

Société de Calcul Mathématique, S. A.
Algorithmes et Optimisation



Qualité des cours d'eau en Artois Picardie :
Paramètres explicatifs et recherche d'anomalies

Résumé du travail réalisé pour

l'Agence de l'Eau Artois-Picardie

par la

Société de Calcul Mathématique S. A.

Janvier 2009

La Société de Calcul Mathématique a réalisé une étude pour l'Agence de l'Eau Artois Picardie (AEAP) : l'objectif est d'évaluer la qualité des eaux à l'aide de paramètres externes, pour expliquer la non-conformité, par rapport aux normes européennes, qui est observée dans certains cas.

Ces paramètres externes sont faciles à mesurer, et donc le bénéfice est clair : on peut limiter les mesures aux endroits que les paramètres externes auront désignés comme fragiles. Dans le cas de l'AEAP, la réduction serait d'environ 30 % du nombre des mesures, pour chacun des trois polluants auxquels nous avons appliqué la méthode.

Nous présentons ici une synthèse de la méthode mise en place et du travail réalisé.

1. Principe de la méthode

La méthode consiste à étudier un certain nombre de paramètres externes (liés à la nature des cours d'eau, des terrains et de la densité de population), susceptibles d'avoir une influence sur la qualité de l'eau. L'analyse croisée de données relatives à ces paramètres et à la qualité de l'eau permet de déterminer les paramètres influents.

Ces données permettent de définir les Situations Génériques de Bonne Qualité de certains polluants, en termes de paramètres influents : lorsque certaines conditions concernant les valeurs des paramètres externes prépondérants sont remplies, la probabilité que l'eau soit de bonne qualité est élevée.

On peut alors obtenir une « cartographie de la sensibilité » de la zone étudiée, en connaissant uniquement la valeur des paramètres externes influents : on peut repérer des zones tranquilles, où l'eau sera de bonne qualité avec une forte probabilité. Sur ces zones, on pourra diminuer le nombre de points de mesures, ou la fréquence des relevés. D'autres points seront plus fragiles : on peut prévoir que l'eau y sera de mauvaise qualité. Ces points mériteront une attention particulière.

La méthode n'est pas parfaite, en ce sens que subsistent des "anomalies" : points que les paramètres externes désignent comme fragiles, alors qu'aucune pollution n'y est réellement observée, et à l'inverse points qui ne sont pas fragiles mais qui sont pollués. Dans l'étude menée pour l'AEAP, les anomalies de première espèce sont peu nombreuses et proches du seuil : on en conclut que la méthode que nous avons introduite est sévère ; on ne pourra la considérer comme laxiste. Elle permet la définition d'une "Situation Générique de Bonne Qualité", de manière simple, économique et explicite. Elle ne se base que sur des données externes : ces informations sont faciles à observer et à stocker, et indépendantes de la mesure elle-même.

Elle repose sur un historique de la pollution (on se sert des données telles qu'elles sont) : aucun ajustement ni test statistique ne vient « biaiser » l'analyse.

2. Données nécessaires

Il est nécessaire de disposer de l'ensemble des données suivantes, en différents points de mesure de la zone à étudier :

- Les teneurs en polluant ;
- La valeur des paramètres externes dont on cherche à évaluer l'influence.

Ces données servent à déterminer les paramètres influents et à définir les Situations Génériques de Bonne Qualité. Ceci permet par la suite de réduire le nombre de mesures nécessaires, en supprimant des points de mesure dans les zones remplissant les conditions des SGBQ.

3. Travail réalisé

L'Agence de l'Eau Artois Picardie dispose de données relatives à la qualité de l'eau en 105 points de mesure du bassin Artois-Picardie, pour certains polluants (DBO5, DCO, NH₄, NTK, NO₃ et PO₄). Voici la carte du réseau :

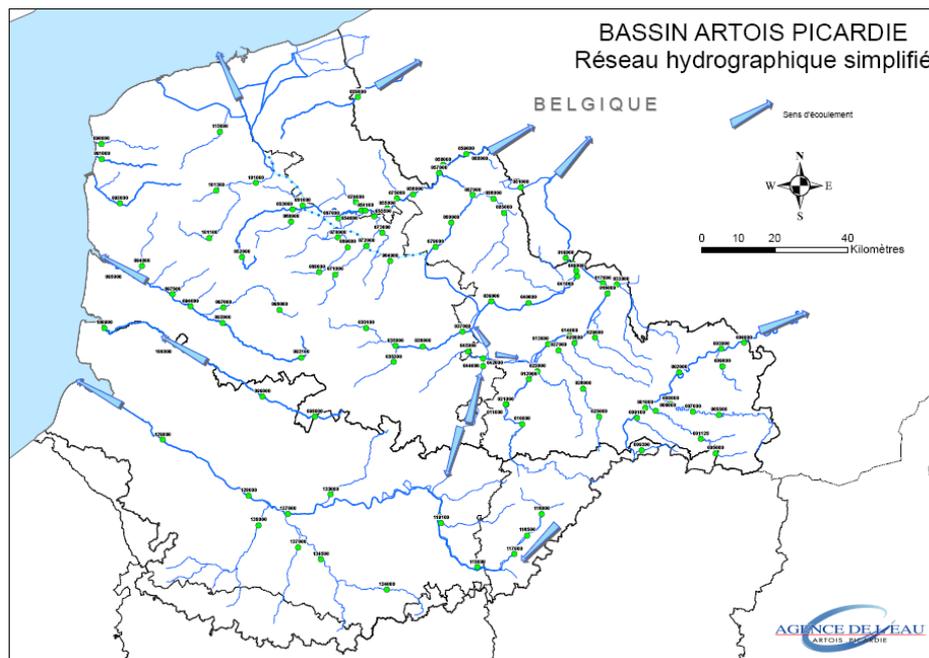


Figure 1 : Réseau hydrographique du bassin Artois Picardie et points de mesure

La qualité des mesures, par polluant, se répartit comme suit :

Nombre de mesures	DBO5	DCO	NH ₄	NTK	NO ₃	PO ₄
Très bonne qualité	75	1	20	44	5	7
Bonne qualité	28	90	48	35	99	61
Mauvaise qualité	2	14	37	26	1	37

Tableau 1 : Nombre de mesures par qualité pour les 6 polluants

3.1. Paramètres influents

Les paramètres externes étudiés sont les suivants :

	Paramètres pris en compte :
Caractéristiques de la population	<ul style="list-style-type: none"> • densité de population
Caractéristiques du cours d'eau	<ul style="list-style-type: none"> • débit moyen • débit spécifique par unité de surface • débit moyen par million d'habitants • régime hydrologique (régulier ou irrégulier) • nature du cours d'eau (fleuve, canal)
Nature du terrain	<ul style="list-style-type: none"> • % cultures permanentes • % forêts et milieux semi-naturels • % prairies • % terres arables • % territoires artificiels • % zones agricoles hétérogènes • % zones humides ou surfaces en eau

Tableau 2 : Paramètres pris en compte

En chaque point de mesure, nous connaissons la valeur de ces paramètres, ainsi que la concentration des six polluants retenus.

Le traitement des données est probabiliste. La méthode consiste à diviser la plage de valeurs possibles de chaque paramètre en deux intervalles : valeurs basses (inférieures ou égales à la médiane des valeurs) et valeurs hautes (supérieures à la médiane). Les deux intervalles contiennent le même nombre de mesures. Dans ces deux intervalles, nous calculons le pourcentage de mesures respectant les seuils de bonne et très bonne qualité (pour chaque polluant). Cela revient à construire la fonction de répartition de la concentration en polluant.

Pour déterminer l'influence d'un paramètre sur la teneur en polluant, nous comparons le pourcentage de mesures de bonne (ou très bonne) qualité dans les deux intervalles de valeurs possibles du paramètre. Si ces pourcentages sont du même ordre de grandeur, alors le paramètre n'a pas d'influence sur la qualité de l'eau pour ce polluant ; si au contraire on observe un écart important entre les deux pourcentages, cela signifie que le paramètre a un impact sur la qualité de l'eau.

Les résultats obtenus en Artois Picardie sont donnés dans le tableau suivant :

Polluant	Paramètres qui ont un impact majeur par ordre d'impact décroissant	
	Sur les mesures de bonne qualité	Sur les mesures de très bonne qualité
DBO5	Pas de paramètre prépondérant	<ul style="list-style-type: none"> • Débit moyen par habitant • Pourcentage de territoires artificialisés • Densité de population • Nature du cours d'eau
DCO	Pas de paramètre prépondérant	Pas de paramètre prépondérant
NH ₄	<ul style="list-style-type: none"> • Débit moyen par habitant • Densité de population • Pourcentage de territoires artificialisés • Nature du cours d'eau • Pourcentage de prairies 	<ul style="list-style-type: none"> • Pourcentage de territoires artificialisés • Densité de population
NTK	<ul style="list-style-type: none"> • Débit moyen par habitant • Densité de population • Pourcentage de territoires artificialisés • Régime hydrologique 	<ul style="list-style-type: none"> • Densité de population • Pourcentage de territoires artificialisés • Régime hydrologique • Débit moyen par habitant • Nature du cours d'eau
NO ₃	Pas de paramètre prépondérant	Pas de paramètre prépondérant
PO ₄	<ul style="list-style-type: none"> • Pourcentage de territoires artificialisés • Densité de population • Débit moyen par habitant • Pourcentage de prairies 	Pas de paramètre prépondérant

Tableau 3 : Paramètres influant majoritairement les différents polluants

Cette étude met en avant six paramètres influents en Artois Picardie : la densité de population, le débit moyen par habitant, le pourcentage de prairies, le pourcentage de territoires artificialisés, le régime hydrologique (régulier ou irrégulier), la nature du cours d'eau (canalisé ou non).

Chaque polluant a ses propres paramètres influents. Pour certains polluants, aucun des paramètres externes retenus n'influe sur la qualité ; pour les autres, de deux à cinq paramètres sont influents, parmi les six cités dans le paragraphe précédent.

Les résultats sont obtenus sans aucun test statistique, sans aucun ajustement, linéaire ou non : il s'agit simplement de construire une distribution de probabilité dans les différents cas ; ils sont donc robustes, puisque l'on se contente d'utiliser les données telles qu'elles viennent.

3.2. Situations Génériques de Bonne Qualité

Lorsque l'étude précédente a permis de déterminer des paramètres influents, on peut définir, pour chaque polluant, des critères de "Situation Générique de Bonne Qualité" (SGBQ), en termes de paramètres externes influents.

Considérons par exemple le polluant PO_4 . Les paramètres influençant la qualité de l'eau pour ce polluant sont le débit moyen par habitant, le pourcentage de prairies et le pourcentage de territoires artificialisés. La répartition des mesures pour ces trois paramètres est indiquée dans le tableau suivant :

Paramètres	Valeurs du paramètre	Pourcentage de mesures de bonne qualité
Débit moyen par habitant	$\leq 43 m^3 s^{-1} 10^{-6} hab^{-1}$	39 %
	$> 43 m^3 s^{-1} 10^{-6} hab^{-1}$	89 %
Pourcentage de prairies	$\leq 10\%$	46 %
	$> 10\%$	83 %
Pourcentage de territoires artificialisés	$\leq 6\%$	95 %
	$> 6\%$	44 %

Tableau 4 : Résultats de l'étude de l'influence des paramètres sur les mesures en PO_4

Le tableau se lit comme suit : pour un débit moyen par habitant inférieur à $43 m^3 s^{-1} 10^{-6} hab^{-1}$ (valeur médiane des données), 39 % des mesures de PO_4 sont bonnes. Lorsque le débit moyen par habitant est élevé, 89 % des mesures sont bonnes. Ainsi, lorsqu'un point de mesure est situé dans une zone où le débit moyen par habitant est élevé, on a deux fois plus de chance que l'eau soit de bonne qualité (vis-à-vis du polluant PO_4) que lorsque le débit moyen par habitant est faible.

En effectuant la même analyse sur les trois paramètres, nous obtenons les critères de bonne qualité suivants :

- Débit moyen par habitant élevé (supérieur à $43 m^3 s^{-1} 10^{-6} hab^{-1}$) ;
- Pourcentage de prairies élevé (supérieur à 10 %) ;
- Pourcentage de territoires artificialisés faible (inférieur à 6 %).

C'est ce que l'on appelle la Situation Générique de Bonne Qualité pour le polluant PO_4 : si un point présente ces caractéristiques, la probabilité que l'eau soit de bonne qualité vis à vis du polluant PO_4 est élevée.

Aucun des paramètres externes retenus n'a d'influence sur la bonne qualité des mesures pour les polluants DBO₅, DCO et NO_3 ; nous ne pouvons donc pas définir de SGBQ pour ces trois polluants. Pour les trois autres polluants (NH_4 , NTK et PO_4), les SGBQ sont données dans le tableau suivant :

Polluant	Situation Générique de Bonne Qualité
NH_4	<ul style="list-style-type: none"> • Débit moyen par habitant élevé ($> 43 m^3 s^{-1} 10^{-6} hab^{-1}$) • Pourcentage de territoires artificialisés faible ($\leq 6\%$) • Nature du cours d'eau : fleuve • Pourcentage de prairies élevé ($> 10\%$)
NTK	<ul style="list-style-type: none"> • Débit moyen par habitant élevé ($> 43 m^3 s^{-1} 10^{-6} hab^{-1}$) • Pourcentage de territoires artificialisés faible ($\leq 6\%$) • Régime hydrologique régulier
PO_4	<ul style="list-style-type: none"> • Pourcentage de territoires artificialisés faible $\leq 6\%$ • Débit moyen par habitant élevé ($> 43 m^3 s^{-1} 10^{-6} hab^{-1}$) • Pourcentage de prairies élevé ($> 10\%$)

Tableau 5 : Situation générique de bonne qualité pour les polluants NH_4 , NTK et PO_4

3.3. Anomalies

Cependant, ces situations génériques ne sont pas parfaites. Elles couvrent la plupart des cas, ce qui est satisfaisant, mais présentent deux sortes d'exceptions :

- Des points où les critères de SGBQ ne sont pas satisfaits, mais où la qualité de mesure est cependant bonne (on s'attendrait à une pollution, du fait des circonstances externes, mais elle n'existe pas) ;
- Des points où les critères de SGBQ sont satisfaits, mais où la mesure est de mauvaise qualité : tous les paramètres externes sont "au vert", mais cependant il y a une pollution, que l'on ne sait pas expliquer.

Ces "anomalies " sont décrites dans le tableau suivant :

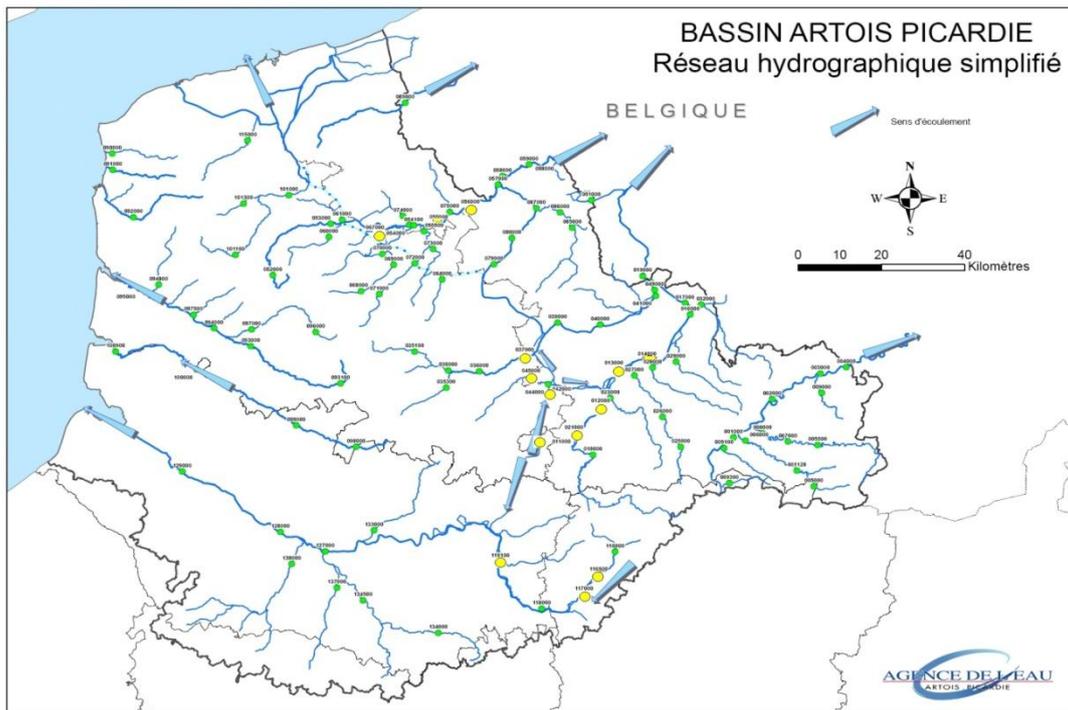
Polluant	NH ₄	NTK	PO ₄
Nombre de mesures remplissant les critères de la SGBQ	31	31	34
Nombre total de points de mesure	105	105	105
% du total	30%	30%	32%
Nombre de mesures de mauvaise qualité remplissant les critères de la SGBQ : pollution inattendue	2	0	1
Nombre total de mesures de mauvaise qualité	37	26	37
% du total	5%	0%	3%
Nombre de mesures de bonne qualité ne remplissant pas les critères de SGBQ : pollution attendue mais inexistante	4	3	10
Nombre total de mesures de bonne qualité et très bonne qualité	68	79	68
% du total	6%	4%	15%

Tableau 6 : Proportion d'anomalies

La méthode appliquée en Artois Picardie permettrait d'économiser près d'un tiers des mesures. Les anomalies sont peu nombreuses : pour les trois polluants étudiés, l'erreur est au plus de 5% (cas où il y a une pollution inattendue parmi les points désignés comme bons par les critères de la SGBQ).

En particulier deux points seulement sont mauvais alors qu'ils devraient être bons au regard des critères de la SGBQ (tous polluants confondus). En outre, dans les deux cas, on est proche du seuil de bonne qualité.

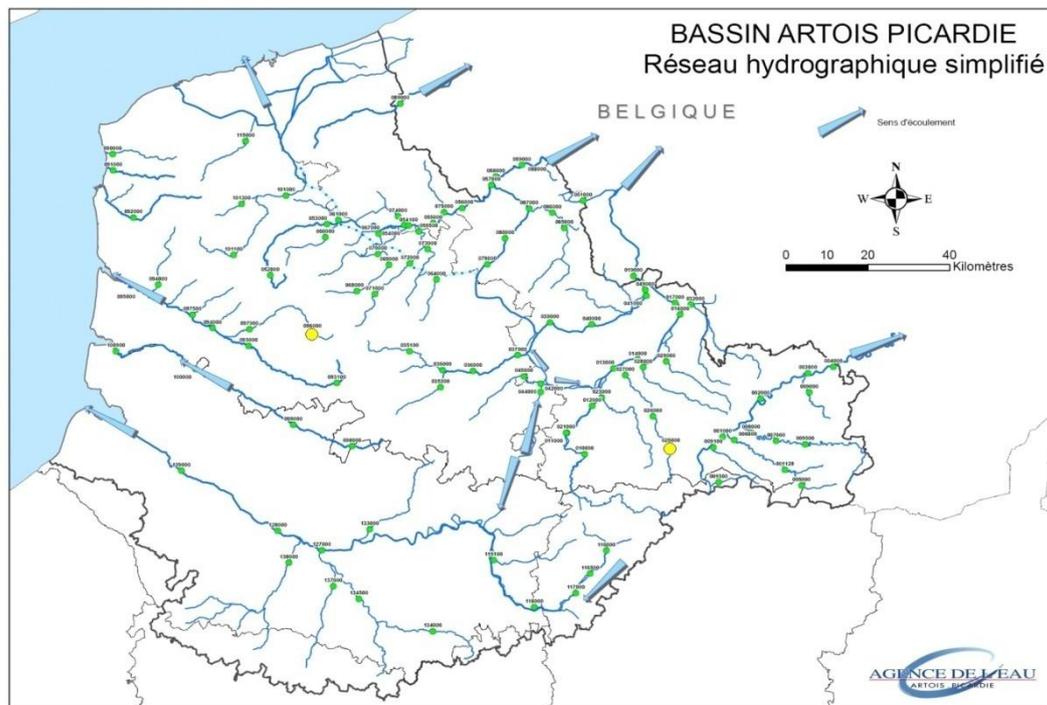
Voici la carte des anomalies "bonnes" :



Légende : ● Points d'anomalies bonnes pour les 3 polluants NT₄, NTK et PO₄

Figure 2 : Mesures désignées comme mauvaises par la méthode de la SGBQ mais non polluées

Et voici la carte des points particuliers "mauvais" :



Légende : ● Points d'anomalies mauvaises pour les 3 polluants NT₄, NTK et PO₄

Figure 3 : Mesures désignées comme bonnes par la méthode de la SGBQ mais polluées

Il y a assurément des anomalies de bonne qualité : points qui pourraient être mauvais au regard de nos critères, alors qu'ils sont bons. Mais ceci indique que nos critères sont stricts et nullement laxistes.

On peut en conclure que la notion de Situation Générique de Bonne Qualité, définie grâce à des paramètres externes faciles à mesurer, est un bon indicateur du niveau de pollution.

4. Variations et applications possibles

Cette méthode est applicable à tout polluant, et à toute zone d'étude. Pour l'utiliser, il suffit de connaître les teneurs en polluant et les valeurs de quelques paramètres externes en plusieurs points de mesure.

On peut prendre en compte d'autres paramètres que ceux étudiés pour l'AEAP. Il est également possible d'utiliser un historique de ces données sur plusieurs années.

La méthode pourrait s'appliquer à d'autres pollutions (air, sols, etc.) ou à l'évaluation d'autres flux (trafic routier ou maritime, etc.).