

NIVEAU DE CONFIANCE RELATIF A L'INDICATEUR REGLEMENTAIRE DE L'ETAT CHIMIQUE D'UNE MASSE D'EAU DE SURFACE : APPLICATION A DES JEUX DE DONNEES REELLES

Recommandation, Aide à la décision

A. YARI, J-P. GHESTEM (BRGM)

Avec la collaboration de
A. LEYNET (BRGM), B. LALERE, N. FISCHER (LNE),
F BOTTA (INERIS)

Décembre 2015

Programme scientifique et technique
Année 2015

Document final



Avec le soutien de



et de



Contexte de programmation et de réalisation

Ce rapport a été réalisé dans le cadre du programme d'activité AQUAREF pour l'année 2015 ...

Auteur (s) :

Anice Yari
BRGM
a.yari@brgm.fr

Jean Philippe GHESTEM
BRGM
jp.ghestem@brgm.fr

avec la collaboration de

Béatrice Lalère
Beatrice.lalere@lne.fr
Nicolas Fischer
Nicolas.fischer@lne.fr
LNE

Fabrizio Botta
Fabrizio.botta@ineris.fr
INERIS

Vérification du document :

Nolwenn Croiset
n.croiset@brgm.fr
BRGM

Les correspondants

Onema : I Barthe Franquin, DCIE, isabelle.barthe-franquin@onema.fr

Etablissement : JP Ghestem, BRGM, jp.ghestem@brgm.fr

Référence du document : A. Yari., JP. Ghestem, A. Leynet (BRGM), B. Lalère, N. Fischer (LNE), F. Botta (INERIS) (2015) - Niveau de confiance relatif à l'indicateur réglementaire de l'état chimique d'une masse d'eau de surface : application à des jeux de données réelles. Rapport AQUAREF 2015 - BRGM/RP-65033-FR, 35 p.

Droits d'usage :	Accès libre
Couverture géographique :	International
Niveau géographique :	National
Niveau de lecture :	Professionnels, experts
Nature de la ressource :	Document

1. Introduction	7
1.1. CONTEXTE	7
1.2. RAPPEL DES TRAVAUX AQUAREF 2013 [4]	8
1.3. OBJECTIFS	10
2. Descriptif des méthodes statistiques employées	11
2.1. METHODE « ANALYTIQUE »	11
2.2. METHODE DE CYRANO (« BOOTSTRAP ») [9]	12
3. Présentation des données étudiées	13
3.1. PRESENTATION GENERALE	13
3.2. DIFFICULTES RENCONTREES AVEC LES DONNEES DISPONIBLES	14
3.2.1. Absence des incertitudes de mesure	14
3.2.2. Absence et/ou multiplicité de la limite de quantification	15
3.2.3. Nombre variable de données mensuelles	16
3.3. CRITERES DE SELECTION DES DONNEES	16
3.3.1. Critère lié à la valeur de la LQ par rapport aux NQE	16
3.3.2. Critère lié au nombre de données par substance	17
4. Résultats et interprétation	19
4.1. BILAN DES RESULTATS OBTENUS POUR LE JEU DE DONNEES 1	19
4.2. BILAN DES RESULTATS OBTENUS POUR LE JEU DE DONNEES 2	22
4.3. BILAN DES RESULTATS OBTENUS POUR LE JEU DE DONNEES 3	25
5. Conclusion	29
6. Bibliographie	31

NIVEAU DE CONFIANCE RELATIF A L'INDICATEUR REGLEMENTAIRE DE L'ETAT CHIMIQUE D'UNE MASSE D'EAU DE SURFACE : APPLICATION A DES JEUX DE DONNEES REELLES
AUTEUR(S) : A YARI (BRGM) JP GHESTEM (BRGM)

RESUME

La Directive Cadre Européenne sur l'Eau (DCE) fixe aux Etats membres des objectifs de bon état chimique des masses d'eau. Cet état chimique réglementaire est défini entre autres grâce à la comparaison entre des concentrations de polluants (moyennes ou maximales) et des valeurs seuils réglementaires (NQE). Ces concentrations sont le résultat d'un processus d'acquisition de mesure incluant des étapes d'échantillonnage et d'analyse qui sont sources d'erreurs et d'incertitudes et entraînent in fine une incertitude sur l'évaluation de l'état chimique.

Le travail proposé s'inscrit dans le programme d'action AQUAREF pour l'année 2015. Il est réalisé en partenariat par le BRGM et le LNE avec la collaboration de l'INERIS dans le cadre de conventions de partenariat avec l'ONEMA. Il a pour objectif de poursuivre les travaux entamés en 2013 qui visent à proposer aux gestionnaires un indicateur du « niveau de confiance » relatif à l'évaluation réglementaire de l'état chimique d'une masse d'eau.

Dans le cadre du champ d'action d'AQUAREF, la réflexion proposée s'intéresse principalement à la contribution des incertitudes liées aux données acquises (incertitudes analytiques et dispersion entre données). L'objectif de cet « indicateur » est d'estimer :

- la probabilité que l'état chimique de la masse d'eau soit bien celui qui a été estimé
- et réciproquement, la probabilité que l'état chimique de la masse d'eau soit différent de celui qui a été estimé.

Le travail de 2013 proposait des principes et des méthodes statistiques. Le présent travail a pour objectif d'appliquer certaines de ces méthodes sur un nombre plus important de données et des données réelles afin d'être plus illustratif des différentes situations rencontrées par les gestionnaires. Parmi les 6 méthodes statistiques citées en 2013, deux méthodes ont été retenues pour ce travail : il s'agit d'une méthode « analytique » simple et de la méthode dite « Cyrano » (bootstrap).

L'évaluation de l'état chimique des eaux de surface, en moyenne annuelle (MA) ou en concentration maximale admissible (CMA), a servi de base pour ce travail. Les données utilisées dans ce rapport sont des données de surveillance provenant de 3 agences de l'eau. La première difficulté rencontrée a concerné les données et métadonnées disponibles. Quelle que soit la méthode statistique, les traitements se heurtent :

- à l'absence d'une information systématique sur l'incertitude de mesure au niveau du résultat,
- à l'absence d'une information systématique concernant la limite de quantification associée à chaque résultat.

Par ailleurs, la méthode Cyrano dans son application actuelle nécessite des jeux de données homogènes en ce qui concerne le nombre de données par substance. C'est une limite par rapport à la méthode « analytique ».

Les deux méthodes statistiques utilisées ont globalement donné des résultats concordants à la fois en moyenne annuelle et en concentration maximale admissible, qu'il s'agisse de la définition de l'état de la masse d'eau mais aussi de l'estimation de l'indice de confiance sur cet état. De façon générale, les niveaux de confiance estimés par les 2 méthodes permettraient donc d'atteindre l'objectif d'une utilisation « qualitative » de cet indice de confiance par les gestionnaires.

Il faut cependant mentionner quelques différences qui renvoient aux différences de principe des méthodes statistiques choisies. La méthode Cyrano que ce soit pour la moyenne annuelle ou pour la concentration maximale admissible donne des estimations qui ne correspondent pas strictement aux moyennes et concentrations maximales simplement définies de façon arithmétique. Ainsi, la méthode Cyrano a tendance à favoriser une estimation de « bon état » en diminuant la concentration maximale prise en compte.

Enfin, la méthode Cyrano présente la caractéristique par rapport à la méthode « analytique » de prendre en compte les corrélations entre substances dans l'estimation des indices de confiance. La méthode « analytique » considère, comme le prévoit la DCE, l'état chimique par substance sur une

année avant de déclarer l'état global de la station (que ce soit pour la moyenne ou pour l'indice de confiance). A l'inverse, la méthode Cyrano donne plus de poids à l'état de la masse d'eau à chaque période d'échantillonnage, notamment pour l'estimation de l'indice de confiance. Dans cette dernière méthode, plusieurs substances déclassant simultanément une station ont moins de poids que ces mêmes substances déclassant la station à deux dates différentes.

Les méthodes statistiques testées dans le présent rapport prennent en compte à la fois :

- des incertitudes d'origine analytique des mesures individuelles,
- la variabilité du milieu à travers l'utilisation de n mesures annuelles et de la dispersion de ces mesures.

Une hypothèse forte de la démarche est donc de considérer que les données de surveillance utilisées pour le calcul de l'indicateur réglementaire d'état permettent de rendre compte de façon fiable de la variabilité du milieu. Il n'est pas dans l'objectif de ce rapport d'étudier cette hypothèse et ses conséquences en termes de niveau de confiance sur l'état. Il convient donc de garder à l'esprit que les résultats présentés constituent une estimation basse de l'incertitude réelle affectant l'évaluation de l'état chimique, sans qu'il soit possible d'estimer l'écart ainsi impliqué.

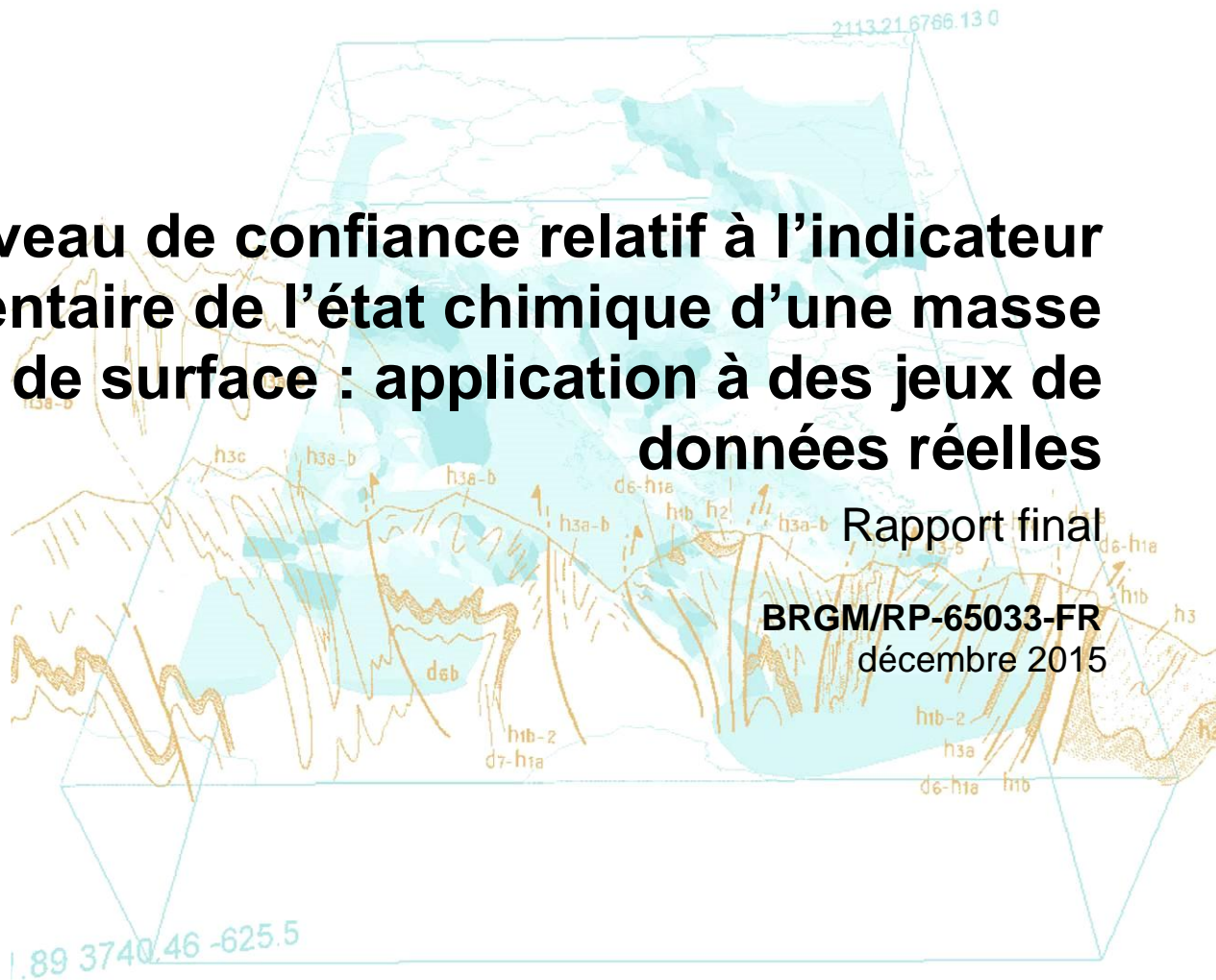
Mots clés (thématique et géographique) : niveau de confiance, masse d'eau, état chimique, incertitude



Niveau de confiance relatif à l'indicateur réglementaire de l'état chimique d'une masse d'eau de surface : application à des jeux de données réelles

Rapport final

BRGM/RP-65033-FR
décembre 2015



Niveau de confiance relatif à l'indicateur réglementaire de l'état chimique d'une masse d'eau de surface : application à des jeux de données réelles

Rapport final

BRGM/RP-65033-FR
décembre 2015

A. YARI, J-P. GHESTEM (BRGM)

Avec la collaboration de

A LEYNET (BRGM) B LALERE, N FISCHER (LNE) F BOTTA (INERIS)

1.89 3740.46 -625.5 Vérificateur :

Nom : N CROISET

Fonction : Chef de projet

Date : 11/01/2016

Signature :

Approbateur :

Nom : L AMALRIC

Fonction : Responsable d'unité

Date : 12/01/2016

Signature :

Le système de management de la qualité et de l'environnement est certifié par AFNOR selon les normes ISO 9001 et ISO 14001.



Géosciences pour une Terre durable

brgm

Mots-clés : incertitude, masse d'eau, état, fiabilité, statistiques, niveau de confiance

En bibliographie, ce rapport sera cité de la façon suivante :

YARI A., GHESTEM J-P. (2015) – Niveau de confiance relatif à l'indicateur réglementaire de l'état chimique d'une masse d'eau de surface : application à des jeux de données réelles. Rapport BRGM/RP-65033-FR, 37p., 1 ill, 5 tab, 1 ann.

Synthèse

La Directive Cadre Européenne sur l'Eau (DCE) fixe aux Etats membres des objectifs de bon état chimique des masses d'eau. Cet état chimique réglementaire est défini entre autres grâce à la comparaison entre des concentrations de polluants (moyennes ou maximales) et des valeurs seuils réglementaires (NQE). Ces concentrations sont le résultat d'un processus d'acquisition de mesure incluant des étapes d'échantillonnage et d'analyse qui sont sources d'erreurs et d'incertitudes et entraînent *in fine* une incertitude sur l'évaluation de l'état chimique.

Le travail proposé s'inscrit dans le programme d'action AQUAREF pour l'année 2015. Il est réalisé en partenariat par le BRGM et le LNE avec la collaboration de l'INERIS dans le cadre de conventions de partenariat avec l'ONEMA. Il a pour objectif de poursuivre les travaux entamés en 2013 qui visent à proposer aux gestionnaires un indicateur du « niveau de confiance » relatif à l'évaluation réglementaire de l'état chimique d'une masse d'eau.

Dans le cadre du champ d'action d'AQUAREF, la réflexion proposée s'intéresse principalement à la contribution des incertitudes liées aux données acquises (incertitudes analytiques et dispersion entre données). L'objectif de cet « indicateur » est d'estimer :

- la probabilité que l'état chimique de la masse d'eau soit bien celui qui a été estimé
- et réciproquement, la probabilité que l'état chimique de la masse d'eau soit différent de celui qui a été estimé.

Le travail de 2013 proposait des principes et des méthodes statistiques. Le présent travail a pour objectif d'appliquer certaines de ces méthodes sur un nombre plus important de données et des données réelles afin d'être plus illustratif des différentes situations rencontrées par les gestionnaires. Parmi les 6 méthodes statistiques citées en 2013, deux méthodes ont été retenues pour ce travail : il s'agit d'une méthode « analytique » simple et de la méthode dite « Cyrano » (bootstrap).

L'évaluation de l'état chimique des eaux de surface, en moyenne annuelle (MA) ou en concentration maximale admissible (CMA), a servi de base pour ce travail. Les données utilisées dans ce rapport sont des données de surveillance provenant de 3 agences de l'eau. La première difficulté rencontrée a concerné les données et métadonnées disponibles. Quelle que soit la méthode statistique, les traitements se heurtent :

- à l'absence d'une information systématique sur l'incertitude de mesure au niveau du résultat,
- à l'absence d'une information systématique concernant la limite de quantification associée à chaque résultat.

Par ailleurs, la méthode Cyrano dans son application actuelle nécessite des jeux de données homogènes en ce qui concerne le nombre de données par substance. C'est une limite par rapport à la méthode « analytique ».

Les deux méthodes statistiques utilisées ont globalement donné des résultats concordants à la fois en moyenne annuelle et en concentration maximale admissible, qu'il s'agisse de la définition de l'état de la masse d'eau mais aussi de l'estimation de l'indice de confiance sur cet état. De façon générale, les niveaux de confiance estimés par les 2 méthodes permettraient donc d'atteindre l'objectif d'une utilisation « qualitative » de cet indice de confiance par les gestionnaires.

Il faut cependant mentionner quelques différences qui renvoient aux différences de principe des méthodes statistiques choisies. La méthode Cyrano que ce soit pour la moyenne annuelle ou pour la concentration maximale admissible donne des estimations qui ne correspondent pas strictement aux moyennes et concentrations maximales simplement définies de façon arithmétique. Ainsi, la méthode Cyrano a tendance à favoriser une estimation de « bon état » en diminuant la concentration maximale prise en compte.

Enfin, la méthode Cyrano présente la caractéristique par rapport à la méthode « analytique » de prendre en compte les corrélations entre substances dans l'estimation des indices de confiance. La méthode « analytique » considère, comme le prévoit la DCE, l'état chimique par substance sur une année avant de déclarer l'état global de la station (que ce soit pour la moyenne ou pour l'indice de confiance). A l'inverse, la méthode Cyrano donne plus de poids à l'état de la masse d'eau à chaque période d'échantillonnage, notamment pour l'estimation de l'indice de confiance. Dans cette dernière méthode, plusieurs substances déclassant simultanément une station ont moins de poids que ces mêmes substances déclassant la station à deux dates différentes.

Les méthodes statistiques testées dans le présent rapport prennent en compte à la fois :

- des incertitudes d'origine analytique des mesures individuelles,
- la variabilité du milieu à travers l'utilisation de n mesures annuelles et de la dispersion de ces mesures.

Une hypothèse forte de la démarche est donc de considérer que les données de surveillance utilisées pour le calcul de l'indicateur réglementaire d'état permettent de rendre compte de façon fiable de la variabilité du milieu. Il n'est pas dans l'objectif de ce rapport d'étudier cette hypothèse et ses conséquences en termes de niveau de confiance sur l'état. Il convient donc de garder à l'esprit que les résultats présentés constituent une estimation basse de l'incertitude réelle affectant l'évaluation de l'état chimique, sans qu'il soit possible d'estimer l'écart ainsi impliqué.

Sommaire

1. Introduction	7
1.1. CONTEXTE	7
1.2. RAPPEL DES TRAVAUX AQUAREF 2013 [4].....	8
1.3. OBJECTIFS	10
2. Descriptif des méthodes statistiques employées	11
2.1. METHODE « ANALYTIQUE ».....	11
2.2. METHODE DE CYRANO (« BOOTSTRAP ») [9].....	12
3. Présentation des données étudiées.....	13
3.1. PRESENTATION GENERALE	13
3.2. DIFFICULTES RENCONTREES AVEC LES DONNEES DISPONIBLES.....	14
3.2.1. Absence des incertitudes de mesure	14
3.2.2. Absence et/ou multiplicité de la limite de quantification	15
3.2.3. Nombre variable de données mensuelles	16
3.3. CRITERES DE SELECTION DES DONNEES	16
3.3.1. Critère lié à la valeur de la LQ par rapport aux NQE	16
3.3.2. Critère lié au nombre de données par substance.....	17
4. Résultats et interprétation	19
4.1. BILAN DES RESULTATS OBTENUS POUR LE JEU DE DONNEES 1	19
4.2. BILAN DES RESULTATS OBTENUS POUR LE JEU DE DONNEES 2	22
4.3. BILAN DES RESULTATS OBTENUS POUR LE JEU DE DONNEES 3.....	25
5. Conclusion.....	29
6. Bibliographie	31

Liste des illustrations

<i>Illustration 1 : représentation schématique de la distribution de la variable Y (en bleu), associée à son incertitude $U(Y)$ et de la valeur seuil réglementaire V_s (zone A : bon état et zone B : mauvais état)</i>	<i>33</i>
--	-----------

Liste des tableaux

<i>Tableau 1 – Résultats obtenus pour les 15 stations du jeu de donnée 1 (NC : niveau de confiance)</i>	21
<i>Tableau 2 – Résultats obtenus pour les 13 stations du jeu de donnée 2 (NC : niveau de confiance)</i> <i>(1) sur cette station le traitement Cyrano ne prend pas en compte le Tributylétain (cf</i> <i>texte)</i>	24
<i>Tableau 3 – Résultats obtenus pour les 17 stations du jeu de donnée 3 (méthode analytique</i> <i>seulement) (NC : niveau de confiance)</i>	25
Tableau 4 – Données des deux HAP utilisées pour l'utilisation de la méthode de Cyrano sur la station 6416910.....	26
Tableau 5 – Comparaison des résultats obtenus par les deux méthodes en fonction de la position relative des maxima annuels	26

Liste des annexes

Annexe 1	32
-----------------------	----

1. Introduction

1.1. CONTEXTE

Dans un objectif d'amélioration de la qualité des masses d'eau européennes, la directive cadre européenne sur l'eau [1] a défini des concepts et méthodologies à mettre en place par les Etats membres. La notion d'« état d'une masse d'eau » est un des concepts fondamentaux de la directive. Il permet notamment de fixer des objectifs en termes de « bon état » des masses d'eau dans chaque Etat membre. Ainsi, le premier objectif fixé par la directive en 2000 était un objectif de « bon état » de l'ensemble des masses d'eau européennes à l'horizon 2015.

La notion d'« état chimique d'une masse d'eau » est variable suivant le type de masse d'eau auquel on s'intéresse. Les substances qui participent à la définition de l'état d'une masse d'eau souterraine ou d'une masse d'eau de surface, ainsi que les seuils à prendre en compte pour définir le bon ou le mauvais état sont variables. Les méthodologies à appliquer sont décrites dans différents textes réglementaires, pour les eaux souterraines [2] et pour les eaux de surface [3].

Pour chaque type d'eau, ces documents décrivent les différentes méthodes d'agrégation des données permettant d'aboutir à la définition d'un indicateur de l'état chimique d'une masse d'eau. Les données prises en compte sont issues de programmes de surveillance des masses d'eau. Elles correspondent à des campagnes d'échantillonnage et d'analyse permettant de connaître les concentrations de polluants dans le milieu considéré. Ces données sont ensuite exploitées pour définir un indicateur caractérisant l'état chimique des masses d'eau suivant des règles bien définies et harmonisées au niveau national et européen. De façon simplifiée, les masses d'eau sont déclarées en bon état si un certain nombre de valeurs réglementaires (normes de qualité environnementale NQE) ne sont pas dépassées. Dans le cas des eaux de surface, la masse d'eau est en bon état chimique si les NQE ne sont dépassées pour aucune des substances prioritaires surveillées.

Suite à ces évaluations d'état des masses d'eau, les gestionnaires de bassin doivent s'attacher à mettre en place des actions (programmes de mesure) permettant de comprendre les situations de « mauvais état » puis, autant que possible, permettre de retourner à une situation de « bon état ». Ceci implique souvent des enjeux financiers très importants. Il semble donc nécessaire de disposer du maximum d'éléments objectifs permettant de prendre des décisions adaptées à partir des données acquises dans les programmes de surveillance des masses d'eau.

Afin de prendre des décisions adaptées, par exemple suite à une déclaration de mauvais état d'une masse d'eau, il est nécessaire de s'interroger sur la fiabilité des données, les incertitudes et les niveaux de confiance des informations sur lesquelles seront basées ces décisions.

Comme dans chaque processus d'acquisition de données, des erreurs sont introduites à chaque niveau de l'acquisition. Les processus d'échantillonnage et d'analyse n'échappent pas à cette règle et à chaque donnée des programmes de surveillance est associée une incertitude plus ou moins bien maîtrisée et/ou connue. Ces données étant utilisées pour évaluer l'état chimique de la masse d'eau, il est logique de

considérer qu'à l'évaluation de cet état chimique est aussi associée une incertitude. On peut ainsi chercher à définir, pour l'évaluation de cet état, un « niveau de confiance », c'est à dire quantifier :

- la probabilité que l'état chimique de la masse d'eau soit bien celui qui a été estimé en comparant les concentrations aux NQE,
- et donc ainsi, la probabilité que l'état chimique de la masse d'eau soit différent de celui qui a été estimé.

La notion de niveau de confiance est introduite dans différents documents européens sans pour autant que la notion ou les méthodes soient précisément définies [4]. Au niveau national, le document décrivant de façon détaillée les règles d'évaluation de l'état d'une masse d'eau de surface continentale [5] introduit cette notion de niveau de confiance pour l'état chimique et aussi pour l'état écologique (le document sera très certainement revu pour prendre en compte le nouvel arrêté « évaluation » [6]). En ce qui concerne l'état chimique, le niveau de confiance défini dans le guide n'utilise pas les incertitudes sur les données de surveillance. Il prend en compte uniquement la quantité d'information existante (par exemple la proportion de substances de l'état chimique sur lesquelles les données disponibles permettent de se prononcer).

Concernant les incertitudes analytiques, ce guide rappelle que le suivi des polluants dans les eaux ne permet pas d'obtenir une valeur exacte de leur concentration mais un encadrement de cette valeur :

- d'une part, parce qu'à toute mesure est liée une incertitude analytique (U)
- d'autre part, parce que la résolution analytique des laboratoires est limitée : en-dessous d'un certain niveau, la concentration d'un polluant ne peut plus être quantifiée, il s'agit de la limite de quantification (LQ).

Incertitude analytique et limite de quantification varient en fonction des capacités des laboratoires mais aussi et surtout en fonction des polluants à analyser. Le guide insiste également sur le fait qu'aujourd'hui l'information relative à l'incertitude analytique n'est pas toujours disponible car pas systématiquement transmise par les laboratoires. En conséquence, elle n'est pas prise en compte dans l'évaluation de l'état chimique (elle est par défaut considérée comme inférieure ou égale à 50% - cf. article 2 de l'arrêté du 08/07/10 modifiant l'arrêté « évaluation » du 25 janvier 2010). Cependant, à court terme, le guide indique que cette dimension devra être intégrée dans l'évaluation de l'état des eaux. Il convient donc d'exiger des laboratoires qu'ils transmettent systématiquement l'incertitude associée aux résultats d'analyses. On peut également ajouter que cette information devrait être bancarisée par les gestionnaires en même temps que le résultat.

1.2. RAPPEL DES TRAVAUX AQUAREF 2013 [4]

En 2013, dans le cadre de son champ d'action relatif à la fiabilité des mesures, AQUAREF a entamé des réflexions sur la notion de niveau de confiance de l'état chimique des masses d'eau. L'objectif était de proposer aux gestionnaires un outil améliorant l'information disponible sur les données relatives à l'état chimique d'une masse d'eau de surface en introduisant des concepts statistiques simples. Ces nouvelles informations devraient permettre des prises de décision plus fiables et offrir de nouvelles possibilités de communication sur l'état des masses d'eau et l'évolution de cet état. De façon plus précise, il s'agissait de proposer une méthodologie de calcul/d'évaluation d'un niveau de confiance sur l'indicateur réglementaire de l'état

chimique d'une masse d'eau utilisant notamment l'incertitude sur chaque donnée et la dispersion entre les données.

L'état chimique réglementaire se fonde principalement, pour chaque substance, sur la moyenne annuelle des concentrations, et sa comparaison à une valeur seuil. On peut distinguer deux grands types de variabilités (et donc d'incertitudes) impliquées par cette approche :

- l'incertitude sur la moyenne annuelle résultant des incertitudes propres à chaque mesure individuelle. L'incertitude analytique est une contribution à ces incertitudes sur les mesures individuelles, mais il y en a d'autres (liées au prélèvement, notamment mais celles-ci sont pour l'instant peu documentées).
- l'incertitude qui procède du fait que la moyenne annuelle est estimée à partir d'un nombre limité de prélèvements, posant la question de la représentativité temporelle de la stratégie d'échantillonnage. Même en supposant que les incertitudes des mesures individuelles soient nulles, la valeur obtenue en moyennant n prélèvements annuels sera différente selon les dates de prélèvement choisies. Cette variabilité sera d'autant plus forte que les variations temporelles de composition chimique de la masse d'eau seront élevées.

Les méthodes statistiques testées dans le présent rapport prennent en compte à la fois :

- **des incertitudes d'origine analytique des mesures individuelles,**
- **la variabilité du milieu, mais avec l'hypothèse forte que la dispersion des n mesures annuelles utilisées pour chaque situation est représentative de cette variabilité.**

Il n'est pas dans l'objectif de ce rapport de vérifier la véracité d'une cette dernière hypothèse, et il convient donc de garder à l'esprit que les résultats présentés constituent une estimation basse de l'incertitude réelle affectant l'évaluation de l'état chimique, sans qu'il soit possible d'estimer l'écart ainsi impliqué.

Les réflexions conduites dans le cadre de l'étude AQUAREF 2013 [4] ont porté sur les eaux de surface. Ces réflexions ont pris en compte les règles actuelles les plus simples de l'évaluation de l'état chimique d'une masse d'eau de surface sans les remettre en cause, notamment sans remise en cause des fréquences d'échantillonnage, de la représentativité des stations, des périodes d'échantillonnage... Ces différents paramètres restent bien évidemment des paramètres très importants pour une détermination fiable de l'état d'une masse d'eau. Leur influence est complexe mais elle semble indispensable à étudier en complément des concepts proposés dans ce rapport.

Notion de niveau de confiance

La comparaison d'une donnée à une valeur seuil permet de conclure sur l'état de la masse d'eau. Ainsi, si la donnée est inférieure à la valeur seuil, on conclut en général à un bon état. Cependant, compte tenu des incertitudes sur la donnée, il existe une probabilité plus ou moins grande que la conclusion faite ne soit pas la bonne. Pour reprendre l'exemple ci-dessus du bon état, il existe une probabilité plus ou moins grande que la masse d'eau soit en mauvais état alors que la donnée est inférieure à la valeur seuil. **Le « niveau de confiance » considéré dans ce rapport correspond à la probabilité en % que l'état de la masse d'eau soit celui qui a été estimé. Ainsi, pour une masse d'eau déclarée en bon état, l'indice de confiance correspond à la**

probabilité que la masse d'eau soit en bon état sur la base des hypothèses décrites plus haut. Pour une masse d'eau déclarée en mauvais état, l'indice de confiance correspond à la probabilité que la masse d'eau soit en mauvais état. Le niveau de confiance est donc sauf exception (cf 4.3) compris entre 50 et 100%.

Exemples

- pour une masse d'eau en bon état et pour laquelle la probabilité du bon état est de 80%. L'indice de confiance est de 80% et il y a une probabilité de 20% que la masse d'eau soit en mauvaise état.
- Pour une masse d'eau en mauvais état et pour laquelle la probabilité du bon état est de 20%, l'indice de confiance est de 80%.

L'étude de 2013 a abordé les principes de la notion de niveau de confiance et a testé six méthodes statistiques. Chaque méthode présentait ses propres hypothèses, ses propres avantages et ses propres inconvénients. Les tests de ces méthodes statistiques ont été conduits sur 4 substances uniquement et sur un jeu de données provenant d'une seule station.

Les principes de base de la méthode sont rappelés en annexe 1.

1.3. OBJECTIFS

Le présent travail s'inscrit dans le programme d'action AQUAREF pour l'année 2015. Il est réalisé en partenariat par le BRGM et le LNE avec la collaboration de l'INERIS dans le cadre de conventions de partenariat avec l'ONEMA. L'objectif est de poursuivre les travaux entamés en illustrant de façon plus concrète les principes et méthodes de l'action 2013. Pour cela, il est proposé de tester la méthodologie sur un nombre de données plus large en termes de substances et de stations. Ces données sont issues de programmes de surveillance DCE.

2. Descriptif des méthodes statistiques employées

Parmi les 6 méthodes statistiques décrites en 2013, 3 donnaient des résultats proches et relativement concordants sans pour autant que les autres méthodes puissent être définitivement écartées. Dans le cadre de ce rapport, 2 de ces 3 méthodes ont pu être testées. Ces méthodes sont décrites dans la référence [4]. Ci-dessous ne sont reprises que les principales caractéristiques de ces méthodes.

2.1. METHODE « ANALYTIQUE »

Moyenne annuelle

Cette méthode est celle de la moyenne arithmétique avec prise en compte de la dispersion « inter-mois » des résultats d'analyse. Le calcul de l'incertitude sur la moyenne prend en compte les incertitudes analytiques ainsi que la dispersion des données.

On note respectivement x_i et u_i le résultat de la mesure i et l'incertitude-type associée pour $i = 1, \dots, n$ et $n = 12$.

A chaque résultat de mesure on associe la variable aléatoire X_i dont la valeur centrale est le résultat de mesure x_i et dont l'étendue des valeurs possibles est déterminée par l'incertitude-type u_i . Dans cette méthode, on prend en compte la dispersion des données en plus de l'incertitude analytique sur chaque donnée. De façon simplifiée, cette dispersion est prise en compte dans le calcul en ajoutant un terme issu d'une estimation d'un intervalle de confiance sur une moyenne.

$$u(\bar{x}) = \sqrt{\frac{\sum u_i^2}{n^2} + \frac{s^2}{n}}$$

Avec s écart-type des n résultats pris en compte et \bar{x} moyenne des x_i .

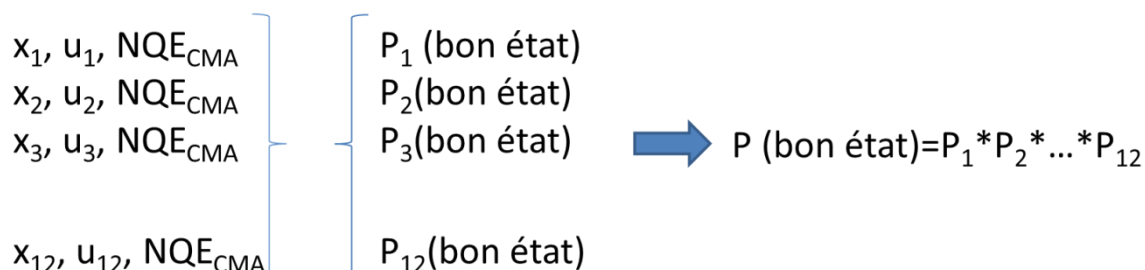
Avec ces hypothèses, les calculs sont ensuite conduits comme décrit en annexe 1. Ils fournissent une probabilité de « bon état » pour chaque substance. La probabilité de « bon état » de la station est estimée comme le produit des probabilités de « bon état » pour chaque substance. Pour cela une hypothèse d'indépendance/absence de corrélation des résultats obtenus sur les différentes substances est faite. Nous verrons au § 4.3 que cette hypothèse peut présenter quelques inconvénients.

Avertissement : dans l'incertitude fournie par les laboratoires il existe des contributions aléatoires (fidélité) et des contributions fixes appelées biais. La part aléatoire de l'incertitude analytique contribue donc également à la dispersion des données. Ainsi, dans le calcul de l'incertitude sur la moyenne tel que conduit ci-dessus, certains termes sont pris en compte deux fois. Toutefois, étant donné qu'il n'est pas possible de disposer de façon simple de l'information sur la part relative du biais et de la fidélité au

sein de l'incertitude déclarée par le laboratoire, la totalité de l'incertitude a été prise en compte. Cette hypothèse peut avoir comme conséquence une légère surestimation de l'incertitude sur la moyenne annuelle (sous estimation de l'indice de confiance sur l'état de la masse d'eau : sous estimation de l'indice de confiance « bon état » pour une masse d'eau en bon état et surestimation de l'indice de confiance « bon état » pour une masse d'eau en mauvais état).

Concentration maximale admissible (CMA)

Pour chaque substance et pour chacun des 12 résultats x_i associés à leur incertitude u_i cette méthode prend en compte la probabilité de dépassement de la valeur maximale NQE_{CMA} (cf annexe 1) Elle considère ensuite que, pour que la masse d'eau soit en bon état pour la substance considérée, il faut que le bon état soit assuré pour les 12 résultats. En supposant l'indépendance des 12 mesures, on calcule le produit de chacune des 12 probabilités de « bon état » (P_1 à P_{12}) pour obtenir la probabilité P de bon état de la masse d'eau pour la substance considérée.



2.2. METHODE DE CYRANO (« BOOTSTRAP ») [9]

Les techniques de bootstrap estiment une statistique (par exemple moyenne ou dispersion) à partir de la seule information des données, et avec un minimum d'hypothèses. Elles sont basées sur des simulations engendrant de « nouveaux échantillons » obtenus par « tirage avec remise » à partir de l'échantillon initial (on parle de rééchantillonnage). Cette méthode permet dans notre cas de traiter de façon relativement simple les aspects d'estimation de la moyenne, de l'incertitude sur la moyenne, de proportion de dépassement d'une valeur seuil et aussi de traiter les aspects liés au dépassement du maximum.

On considère une série de 12 mesures mensuelles sur une liste de N substances, avec les incertitudes de mesure associées (u_i). Les hypothèses faites sont principalement :

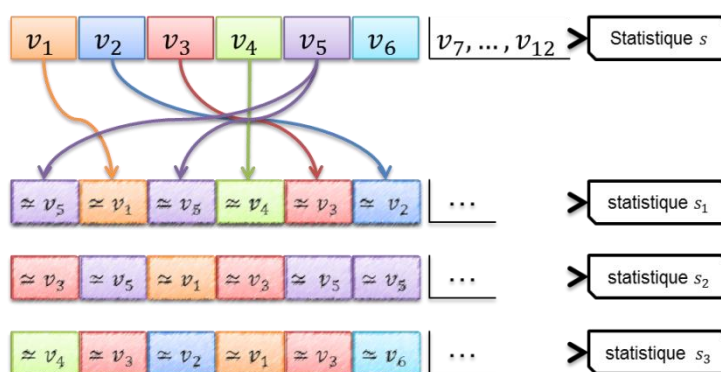
- Mesures mensuelles régulières, indépendantes et identiquement distribuées
- Incertitudes de mesure de type gaussien

Afin de prendre en compte les éventuelles corrélations entre éléments, le choix a été fait de considérer non pas les valeurs individuelles, substance par substance, mais de considérer des vecteurs de N valeurs pour chaque mois. Les mesures sont donc les 12 vecteurs de N valeurs v_1 à v_{12} .

La méthode consiste à réaliser 10 000 tirages aléatoires (avec remise) de 12 vecteurs. « Avec Remise » signifie que pour un tirage, il est possible de trouver plusieurs fois le même vecteur. À chaque tirage, l'incertitude analytique est prise en compte par application d'un bruit gaussien. Pour chaque tirage k des 10 000, la statistique recherchée $s_{k, \text{substance}}$ (la moyenne ou le maximum) sur les 12 valeurs est calculée pour chaque substance. On calcule enfin les paramètres :

- moyenne pour la statistique = $\text{moy}(s_1, \dots, s_{10\,000})$
- écart-type pour la statistique = $\sigma(s_1, \dots, s_{10\,000})$

La méthode est présentée schématiquement ci-dessous.



Il faut noter que cette méthode permet d'aborder à la fois la problématique des moyennes annuelles et celle des maximums annuels.

Le code de calcul a été développé sous *R* (logiciel disponible gratuitement sur Internet à l'adresse <http://www.r-project.org>) lors du précédent travail.

3. Présentation des données étudiées

3.1. PRESENTATION GENERALE

Les données étudiées sont issues de programmes de surveillance nationaux de masses d'eau de surface (cours d'eau) : elles ont été transmises à l'INERIS par les Agences de l'Eau dans le cadre des travaux de priorisation des substances et pour la validation des indicateurs SEEE concernant les eaux de surface.

Plus précisément, il s'agit de 3 jeux de données qui seront identifiés de la façon suivante dans le reste du rapport :

- Jeu 1 : résultats obtenus en 2013 sur 15 stations de prélèvement dépendant de l'Agence de l'eau Adour-Garonne (AEAG)
- Jeu 2 : résultats obtenus en 2012 sur 13 stations de prélèvement dépendant de l'Agence de l'eau Rhin-Meuse (AERM),
- Jeu 3 : résultats obtenus en 2013 sur 17 stations de prélèvement dépendant de l'Agence de l'eau Rhône-Méditerranée-Corse (AERMC).

Ces jeux de données correspondent à des stations sélectionnées sur la base :

- du nombre de substances présentant des résultats quantifiés.
- du nombre total de résultats quantifiés sur la station

Ce choix a été fait afin d'être illustratif pour ce travail et ne pas prendre en compte des stations sans aucune ou avec un nombre faible de substance quantifiée.

On dispose donc de 45 jeux de données. Ces données sont relatives aux concentrations de substances prioritaires de la DCE. Ces concentrations sont utilisées pour déterminer l'état chimique réglementaire de la masse d'eau. Le nombre de substances dans chaque jeu de données est variable, bien qu'il soit généralement compris entre 50 et 60. Une station seulement ne reprend que 15 substances, et deux n'en reprennent qu'une quarantaine.

3.2. DIFFICULTES RENCONTREES AVEC LES DONNEES DISPONIBLES

Dans le cadre du précédent travail, un traitement statistique avait été appliqué sur un jeu de données dit « complet ». Ce jeu de données était constitué de 12 résultats mensuels, ainsi que les incertitudes associées, obtenus pour 41 substances, pour lesquelles on disposait des limites de quantification (LQ) et les normes de qualité environnementales en moyenne annuelle (NQE_{MA}) et en concentration maximale annuelle (NQE_{CMA}). Un tel jeu de données ne posait aucun problème pour l'application de l'une ou l'autre des méthodes statistiques.

Les jeux de données qui ont été utilisés dans le cadre du présent travail, en revanche, ne se présentent pas systématiquement sous cette forme. Il n'est ainsi pas rare de devoir composer avec des jeux de données qui s'avèrent plus ou moins problématiques au niveau de leur exploitation. Les paragraphes suivants décrivent les principales difficultés rencontrées et les solutions qui ont été utilisées pour les surmonter. Ces difficultés sont liées principalement aux insuffisances des données disponibles dans les bases et notamment au manque d'informations systématiques concernant les incertitudes et les limites de quantification.

3.2.1. Absence des incertitudes de mesure

La méthodologie mise en place prend en compte à la fois la dispersion entre les résultats (au sein d'une chronique annuelle) mais aussi les incertitudes sur chaque résultat. Ainsi, l'incertitude de mesure sur chaque résultat est une donnée indispensable pour appliquer au mieux la méthode. Malheureusement, les incertitudes de mesure n'ont jamais été disponibles dans les jeux de données que nous avons utilisées. Comme dans le précédent rapport, il a donc été nécessaire de modéliser une

incertitude en utilisant un modèle de type « puissance » présentant au niveau de la limite de quantification une incertitude élargie U de l'ordre de 75% (k = 2) et une incertitude U de l'ordre de 30% (k = 2) à partir de 10 fois la limite de quantification.

Ce modèle est le suivant :

$$U(\%)=0.735* (C/LQ)^{0.5975}$$

Cette méthode d'attribution d'une incertitude est très simplificatrice. Elle n'est utilisée que pour pallier l'absence de cette information dans les jeux de données disponibles. Par ailleurs, il est quand même nécessaire de disposer de l'information relative à la limite de quantification. En effet la LQ est indispensable pour déterminer l'incertitude et cette information n'est pas toujours facilement disponible.

La première condition et recommandation pour appliquer correctement la méthode « analytique » serait donc de disposer systématiquement dans les bases de données, de l'incertitude de mesure associée au résultat.

3.2.2. Absence et/ou multiplicité de la limite de quantification

Dans ce travail, la limite de quantification est utilisée pour :

- Remplacer les valeurs non quantifiées, en effet, les règles de la DCE imposent de remplacer un résultat non quantifié par la limite de quantification divisée par 2
- Estimer des incertitudes par modélisation (voir paragraphe précédent)
- Comparer les valeurs quantifiées à la NQE (un des critères retenus étant $LQ < NQE$ - cf §3.3.1)

Dans les jeux de données 1 et 2, la limite de quantification n'est pas explicitement fournie pour chaque résultat. Seule l'information sur le « code remarque de l'analyse » permet d'avoir une indication sur cette limite de quantification : en effet, lorsque l'analyte n'est pas quantifié (code remarque égal à 10), la valeur donnée en guise de résultat correspond à la limite de quantification. Ainsi, il est possible d'affecter à partir de ces données une limite de quantification à une substance donnée sauf dans le cas où tous les résultats sont quantifiés (cette situation ne s'est pas présentée dans les jeux de données exploités).

Cependant, une difficulté supplémentaire survient lorsque, pour une même substance au sein d'une même station, deux (voire trois) valeurs différentes correspondent à un code remarque égal à 10. Dans ce cas, il n'est pas facile d'attribuer aux résultats quantifiés une LQ afin de déduire une incertitude sur le résultat. Dans un souci de facilité, nous nous sommes ramenés à une LQ unique (la plus récente, c'est-à-dire celle qui correspond au résultat obtenu sur le dernier prélèvement de l'année).

Il s'agit là encore, comme pour les incertitudes, d'approximations qui ont été nécessaires du fait de l'absence de ces informations dans les bases. La recommandation pour une bonne application de la méthode est donc également de pouvoir disposer de façon systématique de l'information relative à la LQ pour chaque résultat, même lorsque le résultat est quantifié. A l'heure actuelle, ceci ne semble pas assuré de façon automatique dans les bases nationales.

Il est important de rappeler que des critères concernant la limite de quantification de même que l'incertitude sont mentionnés par la directive 2009/90/CE dite directive QA/QC (limite de quantification inférieure à 30% de la NQE et incertitude élargie inférieure à 50% $k=2$ au niveau de la NQE) et qu'il sera certainement nécessaire de justifier de leur respect auprès de la commission européenne. Ces critères sont repris dans l'arrêté agrément des laboratoires et pourront donc être considérés comme respectés dès lors que des laboratoires agréés auront été sélectionnés et/ou que les résultats auront été rendus sous agrément. Toutefois, il n'est pas ou difficilement possible de vérifier que ces critères sont vérifiés à partir de la simple exploitation des bases de données.

3.2.3. Nombre variable de données mensuelles

Contrairement au jeu de données utilisé lors du précédent travail, les jeux de données utilisées ne présentent pas systématiquement 12 résultats d'analyse pour chaque substance.

En fonction de la méthode de traitement utilisée, certaines substances devront être ignorées afin de disposer d'un jeu de données équilibré, c'est-à-dire présentant pour toutes les substances le même nombre de résultat. La méthode Cyrano nécessite par exemple un nombre de données identique pour toutes les substances car elle prend en compte les corrélations entre substances. Pour ce faire, elle ne travaille pas individuellement avec les résultats sur chaque substance mais en considérant le « vecteur » constitué par l'ensemble des résultats obtenus pour toutes les substances à une date donnée.

3.3. CRITERES DE SELECTION DES DONNEES

Afin de garantir l'efficacité de la méthode ainsi que l'obtention de résultats exploitables, il a été nécessaire de procéder à une sélection des données au moyen de critères décrits ci-dessous. Cette sélection entraîne évidemment une perte d'information, que nous avons tenté de minimiser pour conserver une certaine pertinence du traitement des données. Certains critères sont par ailleurs déjà utilisés ou proposés dans certains documents nationaux sur l'évaluation de l'état [5].

3.3.1. Critère lié à la valeur de la LQ par rapport aux NQE

Pour certaines substances, il peut arriver que la limite de quantification soit supérieure à la NQE. Dans ce cas, même si tous les résultats d'analyse sont répertoriés comme non quantifiés, les méthodes statistiques appliquées peuvent conclure à un dépassement systématique de la NQE.

Ainsi, afin d'éviter un déclassement injustifié d'un trop grand nombre de stations, il a été décidé que les substances pour lesquelles la limite de quantification est supérieure ou égale à la NQE ne seraient pas prises en compte dans le traitement (qu'ils soient ou non quantifiés).

Il s'agit là d'un choix effectué dans le cadre de ce rapport pour une interprétation plus facile des résultats. D'autres choix sont bien évidemment possibles. Ils pourraient être mis en application de façon automatique dans une gestion automatisée des données. Cependant, les difficultés liées à la déclaration de bon ou mauvais état d'une substance pour laquelle la limite de quantification est supérieure à la valeur réglementaire demeureront dans la plupart des cas.

3.3.2. Critère lié au nombre de données par substance

Comme indiqué dans le paragraphe 2.2.3., il peut arriver que dans un même jeu de données, le nombre de données disponible ne soit pas identique pour toutes les substances. Il peut s'agir par exemple de jeux de données où la majorité des substances présentent 12 résultats et quelques substances n'en présentent que 6. En fonction de la méthode statistique (cf §2), ces substances sont traitées différemment :

- La méthode analytique telle qu'elle a été employée permet de prendre en compte chaque substance indifféremment du nombre de données qu'elle présente.
- La méthode de Cyrano, à l'inverse, requiert que chaque substance présente le même nombre de données pour pouvoir effectuer les tirages des « vecteurs » et ainsi établir correctement les corrélations entre les substances. Nous avons donc décidé de ne considérer que les substances dont le nombre de données revient le plus souvent au sein d'un même jeu de données (en général 12 ou 13 données) Une autre solution envisageable aurait été de ne considérer, pour chaque substance, qu'un nombre de données égal à la valeur du nombre de données de la substance présentant le plus faible nombre de données, mais cela revenait généralement à supprimer un grand nombre de données.

4. Résultats et interprétation

Il n'existe pas a priori de caractéristiques spécifiques aux 3 jeux de données étudiés. Une exploitation globale sur l'ensemble des stations pourrait donc être envisagée. Cependant, les résultats sont présentés pour chaque jeu de données car il existe des spécificités essentiellement au niveau des données et méta données disponibles. Pour chaque jeu de donnée les tableaux de synthèse présentent :

- L'état global de la station en moyenne annuelle et en concentration maximale admissible.
- L'indice de confiance estimé sur l'état de la station (cf §2.1 et .2.2)
- Le paramètre déclassant le cas échéant.

4.1. BILAN DES RESULTATS OBTENUS POUR LE JEU DE DONNEES 1

Sur le jeu de données 1, la concordance entre les méthodes utilisées pour estimer les états est très bonne. Ce constat est logique pour les moyennes annuelles mais en ce qui concerne les concentrations maximales admissibles, la méthode d'estimation est un peu différente et des différences sont possibles. Sur ce jeu de données, cela n'entraîne pas de différence sur l'état global.

En ce qui concerne les indices de confiance estimés suivant les deux méthodes, ils sont également très concordants avec des écarts en général n'excédant pas une dizaine de % (au maximum 13% pour la station 5156950) ce qui apparaît tout à fait satisfaisant pour une exploitation qualitative d'une telle information. Les écarts sont cependant plus importants pour certains indices de confiance en concentration maximale admissible. Ceci rejoint la remarque précédente concernant la différence de méthode d'estimation de l'état entre la méthode « analytique » et la méthode « Cyrano ». La méthode de Cyrano a tendance à minimiser l'influence des fortes concentrations et donc à surestimer le niveau de confiance sur le bon état de la station par rapport à la méthode analytique. La méthode analytique est basée de façon stricte sur l'utilisation de la valeur maximale alors que la méthode Cyrano revient à une utilisation d'un percentile sur les données. Il faut préciser que cette notion de percentile qui avait été longuement discutée au moment de la mise en place de la directive NQE (directive 2008/105/CE) n'est pas la méthodologie appliquée en France à l'heure actuelle.

Ainsi, en concentration maximale admissible, les niveaux de confiance sur le bon état de la station calculés par la méthode de Cyrano sont systématiquement supérieurs à ceux calculés par la méthode analytique, et parfois dans des proportions importantes (jusqu'à près de 40% pour la station 5131000, par exemple). Cela ne remet pas en cause l'évaluation de ce niveau de confiance mais traduit les modes différents d'estimation de l'état concernant la CMA.

Niveau de confiance relatif à l'indicateur réglementaire de l'état chimique d'une masse d'eau de surface

Station	Méthode analytique						Méthode Cyrano					
	En moyenne annuelle			En concentration maximale admissible			En moyenne annuelle			En concentration maximale admissible		
	Etat global	NC	Paramètres déclassants	Etat global	NC	Paramètres déclassants	Etat global	NC	Paramètres déclassants	Etat global	NC	Paramètres déclassants
5089000	Bon	100		Bon	100		Bon	100		Bon	100	
5218000	Bon	70		Bon	99		Bon	27		Bon	100	
5110000	Bon	54		Mauvais	82	Cadmium	Bon	60		Mauvais	65	Cadmium
5216210	Bon	100		Bon	100		Bon	99		Bon	100	
5131000	Bon	93		Bon	61		Bon	92		Bon	100	
5115000	Bon	100		Bon	100		Bon	100		Bon	100	
5158200	Mauvais	74	Mercure	Mauvais	100	Mercure	Mauvais	67	Mercure	Mauvais	67	Mercure
5163700	Bon	79		Bon	100		Bon	77		Bon	100	
5107000	Bon	100		Bon	100		Bon	100		Bon	100	
5161950	Mauvais	63	Benzo(b)fluoranthène	Bon	100		Mauvais	76	Benzo(b)fluoranthène	Bon	100	
5159000	Mauvais	79	4-tert-octylphénol	Bon	100		Mauvais	68	4-tert-octylphénol	Bon	100	
5029000	Bon	100		Bon	100		Bon	100		Bon	100	
5156950	Mauvais	72	Benzo(b)fluoranthène	Bon	100		Mauvais	59	Benzo(b)fluoranthène	Bon	100	
5158700	Mauvais	62	Benzo(b)fluoranthène	Mauvais	89	Mercure	Mauvais	66	Benzo(b)fluoranthène	Mauvais	66	Mercure
5209000	Bon	78		Bon	100		Bon	73		Bon	100	

Tableau 1 – Résultats obtenus pour les 15 stations du jeu de donnée 1 (NC : niveau de confiance)

On peut également étudier ces résultats du point de vue de la plus-value apportée par cet indice de confiance. Ainsi, pour la station 5110000, l'état global de la station est bon mais avec un indice de confiance de 54%. Cela permet de mettre en lumière que cette station est en bon état mais que :

- En raison des incertitudes analytiques ou des incertitudes liées au calcul de l'indicateur, la probabilité que cette station soit en mauvais état est de l'ordre de 46%.
- L'acquisition de prochaines données de surveillance sur un ou plusieurs paramètres pourrait provoquer un basculement d'état.

Il s'avère sur cet exemple que le niveau de confiance faible est lié au cadmium pour lequel une moyenne annuelle de 75 ng/L est calculée pour une NQE de 80 ng/L.

A l'inverse, pour la station 5158700, la station est en mauvais état mais il existe une probabilité non négligeable qu'en réalité elle soit en bon état (environ 30-40%).

4.2. BILAN DES RESULTATS OBTENUS POUR LE JEU DE DONNEES 2

Pour ce jeu de données, il faut tout d'abord noter que pour deux stations (99206000 et 2103800) les données n'ont pas pu être traitées par la méthode de Cyrano. En effet, il n'a pas été possible, pour ces stations, de satisfaire au critère sur le nombre de données identiques pour toutes les substances sans ignorer un trop grand nombre de données. La station 99206000 présentait des résultats pour 15 substances, 8 d'entre elles ayant fait l'objet de 12 mesures et 7 d'entre elles ayant fait l'objet de 13 mesures. Quant à la station 2103800, elle concernait 57 substances, 25 substances présentant 4 résultats et 32 substances présentant 6 résultats. Ces stations ont donc été ignorées. Par ailleurs, pour la station 2104000, le cas de la méthode Cyrano n'est pas présenté du fait d'un faible nombre de données pour le Tributylétain. Or cette substance est la substance qui provoque le déclassement de la masse d'eau. Un traitement éliminant cette substance aurait entraîné, de façon artificielle, une forte différence de conclusion sur l'état entre les deux méthodes statistiques. Il est possible de noter cependant que, après suppression du TBT de la liste des substances de cette station, les traitements « analytique » et « Cyrano » donnent des résultats très similaires à savoir un niveau de confiance sur le bon état de 100% en MA et de 85% et 64% respectivement en CMA; on retrouve la même tendance de la méthode « Cyrano » à favoriser le « bon état » pour la comparaison à la CMA.

Comme pour le jeu de données n°1, les résultats en termes d'état ou d'indice de confiance sont globalement concordants entre les deux méthodes statistiques utilisées. La comparaison entre les deux méthodes ne peut se faire que sur 10 stations. Pour ces 10 stations, l'écart en valeur absolue entre les résultats fournis par les deux méthodes sur les niveaux de confiance est en moyenne, en MA comme en CMA, de l'ordre de 10% ce qui reste très faible au regard d'objectifs qualitatifs d'utilisation de cet indice.

Pour huit stations, une cohérence parfaite au niveau des états globaux fournis par les deux méthodes est observée. Pour 2 stations (2100150 et 2106410) la différence de principe sur l'évaluation de la CMA entraîne un état différent entre les deux méthodes,

la méthode Cyrano étant plus favorable au bon état comme indiqué au §4.1. Par ailleurs, pour les mêmes raisons, les stations 2095600 et 2105000 présentent des niveaux de confiance de « bon état » en concentration maximale admissible significativement plus élevés avec la méthode Cyrano mais sans changement de l'état.

Niveau de confiance relatif à l'indicateur réglementaire de l'état chimique d'une masse d'eau de surface

Station	Méthode analytique						Méthode Cyrano					
	En moyenne annuelle			En concentration maximale admissible			En moyenne annuelle			En concentration maximale admissible		
	Etat global	NC	Paramètres déclassants	Etat global	NC	Paramètres déclassants	Etat global	NC	Paramètres déclassants	Etat global	NC	Paramètres déclassants
2070250	Bon	100		Bon	100		Bon	100		Bon	100	
2067150	Bon	83		Bon	100		Bon	85		Bon	100	
99206000	Mauvais	100	Benzo(b)fluoranthène Cadmium	Mauvais	58	Cadmium	Non traité					
2105000	Bon	100		Bon	84		Bon	100		Bon	100	
2106410	Bon	100		Mauvais	58	Mercure	Bon	100		Bon	60	
2109000	Bon	100		Bon	100		Bon	100		Bon	100	
2103800	Mauvais	76	Benzo(b)fluoranthène	Bon	100		Non traité					
2104000	Mauvais	68	Tributylétain cation	Mauvais	100	Mercure Tributylétain cation	Non traité					
2116000	Bon	100		Bon	100		Bon	100		Bon	100	
2124000	Bon	52		Bon	100		Bon	59		Bon	100	
2072700	Bon	100		Bon	100		Bon	100		Bon	100	
2100150	Bon	100		Mauvais	50	Mercure	Bon	100		Bon	59	
2095600	Bon	95		Mauvais	100	Mercure	Bon	95		Mauvais	64	Mercure

Tableau 2 – Résultats obtenus pour les 13 stations du jeu de donnée 2 (NC : niveau de confiance)
(1) sur cette station le traitement Cyrano ne prend pas en compte le Tributylétain (cf texte)

4.3. BILAN DES RESULTATS OBTENUS POUR LE JEU DE DONNEES 3

Dans le cas du jeu de données n°3, le choix a été fait de ne pas appliquer la méthode de Cyrano. En effet, sur ce jeu de données, il n'était pas possible de respecter le critère sur le nombre de données par substances (il aurait fallu ignorer un nombre significatif de substances et donc supprimer un trop grand nombre de données). Dans l'état actuel des méthodes utilisées, il s'agit là d'un inconvénient important de la méthode Cyrano qui nécessite un jeu de donnée équilibré concernant le nombre de résultat par substance. A l'inverse, il faut mentionner qu'un des avantages de la méthode Cyrano est de prendre en compte les corrélations entre substances, ce que ne fait pas la méthode « analytique ».

Sur ce jeu de données, le traitement a été effectué pour toutes les stations par la méthode analytique. Les résultats sont reproduits dans le Tableau 3 ci-dessous.

Station	Méthode analytique					
	En moyenne annuelle			En concentration maximale admissible		
	Etat global	NC	Paramètres déclassants	Etat global	NC	Paramètres déclassants
6047500	Bon	100		Bon	98	
6044900	Bon	61		Bon	98	
6000997	Bon	100		Bon	98	
6016940	Bon	100		Bon	99	
6012080	Bon	100		Bon	99	
6039960	Bon	100		Bon	98	
6042500	Bon	100		Bon	99	
6049000	Bon	72		Bon	99	
6131900	Bon	78		Bon	98	
6820013	Bon	86		Bon	98	
6471450	Mauvais	100	Benzo(a) pyrène Benzo(b) fluoranthène Benzo(k) fluoranthène	Mauvais	100	Benzo(a)pyrène
6080975	Mauvais	84	BDE 99	Bon	100	
6127900	Mauvais	98	Cadmium Chloroforme	Mauvais	100	Cadmium
6115090	Bon	74		Bon	100	
6001180	Mauvais	70	Benzo(b) fluoranthène	Bon	98	
6046000	Bon	100		Bon	99	
6416910	Bon	46		Bon	98	

Tableau 3 – Résultats obtenus pour les 17 stations du jeu de donnée 3 (méthode analytique seulement)
(NC : niveau de confiance)

Sur ce jeu de données, le cas de la station 6416910 est intéressant. Pour cette station, l'état global en moyenne annuelle est bon mais l'indice de confiance est inférieur à 50%. Ceci est lié au fait que 2 substances impactent cet indice de confiance. Il s'agit de deux HAP (désignés ci-après HAP1 et HAP2) pour lesquels les indices de confiance individuels sont respectivement de 52% (HAP1) et 88% (HAP2), donc supérieurs à 50% : la station est donc déclarée en bon état pour chacune de ces substances, et donc globalement en bon état. Cependant, l'indice de confiance global sur la station a été défini pour cette méthode « arithmétique » comme le produit des indices de confiance individuels sans tenir compte des éventuelles corrélations : dans le cas présent, on obtient un indice de confiance sur le bon état inférieur à 50%, alors même que la station est déclarée en bon état.

Afin de discuter de façon plus précise ce cas particulier, la méthode de Cyrano a été appliquée sur les données obtenues pour cette station. Les deux substances concernées présentant un nombre de données différent (6 pour le HAP1 et 7 pour le HAP2), la donnée supplémentaire du HAP2 a été retirée (dans le respect des dates de prélèvement, pour ne conserver que les données obtenues pour les mêmes dates) : il s'agissait d'une donnée inférieure à la LQ de la substance en question. Cette modification n'a que peu d'incidence sur les résultats de la méthode analytique : l'indice de confiance pour le HAP2 passe de 88% à 81% (la moyenne annuelle augmentant légèrement suite à la modification) et l'indice de confiance global passe de 46% à 42%.

La méthode de Cyrano a ainsi pu être appliquée sur un jeu de données modifié ne conservant que les substances présentant 6 données (les substances autres que les deux HAP présentent toutes un bon état avec un indice de confiance de 100%). Les valeurs retenues pour les deux HAP sont présentées dans le Tableau 4 ci-dessous (les données inférieures à la LQ sont fixées à LQ/2).

Substance	LQ	NQE	Donnée 1	Donnée 2	Donnée 3	Donnée 4	Donnée 5	Donnée 6
HAP 1	0,005	0,015	0,0051	0,0025	0,0025	0,0025	0,0703	0,0025
HAP 2	0,005	0,015	0,0025	0,0025	0,0025	0,0025	0,0422	0,0025

Tableau 4 – Données des deux HAP utilisées pour l'utilisation de la méthode de Cyrano sur la station 6416910 ($\mu\text{g/l}$)

Dans ce jeu de données, les maxima annuels sont obtenus à la même date (donnée 5). Afin d'évaluer l'impact de la corrélation sur les résultats obtenus (ce qui ne peut être fait avec la méthode analytique), il a été décidé de faire varier la position relative des maxima et d'étudier la différence entre les résultats lorsque les maxima sont observés à la même date et lorsqu'ils sont observés à des dates différentes. Le Tableau 5 récapitule les résultats obtenus.

Position des maxima	Méthode	Indice de confiance HAP1	Indice de confiance HAP2	Indice de confiance global
Maxima à la même date	Analytique	53%	81%	42%
	Cyrano	72%	74%	71%
Maxima décalés	Analytique	53%	81%	42%
	Cyrano	71%	74%	49%

Tableau 5 – Comparaison des résultats obtenus par les deux méthodes en fonction de la position relative des maxima annuels

Ces simulations mettent en évidence une différence significative entre les résultats obtenus selon la position relative des maxima : lorsque les maxima sont obtenus à la même date, la station est déclarée en bon état avec un indice de confiance proche de 70% suivant la méthode Cyrano alors que l'indice de confiance n'est que de 42% par la méthode « analytique ». En revanche, lorsque les maxima sont décalés dans le temps, la méthode de Cyrano identifie deux « sources » potentielles distinctes de dépassement : l'indice de confiance global est donc plus faible et chute aux alentours de 50%, alors que les indices de confiance individuels, eux, restent au même niveau que si les maxima étaient à la même date. L'exemple de cette station illustre ainsi l'importance de la prise en compte de la corrélation entre les substances lorsqu'il s'agit de fournir un niveau de confiance fiable sur l'état chimique d'une masse d'eau.

A noter cependant que ces derniers éléments renvoient à des principes différents qui devront être discutés si un tel indice de confiance devait être utilisé par les gestionnaires. La méthode « analytique » considère, comme le prévoit la DCE, l'état chimique par substance sur une année avant de déclarer l'état global de la station (que ce soit pour la moyenne ou pour l'indice de confiance). A l'inverse, la méthode Cyrano, donne plus de poids à l'état de la masse d'eau à chaque période d'échantillonnage, notamment pour l'estimation de l'indice de confiance. Plusieurs substances déclassant simultanément une station ont moins de poids que ces mêmes substances déclassant la station à deux dates différentes.

5. Conclusion

Ce rapport s'inscrit dans la continuité d'un précédent travail d'AQUAREF réalisé en 2013 par le BRGM et le LNE concernant une réflexion sur la détermination d'un niveau de confiance sur l'indicateur réglementaire de l'état chimique d'une masse d'eau.

En 2013, des principes méthodologiques et des méthodes statistiques avaient été proposés. Le présent rapport a pour objectif d'illustrer de manière plus complète le potentiel de ces méthodes et de souligner certaines de leurs limites. Il a ainsi permis de tester deux des méthodes statistiques issues du rapport 2013 (méthode « analytique » et méthode Cyrano). Ces tests ont été effectués sur un nombre de jeu de données plus important afin de varier les types de données et leurs caractéristiques.

Les données utilisées dans ce rapport sont des données de surveillance d'eau de surface provenant de 3 agences de l'eau. La première difficulté rencontrée a concerné les données et métadonnées disponibles. Quelle que soit la méthode statistique, les traitements se heurtent :

- à l'absence d'une information systématique sur l'incertitude de mesure au niveau du résultat,
- à l'absence d'une information systématique concernant la limite de quantification associée à chaque résultat,
- à la variabilité du nombre de données disponibles.

Par ailleurs la méthode Cyrano présente l'avantage de prendre en compte les corrélations entre substances dans l'estimation des indices de confiance. La contrepartie est que la méthode dans son application actuelle nécessite des jeux de données homogènes en ce qui concerne le nombre de données par substance. Ce nombre doit être identique pour toutes les substances de la station considérée. Sur certains jeux de données, ce critère n'était pas respecté et la méthode n'a pas pu être appliquée. Cependant, un exemple détaillé dans le rapport montre bien l'importance que peut prendre la corrélation inter-substances concernant le niveau de confiance accordé sur la fiabilité de l'état chimique, et pourrait inviter à se tourner vers des méthodes permettant sa prise en compte tout en limitant le désavantage lié à la présentation des données.

Les deux méthodes statistiques utilisées ont qualitativement donné des résultats très concordants à la fois en moyenne annuelle et en concentration maximale admissible, qu'il s'agisse de la définition de l'état de la masse d'eau mais aussi de l'estimation de l'indice de confiance sur cet état. De façon générale, les niveaux de confiance estimés par les deux méthodes sont très proches et permettraient d'atteindre l'objectif d'une utilisation « qualitative » de cet indice de confiance par les gestionnaires, en tenant compte des limites de ces méthodes et de la qualité et / ou quantité des données.

Il faut cependant mentionner quelques différences qui renvoient aux différences de principe des méthodes statistiques choisies. Ainsi, la méthode Cyrano, que ce soit pour la moyenne annuelle ou pour la concentration maximale admissible, donne des estimations qui ne correspondent pas strictement aux moyennes et concentrations maximales simplement définies de façon analytique comme actuellement pratiqué au niveau national. Par exemple, en ce qui concerne la concentration maximale, la méthode Cyrano a systématiquement tendance à favoriser une estimation de « bon état » en diminuant la concentration maximale prise en compte. Cela rejoint des discussions qui avaient eu lieu au niveau européen il y a quelques années concernant la

prise en compte non pas de la concentration maximale annuelle mais d'un percentile des données. A l'heure actuelle, c'est la concentration maximale qui est utilisée.

Enfin, il est important de rappeler que les méthodes statistiques testées dans ce rapport prennent en compte:

- des incertitudes d'origine analytique des mesures individuelles,
- la variabilité du milieu à travers l'utilisation de n mesures annuelles et de la dispersion de ces mesures.

Une hypothèse forte de la démarche est donc de considérer que les données de surveillance utilisées pour le calcul de l'indicateur réglementaire d'état permettent de rendre compte de façon fiable de la variabilité du milieu. Il n'était pas dans l'objectif de ce rapport d'étudier cette hypothèse et ses conséquences en termes de niveau de confiance sur l'état. Il convient donc de garder à l'esprit que les résultats présentés constituent une estimation basse de l'incertitude réelle affectant l'évaluation de l'état chimique, sans qu'il soit possible d'estimer l'écart ainsi impliqué.

6. Bibliographie

- [1] Directive Cadre Européenne sur l'Eau (DCE) 2000/60/CE.
- [2] Arrêté du 17 décembre 2008 établissant les critères d'évaluation et les modalités de détermination de l'état des eaux souterraines et des tendances significatives et durables de dégradation de l'état chimique des eaux souterraines
- [3] Arrêté du 25 janvier 2010 relatif aux méthodes et critères d'évaluation de l'état écologique, de l'état chimique et du potentiel écologique des eaux de surface pris en application des articles R. 212-10, R. 212-11 et R. 212-18 du code de l'environnement
- [4] GHESTEM JP., LEYNET A., DEMEYER S., FISHER N., LALERE B.(2013) – Niveau de confiance relatif à l'indicateur réglementaire de l'état chimique d'une masse d'eau. Rapport final. BRGM/RP-63228-FR, 45 p., 10 ill.
- [5] Guide technique Évaluation de l'état des eaux de surface continentales (cours d'eau, canaux, plans d'eau)-Décembre 2012 MEDDE.
- [6] Arrêté du 27 juillet 2015 modifiant l'arrêté du 25 janvier 2010 relatif aux méthodes et critères d'évaluation de l'état écologique, de l'état chimique et du potentiel écologique des eaux de surface pris en application des articles R. 212-10, R. 212-11 et R. 212-18 du code de l'environnement
- [7] Efron, B. (1979). "Bootstrap methods: Another look at the jackknife". *The Annals of Statistics* 7 (1): 1–26. doi:10.1214/aos/1176344552.

Annexe 1

Définition de l'état d'une masse d'eau et niveau de confiance associé

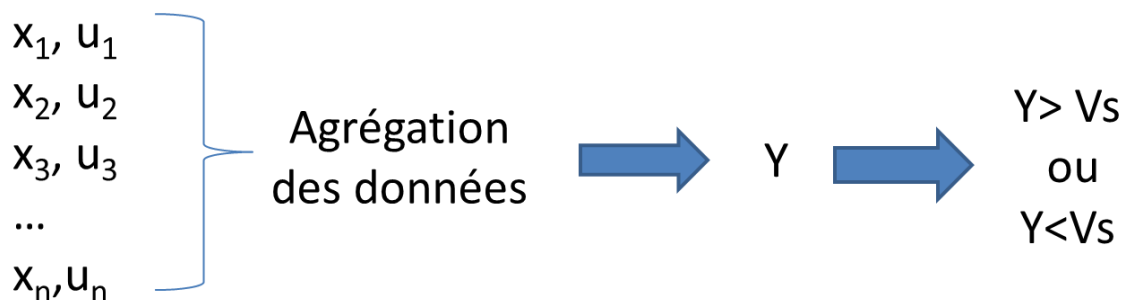
Cette annexe est extraite du rapport [4].

PRESENTATION SIMPLIFIEE

De façon schématique, le problème posé est le suivant : à chaque résultat d'analyse environnementale (x_i), est associée une incertitude de mesure (u_i). Dans le cadre de ce rapport il est considéré que l'incertitude est seulement due à l'incertitude analytique (cf remarque ci-dessous).

Des règles sont fixées au niveau national ou européen afin de prendre des décisions à partir de ces données environnementales. Par exemple dans le contexte DCE, il peut s'agir de déterminer un indicateur de l'état chimique de la masse d'eau.

Pour cela, les données individuelles (x_i, u_i) sont agrégées suivant des règles précises [7] (il peut s'agir par exemple d'effectuer la moyenne des données). Le résultat Y de cette agrégation de données est comparé à une valeur seuil réglementaire V_s . Le résultat de cette comparaison permet de définir un « bon » ou un « mauvais » état.



Remarque : le choix de ne considérer dans un premier temps que l'incertitude analytique est une simplification. On pourrait considérer que l'incertitude associée au résultat intègre d'autres sources d'incertitude (échantillonnage par exemple). Ceci ne modifierait rien au raisonnement. Cependant, et malgré les gros efforts mis ces dernières années sur la qualité des opérations d'échantillonnage, la détermination des incertitudes liées à ces données reste un exercice très difficile et les données disponibles sont, à ce jour, rares.

DEFINITION D'UN NIVEAU DE CONFIANCE

Compte tenu des incertitudes u_i qui affectent les données individuelles x_i , Y n'est pas une valeur déterminée mais une variable qui suit une loi de distribution (gaussienne ou autre) et dont on peut chercher à donner une représentation via le calcul de statistiques (ou de métriques).

Dans ce schéma conceptuel et afin de définir des niveaux de confiance, les deux principales questions qui se posent sont les suivantes :

1. Estimer Y (suivant les règles fixées par les textes réglementaires, par exemple moyenne, maximum, ...) et son incertitude $U(Y)$ en utilisant les résultats individuels x_i et leurs incertitudes u_i . La détermination de U constitue le point le plus délicat du point de vue statistique.
2. Connaissant Y et son incertitude $U(Y)$, déterminer les niveaux de confiance relatifs au dépassement ou non de la valeur seuil réglementaire

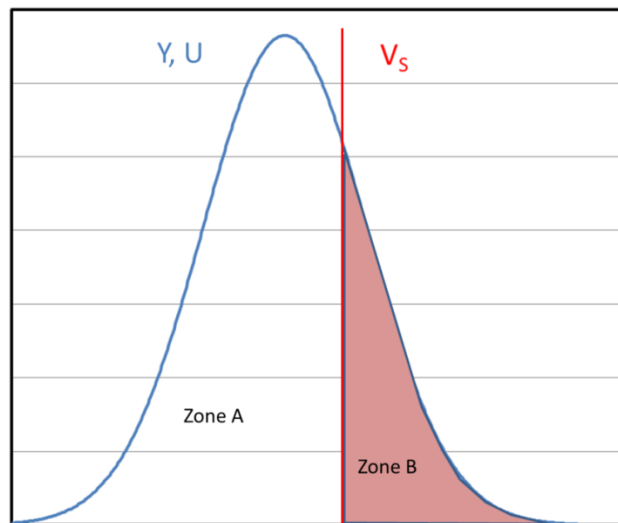


Illustration 1 : représentation schématique de la distribution de la variable Y (en bleu), associée à son incertitude $U(Y)$ et de la valeur seuil réglementaire V_s (zone A : bon état et zone B : mauvais état)

En utilisant l'Illustration 1, les deux objectifs principaux se résument à :

1. Estimer Y et $U(Y)$. L'estimation de Y est déjà classiquement faite par les Etats membres et, en France par les gestionnaires de bassin pour estimer l'état des masses d'eau. Cette estimation sera cependant effectuée aussi dans ce rapport avec les méthodes statistiques retenues.
2. Estimer la probabilité de dépassement de la valeur seuil : connaissant Y et U , calculer le rapport $B/(A + B)$ (où A et B sont les aires des zones A et B). Ce rapport donnera la probabilité que la variable Y dépasse la valeur seuil V_s . Dans le cas présenté ci-dessus la variable Y est en moyenne inférieure à V_s . En faisant l'hypothèse que ceci correspond à une situation de conformité, le rapport $A/(A + B)$ fournit un niveau de confiance (probabilité de « non dépassement ») associé à cette situation de conformité. A l'inverse, $B/(B + A)$ donne la probabilité que la situation soit non conforme alors que la conclusion réglementaire est la conformité.

Attention : le graphique ci-dessus présente un modèle gaussien (loi normale) pour la variable Y . Il s'agit d'une hypothèse simplificatrice. La loi de distribution de la variable Y peut suivre d'autres modèles.

APPLICATION A LA DETERMINATION DE L'ETAT CHIMIQUE D'UNE STATION D'EAU DE SURFACE

Les données de surveillance environnementale, les fréquences d'acquisition et les modes d'agrégation de données sont variables en fonction du type de masse d'eau auquel on s'intéresse. Dans ce rapport, nous prendrons comme exemple le cas de l'évaluation de l'état chimique d'une station d'eau de surface continentale pour appliquer les concepts statistiques retenus.

Il ne s'agit pas d'entrer dans tous les détails de l'évaluation de l'état. Nous avons choisi de nous baser sur quelques règles simples qui permettent d'illustrer les méthodes et principes proposés. Le cas de l'évaluation de l'état chimique d'une masse d'eau de surface sur une station définie peut se résumer de la façon suivante :

- Sur cette station, acquisition tous les mois sur une année des concentrations de n substances prioritaires (par exemple $n=41$ soit 41×12 résultats individuels)
- A chaque concentration doit être associée une incertitude de mesure u_i

Avec ces données, deux évaluations sont réalisées afin de définir l'état chimique de la masse d'eau. Une évaluation concerne la moyenne des concentrations et l'autre, le maximum. Ces évaluations sont résumées dans le tableau ci-dessous.

Moyenne annuelle (MA)	Concentration maximale (CMA)
<p>Pour chaque substance :</p> <ul style="list-style-type: none"> - remplacement des concentrations inférieures à LQ par LQ/2 - calcul de la moyenne annuelle Y des 12 données 	<p>Pour chaque substance :</p> <ul style="list-style-type: none"> - recherche de la concentration maximale Y sur l'année
<p>Pour chaque substance : comparaison de la moyenne annuelle Y avec la valeur réglementaire appelée NQE_{MA}</p> <ul style="list-style-type: none"> • Si $Y < NQE_{MA}$: bon état pour la substance • Si $Y > NQE_{MA}$: mauvais état pour la substance 	<p>Pour chaque substance : comparaison de la valeur maximale avec la valeur réglementaire appelée NQE_{CMA}</p> <ul style="list-style-type: none"> • Si $Y < NQE_{CMA}$: bon état pour la substance • Si $Y > NQE_{CMA}$: mauvais état pour la substance
<p>La station est déclarée en bon état si les états sont bons pour les n substances</p>	<p>La station est déclarée en bon état si les états sont bons pour les n substances</p>
<p>Elle est en mauvais état si l'état est mauvais pour au moins une substance</p>	<p>Elle est en mauvais état si l'état est mauvais pour au moins une substance</p>

Dans la suite du rapport seules ces deux situations seront étudiées.

Les règles détaillées d'évaluation de l'état d'une masse d'eau de surface sont décrites dans le « Guide technique Évaluation de l'état des eaux de surface continentales (cours d'eau, canaux, plans d'eau) » de décembre 2012. [5].



Centre scientifique et technique
Direction des laboratoires
3, avenue Claude-Guillemin
BP 36009 – 45060 Orléans Cedex 2 – France – Tél. : 02 38 64 34 34
www.brgm.fr