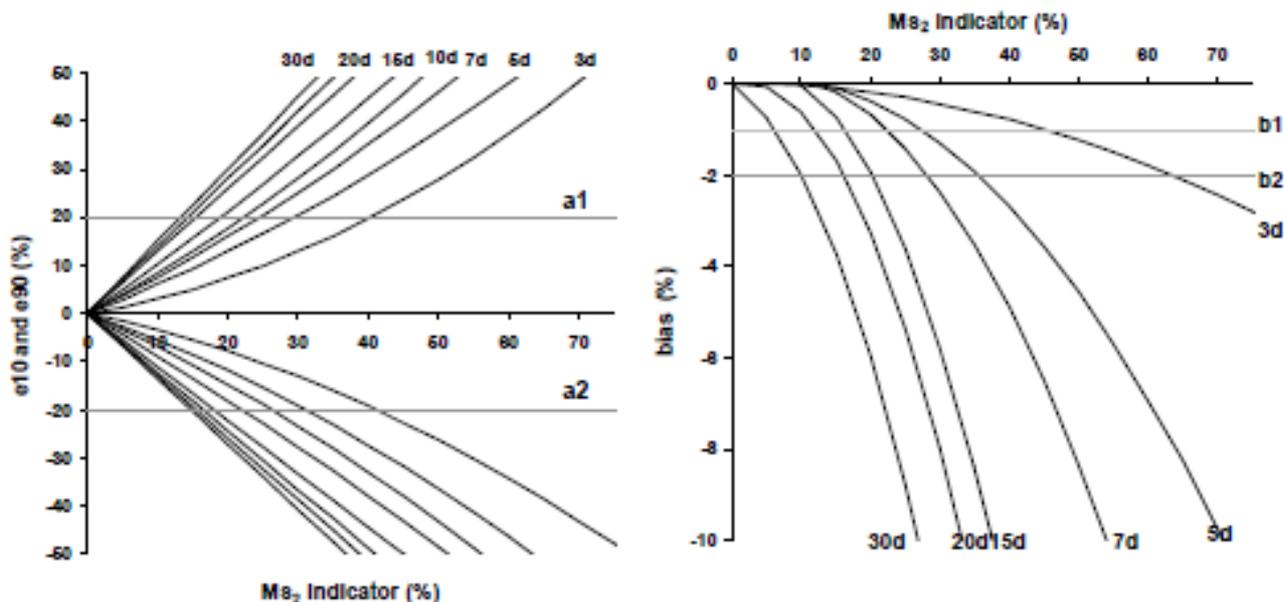


Figure 1-11 : Abaque des incertitudes des flux annuels de MES (méthode de la concentration moyenne pondérée par les débits) : a) déciles inférieurs et supérieurs des erreurs relatives pour différents intervalles de suivis : 3, 5, 7, ..., 30 jours ; b) biais pour différents intervalles de suivis (a1 et a2 = limites imposées de précision à $\pm 20\%$; b1 et b2 = limites imposées de biais à -1 et -2% ; d'après Moatar et al., 2006)

Ces abaques montrent que, pour avoir des flux annuels estimés avec des biais inférieurs à -1 ou 2% et des imprécisions à $\pm 20\%$, les intervalles requis entre deux mesures consécutives sont les suivants : < 3 jours pour des rivières ayant un $M_{2\%} > 40\%$; entre 3 et 5 jours pour des rivières avec $30\% < M_{2\%} < 40\%$, entre 5 et 12 jours pour des rivières avec $20\% < M_{2\%} < 30\%$, entre 20 et 12 jours pour $15\% < M_{2\%} < 20\%$ (Figure 1.11). Cet abaque permet ainsi de déterminer le biais et l'imprécision sur le flux de MES pour cette méthode en fonction de la fréquence d'échantillonnage. Moatar et Meybeck (2007) ont ensuite démontré qu'il était applicable aux autres types de matériaux (ions majeurs, nutriments dissous et totaux). Ce descripteur clé dans la compréhension des incertitudes n'est pas disponible à partir d'un suivi discret car il nécessite les concentrations en continu (au moins journalier). D'autres descripteurs ont donc été élaborés pour décrire cette variabilité à partir d'un suivi discret. Moatar et al., (2006) ont utilisé le descripteur $W_{2\%}$ qui définit le pourcentage du flux d'eau qui s'écoule en 2% du temps à partir des débits en continu. Comme dans le cas de $M_{2\%}$, plus la valeur de $W_{2\%}$ est grande, plus la variabilité hydrologique du cours d'eau est importante. A l'aide de cet indicateur, $W_{2\%}$, Birgand et al. (2009) ont élaboré des abaques permettant de calculer la fréquence minimale de mesure de la teneur en nitrate à respecter en fonction d'objectifs d'incertitudes données sur les flux. La méthode utilisée est la méthode des moyennes des concentrations pondérées par les débits et le débit moyen (M_{18} selon Phillips et al., 1999). Ces abaques ont été élaborés à partir de neuf bassins versants bretons de petites tailles (12 à 252 km²) où les relations C-Q étaient de type dilué (Figure 1-12).

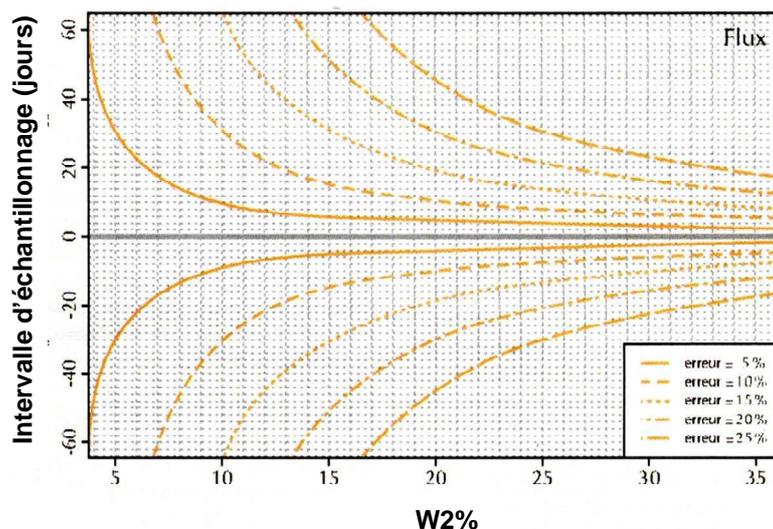


Figure 1-12 : Abaqués élaborés à partir de neuf rivières bretonnes (12 à 252 km²) permettant de calculer la fréquence minimale de mesure de la teneur en nitrate à respecter en fonction de l'indicateur de durée des flux d'eau, $W_{2\%}$, pour des objectifs cibles d'incertitudes (d'après Birgand et al., 2010)

Le comportement des concentrations en fonction des débits diffère selon les types de matériaux. Meybeck et Moatar (2010) ont donc défini une nouvelle typologie des relations C-Q. Neufs types de comportements sont déterminés par les valeurs des pentes de régression pour les valeurs inférieures (b_{50inf}) et supérieures (b_{50sup}) au débit médian : Q_{50} (Figure 1-13). La première lettre « s » de « s-C » se rapporte à la relation C-Q pour des débits $Q < Q_{50}$, ici relation stable pour les faibles débits ; la deuxième lettre « C » de « s-C » se rapporte à la relation C-Q pour les débits $Q > Q_{50}$, dans ce cas concentrations des constituants pour les débits les plus forts.

Si $-0.2 < b_{50 (sup ou inf)} < 0.2$, la relation est considérée comme stable (s et S), si $b_{50 (sup ou inf)} < -0.2$, la relation est considérée comme une dilution (d et D) et si $b_{50 (sup ou inf)} > 0.2$, la relation est considérée comme une concentration (c et C).

- 1 - s-S : Les concentrations sont stables pour toutes les gammes de débits.
- 2 - s-C : Les concentrations sont stables puis augmentent pour les débits supérieurs au débit médian.
- 3 - s-D : Les concentrations sont stables puis se diluent avec l'augmentation des débits.
- 4 - d-S : Tout d'abord une dilution puis une stabilisation des concentrations lorsque les débits augmentent.
- 5 - d-C (« type en U ») : Il ya d'abord une dilution puis une augmentation des concentrations pour les valeurs supérieures au débit médian.
- 6 - d-D : les concentrations en polluants se diluent tout le long de la gamme de débits.
- 7 - c-S : Les concentrations augmentent avec le débit puis se stabilisent pour les valeurs supérieures au débit médian.
- 8 - c-C : Les concentrations augmentent pour toute la gamme des débits.
- 9 - c-D (« type en chevron ») : Les polluants se concentrent puis se diluent.

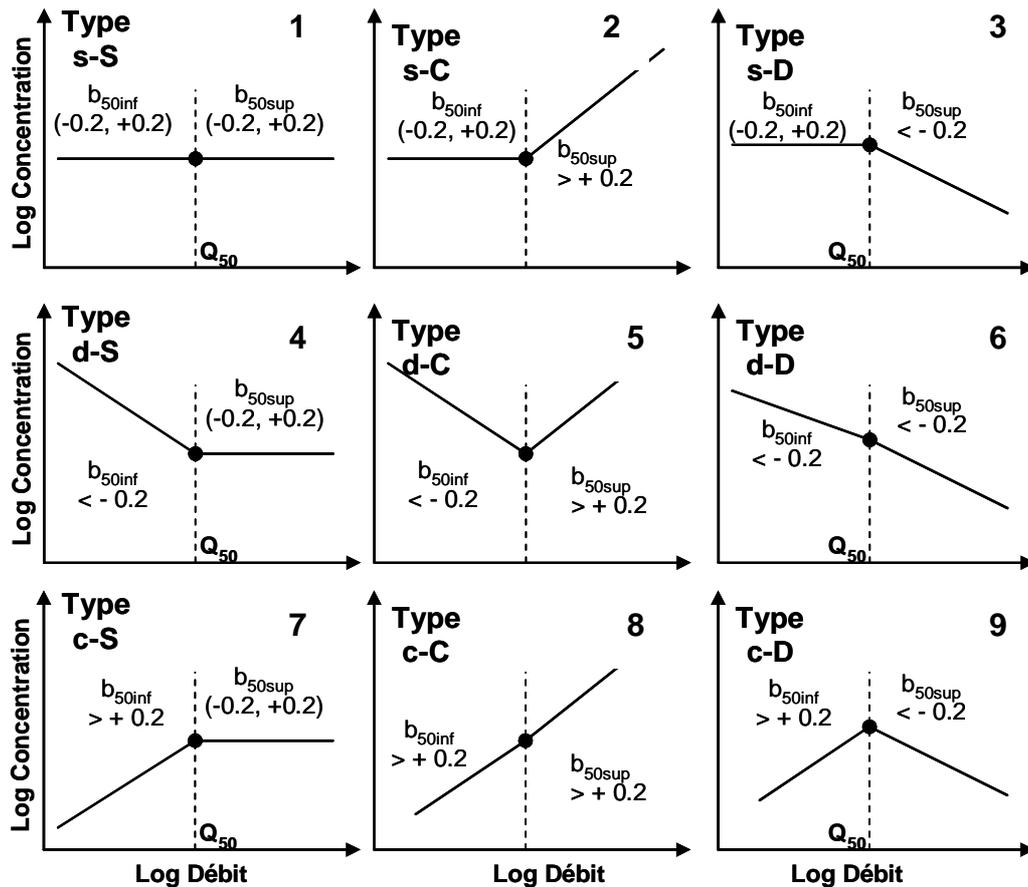


Figure 1-13 : Représentation des neuf typologies décrites par la segmentation au débit médian (Q_{50}) (Meybeck et Moatar, 2010).

La typologie utilisée dans ce manuscrit sera donc celle définie par Meybeck et Moatar à la figure 1-13. Meybeck et Moatar (2010) ont également observé une relation entre le type de matériaux et le b_{50sup} (Figure 1-14). Le b_{50sup} est retenu dans cette thèse comme un indicateur de la variabilité des concentrations et peut correspondre à 65% du flux de SDT et 99% du flux de MES (d'après Meybeck et Moatar, 2010). Ainsi les valeurs négatives, correspondent aux éléments dilués comme les SDT, les valeurs proches de 0 correspondent aux nutriments et les valeurs positives correspondent aux éléments concentrés ou en suspension comme les MES.

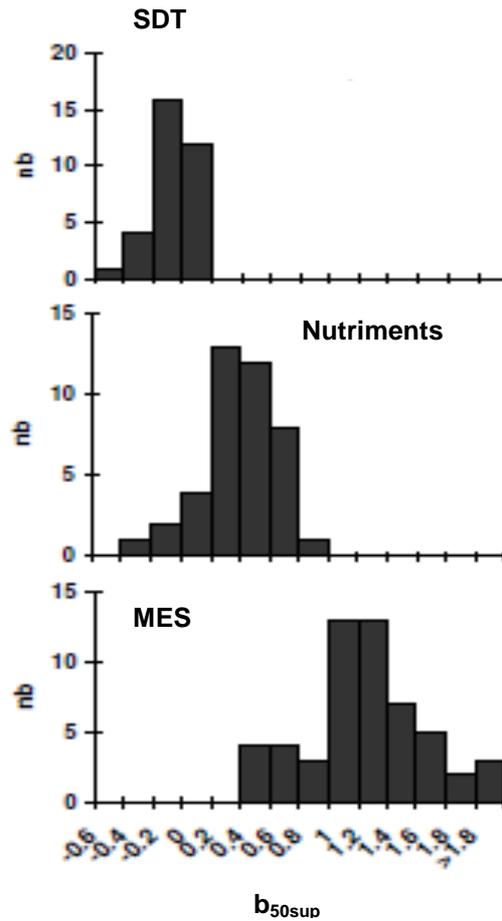


Figure 1-14 : Distribution des exposants b_{50sup} par type de matériaux (Meybeck et Moatar, 2010)

1.4 Synthèse

Les flux de matière, produits des concentrations et des débits, présentent des variabilités liées à ces deux variables. Leur étude n'a pas été explorée, surtout par manque de chroniques à haute résolution temporelle. Pourtant l'analyse fine peut mettre en évidence des variations liées au mélange des eaux, de tributaires d'origine variée, des impacts anthropiques comme les rejets d'origine ponctuelle ou diffuse, la rétention et la régulation par les barrages.

Walling et Fang (2003) ont pu rassembler pour l'étude des tendances des flux de MES à l'échelle mondiale, seulement 145 cours d'eau avec des chroniques de plus de 25 ans mais ils ne prennent pas en compte les incertitudes. De la même manière, les études des tendances sur les chroniques de nitrate en Amérique du Nord et en Europe, ont montré la faiblesse des données historiques (Heathwaite et al., 1997). Là encore les incertitudes sur les flux sont absentes de l'analyse.

Les différents travaux réalisés ont porté régulièrement sur un nombre limité de matériau, de bassin versants et de méthodes. La variabilité d'indicateur classique ou innovant a été reliée la plupart du temps aux incertitudes générées par une méthode (moyenne ou régression) mais les résultats produits sont rarement extrapolables à d'autres bassins, méthodes de calcul ou matériaux qui ne permet pas d'entrevoir une optimisation du réseau de surveillance pour des bassins versant de nature contrastée. Aucune étude ne s'est encore proposé de combiner les incertitudes de plusieurs

méthodes, pour plusieurs types de matériaux et sur un grand nombre de bassins versants afin de définir un cadre d'optimisation du réseau de surveillance.

L'objectif général de cette thèse est donc d'établir un cadre pour définir les incertitudes sur les flux annuels, de faire une étude statistique de ces incertitudes suivant plusieurs facteurs de contrôles afin d'optimiser la surveillance et l'utilisation des méthodes de calculs. Pour cela on se base sur un très grand nombre de données regroupant des bassins versant de tailles moyenne à grande (600 à 1 000 000km²) et pour différents types de matériaux (SDT, MES, nutriments dissous et totaux).

Les objectifs de ce travail sont de traiter les données disponibles afin de :

- i) déterminer et quantifier les incertitudes en fonction du type de polluants et des méthodes de calculs utilisées ;
- ii) analyser les incertitudes associées aux différentes méthodes d'estimation en fonction de la variabilité hydrologique et/ou de la variabilité des concentrations et/ou de la typologie des relations concentrations-débit ;
- iii) analyser l'influence de la fréquence d'échantillonnage sur les incertitudes pour toutes les méthodes de calcul des flux étudiées ;
- iv) analyser l'influence de la durée de calcul des flux sur les incertitudes : vaut-il mieux calculer un flux annuel, sur trois ans, six ans ou dix ans ? ;
- v) déterminer les conditions optimales d'utilisation des méthodes de calcul ;
- vi) proposer la base d'un système expert pour permettre d'évaluer au mieux la qualité de l'eau avec une gamme d'incertitude prédéfinie ;

L'application de ce travail vise à combiner et intégrer ces différentes informations pour optimiser le réseau de surveillance de la qualité de l'eau et ainsi construire un outil facilement utilisable par les gestionnaires et les professionnels. Cette thèse propose donc une approche originale, basée sur une base de données conséquente avec des métriques nouvelles.