

# 1 **Caractérisation spatiale et temporelle des « Masses d'Eau Cours d'Eau ».**

2 **E. Polus, C. de Fouquet, N. Flipo, M. Poulin**

3 *Mines-ParisTech – Centre de Géosciences*

## 4 **Résumé**

5 Focalisé sur les indicateurs physico-chimiques soutenant la biologie, cet article examine l'interpolation  
6 des mesures dans le temps et l'espace, afin de calculer les indices requis par la Directive Cadre sur  
7 l'Eau. En effet, l'application de cette directive soulève quelques questions. Le Réseau National de  
8 Bassin français fournit-il suffisamment d'information pour une cartographie cohérente ? Comment  
9 l'indicateur temporel – le quantile 90 en l'occurrence – varie-t-il spatialement ?

10 Les sorties du modèle déterministe ProSe appliqué à la Seine sont utilisées pour comparer deux  
11 méthodes de calcul du quantile 90 : la méthode classique fondée sur la fonction de quantile empirique  
12 en escalier, et une méthode visant à réduire les biais d'estimation, qui comprend une pondération  
13 temporelle et une linéarisation de la fonction de quantile. Cette seconde méthode, bien que  
14 légèrement plus complexe à mettre en œuvre, permet d'atténuer les biais induits par des mesures  
15 irrégulières ou lacunaires, ainsi que ceux découlant du faible nombre de mesures.

16 Trois méthodes de « spatialisation » des quantiles ont été testées afin d'obtenir des pourcentages  
17 d'occurrence des quantiles par classe de qualité. La première, basée sur le « principe de  
18 défaillance », consiste à ne retenir que la station la plus dégradante ; la deuxième respecte la  
19 proportion de quantiles situés dans chaque classe de qualité, tandis que la dernière attribue un  
20 segment d'influence à chaque station. La spatialisation par segments d'influence des quantiles  
21 temporels au sein des « Masses d'Eau Cours d'Eau » permet d'améliorer nettement les estimations  
22 des pourcentages d'occurrence et révèle la nécessité de la prise en compte de la configuration  
23 spatiale des stations lors du calcul d'un indice de qualité.

## 24 **1 INTRODUCTION**

25 Adoptée par le parlement européen le 23 octobre 2000, la Directive Cadre sur l'Eau (DCE) vise à  
26 préserver la qualité des eaux ou à en restaurer le bon état écologique à l'horizon 2015 (Circulaire  
27 DCE\_2005\_12, 2005). Afin d'en faciliter la caractérisation, les eaux superficielles et souterraines  
28 ont été divisées en unités d'évaluation ou « masses d'eau », dites « Masses d'Eau Cours d'Eau »  
29 ou MECE pour les rivières. La délimitation est fondée sur un critère d'homogénéité biologique et  
30 physico-chimique du milieu (Site du Sandre, 2008).

31 Pour les cinq « indicateurs physico-chimiques soutenant la biologie » considérés dans ce travail, les  
32 nitrates ( $\text{NO}_3^-$ ), l'ammonium ( $\text{NH}_4^+$ ), les orthophosphates ( $\text{PO}_4^{3-}$ ), l'oxygène dissous ( $\text{O}_{2d}$ ) et les  
33 matières en suspension (MES), l'indice recommandé par la DCE est le quantile 90 annuel. Or il a été  
34 montré (Bernard-Michel et al., 2005 ; Bernard-Michel, 2006) qu'aux stations, le calcul de ce quantile  
35 était très biaisé et très imprécis lorsque les prélèvements sont peu nombreux (moins d'une mesure  
36 par mois) ou préférentiels durant l'année (6 mesures en hiver et 12 en été, par exemple).

37 Nous examinons différents aspects techniques posés par le calcul d'un indice statistique (moyenne,  
38 quantile) à partir d'une information très lacunaire : comment le quantile 90 temporel varie-t-il  
39 spatialement? Quelle est l'influence de la méthode de calcul sur les indices de qualité obtenus ? Le  
40 Réseau National de Bassins français (RNB) fournit-il suffisamment d'informations pour caractériser  
41 fidèlement l'état des eaux ?

42 Pour une portion du réseau hydrographique de la Seine, les mesures aux stations du RNB ont été  
43 comparées aux résultats du modèle ProSe, qui en simule le fonctionnement écologique. La  
44 concordance est suffisamment satisfaisante (Poulin, 2006) pour retenir ProSe comme « maquette »,  
45 et reconstituer l'échantillonnage actuel aux stations RNB, avec environ une mesure par mois.

46 Les indices de qualité calculés à partir d'un tel échantillonnage sont comparés à l'indice de référence,  
47 calculé sur l'ensemble des valeurs ProSe (qui présentent une haute fréquence temporelle et spatiale).  
48 Cela permet une quantification empirique de la précision des indicateurs actuels. Enfin, une étude de  
49 sensibilité au nombre de stations par MECE est effectuée.

## 50 **2 METHODES**

51 Le modèle déterministe ProSe décrit plusieurs centaines de kilomètres du linéaire de la Seine, avec  
52 un maillage spatio-temporel très fin en comparaison des données disponibles. Les résultats de ce  
53 modèle fournissent des indicateurs exhaustifs du quantile 90, temporel et spatial, des différentes  
54 grandeurs retenues. Simulant l'échantillonnage actuel aux stations du RNB à l'aide de ProSe, nous  
55 comparons deux méthodes de calcul du quantile 90 temporel et trois méthodes de spatialisation de ce  
56 quantile (figure 1), et nous examinons les écarts par rapport à la « référence ProSe ».

57 Les différentes méthodes sont expliquées en détail dans ce paragraphe, après une brève description  
58 du modèle ProSe.

## Méthode de calcul du quantile 90

| Spatialisation | classique   | améliorée |
|----------------|-------------|-----------|
|                | pénalisante | 2         |
| équi-pondérée  | 1           | /         |
| pondérée       | 2           | 1, 2      |

1 Utilisée pour la "Confrontation visuelle des résultats"  
2 Utilisée pour la "Confrontation quantitative des résultats"

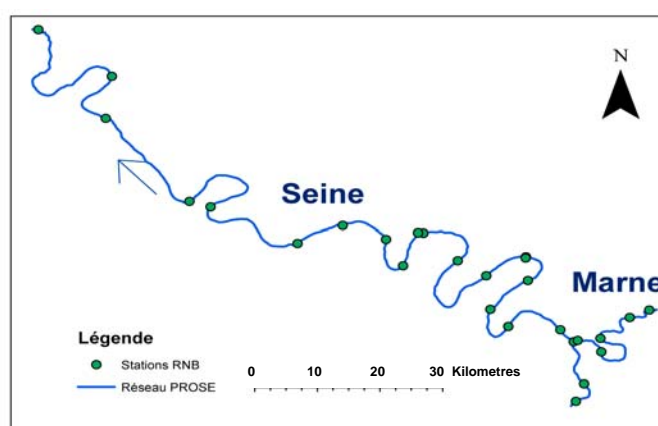
59 **Figure 1 : Méthodes de calcul du quantile 90 et de spatialisation de ce quantile.**

60 *Figure 1 : Methods to calculate 90-percentile and to spatialize it.*

### 61 2.1 Le modèle déterministe ProSe

62 Le modèle ProSe (Even et al., 1998, 2004, 2007 ; Flipo, 2005 ; Flipo et al., 2004, 2007) est composé  
63 de trois modules complémentaires : hydrodynamique, transport et biogéochimie. Le module  
64 hydrodynamique est basé sur une formulation longitudinale monodimensionnelle des équations de  
65 Saint-Venant, résolue par différences finies. Le module de transport simule l'advection et la dispersion  
66 des substances conservatives et réactives, qu'elles soient dissoutes ou particulières. Le module  
67 biogéochimique consiste en une adaptation du modèle RIVE (Billen et al, 1994 ; Garnier et al. 1995).

68 Plus de 300 km de linéaire de la Seine, complétés par quelques kilomètres de la Marne à l'amont de  
69 sa confluence avec la Seine (figure 2) sont simulés. ProSe fournit des concentrations en différentes  
70 substances – dénommés « valeurs ProSe » – toutes les 20 minutes et tous les 500 mètres voire  
71 moins à proximité des singularités.



72

73 **Figure 2 : Réseau modélisé par ProSe et localisation des stations du RNB.**

74 *Figure 2 : Hydrographic network modeled by ProSe and localisation of RNB measurement sites.*

75 L'ensemble de ces valeurs ProSe constitue la référence exhaustive. L'échantillonnage ProSe simule  
 76 l'échantillonnage aux stations du RNB (figure 2). Pour ce faire, seules les valeurs ProSe  
 77 correspondant aux localisations des stations RNB et aux dates effectives de mesure sont retenues.  
 78 Deux jeux de données issus des valeurs ProSe sont donc disponibles : l'ensemble des valeurs ProSe,  
 79 qui constitue la référence exhaustive, et l'échantillonnage de ces valeurs selon le RNB. Des indices de  
 80 qualité sont calculés pour chacun de ces ensembles de données.

81 La DCE préconise l'attribution d'un indice de qualité à chaque MECE. Pour une étude détaillée des  
 82 performances des indicateurs, nous préférons indiquer ici un **pourcentage d'occurrence spatial**  
 83 dans chaque classe de qualité, afin de mettre en évidence la variabilité de l'indice temporel au sein  
 84 d'une MECE. Le pourcentage d'occurrence de référence est égal à la proportion de points  
 85 appartenant à chaque classe de qualité. En revanche, concernant l'échantillonnage ProSe, le quantile  
 86 temporel est d'abord calculé par station puis spatialisé selon différentes méthodes (cf. figure 1), afin  
 87 d'attribuer un pourcentage d'occurrence dans chaque classe de qualité.

## 88 **2.2 Les méthodes de calcul du quantile 90 temporel**

89 Rappelons que le quantile  $k$  d'une population est défini comme la valeur telle que  $k$  % de la population  
 90 lui sont inférieurs (équation 1). Nous présentons ici deux méthodes de calcul du quantile 90 temporel.

$$91 \quad q_k = \alpha \Leftrightarrow P(Z(x) \leq \alpha) = k \quad \text{Equation 1}$$

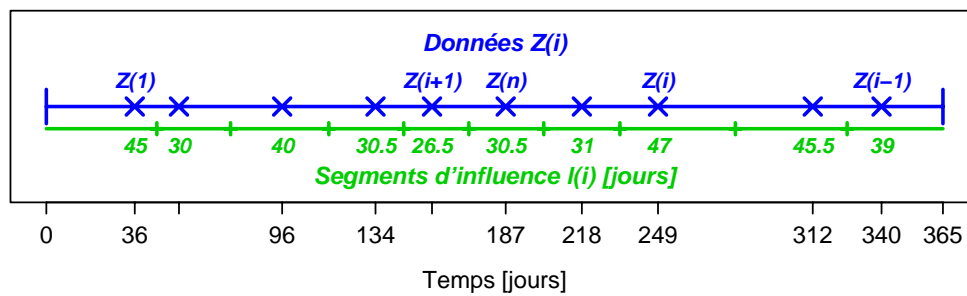
### 93 **2.2.1 Méthode classique**

94 En attendant une méthode universelle (annoncée pour 2009 par la DCE), chaque agence de l'eau  
 95 utilise actuellement sa propre méthode, essentiellement basée sur les dires d'experts, pour le calcul  
 96 du quantile 90 temporel. Nous avons retenu comme « méthode classique » un calcul fondé sur les  
 97 statistiques classiques. En chaque station, le quantile 90 temporel est obtenu par inversion de la  
 98 fonction de quantile empirique, qui est une fonction en escalier.

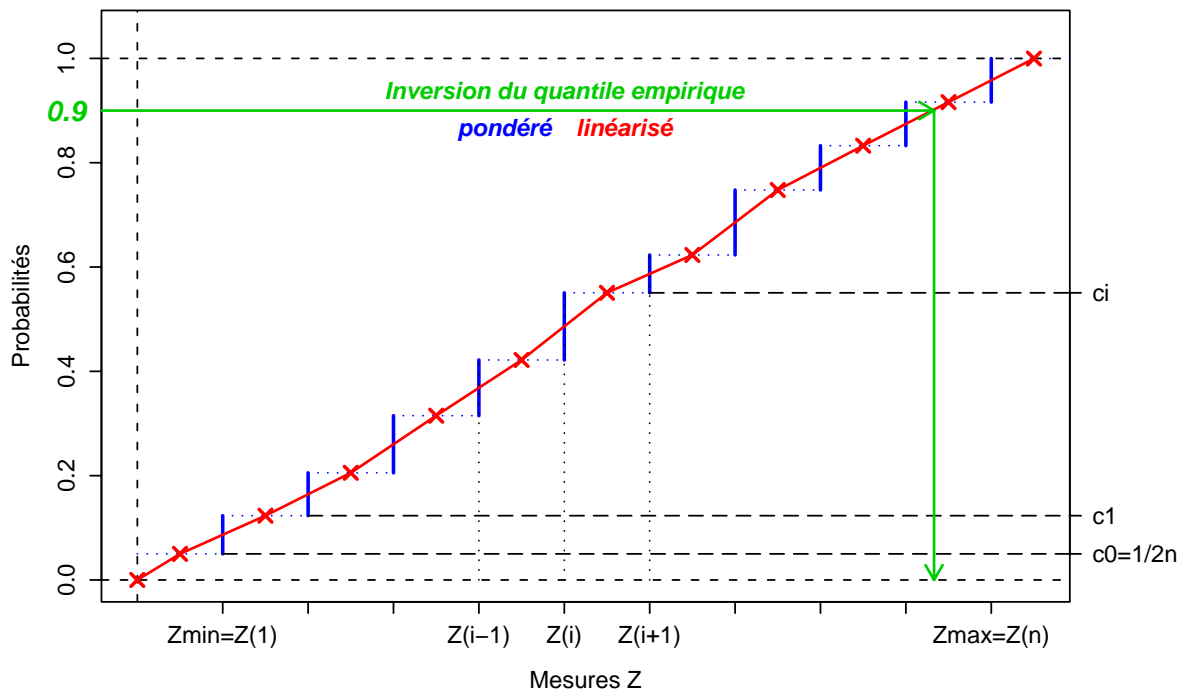
### 99 **2.2.2 Méthode améliorée**

100 Le calcul du quantile 90 par station peut être amélioré pour tenir compte des irrégularités et des  
 101 lacunes de l'échantillonnage temporel (Bernard-Michel, 2006). La « méthode améliorée » comprend  
 102 deux étapes illustrées par la figure 3.

## Pondération des données par segments d'influence



## Méthode améliorée avec au moins 10 mesures par an



103

104 **Figure 3 : Principe de calcul du quantile 90 selon la « méthode améliorée ».**105 **Figure 3 : Principle of 90-percentile calculation by the “improved method”.**

106 Tout d’abord, les données sont pondérées par des segments d’influence qui représentent la « durée  
 107 d’influence » de chaque point. Les limites des segments sont les dates de début et de fin de l’année  
 108 considérée, ainsi que les médiatrices entre dates de mesures consécutives. Le poids d’une mesure  
 109 est égal à la longueur de son segment d’influence rapportée à l’année considérée. La pondération par  
 110 segments d’influence correspond, en fait, à une simplification du krigeage temporel, dont on a montré  
 111 qu’elle est suffisante en pratique pour des calculs à une dimension (Bernard-Michel et al., 2005). Cela  
 112 permet de tenir compte des lacunes temporelles ou des échantillonnages préférentiels (équations 2).  
 113 La fonction de quantile empirique est ensuite calculée à partir des données pondérées, puis linéarisée  
 114 par morceaux (équations 3 et 4).

115 Remarque : afin de diminuer le biais dû au faible nombre de mesures (le quantile est en effet une  
 116 statistique basée sur la loi des grands nombres), la fonction de linéarisation permet de fixer  
 117 arbitrairement un minimum et maximum auxquels un segment d'influence est de longueur nulle est  
 118 associé. En l'absence de bornes spécifiées par l'utilisateur, le minimum est fixé à 0 et le maximum est  
 119 arbitrairement calculé à partir du plus grand écart observé entre deux valeurs successives supérieures  
 120 à la médiane :  $\max = Z_n + \alpha \max_{i \geq med} (Z_{i+1} - Z_i)$  avec  $\alpha = 1.5$

121 Cependant, vu qu'ici nous nous intéressons au quantile 90 et que nous disposons toujours d'au moins  
 122 10 données, les minimum et maximum n'interviennent pas.

$$123 \quad p_0 = p_n = \frac{1}{2n} \quad p_i = \frac{l_i}{l_{tot}} \left(1 - \frac{1}{n}\right) \quad \text{Equations 2}$$

$$124 \quad c_0 = \frac{1}{2n} \quad c_n = 1 - \frac{1}{2n} \quad c_i = \frac{1}{2n} + \frac{1 - \frac{1}{n}}{l_{tot}} \sum_{j=1}^i l_j \quad \text{Equations 3}$$

$$127 \quad q_k = \frac{Z_{i-1} + Z_i}{2} + \left( \frac{Z_{i+1} - Z_{i-1}}{2} \right) \frac{c_k - c_{i-1}}{c_i - c_{i-1}} \quad \text{pour } c_k \in [c_{i-1}; c_i] \quad \text{Equation 4}$$

128 où  $\left\{ \begin{array}{l} p_i \text{ sont les probabilités associées aux données } Z_i \\ c_i \text{ sont les probabilités cumulées associées aux données } Z_i \\ l_i \text{ sont les « durées d'influence » associées aux données } Z_i \\ i \in \{1, \dots, n\} \\ n \text{ est le nombre de données} \\ q_k \text{ est le quantile } k \end{array} \right.$

132 Si la « méthode classique » présente le grand avantage d'être simple et donc facilement applicable  
 133 par tout acteur de la qualité de l'eau, elle n'en présente pas moins plusieurs inconvénients auxquels la  
 134 « méthode améliorée » permet de remédier, moyennant une mise en œuvre un peu plus complexe.  
 135 C'est par exemple le cas des données irrégulières (échantillonnage préférentiel, lacunes, etc.) ou peu  
 136 nombreuses, pour lesquelles les biais importants induits par la « méthode classique » peuvent être  
 137 réduits par la « méthode améliorée » (Bernard-Michel, 2006).

### 138 2.3 Méthodes de spatialisation des quantiles temporels

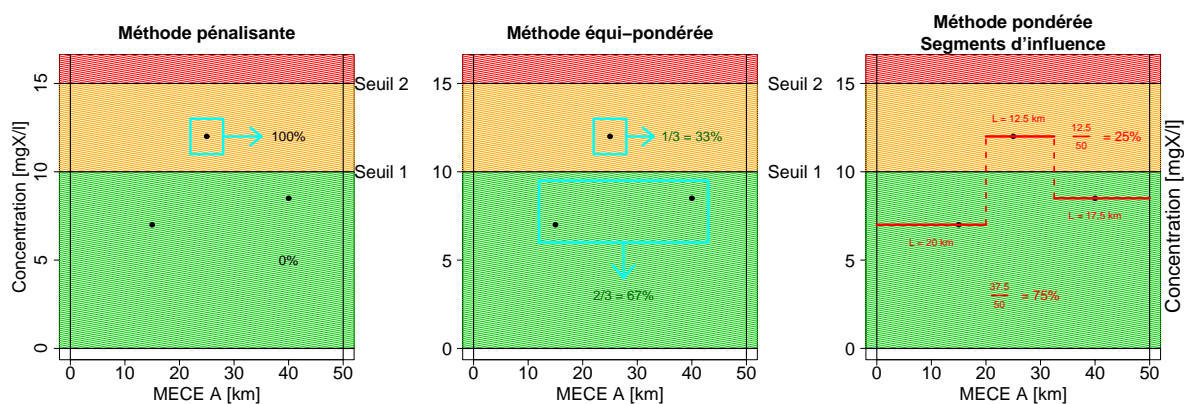
139 Afin d'attribuer un pourcentage d'occurrence des quantiles temporels par classe de qualité, il est  
 140 nécessaire de les spatialiser, c'est-à-dire de définir comment interpoler ou extrapoler spatialement les

141 quantiles temporels au sein d'une MECE. Nous proposons trois méthodes de spatialisation de  
 142 complexité croissante (figure 4).

143 La première méthode, fondée sur un « principe de défaillance », consiste à attribuer le quantile  
 144 temporel le plus défavorable à toute la masse d'eau ; elle est appelée « méthode **pénalisante** ».

145 En référence aux calculs statistiques usuels, la seconde méthode attribue un même poids à toutes les  
 146 stations, qu'elles soient proches ou éloignées les unes des autres : c'est la « méthode **équi-**  
 147 **pondérée** ». Le pourcentage d'occurrence dans une classe de qualité est égal à la proportion de  
 148 stations (situées dans la MECE considérée) pour lesquelles le quantile 90 temporel appartient à cette  
 149 classe.

150 Enfin, la « méthode **pondérée** », généralise la pondération par segment d'influence. Un segment  
 151 d'influence (une longueur d'influence, cette fois) est associée à chaque station, afin d'intégrrer la  
 152 distance – variable – entre stations au calcul du pourcentage d'occurrence par classe de qualité. Le  
 153 pourcentage d'occurrence dans une classe de qualité est alors égal à la longueur totale des segments  
 154 d'influence appartenant à cette classe divisée par la longueur de la MECE.



155  
 156 **Figure 4 : Les différentes méthodes de spatialisation des quantiles 90 pour obtenir les**  
 157 **pourcentages d'occurrence dans chaque classe de qualité.**

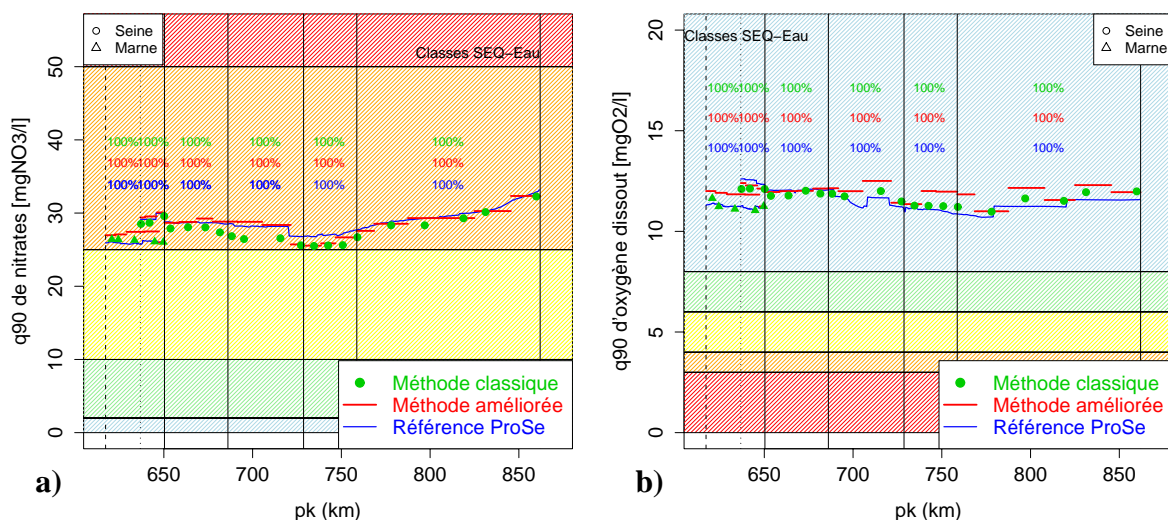
158 *Figure 4 : Different methods to spatialize 90-percentiles in order to attribute occurrence*  
 159 *percentages to each quality class.*

160 Remarque : de façon générale, une pondération spatio-temporelle peut être définie pour des stations  
 161 irrégulièrement espacées et pour des mesures non nécessairement synchrones. Cette généralisation  
 162 n'est pas examinée ici.

163 **3 RESULTATS**

164 Le modèle ProSe est utilisé comme maquette pour valider le calcul du quantile 90 temporel proposé  
 165 par Bernard-Michel (2006) et sa spatialisation. A partir de l'échantillonnage ProSe, qui reproduit  
 166 l'échantillonnage aux stations du RNB, les résultats obtenus par les deux méthodes (« méthode  
 167 classique » et « méthode améliorée ») sont comparés et confrontés aux quantiles exhaustifs des  
 168 valeurs ProSe (référence).

169 **3.1 Confrontation visuelle des résultats**



170  
 171 **Figure 5 : Caractérisation de l'état des eaux selon le SEQ-Eau pour les nitrates (a) et l'oxygène**  
 172 **dissous (b) ; les pourcentages d'occurrence des quantiles 90 temporels dans chaque classe de**  
 173 **qualité sont indiqués par MECE.**

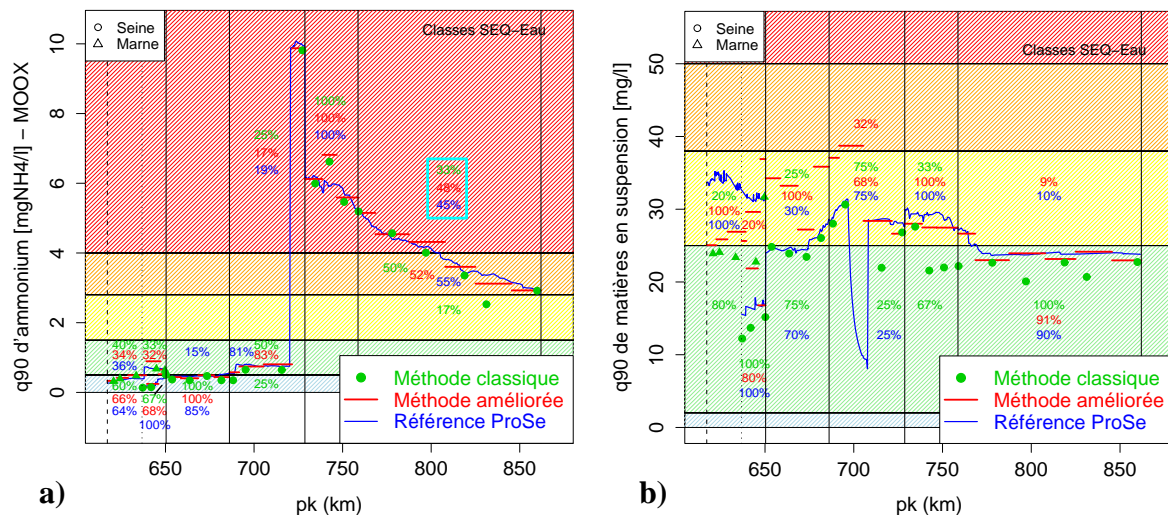
174 *Figure 5 : Characterization of water quality according to SEQ-Eau system for nitrates (a) and*  
 175 *dissolved oxygen (b) ; occurrence percentages of 90-percentiles in each quality class are given*  
 176 *for each « River Water Body ».*

177 Les illustrations se présentent de manière identique pour toutes les variables. Les pourcentages  
 178 d'occurrence sont indiqués dans chacune des classes du SEQ-Eau illustrées par les couleurs  
 179 conventionnelles (bleu : excellent état, vert : bon état, jaune : état moyen, orange : mauvais état,  
 180 rouge : très mauvais état, selon les limites de classes réglementaires). Afin de visualiser la variabilité  
 181 spatiale du quantile temporel au sein d'une MECE, les limites des masses d'eau cours d'eau  
 182 concernées sont matérialisées par des lignes verticales tout au long du domaine modélisé. Les

183 pourcentages d'occurrence obtenus dans chaque classe à l'aide de la « méthode classique équi-  
 184 pondérée » (cf. figure 1) sont indiqués en noir, ceux fournis par la « méthode améliorée » en vert,  
 185 alors que le pourcentage d'occurrence de référence est, lui, repris en bleu.

186 Pour les nitrates et l'oxygène dissous (figure 5), tous les quantiles 90 temporels se situent non  
 187 seulement dans une même classe de qualité, mais ils affichent une très faible variabilité spatiale. Les  
 188 pourcentages d'occurrence fournis par la « méthode classique équi-pondérée », ainsi que ceux  
 189 obtenus par la « méthode améliorée » sont égaux aux pourcentages de référence.

190 Pour les paramètres variant de manière significative sur le domaine, tels l'ammonium et les matières  
 191 en suspension (figure 6), les résultats sont quelque peu différents.



192  
 193 **Figure 6 : Caractérisation de l'état des eaux selon le SEQ-Eau pour l'ammonium (a) et les**  
 194 **matières en suspension (b) ; les pourcentages d'occurrence des quantiles 90 temporels dans**  
 195 **chaque classe de qualité sont indiqués par MECE.**

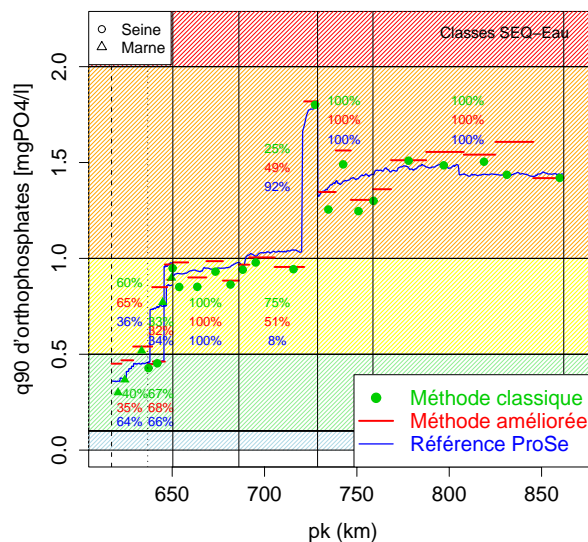
196 **Figure 6 : Characterization of water quality according to SEQ-Eau system for ammonium (a)**  
 197 **and suspended matter (b) ; occurrence percentages of 90-percentiles in each quality class are**  
 198 **given for each « River Water Body ».**

199 Pour l'ammonium, les pourcentages d'occurrence calculés via la « méthode améliorée » sont  
 200 généralement plus proches des pourcentages de référence. C'est ce que l'on observe par exemple  
 201 pour la dernière MECE (figure 6a) : selon la « méthode classique équi-pondérée », 33 % de la MECE  
 202 présentent un très mauvais état écologique, contre 48 % suivant la « méthode améliorée », alors que  
 203 le pourcentage d'occurrence de référence dans cette classe de qualité est de 45 %. Dans ce cas, la

204 « méthode améliorée » apporte donc une précision non négligeable par rapport à la « méthode  
205 classique équi-pondérée ».

206 La même observation peut être faite pour les matières en suspension, lorsque le quantile temporel ne  
207 varie que légèrement au sein de la MECE. Dans le cas contraire, c'est souvent la « méthode  
208 classique équi-pondérée » qui fournit des pourcentages d'occurrence plus proches des pourcentages  
209 de référence.

210 Pour les orthophosphates, l'allure générale du quantile temporel de référence est approximativement  
211 esquissée par les deux méthodes, et il en est de même des pourcentages d'occurrence par classe de  
212 qualité (figure 7).



213  
214 **Figure 7 : Caractérisation de l'état des eaux selon le SEQ-Eau pour les orthophosphates ; les**  
215 **pourcentages d'occurrence des quantiles 90 temporels dans chaque classe de qualité sont**  
216 **indiqués par MECE.**

217 **Figure 7 : Characterization of water quality according to SEQ-Eau system for**  
218 **orthophosphates ; occurrence percentages of 90-percentiles in each quality class are given for**  
219 **each « River Water Body ».**

220 Mais les changements de classe du quantile temporel ne sont pas bien reproduits par ces deux  
221 méthodes. Or ce sont justement ces changements de classe, jusqu'alors déterminés principalement  
222 par expertise, que l'on cherche à positionner le plus fidèlement possible par rapport à la réalité. La  
223 « méthode améliorée » fournit toutefois des pourcentages d'occurrence plus proches des

224 pourcentages de référence, ce qui témoigne de l'intérêt de cette méthode par rapport à la « méthode  
225 classique ».

226 Cette première analyse montre donc que les résultats sont analogues ( $\text{NO}_3^-$ ,  $\text{O}_{2d}$ ), ou quand ils sont  
227 différents ( $\text{NH}_4^+$ ,  $\text{PO}_4$ , MES), c'est globalement la « méthode améliorée » qui donne des pourcentages  
228 d'occurrence plus proches de la référence.

### 229 3.2 Confrontation quantitative des résultats

230 Afin d'affiner l'analyse, le tableau 1 présente la comparaison quantitative des quantiles 90 obtenus à  
231 l'aide des deux méthodes par rapport au quantile de référence. Outre les différences par rapport à la  
232 référence, ce tableau reprend également les écarts (valeur absolue des différences) relatifs au  
233 quantile de référence ainsi que les écarts quadratiques absolus.

| [X]                                | <i>Méthode classique pénalisante</i> |                   |  | <i>Méthode classique pondérée</i> |                   |  | <i>Méthode améliorée pondérée</i> |                   |  |
|------------------------------------|--------------------------------------|-------------------|--|-----------------------------------|-------------------|--|-----------------------------------|-------------------|--|
|                                    | E rel.<br>[mgX/l]                    | D rel.<br>[mgX/l] | E <sup>2</sup> abs.<br>[(mgX/l) <sup>2</sup> ] | E rel.<br>[mgX/l]                 | D rel.<br>[mgX/l] | E <sup>2</sup> abs.<br>[(mgX/l) <sup>2</sup> ] | E rel.<br>[mgX/l]                 | D rel.<br>[mgX/l] | E <sup>2</sup> abs.<br>[(mgX/l) <sup>2</sup> ] |
| <b>NO<sub>3</sub><sup>-</sup></b>  | 0,036                                | 0,017             | 2,073  | 0,023                             | -0,014            | 0,676  | 0,027                             | 0,016             | 0,808  |
| <b>NH<sub>4</sub><sup>+</sup></b>  | 1,449                                | 1,438             | 6,812  | 0,142                             | -0,032            | 0,271  | 0,149                             | 0,089             | 0,240  |
| <b>PO<sub>4</sub><sup>3-</sup></b> | 0,496                                | 0,491             | 0,132  | 0,072                             | -0,023            | 0,008  | 0,098                             | 0,060             | 0,010  |
| <b>O<sub>2d</sub></b>              | 0,038                                | 0,028             | 0,249  | 0,021                             | 0,003             | 0,082  | 0,044                             | 0,036             | 0,332  |
| <b>MES</b>                         | 0,091                                | 0,001             | 10,536   | 0,203                             | -0,139            | 45,065   | 0,199                             | 0,058             | 41,669   |

234 **Tableau 1 : Ecart relatifs (E.rel.), différences relatives (D.rel.) et écarts quadratiques absolus**  
235 **(E<sup>2</sup>.abs.) moyens entre quantiles 90 calculés et quantiles de référence.**

236 **Table 1 : Average relative deviations (E. rel), relative differences (D. rel) and absolute squared**  
237 **deviations (E<sup>2</sup>. abs.) between 90-percentiles calculated and 90-percentile of reference.**

238 Ces statistiques sont présentées en moyenne sur tout le domaine étudié. Pour garantir la pertinence  
239 d'une telle comparaison, la spatialisation des quantiles calculés par les deux méthodes s'est avérée  
240 nécessaire étant donné le faible nombre de stations par MECE. Si les quantiles calculés à l'aide de la  
241 « méthode améliorée pondérée » sont connus en tout point via les segments d'influence spatiaux, ce

242 n'est pas le cas des quantiles fournis par la « méthode classique », uniquement disponibles aux  
243 stations. Deux méthodes de spatialisation de ces derniers quantiles ont été testées : la « méthode  
244 pénalisante » et la « méthode pondérée » (cf. figures 1 & 6).

245 Dès lors, la comparaison des deux premières colonnes (« méthode classique pénalisante » vs.  
246 « méthode classique pondérée ») permet de juger de l'importance de la *spatialisation* des quantiles  
247 temporels, tandis que la confrontation des deux dernières (« méthode classique pondérée » vs.  
248 « méthode améliorée pondérée ») traduit l'influence de la *pondération* des données et de la  
249 *linéarisation* de la fonction de quantile empirique.

250 Toutes variables confondues, la « méthode pondérée » donne généralement de meilleurs résultats en  
251 termes d'écart au quantile de référence, même si l'amélioration reste modeste lorsque la « méthode  
252 pénalisante » fournit déjà de faibles valeurs. C'est par exemple le cas des nitrates et de l'oxygène  
253 dissous, variables qui, rappelons-le, fluctuent peu sur le domaine étudié.

254 En revanche, pour des indicateurs avec une forte variabilité spatiale, tels que les orthophosphates ou  
255 l'ammonium, qui présentent des écarts relatifs plus importants, l'apport de la « méthode pondérée »  
256 devient plus notable : les écarts sont réduits d'un ordre de grandeur. Pour l'ammonium par exemple,  
257 l'écart relatif moyen qui était égal à 145 % avec la « méthode classique pénalisante », ne vaut plus que  
258 14 % avec la « méthode classique pondérée ».

259 Ces observations montrent la grande importance de la spatialisation des indices temporels.

260 Par contre, pour les matières en suspension, les quantiles 90 sont plus proches des quantiles de  
261 référence lorsqu'ils sont calculés par la « méthode classique pénalisante ». Ceci pourrait s'expliquer  
262 par la dynamique contrastée des matières en suspension, qui présentent une concentration faible la  
263 plupart du temps, mais des concentrations très élevées épisodiquement, notamment lors  
264 d'évènements pluvieux. Pour que le quantile 90 calculé soit représentatif du quantile de référence, il  
265 faudrait que ces épisodes aient été échantillonnés, ce qui est loin d'être évident avec une seule  
266 mesure par mois, ce qui correspond à la fréquence d'échantillonnage moyenne du RNB. Dès lors,  
267 contrairement à la « méthode pondérée », qui risque de faire intervenir des quantiles 90 sous-estimés  
268 lorsqu'un épisode n'a pas été échantillonné, la « méthode pénalisée », qui ne retient que le quantile  
269 90 le plus dégradant, a plus de chance d'être proche du quantile de référence, ce dernier étant  
270 influencé par de plus nombreux pics de concentrations.

271 Enfin, les écarts par rapport à la référence obtenus respectivement à l'aide des méthodes  
272 « classique » et « améliorée » présentent approximativement le même ordre de grandeur. Les  
273 matières en suspension mises à part, c'est généralement la « méthode classique » qui fournit des  
274 écarts légèrement plus faibles.

275 En conclusion, ces résultats quantitatifs montrent l'adéquation de la « méthode pondérée » pour la  
276 spatialisation du quantile 90. Bien que la « méthode améliorée » soit théoriquement la plus adaptée  
277 (Bernard-Michel, 2006), l'amélioration induite par la linéarisation et la pondération du quantile  
278 empirique est ici moins flagrante étant donné l'échantillonnage temporel presque régulier – avec  
279 quelques lacunes – des variables sur le RNB. Avec 10 mesures ou plus par an, l'irrégularité des dates  
280 de prélèvement engendre des conséquences beaucoup plus graves sur le calcul du quantile que leur  
281 faible nombre, et donc la pondération est plus importante que la linéarisation.

282 En outre, cette analyse a également permis de montrer l'importance de la localisation spatiale des  
283 stations, et donc la nécessité de spatialiser les quantiles temporels.

### 284 **3.3 Sur un réseau moins informé**

285 Le RNB étant particulièrement dense aux alentours de Paris sur le domaine modélisé par ProSe, il est  
286 intéressant de se ramener au cas d'un réseau plus clairsemé, plus proche des conditions habituelles  
287 de calcul d'un indice de qualité d'une MECE. La plupart des masses d'eau ne comptant en effet  
288 qu'une ou deux stations informées mensuellement, nous avons arbitrairement 'écarté' une station sur  
289 deux sur le domaine d'étude.

290 Le tableau 2 présente les moyennes des écarts, différences et écarts quadratiques obtenus à partir de  
291 ce réseau moins informé (une station sur deux), pour les deux méthodes de calcul du quantile  
292 temporel (« classique » et « améliorée ») et pour deux méthodes de spatialisation (« pénalisante » et  
293 « pondérée », cf. figures 1 & 6). Par rapport à l'échantillonnage effectivement disponible, la  
294 dégradation des résultats est sensible : les écarts par rapport à la référence sont plus marqués  
295 lorsque l'on ne retient qu'une station sur deux.

296 La « méthode pondérée » montre des écarts nettement plus faibles que la « méthode pénalisante »  
297 pour l'ammonium, les orthophosphates et également pour l'oxygène dissous. C'est encore une fois  
298 l'inverse pour les matières en suspension, tandis que les écarts en nitrates ne semblent pas affectés  
299 par la méthode de spatialisation.

300 Les méthodes « classique » et « améliorée » fournissent des résultats quasiment semblables et très  
 301 proches de la référence. Cela s'explique comme précédemment par la régularité des données.

| [X]                                | <i>Méthode classique pénalisante</i> |                   |  | <i>Méthode classique pondérée</i> |                   |  | <i>Méthode améliorée pondérée</i> |                   |  |
|------------------------------------|--------------------------------------|-------------------|--|-----------------------------------|-------------------|--|-----------------------------------|-------------------|--|
|                                    | E rel.<br>[mgX/l]                    | D rel.<br>[mgX/l] | E <sup>2</sup> abs.<br>[(mgX/l) <sup>2</sup> ] | E rel.<br>[mgX/l]                 | D rel.<br>[mgX/l] | E <sup>2</sup> abs.<br>[(mgX/l) <sup>2</sup> ] | E rel.<br>[mgX/l]                 | D rel.<br>[mgX/l] | E <sup>2</sup> abs.<br>[(mgX/l) <sup>2</sup> ] |
| <b>NO<sub>3</sub><sup>-</sup></b>  | 0,028                                | -0,001            | 0,946  | 0,028                             | -0,018            | 1,047  | 0,028                             | 0,008             | 1,014  |
| <b>NH<sub>4</sub><sup>+</sup></b>  | 1,442                                | 1,375             | 6,779  | 0,441                             | 0,240             | 1,854  | 0,410                             | 0,311             | 1,841  |
| <b>PO<sub>4</sub><sup>3-</sup></b> | 0,495                                | 0,479             | 0,131  | 0,133                             | 0,022             | 0,025  | 0,147                             | 0,107             | 0,029  |
| <b>O<sub>2d</sub></b>              | 0,036                                | 0,024             | 0,224  | 0,024                             | 0,006             | 0,097  | 0,046                             | 0,039             | 0,383  |
| <b>MES</b>                         | 0,109                                | -0,017            | 14,394   | 0,202                             | -0,123            | 43,630   | 0,229                             | 0,093             | 52,095   |

302 **Tableau 2 : Ecart E, différences D, et écarts quadratiques E2 moyens entre quantiles 90**  
 303 **calculés et quantiles de référence en conservant une station du RNB sur deux.**

304 *Table 2 : Average relative deviations (E. rel), relative differences (D. rel) and absolute squared*  
 305 *deviations (E<sup>2</sup>. abs.) between 90-percentiles calculated and 90-percentile of reference when*  
 306 *considering half measurement sites.*

#### 307 **4 DISCUSSION ET CONCLUSION**

308 Dans ce travail, nous nous sommes intéressés à l'évolution longitudinale des quantiles 90 temporels,  
 309 et avons testé différentes méthodes de spatialisation des quantiles pour tenter de reconstituer au  
 310 mieux cette évolution spatiale. Les statistiques moyennes sur le domaine modélisé par ProSe, toutes  
 311 MECE confondues, montrent que la « méthode pondérée » fournit presque systématiquement des  
 312 écarts moindres par rapport à la référence, notamment lorsque la « méthode pénalisante » conduit à  
 313 des écarts non négligeables. En revanche, aucune différence notable n'a été observée entre la  
 314 « méthode classique » et la « méthode améliorée », en raison de la régularité temporelle des  
 315 mesures. La « méthode améliorée » reste cependant recommandée en vertu de sa robustesse aux  
 316 lacunes et aux échantillonnages préférentiels (Bernard-Michel, 2006).

317 Les pourcentages d'occurrence dans les différentes classes de qualité, plus riches en information  
 318 qu'un unique indice par MECE, sont quant à eux mieux estimés (plus proches des pourcentages de

319 référence) lorsqu'ils sont calculés via la « méthode améliorée pondérée ». Cette amélioration est  
320 d'autant plus marquée que les quantiles 90 sont variables au sein d'une MECE.

321 Ce travail a porté sur la caractérisation de la variation spatiale des quantiles 90 temporels au sein des  
322 MECE. Nous disposons donc de l'évolution longitudinale de ces quantiles sur la portion de réseau  
323 hydrographique retenue pour ce travail. Cette portion de réseau comporte 6 MECE, pour certaines  
324 desquelles le quantile 90 fluctue entre différentes classes de qualité. Or la DCE demande de  
325 caractériser chacune de ces MECE par un unique indice de qualité. Se pose alors la question de  
326 l'indice à choisir pour synthétiser cette information spatiale : sur quels principes en fonder le calcul ?

327 Faut-il retenir une moyenne spatiale des quantiles temporels par MECE ? Dans ce cas, on pourrait se  
328 demander pourquoi ne pas retenir une moyenne temporelle de quantiles spatiaux (par MECE).

329 Une autre perspective serait d'appliquer le « principe de défaillance » temporellement et spatialement.  
330 En effet ce « principe de défaillance » conduit actuellement à retenir le quantile 90 (tri)annuel comme  
331 indicateur statistique pour les paramètres physico-chimiques en chaque station (Circulaire  
332 DCE\_2005\_12, communication personnelle Louis-Charles OUDIN, 2003). Retenir le quantile 90  
333 temporel en une station revient, en termes de qualité d'eau, à retenir une valeur plus exigeante que la  
334 moyenne. Faut-il élargir ce « principe de défaillance » spatialement au sein d'une MECE et calculer un  
335 quantile spatio-temporel ?

336 Enfin, la forte variabilité spatiale de certains quantiles temporels est quelque peu surprenante étant  
337 donné le critère d'homogénéité sur lequel est basé le découpage des MECE. On pourrait dès lors se  
338 demander si le découpage des MECE ne devrait pas être évolutif, c'est-à-dire s'adapter à l'apparition  
339 ou à la disparition de singularités ? La station d'épuration Seine Aval, par exemple, mériterait d'être  
340 considérée comme une discontinuité en raison de la variation abrupte des concentrations en  
341 ammonium et orthophosphates qu'elle provoque (cf figures 6a) & 7, pk ≈ 720).

## 342 **5 REMERCIEMENTS**

343 Les auteurs souhaitent exprimer leur gratitude au personnel du Service de Navigation de la Seine  
344 (SNS), ainsi qu'à celui de la Compagnie Générale des Eaux (CGE) et du Syndicat Interdépartemental  
345 d'Assainissement de l'Agglomération Parisienne (SIAAP), pour avoir fourni les données nécessaires à  
346 ce travail.

347 **6 REFERENCES BIBLIOGRAPHIQUES**

348 BERNARD-MICHEL C. et C. de FOUQUET (2005) Geostatistical indicators of waterway quality for  
349 nutrients, *VIIIth Congress of Geostatistics*, du 26 septembre au 1 octobre 2004, Banff, Alberta,  
350 Canada, Compte rendu, n°2, p. 907-912.

351 BERNARD-MICHEL C. (2006) *Indicateurs géostatistiques de la pollution dans les cours d'eau*. Thèse  
352 de Doctorat, spécialité géostatistique, ENSMP, 161 p.

353 BILLEN G., J. GARNIER et P. HANSET (1994) Modelling phytoplankton development in whole  
354 drainage networks: the RIVERSTRAHLER model applied to the Seine river system, *Hydrobiologia*,  
355 n°289, p. 119–137.

356 EVEN S., M. POULIN, J. GARNIER, G. BILLEN, P. SERVAIS, A. CHESTERIKOFF et M. COSTE  
357 (1998) River ecosystem modeling: application of the PROSE model to the Seine River (France),  
358 *Hydrobiologia*, n°373/374, p. 27-45.

359 EVEN S., M. POULIN, J.-M. MOUCHEL, M. SEIDL et P. SERVAIS (2004) Modelling oxygen deficits in  
360 the Seine river downstream of combined sewer overflows. *Ecol. Model.* n°173, p. 177-196.

361 EVEN S., J.-M. MOUCHEL, P. SERVAIS, N. FLIPO, M. POULIN, S. BLANC, M. CHABANEL et C.  
362 PAFFONI (2007) Modeling the impacts of Combined Sewer Overflows on the river Seine water quality.  
363 *Sci. Total. Environ.*, n°375, p.140–51.

364 FLIPO N. (2005) *Modélisation intégrée des transferts d'azote dans les aquifères et les rivières :  
365 Application au bassin du Grand Morin*. Thèse de Doctorat, spécialité hydrologie et hydrogéologie  
366 quantitatives, ENSMP, 178 p. [http://pastel.paristech.org/1432/01/these\\_flipo.pdf](http://pastel.paristech.org/1432/01/these_flipo.pdf)

367 FLIPO N., S. EVEN, M. POULIN, M.-H. TUSSEAU-VUILLEMIN, T. AMEZIANE et A. DAUTA  
368 (2004).Biogeochemical modelling at the river scale: plankton and periphyton dynamics (Grand Morin  
369 case study, France), *Ecol. Model.*, n°176(3-4), p. 333-347.

370 FLIPO N., C. RABOUILLE, M. POULIN, S. EVEN, M.-H. TUSSEAU-VUILLEMIN et M. LALANDE  
371 (2007) Primary production in headwater streams of the Seine basin: the Grand Morin river case study,  
372 *Sci. of Total. Environ.*, n°375, p. 98-109.

373 GARNIER J., G. BILLEN et M. COSTE (1995) Seasonal succession of diatoms and chlorophyceae in  
374 the drainage network of the river Seine: observations and modelling, *Limnol. Oceanogr.*, n°40(4), p.  
375 750–765.

376 POULIN M. (2006) *Réalisation de simulations PROSE, Année de référence 2003*, rapport de l'Ecole  
377 des Mines de Paris pour le Syndicat Interdépartemental d'Assainissement de l'Agglomération  
378 Parisienne, Référence R060629, 83 p.

379 Préfecture de l'Ain, Directive Cadre sur l'Eau. Circulaire DCE\_2005\_12 :  
380 [www.ain.pref.gouv.fr/DDAF/ODE/politique/docs/circulaire-dce-bee.pdf](http://www.ain.pref.gouv.fr/DDAF/ODE/politique/docs/circulaire-dce-bee.pdf) (consultation en novembre 2007)

381 Service d'Administration Nationale des Données et Référentiels sur l'Eau (Sandre), Directive Cadre  
382 Eau : [http://sandre.eaufrance.fr/rubrique.php3?id\\_rubrique=21](http://sandre.eaufrance.fr/rubrique.php3?id_rubrique=21) (consultation en janvier 2008)