

# Éléments de synthèse collective des travaux du groupe de travail « Données sur l'eau »

■ Groupe de travail Astee

## Introduction

La mise en place du groupe de travail « Données » au sein de la CREMA répondait davantage à un besoin d'échanges que de livrer un rapport bien particulier. Les réunions se sont donc tenues autour de présentations synthétiques d'expériences diverses dans l'acquisition, le traitement, l'utilisation et la diffusion des données sur l'eau.

De ces échanges et des contributions à ce dossier, nous avons essayé de rassembler quelques grands axes de réflexion partagés et qui peuvent constituer une expression des besoins d'une partie des acteurs de la gestion de l'eau dans ce domaine. De même, le compte-rendu de la réunion locale, organisée grâce à l'appui de la section Rhône-Alpes en mai 2006, constitue une bonne synthèse des interrogations et des attentes qui s'expriment, au moment où le paysage et les besoins évoluent et se structurent.

## 1. Beaucoup de données mais peu d'information

De façon récurrente et très partagée, les membres du groupe de travail ont exprimé le décalage existant entre l'acquisition des données et leur utilisation. Trop de données sont sous-valorisées, tandis que dans de nombreux cas, des investigations nécessaires, car préalables à la prise de décision, se heurtent à des lacunes tant spatiales que temporelles, ou tout au moins à la difficulté de rassembler les données d'un même territoire et d'en tirer l'information nécessaire. Plusieurs aspects sont à considérer plus particulièrement.

### 1.1. Il faut éviter les redondances, la réplication d'une même donnée, mais valoriser la donnée

Cette dimension domine dans la refonte des réseaux de surveillance demandés par la directive cadre, à la

fois au niveau des contrôles de surveillance, autour des réseaux pilotés par les services de l'État et ses établissements publics, et dans le cadre des réseaux de contrôle opérationnel, à partir des réseaux d'usage développés par les différents acteurs. Mais comme le souligne le cahier des charges sur le réseau de suivi des eaux souterraines (septembre 2003), « la sélection de sites de surveillance supplémentaires pour le contrôle opérationnel nécessite un minimum de précautions. En effet, à l'issue de cet exercice, le réseau sera composé de points dont les objectifs sont très variables (suivi d'une installation classée, d'une pollution diffuse...). Il convient donc de rester prudent et d'évaluer la pertinence des informations apportées par chaque site. » Cet aspect est plus particulièrement illustré dans l'article de ce dossier sur la mise en place d'un observatoire de la qualité de l'eau de Seine dans l'agglomération parisienne.

### 1.2. Difficulté pour définir le bénéfice apporté par le traitement approprié des données : besoin d'études de cas chiffrées

Si la constitution d'historiques validés est aujourd'hui reconnue par de nombreux acteurs comme un bénéfice important de la collecte des données pour l'évaluation de la qualité des cours d'eau, il est souvent difficile de justifier *a priori* la mise en place de réseaux de mesure, *a fortiori* lorsque ceux-ci font appel à de la mesure en continu. Même si le coût de cette dernière ne représente pas en soi un frein à son utilisation, il y a lieu de bien montrer l'importance d'une approche de coût global dans l'élaboration d'une stratégie de surveillance et les complémentarités possibles par exemple entre mesures ponctuelles et mesures en continu.

Il y a surtout une mauvaise perception des bénéfices (ou des coûts évités !) apportés par une meilleure

connaissance car peu d'études chiffrées ont été faites sur ce sujet. L'analyse économique demandée par la mise en œuvre de la directive cadre incite de plus en plus à la mise en relation entre l'évolution de la qualité de l'eau (dégradation ou amélioration) et les pressions exercées sur le milieu ainsi que les efforts consentis par les usagers en termes de dépollution. On peut citer à cet égard la tentative faite par l'agence de l'eau Seine Normandie de lier les efforts effectués par l'ensemble des acteurs sur le bassin avec l'évolution des paramètres de la qualité de l'eau à l'échelle de son 8<sup>e</sup> programme d'intervention au travers notamment de l'utilisation des modèles développés à l'échelle du bassin par le programme de recherche PIREN Seine.

### 1.3. Favoriser l'accès à des méthodes et des outils de traitement des données

Il est important de poursuivre la mise à disposition du plus grand nombre d'outils de gestion et de traitement des données. Même si le SEQ s'est largement imposé comme méthode et outil d'analyse des données sur l'eau, sa généralisation s'est heurtée à la lourdeur de la méthodologie (comme l'importation des données) et à des possibilités limitées d'utilisation dans la mise en relation avec les usages (interprétation des données de mesure en continu, altérations incomplètes). L'absence d'officialisation de la version 2 du SEQ, largement inspirée des remontées des utilisateurs, a également présenté un frein à sa généralisation et a limité l'utilisation majeure de cet outil à l'analyse des données du « Réseau national de bassin » (voir la synthèse des réflexions du groupe de travail sur le SEQ dans ce dossier). Aujourd'hui, les utilisateurs sont en attente de la mise en place des outils qui permettront d'évaluer l'état écologique des cours d'eau.

On peut saluer à l'inverse les initiatives de mise à disposition d'outils comme ceux du portail ADES (MOLOSSE, PIEZEAU, SEQ, SYSIPHE), en plus des données brutes elles-mêmes.

L'harmonisation recherchée par l'architecture du SIE cherche à éviter que ne se reproduise le développement, notamment par des services de l'État, d'applications destinées par exemple à des évaluations à une échelle spécifique, par exemple à l'échelle régionale

(ex. : GDES), sans objectif de mutualisation et de partage à l'échelle du SIE (ex. : NOPOLU, développé pour l'IFEN). De même, les « Schémas directeurs des données sur l'eau des bassins » préconisent que les structures locales de gestion, telles que les CLE, doivent utiliser les outils mis à leur disposition pour « bancariser » et valoriser leurs données sans développer leur propre système d'information.

### 1.4. Besoin de raccourcir les délais de traitement

Le « Livre vert » sur l'architecture du système d'information sur l'eau, édité par le ministère de l'Écologie en janvier 2005, pointe particulièrement cette insuffisance de l'utilisation des données sur l'eau : « pour des études de synthèse sur la qualité des eaux, l'évaluation se heurte à la difficulté de réunir les données à partir de banques de données multiples, voire inexistantes, et dans des délais raisonnables... La disponibilité des données collectées au sein de banques de bassin reste hétérogène : l'accès aux données qualitatives sur les cours d'eau est variable à la fois en termes d'information mise à disposition (brute ou interprétée) et en termes de données géographiques ; les délais nécessaires pour la réalisation de synthèses nationales par extraction de bases locales sont excessifs. »

Un tel constat doit s'accompagner d'une interrogation sur la validation et la valorisation de ces données :

- qui valide les données : le producteur, le « collecteur » ou le futur utilisateur (sachant que le besoin de validation dépend souvent de l'usage envisagé) ?
- pour qui cette validation est-elle effectuée : acteurs de l'eau, grand public ? car le format des données (pas de temps par exemple ou consolidation) en dépend ainsi que les délais requis ;
- enfin, avec quels moyens (humains et donc financiers) cette validation est-elle faite ? On en revient au bénéfice de la donnée décrit ci-dessus qui est largement fonction de la qualité ou plutôt de la fiabilité de celle-ci.

En ce qui concerne les données utilisées dans les réseaux du contrôle de surveillance DCE, les schémas directeurs des données sur l'eau devraient préciser la nature des procédures de validation effec-

tuées au niveau du bassin et le délai autorisé pour la validation. Le « Livre vert » suggère dès lors que « si les procédures de validation ne peuvent être achevées dans ce délai autorisé, les données non validées doivent être publiées au plus tard un mois après l'expiration de ce délai, accompagnées de la mention de leur non-validation. » Cette diffusion de données non validées fait largement débat mais dépend fortement des réponses aux questions posées ci-dessus. Si des règles précises doivent être établies pour les bases nationales, la situation est à considérer au cas par cas, mais sans *a priori*, au niveau local, par les acteurs concernés.

## 2. Des données à considérer dans leur contexte

Si le SEQ a permis de formaliser l'analyse des données sur l'eau en évaluant la qualité de l'eau à partir de groupes de paramètres définis (les altérations), on constate un certain désintérêt pour des analyses de paramètres élémentaires (conductivité, pH, UV), le croisement des données et l'utilisation des informations issues du terrain, autres que les données climatologiques auxquelles les références ne sont pas encore systématiquement faites.

De nombreuses analyses réalisées à partir de paramètres élémentaires permettent d'obtenir un grand nombre d'informations sur l'état du milieu en un point donné et son environnement : diagramme de Piper<sup>1</sup> pour l'analyse des faciès chimiques et la définition de typologies locales, observation et analyse des variances, utilisation du « krigage » (modélisation statistique de données spatiales) pour l'optimisation des fréquences de mesure ou l'interpolation spatiale, corrélation multi-paramètres (exemple : conductivité et teneur en différents cations).

Il s'agit bien de valider les données entre elles (dans leur contexte événementiel d'où l'importance de l'interprétation terrain), de les corréler et d'en déduire les éléments d'appréciation suivants.

<sup>1</sup> Représentation graphique simplifiée de la concentration relative en ions positifs et négatifs d'une eau.

### • L'évaluation de l'état global du système à une échelle donnée

Une situation normale peut ainsi être définie à partir de la connaissance de l'état de plusieurs paramètres. L'analyse multi-paramètres permet de repérer les situations anormales de façon beaucoup plus fiable et précise qu'une détection sur seuils<sup>2</sup> ou que l'évaluation d'une altération.

### • Des stratégies de surveillance dans lesquelles peut être recherchée la complémentarité entre les mesures ponctuelles et les mesures en continu

Si les mesures ponctuelles ont en effet montré au fil du temps l'impact positif des politiques publiques (assainissement, épuration) sur la qualité physico-chimique des milieux, les défis à relever concernent principalement la gestion des flux et des pollutions en temps de pluie, ce qui nécessite un pilotage fin et en temps réel des réseaux et du milieu. On observe cependant aujourd'hui encore peu de convergence entre le suivi météorologique des réseaux et des cours d'eau (voir l'article de J.-L. COMBRISSON et *al.* dans ce dossier). Cette complémentarité peut intervenir au moment de la mise en place des réseaux de surveillance (choix des points, fréquence en fonction de la variabilité) mais également en cours de fonctionnement, en particulier pour relier les données issues de réseaux de connaissance et d'usage. L'ensemble des données collectées ponctuellement doit pouvoir être replacé dans son environnement spatial et temporel (connaissance de l'écart-type des valeurs réelles de contamination du milieu). La mesure en continu peut être un complément permettant de dresser un tel « paysage » avec des impératifs moindres en termes de validation (connaissance de l'incertitude).

Trois évolutions nécessaires sont toutefois à prendre en compte :

- on rencontre encore de nombreuses difficultés à lier les mesures de qualité des eaux et les données relatives aux pressions sur les milieux, comme les rejets des collectivités ou des activités économiques (ex. : rejets ponctuels ou diffus des activités économiques) ;

<sup>2</sup> GRAPIN : "Détection et localisation des rejets industriels diffus en assainissement", *Hydrosciences* 146, septembre 2004.

- il existe un important besoin de données spatialisées pour les usages, ainsi que de données pédologiques, parent pauvre au regard de l'information sur l'hydrographie de surface (BD-Carthage), la topographie (BD-Topo), l'occupation des sols (Corine Land Cover), ou l'hydrogéologie (Référentiel hydrogéologique français) ;

- les connaissances restent encore largement à approfondir en ce qui concerne les mécanismes qui régissent l'occurrence des contaminations observées dans le milieu : biodisponibilité des divers polluants, remise en solution à partir des stocks piégés par les sédiments, impacts écotoxicologiques sur les organismes vivants, processus de transfert sol-aquifère, en particulier au niveau de la zone non saturée<sup>3</sup>.

Ce contexte événementiel est également essentiel pour apprécier la qualité des données. À ce titre, la qualité des données fait moins référence à la précision de la donnée qu'à la fiabilité de celle-ci, même si les deux méritent d'être connues, adaptées au besoin et maîtrisées. Mais il est indispensable de se donner les moyens d'obtenir des résultats de mesure représentatifs des événements et une qualité de données constante<sup>4</sup>. La fiabilité illustre ainsi la disponibilité et la capacité d'ajustement de la précision de la mesure (qu'il s'agisse de mesures ponctuelles ou en continu). Les deux articles (Ph. NAMOUR et al. ; L. GRIMALDI et al.) présentés dans ce dossier montrent également que la motivation concernant la qualité des données est souvent d'ordre scientifique (données utilisées dans le cadre de la recherche), mais devient également souvent économique, liée soit au besoin de justifier des investissements réalisés, soit à celui d'optimiser un usage (dans ce cas la production d'électricité).

Enfin, on peut conclure en évoquant le besoin de définir les conditions d'une évolution plus rapide de la réglementation pour admettre des techniques de mesure innovantes, capables de fournir des données à un coût raisonnable. Il convient de créer les condi-

tions de l'expansion du marché en favorisant les synergies entre pouvoirs publics, chercheurs, fabricants et utilisateurs (voir à ce sujet l'article de C. GONZALES et al.).

### 3. Des besoins et des initiatives à reconnaître : problématique de la diffusion des données

La circulaire SDDE de décembre 2004 prévoit que « des réseaux ou des sites d'autres organismes (collectivités locales, associations, universités, CNRS...) peuvent être utilisés dans l'élaboration des réseaux de contrôle opérationnel. Il est nécessaire de veiller à ce qu'une organisation partenariale favorise alors une optimisation des dispositifs de contrôle opérationnel des différents acteurs impliqués dans la production des données. Des conventions sont particulièrement nécessaires avec des collectivités locales, des associations ou des entreprises qu'il est jugé opportun d'associer de façon durable à la production des données. »

Outre les précautions à prendre dans la consolidation d'un réseau à partir de données de nature différente, le « Livre vert » cité plus haut fait remarquer que « l'implication, souhaitable, d'un nombre croissant d'organisations (en particulier les collectivités territoriales) va entraîner la croissance des flux de données et le développement de nouveaux systèmes d'information. »

Il est apparu au cours des discussions que les échanges de données entre les différents interlocuteurs se trouvent limités principalement pour les raisons suivantes :

- problèmes de confiance (quelle fiabilité des données ?), de responsabilité (quelles conséquences à la libre disposition de mes données ?), de reconnaissance (quel bénéfice tirer de la mise à disposition de mes données ?) ;
- existence de freins techniques, organisationnels, économiques à la mise en commun : les données sont-elles pertinentes ? à quelle interprétation vont-elles pouvoir donner lieu ? dans quels délais peut-on espérer récupérer ces données ? à qui appartiennent-elles et à qui incombent les responsabilités liées à leur utilisation ?

<sup>3</sup> AESN : "État du milieu naturel, bilan à partir des observations du milieu", 2007.

MAILLOT : "Protection des captages d'eau de distribution publique", TSM n°4, 2005.

<sup>4</sup> NAMOUR : "Métrologie de terrain et qualité des données", OTHU FT7, 2002.

- les nouvelles « obligations » de mise à disposition de l'information (convention d'Aarhus, droit d'accès à l'information relative à l'environnement, DCE...) sont encore mal connues des producteurs de données et les implications, notamment économiques et juridiques, mal appréhendées.

En tout état de cause, compte tenu des transformations subies par une donnée avant comme après sa publication, il est indispensable, comme le suggère le « Livre vert », de préconiser l'emploi de métadonnées indiquant la traçabilité des données. Il pose également la question de la forme d'accès aux données : consultation et/ou déchargement. Dans la mesure du possible, les deux modes d'accès aux données devraient être proposés. Cette problématique est tout à fait transférable à tout type de données produites, aussi nous reproduisons les recommandations du « Livre vert » :

- « La présentation des données dans une page web contenant du texte et des images ou dans un document pdf permet à l'utilisateur de « voir » une information, éventuellement de la conserver pour la revoir ultérieurement, ou de l'imprimer pour la voir sur un autre support, mais ne lui permet aucune autre utilisation.

- Certains usagers peuvent avoir besoin des données pour les « utiliser » : leur appliquer un traitement particulier, pour les comparer à des données provenant d'autres sources, pour publier les résultats de ces traitements ou comparaisons.

- Toutes les données brutes doivent être disponibles au moins sous une forme « à utiliser » et de préférence sous les deux formes, « à voir » et « à utiliser ».

- Toutes les données élaborées doivent être disponibles au moins sous une forme « à voir », surtout si elles sont accompagnées de commentaires, et de préférence sous les deux formes, du moins pour leurs aspects quantitatifs. Les formes à voir doivent inclure un lien vers une forme à utiliser. »

Une analyse succincte, dans le cadre des travaux du groupe, de quelques sites Internet consacrés à la mise à disposition de données révèle une très grande richesse assortie d'une grande variété de la nature des données disponibles et de leur qualité. Les plus performants sont dotés d'une interface utilisateur permettant d'explorer les données disponibles à l'aide

d'un fond cartographique<sup>5</sup> ou proposent le chargement de logiciels de gestion et de traitement des données.

Une typologie des données disponibles se dégage :

- les données brutes, soit statiques (par exemple les caractéristiques d'un point d'eau), soit dynamiques (comme l'évolution de la cote d'une nappe) ;

- les données élaborées, soit non interprétées (par exemple le débit moyen mensuel d'une rivière), soit interprétées (comparaison d'un débit d'une rivière avec une valeur normale).

Globalement, il se révèle également très difficile d'évaluer la qualité des données issues de ces sites : ils fournissent rarement les méthodes d'évaluation ou le format utilisés, même si le format SANDRE se généralise progressivement, mais sans être systématiquement renseigné. A minima, nous appuyons la volonté de renseigner les fiches du dispositif de collecte<sup>6</sup> permettant de recenser de façon exhaustive les réseaux de mesure existants et leur utilisation.

Il faut insister sur le lien entre données et cartographie : pour faciliter l'exploitation et la valorisation des bases de données, elles doivent être géoréférencées. Cette exigence se révèle particulièrement importante dans le cadre de l'élaboration des SAGE et plus généralement des programmes de mesure et plans de gestion de la directive cadre.

Le groupe s'est attaché à mettre en valeur des initiatives collectives de mise en commun des données, tout en cherchant à comprendre les moteurs (comme l'atteinte d'une connaissance globale et intégrée des hydrosystèmes) et surtout les freins à cette mise en commun :

- une exigence de qualité imposée (délais, exigences métrologiques) ;

- le coût de l'acquisition et /ou de la mise en commun ;

- le traitement des données ;

- l'hétérogénéité des systèmes et de leur finalité.

<sup>5</sup> Ex. : SIGORE, le SIG de l'Observatoire régional de l'environnement de la région Poitou-Charentes ([sigore.observatoire-environnement.org](http://sigore.observatoire-environnement.org)) ; L'eau dans l'Allier ([eau-dans-allier.cg03.fr](http://eau-dans-allier.cg03.fr)) ; SIGESAQI, Système d'information pour la gestion des eaux souterraines en Aquitaine ([sigesaqi.brgm.fr](http://sigesaqi.brgm.fr))

<sup>6</sup> <http://dispositif.rnde.tm.fr/RESEAUX/presentation/main/index.php>

Deux exemples sont plus particulièrement présentés dans ce dossier.

- Le portail ADES dont l'objectif est de constituer un outil de collecte et de conservation de toutes les données sur les eaux souterraines, mobilisable par un large ensemble de partenaires. Le succès de cet outil réside dans la reconnaissance, connue et reconnue, de son rôle centralisateur qui constitue un gain de temps pour la recherche de l'information. Le développement de l'outil s'est également associé dès l'origine à une démarche de service avec fourniture d'information en plus de la donnée : localisation, lien avec les masses d'eau, référentiel hydrogéologique, historique des données, outil d'extraction sélective. Des problèmes néanmoins demeurent comme notamment, celui d'alimentation de la base (fréquence de chargement très hétérogène), d'où la nécessité là encore d'imaginer des processus de conventionnement, notamment avec des collectivités.
- L'APRONA est une association créée en 1995 à l'initiative de la région Alsace, des départements du Haut-Rhin et du Bas-Rhin, de la préfecture de région et de l'agence de l'eau Rhin Meuse. Elle gère le réseau piézométrique et le réseau qualité ainsi que la station d'alerte sur le Rhin, mais ses missions sont beaucoup plus vastes (voir l'article de M. HERR). Une telle organisation partenariale sous le mode associatif est assez particulière dans le contexte français de suivi des données sur l'eau et s'apparente sur bien des points aux missions et au fonctionnement des associations agréées pour le suivi de la qualité de l'air.

#### 4. Propositions d'évolution / recommandations

Les discussions au sein du groupe de travail ont été riches et variées. Elles ont avant tout montré le besoin d'information et d'échanges des producteurs de données, qu'ils se situent dans le « premier cercle » des partenaires du SIE aussi bien que dans la multitude de producteurs de données (le catalogue des dispositifs recense aujourd'hui plusieurs centaines de systèmes de collecte de données et d'information sur l'eau).

Concernant les réseaux de connaissance, les recommandations émises à l'issue de l'étude sur les dispo-

sitifs de collecte menés de 2001 à 2003 par la direction de l'Eau restent d'actualité même si certaines ont très largement progressé depuis :

- poursuite de la mise en œuvre des réseaux nationaux de connaissance des eaux souterraines ;
- mise en place d'un réseau sur les plans d'eau ;
- organisation du suivi des masses d'eau utilisées pour l'alimentation en eau potable et abandonnées ;
- mise au point d'indicateurs de qualité et écologiques plus complets : évolution du SEQ en système d'évaluation de l'état ;
- recueil des données relatives aux diverses pressions qui agissent sur la qualité de l'eau (occupation des sols, intrants utilisés en agriculture, installations classées, zones protégées...) et économie.

Des éléments qui précèdent, il nous semble également important de préciser que la qualité des données, leur exploitation et leur traitement sont un préalable à la densification des réseaux.

Dans tous les cas, des progrès sont à réaliser dans la fiabilisation de la production des données, le développement de démarches qualité pour l'exploitation des réseaux (validation des données) et surtout, le traitement des données pour en tirer toute l'information nécessaire à la gestion intégrée des hydrosystèmes.

Ceci implique d'engager rapidement les discussions entre services de l'État et maîtres d'ouvrage afin d'articuler les besoins et les contraintes réglementaires des acteurs locaux (réseaux d'usage) et les exigences de la DCE (réseau de contrôle opérationnel), dans une perspective de co-construction.

En ce qui concerne la mesure en continu, on peut distinguer :

- le cas où des stations de mesure existent, développées dans un contexte particulier avec des objectifs bien précis : besoin de mieux valoriser les données pour des objectifs multiples à différentes échelles ;
- lorsque les stations n'existent pas, la multifonctionnalité des installations peut être recherchée, ce qui suppose de mieux coordonner les interventions de divers acteurs, notamment par exemple dans le cadre du développement des SAGE ; ceci suppose de définir des règles simples et faciles à mettre en pratique :

- pour la conception des systèmes d'acquisition : installation prioritaire aux exutoires de bassins-versants contre points plus ciblés avec mesures de débit pour calculs de flux ; évaluation de la pérennité de l'installation (stations légères, mobiles, campagnes localisées et ciblées au fil de l'eau) ; évaluation du coût par comparaison avec campagnes ponctuelles à fréquence élevée.

- et pour la diffusion des données à une échelle locale (définition des objectifs, validation des données, mise à disposition).

Dans tous les cas, des progrès sont à réaliser dans la fiabilisation de la production des données, le développement de démarches qualité pour l'exploitation des réseaux (validation des données) et surtout le traitement des données pour en tirer toute l'information nécessaire à la gestion intégrée des hydrosystèmes.