

*Santé environnement*

**Déterminants naturels et humains  
des pollutions fécales accidentelles  
des petites unités d'alimentation  
en eau potable, France, 2003-2004:  
une étude écologique géographique**

# Sommaire

Résumé	2
Introduction	3
<b>1. Matériel et méthodes</b>	<b>4</b>
1.1 Définition d'une "pollution"	4
1.2 Classes de taille des Unités de Distribution	5
1.3 Les covariables	5
1.4 Méthodes statistiques	7
<b>2. Résultats</b>	<b>9</b>
2.1 Nombres de prélèvements, de non conformités et de "pollutions"	9
2.2 Distribution statistique et géographique des variables du modèle	9
2.2.1 Nombre d'unités de distribution ( <i>N_UD</i> )	10
2.2.2 Nombre de pollutions observées ( <i>N_Pol</i> )	11
2.2.3 Nombre de pollutions observées standardisé sur le nombre d'unités de distribution ( <i>N_Pols</i> )	11
2.2.4 Unités de distribution sans désinfection ( <i>sDes</i> )	11
2.2.5 Captages sans périmètre de protection ( <i>sDUP</i> )	12
2.2.6 Précipitations ( <i>Prec</i> )	13
2.2.7 Surface karstifiée ( <i>Karst</i> )	13
2.3 Modèle linéaire	13
2.4 Analyse de la variance	18
2.5 Interactions	18
2.6 Vers un modèle unique	18
<b>3. Discussion</b>	<b>20</b>
3.1 Santé publique : populations résidentes et populations de passage	20
3.2 Corrélations écologiques et causalités	20
3.3 Renforcer l'interprétation des corrélations	22
<b>4. Conclusion</b>	<b>24</b>
Références bibliographiques	25

# Déterminants naturels et humains des pollutions fécales accidentelles des petites unités d'alimentation en eau potable, France, 2003-2004 : une étude écologique géographique

## **Coordination, réalisation et rédaction :**

Pascal Beaudéau (Institut de veille sanitaire)

## **Contributions :**

Direction générale de la santé, bureau de l'eau : Henri Davezac (expertise Sise-Eaux et fourniture des données), Charles Saout, Alban Robin (expertise eau).

Laboratoire de morphodynamique continentale et côtière, Université de Rouen : Danièle Valdes (expertise hydrogéologique, fourniture des données, cartographie), Jean-Paul Dupont (expertise hydrogéologique).

Institut de veille sanitaire : Marie Teyssandier (protocole, documentation, calcul), Damien Mouly, Pascal Beaudéau (protocole, calcul), Véréne Wagner (calcul), Perrine de Crouy-Chanel, Julien Koutcheravenko et Morgane Stempfelet (cartographie).

Validation : Jean-Paul Dupont (Université de Rouen).

# Résumé

La prévention et la surveillance des épidémies d'origine hydrique reposent en grande partie sur la détection de cas groupés de gastro-entérites. Il n'y a cependant eu en France entre 1998 et 2006 aucune épidémie rapportée touchant des unités de distribution (UD) desservant moins de 1 000 usagers. De fait, l'absence d'école et d'établissements susceptibles de concentrer des malades contrarie la détection des épidémies. La surveillance des pollutions microbiologiques accidentelles de l'eau distribuée pourrait compléter utilement la surveillance épidémiologique, notamment dans la recherche des facteurs de risque.

Nous avons défini une pollution microbiologique accidentelle de l'eau distribuée (notée "pollution") comme l'occurrence d'au moins 20 bactéries fécales (*Escherichia coli* et entérocoques) formant colonies parmi les analyses réalisées sur les prélèvements (prise d'essai 100 mL) collectés un même jour sur une même UD. Le secteur d'étude couvre 90 départements de France continentale et la période d'étude les années 2003 et 2004. L'étude s'est limitée à la catégorie des UD de moins de 2000 usagers car cette catégorie bénéficie d'une bonne homogénéité : 9 UD sur 10 sont alimentées par des ressources souterraines traitées au plus par une simple chloration. Trois classes de taille sont distinguées à l'intérieur de la catégorie étudiée (<100, 100-499 et 500-1 999 habitants) afin de contrôler l'effet "taille", notamment les variations de fréquence des "pollutions" (10 à 1) et des prélèvements (1 à 2) entre les classes <100 et 500-1 999. Nous avons exploré les corrélations entre la fréquence des "pollutions" et des indicateurs météorologiques, hydrologiques, topographiques et géologiques, ainsi que des indicateurs de gestion sanitaire de l'eau. Des méthodes de régressions linéaires multiples ont été mises en œuvre pour modéliser la relation entre le nombre standardisé de "pollutions" par département et les covariables significatives.

Le contrôle sanitaire a mis en évidence 2 739 "pollutions" répondant à la définition de cas. La fréquence des "pollutions" varie amplement d'un département à l'autre. Cette surdispersion persiste après standardisation sur la distribution des tailles d'UD qui est très hétérogène entre départements.

Les précipitations sont corrélées très significativement au risque de "pollution" avec un accroissement du risque de 30 % par 100 mm supplémentaires de hauteur d'eau précipitée annuellement. La proportion du territoire occupée par des formations de surface karstifiées, la proportion de ressources sans périmètre de protection et la proportion d'UD dépourvues de désinfection sont aussi corrélées positivement au risque de "pollution". Pour une augmentation de 10 % de chacune de ces trois covariables, on observe un accroissement égal du risque relatif. Ces corrélations sont généralement reproductibles (pentes et significativités) entre les trois classes de taille et conformes aux connaissances scientifiques ou expertes. De nombreux départements présentent encore un excès de risque significatif après contrôle des covariables, ce qui suggère la présence de facteurs de risque de "pollution" non identifiés. Ce phénomène touche plus particulièrement les UD desservant plus de 500 habitants, vraisemblablement à cause de l'apparition dans cette catégorie de traitements plus sophistiqués qui estompe le rapport entre qualité de l'eau brute et qualité de l'eau distribuée.

En dépit d'arguments favorables (reproductibilité, vraisemblance "biologique"), les corrélations écologiques observées ne peuvent être interprétées en termes de causalité, sauf pour les précipitations. Dans la mesure où les prévalences de la désinfection et de la protection des ressources ont régulièrement augmenté dans les années 2000, l'ajustement du modèle sur différentes périodes de temps prises dans cette décennie devrait confirmer le rôle direct ou indirect de ces indicateurs.

# Introduction

Le contrôle réglementaire de la qualité microbiologique de l'eau repose depuis un siècle sur la recherche dans l'eau distribuée d'indicateurs de pollution fécale, tels que *Escherichia coli* et les entérocoques fécaux. Plusieurs études épidémiologiques ont montré une association entre les indicateurs bactériens et l'incidence des gastro-entérites, ainsi qu'un effet dose (Zmirou *et al.* 1987, 1995), sans toutefois dégager de fonction dose-réponse quantifiée. Plusieurs facteurs viennent en effet compliquer cette relation.

La plupart des souches des bactéries indicatrices ne sont pas pathogènes. Leur présence n'indique donc pas celle d'agents pathogènes mais un risque de présence d'agents pathogènes, variable selon le contexte épidémiologique, c'est-à-dire selon le niveau d'excrétion d'agents pathogènes dans les populations humaines ou animales. L'aléa épidémiologique cause donc une perte de spécificité de l'indicateur.

Le traitement de l'eau, quant à lui, entame la sensibilité des indicateurs. La désinfection de l'eau par les produits chlorés, largement répandue en France, inactive les bactéries pathogènes et indicatrices, sans affecter les parasites ni les virus. L'absence de bactéries indicatrices ne garantit donc pas l'absence de risque sanitaire - notamment du risque viral ou parasitaire, tandis que leur présence implique seulement une probabilité de risque sanitaire.

Enfin, on ne dispose pas en pratique de données consistantes sur l'exposition des personnes aux agents pathogènes pendant la période d'exposition aiguë précédant les épidémies, que ce soit au moyen de mesures directes des concentrations en pathogènes (données pratiquement inexistantes) ou à travers des mesures d'indicateur de contamination fécale (données très parcellaires).

En conclusion, dans un contexte épidémiologique donné et en l'absence de traitement, le risque sanitaire d'origine hydrique est une fonction monotone croissante des concentrations en bactéries fécales dans l'eau distribuée, mais la prise en considération d'un grand nombre de situations épidémiologiques et de différents niveaux de traitement introduit une proportion importante de faux positifs et de faux négatifs dans cette relation.

Admettant l'existence d'une relation dose-réponse "moyenne" et sachant qu'il y a 1 500 à 2 000 pollutions fécales des eaux d'alimentation (définition *infra*) rapportées par année en France contre seulement une épidémie documentée en moyenne, il est très avantageux en terme de sensibilité de rechercher les déterminants des pollutions fécales des eaux distribuées plutôt que des épidémies. L'existence de facteurs spécifiques des contaminations parasitaires et virales biaise marginalement la portée des conclusions quant au risque épidémique.

En France, le nombre de pollutions fécales accidentelles de l'eau distribuée (*N\_Pol*) varie très sensiblement d'un département à l'autre (Beaudeau *et al.* 2006). Cette disparité reflète en premier lieu celle de la distribution en taille des UD car la probabilité des pollutions augmente fortement quand la taille de l'UD se réduit (ministère de la Santé et des Solidarités, 2006). La standardisation de *N\_Pol* sur la distribution des UD par classe de taille ne réduit pas la surdispersion<sup>1</sup> de *N\_Pol*. Cela laisse augurer la présence de déterminants naturels ou humains.

L'objet de ce travail est de rechercher ces facteurs à travers une étude écologique géographique, dont l'unité élémentaire est le département, et de quantifier leur effet, avec les précautions inhérentes aux méthodes écologiques employées.

---

<sup>1</sup> La distribution aléatoire de référence étant une distribution de Poisson.

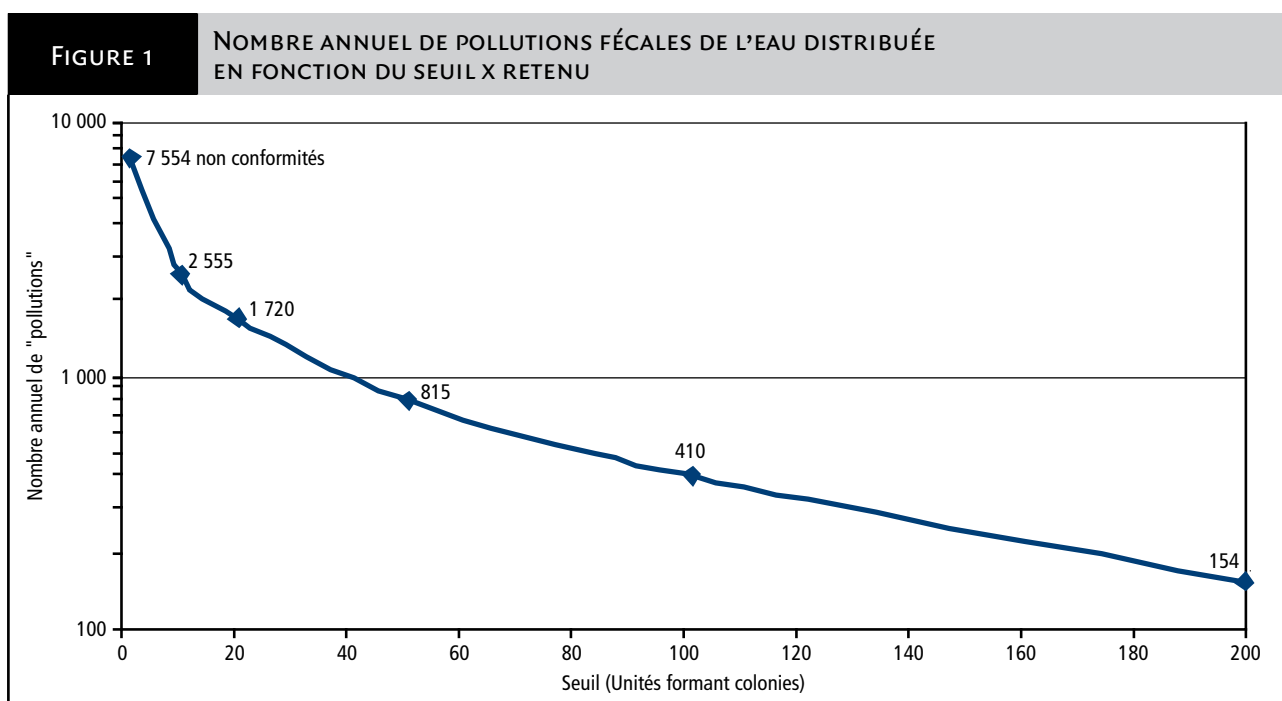
# 1. Matériel et méthodes

## 1.1 DÉFINITION D'UNE "POLLUTION"

Une pollution fécale accidentelle des eaux d'alimentation, qu'on appellera par la suite "pollution", est définie comme l'occurrence de résultats microbiologiques tels que la somme des unités formant colonies décomptées sur les deux indicateurs spécifiques de pollution fécale (*Escherichia coli* "EC" plus entérocoques "En") et sur l'ensemble des échantillons prélevés un même jour sur un même réseau, égale ou excède une valeur seuil de 20. Les méthodes analytiques sont imposées par la réglementation et uniformes : filtration sur membrane de 100 mL, culture sur milieu sélectif en boîte de Pétri et décompte des unités formant colonies après une incubation de 48h à 44°C, selon

les normes internationales ISO 9308-1:2000 (*Escherichia coli*) et ISO 7889-1:1998 (entérocoques).

Le critère de non-conformité microbiologique (absence de bactéries fécales cultivables) n'a pas été retenu car les contaminations de faible intensité peuvent être dues à une contamination fortuite lors du prélèvement ou à l'érosion du biofilm causée par un à coup hydraulique et n'indiquent généralement pas un problème généralisé au réseau de distribution. Plus le seuil est élevé, plus la probabilité d'épidémie est importante (Cf. introduction). On a donc cherché à maximiser la valeur du seuil sans compromettre la puissance statistique qui diminue rapidement quand le seuil augmente (figure 1).



Source : Sise-Eaux, Ddass, Drass, Ministère chargé de la Santé, données 2003-2004.

Une pollution est définie comme l'occurrence de résultats microbiologiques tels que la somme des unités formant colonies décomptées sur les deux indicateurs (*Escherichia coli* "EC" et entérocoques "En") et sur l'ensemble des échantillons prélevés un même jour sur un même réseau, égale ou excède une valeur seuil de  $x$ .

Les données microbiologiques ont été tirées de la base Sise-Eaux du ministère chargé de la Santé, qui rassemble de façon exhaustive les résultats analytiques obtenus au titre du contrôle sanitaire de l'eau potable, mais aussi un grand nombre d'informations structurelles sur la ressource, le traitement, le réseau, la production, l'opérateur et la population desservie. Les prélèvements "sortie usine" et "réseau" sont comptabilisés dans le calcul de  $N_{Pol}$ , à l'exception (i) des analyses "non représentatives" au sens de Sise-Eaux, c'est-à-dire réalisées sur des échantillons prélevés sur de l'eau non distribuée, (ii) des analyses recommencées consécutivement à une non-conformité

microbiologique ou plus généralement (iii) des analyses réalisées pour d'autres motifs que le contrôle sanitaire planifié (motif  $\neq$  CS, selon la nomenclature Sise-Eaux).

Les départements de la petite couronne parisienne ont été exclus car leurs ressources sont situées hors département. Les départements d'Outre-mer et la Corse sont aussi exclus car les pollutions fécales de réseaux y sont beaucoup plus fréquentes qu'ailleurs en France. De plus, on peut suspecter la présence de facteurs de risque particuliers défavorables à une analyse conjointe.

## 1.2 CLASSES DE TAILLE DES UNITÉS DE DISTRIBUTION

Trois classes de tailles d'UD sont distinguées : les UD desservant une population fixe de moins de 100 habitants, c'est-à-dire des hameaux isolés en secteurs montagneux (classe 1), les petites UD de 100 à 499 habitants (classe 2) et les UD de 500 à 2 000 habitants (classe 3). Les trois classes correspondent à des populations cumulées de 0,29M, 1,82M et 5,87M d'habitants, respectivement.

Deux critères ont prévalu pour le choix des limites de classes :

- l'homogénéité des classes quant à la nature de l'eau dans l'optique d'assurer une meilleure interprétation des résultats. Les trois classes de taille retenues sont homogènes du point de vue du type de ressource et du type de traitement. Entre 97 % et 94 % des ressources utilisées sont en effet des ressources souterraines, selon la classe de taille. 87 % des traitements mis en œuvre se réduisent à une désinfection avec éventuellement, un traitement physique simple tel qu'une filtration rapide sur sable (tableau 1). Globalement, le parc d'UD étudié est essentiellement alimenté par des ressources souterraines simplement désinfectées ;

- l'homogénéité intraclasse de la fréquence d'échantillonnage. La fréquence que prévoit la réglementation est 3 à 5 prélèvements annuels pour la classe 1 ; 5 à 6 pour la classe 2 et 8 pour la classe 3.

La catégorie des UD de 2 000 et plus est extrêmement hétérogène puisqu'elle inclut à la fois des gros villages et des réseaux urbains desservant de centaines de milliers d'usagers, des ressources souterraines et des ressources de surface (selon la sous-catégorie de taille 21 % à 79 % des ressources utilisées sont des eaux de surface) et enfin, une grande variété de nature et de niveau de traitement (tableau 1). Cette catégorie n'est pas étudiée.

La variable à modéliser est le nombre annuel de "pollutions" constatées sur la période 2003-2004, standardisé sur l'effectif départemental dans la classe de taille :

$$N\_Pol_{ij} = N\_Pol_{ij} \times \frac{N\_UD_j}{N\_UD_{ij}}$$

où  $i$  indice le département et  $j$  la classe de taille d'UD.  $N\_UD_j$  est le nombre moyen d'UD par département pour la classe  $j$ .

TABLEAU 1		RÉPARTITION DES UNITÉS DE DISTRIBUTION PAR TYPE DE TRAITEMENT				
	N	A1	A2	A3	Total	
<100 (classe 1)	9,6 %	86,6 %	3,0 %	0,7 %	100,0 %	
100-1 999 (classes 2+3)	2,1 %	87,4 %	8,4 %	2,2 %	100,0 %	
2 000-9 999	0,6 %	63,0 %	21,9 %	14,6 %	100,0 %	
10 000-49 999	0,9 %	43,8 %	31,3 %	24,0 %	100,0 %	
50 000-99 999	0,0 %	30,0 %	10,0 %	60,0 %	100,0 %	
≥100 000	0,0 %	7,7 %	7,7 %	84,6 %	100,0 %	

N : absence de désinfection ; A1 : désinfection et, éventuellement, traitement physique simple ; A2 : traitement complet ; A3 : traitement complet avec affinage. Estimations basées sur le rapport du ministère chargé de la Santé et des Solidarités, 2006.

## 1.3 LES COVARIABLES

Les variables pressenties pour indiquer des facteurs du risque de "pollution" (tableau 2), qu'on appellera covariables, ont été déterminées à dire d'experts : personnel du ministère (Direction générale de la santé, Direction départementale des affaires sanitaires et sociales...) et hydrogéologues de l'université de Rouen. On peut distinguer :

- des covariables de gestion sanitaire telles que la prévalence de l'absence de désinfection ( $sDes$ ) ou de l'absence de périmètre de protection officialisé par une déclaration d'utilité publique<sup>2</sup> ( $sDUP$ ) ;
- des covariables indiquant le statut (mode de gestion délégué  $Priv$ ) et le coût ( $Prix$ ) de production de l'eau, supposées en lien avec la qualité de la gestion technique selon certains experts ;
- des covariables indiquant le morcellement de la production qui induit des difficultés de gestion technique et sanitaire, soit directement

comme le nombre d'UD ( $N\_UD$ ) et le nombre de captages par UD ( $Cpt/UD$ ), soit en lien avec le relief (l'altitude moyenne  $Alt$ ) ;

- des covariables météorologiques (précipitation moyennes  $Prec$  et écart relatif à la normale  $dPrec$ ), hydrologiques (densité du réseau hydrographique de surface  $DRHS$ ) et géologiques (proportion de la surface occupée par des formations karstiques  $Karst$ ) – Cf. Encart p.6 ;
- une covariable d'occupation du sol (proportion des terres en culture  $Cult$ ).

L'historique des données structurelles sur les installations d'eau potable n'est pas conservé dans Sise-Eaux. Compte tenu des délais dans le porté à connaissance des modifications structurelles apportées aux installations, les données non analytiques relatives à la gestion de l'eau potable ( $sDUP$ ,  $sDes$ ...), extraites de Sise-Eaux en juillet 2007, représentent la situation 2006.

<sup>2</sup> La notion des périmètres de protection est apparue dans la loi française dès 1902. Leur instauration est promue par l'Organisation mondiale de la santé (Water Safety Plans, 2004) et obligatoire en France depuis 1964. Elle vise à protéger les captages des pollutions accidentelles par une certaine disposition réglementaire, notamment l'acquisition des terrains autour du captage, la réglementation des activités à risque dans le secteur correspondant au cône d'appel du pompage, telles que l'épandage de lisier ou le forage. La régularisation des captages non protégés est en cours en France (Plan national santé environnement).

## Encart - Le karst, les formations karstiques en France et les cartes disponibles

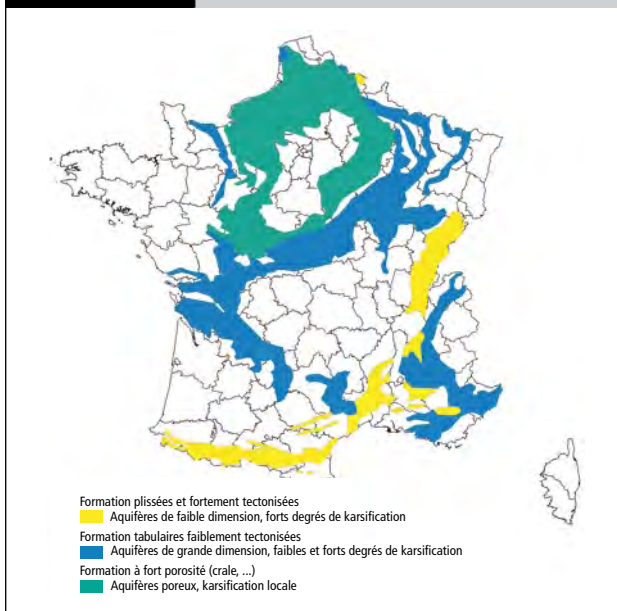
Les formations et les paysages karstiques résultent d'un processus d'érosion dominé par la dissolution du substrat carbonaté. L'hydrologie se caractérise par le développement d'un réseau de drainage souterrain (fissures, pertes, rivières souterraines). Les eaux qui s'introduisent dans les pertes sont restituées sans filtration au niveau des sources et des résurgences. Les eaux karstiques sont donc particulièrement vulnérables aux pollutions particulières et microbiologiques. Les pluies intenses accroissent plus que pour les autres types d'aquifères les risques de pollution. Le substrat karstifié peut être couvert de formations non karstifiables (Nicod 1994), percées au niveau des pertes. Ces formations qui, par temps de pluie modérée, contribuent à la filtration des eaux, peuvent, en cas de pluies intenses, augmenter le risque de pollution à cause des ruissellements de surface.

Ainsi, la Craie de Haute-Normandie est recouverte d'argile à silex et surmontée d'un sol limoneux "battant". Les pluies intenses engendrent alors des ruissellements de surface qui traversent le karst et sont restituées en quelques heures aux exutoires. Ces crues turbides sont repérables par leurs eaux peu minéralisées (transit rapide des eaux météoriques) et par leur charge en matières solides d'origine terrigène.

Un tiers du territoire français est occupé par des formations superficielles propices au développement de karsts. Depuis les années 1970, de nombreux auteurs ont participé à la cartographie des secteurs karstiques en France. Marsaud<sup>3</sup> (1995) distingue trois catégories de formations karstiques selon la tectonique et la porosité du substrat carbonaté (figure 2). Cette catégorisation n'a pas été conçue spécifiquement pour la prédiction de la vulnérabilité des eaux aux pollutions microbiologiques. Elle n'a pas été retenue dans la formation de la covariable *Karst* (proportion de surface karstifiée) qui traite uniformément des trois secteurs.

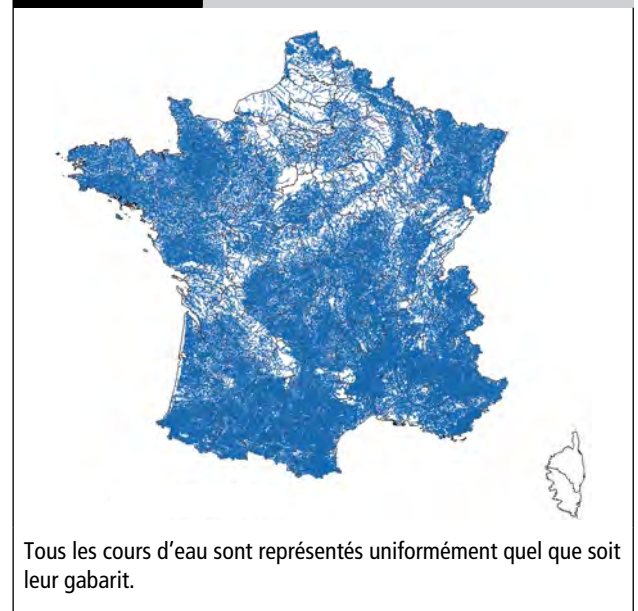
À drainage total constant, c'est-à-dire à précipitations efficaces constantes, la densité du réseau hydrographique de surface (*DRHS*; figure 3) est inversement proportionnelle à la densité du réseau de drainage souterrain. Étant donné la part importante que prennent les réseaux karstiques dans le drainage souterrain, on attend une correspondance "en négatif" entre la carte géologique (figure 2) et la carte hydrographique (figure 3). Si cette correspondance ressort clairement dans le pourtour du Bassin Parisien, d'autres secteurs y échappent : la Beauce (substrat calcaire peu évolué), les Landes (substrat sableux très perméable) et les secteurs de haute montagne où coexistent des karsts développés et un important réseau de surface.

FIGURE 2 LE KARST EN FRANCE



Source : Marsaud, 1997.

FIGURE 3 RÉSEAU HYDROGRAPHIQUE DE SURFACE



Source : École des Mines, Paris.

<sup>3</sup> Différents documents ont été utilisés par Marsaud pour dresser cette carte synthétique :

- Carte géologique de la France 1/1 500 000 ;
- Carte hydrogéologique de la France 1/1 500 000 (Margat, 1986) ;
- Cartes des formations karstiques (Gèze, 1973, Crampon, 1994).



**TABLEAU 2 LES VARIABLES ET LES DONNÉES**

Variables à expliquer		Détail <sup>a</sup>	Données manquantes	Source
<i>N_Pol</i>	Nombre de "pollutions" observées, années 2003-2004	Oui	Non	Sise-Eaux, Min Santé
<i>N_Pols</i>	Idem standardisé sur le nombre d'UD	Oui	Non	Sise-Eaux, Min Santé
<b>Covariables retenues</b>				
<i>sDes</i>	Proportion d'unités de distribution sans désinfection (type traitement = N)	Oui	Dpts 03 et 06 (2 %)	Sise-Eaux, Min Santé
<i>sDUP</i>	Proportion de captages sans périmètre de protection officialisé par une déclaration d'utilité publique	Oui	<i>sDUP1</i> : 4/87 (5 %)	Sise-Eaux, Min Santé
<i>Karst</i>	Proportion de surface karstifiée		Non	Marsaud, 1997
<i>Prec</i>	Précipitations moyennes annuelles, années 2003-2004 (mm)		Non	Météo-France
<b>Indicateurs testés non retenus comme covariables du modèle</b>				
<i>DRHS</i>	Densité du réseau hydrographique de surface (m/km <sup>2</sup> )		Non	École des Mines, Paris
<i>Alt</i>	Altitude médiane des communes (m)		Non	Institut Géographique National
<i>dPrec</i>	Écart des précipitations aux normales (1971-2000)		Non	Météo-France
<i>N_UD</i>	Nombre d'UD	Oui	Non	Sise-Eaux, Min Santé
<i>Cpt/UD</i>	Nombre moyen de captages par UD	Oui	Non	Sise-Eaux, Min Santé
<i>Cult</i>	Proportion de terre arable		Non	Corine Land Cover
<i>Prix</i>	Prix moyen de l'eau incluant l'assainissement (€/m <sup>3</sup> )	Non	Non	Institut Français environnement
<i>Priv</i>	Proportion d'UD en gestion déléguée	Oui	Non	Sise-Eaux, Min Santé
<b>Classes de taille</b>				
1	UD < 100 habitants			
2	100 ≤ UD < 500 habitants			
3	500 ≤ UD < 2 000 habitants			

UD = Unités de distribution.

<sup>a</sup> Variable détaillée par classe de taille d'UD.

## 1.4 MÉTHODES STATISTIQUES

Des données manquantes affectent les covariables *sDUP* et *sDes*. Elles ont été remplacées par les valeurs moyennes de la classe d'appartenance quand les données manquent dans les trois classes de taille. Quand la donnée est absente pour une seule classe de taille, elle est remplacée par la valeur moyenne observée dans la classe corrigée de l'écart à la moyenne constaté dans la classe adjacente, par exemple

$$sDUP1_{dm} = \overline{sDUP1x} \frac{sDUP2}{sDUP2}$$

Cela se justifie par le fait que les deux covariables relèvent de la gestion sanitaire mise en œuvre par les mêmes acteurs départementaux et que leurs déclinaisons par catégorie de taille sont donc corrélées (Cf. Résultats).

Des méthodes de régression linéaires ont été mises en œuvre sous Stata<sup>TM</sup>.

Dans la phase exploratoire de recherche des covariables significatives, des régressions pas à pas (*stepwise*) sont mises en œuvre. Trois critères sont examinés pour la sélection des covariables : le degré de significativité

de l'association avec *N\_Pols*, la vraisemblance de l'association au regard de l'expertise et la cohérence interclasses de la relation. Les trois conditions correspondantes doivent être remplies conjointement : corrélations partielles significatives à  $p=0,05$  ; signe du coefficient de la droite de régression conforme à l'attendu ; reproductibilité entre classes du coefficient et de la probabilité associée ( $p<0,10$ ).

La validation des modèles finalisés repose sur l'examen des représentations graphiques des droites de régression et des observations, sur l'étude et la normalisation de la distribution des résidus et sur le test de la robustesse du modèle aux différentes options de calcul (Cf. *infra*).

L'option "*robust*" est retenue pour prendre en compte l'hétéroscédasticité. La représentation des classes de taille varie énormément selon le département, notamment pour la classe [0;100] (0 à 588 UD). Pour réduire l'influence des départements pour lesquels la classe de taille est faiblement représentée, c'est-à-dire pour lesquels l'estimation de la fréquence des "pollutions" est imprécise,  $N\_Pols_{ij}$  est pondérée par la racine carrée du nombre d'UD<sup>4</sup>.

<sup>4</sup> Par définition de  $N\_Pol_{s_{ij}}$  :  $Var(N\_Pol_{s_{ij}}) = (N\_UD_i / N\_UD_{ij})^2 \times Var(Pol_{ij})$ .

Si le processus est Poissonnien, alors  $var(Pol_{ij}) = moy(Pol_{ij})$  ; si, de plus, il est homogène sur l'ensemble des départements, alors :

$var(Pol_{ij}) = moy(Pol_{ij}) = cste_1 \times N\_UD_{ij}$

D'où  $Var(N\_Pol_{s_{ij}}) = cste_2 / N\_UD_{ij}$

En l'absence de pondération, chaque département contribue à la variance résiduelle en raison inverse du nombre d'UD. On rend égal le niveau de contribution des départements en pondérant  $N\_Pol_{s_{ij}}$  par la racine carrée de  $N\_UD_{ij}$ .

La normalité des résidus de modélisation a été testée par le test de Kolmogorov-Smirnov. L'écart à la normalité a été réduit par ajout de covariables indicatrices (0/1) des départements qui s'écartent très significativement du modèle. Ces "outliers" ont été définis par une probabilité d'occurrence inférieure à  $10^{-3}$ , c'est-à-dire une différence entre valeurs observée et attendue supérieure à 3 fois l'écart-type du résidu ( $|Obs-Exp| > 3s$ ). Les valeurs prises par les coefficients représentent le nombre de "pollutions" en excès par rapport à la prévision du modèle national.

Les corrélations entre les trois déclinaisons catégorielles de la variable *N\_Pols* (variables d'entrée et résidus de modélisation) sont calculées pour identifier les éventuelles colinéarités et effets de confusions qui pourraient en découler entre *N\_Pols* d'une part, et les covariables d'autre part.

## 2. Résultats

### 2.1 NOMBRES DE PRÉLÈVEMENTS, DE NON CONFORMITÉS ET DE "POLLUTIONS"

Le nombre de prélèvements pris en compte pour la recherche des "pollutions" (France continentale sauf Paris et la petite couronne, années 2003 et 2004) s'élève approximativement<sup>5</sup> à 246 000. Le nombre de "pollutions" est de 2 739. Le taux de "pollution", qui est homogène au risque individuel d'exposition quotidienne d'un consommateur d'eau du robinet pour la boisson, est 10 fois plus

important en classe 1 qu'en classe 3, et 100 fois plus important que celui associé aux plus grandes UD (tableau 3). Il est très probable qu'un buveur d'eau du robinet résidant en secteur rural montagnoux soit exposé au cours d'une année à une "pollution". La population résidente concernée reste cependant modeste en nombre.

À l'instar des "pollutions", les non-conformités sont 10 fois plus fréquentes parmi les UD < 500 (réunion des classes 1 et 2) que parmi les autres (classe 3). Elles y sont aussi plus graves avec des "pollutions" 2 fois plus fréquentes parmi les non-conformités (tableau 4).

	UD < 100 (classe 1)	100 ≤ UD < 500 (classe 2)	500 ≤ UD < 2 000 (classe 3)	2 000 ≤ UD < 10 000 <sup>c</sup>	10 000 ≤ UD < 50 000 <sup>c</sup>
Population résidente desservie (millions) <sup>a</sup>	0,29	1,82	5,87	16,5	16,9
Nombre réglementaire de prélèvements ((UD.an) <sup>-1</sup> ) <sup>b</sup>	4	6	8	15	50
Nombre d'UD <sup>a</sup>	8 228	7 487	5 636	3 839	841
Nombre de prélèvements (an <sup>-1</sup> ) <sup>a</sup>	32 912	44 922	45 088	57 585	42 050
Nombre de "pollutions" (an <sup>-1</sup> )	950	502	155	87	15
Taux de "pollutions"	0,0289	0,0112	0,0034	0,0015	0,0004
Probabilité d'au moins une "pollution" (an <sup>-1</sup> )	>0,99	0,98	0,72	0,43	0,12

<sup>a</sup> Ministère de la Santé et des Solidarités, 2006.

<sup>b</sup> Estimations basées sur l'article R.1321-15 du code de la santé publique et l'arrêté du 11 janvier 2007 relatif au programme de prélèvements et d'analyses du contrôle sanitaire pour les eaux fournies par un réseau de distribution.

<sup>c</sup> Pour mémoire. Au-delà de 50 000, le taux de pollution se stabilise autour de 0,0003.

	UD < 500 (classe 1+2)	500 ≤ UD < 2 000 (classe 3)	UD ≥ 2 000 <sup>c</sup>	Total <sup>c</sup>
Nombre de prélèvements <sup>a</sup>	70 000	53 000	211 000	334 000
Taux de non-conformité (NC) <sup>a, b</sup>	14,16 %	3,58 %	0,82 %	4,07 %
Taux de "pollutions" parmi les NC	26,19 %	14,84 %	11,67 %	22,77 %

Source : Sise-Eaux, Ddass, Drass, ministère chargé de la Santé.

<sup>a</sup> Ministère de la Santé et des Solidarités, 2006.

<sup>b</sup> NC : un prélèvement est non-conforme si au moins 1 E. coli ou 1 entérocoque a été détecté. Estimations basées sur le rapport du ministère de la Santé et des Solidarités, 2006.

<sup>c</sup> Pour mémoire.

### 2.2 DISTRIBUTION STATISTIQUE ET GÉOGRAPHIQUE DES VARIABLES DU MODÈLE

Les covariables testées et les covariables retenues dans le modèle linéaire sont fournies en tableau 6. Les covariables relatives au relief et au mode de gestion de l'eau sont éliminées, ainsi que la proportion de terres labourées. L'absence de corrélation entre Karst et la densité

du réseau hydrographique de surface (DRHS) dément la réputation d'équivalence des deux covariables. Karst, qui assure de meilleures corrélations avec *N\_Pols*, a été préféré.

La distribution statistique des variables retenues est donnée en tableau 5 et leur distribution géographique est cartographiée en figures 4 à 8.

<sup>5</sup> Les modalités d'extraction des données de la base Sise-Eaux ne comprenaient pas le décompte des prélèvements et des non conformités.

TABLEAU 5

DISTRIBUTION DU NOMBRE D'UNITÉS DE DISTRIBUTION PAR DÉPARTEMENT ET DES VARIABLES RETENUES

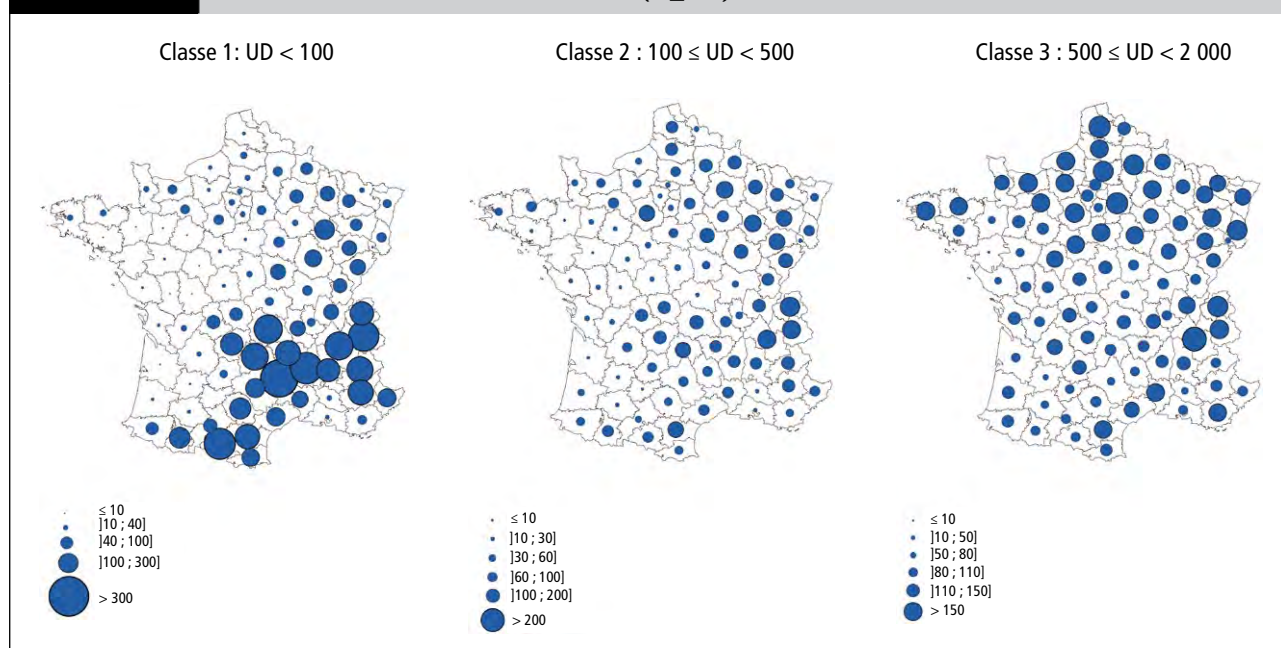
Variable	N	Min	P10	P25	P50	P75	P90	Max	Moyenne	Écart-type
<i>N_UD1</i>	90	0	1	6	36	113	262	588	86,74	118,10
<i>N_UD2</i>	90	0	13	22	68	117	183	286	80,44	65,61
<i>N_UD3</i>	90	2	22	32	53	94	106	138	61,19	35,64
<i>N_Pol1</i>	90	0	0	0	1	25	59	141	17,92	31,30
<i>N_Pol2</i>	90	0	0	0	3	11	29	57	8,84	13,11
<i>N_Pol3</i>	90	0	0	0	1	5	8	14	2,90	3,71
<i>N_Pols1</i>	87	0,00	0,00	0,00	4,34	18,44	29,29	53,25	10,10	13,08
<i>N_Pols2</i>	89	0,00	0,00	0,00	4,47	9,46	21,45	29,91	6,61	7,65
<i>N_Pols3</i>	90	0,00	0,00	0,00	1,53	4,37	8,83	20,62	2,91	3,86
<i>sDes1</i>	87	0,00	0,00	0,00	0,24	0,57	0,80	1,00	0,321	0,310
<i>sDes2</i>	89	0,00	0,00	0,00	0,13	0,38	0,54	0,81	0,207	0,218
<i>sDes3</i>	90	0,00	0,00	0,00	0,06	0,15	0,31	0,74	0,110	0,144
<i>sDUP1</i>	87	0,00	0,33	0,48	0,70	0,87	1,00	1,00	0,648	0,267
<i>sDUP2</i>	89	0,00	0,20	0,34	0,57	0,75	0,90	1,00	0,546	0,259
<i>sDUP3</i>	90	0,06	0,15	0,25	0,41	0,59	0,74	0,89	0,436	0,218
<i>Karst</i>	90	0,00	0,00	0,06	0,38	0,61	0,83	0,99	0,370	0,310
<i>Prec</i>	90	537	633	698	790	918	1 093	1 288	823	168

### 2.2.1 Nombre d'unités de distribution (*N\_UD*)

Le nombre d'UD varie beaucoup en fonction du département (figure 4). Les petites unités (classe 1 et 2) se concentrent dans les départements à dominance montagnaise ou rurale. Les UD de classe 1 sont particulièrement nombreuses au sud d'une ligne joignant la Savoie au Pays Basque. La représentation de la classe 2 reprend

celle de la classe 1 avec un déplacement des zones denses vers le nord. En classe 3, la répartition tend à s'uniformiser sur le territoire. Dans un vaste secteur englobant le Centre, l'Ouest et le Sud-Ouest (à l'exception des secteurs vallonnés), les très petites UD (classe 1 et dans une moindre mesure classe 2) sont en nombre insuffisant (<10; figure 4) pour disposer de statistiques suffisamment précises. *N\_Pols* et variables rapportées au nombre d'UD n'y sont pas cartographiées.

FIGURE 4

NOMBRE D'UNITÉS DE DISTRIBUTION (*N\_UD*)

Source : Sise-Eaux, Ddass, Drass, ministère chargé de la Santé.

## 2.2.2 Nombre de pollutions observées ( $N_{Pol}$ )

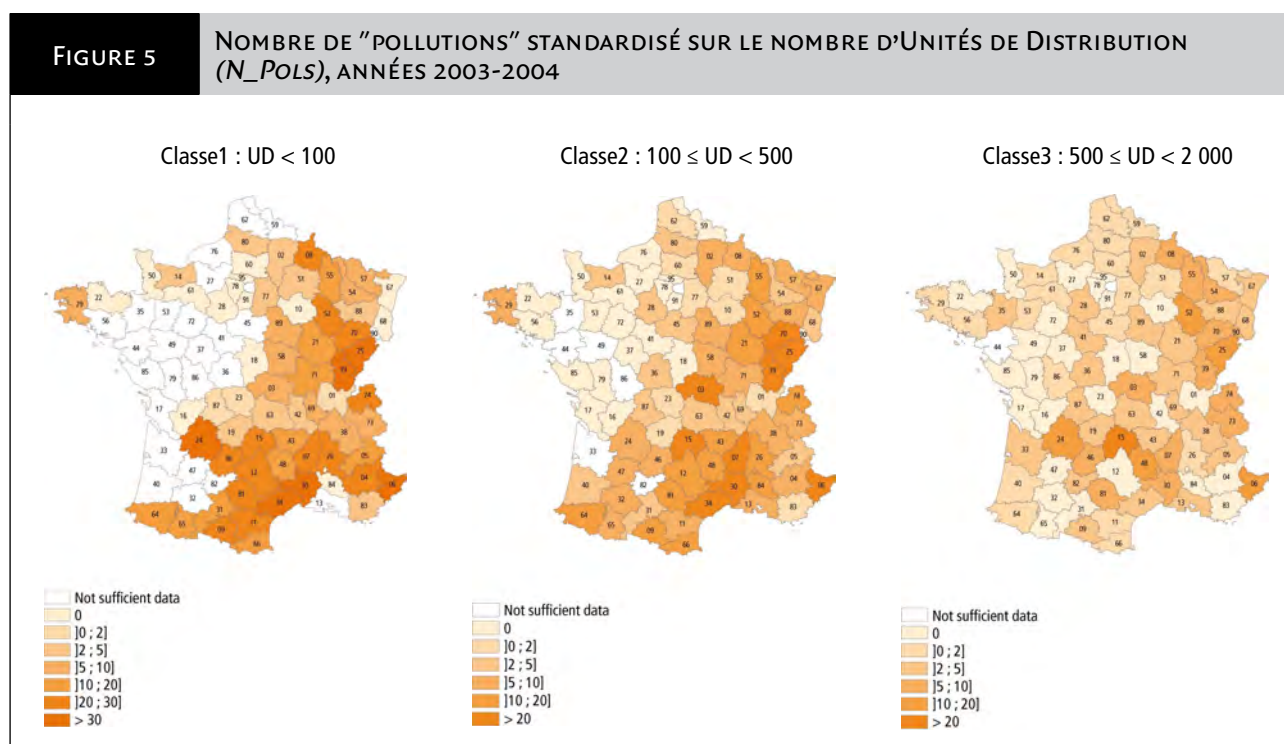
Soixante-seize départements sur les 90 étudiés (84%) ont éprouvé au moins une "pollution" durant la période d'étude, 53 (59%) si on ne prend en considération que la classe 1, 64 (71%) la classe 2 et 62 (69%) la classe 3. Le nombre médian de "pollutions" s'établit à 1, 3 et 1 pour les classes de taille 1, 2 et 3, respectivement. Les distributions sont étirées à droite avec des moyennes beaucoup plus élevées que les médianes (17,9; 8,8 et 2,9 respectivement) et des valeurs maximales élevées atteignant respectivement : 141 ; 57 et 14. La standardisation atténue sensiblement la dissymétrie de la distribution en classe 1 (moyennes de  $N_{Pols}$  : 10,1; 6,6 et 2,9 respectivement).

## 2.2.3 Nombre de pollutions observées standardisé sur le nombre d'unités de distribution ( $N_{Pols}$ )

Le nombre de "pollutions" standardisé est plus faible dans la moitié nord-ouest quand on sépare la France selon une diagonale allant des Ardennes au Pays Basque (figure 5). La situation reste cependant hétérogène au sein des deux ensembles.

La variance de  $N_{Pols1}$  est beaucoup plus importante que celle de  $N_{Pols2}$  et  $N_{Pols3}$ , principalement parce que la moyenne est plus élevée que pour les autres classes, mais aussi parce que la surdispersion est plus forte (variance rapportée à la moyenne  $s^2/m = 17$  pour la classe 1 vs. 9 et 5 pour les classes 2 et 3, respectivement).

Avec des corrélations entre les variables  $N_{Pols}$  s'établissant entre 0,67 et 0,74; l'effet "département" représente environ la moitié de la variance expliquée.



Source : Sise-Eaux, Ddass, Drass, ministère chargé de la Santé.

## 2.2.4 Unités de distribution sans désinfection ( $sDes$ )

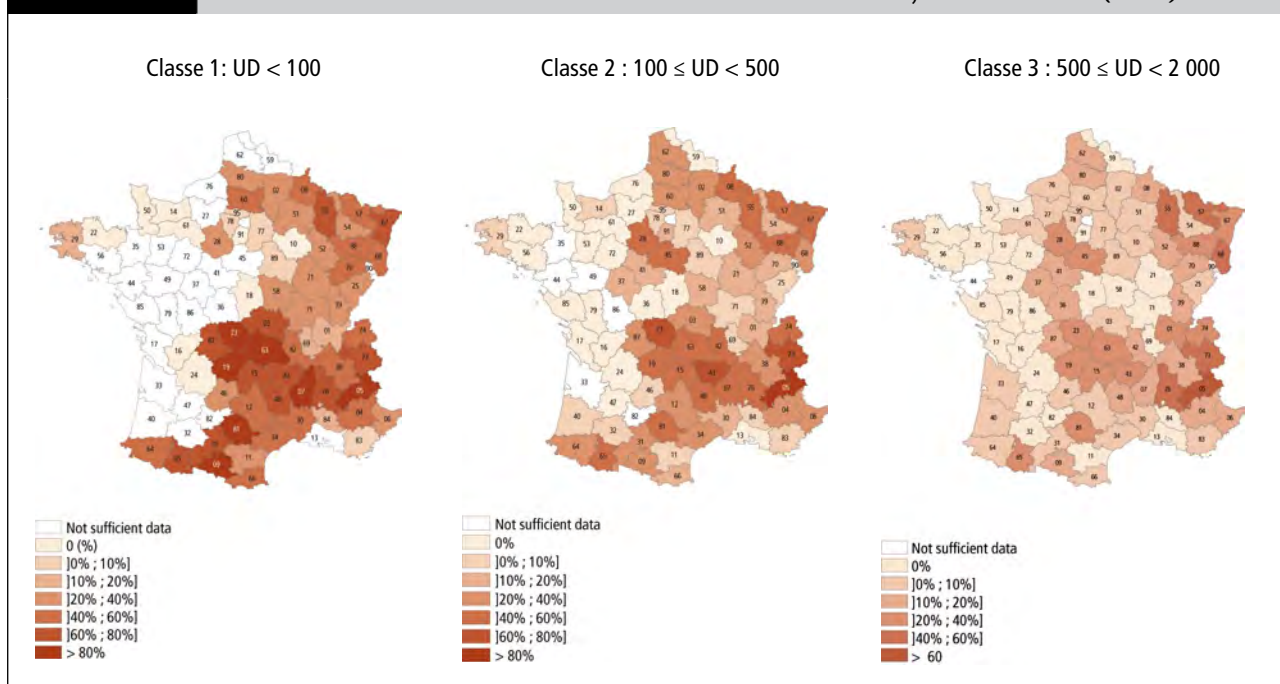
La proportion d'absence de désinfection avant distribution décroît quand la taille des UD augmente : la médiane passe de 22 % pour la classe 1 à 6 % pour la classe 3. Là encore, il existe une grande disparité entre départements (par exemple 0 à 100 % en classe 1). Le taux de données manquantes de  $sDes$  est 2 %.

L'absence de désinfection est plus importante à l'est qu'à l'ouest (figure 6). Cette tendance est la plus marquée en classe 3, notamment en Alsace et dans la moitié nord des Alpes. En ce qui concerne les classes 1 et 2, elle touche les départements de montagne et certains départements ruraux.

Les corrélations entre les variables  $sDes$  s'établissent entre 0,71 et 0,88. Comme pour  $N_{Pols}$ , la force de ces corrélations reflète la cohérence interclasses du niveau d'attention que porte l'interprofession départementale sur la généralisation de la désinfection.

FIGURE 6

PROPORTION D'UNITÉS DE DISTRIBUTION SANS DÉSINFECTION, NI TRAITEMENT (*sDes*)



Source : Sise-Eaux, Ddass, Drass, ministère chargé de la Santé.

### 2.2.5 Captages sans périmètre de protection (*sDUP*)

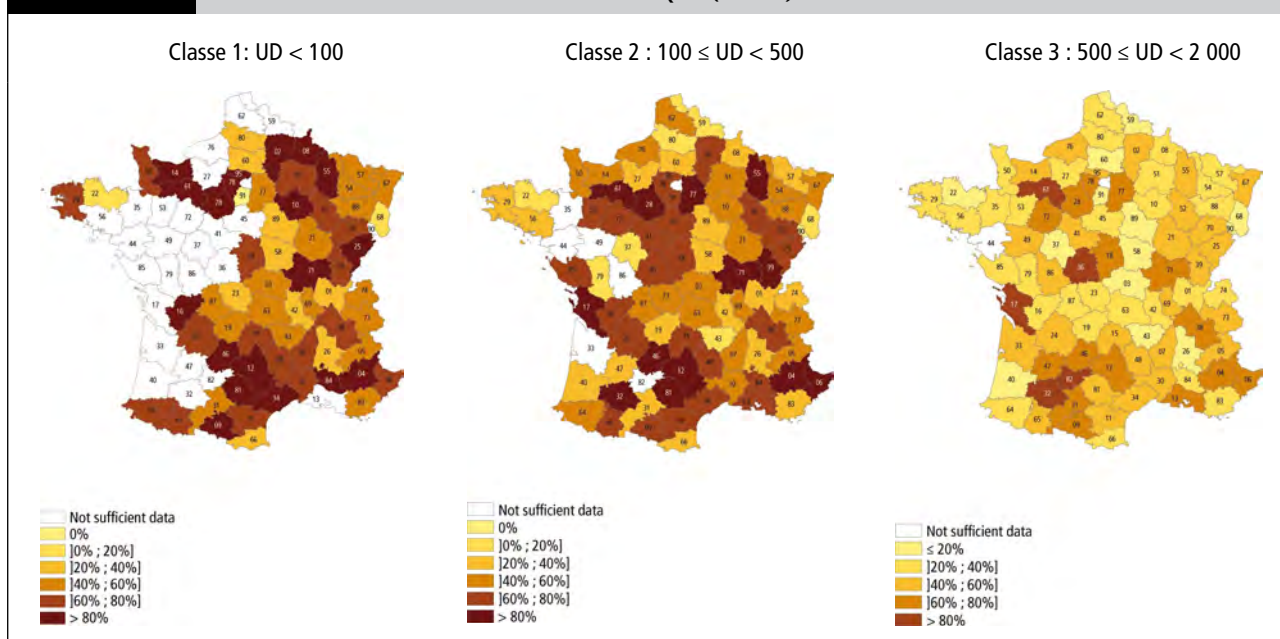
Le taux moyen d'absence de protection réglementaire des captages décroît de 70 % à 41 % entre la classe 1 et la classe 3. Le taux de données manquantes de *sDUP* est 5%. La dispersion intraclasse de la covariable est forte avec des variations de 0 à 100 % pour les classes 1 et 2. Contrairement à la désinfection, il n'apparaît pas

de gradient spatial. Quelques régions semblent rassembler des départements à problème : Basse-Normandie, Région parisienne, Région Centre et sud du Massif Central (figure 7).

Les corrélations entre les variables *sDUP* s'établissent entre 0,55 et 0,73. Là encore, ces niveaux reflètent la cohérence du niveau de prise en charge départementale de la protection des ressources parmi les trois classes de taille d'UD.

FIGURE 7

PROPORTION DE CAPTAGES SANS PÉRIMÈTRE DE PROTECTION OFFICIAISÉ PAR UNE DÉCLARATION D'UTILITÉ PUBLIQUE (*sDUP*)



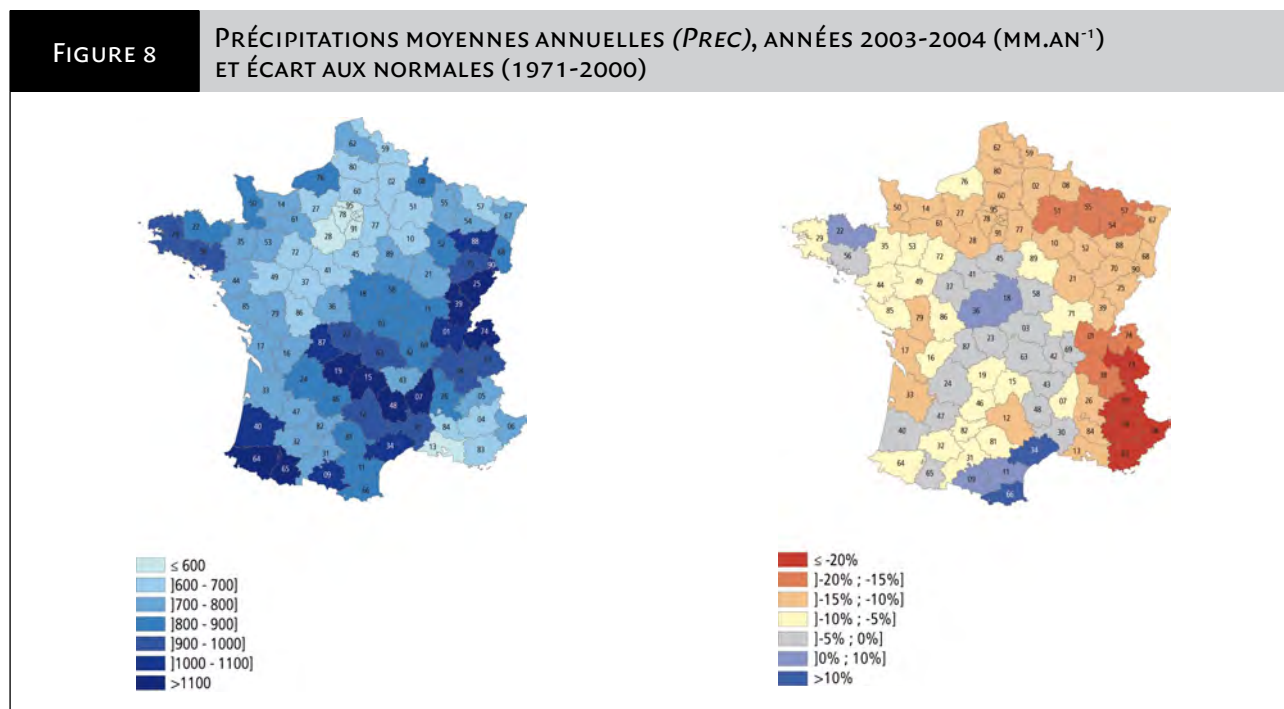
Source : Sise-Eaux, Ddass, Drass, ministère chargé de la Santé.

## 2.2.6 Précipitations (*Prec*)

Les précipitations moyennes sur la période d'étude se sont élevées à 780 mm/an. Elles varient entre 537 mm dans les Bouches-du-Rhône et 1 288 dans les Pyrénées-Atlantiques (figure 8). Elles ont été déficitaires durant la période d'étude dans 80 des 90 départements étudiés avec un déficit moyen de -9 % par rapport à la période de référence (1971-2000). Des précipitations excédentaires sont toutefois constatées dans la région située entre les Hautes-Pyrénées et l'Hérault

(maximum : Pyrénées Orientales, +19%). Les déficits les plus marqués (<-13 %) ont été enregistrés en Champagne-Ardennes et surtout dans la moitié est des Alpes avec un maximum dans le Var (-29 %).

L'année 2003 se caractérise par une sécheresse quasi généralisée (-13 %). Seule, la partie côtière du Languedoc-Roussillon a connu un excès très net (+28 %). L'année 2004 est légèrement sèche (-5 %) avec des précipitations excédentaires dans le Centre (+17 %) et un déficit très marqué dans la pointe Sud-Est (-34 %).



## 2.2.7 Surface karstifiée (*Karst*)

Le pourcentage moyen de surface karstifiée est de 37 % avec des variations de 0 % à 99 %. 22 % des départements n'ont aucune surface karstifiée. La distribution est donnée en Encart, figure 2.

occupée par des formations karstiques (*Karst*), et deux indicateurs de facteurs humains, la prévalence des mesures de protection des ressources (*sDUP*) et la prévalence de la désinfection (*sDes*).

## 2.3 MODÈLE LINÉAIRE

Parmi les covariables testées, 4 covariables répondent aux critères d'inclusion et ont été retenues au terme de la modélisation de *N\_Pols*. Elles comprennent deux indicateurs des facteurs naturels du risque de "pollution", les précipitations totales (*Prec*) et la proportion de la surface

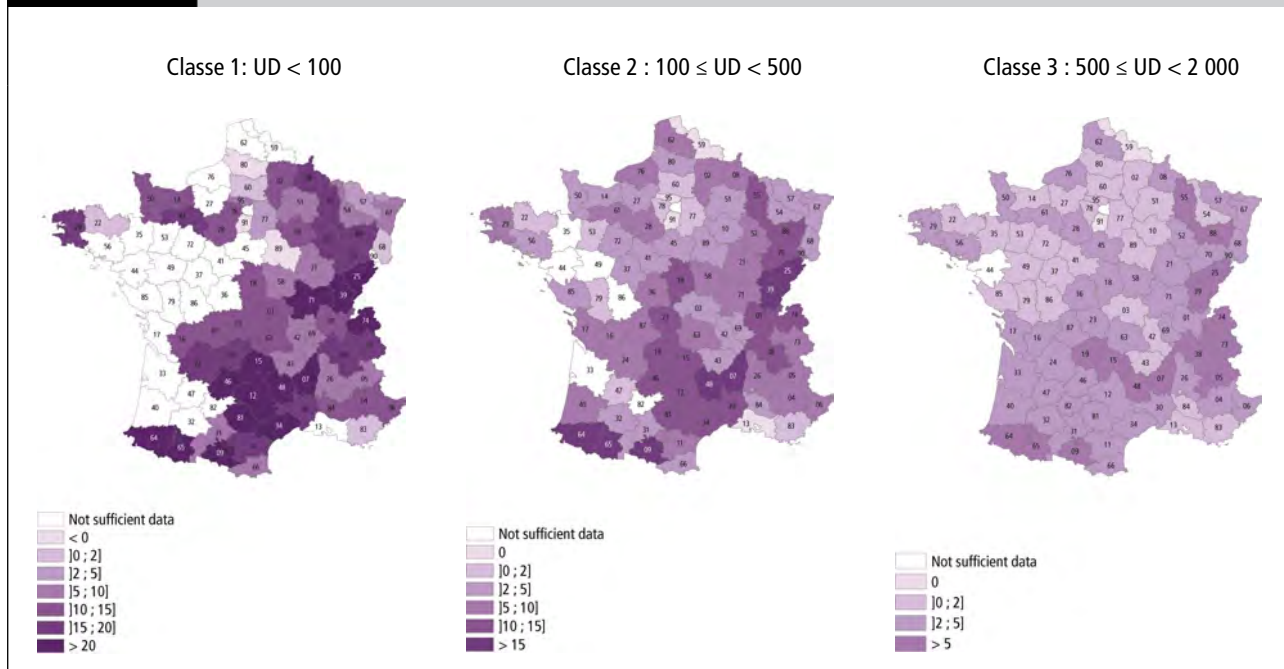
En supposant que la distribution des résidus soit Normale (Cf. *infra*), les coefficients associés à *Prec* sont très significatifs, ceux associés à *Karst* et *sDUP* significatifs et celui associé à *sDes* peu significatifs (tableau 6). Les droites de régression sont données en figure 10 et les valeurs prévues par le modèle pour les années 2003-2004 sont cartographiées en figure 9. Il s'agit de la somme des covariables pondérées par les coefficients du modèle. La prépondérance des précipitations totales transparaît dans cette représentation.

Covariable	Classe 1		Classe 2		Classe 3	
		IC 95%		IC 95%		IC 95%
<i>sDes</i>	7,7*	-0,481 15,9	5,26*	-0,814 11,3	4,29**	0,885 7,7
<i>sDUP</i>	17,5***	9,3 25,7	6,83**	1,98 11,7	3,09**	0,572 5,6
<i>Karst</i>	16,4***	8,06 24,8	5,35**	1,19 9,52	1,78*	-0,125 3,69
<i>Prec</i>	0,0338***	0,0161 0,0513	0,0247***	0,0158 0,0336	0,0082***	0,0039 0,0125

\*  $p < 0,1$ ; \*\*  $p < 0,02$ ; \*\*\*  $p < 0,001$ .

FIGURE 9

NOMBRE DE "POLLUTIONS" STANDARDISÉ SUR LE NOMBRE D'UNITÉS DE DISTRIBUTION ( $N\_POLs$ ), ANNÉES 2003-2004, PRÉVISION DU MODÈLE



Source : Sise-Eaux, Ddass, Drass, ministère chargé de la Santé.  
Le modèle utilisé ici n'inclut pas les indicatrices des départements.

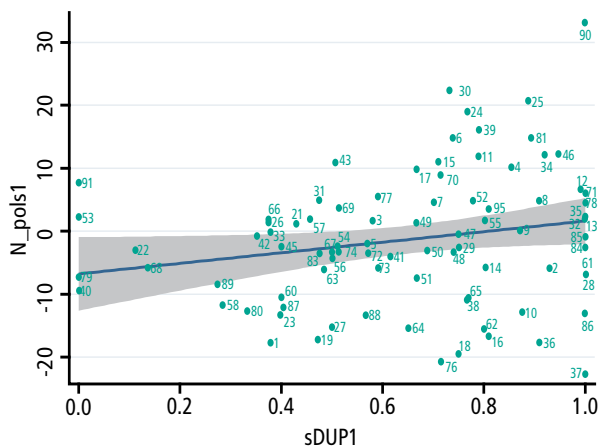


FIGURE 10

RÉGRESSION DE  $N\_POLS$  (CENTRÉE) SUR LES COVARIABLES RETENUES DANS LE MODÈLE APRÈS CONTRÔLE DES AUTRES COVARIABLES

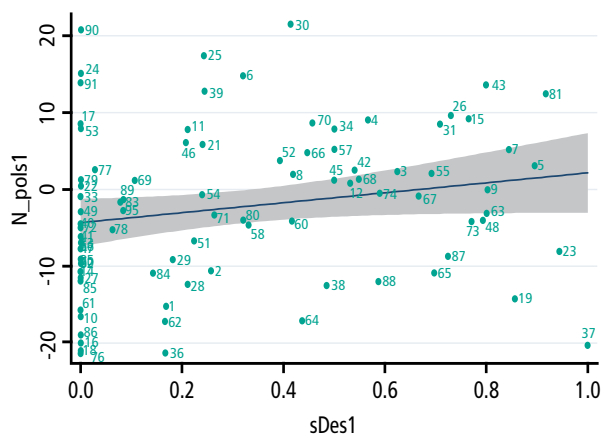
Absences de périmètre de protection officialisé par une DUP

Après contrôle des autres covariables  
UD < 100 habitants

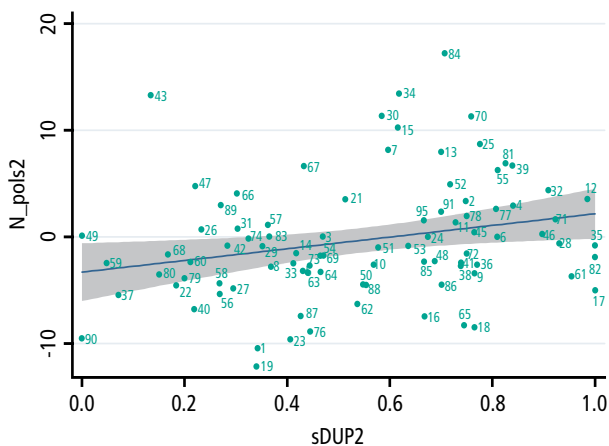


Absence de désinfection

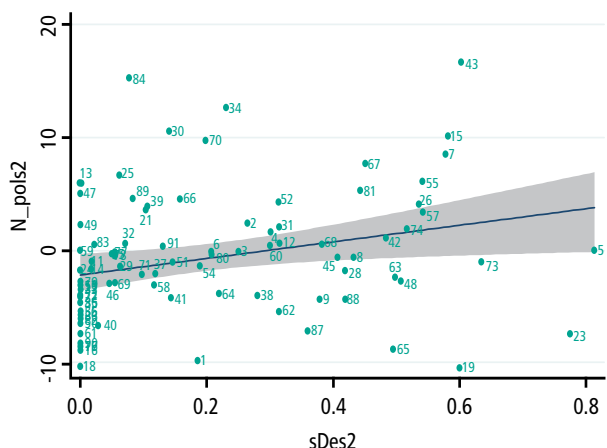
Après contrôle des autres covariables  
UD < 100 habitants



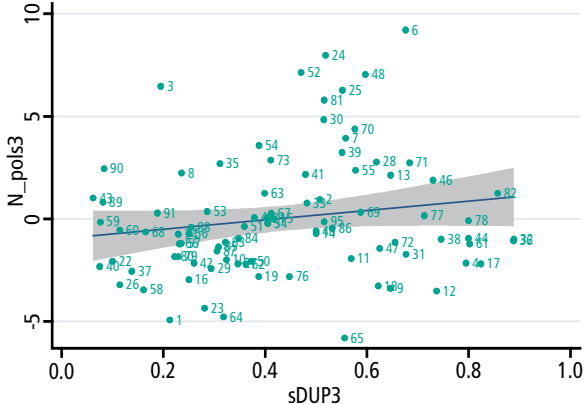
Après contrôle des autres covariables  
100 ≤ UD < 500 habitants



Après contrôle des autres covariables  
100 ≤ UD < 500 habitants



Après contrôle des autres covariables  
500 ≤ UD < 2000 habitants



Après contrôle des autres covariables  
500 ≤ UD < 2000 habitants

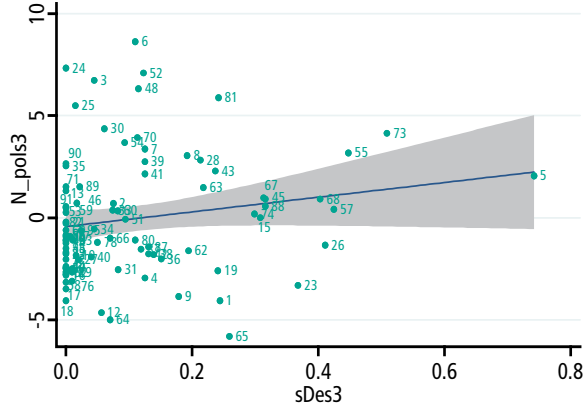
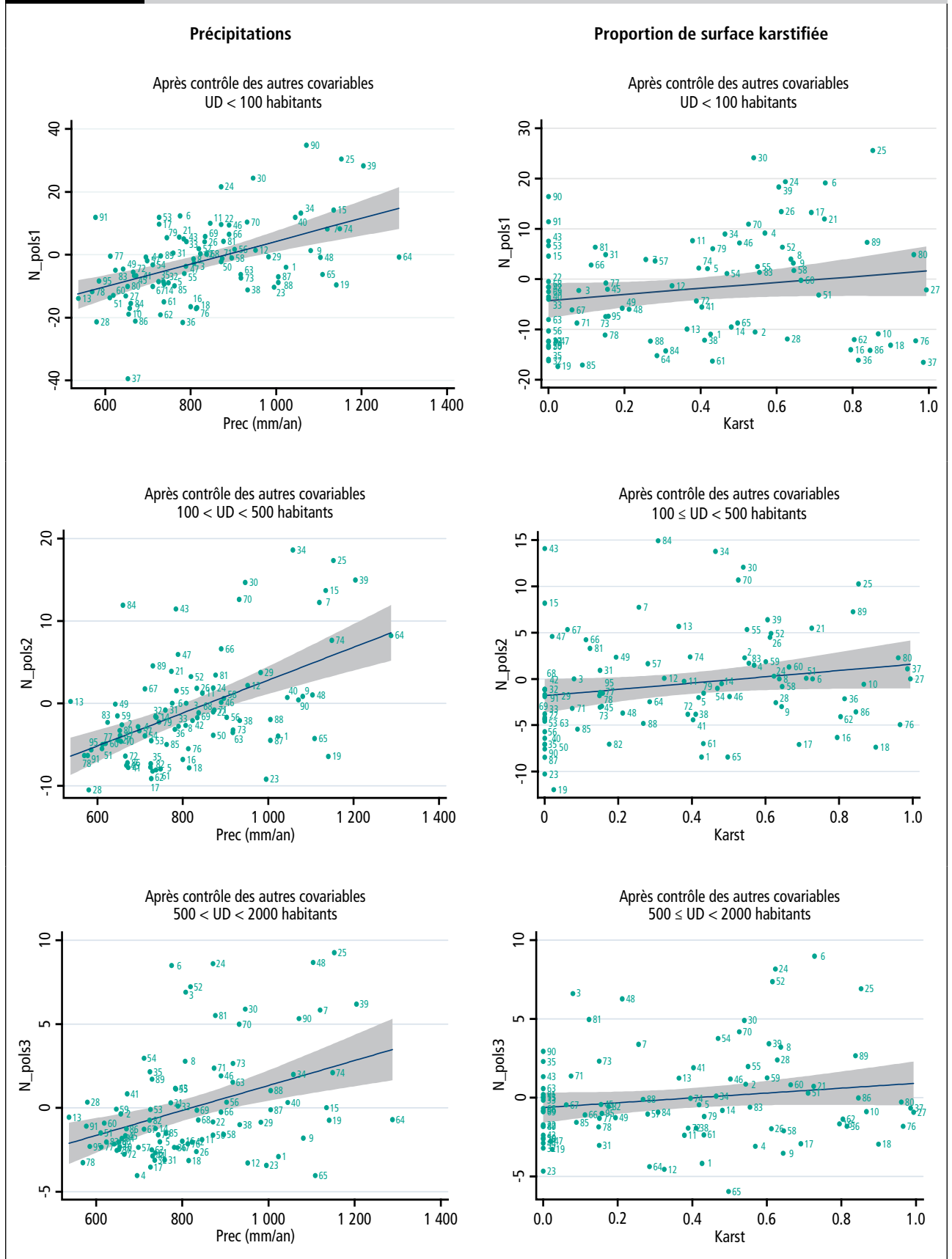


FIGURE 10

RÉGRESSION DE  $N\_POLS$  (CENTRÉE) SUR LES COVARIABLES RETENUES DANS LE MODÈLE APRÈS CONTRÔLE DES AUTRES COVARIABLES (SUITE)



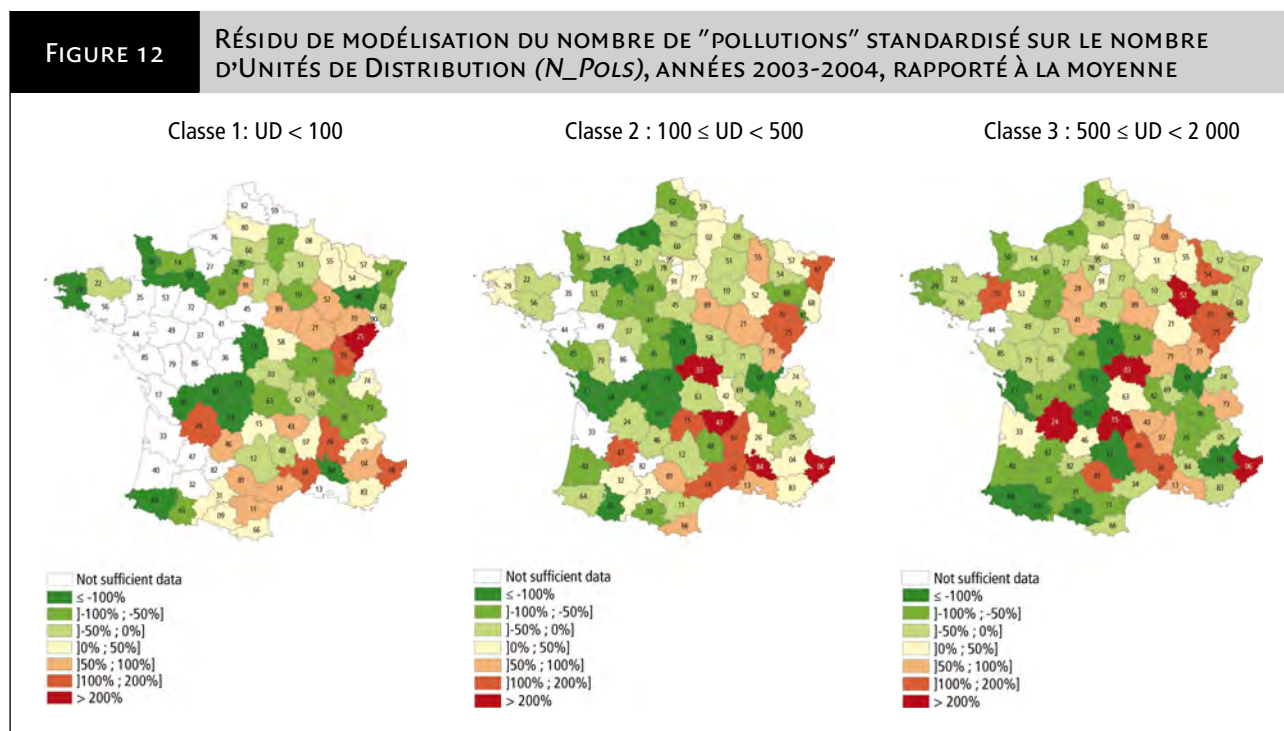
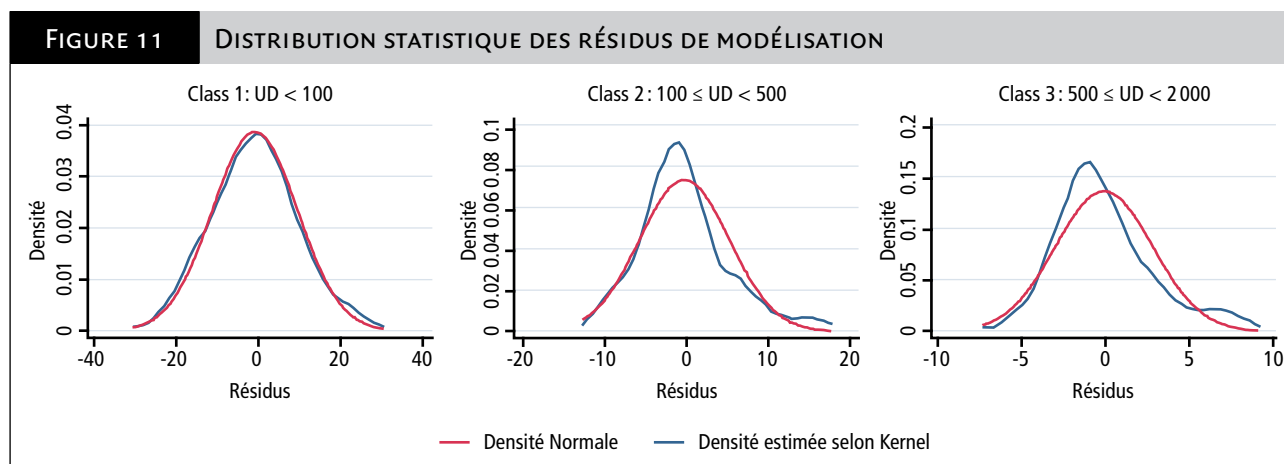
Des covariables supplémentaires indicatrices des départements mal ajustés ("outliers" situés à plus de trois écarts-types de la valeur attendue) ont dû être introduites pour un ajustement plus correct des modèles : pour la classe 3, le Cantal subissant 16 "pollutions" en

excès par rapport à la valeur attendue par modélisation, soit plus de 5 fois la moyenne de  $N\_Pols3$ ; pour la classe 2, les Alpes-Maitimes (06) et l'Allier (03) qui présentent un excès de risque de 21 et de 19 "pollutions", soit environ 3 fois la moyenne de  $N\_Pols2$ .

Les résultats (valeurs des coefficients et significativité) ne sont pas sensiblement altérés par l'application ou non des différentes options du calcul d'ajustement du modèle : avec ou sans les indicatrices des départements, avec ou sans l'option "robust" de gestion de l'hétéroscédasticité, et avec ou sans la pondération des observations (données non présentées). Ces résultats argumentent en faveur de la robustesse du modèle.

La distribution du résidu de modélisation de  $N\_PolS1$  ne diffère pas d'une gaussienne (test de Kolmogorov-Smirnov non significatif). Ce n'est pas le cas de  $N\_PolS2$  ( $p < 0,05$ ) et  $N\_PolS3$  ( $p < 0,01$ ) (figure 11). L'introduction de variables indicatrices des départements

diminue l'écart à la normale. La réduction de la variance résiduelle qu'elle procure provoque, par ailleurs, l'émergence de nouvelles observations excentrées (distributions bimodales de la figure 11). Il est possible de normaliser les résidus des modèles (c'est-à-dire obtenir  $p > 0,05$ ) pour les classes 2 et 3 en introduisant parmi les covariables, d'une part, des indicatrices supplémentaires de département excentrés, et d'autre part, le résidu du modèle 1 ( $p < 10^{-4}$ ), qui incarne l'effet "département" résiduel estimé sur la classe 1 (données non présentées). Ces modalités renforcées du contrôle de l'effet département conduit à une diminution systématique des probabilités associées aux coefficients, ce qui corrobore la réalité de l'effet de *Prec*, *Karst*, *sDUP* et *sDes*.



Le modèle utilisé ici n'inclut pas les indicatrices des départements.

En dépit de la perte de puissance statistique qu'entraîne le recours à un test de corrélation non paramétrique (test des rangs de Spearman, tableau 7), ce test confirme le lien entre  $N\_PolS$  et *Prec*, à un moindre degré entre  $N\_PolS$  et *sDUP*, et corrobore la réalité du lien entre  $N\_PolS$  et *sDes*, qui ressortait modérément de l'approche paramétrique. Une corrélation non significative apparaît entre  $N\_PolS$  et *Karst* dans les 3 classes de taille.

En résumé, la covariable *sDes* est très significative selon l'approche non paramétrique; *Karst* et *sDUP* sont très significatives selon l'approche paramétrique pour la classe 1, seule classe pour laquelle le résidu de modélisation est normal; enfin, *Prec* est très significative quelles que soient l'approche et la classe de taille.

TABLEAU 7	PROBABILITÉS ASSOCIÉES AUX COEFFICIENTS DES RANGS DE SPEARMAN		
	<i>N_Pols1</i>	<i>N_Pols2</i>	<i>N_Pols3</i>
<i>sDUP</i>	0,0364	0,0049	0,0786
<i>sDes</i>	0,0148	0,0005	0,0337
<i>Prec</i>	0,0000	0,0000	0,0003
<i>Karst</i>	0,1451	0,2556	0,1603

*N\_Pols* est préalablement ajustée sur les trois covariables non testées.

## 2.4 ANALYSE DE LA VARIANCE

La part de variance de *N\_Pols* expliquée par les covariables *Prec*, *Karst*, *sDUP* et *sDes* diminue quand la taille des UD augmente (tableau 8). La part de variance expliquée par les variables naturelles est plus forte que celle expliquée par les variables humaines.

TABLEAU 8	CONTRIBUTION DES COVARIABLES À LA VARIANCE DU NOMBRE STANDARDISÉ DE "POLLUTIONS" ( <i>N_POLS</i> )			
		Classe 1	Classe 2	Classe 3
Variance de <i>N_Pols</i>		171	59	15
Départements "outliers"			06, 03	15
<b>Part de variance expliquée</b>				
Par les covariables "naturelles" (2 covariables: <i>Prec</i> , <i>Karst</i> )		37 %	32 %	19 %
Par les covariables "humaines" (2 covariables: <i>sDUP</i> , <i>sDes</i> )		23 %	12 %	9 %
<b>Par le sous-ensemble</b> (4 covariables)		<b>47 %</b>	<b>39 %</b>	<b>26 %</b>
Par les départements "outliers" (0 à 2 covariables)		0 %	12 %	23 %
<b>Par l'ensemble</b> (4 à 6 covariables)		<b>47 %</b>	<b>50 %</b>	<b>42 %</b>

## 2.5 INTERACTIONS

Des interactions entre les facteurs de risque de "pollution" ont été recherchées grâce à des tests de variables construites comme le produit des variables primaires, telles que *Prec x Karst* qui représente la synergie entre la présence de Karst et les écoulements. Aucune interaction stable n'a pu être mise en évidence. En ce qui concerne la classe 1, le modèle  $N\_Pols1 = a.DUP1 \times Prec + b.sDes1 + c.Karst + d$  explique la même proportion de variance que le modèle présenté.

*N\_Pols* (tableau 9). La grande différence de variance de *N\_Pols* entre les classes de taille (tableau 8) ne permet cependant pas de proposer un ajustement des coefficients sur l'ensemble des données.

On a simulé l'effet sur le nombre de "pollutions" de différents scénarios d'augmentation des covariables: +10% sur *Karst*, *sDUP* et *sDes*, et un excédent de précipitation de 100 mm, qui correspond à une variation relative de +12 % par rapport à la lame d'eau précipitée moyenne. *N\_Pols* répond aux variations relatives de *Karst*, *sDUP* et *sDes* par des variations relatives de même niveau. *N\_Pols* est beaucoup plus sensible à des pluies excédentaires avec des variations relative d'environ 30 % pour un excès de 10 % sur la moyenne annuelle.

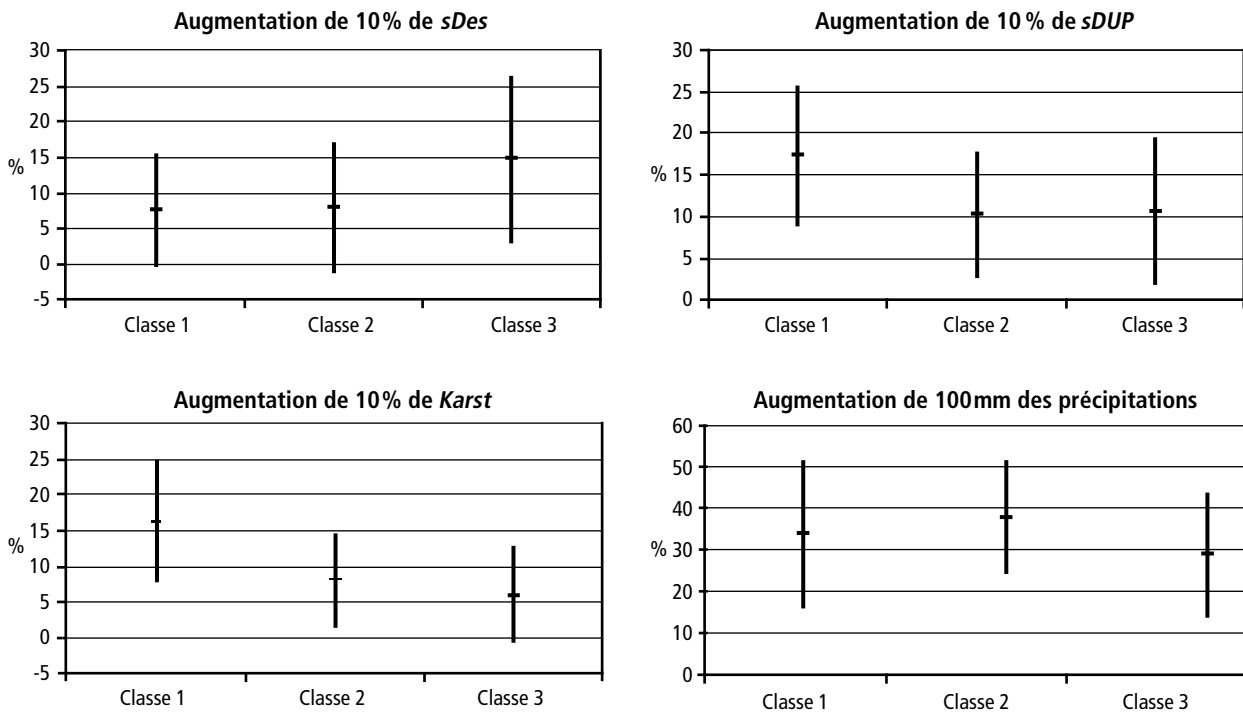
## 2.6 VERS UN MODÈLE UNIQUE

Les coefficients des modèles ne diffèrent pas significativement selon la classe de taille, si on les rapporte à la moyenne de la variable à expliquer

TABLEAU 9	COEFFICIENTS DES MODÈLES RAPPORTÉS À LA MOYENNE DE <i>N_POLS</i> (A)								
	Covariable	Classe 1			Classe 2			Classe 3	
			IC 95%			IC 95%			IC 95%
<i>sDes</i>	<b>0,76</b>	-0,05	1,57	<b>0,80</b>	-0,12	1,71	<b>1,47</b>	0,30	2,65
<i>sDUP</i>	<b>1,73</b>	0,92	2,54	<b>1,03</b>	0,30	1,77	<b>1,06</b>	0,20	1,92
<i>Karst</i>	<b>1,62</b>	0,80	2,46	<b>0,81</b>	0,18	1,44	<b>0,61</b>	-0,04	1,27
<i>Prec</i>	<b>0,0033</b>	0,0016	0,0051	<b>0,0037</b>	0,0024	0,0051	<b>0,0028</b>	0,0013	0,0043

FIGURE 13

SIMULATION DES MODÈLES : AUGMENTATION RELATIVE DU NOMBRE DE "POLLUTIONS"  
POUR DIFFÉRENTS SCÉNARIOS ET INTERVALLES DE CONFIANCE À 95 %



## 3. Discussion

La discussion porte successivement sur l'enjeu de santé publique, sur l'interprétation des corrélations écologiques observées et enfin, sur le prolongement d'étude qui permettrait de corroborer un rapport de causalité.

### 3.1 SANTÉ PUBLIQUE : POPULATIONS RÉSIDENTES ET POPULATIONS DE PASSAGE

Dans tous les pays développés, la qualité microbiologique de l'eau des très petites UD pose problème. Du fait du manque de moyens financiers, techniques et humains, la présence d'une ressource vulnérable entraîne mécaniquement des contaminations au robinet.

En France, la fraction exposée de la population résidente reste relativement faible (tableau 3). L'autorité sanitaire dispose de différents moyens pour protéger cette population. La collectivité responsable ou le délégataire sont tenus d'informer chaque année la population résidente de la situation locale par un courrier joint à la facture d'eau. D'autre part, l'autorité sanitaire interdit la consommation alimentaire de l'eau pendant les périodes de contamination identifiées ; le seuil critique de contamination est adapté à la situation moyenne du département et dépasse généralement la valeur limite de non-conformité. L'expérience montre que ces interdictions sont bien suivies durant les épidémies (Di Palma *et al.* 2003, Tillaut *et al.* 2004). Il n'existe à notre connaissance pas de données de compliance pour des interdictions promulguées hors contexte épidémique (non-conformité microbiologique).

Les populations de passage ne bénéficient pas ou peu de cette information. La France se distingue pourtant parmi les autres pays développés par le niveau élevé de sa fréquentation touristique : 1,3 milliard de nuités par an dont la moitié en secteur rural (32 %) ou montagnoux (19 %). 38 % des nuités sont achetées par des Européens non Français et des étrangers (données 2006, ministère de l'Économie, des Finances et de l'Emploi, 2007). On peut suspecter que cette population de passage soit non seulement peut informée du niveau du risque microbiologique local, mais encore, qu'elle perçoive avec une confiance excessive la qualité de l'eau collectée en secteur rural et surtout montagnoux. Sans disposer de données directes sur le sujet, il est à noter que la carte de non-consommation de l'eau du robinet (proportion d'abstinents déclarés (Beaudeau *et al.*, 2003)) possède plus de correspondances avec celle des pollutions des eaux naturelles engendrées par l'agriculture intensive (pesticides, nitrates...)

qu'avec celle de la contamination microbiologique. Il existe donc un faisceau d'arguments désignant les touristes comme une population à risque : connaissance insuffisante du risque et des situations locales, perception paradoxale du risque, faible niveau d'immunisation acquise vis-à-vis des agents pathogènes d'origine entérique, comparativement aux populations résidentes plus régulièrement exposées. Compte tenu de l'enjeu sanitaire et économique, il est important pour la France d'explorer les possibilités offertes par des systèmes d'information tels que celui développé par l'Assurance maladie pour la liquidation des soins de santé, car ces systèmes permettent de mesurer l'impact des pollutions microbiologiques des eaux distribuées sur les populations de passage, voire de détecter des épidémies (Beaudeau *et al.*, 2006). Il s'agit aussi d'acquérir des données locales de consommation alimentaire de l'eau de robinet dans des circonstances particulières (épidémies, avis de restriction de la consommation alimentaire, non-conformité passagère ou chronique, d'ordre chimique ou microbiologique), sur le moment et dans le temps, parmi la population résidente et la population touristique (Français, étrangers).

### 3.2 CORRÉLATIONS ÉCOLOGIQUES ET CAUSALITÉS

L'étude présentée ici est une étude "écologique" au sens de l'épidémiologie, car l'unité de base n'est pas l'individu, en l'occurrence l'UD, mais un groupe d'individus, en l'occurrence le département.

Le choix du département comme unité statistique se justifie non seulement par l'organisation pratique de la collecte de données mais aussi par le caractère départemental de la gestion de l'eau (Ddass, DDAF et Conseil général). La gestion privée est organisée à l'échelle de grandes régions et non du département, mais elle reste minoritaire dans les catégories de taille d'UD étudiées. Par exemple, en classe 1 et 2, 20 % seulement des UD sont en gestion déléguée (ministère de la Santé et des Solidarités, 2006).

Les études écologiques géographiques aident à la formulation ou à la consolidation d'hypothèses mais les corrélations écologiques seules sont généralement insuffisantes pour mettre en évidence un rapport de causalité, notamment à cause des biais écologiques et des effets de confusion potentiels. Les résultats présentés ici s'appuient cependant non seulement sur l'existence de corrélations significatives, mais aussi sur leur reproductibilité entre trois classes de taille d'UD et sur leur vraisemblance au regard de l'expertise (tableau 10).

**TABEAU 10 ARGUMENTAIRE POUR L'INTERPRÉTATION DES CORRÉLATIONS EN TERME DE CAUSALITÉ**

	Limites	Commentaire
Corrélations statistiques significatives	Distributions des résidus non gaussiennes pour les classes 2 et 3 Autocorrélation géographique non prise en compte	Normalisable par un contrôle renforcé de l'effet "département" La taille importante des unités géographique limite les corrélations entre unités adjacentes Pas d'autocorrélation spatiale <i>a priori</i> des facteurs de risque liés à la gestion (organisation administrative départementale) Tests non paramétriques de corrélation OK, sauf pour <i>Karst</i>
Reproductibilité des résultats entre les 3 classes de taille	Tests de reproductibilité limités (autres périodes, autres pays ?) <i>N_Pols*</i> corrélées entre classes de taille; idem pour <i>sDes*</i> et <i>sDUP*</i> : possibilité d'effet de confusion	À faire (autres périodes) La distribution géographique de <i>N_UD</i> diffère sensiblement selon la classe de taille (figure 4): ce ne sont pas les mêmes départements qui contribuent aux trois modèles
Accord avec l'expertise	OK	Base de faits +/- consistante selon la covariable (Cf. texte)

Le protocole écologique tire les résultats des tests vers l'hypothèse nulle: les UD jugées exposées au risque de "pollution" au sein d'un département donné ne sont peut être pas celles qui ont été polluées. Le caractère approximatif des covariables et les erreurs de classement accentuent l'affaiblissement des corrélations. Les données ne permettent pas par exemple de distinguer les pollutions introduites au niveau de la distribution (24 % du total des pollutions de l'eau distribuée

en France (Nedellec *et al.*, 1996), et les pollutions de la ressource, alors que les covariables étudiées concernent exclusivement les pollutions de la ressource. Le tableau 11 fait le bilan des sources de variance et propose une appréciation de leur importance relative. Compte tenu de l'importance des sources irréductibles de variance (mesure, protocole écologique), les facteurs de risque de "pollution" identifiés expliquent une part substantielle de la variance explicable.

**TABEAU 11 SOURCES DE VARIANCE (V) DE *N\_POLS***

Source	Importance	Commentaire
Mesure: prélèvement	+	Peu documenté, V réputée supérieure à V (analyse)
Mesure: analyse	+	Aléa de Poisson sur le comptage des colonies
Erreurs, données manquantes	-	DM absentes ou rares ( <i>sDUP</i> et <i>sDes</i> )
Définition de cas non spécifique des pollutions de la ressource	+	1/4 des "pollutions" seraient des retours d'eau
Protocole écologique	++	V importante vu la grande taille de l'unité géographique
Facteurs de risque identifié	+++	26 % à 47 % de la variance totale
Facteurs de risque non identifiés d'ordre naturel	++	19 % à 37 % de la variance totale
Facteurs de risque non identifiés liés à la gestion	+	9 % à 23 % de la variance totale

Les facteurs naturels et humains expliquent d'autant plus la survenue des "pollutions" que la taille des UD est petite (tableau 8). Cette tendance est à rapprocher de la sophistication des traitements mis en œuvre qui croît avec la taille de l'UD (tableau 1). Ces traitements, en corrigeant les défauts de l'eau brute, estompent en effet le rapport entre les pollutions de réseau et les épisodes de détérioration de la qualité des eaux brutes. Ainsi, coexisteraient en classe 3 des systèmes sans clarification, pour lesquels les facteurs de risque de "pollution" sont inféodés à la qualité de la ressource à l'instar des classes 1 et 2, et des systèmes incluant une clarification qui affranchit plus ou moins la qualité de l'eau traitée de celle de l'eau brute. Dans ce dernier cas, *Prec*, *Karst* et *sDUP*, directement liées à la qualité de l'eau brute, ne peuvent expliquer les "pollutions".

Lors de la modélisation, l'émergence de départements "singuliers", présentant un important excès de risque par rapport à la prévision du modèle, découle aussi vraisemblablement de facteurs locaux, en l'occurrence d'un manque d'attention de certains acteurs. Les données

et les méthodes employées dans ce travail sont inappropriées pour l'interprétation des particularismes locaux.

Les relations mises en évidence entre *N\_Pols* et les quatre covariables sont plus au moins documentées dans la littérature.

Le nombre de "pollutions" *N\_Pols* est très sensible aux précipitations, ce qui n'est pas surprenant car la pollution de la ressource est plus souvent imputable à la remise en circulation des bactéries par les pluies (débordements de systèmes d'assainissement, ruissellements entraînant les déjections animales...) qu'à des pollutions permanentes (rejets de temps sec de systèmes d'assainissement). Le lien entre la survenue d'épidémies et les précipitations a été mis en évidence aux USA (Curriero *et al.* 2001). En France, 3 des 5 épidémies de gastro-entérites dues à une pollution de ressources entre 1998 et 2006 ont été conditionnées par des événements pluvieux (Beaudeau *et al.*, 2008).

La corrélation entre  $N\_Pols$  et la proportion de surface karstifiée  $Karst$  est reproductible entre les classes de taille d'UD. Le karst est une caractéristique hydrogéologique structurelle qui potentialise le pouvoir polluant des ruissellements de surface par l'existence de circulations rapides et une absence de filtration. Des études ont mis en évidence la vulnérabilité microbiologique des eaux karstiques de Haute-Normandie (Dussart-Baptista *et al.*, 2003), mais il n'existe pas à notre connaissance de données épidémiologiques nationales ou internationales spécifiques<sup>6</sup>. Le cas des eaux karstiques est pris en compte dans les législations tant américaines qu'euro-péennes sous la dénomination "eaux souterraines influencées par les eaux de surfaces", qui recouvre aussi d'autres situations hydrogéologiques (nappes alluviales subaffleurantes...). Ces législations, qui imposent aux eaux "influencées" des dispositions de protection et de surveillance concernant à l'origine les eaux de surface, témoignent d'une reconnaissance implicite du risque karstique.

La protection de la ressource et la post-désinfection sont des mesures de santé publique centenaires et largement confirmées par l'usage. Dans la mesure où l'établissement des périmètres de protection procède d'interdiction de dépôt de matières dangereuses (fumiers, lisiers...) susceptibles de s'introduire dans le captage et s'accompagne de moyens efficaces pour y parvenir, comme l'appropriation de terrains, il n'est pas étonnant de constater un effet positif sur la fréquence des pollutions accidentelles. Des mesures de protection de la ressource sont adoptées dans de nombreux pays et constamment encouragées par l'OMS (World Health Organization 2004). En France, le Plan national santé-environnement 2004-2008 compte parmi ses objectifs la généralisation des périmètres de protection en 2010 (ministère de la Santé et de la Protection sociale *et al.* 2004). La post-chloration élimine les bactéries pathogènes et procure à l'eau en distribution un pouvoir désinfectant rémanent (très limité) vis-à-vis de micro-intrusions d'eau contaminée dans le réseau de distribution. La post-chloration, bien qu'adoptée par une majorité de pays, ne bénéficie pas comme l'établissement de périmètres de protection d'un tel consensus (Ainsworth 2004). La législation française ne prévoit pas d'obligation en la matière, mais l'autorité sanitaire a cependant préconisé la généralisation et le renforcement de la post-désinfection après 2001 dans le cadre de la prévention d'actions terroristes ou malveillantes.

Au terme de cette revue des arguments relevant de la statistique et de l'expertise scientifique, il apparaît nécessaire de moduler l'interprétation des corrélations entre  $N\_Pols$  et les covariables en fonction de la covariable. Le rôle de la pluie est à la fois très fort, reproductible et bien identifié par ailleurs : le rapport de causalité s'impose. Le rôle du karst est à confirmer à travers une meilleure spécification des covariables (Cf. Infra). La mise en lumière d'un lien direct entre  $sDUP$  et  $sDES$  représente un enjeu important car cela touche directement au travail du gestionnaire. Dans l'hypothèse d'un lien causal, le gain marginal à étendre les DUP ou la désinfection de l'eau avant distribution s'élève à 1 % de "pollutions" en moins pour 1 % de ressource protégée ou 1 % de désinfection en plus. L'estimation est cependant plus solide pour  $sDUP$  que pour  $sDes$ . La protection de la ressource couvre tous les agents pathogènes tandis que la désinfection ne prémunit que du risque bactérien, et dans une moindre mesure du risque viral. Dans l'hypothèse d'un lien de causalité et sans prendre en considération le coût, il paraît plus important de progresser d'un point sur les périmètres de protection que sur la désinfection. En l'absence d'un lien de causalité,  $sDUP$  n'en demeurerait pas moins le meilleur indicateur des bonnes pratiques de gestion sanitaire de l'eau, parmi les indicateurs testés.

### 3.3 RENFORCER L'INTERPRÉTATION DES CORRÉLATIONS

Parmi les développements envisageables de ce travail, ceux qui peuvent conduire à la confirmation d'un rôle direct de  $sDUP$  et  $sDes$  sur  $N\_Pols$  sont prioritaires. La période 2002-2010 est une période cruciale qui devrait voir une évolution sensible du taux de protection des ressources (Plan national santé environnement) et du taux de désinfection (mesure "Vigipirate"). L'étude de l'évolution conjointe des 3 variables au cours de cette période offre une base expérimentale de premier ordre pour trancher dans un sens ou dans l'autre.

D'autres développements, tels qu'une meilleure spécification des variables et des interactions, seraient utiles à l'affinement des conclusions de cette étude écologique.

**Precipitations.** Dans la mesure où les ruissellements de surface fournissent le gros des flux en bactéries pénétrant les ressources, il serait préférable d'utiliser la lame d'eau ruisselée en surface plutôt que les précipitations totales pour saisir le lien avec les "pollutions". Cette donnée n'est cependant pas disponible au niveau départemental. Les pluies efficaces présentent un intérêt intermédiaire, mais ces données, qui résultent de calculs complexes, ne sont pas disponibles en routine sur l'ensemble du territoire.

Le **karst** est une donnée structurelle qui potentialise le pouvoir polluant des ruissellements de surface par l'absence de filtration (Cf. encart). Les cartes utilisées représentent cependant mal la vulnérabilité des sources aux pollutions microbiologiques. La propension à ruisseler des formations de couverture serait à prendre en compte. Une alternative à l'approche géologique serait de dresser une carte sur la base des données de Sise-Eaux concernant la conductivité et la turbidité des eaux souterraines brutes. L'introduction dans le modèle d'une interaction entre  $Karst$  et  $Prec$  (par exemple  $Karst \times Prec$ ) ne représente pas un apport significatif, ce qui traduirait plus les carences des covariables  $Karst$  et  $Prec$ , précédemment évoquées et la relative sécheresse des années d'étude, qu'une absence réelle d'interaction au plan physique. Les crues turbides hivernales, qui sont une expression typique de l'interférence entre karst et ruissellements de surface, ont par exemple fait défaut en Haute-Normandie pendant la période d'étude. Un système d'information géographique permettrait de calculer l'interaction entre karst et écoulements au niveau local, de l'intégrer au niveau départemental et de supprimer ainsi les biais écologiques liés à l'hétérogénéité infra-départementale des covariables primaires.

La proportion d'eau distribuée sans **désinfection** n'intervient pas significativement pour expliquer le risque de "pollution" pour les classes 1 et 2 ( $0,05 < p < 0,1$ ). Cela pourrait s'expliquer par une probabilité de dysfonctionnement du chloromètre supérieure pour les très petites unités que pour les UD de taille supérieure. Il serait théoriquement possible (si l'information est correctement renseignée) d'utiliser l'attribut du prélèvement  $T =$  "eau distribuée désinfectée" ou  $S =$  "eau distribuée sans désinfection" pour estimer le taux moyen de pannes de désinfection par UD dans chaque département, en complément de la covariable  $sDes$ . Cette option représente un intérêt majeur pour reconstituer la progression de la désinfection dans le temps car l'historique du niveau de traitement n'est pas conservé dans Sise-Eaux.

<sup>6</sup> Les études épidémiologiques concernant cette thématique ne sont pas facilement repérables car le thésaurus de Medline ne comporte pas le mot clé "karst".



Le tableau 12 rassemble les questions qui pourraient faire l'objet de nouveaux développements. Parmi les possibilités de développement non encore évoquées, des variations sur les niveaux du seuil utilisés pour la définition de *N\_Pols* permettraient d'examiner si des différences

existent entre les déterminants des non conformités, dont la fréquence conditionne l'appréciation sanitaire officielle sur l'eau distribuée, et ceux des pollutions plus importantes, qui représente mieux, à notre sens, le risque épidémique.

TABLEAU 12 PROPOSITIONS D'ÉVOLUTION DU PROTOCOLE D'ÉTUDE		
Action	Intérêt	Degré de priorité
<b>Améliorer la variable à expliquer (<i>N_Pols</i>)</b>		
Variation du seuil de "20 EC+En"	Comparaison avec les non conformités	3
Intégrer dans la définition de <i>N_Pols</i> la turbidité, le type de ressources (surface vs. souterraine), le niveau de traitement et éventuellement le type d'aquifère	Définitions de cas plus spécifiques	2
<b>Améliorer les covariables</b>		
Estimation précises par année de <i>sDUP</i>		1
Compléter la covariable <i>sDes</i> par le taux de désinfectant indétectable dans les échantillons, par année		1
Précipitations : estimer les précipitations efficaces à partir des <i>Prec</i> et <i>ETP</i> par 24h ; distinguer les saisons	Mieux approcher les ruissellements de surface	1
Améliorer la covariable <i>Karst</i> en introduisant la notion de karst couvert	Idem	2
Améliorer la covariable <i>Karst</i> en élaborant une cartographie basée sur la conductivité et la turbidité des eaux des captages	Idem	1
Utiliser les données hydrologiques (prévalence des crues, variabilité des débits...)	Idem	2
Relief : utiliser des indicateurs de variabilité des altitudes	Idem (notion de pente)	2
Calculer localement les interactions entre le relief, l'occupation du sol et le karst, à partir d'un SIG	Réduire les biais écologiques	3
<b>Actions générales</b>		
Augmenter la durée d'étude (6 années)	Puissance statistique	1
Analyses transversales répétées dans le temps	Reproductibilité, évolution des facteurs de risque ( <i>sDUP</i> et <i>sDes</i> notamment) et impact sur <i>N_Pols</i>	1

## 4. Conclusion

La surveillance des pollutions fécales accidentelles de eaux distribuées et de ses déterminants complète utilement la surveillance des épidémies. La détection des épidémies repose actuellement sur le suivi analytique de l'eau et sur le signalement par les acteurs locaux de cas groupés de maladie. Ces modalités sont inadaptées au cas de petites UD, d'une part parce que les prélèvements du contrôle réglementaire de l'eau sont rares (tableau 3) et d'autre part parce que les populations desservies sont trop peu nombreuses pour qu'un médecin généraliste voit suffisamment de cas pour qu'il suppute la présence d'une épidémie (Beaudeau P. *et al.* 2008). Au contraire, la surveillance des "pollutions" permet d'approcher des facteurs de risque épidémique parmi les petites UD car les "pollutions" identifiées sont 1 000 fois plus nombreuses que les épidémies rapportées dans ces catégories de taille d'UD, la fréquence des "pollutions" compensant largement la rareté des prélèvements. Les UD de ces catégories de tailles sont homogènes tant du point de vue des ressources (souterraines) que des traitements (simple désinfection au chlore) et les relations avec les déterminants étudiés ne sont pas brouillées par l'existence de traitements de clarification, comme c'est le cas pour les UD de plus grande taille. Il apparaît en revanche que cette approche est

mal adaptée à la recherche des déterminants des "pollutions" des grandes UD dont les ressources se partagent entre eaux souterraines et eaux de surface et dont les traitements sont de niveau et de nature très divers.

Selon la classe de taille d'UD, entre le quart et la moitié de la variance du nombre de "pollutions" survenant dans un département métropolitain français s'expliquent par 4 covariables clairement interprétables : les précipitations totales, l'importance des formations karstiques, la prévalence de la protection réglementaire des captages et la prévalence de la désinfection de l'eau distribuée. Compte tenu du caractère approximatif des variables d'ordre naturel et du protocole écologique, la part de variance expliquée est estimée par défaut. Il existe une marge de progrès importante qui réside dans une meilleure formulation de ces covariables mais surtout dans l'examen de plusieurs périodes de temps. Ce dernier développement serait susceptible d'apporter des arguments déterminants pour l'interprétation des corrélations en terme de causalité. L'enjeu porte sur les variables *sDUP* et *sDes*, dont l'évolution appartient essentiellement aux acteurs locaux de la santé environnementale.

# Références bibliographiques

Ainsworth R, 2004. Safe piped water: Managing microbial water quality in piped distribution systems. 168p, World Health Organization, Genova.

Beaudeau P, de Valk H, Vaillant V, Mannshott C, Tillier C, Mouly D, & Ledrans M, 2008. Lessons learned from 10 investigations of waterborne gastroenteritis outbreaks, France, 1998-2006. *J. Water Health*, 6[4], 491-503.

Beaudeau P, Mouly D, & Lauzeille D, 2006. Détection précoce automatisée des épidémies de gastro-entérites d'origine hydrique à partir des données de vente ou de remboursement des médicaments. Étude de faisabilité: choix des données sanitaires et des départements pilotes [Early detection of waterborne gastroenteritis outbreaks based on drug sale or refunding data. Feasibility study: choice of the data and pilot counties] Rapport. Institut de veille sanitaire, Saint-Maurice.

Beaudeau P, Zeghnoun A, Ledrans M, & Volatier J.L, 2003. Consommation d'eau du robinet pour la boisson en France métropolitaine: résultats tirés de l'enquête INCA1. *Environnement, Risques & Santé* 2[3], 147-58.

Curriero F.C, Patz J.A, Rose J.B, & Lele S, 2001. The association between extreme precipitation and waterborne disease outbreaks in the United States, 1948-1994. *Am.J.Public Health* 91[8], 1194-9.

Di Palma M, Carbone S, Beaudeau P, Checlair E & Gallay A, 2003. Épidémie de gastro-entérites à *Cryptosporidium*, Dracy-le-Fort, Saône et Loire, Septembre 2001. Institut de veille sanitaire, Saint-Maurice.

Dussart-Baptista L, Massei N, Dupont J.P, & Jouenne T, 2003. Transfer of bacteria-contaminated particles in a karst aquifer: evolution of contaminated materials from a sinkhole to a spring. *Journal of Hydrology* 284[1-4], 285-95.

Marsaud B, 1997. Structure et fonctionnement de la zone noyée des karsts à partir des résultats expérimentaux. Thèse de doctorat. Université Paris XI Orsay, Bureau de Recherche Géologique et Minière.

Ministère de l'Économie, des Finances et de l'Emploi. 2007. Chiffres clés du tourisme. Ministère de l'Économie, des Finances et de l'Emploi, Paris.

Ministère de la Santé et des Solidarités. 2006. L'eau potable en France, 2002-2004. Paris.

Ministère de la Santé et de la Protection sociale, ministère de l'Écologie et du Développement durable, ministère de l'Emploi, du Travail et de la Cohésion sociale et ministère délégué à la Recherche. 2004. Plan national santé environnement 2004-2008. Ministère de la Santé et de la Protection sociale, Paris.

Nedellec V, Jouan M, Ledrans M, & Tricard D, 1996. Pollutions accidentelles des eaux d'alimentation en France. Rapport du Réseau national de santé publique et du ministère du Travail et des Affaires sociales. Paris.

Nicod J, 1994. Plateaux karstique sous couverture en France, d'après les travaux récents. Caractères géomorphologiques et problèmes d'environnement [Karstic plateaus clay, sand or detrital cover]. *Ann.Geo.*[576], 170-94.

Tillaut H, Encrenaz N, Checlair E, Alexandre-Bird A, Gomes Do Esperito Santo E, & Beaudeau P, 2004. Épidémie de gastro-entérite, Isère, novembre 2002. *BEH*[12], 3-4.

World Health Organization 2004. Guidelines for Drinking-water Quality, third edition. 515p, Genova.

Zmirou D, Ferley J.P, Collin J.F, Charrel M & Berlin J, 1987. A follow-up study of gastro-intestinal diseases related to bacteriologically substandard drinking water. *Am.J.Public Health* 77[5], 582-4.

Zmirou D, Rey S, Courtois X, Ferley J.P, Blatier J.F, Chevallier P, Boudot J, Potelon J.L, & Mounir R. 1995. Residual microbiological risk after simple chlorine treatment of drinking ground water in small community systems. *European Journal of Public Health* 5, 75-81.

## Déterminants naturels et humains des pollutions fécales accidentelles des petites unités d'alimentation en eau potable, France, 2003-2004 : une étude écologique géographique

Les pollutions microbiologiques accidentelles de l'eau distribuée (notée "pollutions") ont été définies comme l'occurrence d'au moins 20 bactéries fécales (*Escherichia coli* et entérocoques) formant colonies parmi les analyses réalisées sur les prélèvements (prise d'essai 100 mL) collectés un même jour sur une même unité de distribution (UD).

2 739 "pollutions" ont frappé les petites UD (<2 000 usagers desservis) entre 2003 et 2004 dans les 90 départements de la France continentale, hors Paris et la petite couronne. Après standardisation sur la distribution des UD par taille (partition en 3 classes), nous avons exploré les corrélations entre le nombre de "pollutions" par département (*N\_Pols*) et des indicateurs météorologiques, hydrologiques, topographiques et géologiques, ainsi que des indicateurs de gestion sanitaire de l'eau.

*N\_Pols* est liée à 4 covariables : (i) les précipitations, de façon très significative, (ii) l'extension des secteurs karstiques, (iii) la prévalence de la désinfection et (iv) celle des mesures de protection des captages. Un excès de 100 mm de hauteur d'eau précipitée annuellement correspond à une augmentation du risque de "pollution" de 30 %. Un accroissement de 10 % des proportions des surfaces karstifiées, des ressources sans périmètre de protection et des UD dépourvues de désinfection s'accompagnent d'un accroissement de 10 % de la fréquence des "pollutions". Les corrélations sont reproductibles entre les 3 classes de taille d'UD et conformes à l'attente des experts.

Il serait intéressant de répéter ces études dans le temps afin de renforcer l'interprétation en termes de causalité, et, le cas échéant, de suivre l'évolution du risque et de ses déterminants.

**Mots clés :** base de données sur la qualité de l'eau, chloration, contrôle de l'eau distribuée, contamination fécale, désinfection, eau souterraine, entérocoque, *Escherichia coli*, étude transversale, France, karst, périmètre de protection, précipitations, protocole écologique, réglementation, eau du robinet.

### Natural and human factors of faecal contamination events of drinking water in small distribution networks, France, 2003-2004: a geographical ecological study

*We defined an accidental faecal contamination event of drinking water ("pollution") as the occurrence of at least 20 forming colony faecal bacteria (Escherichia coli and enterococci) among all the samples collected (specimen test 100mL) within one day from a given drinking water distribution unit (DWUD).*

*2739 "pollutions" hit the small DWUD (< 2,000 people provided), in 2003-2004 within the 90 counties ("departments") in the continental France, Paris area being excluded. Once a size standardization of the DWUD (with 3 size classes), we explored correlations between the number of "pollutions" by county *N\_Pols* and various indicators related to weather, land cover, topography, geology and water health management.*

**N\_Pols* correlates with 4 covariates: (i) precipitations - very significantly -, (ii) the extension of the karstic areas, (iii) the prominence of disinfection and (iv) the rate of drillings areas protected. 100 mm of yearly precipitation in excess increases the pollution risk by 30 %. A 10 % extension of the karstic areas, of unprotected resources and of DWUD free of disinfection entail a higher risk of "pollutions" by 10 %. The correlations are valid for the 3 DWUD size classes and match the expertise.*

*It could be worth repeating this study through time in order to reinforce the interpretation toward causality, and to follow up the evolution of both the risk and its factors.*

Citation suggérée :

Beaudeau P. Déterminants naturels et humains des pollutions fécales accidentelles des petites unités d'alimentation en eau potable, France, 2003-2004 : une étude écologique géographique. Saint-Maurice (Fra) : Institut de veille sanitaire, août 2008, 25 p. Disponible sur : [www.invs.sante.fr](http://www.invs.sante.fr)