

MINISTÈRE DE L'ENVIRONNEMENT
ET DE LA LUTTE CONTRE
LES CHANGEMENTS CLIMATIQUES

CONTAMINATION
BACTÉRIOLOGIQUE DES PETITS
COURS D'EAU EN MILIEU
AGRICOLE : DÉVELOPPEMENT
DE MODÈLES PRÉDICTIFS

2020

Photo de la page couverture :

La rivière Boyer Nord vue de la station d'échantillonnage
(photo : Michel Patoine, MELCC)

Coordination et rédaction

Cette publication a été réalisée par la Direction générale du suivi de l'état de l'environnement du ministère de l'Environnement et de la Lutte contre les changements climatiques (MELCC).

Renseignements

Pour tout renseignement, vous pouvez remplir le formulaire disponible à cette adresse :

www.environnement.gouv.qc.ca/formulaires/renseignements.asp

Téléphone : 418 521-3830
1 800 561-1616 (sans frais)

Ministère de l'Environnement
et de la Lutte contre les changements climatiques
675, boul. René-Lévesque Est, 7^e étage
Québec (Québec) G1R 5V7

Ce document peut être consulté sur le site Web du ministère à l'adresse suivante :

<http://www.environnement.gouv.qc.ca/eau/flrivlac/contamination-bacteriologique-petits-cours-eau.pdf>

ou

Visitez notre site Web :

<http://www.environnement.gouv.qc.ca/eau/flrivlac/index.htm>

Référence à citer

PATOINE, Michel et François D'AUTEUIL-POTVIN (2020). *Contamination bactériologique des petits cours d'eau en milieu agricole : développement de modèles prédictifs*, Québec, ministère de l'Environnement et de la Lutte contre les changements climatiques, Direction de la qualité des milieux aquatiques, 29 pages et 10 annexes.

Dépôt légal – 2020
Bibliothèque et Archives nationales du Québec
ISBN 978-2-550-87111-8 (PDF)

Tous droits réservés pour tous les pays

© Gouvernement du Québec – 2020

ÉQUIPE DE RÉALISATION

Auteurs

Michel Patoine
François D'Auteuil-Potvin

Révision scientifique

Marc Simoneau*
Denis Brouillette
L'équipe de la DAEMH¹
Philippe Cantin^{2, **}

Échantillonnage et suivi auprès des préleveurs d'eau

René Therreault
Manon Ouellet
Marie-Julie Laperrière*
Camil Giasson*

Analyses de laboratoire

Philippe Cantin^{**} et son
équipe²

Cartographie et compilation de données d'utilisation du territoire

Mona Frenette
Suzanne Lavoie
Lyne Blanchet*
Serge Poirier*

Mise en page et graphisme

Claire Cournoyer
Murielle Gravel*

¹. Direction de l'agroenvironnement et du milieu hydrique (DAEMH).

². Centre d'expertise en analyse environnementale du Québec (affectation au moment de la réalisation des analyses et de la révision).

* Ancien employé de la Direction générale du suivi de l'état de l'environnement.

** Affectation au moment de la révision scientifique et de la réalisation des analyses de laboratoire.

Mots clés : coliformes fécaux, qualité de l'eau, bassins versants agricoles, modèles de prédiction du risque.

REMERCIEMENTS

Les auteurs remercient la Direction de l'agroenvironnement et du milieu hydrique, la Direction régionale de l'analyse et de l'expertise de l'Estrie et de la Montérégie, les gouvernements du Québec et du Canada (Plan d'action Saint-Laurent Vision 2000), Environnement et Changement

climatique Canada, Agriculture et Agroalimentaire Canada et la Corporation de l'aménagement de la rivière L'Assomption (CARA) pour leur contribution financière dans le cadre des analyses de qualité de l'eau de certaines stations.

RÉSUMÉ

La contamination microbienne d'origine fécale des cours d'eau en milieu agricole a fait l'objet d'un nombre limité d'études, et peu d'outils simples ont été mis à la disposition des acteurs dans le domaine de l'eau pour orienter leurs décisions. Pourtant, cette forme de contamination est présente et contribue à accroître le risque de transmission par l'eau d'organismes pathogènes aux humains et aux animaux. Afin qu'on puisse développer des modèles prédictifs de ce risque, les données des suivis de coliformes fécaux (CF) ou thermotolérants réalisés à une cinquantaine de stations dans des petits bassins versants agricoles situés dans différentes régions au Québec ont été mises en relation avec des indicateurs de la pression des activités agricoles et humaines dans les bassins versants. Les données de CF mesurées toutes les deux semaines pendant trois années (de 2000 à 2009) à 14 stations spatialement indépendantes ont été retenues pour la construction de modèles reliant celles-ci aux variables de pression du territoire pour des bassins versants de 10 à 100 km². Les données de 2006 à 2009 d'un autre groupe de 14 stations indépendantes ont été retenues pour la validation externe de ces modèles. Les données aux autres stations ont permis de quantifier l'influence de sources de contamination anormales selon les variables de pression du territoire retenues dans les modèles.

Des relations entre les CF dans les cours d'eau et la densité animale et humaine ou la densité d'élevage avec amas de fumier et les superficies recevant des déjections animales ont été développées et validées pour la saison estivale, en raison des usages de l'eau qui y ont lieu. Elles

permettent de prédire, à l'aide d'indicateurs simples et disponibles, des concentrations médianes ou des fréquences de dépassement de critères de qualité de l'eau pour les CF. Elles constituent un outil servant à mieux évaluer le risque de contamination bactériologique dans les cours d'eau et peuvent aider à orienter des stratégies de gestion de ce type de contamination dans les bassins versants agricoles. Elles permettent notamment d'estimer les valeurs estivales de CF dans des petits bassins versants non pourvus de station de qualité d'eau, d'identifier les stations influencées par des sources de contamination anormales ainsi que d'estimer la contribution de certaines activités agricoles et l'effet prévisible de certaines interventions d'assainissement ou de certains développements de l'agriculture dans un petit bassin versant. L'application de ces outils aux stations de la présente étude a notamment montré que la densité animale ou la superficie recevant des déjections animales contribuent davantage aux valeurs de CF observées.

Les relations développées sont applicables à des bassins versants comparables à ceux utilisés pour leur développement et leur validation, dans les limites des variables prédictives, qui ne représentent pas toutes les sources possibles de contamination bactériologique. La présente étude a en outre permis de montrer leur potentiel comme outil de prédiction du risque de contamination. D'autres étapes devront toutefois être franchies pour que cet outil devienne pleinement opérationnel.

TABLE DES MATIÈRES

INTRODUCTION	1		
1 MATÉRIEL ET MÉTHODES	2		
1.1 Sélection des petits bassins versants et des stations de qualité d'eau.....	2		
1.2 Choix des indicateurs microbiens de la qualité de l'eau et du territoire.....	4		
1.3 Échantillonnage de l'eau et analyses bactériologiques	7		
1.4 Méthodes statistiques	7		
1.4.1 Construction des modèles.....	7		
1.4.2 Validation externe des modèles.....	8		
1.4.3 Identification des stations influencées	8		
2 RÉSULTATS ET DISCUSSION.....	9		
2.1 Relations entre les coliformes fécaux dans les cours d'eau et les activités dans les bassins versants	9		
2.1.1 Modèles 2006 basés sur la densité animale et humaine	9		
2.1.2 Modèles 2006 basés sur certaines pratiques agricoles	10		
2.1.3 Validation externe et domaine d'application des modèles	16		
2.2 Utilisation des relations prédictives du risque de contamination.....	18		
2.2.1 Identification des stations influencées.....	18		
2.2.2 Estimation de la contribution des sources animales et humaines	20		
2.2.3 Estimation des valeurs de CF dans les bassins versants sans suivi de la qualité de l'eau et de l'effet des changements dans les bassins versants	23		
2.2.4 Limites des modèles et balises d'interprétation des résultats.....	24		
CONCLUSION	27		
RÉFÉRENCES BIBLIOGRAPHIQUES	28		
ANNEXES	30		

Tableau 1	Description des stations d'échantillonnage et des petits bassins versants	3
Tableau 2	Types d'élevage dans les bassins retenus pour la construction des modèles 2006.....	6

Tableau 3	Stations soumises à des influences anormales.....	19
-----------	---	----

LISTE DES FIGURES

Figure 1	Emplacement des bassins versants et des stations d'échantillonnage2	Figure 7	Dépassement du critère de qualité de l'eau de mai à octobre pour les activités de contact direct et l'irrigation en fonction de la DA et de la DH – modèle 2006RL... 15
Figure 2	Médianes estivales de CF mesurées et prédites à l'aide des DA et DH – modèles 2006 12	Figure 8	Concentrations médianes estivales de CF mesurées à 28 stations et contribution estimée des animaux d'élevage et des humains 21
Figure 3	Médiane estivale de CF en fonction de la DA et de la DH – modèle 2006A 12	Figure 9	Concentrations médianes estivales de CF mesurées à 28 stations et contribution estimée des superficies réceptrices de déjections et du stockage en amas 22
Figure 4	Médiane estivale de CF en fonction de la S_déj ou de la DA_entr et de la DA_amas 13		
Figure 5	Dépassements de mai à octobre des critères de CF mesurés et prédits à l'aide de la DA et de la DH – modèles 2006RL 14		
Figure 6	Dépassement du critère de qualité de l'eau de mai à octobre pour les activités de contact indirect en fonction de la DA et de la DH – modèle 2006RL 14		

LISTE DES ANNEXES

Annexe 1	Modèles 2001 et données utilisées pour leur développement 30	Annexe 7	Utilisations possibles des relations entre les CF et le territoire 56
Annexe 2	Modèles 2006 et données utilisées pour leur développement 38	Annexe 8	Contribution des variables à la médiane de coliformes fécaux et résidus 57
Annexe 3	Validation externe des équations de prédiction de la concentration médiane estivale de coliformes fécaux 51	Annexe 9	Exemples d'utilisations des relations pour prédire le risque de contamination 60
Annexe 4	Résidus des modèles 2006 52	Annexe 10	Références bibliographiques des annexes 63
Annexe 5	Données de pression microbienne du territoire 54		
Annexe 6	Fréquence cumulée des densités animales et humaines des petits bassins versants 55		

INTRODUCTION

Les ruisseaux et les petites rivières sont essentiels à la survie de plusieurs espèces de poissons et autres organismes aquatiques. Ils permettent divers usages comme la pêche, l'irrigation des cultures ou l'abreuvement du bétail, qui peuvent être compromis par la dégradation de la qualité bactériologique de l'eau. Cette dégradation influence également la qualité et les coûts de traitement de l'eau brute aux prises d'eau potable des rivières situées en aval le cas échéant, et augmente les risques de transmission de microorganismes pathogènes. Plusieurs de ces petits cours d'eau drainent des bassins versants où les activités agricoles sont importantes.

Depuis 25 ans, le ministère de l'Environnement et de la Lutte contre les changements climatiques (MELCC) a réalisé des suivis de la qualité bactériologique et physicochimique de l'eau dans une cinquantaine de petits bassins versants agricoles au Québec (Patoine et D'Auteuil-Potvin, 2015). Les coliformes thermotolérants ou coliformes fécaux (CF) ont été retenus comme indicateur de pollution par des matières fécales et de la présence possible de microorganismes pathogènes (CEAEQ, 2011). Les stations échantillonnées ne représentent qu'une faible proportion des petits bassins versants agricoles au Québec. Les résultats de ces suivis ont tout de même montré que malgré une baisse des concentrations de CF consécutive aux interventions d'assainissement à quelques stations, la qualité bactériologique de ces cours d'eau demeure généralement problématique, notamment en saison estivale, durant laquelle les principaux usages de l'eau et les activités agricoles sont plus présents (Patoine et D'Auteuil-Potvin, 2015).

En l'absence de suivi des CF, une approche par modélisation constitue une solution de rechange. Des modèles déterministes comme SWAT (*Soil*

and Water Assessment Tool) ou GIBSI (gestion intégrée des bassins versants à l'aide d'un système informatisé) permettent de modéliser sur une base journalière les CF d'origine agricole et humaine. Ce dernier a été utilisé pour modéliser les CF dans deux petits bassins versants de la rivière Chaudière (Rogel, 2007). Ces modèles permettent aussi de prédire l'effet de certaines mesures correctives à l'échelle du bassin versant. Cependant, l'importance de l'infrastructure requise pour leur fonctionnement limite leur utilisation. Des modèles de régression liant les concentrations de CF ou les dépassements de critères de qualité de l'eau et les variables du territoire liées à la contamination fécale seraient plus simples à utiliser et pourraient aider à mieux comprendre les mécanismes impliqués et aider les divers intervenants de la gestion de l'eau dans leurs décisions.

Cependant, peu de modèles de prédiction des CF existent et de nouveaux outils sont donc nécessaires. Hébert et Blais (2017) ont développé, à l'aide des données de 2009 à 2012 pour les grands bassins versants au Québec, des modèles prédictifs pour le phosphore et l'azote, mais n'ont pas réussi à obtenir un modèle performant pour les CF. Par contre, Patoine (2011) a établi des relations entre les CF et la densité animale et humaine à l'aide des données couvrant la période de 2000 à 2003 dans 11 petits bassins versants agricoles au Québec (annexe 1). Les données plus récentes disponibles permettent de valider ou d'améliorer et de compléter ces relations.

Ce rapport présente le développement et la validation de relations liant les CF mesurés dans les cours d'eau et des indicateurs de pression microbienne des petits bassins versants agricoles au Québec ainsi que certaines utilisations de ces relations prédictives du risque de contamination bactériologique des petits cours d'eau agricoles.

1 MATÉRIEL ET MÉTHODES

1.1 Sélection des petits bassins versants et des stations de qualité d'eau

Les données du suivi de la qualité bactériologique des petits cours d'eau en milieu agricole réalisé à 51 stations d'échantillonnage situées dans une vingtaine de bassins versants primaires de 6 régions hydrographiques du Québec ont été utilisées. Ces stations se situent dans les provinces naturelles des basses-terres du Saint-Laurent, des Appalaches et des Laurentides méridionales et centrales (figure 1). Elles drainent des petits bassins versants dont la superficie varie de 3 à 202 km². La durée du suivi a varié, selon les stations, de trois années à deux décennies alors que la fréquence d'échantillonnage a varié selon les périodes (Patoine et D'Auteuil-Potvin, 2015). La période de suivi de la qualité de l'eau considérée est comprise entre 1999 et 2012. Les petits bassins

versants drainés aux 51 stations supportent des densités animales de 5 à plus de 300 unités animales (UA) par kilomètre carré et des densités humaines de 2 à plus de 100 habitants par kilomètre carré. Le tableau 1 présente les superficies drainées à chaque station ainsi que les densités animales en 2006 (données de Statistique Canada réparties au prorata de la superficie de la municipalité dans le bassin versant) et les densités des populations humaines dont les eaux usées sont rejetées dans les petits bassins versants en 2007. On a obtenu ces densités animales et humaines en divisant le nombre d'UA et de personnes du petit bassin versant par sa superficie.

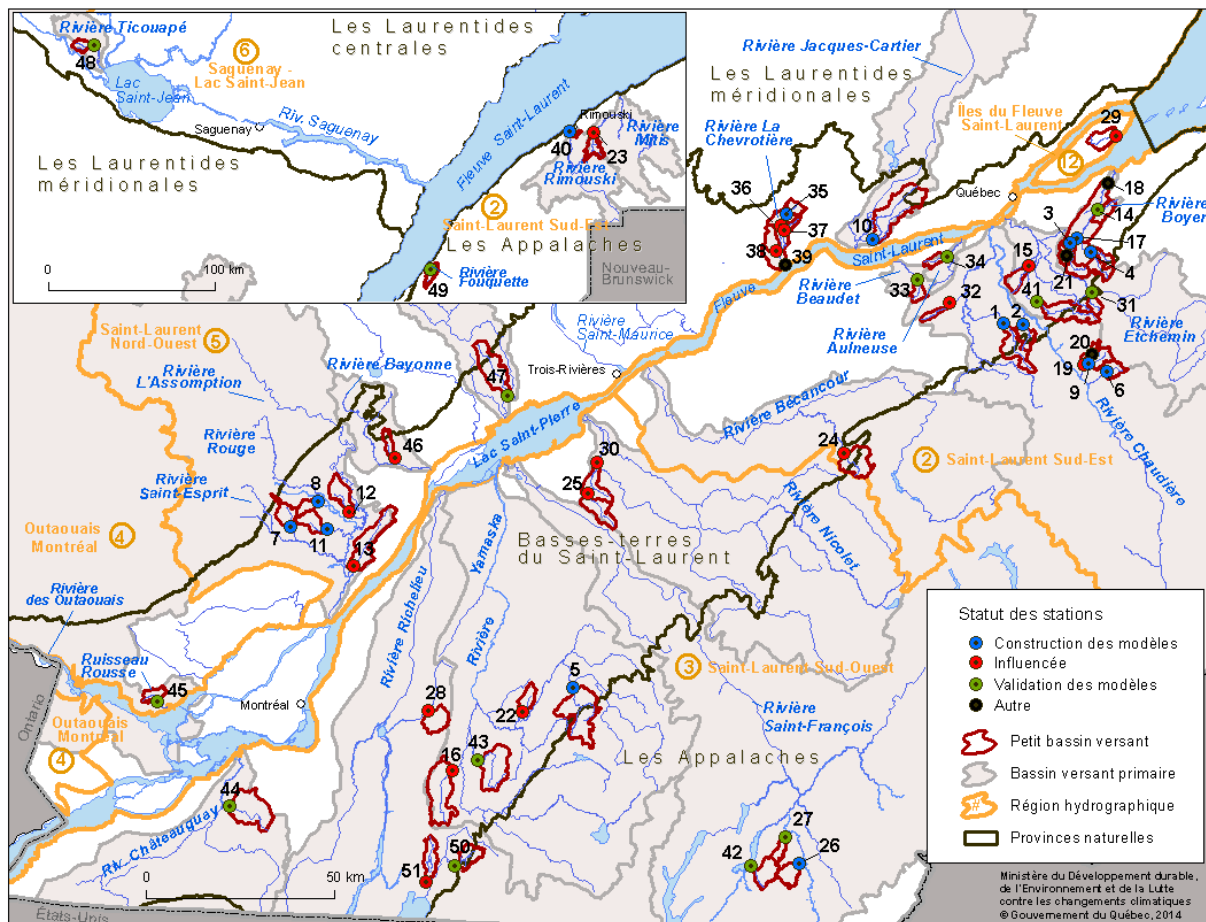


Figure 1 Emplacement des bassins versants et des stations d'échantillonnage

Tableau 1 Description des stations d'échantillonnage et des petits bassins versants

Bassin versant primaire	Cours d'eau	Numéro de station figure 1	Numéro BQMA ^a	SBV ^b (km ²)	DA 2006 ^c (UA/km ²)	DH 2007 ^d (hab./km ²)	Statut des stations
Mitis	Petite rivière Neigette	23	02190010	98	12	11	Influencée
Rimouski	Levasseur	40	02200022	23	7	35	Modèle
Fouquette	Fouquette	49	02E90002	55	34	11	Validation
Boyer	Boyer	18 (includ 17, 3, 14)	02300001	202	103	23	Autre ^e
	Boyer Sud	17 (includ 4)	02300002	65	137	19	Modèle
	Boyer Nord	3 (includ 21)	02300003	36	137	20	Modèle
	Du Portage	14	02300004	17	65	18	Validation
	Honfleur	4	02300005	23	142	18	Modèle (2001) ^e
	Grillade	21	02300013	4	98	22	Autre ^e
Etchemin	Saint-Patrice	15	02330029	21	48	20	Influencée
	Du Moulin Bernier	31	02330048	14	79	8	Validation
	Le Bras	41	02330050	53	96	23	Validation
Chaudière	Des Îles Brûlées	2	02340086	22	321	45	Modèle
	Turmel	19	02340091	7	88	21	Influencée ^e
	Binet	20	02340092	3	74	15	Autre ^e
	Belair	9 (includ 20)	02340093	30	66	14	Modèle
	Morency	6	02340098	18	56	9	Modèle
Beaudet	Bras d'Henri	1	02340099	47	239	15	Modèle
	Laflamme	32	02340119	10	55	26	Influencée
	Bourret	33	02350002	14	24	16	Validation
Bécancour	Bourbon	24	02400048	51	38	9	Influencée
Aulneuse	Des Moulanges	34	02500001	12	22	20	Validation
Nicolet	Saint-Zéphirin	30 (includ 25)	03010046	77	64	13	Influencée
	Saint-Zéphirin amont	25	03010086	40	51	13	Influencée
Saint-François	Brook	42	03020065	35	42	11	Validation
	Veillette	26	03020388	27	61	7	Modèle
	Bradley	27	03020389	13	57	8	Validation
Yamaska	Runnels	5	03030008	98	77	12	Modèle
	À la Barbue	43	03030110	70	179	34	Validation
	Du Sud-Ouest	16	03030237	91	117	18	Influencée
Richelieu	Des Aulnages	22	03030253	21	109	13	Influencée
	Walbridge	50	03040066	29	95	13	Validation
	Ewing	51	03040073	31	71	11	Influencée
	À l'Ours	28	03040195	35	22	19	Influencée
Châteauguay	De l'Esturgeon	44	03090018	96	20	29	Validation
Outaouais	Rousse	45	04310084	17	9	50	Validation
Chevrotière	Le Grand Bras amont	35	05050096	14	5	2	Modèle
	Le Grand Bras	36 (includ 35)	05050097	26	11	2	Influencée
	Le Petit Bras	37	05050098	7	19	26	Influencée ^e
	Du Moulin	38	05050099	20	25	133	Influencée
	La Chevrotière	39 (includ 35 à 38)	05050100	106	19	28	Autre ^e
Jacques-Cartier	Aux Pommes	10	05080053	95	12	45	Modèle
	Du Point du Jour	13	05220063	71	14	65	Influencée
L'Assomption	Vacher	11	05220239	29	46	81	Modèle
	Saint-Pierre	12	05220240	33	28	90	Influencée
	Saint-Esprit	7	05220241	28	81	31	Modèle
	Desrochers	8	05220249	15	40	31	Modèle
Bayonne	Bibeau	46	05240014	15	39	13	Influencée
Du Loup	Chacoura	47	05280053	67	65	15	Validation
Ticouapé	Rouge	48	06200004	62	19	35	Validation
Île d'Orléans	Dauphine	29	12020001	27	24	2	Influencée

a : Numéro de station, Banque de données sur la qualité du milieu aquatique;

b : SBV = Superficie du bassin versant, selon Arc Map 9.3;

c : Densité animale pour l'année 2006 en unités animales par km² ;

d : Densité humaine de la population dont les eaux usées sont rejetées dans le bassin versant pour l'année 2007 en habitants par km²;

e : Station hors du domaine de 10 à 100 km² ou dépendante.

Les 51 petits bassins versants ont été divisés en 3 catégories, soit ceux retenus pour la construction des modèles, ceux pour leur validation et les stations influencées (figure 1 et tableau 1). Les bassins versants retenus pour la construction et la validation des modèles sont spatialement indépendants et drainent une superficie de 10 à 100 km². Afin qu'on puisse obtenir la meilleure précision possible des modèles, les stations sélectionnées pour leur construction comportaient des analyses de CF mesurées toutes les deux semaines pendant trois années, alors que les stations retenues pour la validation pouvaient comporter moins de données. Dans les deux cas, les données de CF les plus proches de l'année du recensement quinquennal sur l'agriculture de 2006 ou de 2001 de Statistique Canada ont été retenues pour une raison de concordance entre les données de qualité d'eau et d'occupation du territoire. Les valeurs de CF manquantes ont été complétées à l'aide de données proches de la période considérée pour que le nombre d'échantillons soit plus homogène entre les stations. Ces stations n'étaient pas influencées par une source anormale de contamination comme un rejet d'un réseau d'égout non traité dans le bassin versant ou un rejet ponctuel en amont rapproché de la station d'échantillonnage (rejet de résidence isolée, bétail au cours d'eau, zone urbaine) au cours de la période retenue.

Les 11 petits bassins versants numérotés de 1 à 11, dont les données de CF de 2000 à 2003 avaient

servi à construire les équations liant les CF à la densité animale et humaine publiées dans l'article de Patoine (2011), ont été utilisés pour présenter de façon plus complète ces relations, nommées « modèles 2001 » (annexe 1). Les données de CF les plus récentes disponibles à 14 stations (numérotées de 1 à 3, 5 à 11, 17, 26, 35 et 40) ont été utilisées pour la construction de nouveaux modèles, nommés « modèles 2006 » (annexe 2). Quatorze autres stations pour lesquelles au moins seize données de CF étaient disponibles de 2006 à 2009, ont été retenues pour la validation externe des modèles (annexe 3).

Quatre stations (numérotées 18, 20, 21 et 39) non indépendantes spatialement et dont la superficie du bassin versant était hors de la gamme de 10 à 100 km² ont servi à la vérification du potentiel d'utilisation des modèles pour des bassins imbriqués ou englobants ou encore à l'extérieur de leur domaine de superficie.

Les 18 autres stations, non retenues pour la construction ou la validation externe des modèles en raison de la présence d'une source de contamination susceptible d'exercer une influence particulière sur les valeurs de CF ou encore de l'indication d'une valeur anormale de CF selon le test de Grubb (annexe 4), ont été utilisées pour la quantification de cette influence.

1.2 Choix des indicateurs microbiens de la qualité de l'eau et du territoire

Les coliformes fécaux ont été retenus comme indicateur du niveau de contamination bactériologique d'origine fécale des cours d'eau et du risque de contamination par des microorganismes pathogènes en raison de leur utilisation à cette fin depuis plusieurs années dans les suivis du MELCC (Patoine et D'Auteuil-Potvin, 2015).

Les sources potentielles de CF et de microorganismes pathogènes dans les bassins versants retenus sont principalement les déjections des animaux d'élevage et les eaux usées des populations humaines. D'autres sources comme la faune et les animaux de compagnie peuvent également être présentes. Une partie des CF contenus dans les déjections animales se retrouve dans les cours d'eau à la suite de leur dépôt direct par le bétail y ayant libre accès, d'écoulements

d'amas de fumier ou du transport par les précipitations en provenance des terres en culture où leur épandage a été effectué. Les CF d'origine humaine se retrouvent également dans les cours d'eau à cause d'installations septiques défectueuses ou de lacunes dans les raccordements des réseaux d'égout ou dans la désinfection des eaux usées (Patoine et D'Auteuil-Potvin, 2015).

Plusieurs indicateurs potentiels de pression microbienne dans les bassins versants ont été compilés à l'aide des données disponibles (annexe 5) provenant des recensements sur l'agriculture de 2006 ou de 2001 de Statistique Canada (densité animale, superficies d'épandage avec ou sans incorporation au sol de fumier, de lisier ou de compost), du Système d'aide à la gestion des opérations (SAGO) du MELCC (stockage des

déjections animales en amas contigu au bâtiment d'élevage ou au champ (DA_amas), densité de bétail ayant accès aux cours d'eau (DA_accès)) et du Répertoire des municipalités du Québec (densité humaine). Une analyse de corrélation de Spearman entre ces indicateurs de pression microbienne pour les bassins versants drainés aux 14 stations retenues pour la construction des modèles (résultats non présentés) a montré que la densité animale dans le bassin versant (DA), la densité animale dont le stockage des déjections se fait en amas (DA_amas) et la superficie recevant des déjections animales (S_déj) étaient corrélées à la médiane estivale de CF. Cette étude a aussi montré que la DA était corrélée à la DA_accès et à la S_déj, alors que la densité humaine (DH) était corrélée à la DA_amas. La présence d'autocorrélation entre certaines variables et le faible nombre de stations ($n = 14$) ont permis d'orienter le choix des combinaisons de variables à inclure dans les modèles.

Les indicateurs de pression des sources microbiennes du territoire retenus pour le premier groupe de modèles plus généraux sont la DA et la DH, dont le lien avec les CF a été démontré dans un article de Patoine (2011). Deux autres indicateurs ont été retenus pour la construction d'un autre modèle plus relié aux pratiques agricoles, soit la S_déj et la DA_amas. La S_déj comprend les pâturages, mais non les déjections ayant subi un traitement de désinfection comme le compostage, alors que la DA_amas inclut les déjections stockées en amas contigu au bâtiment d'élevage ou dans un champ. La densité animale sans entreposage en amas ($DA_{entr} = DA - DA_{amas}$) a aussi été utilisée avec la DA_amas pour la construction d'un modèle. L'utilisation de deux variables explicatives pour la construction des modèles à l'aide de 14 stations permet de maintenir à moins de 15 % l'augmentation approximative de l'erreur due à l'estimation des coefficients de régression, calculée selon la relation proposée par Sawyer (1982).

La DA (exprimée en unités animales) et la DH constituent des indicateurs valables de la pression des CF produits par les élevages et les résidents en lien avec la qualité bactériologique de l'eau dans un bassin versant (Patoine, 2011). Ces densités sont exprimées par unité de superficie du bassin versant à des fins de comparabilité. La production journalière de CF provenant des déjections fraîches correspond à environ $10 \log_{10}$ (UFC) par unité animale pour les principaux types d'élevage (bovin, porcin, ovin, volaille) (Patoine et D'Auteuil-

Potvin, 2015 : annexe 1) et à environ $9 \log_{10}$ (UFC) par personne (Geldriech, 1978).

Les variations locales des pratiques agricoles (entreposage, traitement, épandage, pâturage) entre les types d'élevage et les échanges de déjections animales entre bassins versants peuvent toutefois influencer le lien entre la DA et l'apport de CF en rivière ou le risque de contamination. Le tableau 2 indique, pour les bassins versants drainés aux 14 stations retenues pour la construction des modèles 2006, la proportion des principaux types d'élevage présents. Cette proportion varie de 0 à 90 % pour l'élevage bovin, de 0 à 100 % pour l'élevage porcin, de 0 à 29 % pour la volaille et de 0 à 28 % pour les autres types d'élevage. L'élevage porcin est majoritaire dans neuf bassins versants, alors que l'élevage bovin l'est dans les cinq autres. Les variations de la conformité des installations septiques et leur position dans le bassin versant peuvent aussi influencer le lien entre la DH et l'apport de CF en rivière. Toutefois, les informations disponibles sur les installations septiques sont partielles et souvent compilées à une large échelle. Les apports de la faune et des animaux de compagnie ne sont pas considérés dans la densité animale, mais pourraient aussi avoir une influence dans certains cas particuliers. L'utilisation de la S_déj et de la DA_amas comme indicateurs de pression permet de mieux tenir compte des variations locales des pratiques agricoles et des échanges de déjections entre bassins versants. Toutefois, ces indicateurs ne tiennent pas compte de la quantité de déjections épandues ni de la DH. De plus, aucun des indicateurs retenus ne tient compte directement de l'effet des facteurs climatiques, comme le délai entre l'épandage des déjections animales et les précipitations provoquant un ruissellement.

De façon à améliorer la précision des modèles basés sur la DA, dans le cas des stations retenues pour leur construction, la répartition au bassin versant des données d'élevage de Statistique Canada du recensement de 2001 ou de 2006 disponibles par municipalité a été réalisée à l'aide des données des visites ferme par ferme (réalisées à compter de 2003) des lieux d'élevage répertoriés dans SAGO (tableaux A1.2 et A2.3), plutôt qu'au prorata de la superficie (agricole et non agricole) de la municipalité dans le bassin versant (tableau 1). Cet ajustement a permis de corriger un biais important dans le cas du petit bassin versant de la rivière Bras d'Henri notamment, en raison de la forte densité animale et de la répartition spatiale

non uniforme des élevages dans les municipalités touchant ce bassin versant.

On a obtenu les DH à l'aide des populations par municipalité du Répertoire des municipalités du Québec, en attribuant au bassin versant la proportion du nombre de résidences de la municipalité présente dans celui-ci. Le nombre de résidences provient de la Base de données

topographiques du Québec (BDTQ). Afin que la population dont les eaux usées sont rejetées dans les petits bassins versants corresponde à la réalité, une correction a été apportée pour soustraire les populations situées dans le bassin versant dont les eaux usées sont rejetées à l'extérieur de celui-ci (stations numérotées 3, 10, 12, 13, 18, 28, 44) et ajouter celles provenant de l'extérieur (stations numérotées 2, 9, 38, 48).

Tableau 2 Types d'élevage dans les bassins retenus pour la construction des modèles 2006

Bassin versant à la station (numéro figure 1)	Pourcentage des types d'élevage en 2001 ou en 2006							
	Bovin		Porcin		Volaille		Autre	
	2001	2006	2001	2006	2001	2006	2001	2006
Bras d'Henri (1)	--	35	--	64	--	1	--	1
Des Îles Brûlées (2)	20	--	71	--	9	--	0	--
Boyer Nord (3)	--	30	--	62	--	7	--	1
Boyer Sud (17)	--	29	--	65	--	6	--	0
Runnels (5)	--	25	--	69	--	5	--	1
Morency (6)	36	--	64	--	0	--	0	--
Saint-Esprit (7)	--	51	--	14	--	29	--	6
Desrochers (8)	12	--	80	--	0	--	8	--
Belair (9)	33	--	66	--	0	--	1	--
Aux Pommes (10)	90	--	0	--	4	--	6	--
Vacher (11)	--	58	--	26	--	13	--	3
Veillette (26)	--	57	--	43	--	0	--	0
Le Grand Bras amont (35)	--	0	--	100	--	0	--	0
Levasseur (40)	--	72	--	0	--	0	--	28

Les valeurs présentées sont basées sur les données sur les élevages par municipalité de Statistique Canada (2002, 2006) transformées en unités animales (selon les ratios présentés par Patoine et D'Auteuil-Potvin (2015) : annexe 1) et réparties par bassin versant selon la proportion des animaux dans le bassin, établie à l'aide des données des visites des lieux d'élevage réalisées par le MELCC à compter de 2003 et répertoriées dans SAGO.

La répartition au bassin versant des superficies recevant des déjections animales du Recensement sur l'agriculture de 2006 ou de 2001 de Statistique Canada, disponibles par municipalité, s'est faite au prorata de la superficie (agricole et non agricole) de la municipalité dans le bassin versant, l'emplacement des lieux d'épandage n'étant pas

documenté dans une base de données géoréférencées. Cette méthode de répartition peut entraîner une imprécision lorsque la proportion des superficies d'épandage d'une municipalité n'est pas la même dans le bassin versant et à l'extérieur de celui-ci. Les DA_amas ont été obtenues à l'aide des données des visites ferme par ferme (réalisées

à compter de 2003) des lieux d'élevage de SAGO pour lesquels un amas de déjections animales contigu au bâtiment d'élevage ou au champ était répertorié.

L'année retenue pour les indicateurs de pression a été choisie de façon à se rapprocher le plus possible de la période des données de CF retenue.

1.3 Échantillonnage de l'eau et analyses bactériologiques

Les échantillons d'eau ont été prélevés à une fréquence d'une fois toutes les deux semaines à une fois par mois selon les stations et les années, en fonction des différents projets (annexe 3 de Patoine et D'Auteuil-Potvin, 2015).

Le prélèvement des échantillons d'eau pour l'analyse des CF était effectué à l'aide d'une bouteille stérilisée. Les échantillons ont été conservés dans des glacières contenant des réfrigérants jusqu'à leur arrivée au Centre d'expertise en analyse environnementale du Québec (CEAEQ) pour être analysés dans un délai ne dépassant pas 48 heures après l'échantillonnage.

Les analyses des CF ont été faites par filtration sur membrane de porosité de 0,45 µm avec incubation pendant 24 heures à 44,5 °C, sur milieu m-FC (méthode MA. 700 – Fec.Ec 1.0 du CEAEQ (2011)). Jusqu'en mars 2003, les volumes filtrés étaient de 50, 10 et 1 ml, ce qui permettait de mesurer une gamme de valeurs de 2 à 6 000 unités formatrices de colonies (UFC) par 100 ml. D'avril 2003 à décembre 2012, les volumes filtrés ont été ajustés (10, 1 et 0,1 ml), ce qui a permis de mesurer une gamme de valeurs de 10 à 60 000 UFC par 100 ml.

Les résultats d'analyse ont été validés et intégrés dans la Banque de données sur la qualité du milieu aquatique (BQMA) de la Direction du suivi de l'état de l'environnement du MELCC.

1.4 Méthodes statistiques

1.4.1 Construction des modèles

Les modèles de prédiction des médianes de CF dans les cours d'eau à partir des indicateurs de pression microbienne du territoire ont été construits à l'aide de régressions par la méthode des moindres carrés. Dans le cas des modèles de prédiction des dépassements de critères de qualité de l'eau, la régression logistique a été utilisée. Les modèles 2006 bivariés (DA, DH et S_déj, DA_amas) ont été construits à l'aide de 14 stations spatialement indépendantes (bassins versants non imbriqués les uns dans les autres). Les informations détaillées sur ces stations et leur sélection apparaissent à l'annexe 2.

Les analyses de régression par la méthode des moindres carrés ont été réalisées à l'aide du logiciel SigmaStat pour Windows Version 3.10 de Systat Software, Inc. (SigmaStat). Le résultat du test de normalité de Kolmogorov-Smirnov inclus dans SigmaStat a été validé à l'aide du test de Shapiro-Wilk, mieux adapté aux échantillons de petite taille (Razali et Wah, 2011). La normalité a été acceptée au seuil de 1 %, compte tenu de la faible robustesse de ces deux tests à la non-

normalité pour des échantillons de taille inférieure à 30. La contribution de chaque variable significative à l'explication de la variance de CF a été calculée selon la méthode décrite par Borcard (2009) et Scherrer (1984). Les intervalles de confiance à 95 % (IC95) des droites de régression et les intervalles de prédiction pour une valeur unique au niveau de confiance de 95 % (IP95) ont aussi été obtenus à l'aide de SigmaStat pour les valeurs de DA et de DH ayant servi à la construction des modèles. Le logiciel SAS Version 9.1 (SAS Institute, 2003) a été utilisé pour calculer les IP95 pour les autres valeurs de DA et de DH. Le rapport entre les statistiques PRESS (somme des carrés des résidus pour chaque observation, calculés à l'aide d'un modèle sans cette observation) et SS_{Total} (somme des carrés totaux du modèle) obtenues de SigmaStat soustrait de 1 ($1 - PRESS/SS_{Total}$) a fourni un coefficient de capacité prédictive des relations par validation interne (Q^2_{INT}) qui permet de vérifier leur robustesse. Un Q^2_{INT} supérieur à 0,5 est une condition minimale nécessaire pour que la relation ait une capacité prédictive.

La procédure REG de SAS a été utilisée pour l'analyse de régression forcée à l'origine non pondérée.

Les analyses de régression logistique ont été réalisées à l'aide du logiciel SigmaStat selon la procédure décrite dans un article de Patoine (2011). Les IC95 des pourcentages de dépassement des critères de qualité de l'eau modélisés par régression logistique ont été déterminés à l'aide de la méthode d'ajustement de modèle après transformation des variables explicatives « *model refitting* », décrite par Sofroniou et Hutcheson (2002). Cette méthode fournit un intervalle de confiance à 95 % pour les prédictions des régressions logistiques pour un échantillon de grande taille. La probabilité des tests de Hosmer-Lemeshow (P_{HL}) et du Chi-carré de Pearson (P_{X^2}) de SigmaStat a été retenue pour déterminer la qualité globale de la calibration des relations obtenues. Une valeur de P_{HL} et du P_{X^2} supérieure à 5 % indique un bon ajustement du modèle aux données. L'aire sous la courbe ROC (*Receiver Operating Characteristic*) (A_{ROC}) a aussi été calculée, avec le logiciel TANAGRA (Rakotomalala, 2005), dans le but de vérifier la puissance des relations logistiques obtenues pour discriminer les cas dépassant le critère de qualité de l'eau de ceux respectant le critère. Une valeur de A_{ROC} égale à 1 indique un modèle parfaitement discriminant et une valeur de 0,5, un modèle sans capacité prédictive.

D'autres types de méthodes de régression (modèle à données répétées avec structure d'autocorrélation ajusté par maximum de vraisemblance, modèle pour données censurées) ont été utilisés pour les modèles 2001 dans le but de vérifier les hypothèses d'indépendance temporelle et spatiale des données de CF (annexe 1).

1.4.2 Validation externe des modèles

Quatorze stations spatialement indépendantes de celles utilisées pour la construction des modèles ont été retenues pour la validation externe des modèles, de façon à ce qu'on puisse vérifier leur capacité à prédire les valeurs de CF dans d'autres bassins versants.

Dans le cas des modèles de prédiction de la médiane de CF, obtenus par la méthode des moindres carrés, le coefficient de validation externe (Q^2_{EXT}) proposé par Consonni et autres (2010) a été utilisé pour vérifier la qualité prédictive des

modèles de régression pour des bassins versants différents de ceux ayant servi à leur développement. Ce coefficient a été retenu parce qu'il peut être calculé pour une ou plusieurs stations. Il est obtenu par le calcul du rapport entre $PRESSv/Nv$ (somme des carrés des résidus des stations de validation sur leur nombre) et SS_{Total}/N (somme des carrés totaux des stations de construction du modèle sur leur nombre) soustrait de 1 $[1 - (PRESSv/Nv)/(SS_{Total}/N)]$. Une valeur du Q^2_{EXT} supérieure ou égale à 0,5 indique un modèle ayant une capacité prédictive externe acceptable. Lorsque plusieurs stations sont disponibles pour la validation, la valeur moyenne de leur Q^2_{EXT} individuel devrait être supérieure à 0,5 pour que le modèle puisse être utilisé à des fins de prédiction dans d'autres bassins versants.

Une comparaison graphique des résultats des modèles de prédiction de la médiane de CF et du pourcentage de dépassement de critère de CF aux valeurs mesurées aux stations de qualité d'eau a également été réalisée pour les bassins versants ayant servi à la construction et à la validation de ces modèles. Dans le cas des modèles de prédiction des dépassements des critères de CF, une faible corrélation de Pearson ($r_{Pearson} < 0,40$) entre les valeurs prédites et celles mesurées aux stations retenues pour la validation des modèles a été considérée comme indicatrice d'une absence de capacité de prédiction externe.

1.4.3 Identification des stations influencées

Le test de détection des valeurs anormales de Grubb permet de vérifier si une observation fait partie ou non de la même population statistique qu'un groupe d'observations dont la distribution suit une loi normale. Ce test a été utilisé pour comparer les résidus des modèles de prédiction des concentrations médianes de CF développés à l'aide de la méthode des moindres carrés, qui ont passé le test de normalité de Shapiro-Wilk, au résidu d'une station n'ayant pas servi à la construction des modèles. La station qui ne faisait pas partie de la même population statistique était considérée comme influencée de façon anormale par d'autres facteurs que la DA et la DH.

Dans le cas de la régression logistique, qui ne requiert pas le postulat de normalité, le test des résidus du modèle reliant le dépassement du seuil de 1 000 UFC/100 ml à la DA et à la DH ayant passé le test de Shapiro-Wilk, l'approximation normale a été acceptée et le test de Grubb a également été utilisé.

Pour chaque modèle, on a déterminé le seuil délimitant une valeur anormale pour un test de Grubb bilatéral selon un niveau de probabilité de 5 % (logiciel GraphPad QuickCalcs : outlier calculator, 2012, disponible à <http://graphpad.com/quickcalcs/Grubbs1.cfm>) en ajoutant par essais et erreurs une quinzième valeur aux résidus des 14 stations utilisées pour la construction du modèle. Ce seuil est fixe sur l'ensemble de la distribution des valeurs de DA et de DH.

La différence entre la valeur des CF mesurée aux stations de qualité de l'eau influencées et la valeur des CF calculée à l'aide du modèle pour les valeurs de DA et de DH dans le bassin versant drainé à ces stations a ensuite été comparée au seuil de valeur anormale. Cette approche comporte l'avantage d'utiliser un test de détection des valeurs anormales standardisé, en plus d'être simple d'utilisation.

2 RÉSULTATS ET DISCUSSION

2.1 Relations entre les coliformes fécaux dans les cours d'eau et les activités dans les bassins versants

Les deux premières parties de cette section présentent les nouveaux modèles de prédiction des médianes estivales de CF ou du dépassement de critère de qualité de l'eau, nommés « modèles 2006 », construits à l'aide des données de CF mesurées toutes les deux semaines les plus récentes à 14 stations et des variables de territoire DA et DH ou de nouvelles variables basées sur certaines pratiques agricoles. Les équations ainsi que la comparaison des valeurs de CF modélisées aux valeurs mesurées et la représentation graphique des équations dans les limites de leur domaine d'application sont présentées aux figures 2 à 7. La troisième partie présente la validation des relations à l'aide de 14 stations non utilisées pour leur développement ainsi que le domaine d'application de relations développées.

2.1.1 Modèles 2006 basés sur la densité animale et humaine

Trois modèles 2006 significatifs liant la DA, la DH et la médiane estivale de CF, obtenus à partir des données des 14 petits bassins versants, sont présentés à l'annexe 2 (figure A2.1, figure A2.2 et tableau A2.1). Les informations sur les variables utilisées pour leur développement et sur les stations et les périodes de mesure des CF y sont aussi présentées (tableau A2.3). Les trois modèles présentent un pouvoir de prédiction interne ($Q^2_{INT} > 0,5$), mais l'équation A (figure 2) présente la formulation la plus simple. Elle explique 83 % de la variance des médianes estivales de CF ($R^2 = 0,83$), les variables DA et DH contribuant à expliquer respectivement 77 % et 6 % de cette variance. L'intervalle de confiance à 95 % autour

de la droite de régression de cette équation varie de ± 84 à ± 235 UFC/100 ml, et l'intervalle de prédiction à 95 % varie de ± 311 à ± 380 UFC/100 ml, aux valeurs de DA et de DH des 14 bassins versants (tableau A2.1).

On peut simplifier l'équation A en arrondissant à la dizaine près l'ordonnée à l'origine, puisque sa valeur de -2 (erreur type : ± 68) UFC/100 ml est inférieure à la limite de détection (LD) de la méthode analytique des CF (10 UFC/100 ml), n'est pas significativement différente de zéro ($p = 0,9758$) et n'a pas de signification physique parce qu'elle est négative. Les sources autres que la DA et la DH sont ainsi considérées comme négligeables dans ce modèle, qui devient :

$$CF_{\text{(médiane estivale)}} = 2,0 (\pm 0,3) * DA + 4,0 (\pm 1,8) * DH + 0 (\pm 70) \text{ (Modèle 2006A simplifié)}$$

L'ordonnée à l'origine négative de l'équation A peut aussi être remplacée par la moitié de la LD de la méthode analytique des CF (5 UFC/100 ml), à l'aide d'une régression forcée à l'origine de (CF - 5) en fonction de DA et de DH. Cela suppose que les médianes estivales de CF dans les petits bassins versants sans élevage ni habitation sont sous la LD, ce qui implique des apports de CF de la faune faibles et occasionnels. Le modèle obtenu peut s'écrire comme suit :

$$CF_{\text{(médiane estivale)}} = 1,98 (\pm 0,23) * DA + 3,89 (\pm 1,20) * DH + 5 \text{ (Modèle 2006A forcé à LD/2)}$$

Ce modèle est très significatif ($p < 0,0001$) ainsi que ses variables DA ($p < 0,0001$) et DH ($p < 0,0072$) et il explique 92 % de la variance des valeurs de CF ($R^2 = 0,92$). L'intervalle de confiance à 95 % autour de la droite de régression de cette équation varie de ± 5 à ± 215 UFC/100 ml et l'intervalle de prédiction à 95 % varie de ± 284 à ± 356 UFC/100 ml, aux valeurs de DA et de DH des 14 bassins versants. Les valeurs médianes prédites à l'aide du modèle forcé à LD/2 sont pratiquement identiques à celles du modèle non forcé (résultat comparatif non présenté), mais les intervalles de confiance et de prédiction sont plus faibles (figure A2.2).

Trois modèles liant la DA et la DH au pourcentage de dépassement du critère de qualité de l'eau du MDDEFP (2013) pour les activités de contact indirect (1 000 UFC/100 ml) et les activités de contact direct (200 UFC/100 ml) ainsi qu'au dépassement de la recommandation du CCME (2005) pour l'irrigation des cultures (100 UFC/100 ml) (modèles 2006RL) ont aussi été obtenus par régression logistique (tableau A2.2 et figure 5). Le premier modèle est bien ajusté aux données ($P_{HL} > 0,05$ et $P_{X^2} > 0,05$) et présente une capacité prédictive acceptable ($A_{ROC} > 0,5$) du dépassement du critère de qualité de l'eau pour les activités de contact indirect comme la pêche. Les deux autres modèles de régression logistique ont montré un moins bon ajustement aux données ($P_{HL} < 0,05$; $P_{X^2} > 0,05$), mais une capacité prédictive interne acceptable ($A_{ROC} > 0,5$). L'intervalle de confiance à 95 % autour des pourcentages de dépassement modélisés varie de l'ordre de ± 3 % à ± 15 % selon la DA et la DH pour ces trois modèles (figure A2.4).

2.1.2 Modèles 2006 basés sur certaines pratiques agricoles

Dans le but d'établir un lien plus direct avec les pratiques agricoles, les deux relations suivantes entre la médiane estivale des CF et deux variables disponibles ont également été développées :

$$CF_{\text{(médiane estivale)}} = 6,53 (\pm 2,23) * S_{\text{dég}} + 7,07 (\pm 2,60) * DA_{\text{amas}} + 33 (\pm 76)$$

(Modèle 2006A')

$$CF_{\text{(médiane estivale)}} = 1,44 (\pm 0,24) * DA_{\text{entr}} + 8,05 (\pm 1,49) * DA_{\text{amas}} + 98 (\pm 36)$$

(Modèle 2006A'')

où $CF_{\text{(médiane estivale)}}$ est la concentration médiane estivale (de mai à octobre) des CF en UFC/100 ml, $S_{\text{dég}}$ la superficie du bassin versant cultivée ou en pâturage recevant des déjections animales n'ayant pas subi de traitement, en % ou en ha/km², DA_{amas} , la densité du cheptel dont les déjections sont entreposées en amas contigu au bâtiment d'élevage ou au champ, en UA/km², et DA_{entr} , la densité du cheptel dont les déjections sont entreposées autrement qu'en amas (entreposage étanche et cours d'exercice), obtenue de la différence entre DA et DA_{amas} .

La première relation (2006A') est significative (p du modèle $< 0,0005$; p des variables $< 0,05$) et permet d'expliquer plus des trois quarts de la variance des CF ($R^2 = 0,78$), les variables $S_{\text{dég}}$ et DA_{amas} contribuant à expliquer chacune 39 % de cette variance. Elle présente une capacité prédictive acceptable pour les 14 bassins versants ayant servi à son développement ($Q^2_{\text{INT}} = 0,57$). L'intervalle de confiance à 95 % autour de la droite de régression de cette équation varie de ± 98 à ± 339 UFC/100 ml, et l'intervalle de prédiction à 95 % varie de ± 360 à ± 485 UFC/100 ml, aux valeurs de $S_{\text{dég}}$ et de DA_{amas} des 14 bassins versants (tableau A2.1).

La seconde relation (2006A'') est la plus significative (p du modèle $< 0,0001$; p des variables $< 0,0005$) et permet d'expliquer les neuf dixièmes de la variance des CF ($R^2 = 0,90$), les variables DA_{entr} et DA_{amas} contribuant à expliquer respectivement 48 % et 42 % de cette variance. Elle présente une capacité prédictive acceptable pour les 14 bassins versants ayant servi à son développement ($Q^2_{\text{INT}} = 0,86$). L'intervalle de confiance à 95 % autour de la droite de régression de cette équation varie de ± 61 à ± 221 UFC/100 ml et l'intervalle de prédiction à 95 % varie de ± 234 à ± 316 UFC/100 ml, aux valeurs de $S_{\text{dég}}$ et de DA_{amas} des 14 bassins versants (tableau A2.1).

Trois modèles liant la $S_{\text{dég}}$ et la DA_{amas} au pourcentage de dépassement du critère de qualité de l'eau pour les activités de contact indirect et direct et l'irrigation des cultures (modèles 2006RL') ont aussi été obtenus par régression logistique (tableau A2.1 et figure A2.5). Ces modèles ont montré un ajustement aux données satisfaisant selon le test du Chi-carré de Pearson seulement ($P_{X^2} > 0,05$; $P_{HL} < 0,05$), mais une capacité prédictive interne acceptable ($A_{ROC} > 0,5$). L'intervalle de confiance à 95 % autour des pourcentages de dépassement modélisés varie de

l'ordre de $\pm 3\%$ à $\pm 15\%$ selon la $S_{\text{d}\acute{e}j}$ et la DA_{amas} (figures A2.6, A2.7 et A2.8).

Trois autres modèles liant la DA_{entr} et la DA_{amas} au pourcentage de dépassement du critère de qualité de l'eau pour les activités de contact indirect et direct et l'irrigation des cultures (modèles 2006RL'') ont été obtenus par régression logistique (tableau A2.2). Le premier modèle (figure A2.9) est bien ajusté aux données ($P_{HL} > 0,05$, $P_{X^2} > 0,05$) et présente une capacité prédictive acceptable ($A_{ROC} > 0,5$) du dépassement du critère de qualité de l'eau pour les activités de contact indirect comme la pêche. Le second modèle de régression logistique a montré un moins bon ajustement aux données ($P_{HL} < 0,05$; $P_{X^2} > 0,05$), mais une capacité prédictive interne acceptable ($A_{ROC} > 0,5$). Le troisième modèle a montré le moins bon ajustement aux données ($P_{HL} < 0,05$; $P_{X^2} = 0,05$), mais une capacité prédictive interne acceptable ($A_{ROC} > 0,5$). L'intervalle de confiance à 95 % autour des pourcentages de dépassement modélisés, déterminé pour le critère de contact indirect, varie de l'ordre de $\pm 3\%$ à $\pm 15\%$ selon la DA_{entr} et la DA_{amas} (figure A2.9).

La comparaison graphique des dépassements estivaux de critères de qualité de l'eau modélisés à l'aide de la régression logistique à ceux mesurés en rivière (figure A2.5) tend à confirmer que les modèles 2006RL' et 2006RL'' sont adéquats dans le cas des stations ayant servi à leur construction et peuvent y être utilisés pour prédire des dépassements de critères. Par contre, ces modèles retiennent moins adéquats dans le cas des stations retenues pour leur validation, notamment les modèles de prédiction du dépassement du critère de contact direct et d'irrigation.

L'examen des figures A2.2 et A2.5 et de l'annexe 2 suggère que les modèles 2006A et 2006RL basés sur la DA et la DH présentent un meilleur ajustement que les modèles 2006A' et 2006RL' en raison de leur caractère plus général. Ces modèles tiennent compte indirectement de divers facteurs liés aux apports de CF comme la quantité de déjections animales épandue, la densité d'accès

du bétail aux cours d'eau ou les rejets des résidences isolées. Les modèles 2006A' et 2006RL' utilisant les superficies avec épandages de déjections et la densité animale dont les déjections sont entreposées en amas permettent toutefois de modéliser adéquatement les CF en tenant compte de l'entreposage non étanche et des superficies réceptrices de déjections non traitées. Le modèle 2006A'' de prédiction de la médiane de CF , basé sur la densité animale dont les déjections ne sont pas entreposées en amas (DA_{entr}) et celle dont les déjections sont entreposées en amas (DA_{amas}), présente le meilleur ajustement. Parmi les modèles obtenus par régression logistique, le modèle 2006RL de prédiction des dépassements du critère pour les usages de contact indirect présente le meilleur ajustement, suivi par les modèles 2006RL'' et 2006RL'.

Les courbes représentant les médianes estivales de CF obtenues des modèles bivariés 2006A sont illustrées à la figure 3 et les intervalles de prédiction, à la figure A2.2. Dans le cas des modèles 2006A' et 2006A'', les médianes estivales modélisées sont illustrées à la figure 4 et les intervalles de prédiction, à la figure A2.3. Les courbes représentant les dépassements du critère de qualité de l'eau pour les activités de contact indirect et direct et l'irrigation (1 000, 200 et 100 UFC/100 ml) en fonction de la DA pour différentes valeurs de DH sont illustrées aux figures 6 et 7 et les intervalles de confiance, à la figure A2.4. Celles représentant les dépassements du critère pour les activités de contact indirect et les intervalles de confiance en fonction de la $S_{\text{d}\acute{e}j}$ ou de la DA_{entr} pour différentes valeurs de DA_{amas} sont illustrées aux figures A2.6 et A2.9. Dans le cas des critères pour les activités de contact direct et l'irrigation, les courbes des dépassements prédits et des intervalles de confiance en fonction de la $S_{\text{d}\acute{e}j}$ sont présentées aux figures A2.7 et A2.8. L'ensemble des courbes illustrées sont dans les limites du domaine joint (figure A2.10) des valeurs des variables explicatives ayant servi à construire les modèles 2006 bivariés.

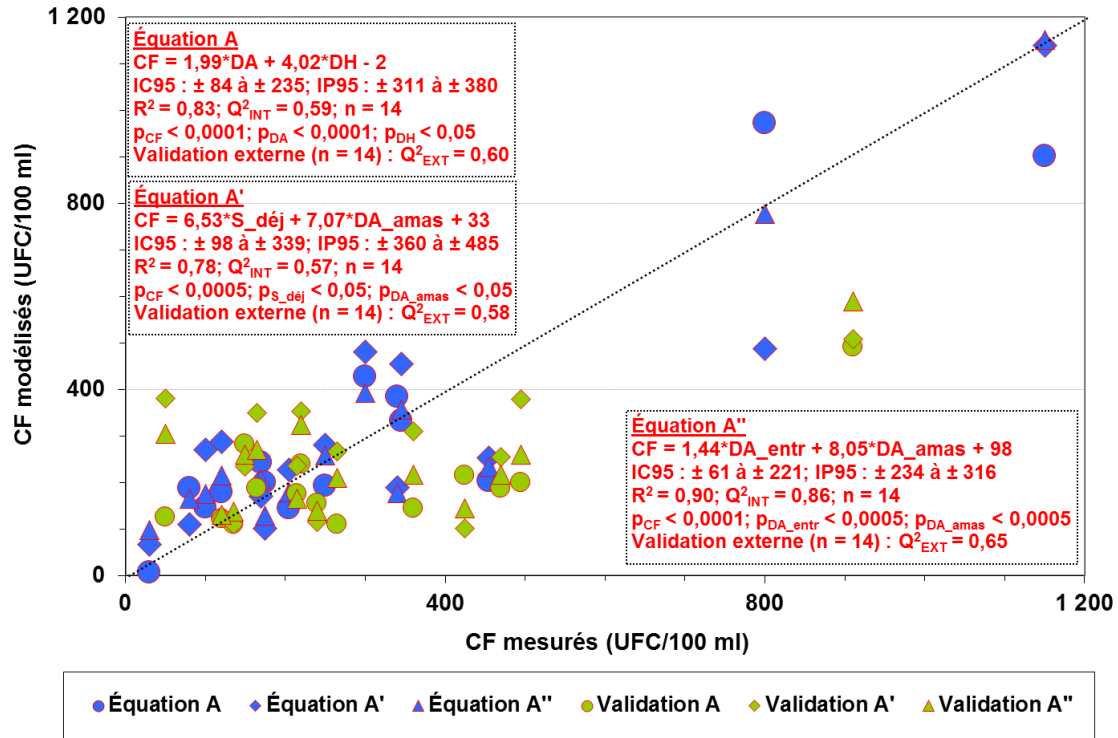


Figure 2 Médiannes estivales de CF mesurées et prédites à l'aide des DA et DH – modèles 2006

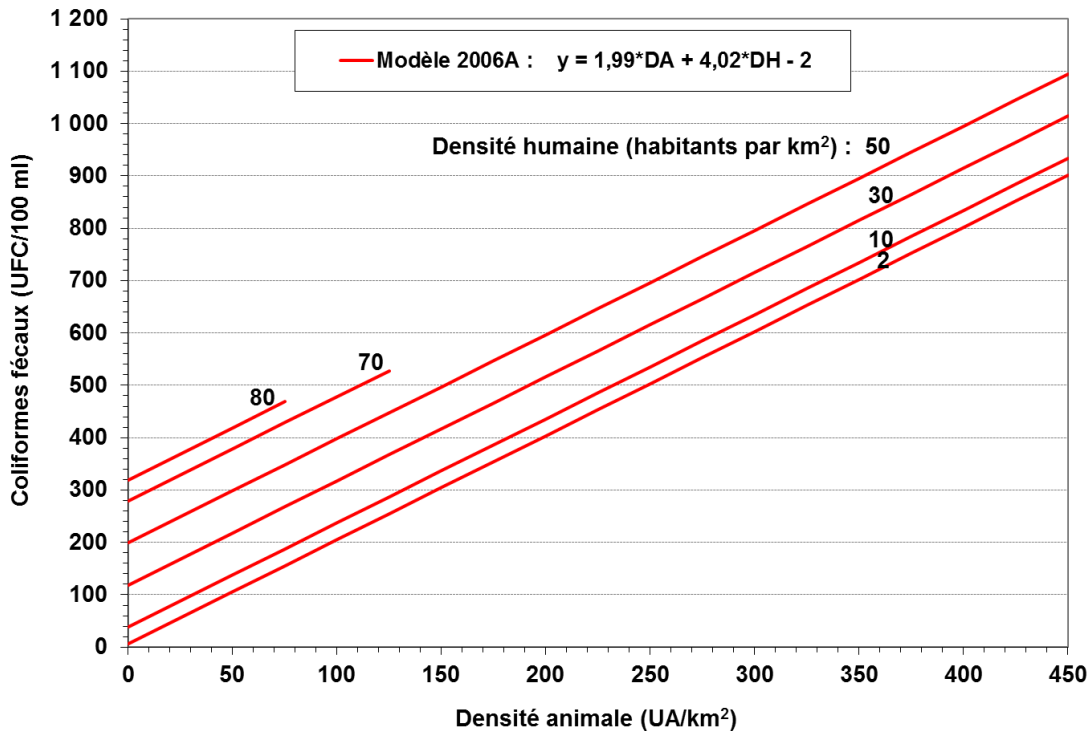
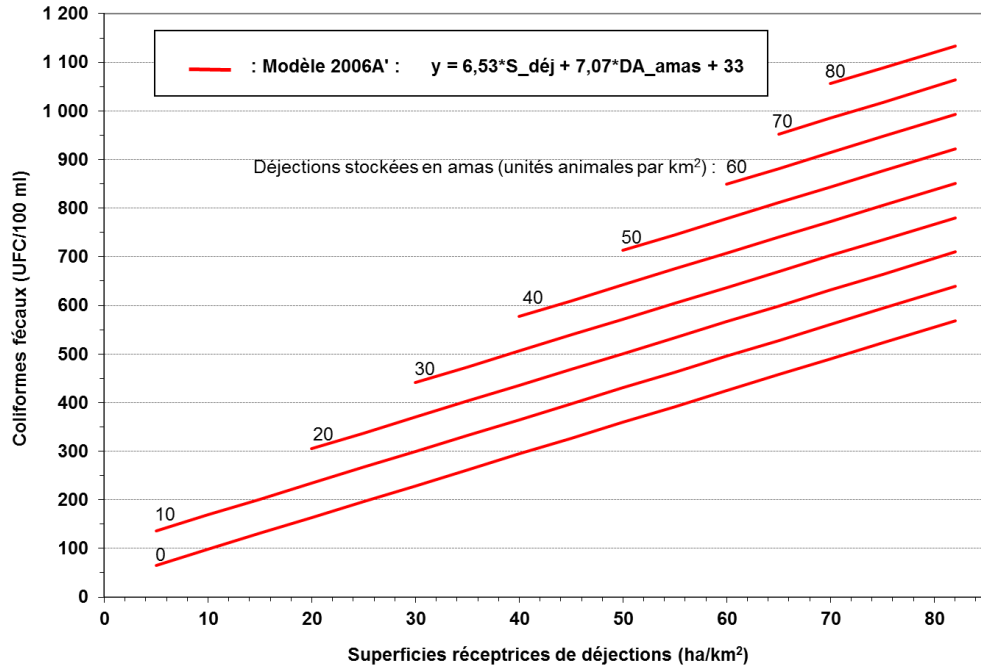
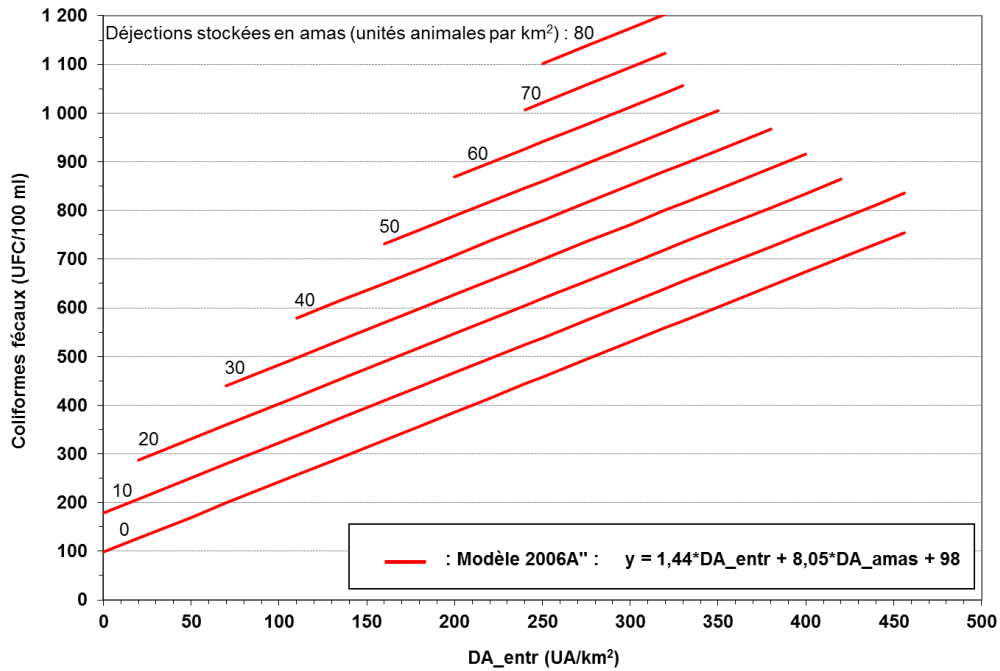


Figure 3 Médiannes estivales de CF en fonction de la DA et de la DH – modèle 2006A



a) Modèle 2006A' – CF médiane en fonction de la S_{déj} et de la DA_{amas}



b) Modèle 2006A'' – CF médiane en fonction de la DA_{entr} et de la DA_{amas}

Figure 4 Médiane estivale de CF en fonction de la S_{déj} ou de la DA_{entr} et de la DA_{amas}

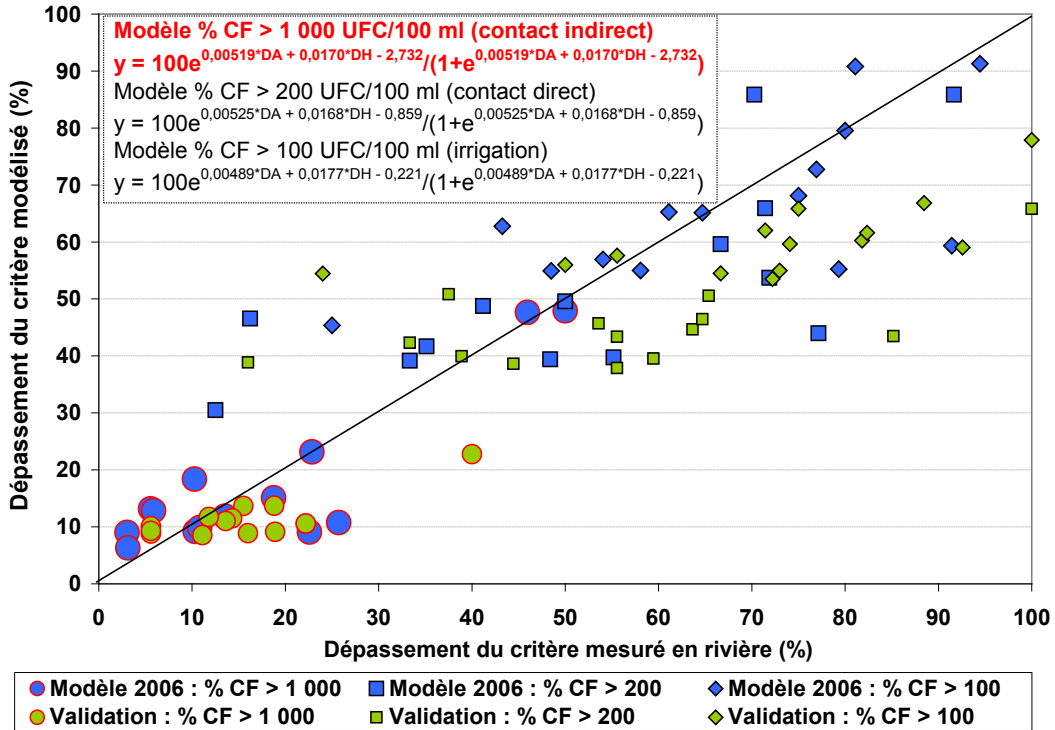


Figure 5 Dépassements de mai à octobre des critères de CF mesurés et prédits à l'aide de la DA et de la DH – modèles 2006RL

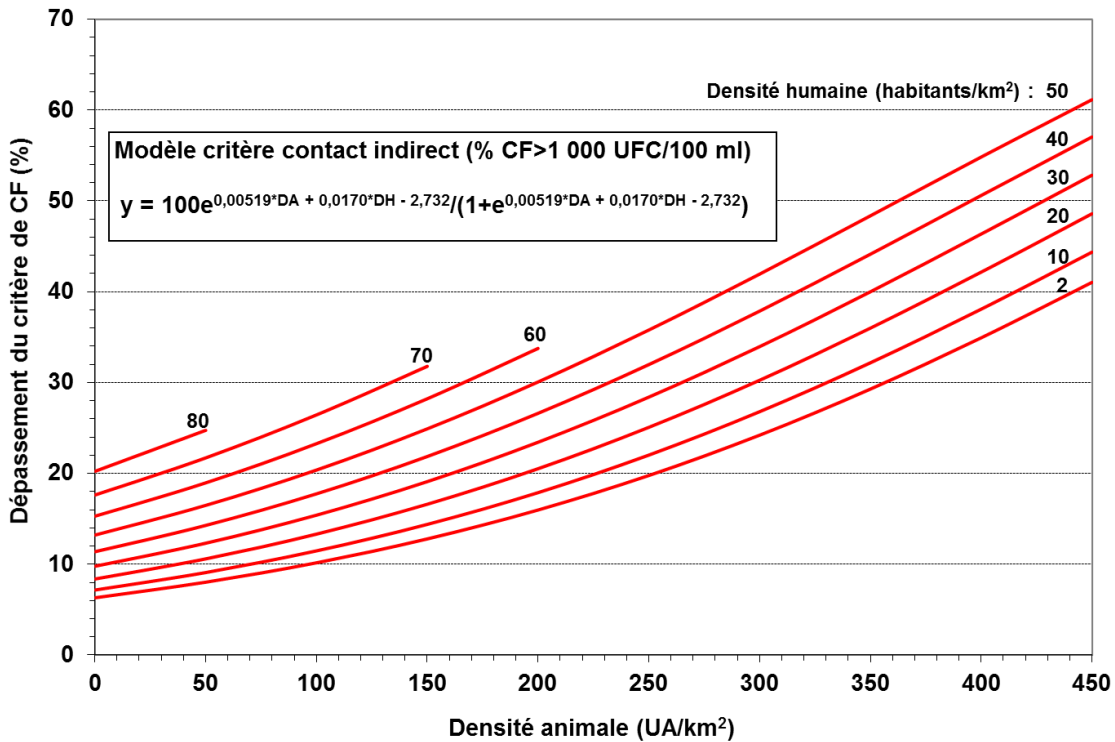
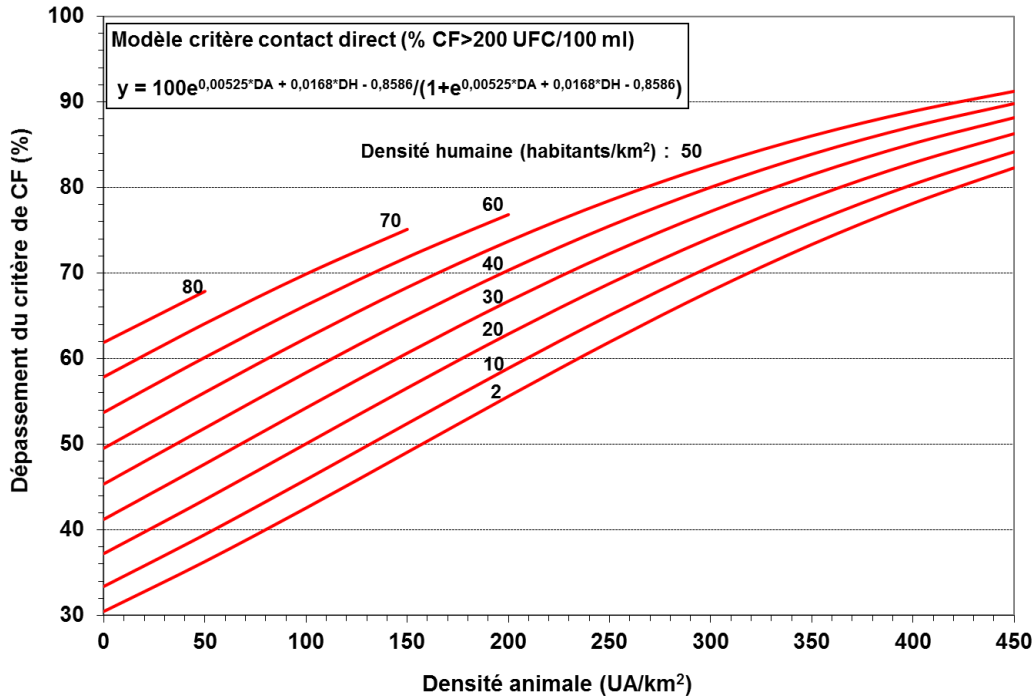
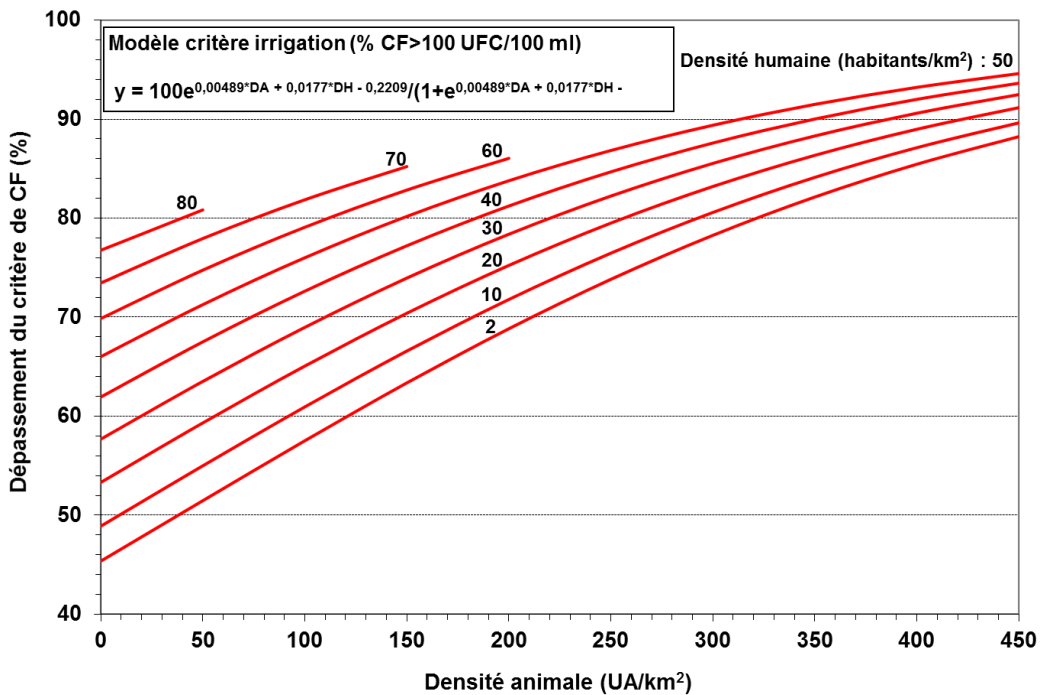


Figure 6 Dépassement du critère de qualité de l'eau de mai à octobre pour les activités de contact indirect en fonction de la DA et de la DH – modèle 2006RL



a) Critère pour les activités de contact direct (200 UFC/100 ml)



b) Critère pour l'irrigation des cultures (100 UFC/100 ml)

Figure 7 Dépassement du critère de qualité de l'eau de mai à octobre pour les activités de contact direct et l'irrigation en fonction de la DA et de la DH – modèle 2006RL

2.1.3 Validation externe et domaine d'application des modèles

Quatorze stations spatialement indépendantes de celles utilisées pour la construction des modèles ont été retenues pour leur validation externe (annexe 3). Ces stations présentent des données moins précises (mesures de CF toutes les deux semaines pendant deux années ou mensuelles pendant trois années; évaluation de la DA par la répartition des données municipales selon le territoire dans le bassin versant). Cette imprécision peut affecter le coefficient de validation externe, mais non les modèles, construits à l'aide de données plus précises. Les 14 stations sélectionnées pour la validation externe des modèles ne sont soumises à aucun rejet non traité d'eaux usées municipales, et leur valeur de CF n'est pas anormale selon le test de Grubb bilatéral au seuil de 5 % appliqué sur les résidus (différence entre la valeur modélisée et celle mesurée) (tableau A4.2).

L'annexe 3 présente le coefficient de validation externe des modèles de prédiction de la médiane de CF ayant une capacité prédictive dans d'autres bassins versants que ceux utilisés pour leur construction, c'est-à-dire les modèles pour lesquels le Q^2_{EXT} moyen des 14 stations de validation est supérieur à 0,5. Les modèles 2006A et 2006A' (équations 2006A et 2006A' à la figure 2) ont tous deux une capacité prédictive acceptable des médianes de CF d'autres petits bassins versants, leur coefficient de validation externe moyen étant supérieur à 0,5 ($Q^2_{EXT} = 0,60$ et $0,58$ respectivement). Le modèle 2006A'' présente la meilleure capacité prédictive ($Q^2_{EXT} = 0,65$). Le forçage à l'origine du modèle 2006A a peu d'effet sur sa capacité prédictive externe ($Q^2_{EXT} = 0,61$). La comparaison graphique des médianes estivales de CF modélisées à l'aide des équations 2006A, 2006A' et 2006A'' à celles mesurées en rivière (figure 2) illustre l'adéquation des modèles dans le cas des stations ayant servi à leur construction et à leur validation. Les modèles 2001 de prédiction des médianes estivales de CF (figures A1.1a et A1.3) présentent également une capacité prédictive externe acceptable pour les cas univariés ($Q^2_{EXT} = 0,57$) et bivariés ($Q^2_{EXT} = 0,58$).

Les équations 2006 de prédiction de la médiane de CF à l'aide de la DA et de la DH donnent des valeurs semblables, mais les IP sont plus petits dans le cas du modèle forcé à l'origine en raison de l'IC supposé nul à LD/2. Cette hypothèse, bien que plausible, dépend de la densité faunique dans les

petits bassins versants étudiés et du lieu de fréquentation des cours d'eau par la faune. Elle n'a pu être validée en raison de l'absence de bassins sans élevage ni habitation.

L'examen des valeurs individuelles de Q^2_{EXT} dans le tableau A3 montre que la station numérotée 43 (rivière à la Barbue), qui comporte l'agriculture la plus intensive parmi les 14 stations de validation, affecte à la baisse le Q^2_{EXT} moyen des modèles 2006 de façon importante, et dans une moindre mesure celui des modèles 2001. Sa meilleure adéquation aux modèles 2001 suggère que les pratiques agricoles dans son bassin versant pourraient être plus comparables à celles qui prévalaient vers l'an 2000 dans le groupe de bassins versants utilisés pour la construction des modèles 2001. L'examen de ses résidus pour le modèle 2006A à l'aide du test de Grubb (tableau A4.2) montre que la valeur médiane de CF de cette station n'est pas anormale au seuil standard de 5 %, mais serait anormalement élevée selon un seuil de 10 %. Ainsi, l'influence d'une source de contamination anormale d'origine indéterminée pourrait aussi être responsable de sa moins bonne adéquation aux modèles 2006A et 2006A'.

La comparaison graphique des dépassements estivaux de critères de qualité de l'eau modélisés à l'aide de la régression logistique à ceux mesurés en rivière (figure A1.2a, figure 5 et figure A2.5) montre dans l'ensemble une meilleure adéquation des modèles 2006RL, 2006RL', 2006RL'' et 2001RL pour le critère de contact indirect (1 000 UFC/100 ml), dans le cas des stations ayant servi à leur construction et à leur validation. L'adéquation est moins bonne pour les activités de contact direct et l'irrigation. Toutefois, les modèles basés sur la DA et la DH présentent suffisamment bien les dépassements de CF pour être employés en dehors des bassins versants utilisés pour leur construction (figure A1.2a et figure 5). Dans le cas des modèles de prédiction des dépassements de critère basés sur la DA_amas et la S_déj ou la DA_entr (tableau A2.2), l'ajustement aux données mesurées (figure A2.5) n'est pas assez bon pour qu'ils soient utilisés à des fins de prédiction externe pour les activités de contact direct et l'irrigation, comme le suggère le faible coefficient de corrélation entre les valeurs modélisées et mesurées ($r_{\text{Pearson}} < 0,40$).

Les relations développées permettent de prédire des valeurs médianes ou des dépassements de critères de qualité de l'eau de surface pour les CF basés sur une période de trois années.

Les relations bivariées avec la DA et la DH (modèles 2006 et 2001) ou avec la S_déj et la DA_amas (modèle 2006A') ou encore avec la DA_entr et la DA_amas (modèle 2006A'') sont valides dans le domaine joint des deux variables explicatives (figures A1.5 et A2.10). Les relations univariées avec la DA obtenues à l'aide des données de 2000 à 2002 (modèles 2001) sont valides dans la gamme des valeurs de 10 à 520 UA/km² observées dans les bassins versants numérotés de 1 à 10, lorsque la DH se situe dans la gamme de 9 à 51 habitants/km².

Également, les relations obtenues sont représentatives des groupes de petits bassins versants du Québec méridional (figure 1) comparables à ceux ayant servi à leur construction, qui représentent une agriculture surtout intensive. L'annexe 6 (figure 6.1) illustre la représentation plus importante des petits bassins versants comportant une densité animale élevée. La représentation des densités humaines des bassins versants se rapproche de celle de l'ensemble du Québec méridional (figure A6.2).

Les modèles sont applicables dans la gamme de superficies des petits bassins versants ayant servi à leur construction (de 14 à 98 km²). Leur validation externe pour la période estivale (annexe 3) montre que les relations ne peuvent être utilisées pour les bassins versants de moins de 10 km², mais présentent un potentiel pour ceux de 100 à 200 km². Elle suggère de plus que ces relations peuvent servir à prédire les médianes de CF d'autres bassins versants que ceux ayant servi à leur construction et pour des périodes plus récentes.

Les modèles 2006 tiennent compte d'améliorations dans certaines pratiques agricoles, notamment le retrait du bétail des cours d'eau (annexe 2 de Patoine et D'Auteuil-Potvin, 2015), alors que les modèles 2001 basés sur la DA permettent de prédire les CF dans les territoires ayant des pratiques agricoles (stockage et épandage des déjections animales, pâturage, etc.) comparables à celles des années 2000 ou favorisant davantage la contamination bactériologique.

La capacité prédictive externe des modèles 2006 et 2001 suggère que le groupe de 14 bassins versants retenus pour la validation externe, dont la densité animale est comprise entre 9 et 179 UA/km², comporte des pratiques agricoles qui se situent dans la gamme de celles observées en 2006 et en 2001. Elle suggère aussi que ces modèles demeurent représentatifs des conditions actuelles, compte tenu de l'absence de tendance significative dans les concentrations estivales de CF à la plupart des stations en milieu agricole (Patoine et D'Auteuil-Potvin, 2015) et de la gamme de densité animale généralement rencontrée.

La différence plus importante entre les modèles 2006 et 2001 pour les fortes densités animales peut s'expliquer notamment par la tendance à la baisse des valeurs estivales de CF observée à la station numérotée 1 utilisée pour leur développement, qui comporte la plus forte DA, et dans une moindre mesure, par la forme des modèles (transformation des variables). Le nombre limité de bassins versants comportant des DA supérieures à 250 UA/km² peut entraîner une certaine imprécision des modèles pour les plus fortes densités animales. Toutefois, les tests d'influence des analyses de régression n'ont pas fait ressortir de problème pour ces bassins versants.

Par ailleurs, la précision des modèles 2006 bivariés est plus grande que celle des modèles 2001 en raison de l'utilisation de trois stations additionnelles pour leur construction et de l'évaluation plus précise des DH. L'utilisation des modèles 2006 à des fins de prédiction devrait donc être privilégiée dans les cas où les pratiques agricoles ont évolué normalement ou en l'absence d'information suffisante sur celles-ci, alors que l'utilisation des modèles 2001 devrait se limiter aux cas où l'assainissement est moins avancé. Dans ce cas, le modèle 2001 univarié devrait être préféré au modèle 2001 bivarié lorsque les DH sont dans la gamme de 9 à 51 habitants/km².

2.2 Utilisation des relations prédictives du risque de contamination

Cette section présente certaines utilisations des relations développées pour prédire le risque de contamination des petits cours d'eau agricoles ainsi que les limites d'utilisation et les balises d'interprétation des résultats. Les équations obtenues peuvent servir notamment à identifier les stations soumises à une influence anormale, lorsque des données de CF sont mesurées, ou à estimer la contribution relative des sources animales et humaines de CF dans un bassin versant. Elles permettent aussi de prédire les valeurs de CF d'un bassin versant non pourvu d'un suivi de la qualité de l'eau en fonction de l'occupation du territoire actuel ou projeté ou d'estimer l'effet d'une intervention d'assainissement ou d'un projet de développement. L'annexe 7 résume les utilisations possibles des modèles développés à des fins de prédiction ou de diagnostic et identifie les principaux utilisateurs potentiels. Seules les relations possédant une capacité de prédiction à l'extérieur des bassins versants ayant servi à leur construction y sont indiquées.

2.2.1 Identification des stations influencées

Les équations obtenues peuvent aider à identifier les stations de qualité d'eau influencées par des sources de contamination bactériologique anormales, compte tenu de la DA et de la DH dans le bassin versant. L'approche proposée consiste à comparer au seuil du test de détection des valeurs anormales (test de Grubb) la différence entre la valeur estivale de CF mesurée à une station de qualité d'eau et celle modélisée à l'aide des variables de territoire pour le bassin versant (résidus). Les valeurs dépassant ce seuil sont considérées comme anormales selon le test de Grubb, ce qui indique une influence anormale probable d'une autre source de contamination que les apports reliés aux variables du modèle dans le bassin versant. Aux stations soumises à une influence anormale, la valeur des résidus, à différentes périodes, ainsi que la valeur du seuil de Grubb, sont présentées dans le tableau 3 et à l'annexe 4. L'utilisation des modèles basés sur la DA et la DH présentant une capacité de prédiction externe (2006A et 2006RL) devrait être privilégiée. On pourrait aussi utiliser les modèles basés sur d'autres variables, en connaissant la valeur du seuil de test de Grubb, qui est respectivement de 537 et 349 UFC/100 ml pour les modèles 2006A' et 2006A'' et de 38 et 27 % pour les modèles 2006RL' et 2006RL'' (contact indirect).

Une fois les stations anormalement influencées identifiées, une recherche de l'origine de la contamination dans le bassin versant peut être entreprise. Cette recherche peut se faire à l'aide des informations disponibles dans les bases de données géoréférencées ou à l'aide d'observations sur le terrain prises de l'aval vers l'amont. Les influences anormales peuvent provenir notamment d'un rejet direct au cours d'eau de matières fécales en amont rapproché d'une station d'échantillonnage (résidence isolée sans installation septique ou bétail au cours d'eau), d'un rejet municipal non traité dans le bassin versant ou encore d'une zone urbaine comportant des branchements d'égouts domestiques sur une conduite pluviale. Dans certains cas où les bassins versants sont de faible superficie, une influence anormale peut s'expliquer par l'épandage de déjections provenant de lieux d'élevage situés dans un bassin versant voisin, en raison d'une configuration des terres particulière. D'autres facteurs comme la présence d'un lac ou encore une concentration élevée de certaines espèces fauniques comme le rat musqué ou le chevreuil en amont rapproché d'une station pourraient influencer les CF mesurés en rivière.

L'utilisation de cette approche est présentée à titre d'exemple et d'information comme aide à la décision pour les stations non retenues pour la construction et la validation des modèles parce que soumises à une influence particulière. Le tableau 3 indique, pour la période estivale, les stations dont les médianes de CF présentent un écart (résidus) anormal par rapport aux modèles 2006A et 2006RL (% CF > 1 000) ainsi que les petits bassins versants comportant des sources particulières de contamination bactériologique. L'écart de la médiane mesurée par rapport au modèle (résidus au tableau 3) permet aussi de quantifier l'ordre de grandeur de l'effet observé des sources particulières de contamination identifiées. Les influences particulières identifiées sont des traverses d'animaux ou des points d'accès du bétail au cours d'eau, des rejets de résidences isolées ou d'agglomérations ou un secteur urbain en amont rapproché de la station ou encore un rejet non traité d'un réseau d'égout municipal. À la station du Grand Bras, l'influence provient vraisemblablement de l'épandage des déjections de plusieurs lieux d'élevage non comptabilisés dans la DA de son bassin versant parce que situés dans le bassin voisin (Le Petit Bras). La situation du bassin versant du Grand Bras est particulière

puisqu'il contient seulement deux lieux d'élevage, mais la majeure partie des terres cultivées sont rattachées aux lieux d'élevage du bassin voisin. Pour certaines stations, la cause probable de l'influence demeure indéterminée. L'effet de l'enlèvement d'une source de contamination en amont immédiat d'une station d'échantillonnage ou d'un point d'un cours d'eau servant à un usage particulier pourrait aussi être estimé à l'aide des équations obtenues et des facteurs d'influence présentés dans le tableau 3.

Cet outil pourrait, entre autres, permettre de mieux orienter les priorités d'action dans le cadre de la gestion de l'eau sur le territoire. Son application est toutefois limitée aux bassins versants pour lesquels suffisamment de données estivales de CF sont disponibles, par exemple sur une base mensuelle pour une période minimale de trois ans.

Tableau 3 Stations soumises à des influences anormales

Station BQMA ^a	Coliformes fécaux (mai-octobre)			Cours d'eau	Résidus ^b			Influence particulière ^d	
	Période	Médiane (UFC/100 ml)	> 1 000 (%)		Modèles 2006		Fi ^c		
					A	RL			
02300004	2000-2002	4 200	85,7	35	Du Portage	3 953	73	17	RR, TA, ACE
05050098	2006-2008	3 050	85,0	40	Le Petit Bras	2 909	75	22	RR, écoulement d'amas de fumier
05220240	2000-2002	2 800	80,6	36	St-Pierre	2 426	57	7,5	ACE, SU
05220240	2005-2008	815	44,7	38	St-Pierre	399	19	2,0	SU
03010086	2006-2009	1 650	62,5	32	St-Zéphirin amont	1 498	53	11	SU, écoulement d'amas de fumier
03040195	2006-2008	1 400	62,2	37	À l'Ours	1 282	53	12	SU
03040073	2005-2007	1 300	58,8	17	Ewing	1 117	49	7,1	Indéterminée
06200004	2008-2009	1 200	50,0	26	Bibeau	1 069	41	9,2	Indéterminée
03030253	2006-2008	1 300	61,1	36	Des Aulnages	1 033	49	4,9	Indéterminée
03030237	2000-2002	1 350	61,8	34	Du Sud-Ouest	1 011	47	4,0	RNT 835 Sainte-Brigide-d'Iberville
03030237	2005-2008	1 300	56,3	16	Du Sud-Ouest	997	42	4,3	Indéterminée
02330029	2002-2005	1 200	55,9	34	St-Patrice	976	44	5,4	Groupe de chalets construits avant 1981
05050099	2006-2008	1 500	65,6	32	Du Moulin	917	24	2,6	SU
05220063	2000-2002	900	40,5	37	Du Point du Jour	593	22	2,9	RR
12020001	2006-2008	435	22,2	18	Dauphine	381	15	8,1	Indéterminée (influence : modèle B)
02340119	2006-2008	590	44,4	18	Laflamme	378	33	2,8	Indéterminée
02340091	2000-2002	700	42,4	33	Turmel	373	27	2,1	Indéterminée
05050097	2006-2008	355	16,7	36	Le Grand Bras	327	10	12,7	Déjections du bassin versant voisin
02400048	2006-2008	415	26,3	38	Bourbon	305	18	3,8	RNT 211 Sainte-Sophie-d'Halifax
03010046	2010-2012	305	16,7	18	St-Zéphirin	127	7	1,7	RNT 280 Saint-Zéphirin-de-Courval RNT 74 La Visitation-de-Yamaska
02190010	2006-2008	9	0,0	26	Petite riv. Neigette	-57	-8	0,1	Lac en amont de la station
Seuil selon le test de Grubb (standard : bilatéral 5 %)						± 464	± 26		
<i>Seuil selon le test de Grubb (bilatéral 10 %)</i>						<i>± 417</i>	<i>± 23</i>		
Unités des résidus exprimés en UFC/100 ml pour le modèle 2006A et en pourcentage des analyses de CF > 1 000 UFC/100 ml pour le modèle de régression logistique 2006RL.									

a : Numéro de station, Banque de données sur la qualité du milieu aquatique.

b : Les résidus correspondent à la différence entre les valeurs estivales de coliformes fécaux mesurées dans le cours d'eau et celles calculées à l'aide des modèles 2006A et 2006RL (% CF > 1 000). Les résidus anormaux selon le test de Gubb bilatéral en considérant un seuil standard de 5 % sont indiqués en caractère gras. Les densités animales et humaines utilisées pour le calcul des résidus proviennent du tableau 1 pour les périodes de 2005 à 2009 et de l'annexe 1 pour les périodes de 2000 à 2005, sauf pour le ruisseau Turmel (non présenté à

l'annexe 1), où les valeurs utilisées sont de 123 UA/km² et de 21 habitants/km².

c : Fi = Facteur d'influence selon le modèle A (valeur mesurée/valeur estimée par le modèle A).

d : RR = Rejet de résidence isolée; TA = Traverse d'animaux; ACE = Accès du bétail au cours d'eau; SU = Secteur urbain en amont rapproché de la station; RNT XXX= Rejet non traité d'un réseau d'égout municipal desservant XXX personnes.

N = Nombre de données utilisées pour calculer la médiane estivale de coliformes fécaux.

2.2.2 Estimation de la contribution des sources animales et humaines

Les équations obtenues ou leur représentation graphique aident à illustrer les problématiques de contamination bactériologique en lien avec les sources animales et humaines. Par exemple, l'équation du modèle 2006A (équation « Modèle 2006A simplifié » ou équation aux figures 2 et 3) montre qu'en période estivale, chaque UA et chaque habitant par kilomètre carré de bassin versant ajoutent respectivement environ 2,0 (erreur type : $\pm 0,3$) UFC/100 ml et 4,0 (erreur type : $\pm 1,8$) UFC/100 ml à la médiane de CF dans les cours d'eau étudiés. L'équation du modèle 2006A' à la figure 4 montre pour sa part qu'en période estivale, chaque hectare recevant des déjections et chaque UA avec entreposage en amas par kilomètre carré de bassin versant ajoutent respectivement environ 6,8 (erreur type : $\pm 2,2$) UFC/100 ml et 7,5 (erreur type : $\pm 2,5$) UFC/100 ml à la médiane de CF dans les cours d'eau étudiés. L'équation du modèle 2006A'' montre un ajout à la médiane de CF dans les cours d'eau étudiés d'environ 8,1 (erreur type : $\pm 1,5$) UFC/100 ml et 1,4 (erreur type : $\pm 0,2$) UFC/100 ml pour chaque UA par kilomètre carré de bassin versant avec stockage en amas et autrement qu'en amas (entreposage étanche et cours d'exercice) respectivement. Ce dernier résultat suggère que pour un nombre d'UA donné, la contribution à la médiane de CF des lieux d'élevage avec stockage en amas (y compris les pertes de CF au stockage et à l'épandage) est environ cinq fois plus élevée que celle des lieux d'élevage avec stockage autrement qu'en amas, dont l'apport provient essentiellement de l'épandage des déjections animales.

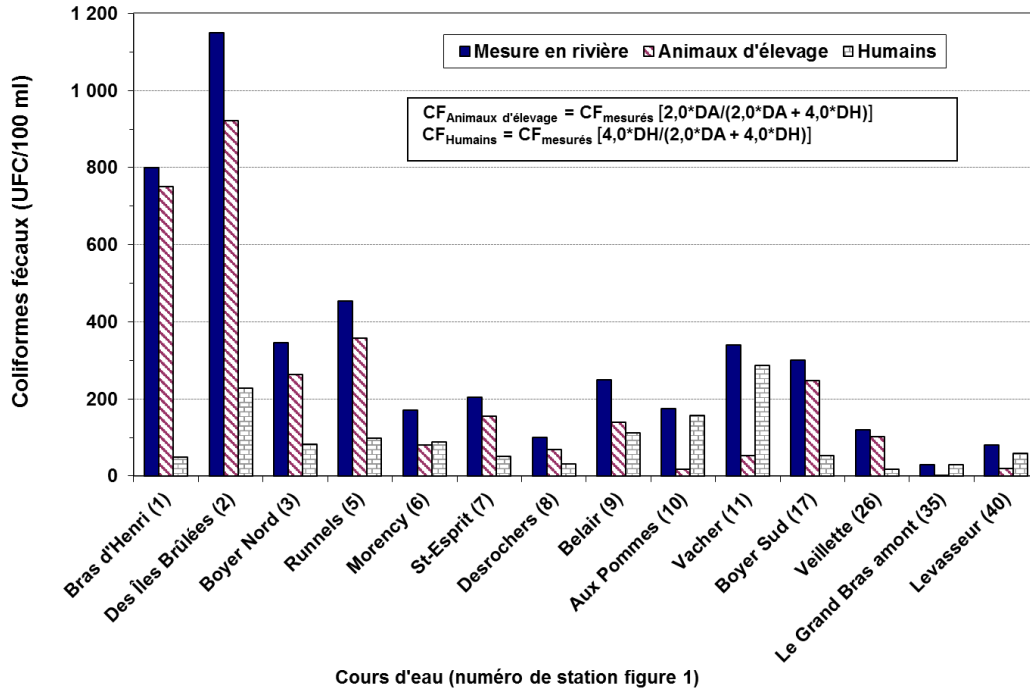
Les équations bivariées obtenues permettent également d'estimer, pour chaque cours d'eau, la proportion des CF qui proviennent des animaux d'élevage ou des humains. Le modèle 2006A ou sa forme simplifiée (modèle 2006A simplifié) sont plus faciles à utiliser à cette fin en raison de l'absence de transformation des données et de leur validité pour les valeurs de DA et de DH proches de zéro. Le modèle 2006A' permet pour sa part d'estimer la contribution des S_déj et de la DA_amas, la constante de l'équation 2006A' représentant la contribution des autres sources comme la DH et la faune. Deux approches peuvent être retenues.

La première approche consiste à calculer le rapport entre le terme correspondant à chaque source et la valeur de CF estimée à l'aide des équations. Ce

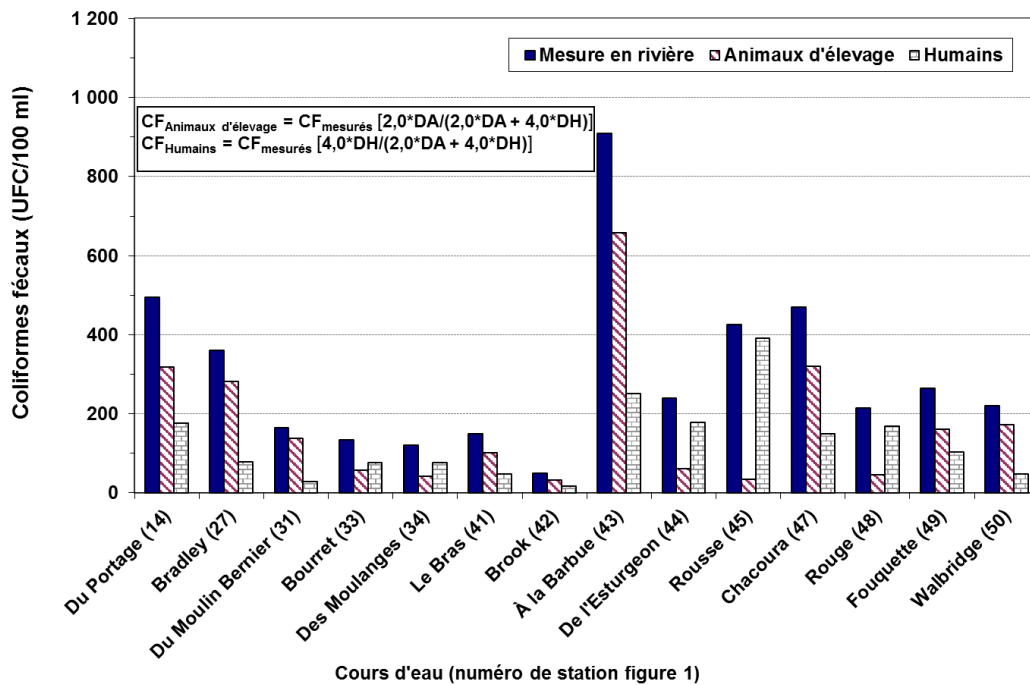
rapport peut être multiplié par la concentration de CF mesurée au cours d'eau pour l'obtention d'une contribution exprimée en UFC/100 ml. Cette approche considère que les erreurs résiduelles sont réparties proportionnellement à la contribution de chaque source. Par conséquent, elle ne peut être utilisée qu'aux stations non influencées par des sources anormales de contamination. Les concentrations médianes estivales de CF mesurées dans les 28 cours d'eau ayant servi à la construction et à la validation des modèles ainsi que les contributions des différentes sources estimées à l'aide de cette approche sont présentées aux figures 8 et 9.

La seconde approche consiste à calculer la contribution de chaque source à l'aide des coefficients des modèles et à illustrer la concentration résiduelle non expliquée par le modèle (annexe 8). La figure A8.1 illustre les contributions de la DA et de la DH et les résidus aux stations utilisées pour la construction et la validation des modèles. Dans ce cas, les résidus indiquent la contribution non expliquée par le modèle. Les contributions des S_déj et de la DA_amas peuvent également être représentées à l'aide du modèle 2006A' (figure A8.2). La figure A8.3 illustre les contributions des animaux d'élevage et des humains obtenues selon cette approche à l'aide du modèle 2006A simplifié, dans le cas des stations influencées. Le dépassement des résidus par rapport au seuil de Grubb fournit une indication de la contribution des sources anormales compte tenu des valeurs de DA et de DH dans le bassin versant.

Dans une analyse de stations influencées, cette seconde approche permet de comparer l'importance des sources anormales de contamination (dépassement du seuil de Grubb par les résidus) à la contribution estimée des élevages d'animaux et des humains dans un bassin versant. Cette information peut aider à prioriser les actions en fonction de l'effet des différentes sources sur la concentration de CF. La figure A8.3 montre qu'en amont des stations des rivières du Sud-Ouest et Saint-Zéphirin amont et des ruisseaux Saint-Patrice, Des Aulnages, à l'Ours, Le Petit Bras, Bibeau et Ewing, les sources anormales de contamination devraient être identifiées et éliminées d'abord en raison de leur importance et des bénéfices attendus. Cette approche montre en outre l'importance de porter une attention particulière aux sources de contamination en amont rapproché des usages de l'eau.

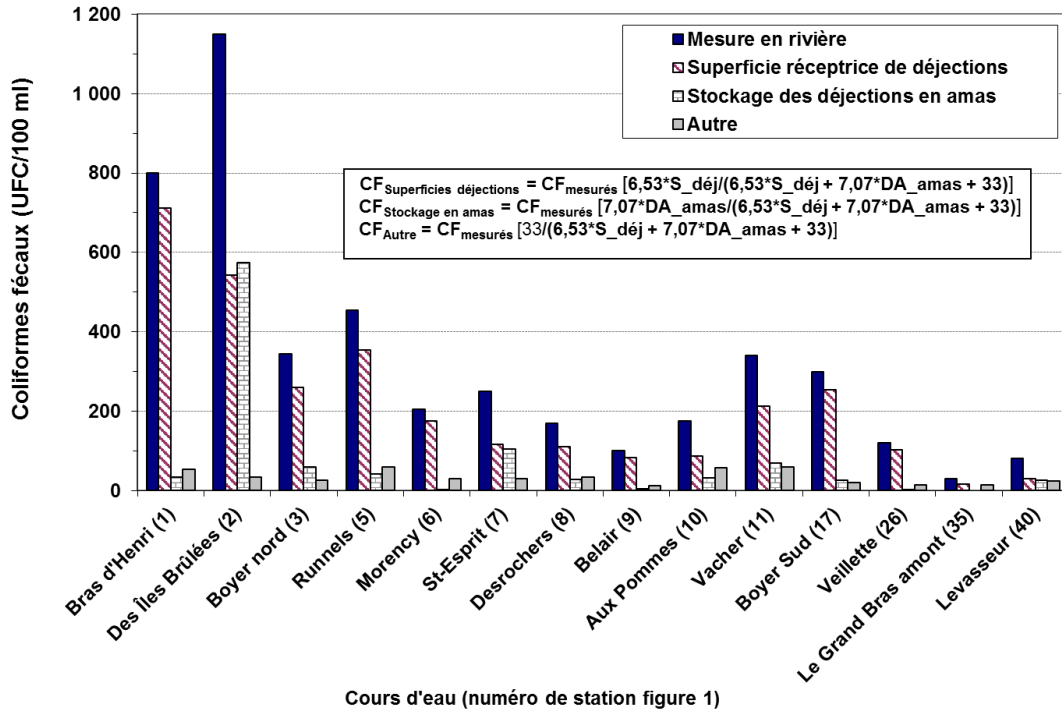


a) Stations utilisées pour la construction des modèles 2006

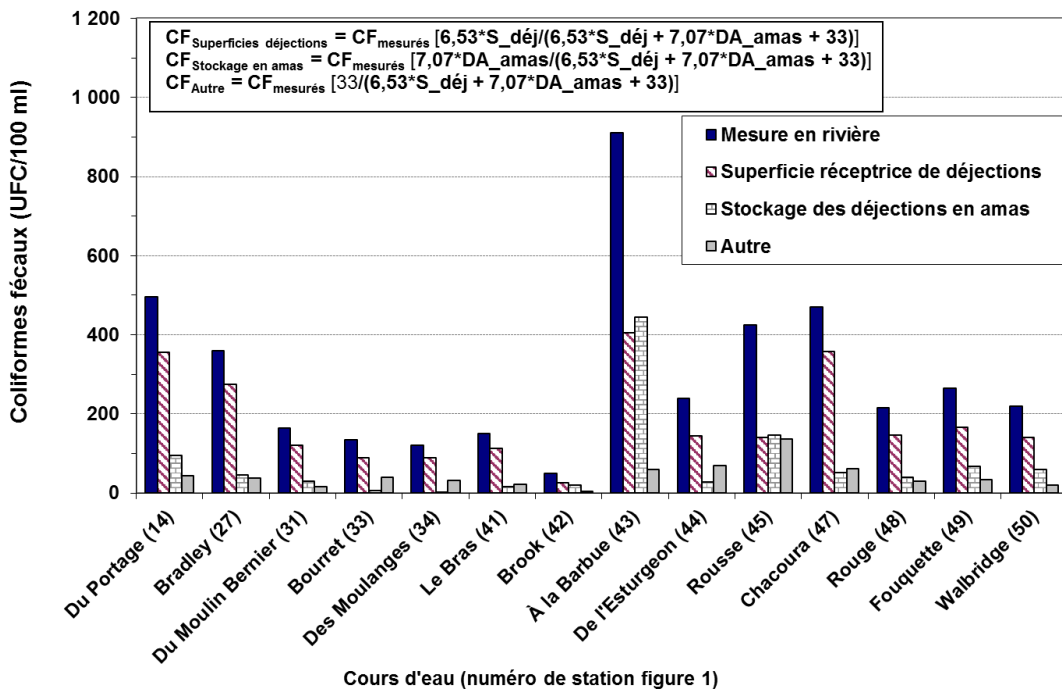


b) Stations utilisées pour la validation des modèles 2006

Figure 8 Concentrations médianes estivales de CF mesurées à 28 stations et contribution estimée des animaux d'élevage et des humains



a) Stations utilisées pour la construction des modèles 2006



b) Stations utilisées pour la validation des modèles 2006

Figure 9 Concentrations médianes estivales de CF mesurées à 28 stations et contribution estimée des superficies réceptrices de déjections et du stockage en amas

Les résultats des deux approches montrent en outre que la principale source de contamination varie selon le bassin versant, mais que la densité animale ou les superficies recevant des déjections animales contribuent généralement le plus aux valeurs médianes de CF aux stations non soumises à des influences anormales de la présente étude.

2.2.3 Estimation des valeurs de CF dans les bassins versants sans suivi de la qualité de l'eau et de l'effet des changements dans les bassins versants

Les modèles indiqués à l'annexe 7 permettent d'estimer les médianes estivales de CF et les dépassements de critères dans des bassins versants sans station de mesure de la qualité de l'eau. Ceux basés sur la DA et la DH permettent également d'estimer l'effet de certaines mesures appliquées à une partie des UA d'un petit bassin versant (comme le traitement de déjections animales), de l'ajout ou de la diminution du nombre d'animaux dans un petit bassin versant, ou encore d'un changement dans la densité des résidents non raccordés à un réseau d'égout. Celui basé sur la S_déj et la DA_amas permet aussi d'estimer l'effet d'un changement dans les superficies réceptrices de déjections animales ou dans le stockage en amas.

L'effet d'autres changements comme l'exportation de déjections animales hors du bassin ou la réception de déjections animales d'autres bassins pourrait aussi être estimé à l'aide de l'un ou l'autre des modèles, dans la mesure où le changement peut être exprimé sous la forme des variables d'entrée du modèle.

Toutefois, les estimations réalisées sont valables dans la mesure où la proportion des pratiques agricoles demeure dans la gamme de celle observée dans les bassins versants ayant servi à la construction du modèle utilisé. Ces pratiques agricoles peuvent être influencées par la proportion des types d'élevage présents dans le bassin versant (tableau 2).

Des projets de développement de l'agriculture ou des interventions d'assainissement dans certains petits bassins versants pourraient aussi changer les proportions de certaines pratiques agricoles et influencer le risque de contamination microbienne.

Les meilleurs modèles présentent en outre des possibilités intéressantes quant à la prédiction du risque de contamination bactériologique de l'eau de surface. Le modèle 2006A de calcul des médianes estivales de CF basé sur la densité animale et humaine pourrait être utilisé pour classer ou cartographier les petits bassins versants selon différents niveaux de risque. Dans le cas des bassins versants pourvus d'une station de mesure de la qualité de l'eau, ce modèle permet aussi d'identifier les stations influencées par une contamination fécale anormale et d'estimer la contribution relative des humains et des animaux d'élevage. Les modèles 2006RL pourraient également être utilisés pour classer ou cartographier les petits bassins versants selon les dépassements de critères de qualité de l'eau pour les usages de contact indirect et direct et pour l'irrigation. Les modèles 2006A' ou 2006A'' basés sur certaines pratiques agricoles (stockage en amas, superficie réceptrice de déjections animales) pourraient être utilisés pour estimer l'effet de l'entreposage étanche et du traitement des déjections animales.

Trois exemples d'utilisations potentielles des relations développées pour prédire le risque de contamination sont présentés à l'annexe 9; le premier estime les valeurs de CF dans un bassin versant sans station de mesure de la qualité de l'eau, le second estime l'effet de l'ajout du traitement des déjections animales et le troisième l'effet de l'ajout de cheptels dans un petit bassin versant. Ils permettent également d'illustrer la procédure de calcul et de comparer le résultat obtenu à l'aide des différents modèles.

Les relations permettant de prédire le pourcentage de dépassement du critère de qualité de l'eau pour une activité de contact indirect comme la pêche (figures 6 et A2.4 ou figures A2.6 ou A2.9 ou figure A1.2) présentent un intérêt particulier puisque la pêche dans les cours d'eau en milieu agricole constitue soit une activité fréquente, un usage à préserver ou encore un usage à récupérer qui sert souvent d'objectif mobilisateur pour que des actions collectives de dépollution soient entreprises. La prédiction de l'effet de l'entreposage étanche des amas de déjections animales sur les valeurs de CF présente également un intérêt à des fins de gestion de l'eau.

2.2.4 Limites des modèles et balises d'interprétation des résultats

Les modèles proposés ont été obtenus par régression entre certains indicateurs de la pression microbienne (DA et DH ou S_déj et DA_amas ou DA_entr et DA_amas) de petits bassins versants supportant une agriculture surtout intensive et les valeurs de CF dans les cours d'eau. Leur précision dépend de celle des indicateurs ayant servi à leur construction ou utilisés à des fins de prédiction. Pour la construction des modèles, les données de DA comparables les plus précises disponibles pour 2001 et 2006 ont été utilisées (cheptel par subdivision de recensement unifiée (SRU) de Statistique Canada 2001 et 2006 attribué au bassin versant à l'aide des données des lieux d'élevage de SAGO des années 2003 et suivantes). L'utilisation des données par SRU de Statistique Canada (annexe 5) devrait se limiter aux années 2001, 2006 et 2016, celles du recensement de l'agriculture de 2011 n'étant pas disponibles sur une base comparable. Les prédictions obtenues avec les modèles basés sur la DA pour les années postérieures à 2006 devraient se faire à l'aide des données du recensement de l'agriculture 2016 ou à l'aide de données de cheptel plus précises. Les données de SAGO du MELCC ou des fiches d'enregistrement des exploitations agricoles du MAPAQ présentent un potentiel à cette fin. Toutefois, la disponibilité de ces données et la capacité de traiter l'information par petit bassin versant pour des utilisateurs potentiels sont très limitées présentement.

Dans le cas de la S_déj, disponible à l'échelle municipale seulement, la répartition au prorata de la superficie totale dans le bassin versant aurait pu être légèrement améliorée par l'utilisation du prorata de la superficie cultivée. L'utilisation des modèles basés sur la S_déj devrait se limiter à l'évaluation de l'effet d'interventions d'assainissement ou à des prédictions pour les années 2001, 2006 ou 2016 parce la méthode de compilation par SRU des données permettant le calcul de cette variable a changé dans le Recensement sur l'agriculture de 2011 de Statistique Canada et qu'il n'existe pas d'autres bases de données permettant de la calculer. La DA_amas, établie à l'aide des informations répertoriées dans SAGO par lieu d'élevage, aurait pu être déterminée avec plus de précision à l'aide des données à la ferme, si elles avaient été disponibles. La détermination à l'aide de SAGO du nombre d'UA pour lesquelles les déjections sont entreposées en amas peut en effet être imprécise

dans les cas où l'amas ne provient pas de la totalité des cheptels d'un lieu d'élevage. L'utilisation des modèles basés sur la DA_amas à des fins de prédiction pour les années postérieures à 2006 est toutefois facilitée par la mise à jour régulière de cette base de données du MELCC.

Deux périodes ont été retenues pour la construction des modèles : une première voisine de 2001 (modèles 2001), une seconde voisine des données 2006 lorsque disponibles (modèles 2006). Les modèles 2006 prédisent des valeurs de CF généralement moins élevées que les modèles 2001 en raison notamment d'une amélioration, entre les deux périodes, de certaines pratiques agricoles non incluses dans les variables explicatives, comme l'accès du bétail aux cours d'eau. L'utilisateur doit être conscient de cette réalité dans l'interprétation des résultats qu'il obtient à l'aide des modèles proposés. En l'absence d'information suffisante sur les pratiques agricoles, l'utilisateur devrait privilégier les modèles 2006. Il doit également demeurer à l'intérieur du domaine d'application des modèles (section 2.1.3) pour faire une interprétation fiable des résultats. Les modèles présentant une capacité prédictive externe acceptable (annexe 7) peuvent être utilisés dans l'ensemble des petits bassins versants du Québec méridional (figure 1).

Cet outil de prédiction est cependant limité par le fait que les variables des modèles (DA et DH, S_déj et DA_amas ou DA_entr et DA_amas) ne représentent pas tous les aspects des sources anthropiques de contamination bactériologique des bassins versants. Plusieurs facteurs comme la position des sources anthropiques de contamination dans le bassin versant, le type d'élevage, le mode d'épandage des déjections animales ou l'accès du bétail au cours d'eau ne sont pas directement pris en compte et contribuent à l'imprécision ou à la variance non expliquée des modèles. La faune et les animaux de compagnie ne sont pas non plus pris en compte distinctement dans les modèles développés. Également, les épisodes hydrologiques, souvent associés à un plus fort risque (CAC, 2013), ne sont pas considérés. Finalement, le nombre limité de stations et la taille des bassins versants utilisés pour leur construction rendent les modèles plus sujets à des particularités locales et à des changements de pratiques. L'ensemble des facteurs non pris en compte contribuent à accroître l'amplitude des intervalles de prédiction des modèles.

L'utilisation des relations pour estimer le changement de CF à la suite d'une intervention consistant en une baisse ou en un ajout de DA ou de DH dans un bassin versant est valable dans la mesure où les conditions autres que la DA et la DH qui peuvent influencer les CF, comme la proportion de différents types d'élevage (tableau 2) ou de certaines pratiques agricoles, demeurent dans la gamme de celles du groupe de bassins ayant servi à leur développement et ne changent pas de manière importante à la suite de l'intervention.

Aucun des modèles développés n'est complet en raison du nombre limité de variables explicatives qu'ils comportent. Les modèles basés sur la DA et la DH ainsi que ceux basés sur la DA_amas et la S_déj ou la DA_entr peuvent toutefois être utilisés de façon complémentaire à des fins de prédiction pour une situation donnée. Le modèle basé sur la DA et la DH est plus général, mais celui basé sur la DA_amas et la S_déj ou la DA_entr, même s'il n'intègre pas la DH, permet de prédire l'effet de certaines pratiques agricoles comme le traitement des déjections animales ou leur exportation à l'extérieur d'un bassin versant ou encore l'effet de l'entreposage étanche des amas de déjections animales. Du fait que les modèles n'intègrent pas plus de deux variables à la fois, en raison notamment du nombre restreint de stations disponibles pour leur développement, l'effet prédit à l'aide d'une variable peut aussi inclure l'effet d'autres variables corrélées à celle-ci, non incluses dans le modèle.

L'utilisateur doit finalement être conscient des limites liées à la forme des modèles. Les modèles linéaires de prédiction des médianes de CF (2006A, 2006A' et 2006A'') ont une forme plus simple et fournissent des valeurs réalistes. Par contre, la borne inférieure de l'IP à 95 % de leur médiane prédite peut présenter une valeur négative dans certains cas. La borne inférieure de l'IP doit alors être considérée uniquement comme une estimation de l'incertitude de la médiane prédite puisqu'une valeur négative n'a pas de signification physique. Les modèles de prédiction des médianes utilisant une transformation logarithmique des données de CF (modèles 2001 et modèles 2006B et 2006C) ne présentent pas ce problème d'IP négatif. Par contre, la signification physique des coefficients de leurs variables est plus difficile à interpréter et leur ajustement aux données des stations ayant servi à leur construction et à leur validation externe (Q^{2INT} , Q^{2EXT}) est généralement moins bon. Les modèles 2006B et 2006C ont aussi montré une moins bonne

capacité prédictive des valeurs médianes de CF que les modèles linéaires.

Les modèles de prédiction des dépassements de critères de qualité de l'eau (2006RL, 2006RL', 2006RL'' et 2001RL) ont tendance à excéder les plus faibles valeurs de CF mesurées et à sous-estimer les plus fortes valeurs dans le cas des stations retenues pour la validation. La forme de la fonction utilisée (logit), qui ne s'ajuste pas toujours parfaitement aux données, pourrait contribuer à expliquer la différence. Une autre explication est la présence possible de sources anormales de contamination en amont des stations présentant des dépassements plus fréquents de certains seuils de CF. L'intérêt de ces modèles réside notamment dans leur capacité à prédire des dépassements de critères de qualité de l'eau avec des intervalles de confiance toujours positifs.

L'outil proposé présente un potentiel d'utilisation intéressant. Afin qu'il devienne pleinement opérationnel, d'autres étapes devraient être franchies, comme une démarche de transfert de l'information basée sur une identification plus précise des besoins des utilisateurs et des balises d'interprétation des résultats. Une telle démarche aiderait aussi à préciser les besoins de mise à jour des modèles. Le développement d'équations incluant davantage de paramètres pourrait être rendu possible à l'aide de mesures de bactéries indicatrices de la contamination fécale dans un plus grand nombre de bassins versants et de données plus complètes sur les pratiques agricoles et l'occupation du territoire. La construction de modèles à l'aide de données postérieures à l'élimination de certaines pratiques comme l'accès du bétail aux cours d'eau permettrait d'enlever leur effet potentiel sur les valeurs de CF prédites.

Par ailleurs, la détermination du risque de contamination à l'aide des CF ne prend pas en compte la présence réelle des organismes pathogènes et ne permet pas d'en déterminer directement l'origine. Les CF peuvent ne pas bien représenter certaines caractéristiques des différents organismes pathogènes, comme leur prédominance chez certaines espèces animales ou certains groupes d'âge, leur pathogénicité, leur caractère épisodique et leur survie.

Le suivi de certains organismes pathogènes dans les petits tributaires agricoles aiderait à améliorer grandement les outils de gestion du risque de contamination bactériologique. Finalement, l'utilisation d'autres approches complémentaires,

comme des méthodes de recherche des sources de contamination (Dufour et autres, 2012), contribuerait à mieux cibler les interventions d'assainissement.

Le potentiel d'utilisation des relations à de plus grands bassins versants pourrait aussi être

davantage exploré. Delpla et Rodriguez (2014) ont d'ailleurs conçu des modèles de régression liant les CF à certaines variables du territoire provenant surtout de grands bassins versants au Québec, qu'ils ont utilisé pour montrer une hausse prévisible des CF en rivière à la suite des changements futurs dans l'utilisation des terres et le climat.

CONCLUSION

Des relations entre les CF dans les cours d'eau et les activités agricoles et humaines ont été développées pour les petits bassins versants. Ces relations permettent d'estimer les concentrations médianes estivales de CF ou les dépassements des critères de qualité de l'eau pour les usages de contact indirect et direct et pour l'irrigation des cultures, à l'aide de variables descriptives des bassins versants comme la densité animale et humaine ou la superficie recevant des déjections animales et la densité animale dont les déjections sont stockées en amas. Elles constituent un outil potentiel qui permet, à l'aide d'indicateurs simples et disponibles, de mieux évaluer le risque de contamination bactériologique des cours d'eau. Ces relations pourraient contribuer à améliorer la gestion de ce type de contamination dans les bassins versants agricoles. Elles sont toutefois limitées par le fait que leurs variables explicatives ne représentent pas tous les aspects des sources agricoles et anthropiques de la contamination bactériologique.

D'autres étapes devraient en outre être franchies pour que cet outil devienne pleinement opérationnel, comme une démarche de transfert de l'information basée sur une identification plus précise des besoins des utilisateurs et des balises d'interprétation des résultats obtenus à l'aide des modèles. Une telle démarche aiderait également à préciser les besoins de mise à jour des modèles proposés. Le besoin de développement d'autres approches complémentaires, permettant d'estimer encore mieux ce risque, comme la mesure de microorganismes pathogènes, ou de mieux cibler les interventions, comme l'utilisation de méthodes de dépistage des sources de contamination microbienne, pourrait aussi être évalué.

RÉFÉRENCES BIBLIOGRAPHIQUES

- BORCARD, D. 2002. *Régression et corrélation multiples et partielles*, Université de Montréal, Département de sciences biologiques, [<http://biol09.biol.umontreal.ca/borcardd/r2p/artiel.pdf>].
- Conseil canadien des ministres de l'environnement (CCME), 2005. Recommandations canadiennes pour la qualité des eaux : protection des utilisations de l'eau à des fins agricoles – tableau sommaire, mis à jour en octobre 2005, dans *Recommandations canadiennes pour la qualité de l'environnement*, 1999, Winnipeg, Le Conseil, [<http://www.obvcapitale.org/wp-content/uploads/2012/07/CCME-2005-Crit%C3%A8res-protection-des-eaux-%C3%A0-des-fins-agricoles.pdf>].
- Centre d'expertise en analyse environnementale du Québec (CEAEQ), 2011. *Recherche et dénombrement des coliformes fécaux (thermotolérants) et confirmation à l'espèce Escherichia coli : méthode par filtration sur membrane*. MA. 700 – Fec.Ec 1.0, Rév. 3, 21 p., [<http://www.ceaeq.gouv.qc.ca/methodes/pdf/MA700FecEc10.pdf>].
- CONSONNI, V., D. Ballabio et R. Todeschini, 2010. "Evaluation of model predictive ability by external validation techniques", *J. Chemometrics* [<http://www.interscience.wiley.com>] DOI 10.1002/cem.1290.
- DELPHA, I., et M.J. RODRIGUEZ, 2014. "Effects of future climate and land use scenarios on riverine source "water quality", *Science of the Total Environment*, vol. 493, p. 1014-1024.
- DUFOUR, A., J. Bartram, R. Bos et V. Gannon, 2012. *Animal waste, water quality and human health*, London, IWA Publishing, 476 p.
- GELDREICH, E. E., 1978. "Bacterial populations and indicators concepts in feces, sewage, stormwater and solid wastes.", dans *Indicators of Viruses in Water and Food*, Gerald Berg, Ed. Ann Arbor Science, Ann Arbor, Michigan, p. 51-96.
- HÉBERT, S., et D. BLAIS, 2017. *Territoire et qualité de l'eau : développement de modèles prédictifs*, Québec, ministère du Développement durable, de l'Environnement et de la Lutte contre les changements climatiques, Direction générale du suivi de l'état de l'environnement et Direction de l'expertise en biodiversité, 30 p.
- Ministère du Développement durable, de l'Environnement, de la Faune et des Parcs (MDDEFP), 2013. *Critères de qualité de l'eau de surface*, 3^e édition, Québec, Direction du suivi de l'état de l'environnement, 510 p. + 16 annexes, [http://www.mddelcc.gouv.qc.ca/eau/criteres_eau/criteres.pdf].
- PATOINE, M., 2011. « Influence de la densité animale sur la concentration de coliformes fécaux dans les cours d'eau du Québec méridional, Canada », *Revue des sciences de l'eau*, vol. 24, n° 4, p. 421-435, [<http://www.erudit.org/revue/rseau/2011/v24/n4/1007628ar.html?vue=resume>].
- PATOINE. M., et F. D'Auteuil-Potvin, 2015. *Contamination bactériologique des petits cours d'eau en milieu agricole : état et tendances*, Québec, ministère du Développement durable, de l'Environnement et de la Lutte contre les changements climatiques, Direction du suivi de l'état de l'environnement, 35 p. + 8 annexes.
- RAKOTOMALALA R., 2005. TANAGRA : un logiciel gratuit pour l'enseignement et la recherche, dans : *Actes de EGC'2005*, RNTI-E-3, vol. 2, p. 697-702. Disponible à [<http://eric.univ-lyon2.fr/~ricco/tanagra/fr/tanagra.html>] (consulté le 13 août 2013).
- RAZALI, N.M., et Y.B. Wah, 2011. "Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling tests", *Journal of Statistical Modeling and Analytics*, vol. 2, n° 1, p. 21-33.

- ROGEL, G., 2007. *Développement et intégration au système de modélisation intégré GIBSI d'un modèle de transport et du devenir des coliformes fécaux d'origine agricole à l'échelle du bassin versant*. Mémoire de maîtrise, Université du Québec, INRS-Eau, Terre et Environnement, Québec. Disponible à <http://espace.inrs.ca/449/> (consulté le 19 août 2013).
- SAS INSTITUTE, 2003. *SAS/STAT 9.1 user's guide*. SAS Institute, Cary, NC. Disponible à <http://support.sas.com/91doc/docMainpage.jsp> (consulté le 13 août 2013).
- SAWYER, R., 1982. "Sample size and the accuracy of predictions made from multiple regression equations", *Journal of Educational and Behavioral Statistics*, vol. 7, n°2, p. 91-104.
- SCHERRER, B., 1984. *Biostatistique*, Gaëtan Morin éditeur, Boucherville (Québec), 850 p.
- SOFRONIOU, N., et G.D. Hutcheson, 2002. "Confidence intervals for the predictions of logistic regression in the presence and absence of a variance-covariance matrix", *Understanding statistics*, vol. 1, n° 1, p. 3-18.
- Statistique Canada, 2006. *Recensement de l'agriculture de 2006*. Disponible à <http://www.statcan.gc.ca/pub/95-629-x/2007000/4123856-fra.htm> (consulté le 3 janvier 2013).
- Statistique Canada, 2002. *Recensement de l'agriculture de 2001*, Données sur les exploitations agricoles diffusion complète. CD-ROM numéro 95F0304XCB.

ANNEXES

Annexe 1 Modèles 2001 et données utilisées pour leur développement

Cette annexe complète les relations entre les CF et les densités animales et humaines publiées dans la Revue des sciences de l'eau (Patoine, 2011), ici nommées « modèles 2001 », en y ajoutant des intervalles de confiance, les résultats de modèles à données répétées et à données censurées ainsi que les données utilisées pour le développement de ces modèles.

Les données de CF mesurées toutes les deux semaines dans 16 cours d'eau du Québec méridional de 2000 à 2003 ont permis de développer les modèles 2001 qui permettent de prédire les concentrations médianes de CF ou le dépassement des critères de qualité de l'eau ou encore d'illustrer l'influence particulière de certaines sources de contamination. Les stations numérotées de 1 à 10 ont servi à développer les modèles univariés utilisant la DA comme variable d'entrée et les stations numérotées de 1 à 11, les modèles bivariés utilisant la DA et la DH. Les stations 12 à 16 ont servi à quantifier l'influence des sources locales de CF. Patoine (2011) a présenté, en utilisant les médianes de CF calculées sur trois années, des équations de prédiction des médianes annuelles, estivales et hivernales de CF, des facteurs d'influence locale en période estivale et hivernale ainsi que des équations de prédiction du dépassement des valeurs de 100, 200 et 1 000 UFC/100 ml en période estivale. Une gamme plus complète d'équations pour les périodes estivale (de mai à octobre) et hivernale (de novembre à avril) est présentée dans le tableau A1.1. Les valeurs de CF et des variables utilisées pour leur développement y sont aussi présentées (tableaux A1.2 et A1.3) ainsi que le domaine joint du modèle bivarié (figure A1.5).

Les régressions linéaires par la méthode des moindres carrés sont accompagnées de plusieurs statistiques qui ont permis de vérifier les hypothèses utilisées et la capacité prédictive des équations obtenues. L'utilisation des médianes de CF calculées sur une période de trois années a fait que leur variabilité interannuelle a été limitée, augmentant ainsi la variance expliquée par les modèles. L'utilisation des médianes de chaque année a permis par contre d'augmenter la fiabilité des tests statistiques de ces régressions. Elle a aussi permis d'ajuster, par maximum de vraisemblance à l'aide de la procédure MIXED de SAS (procédure décrite plus loin), un modèle qui tient compte de la

présence d'une éventuelle corrélation entre les répétitions annuelles des médianes de CF d'une même station et entre les stations drainant des bassins versants adjacents. Son résultat se rapproche, tant sur le plan de la signification que sur celui de la valeur des coefficients, des modèles ajustés par moindres carrés. Ce résultat suggère l'indépendance spatiale et temporelle des données de CF. L'utilisation de bassins versants indépendants (non imbriqués les uns dans les autres) et la nature de la contamination bactériologique ne favorisant pas l'autocorrélation spatiale justifient d'accepter l'indépendance spatiale. De plus, l'analyse des données de CF ne montre généralement pas d'autocorrélation temporelle significative à un pas d'échantillonnage de 14 jours ou plus, ce qui permet d'accepter leur indépendance temporelle. Cette indépendance a permis d'utiliser la procédure LIFEREG de SAS, qui gère les données censurées à gauche ($CF < 2 \text{ UFC}/100 \text{ ml}$) et à droite ($CF > 6\,000 \text{ UFC}/100 \text{ ml}$), pour générer, en utilisant toutes les données de CF, des équations de prédiction de différents centiles, dont la médiane. L'ensemble des données de CF a aussi été utilisé pour la modélisation des dépassements de critères de qualité de l'eau à l'aide de régressions logistiques.

La figure A1.1a illustre les médianes estivales et hivernales de CF mesurées de 2000 à 2002 aux stations numérotées de 1 à 10 et les modèles obtenus par la méthode des moindres carrés en fonction de la DA à l'aide de ces médianes avec leurs intervalles de confiance et de prédiction. Elle montre également les médianes mesurées aux stations numérotées de 11 à 16, plus élevées en raison de densités humaines élevées ou de sources de contamination fécale en amont rapproché de celles-ci. La figure A1.1b illustre la variabilité des données estivales de CF à l'aide de centiles mesurés et modélisés par régression censurée en fonction de la DA, en utilisant toutes les données de CF mesurées de 2000 à 2002 aux stations numérotées de 1 à 10. Les dépassements des critères de qualité de l'eau du MDDEP (2013) modélisés par régression logistique pour certains usages en période estivale (activités de contact direct et indirect et irrigation) à l'aide des mêmes données de CF et les intervalles de confiance à 95 % sont également illustrés (figure A1.2b). La figure A1.2a permet de comparer visuellement, pour les seuils de 1 000, 200 et 100 UFC/100 ml, les dépassements mesurés à ceux

obtenus par régression logistique (modèles 2001RL).

Les modèles 2001 bivariés obtenus par la méthode des moindres carrés à l'aide des données de CF de 2000 à 2002 pour les stations numérotées de 1 à 11, qui permettent de prédire les médianes estivales et hivernales de CF en fonction de la DA et de la DH, sont également illustrés à la figure A1.3 à l'aide des équations présentées par Patoine (2011). Les variables DA et DH contribuent à expliquer respectivement 84 % et 2 % de la variance des concentrations médianes estivales de CF, alors que la variance de la médiane hivernale de CF est davantage expliquée par la DH (44 %) que par la DA (24 %) (Patoine, 2011). Les intervalles de prédiction à 95 % des médianes estivales de CF par le modèle 2001 bivarié peuvent être obtenus à l'aide de la figure A1.4. Dans le cas du modèle hivernal bivarié, les intervalles de prédiction ne sont pas présentés en raison de sa faible capacité prédictive ($Q^2_{INT} < 0,27$). Les courbes représentées aux figures A1.3 et A1.4 sont dans les limites du domaine joint (figure A1.5) des valeurs de DA et de DH ayant servi à construire les modèles 2001 bivariés.

Patoine (2011) indique que l'augmentation approximative de l'erreur due à l'estimation des coefficients de régression, selon la relation proposée par Sawyer (1982), est inférieure à 15 % pour les modèles à une variable explicative (DA) basés sur dix stations comparativement à près de 20 % pour les modèles à deux variables explicatives (DA, DH) obtenus à l'aide de onze stations. Ce résultat montre que l'utilisation de plus de onze stations pour la construction des modèles bivariés serait requise pour que soit obtenue une précision comparable aux modèles univariés construits à l'aide de dix stations.

Les modèles 2001 fournissent en outre un outil de prédiction du risque de contamination microbienne dans les cours d'eau de petits versants supportant une agriculture surtout intensive, en raison d'une représentation plus importante, lors de leur construction, des petits bassins versants comportant une DA élevée. Leurs intervalles de confiance et de prédiction fournissent une indication de leur précision. L'ajout de petits bassins versants comportant une agriculture moins intensive est nécessaire pour la construction de modèles plus précis applicables à une plus grande gamme de conditions d'utilisation du territoire.

L'absence de tendance significative dans les concentrations de CF à la plupart des stations depuis

l'année 2000 (Patoine et D'Auteuil-Potvin, 2015) suggère que les modèles 2001 pourraient être encore valables. Néanmoins, la tendance significative à la baisse des CF pour la rivière Bras d'Henri, qui présente la plus forte DA, justifie que soient recalculées de nouvelles relations avec des données de qualité de l'eau plus récentes. Également, les données de CF plus récentes d'autres bassins versants comportant de faibles DA et DH suggèrent que le modèle 2001 n'est pas adéquat pour les faibles valeurs de CF. Ces données plus récentes d'autres bassins versants permettent d'obtenir de nouvelles relations plus précises avec un domaine plus étendu.

Les données d'autres bassins versants spatialement indépendants permettent également de vérifier la capacité des modèles 2001 présentant un pouvoir de prédiction interne ($Q^2_{INT} > 0,50$) à prédire les valeurs de CF dans d'autres bassins versants que ceux ayant servi à leur construction.

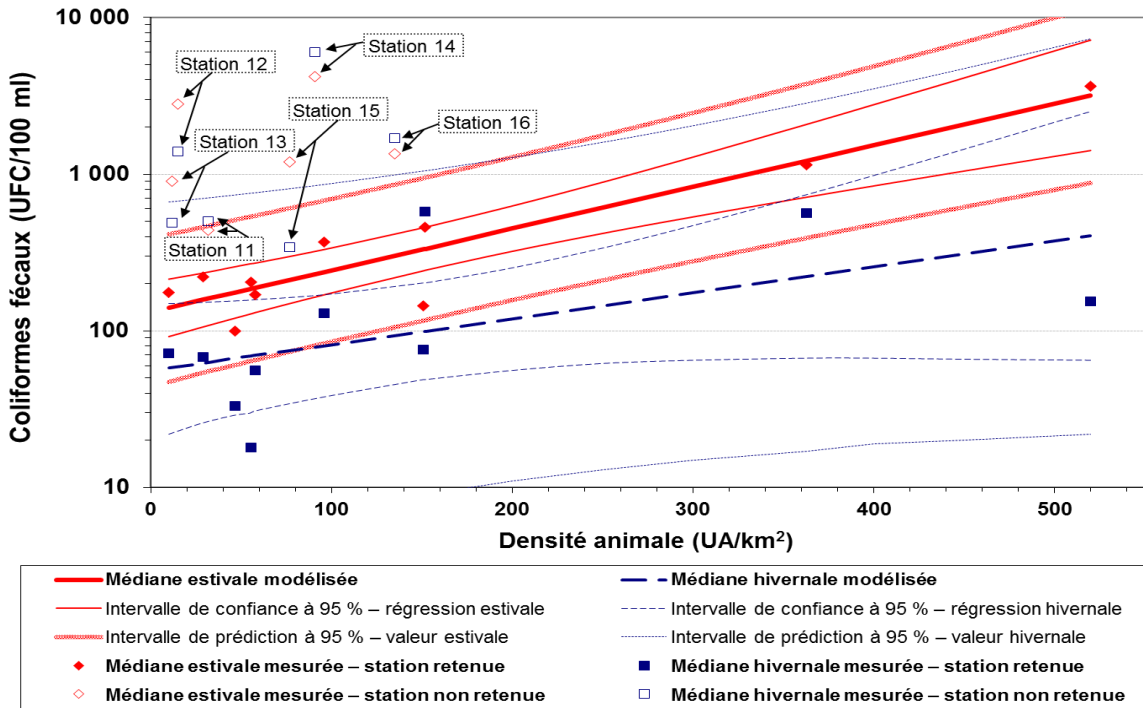
Procédures de SAS utilisées pour les modèles considérant les données répétées et les données censurées

La procédure MIXED de SAS a servi à ajuster, par maximum de vraisemblance, les modèles tenant compte de la présence d'une éventuelle corrélation entre les répétitions annuelles de médianes d'une même station et entre les stations voisines drainant des bassins versants adjacents (sous-îlots). Plus précisément, un modèle incorporant à la fois une structure de variance autorégressive temporelle sur les résidus issus d'une même station et un effet aléatoire pour tenir compte de la dépendance spatiale entre deux sous-îlots a été utilisé. Sa forme peut s'exprimer, dans le cas bivarié, à l'aide de la relation suivante :

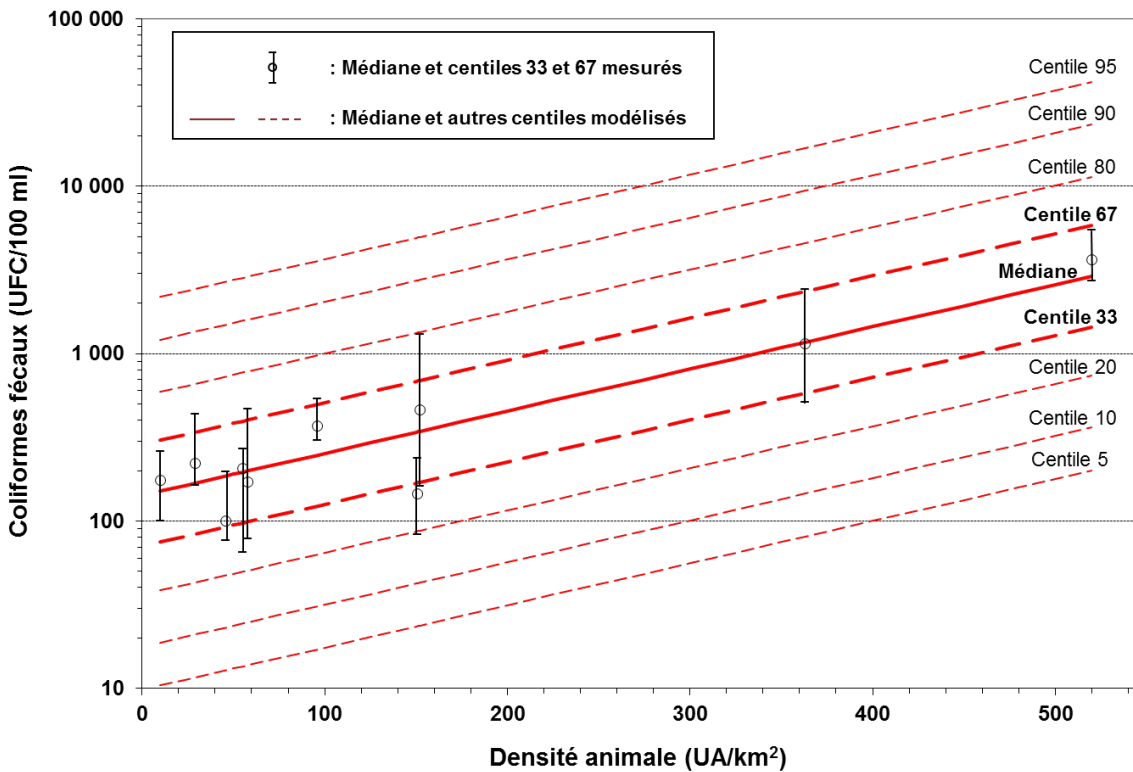
$$\ln(CF)_{j(i)k} = \text{Ordonnée à l'origine} + a \cdot DA_{j(i)} + b \cdot DH_{j(i)} + e_i + e_{j(i)k}$$

où i est l'indice du sous-îlot, $j(i)$ la station j du sous-îlot i , k l'année, e_i l'effet aléatoire sous-îlot, $e_{j(i)k}$ l'erreur résiduelle autorégressive temporelle d'ordre 1, DA la densité animale et DH la densité humaine.

La procédure LIFEREG de SAS a servi à ajuster les modèles tenant compte des données censurées à gauche et à droite (à l'extérieur de la gamme mesurable de 2 à 6 000 UFC/100 ml ou de 10 à 60 000 UFC/100 ml, selon la dilution utilisée).

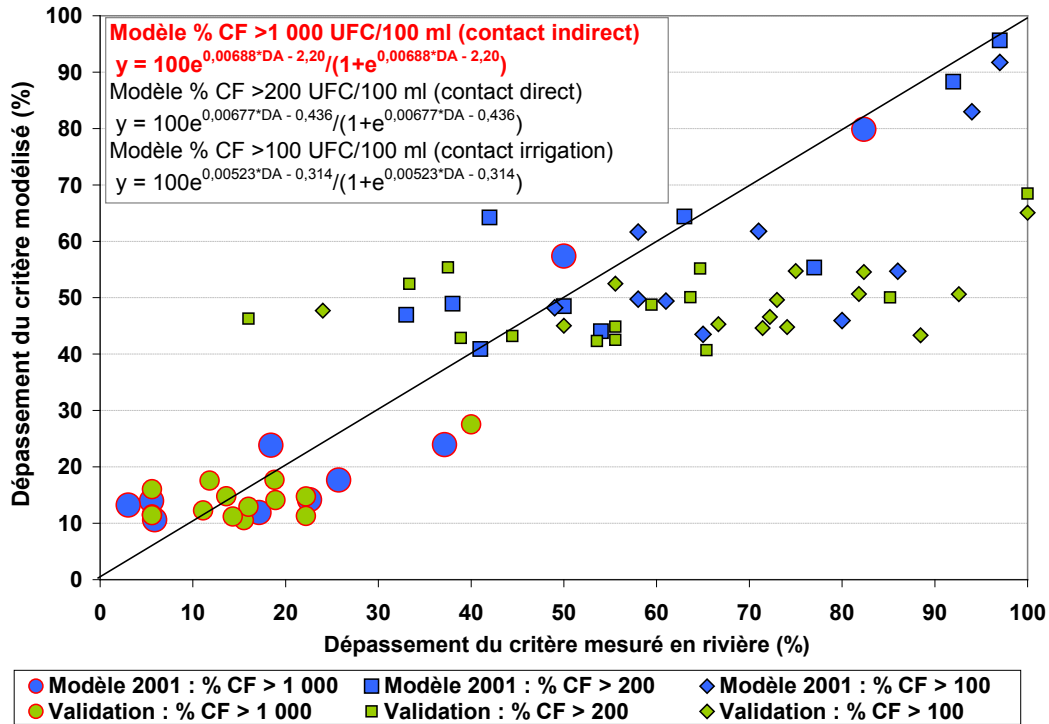


a) Médianes estivales et hivernales mesurées et modélisées (moindres carrés) en fonction de la DA

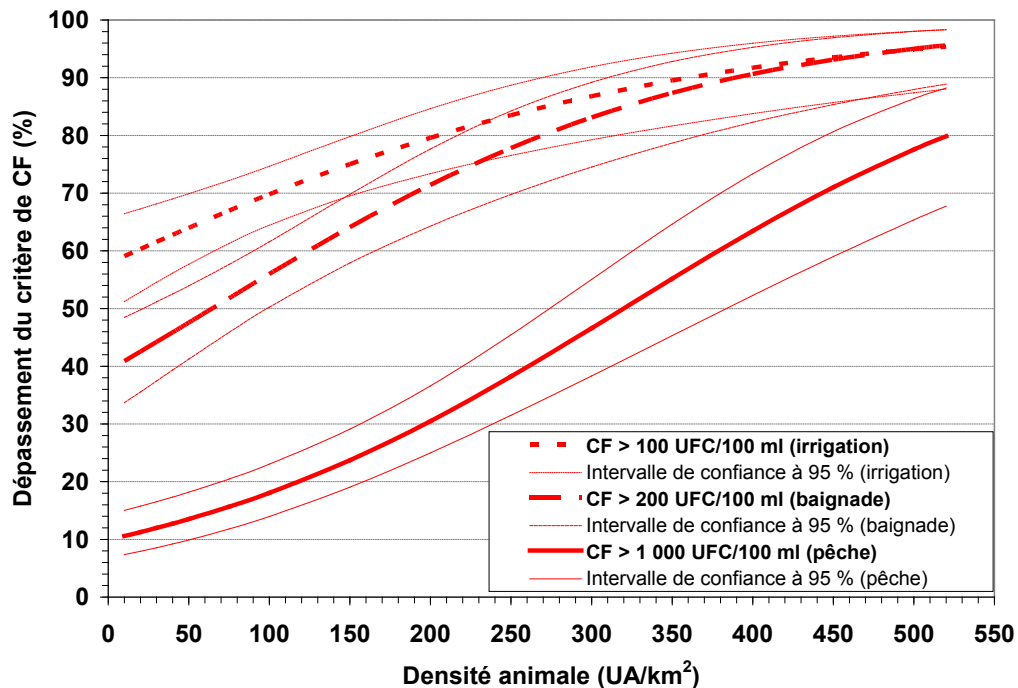


b) Centiles estivaux mesurés et modélisés (régression censurée) en fonction de la DA

Figure A1.1 Coliformes fécaux mesurés et modélisés (DH = 9 à 51 habitants/km²) – modèles 2001

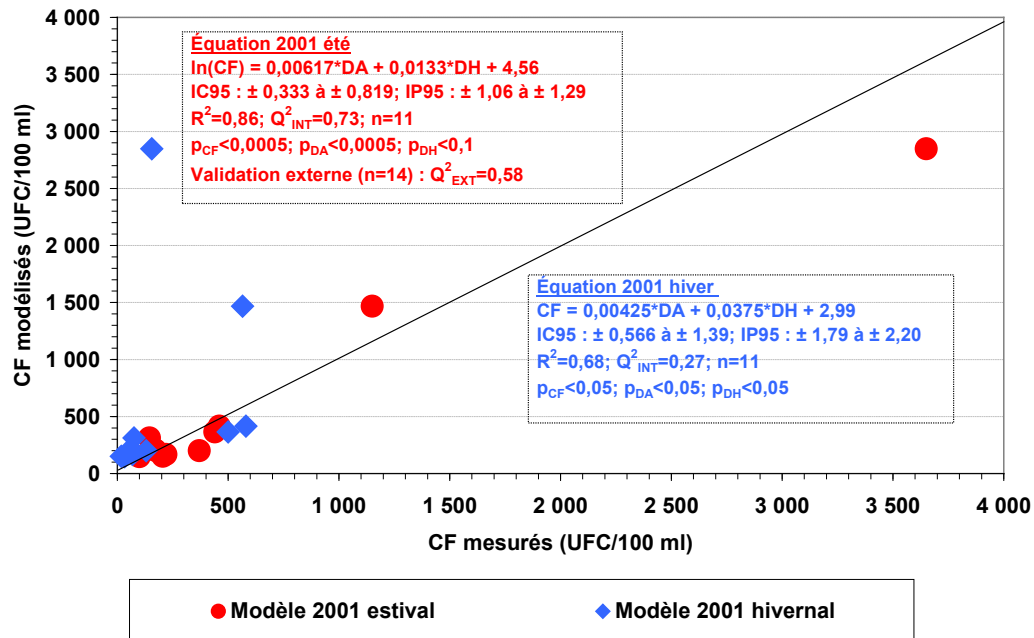


a) Dépassement estival des critères de l'eau des CF mesurés et modélisés

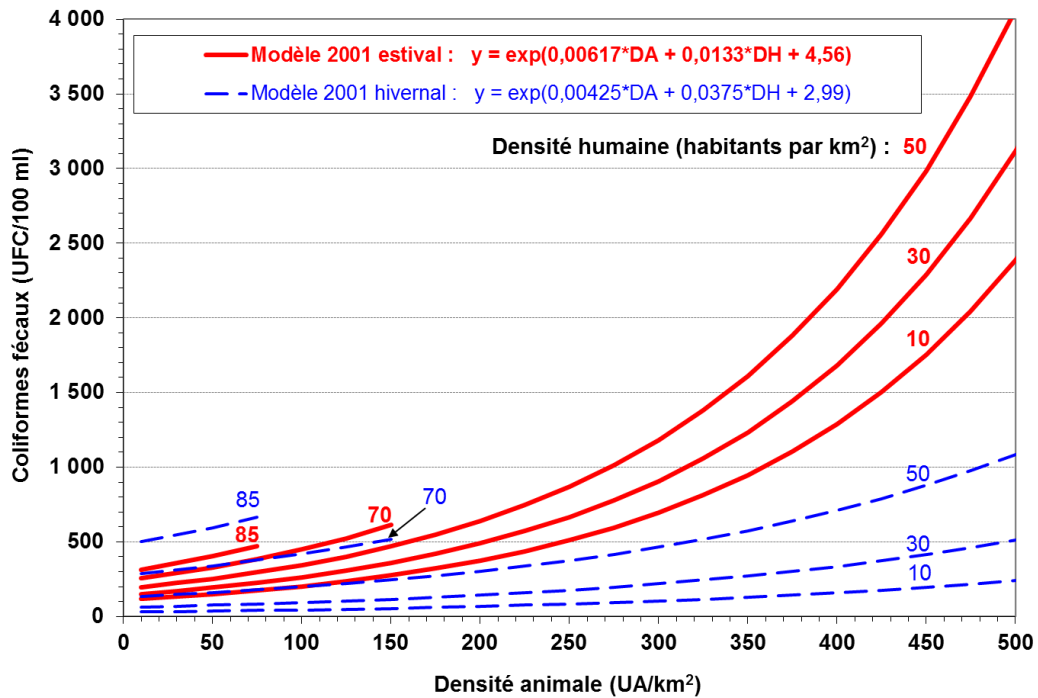


b) Modèles 2001RL de prédiction du dépassement estival des critères de qualité de l'eau

Figure A1.2 Dépassement des critères de qualité de l'eau de mai à octobre modélisé par régression logistique en fonction de la DA (DH = 9 à 51 habitants/km²) – modèles 2001RL



a) Médianes de CF mesurées et prédites à l'aide des DA et DH selon les modèles bivariés 2001



b) Modèles 2001 bivariés – CF médiane en fonction des DA et DH

Figure A1.3 Médianes estivale et hivernale des CF en fonction de la DA et de la DH – modèles 2001

Tableau A1.1 Modèles de prédiction de la concentration médiane et du dépassement de seuils de coliformes fécaux en fonction de la densité animale et humaine

Procédure statistique - données de CF	N	Équation de prédiction de la médiane	$P_{Ln(CF)}$	P_{DA}	P_{DH}	R^2	Q^2_{INT}	D_{Cook}	Note
Modèle univarié (10 stations) – mai à octobre									
RLS de SigmaStat – médianes 3 ans	10	$Ln(CF)_{été} = 0,00613*DA + 4,88$	--	0,00011	--	0,86	0,81		RSE ^a
RLS de SigmaStat – médianes an 1, 2 et 3	30	$Ln(CF)_{été} = 0,00586*DA + 4,94$	--	< 0,0001	--	0,79	0,77		--
Mixed de SAS – médianes an 1, 2 et 3	30	$Ln(CF)_{été} = 0,00592*DA + 4,91$	--	< 0,0001	--	--	--	--	--
Lifereg de SAS – toutes les données	347	$Ln(CF)_{été} = 0,0058*DA + 4,96$	--	< 0,0001	--	--	--	--	--
Modèle bivarié (11 stations) – mai à octobre									
RLM de SigmaStat – médianes 3 ans	11	$Ln(CF)_{été} = 0,00617*DA + 0,0133*DH + 4,56$	0,00037	0,00011	0,0628	0,86	0,73	< 1	RSE ^a
RLM de SigmaStat – médianes an 1, 2 et 3	33	$Ln(CF)_{été} = 0,00585*DA + 0,0121*DH + 4,67$	< 0,0001	< 0,0001	0,0069	0,76	0,72	< 1	--
Mixed de SAS – médianes an 1, 2 et 3	33	$Ln(CF)_{été} = 0,00607*DA + 0,0134*DH + 4,57$	--	< 0,0001	0,0312	--	--	--	--
Lifereg de SAS – toutes les données	384	$Ln(CF)_{été} = 0,0058*DA + 0,0147*DH + 4,61$	--	< 0,0001	0,0001	--	--	--	--
Modèle bivarié (11 stations) – novembre à avril									
RLM de SigmaStat – médianes 3 ans	11	$Ln(CF)_{hiver} = 0,00425*DA + 0,0375*DH + 2,99$	0,0106	0,0215	0,0073	0,68	0,27	< 1,61	RSE ^a
RLM de SigmaStat – médianes an 1, 2 et 3	33	$Ln(CF)_{hiver} = 0,00444*DA + 0,0451*DH + 2,72$	< 0,0001	0,0011	< 0,0001	0,52	0,45	< 1	--
Mixed de SAS – médianes an 1, 2 et 3	33	$Ln(CF)_{hiver} = 0,00448*DA + 0,0474*DH + 2,59$	--	0,0255	0,0007	--	--	--	--
Lifereg de SAS – toutes les données	313	$Ln(CF)_{hiver} = 0,0047*DA + 0,0438*DH + 2,73$	--	< 0,0001	< 0,0001	--	--	--	b

a : Relation de Patoine (2011) publiée dans la Revue des sciences de l'eau 24(4) (2011) 421-435.

b : Terme d'interaction DA*DH significatif (P=0,0009) non retenu pour faciliter la comparaison des coefficients de DA et DH avec les autres équations.

RLS = Régression linéaire simple par moindres carrés ; RLM = Régression linéaire multiple par moindres carrés ; MIXED de SAS = Procédure de SAS par maximum de vraisemblance tenant compte d'une éventuelle corrélation entre les répétitions annuelles de médianes d'une même station et entre les stations voisines drainant des bassins versants adjacents ; LIFEREG de SAS = Procédure de SAS tenant compte des données censurées à gauche et à droite ; N = Nombre de valeurs de CF utilisées pour la régression ; Ln(CF) = Logarithme de la concentration médiane de coliformes fécaux en UFC/100 ml ; DA = Densité animale en unité animale par km carré ; DH = Densité humaine en habitant par km carré ; $P_{Ln(CF)}$ = Probabilité associée au modèle ; P_{DA} = Probabilité associée à la variable DA ; P_{DH} = Probabilité associée à la variable DH ; R^2 = Proportion de la variance expliquée par le modèle ; Q^2_{INT} = Coefficient de validation interne du modèle (capacité de prédiction interne si > 0,5) ; D_{Cook} = Distance de Cook (absence de données influentes si < 1 ; possibilité de données influentes entre 1 et 2).

Équation de prédiction du pourcentage de dépassement ^a	N	P_{WALD}		P_{HL}	P_X^2	A_{ROC}		Note
		DA	DH			DA	DH	
Modèle univarié (10 stations) – mai à octobre								
$\%CF_{été} > 1\ 000 = 100e^{0,00688*DA - 2,200} / (1 + e^{0,00688*DA - 2,200})$	347	< 0,0001	--	0,058	0,538	0,79	--	RSE ^b
$\%CF_{été} > 200 = 100e^{0,00677*DA - 0,436} / (1 + e^{0,00677*DA - 0,436})$	347	< 0,0001	--	0,013	0,613	0,69	--	RSE ^b
$\%CF_{été} > 100 = 100e^{0,00523*DA - 0,314} / (1 + e^{0,00523*DA - 0,314})$	347	< 0,0001	--	0,006	0,632	0,63	--	RSE ^b
Modèle bivarié (11 stations) – mai à octobre								
$\%CF_{été} > 1\ 000 = 100e^{0,00689*DA + 0,0107*DH - 2,476} / (1 + e^{0,00689*DA + 0,0107*DH - 2,476})$	384	< 0,0001	0,0819	0,068	0,544	0,75	0,52	--
$\%CF_{été} > 200 = 100e^{0,00658*DA + 0,0130*DH - 0,7069} / (1 + e^{0,00658*DA + 0,0130*DH - 0,7069})$	384	< 0,0001	0,0085	0,008	0,607	0,65	0,53	--
$\%CF_{été} > 100 = 100e^{0,00520*DA + 0,0161*DH - 0,5595} / (1 + e^{0,00520*DA + 0,0161*DH - 0,5595})$	384	< 0,0001	0,0045	0,008	0,617	0,59	0,56	--
Modèle bivarié (11 stations) – novembre à avril								
$\%CF_{hiver} > 1\ 000 = 100e^{0,00584*DA + 0,0580*DH - 5,317} / (1 + e^{0,00584*DA + 0,0580*DH - 5,317})$	313	< 0,0001	< 0,0001	14E-04	0,989	0,57	0,80	--

a : Modèles obtenus par régression logistique à l'aide du logiciel SigmaStat en utilisant toutes les données de coliformes fécaux.

b : Relation de Patoine (2011) publiée dans la Revue des sciences de l'eau, vol. 24, n° 4, p. 421-435.

$\%CF > X$ = Pourcentage des valeurs de coliformes fécaux dépassant la concentration X, exprimée en UFC/100 ml ; N = Nombre de valeurs de CF utilisées pour la régression ; P_{WALD} = Probabilité que la variable DA ou DH soit non significative selon le test de Wald ; P_{HL} et P_X^2 = Probabilités des tests de Hosmer-Lemeshow et du Chi-carré de Pearson (une valeur > 0,05 indique une bonne adéquation entre les résultats modélisés et observés) ; A_{ROC} = Aire sous la courbe ROC (Receiver Operating Characteristic) calculée avec le logiciel TANAGRA (une valeur égale à 1 indique un modèle parfaitement discriminant, alors qu'une valeur égale ou inférieure à 0,5 indique un modèle sans capacité prédictive).

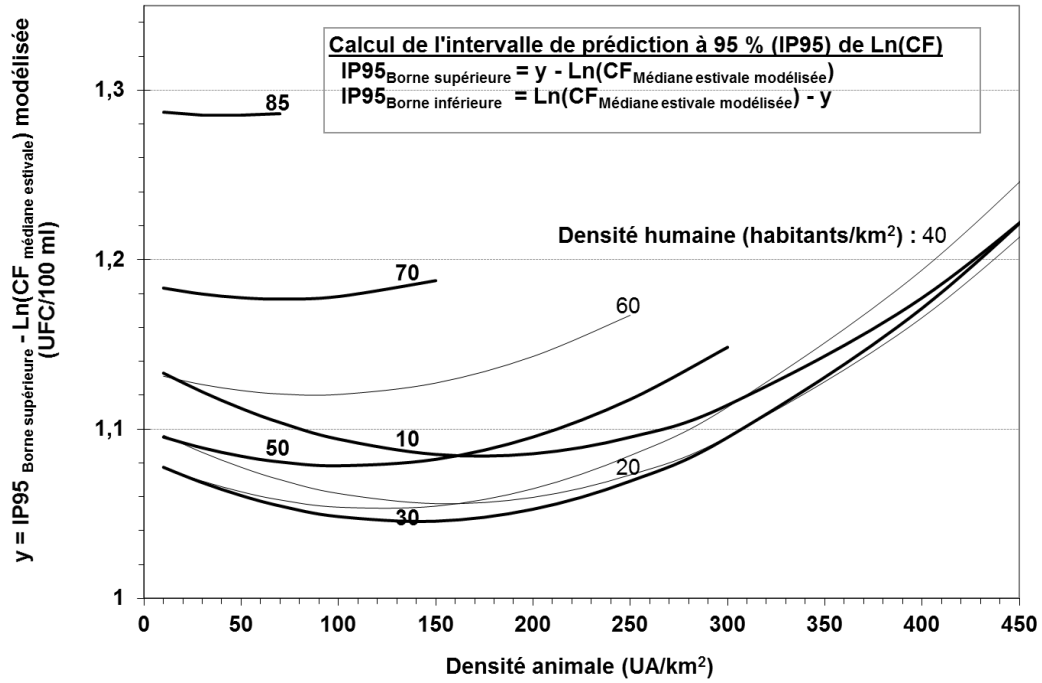


Figure A1.4 Intervalle de prédiction à 95 % de la médiane estivale de CF modélisée selon la densité animale pour des densités humaines (DH) de 10 à 85 habitants par km² – modèle bivarié 2001

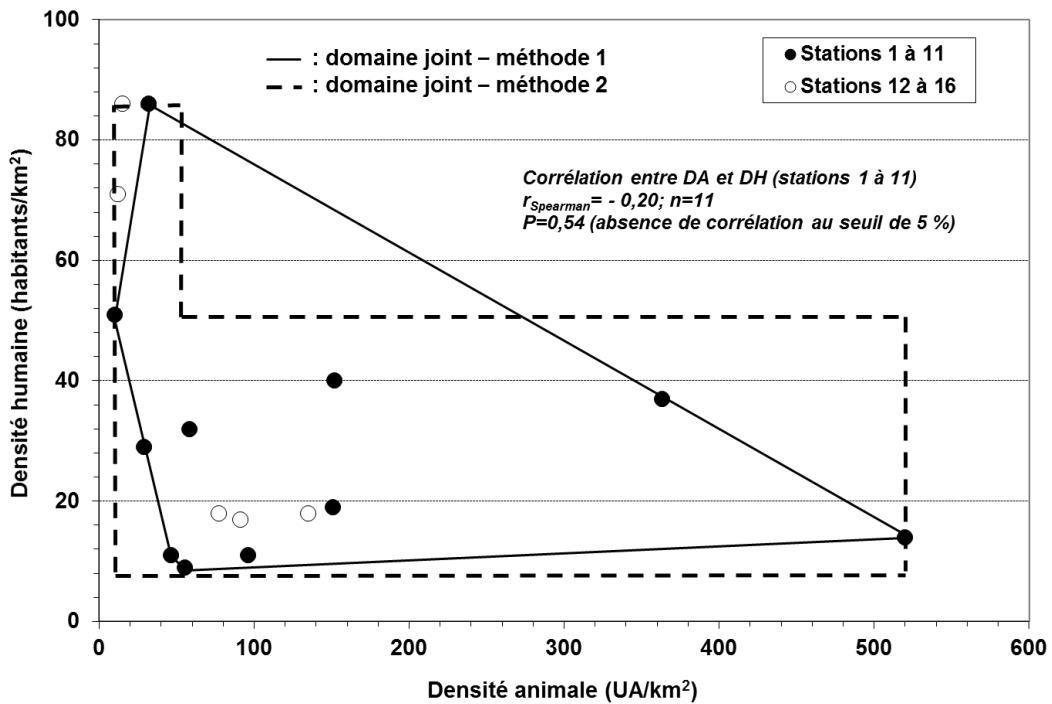


Figure A1.5 Domaine joint des modèles 2001 bivariés et corrélation entre la DA et la DH

Tableau A1.2 Description des petits bassins versants pour la période de 2000 à 2003

Nunéro de station figure 1	Cours d'eau	Superficie du bassin versant (km ²)	Densité animale (UA/km ²) ^{a, b}	Densité humaine (hab./km ²) ^a	Forêt (ha/km ² ou %) ^a	Animaux au cours d'eau (UA/km/km ²) ^{a, c}	Rejet résidentiel observé ou potentiel en amont immédiat de la station d'échantillonnage et rejet municipal
Stations non influencées, retenues pour la construction des modèles 2001							
1	Bras d'Henri	47	520	14	35	3,6	
2	Des Îles Brûlées	22	363	37	21	3,8	Rejet traité de St-Bernard (boues activées)
3	Boyer Nord	36	152	40	30	0,5	
4	Honfleur	23	151	19	27	2,2	Rejet traité de Honfleur (étangs aérés)
5	Runnels	98	96	11	63	3,7	
6	Morency	18	55	9	58	6,6	
7	Saint-Esprit	28	29	29	33	0	
8	Desrochers	15	58	32	20	0	
9	Belair	30	46	11	67	0,9	Rejet traité de Saints-Anges (étangs aérés)
10	Aux Pommes	94	10	51	64	0	
11	Vacher	28	32	86	17	0	Rejet traité de St-Jacques (boues activées)
Stations influencées, non retenues pour la construction des modèles 2001							
12	Saint-Pierre	33	15	86 ^d	16	7,1	Urbanisé à 20 % et trop-plein de Joliette ^e
13	Du Point du Jour	71	12	71	39	0,5	Rejet de résidence isolée
14	Portage	17	91	17	52	20	Rejet de résidence isolée
15	Saint-Patrice	21	77	18	47	0	Groupe de chalets construits avant 1981
16	Du Sud-Ouest	91	135	18	14	0,1	Rejet non traité de Sainte-Brigide-d'Iberville

a : Exprimé par rapport à la superficie du bassin versant, déterminée à l'aide de l'outil géomatique d'analyse spatiale ArcView 3.2.

b : Une unité animale (UA) équivaut à 1 bovin de plus d'un an, 5 veaux, 5 porcs, 4 truies, 75 dindons, 250 volailles autres. Données municipales de cheptel de Statistique Canada (2002) transformées en UA et réparties par bassin versant à l'aide des données des lieux d'élevage dans SAGO.

c : Intensité d'accès des animaux au cours d'eau calculé par la somme des rapports entre la taille du cheptel ayant accès au cours d'eau et la distance de la station de qualité d'eau au point d'accès des animaux au cours d'eau.

d : Densité humaine sans les résidences desservies par la station d'épuration de Joliette dont les eaux traitées sont rejetées en aval de la station 12.

e : Présence d'un poste de pompage à environ 800 m en amont de la station d'échantillonnage sans débordement observé durant la période de suivi des débordements (de janvier 2001 à avril 2003) et d'un régulateur (par lequel passe 10 % du débit de Joliette en période de non-débordement) à environ 2,5 km en amont de la station d'échantillonnage avec débordements fréquents en temps de pluie (8 fois en 2001 et 17 fois en 2002; 17 fois par année en moyenne, selon le suivi hebdomadaire de 2005 à 2007 par le MAMR (2008)).

Tableau A1.3 Concentrations médianes et dépassement de critères de coliformes fécaux mesurés tous les 14 jours à l'embouchure des petits bassins pour la période de mai 2000 à avril 2003^a

Nunéro de station figure 1	BQMA ^b	Statistiques annuelles			Statistiques estivales (mai-octobre)				Statistiques hivernales (novembre-avril)		
		Centile 20	Médiane	Centile 80	Médiane	%>100 ^c	%>200 ^c	%>1000 ^c	N	Médiane	N
		(UFC/100 ml)			(UFC/100 ml) (%)				(UFC/100 ml)		
Stations non influencées, retenues pour la construction des modèles 2001											
1	02340099	124	950	5 260	3 650	97	97	82	34	155	28
2	02340086	326	750	2 500	1 150	94	92	50	36	565	32
3	02300003	150	510	1 860	460	71	63	37	35	580	23
4	02300005	46	105	350	145	58	42	18	38	76	28
5	03030008	96	290	598	370	86	77	26	35	130	27
6	02340098	7	50	258	205	61	50	6	36	18	33
7	05220241	47	170	582	220	80	54	17	35	68	25
8	05220249	29	78	412	170	58	48	23	31	56	28
9	02340093	20	66	170	100	49	33	3	33	33	33
10	05080053	45	110	270	175	65	41	6	34	72	27
11	05220239	200	500	1 800	440	87	73	22	37	500	29
Stations influencées, non retenues pour la construction des modèles 2001											
12	05220240	900	2 200	4 000	2 800	100	97	81	36	1 400	30
13	05220063	300	700	1 580	900	95	89	41	37	490	25
14	02300004	484	5 300	6 000	4 200	97	97	86	35	6 000	25
15	02330029	162	480	1 400	1 200	88	82	56	34	340	29
16	03030237	526	1 400	3 160	1 350	94	91	62	34	1 700	29

a : La période de mesure est différente pour les stations numérotées 3 (de novembre 1999 à février 2005), 8 (de novembre 1999 à avril 2004), 9 (de septembre 1999 à juillet 2002), 10 (de décembre 2001 à avril 2005) et 15 (de janvier 2002 à décembre 2005).

b : Numéro de station, Banque de données sur la qualité du milieu aquatique.

c : Pourcentage des concentrations de coliformes fécaux dépassant 100, 200 et 1 000 unités formatrices de colonies (UFC) par 100 ml.

N = Nombre de données utilisées pour calculer les médianes estivales et hivernales et les pourcentages de dépassement des seuils de 100, 200 et 1 000 UFC/100 ml.

Annexe 2 Modèles 2006 et données utilisées pour leur développement

La construction des nouveaux modèles nommés « modèles 2006 » s'est faite à l'aide de 14 petits bassins versants spatialement indépendants, non influencés et disposant de données estivales de CF représentatives des années voisines de 2006, mesurés toutes les deux semaines pendant trois ans. À cette fin, trois petits bassins versants, comportant de faibles DA et DH et situés dans de nouveaux bassins versants primaires (La Chevrotière, Saint-François et Rimouski), ont été ajoutés aux onze utilisés précédemment pour le modèle 2001, ce qui a permis d'élargir son domaine. La station située près de l'embouchure de la rivière Boyer Sud a été retenue plutôt que celle en amont en raison de la disponibilité de données plus récentes. Des données triennales mesurées toutes les deux semaines proches de l'année 2006 (de 2004 à 2006, de 2005 à 2007 ou de 2006 à 2008) étaient disponibles à huit stations. À cinq stations, les données mesurées toutes les deux semaines disponibles de 2000 à 2002 seulement ont été retenues pour la construction des modèles 2006, compte tenu de l'absence de tendance significative des valeurs mensuelles de CF et de la stabilité des densités animales et humaines de 2001 à 2006. À une station, les données mensuelles de 2002 et 2003 ont été combinées aux données mesurées toutes les deux

semaines de 2008 et 2009, non significativement différentes, pour obtenir un nombre d'échantillons comparables aux autres stations (tableau A2.3). Les indicateurs de pression du territoire retenus pour la construction des modèles 2006 bivariés sont la DA et la DH, ainsi que la DA_amas et la S_déj. Ces indicateurs ont été déterminés à l'aide des données les plus proches de la période de mesure des CF retenue à chaque station de qualité d'eau.

L'utilisation de quatorze stations permet d'obtenir des modèles à deux variables explicatives, dont l'augmentation approximative de l'erreur due à l'estimation des coefficients de régression calculée à l'aide de la relation proposée par Sawyer (1982) est inférieure à 15 %, ce qui est comparable à celle des modèles à une variable construits à l'aide de dix stations (Patoine, 2011).

Les 14 stations sélectionnées pour la construction des modèles ne sont soumises à aucun rejet non traité d'eaux usées municipales et leur valeur médiane de CF n'est pas anormale selon le test de Grubb bilatéral au seuil de 5 % appliqué sur les résidus (différence entre la valeur modélisée et celle mesurée) (tableau A4.2).

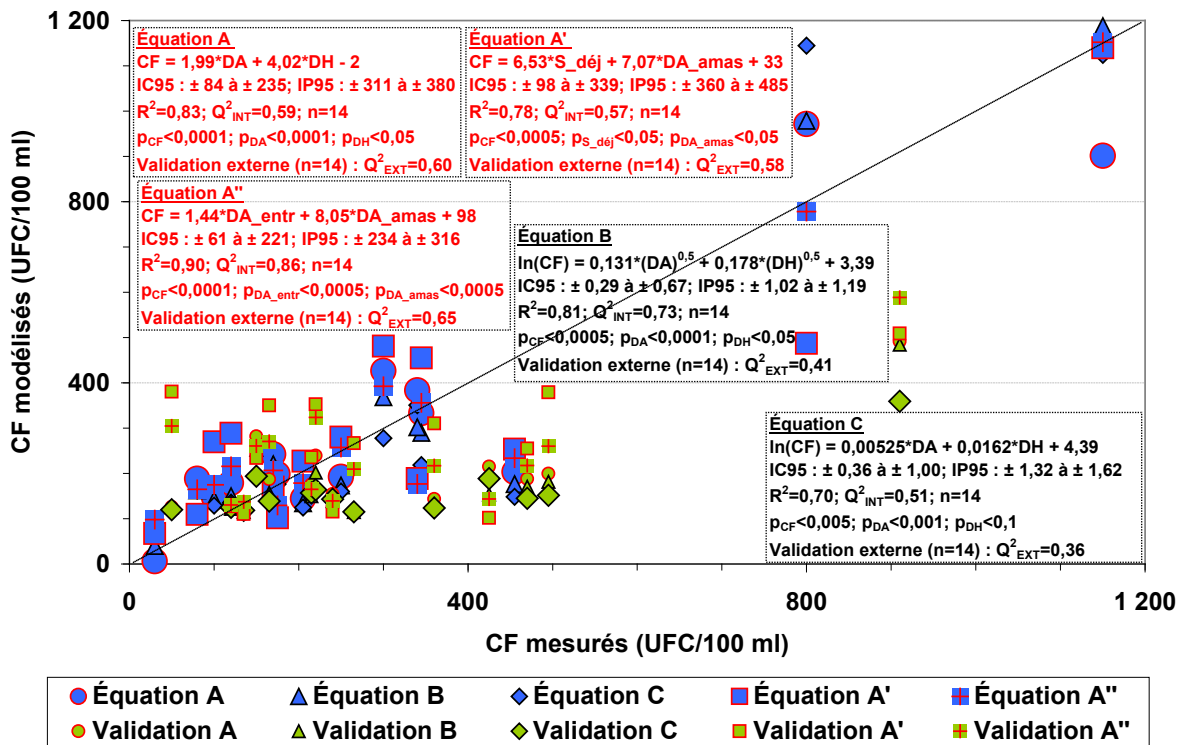


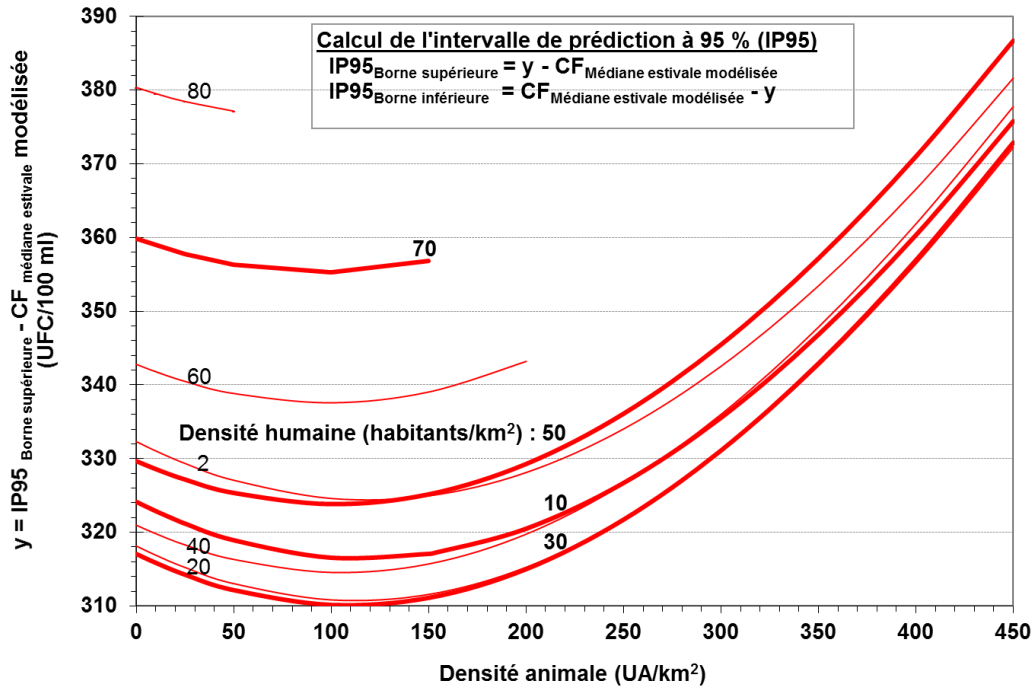
Figure A2.1 Médiannes estivales de CF mesurées et modélisées à l'aide des modèles 2006

Tableau A2.1 Intervalles de confiance et de prédiction autour des médianes estivales de CF modélisées aux stations utilisées pour la construction des modèles 2006

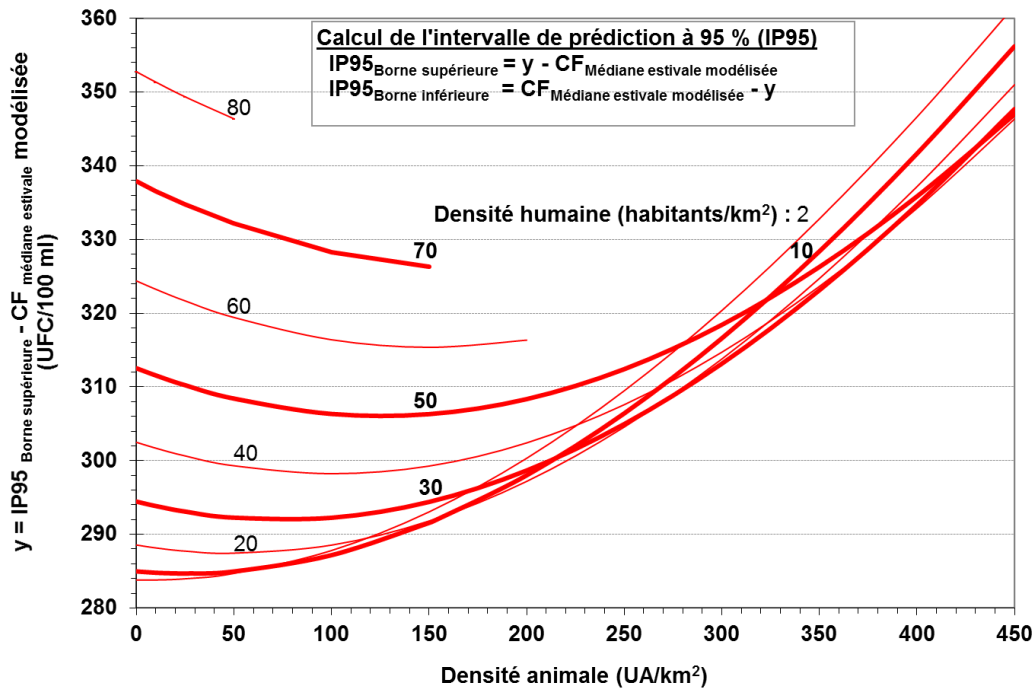
Station (figure 1)	DA (UA/km ²)	DH (hab./km ²)	S _{déj} (ha/km ²)	DA _{amas} (UA/km ²)	Intervalle de confiance à 95 %					Intervalle de prédiction à 95 %				
					2006A	2006B	2006C	2006A'	2006A''	2006A	2006B	2006C	2006A'	2006A''
35	0,1	2	5	0,0	$\pm 144 \pm 0,65$	$\pm 0,61$	± 148	± 79	± 332	$\pm 1,18$	$\pm 1,41$	± 377	± 239	
10	10	45	8	3	$\pm 123 \pm 0,47$	$\pm 0,52$	± 142	± 76	± 324	$\pm 1,09$	$\pm 1,38$	± 374	± 238	
40	24	35	6	5	$\pm 101 \pm 0,37$	$\pm 0,43$	± 150	± 72	± 316	$\pm 1,05$	$\pm 1,34$	± 378	± 237	
11	30	81	18	5	$\pm 235 \pm 0,67$	$\pm 1,00$	± 113	± 71	± 380	$\pm 1,19$	$\pm 1,62$	± 364	± 236	
7	35	31	20	17	$\pm 93 \pm 0,32$	$\pm 0,40$	± 132	± 82	± 314	$\pm 1,03$	$\pm 1,33$	± 371	± 240	
9	46	14	35	2	$\pm 102 \pm 0,31$	$\pm 0,44$	± 106	± 68	± 316	$\pm 1,03$	$\pm 1,34$	± 362	± 236	
6	55	9	30	0,1	$\pm 112 \pm 0,36$	$\pm 0,47$	± 104	± 68	± 319	$\pm 1,05$	$\pm 1,36$	± 362	± 236	
8	58	32	17	4	$\pm 89 \pm 0,30$	$\pm 0,38$	± 115	± 65	± 312	$\pm 1,03$	$\pm 1,33$	± 365	± 235	
26	77	7	38	1	$\pm 113 \pm 0,39$	$\pm 0,48$	± 114	± 66	± 320	$\pm 1,06$	$\pm 1,36$	± 365	± 235	
5	79	12	30	3	$\pm 100 \pm 0,32$	$\pm 0,43$	± 98	± 64	± 316	$\pm 1,03$	$\pm 1,34$	± 360	± 234	
3	128	20	53	11	$\pm 84 \pm 0,29$	$\pm 0,36$	± 129	± 61	± 311	$\pm 1,02$	$\pm 1,32$	± 370	± 234	
17	177	19	62	6	$\pm 94 \pm 0,34$	$\pm 0,40$	± 180	± 75	± 314	$\pm 1,04$	$\pm 1,33$	± 390	± 238	
2	363	45	82	80	$\pm 192 \pm 0,60$	$\pm 0,82$	± 339	± 221	± 355	$\pm 1,15$	$\pm 1,51$	± 485	± 316	
1	459	15	66	3	$\pm 228 \pm 0,66$	$\pm 0,97$	± 209	± 209	± 376	$\pm 1,18$	$\pm 1,60$	± 405	± 308	

DA = Densité animale; DH = Densité humaine; S_{déj} = Superficie réceptrice de déjections animales; DA_{amas} = Densité animale dont les déjections sont stockées an amas; 2006A, 2006B et 2006C = Modèles 2006A, B et C liant la médiane estivale de CF à la DA et à la DH; 2006A' et 2006A'' = Modèles 2006A' et A'' liant la médiane estivale de CF à la DA_{amas} et à la S_{déj} ou à la DA_{entr}.

Les intervalles de confiance (IC) et de prédiction (IP) indiqués correspondent à la différence entre la valeur de la borne supérieure ou inférieure de l'intervalle et la médiane modélisée. Les IC et IP en italique sont exprimés en logarithme népérien (ln).

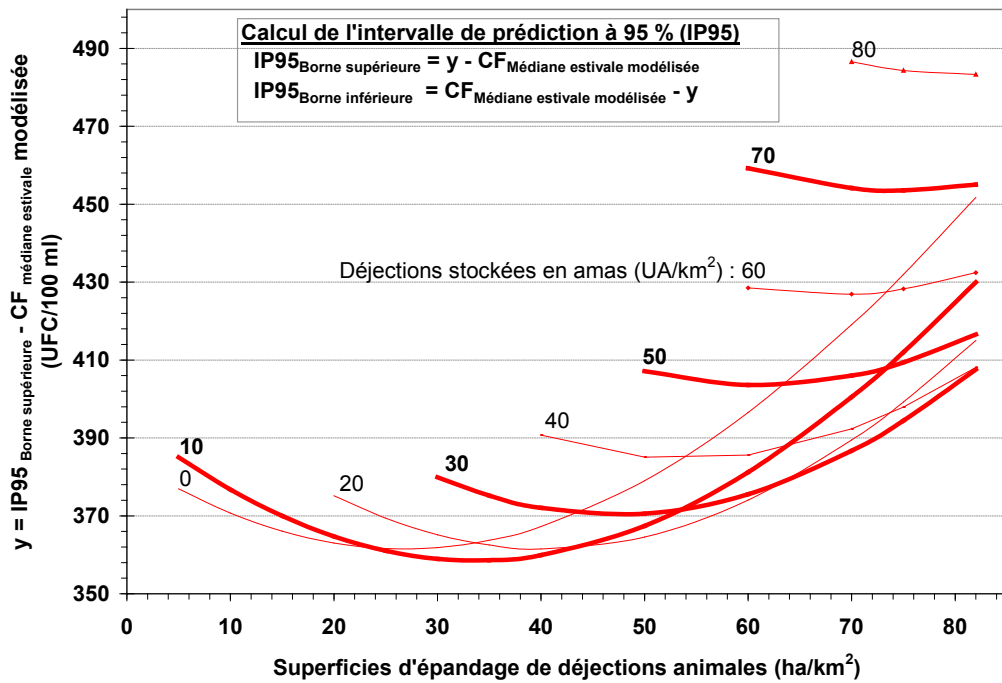


a) Équation 2006A

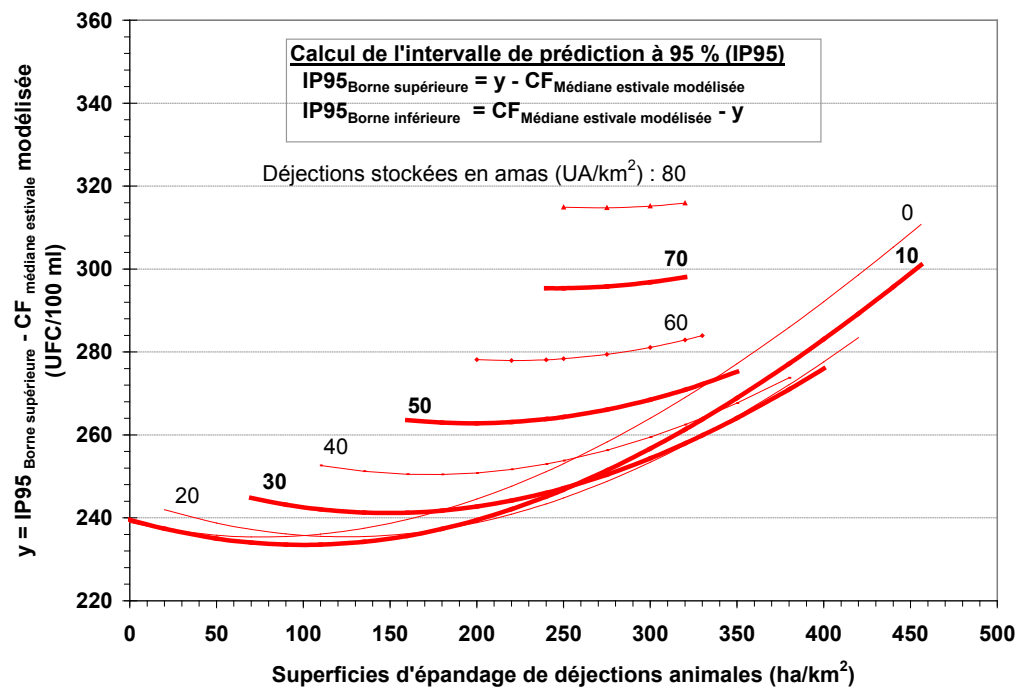


b) Équation 2006A forcée à la moitié de la limite de détection

Figure A2.2 Intervalle de prédiction à 95 % de la médiane estivale de CF modélisée selon la densité animale pour des densités humaines (DH) de 2 à 80 habitants par km² – modèles 2006A



a) Équation 2006A'



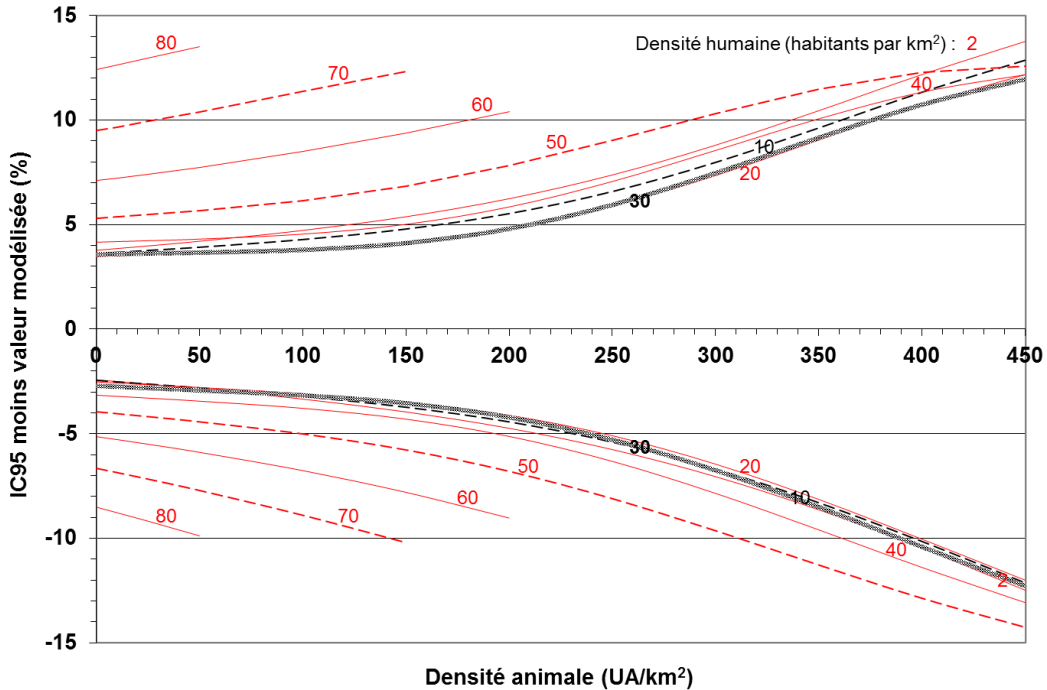
b) Équation 2006A''

Figure A2.3 Intervalle de prédiction à 95 % de la médiane estivale de CF modélisée à l'aide de la S_déj ou de la DA_entr pour des DA_amas de 0 à 80 UA par km² – modèles 2006A' et 2006A''

Tableau A2.2 Modèles de prédiction du dépassement de seuils de coliformes fécaux en fonction de la DA et de la DH, de la S_déj et de la DA_amas et de la DA_entr et de la DA_amas – modèles 2006RL, 2006RL' et 2006RL''

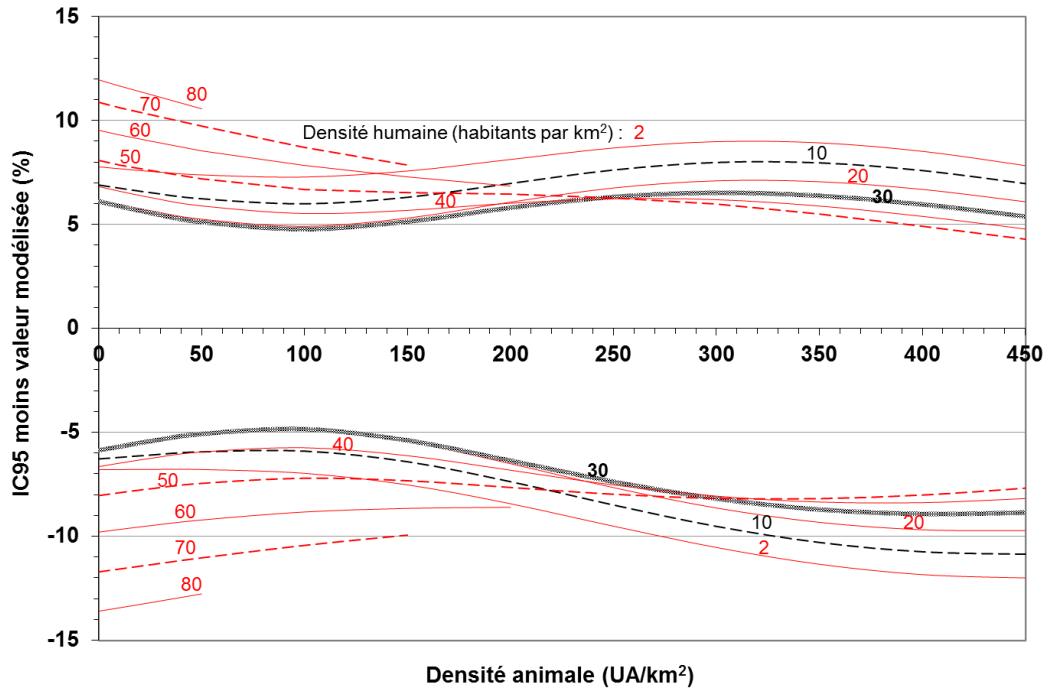
Équation de prédiction du pourcentage de dépassement ^a	N	P _{WALD}		P _{HL}	P _{X²}	A _{ROC}	
		X1	X2			X1	X2
Modèles 2006RL : X1 = DA, X2 = DH (14 stations) – mai à octobre							
%CF _{été} >1 000 = 100e ^{0,00519*X1 + 0,0170*X2 - 2,732} / (1 + e ^{0,00519*X1 + 0,0170*X2 - 2,732})	483	< 0,0001	0,0045	0,347	0,493	0,70	0,60
%CF _{été} >200 = 100e ^{0,00525*X1 + 0,0168*X2 - 0,8586} / (1 + e ^{0,00525*X1 + 0,0168*X2 - 0,8586})	483	< 0,0001	0,0005	< 0,001	0,352	0,69	0,59
%CF _{été} >100 = 100e ^{0,00489*X1 + 0,0177*X2 - 0,2209} / (1 + e ^{0,00489*X1 + 0,0177*X2 - 0,2209})	483	< 0,0001	0,0008	0,003	0,303	0,66	0,60
Modèles 2006RL' : X1 = S_déj, X2 = DA_amas (14 stations) – mai à octobre							
%CF _{été} >1 000 = 100e ^{0,0196*X1 + 0,00972*X2 - 2,397} / (1 + e ^{0,0196*X1 + 0,00972*X2 - 2,397})	483	0,0023	0,1119	0,004	0,421	0,66	0,64
%CF _{été} >200 = 100e ^{0,0246*X1 + 0,0211*X2 - 0,8435} / (1 + e ^{0,0246*X1 + 0,0211*X2 - 0,8435})	483	< 0,0001	0,0283	< 0,001	0,297	0,68	0,66
%CF _{été} >100 = 100e ^{0,0188*X1 + 0,0280*X2 - 0,0776} / (1 + e ^{0,0188*X1 + 0,0280*X2 - 0,0776})	483	0,0004	0,0295	< 0,001	0,088	0,66	0,66
Modèles 2006RL'' : X1 = DA_entr, X2 = DA_amas (14 stations) – mai à octobre							
%CF _{été} >1 000 = 100e ^{0,00430*X1 + 0,0131*X2 - 2,244} / (1 + e ^{0,00430*X1 + 0,0131*X2 - 2,244})	483	< 0,0001	0,0083	0,119	0,452	0,69	0,64
%CF _{été} >200 = 100e ^{0,00350*X1 + 0,0300*X2 - 0,4427} / (1 + e ^{0,00350*X1 + 0,0300*X2 - 0,4427})	483	0,0001	0,0016	< 0,001	0,237	0,69	0,66
%CF _{été} >100 = 100e ^{0,00311*X1 + 0,0344*X2 - 0,1871} / (1 + e ^{0,00311*X1 + 0,0344*X2 - 0,1871})	483	0,0024	0,0082	0,002	0,048	0,65	0,66

a : Modèles obtenus par régression logistique en utilisant toutes les données de coliformes fécaux; %CF > Y = Pourcentage des valeurs de CF dépassant la concentration Y, exprimée en UFC/100 ml; N = Nombre de valeurs de CF utilisées pour la régression; P_{WALD} = Probabilité que la variable X1 ou X2 soit non significative selon le test de Wald ; P_{HL} et P_{X²} = Probabilité des tests de Hosmer-Lemeshow et du Chi-carré de Pearson (une valeur > 0,05 indique une bonne adéquation entre les résultats modélisés et observés); A_{ROC} = Aire sous la courbe ROC (Receiver Operating Characteristic) (une valeur égale à 1 indique un modèle parfaitement discriminant alors qu'une valeur égale ou inférieure à 0,5 indique un modèle sans capacité prédictive).

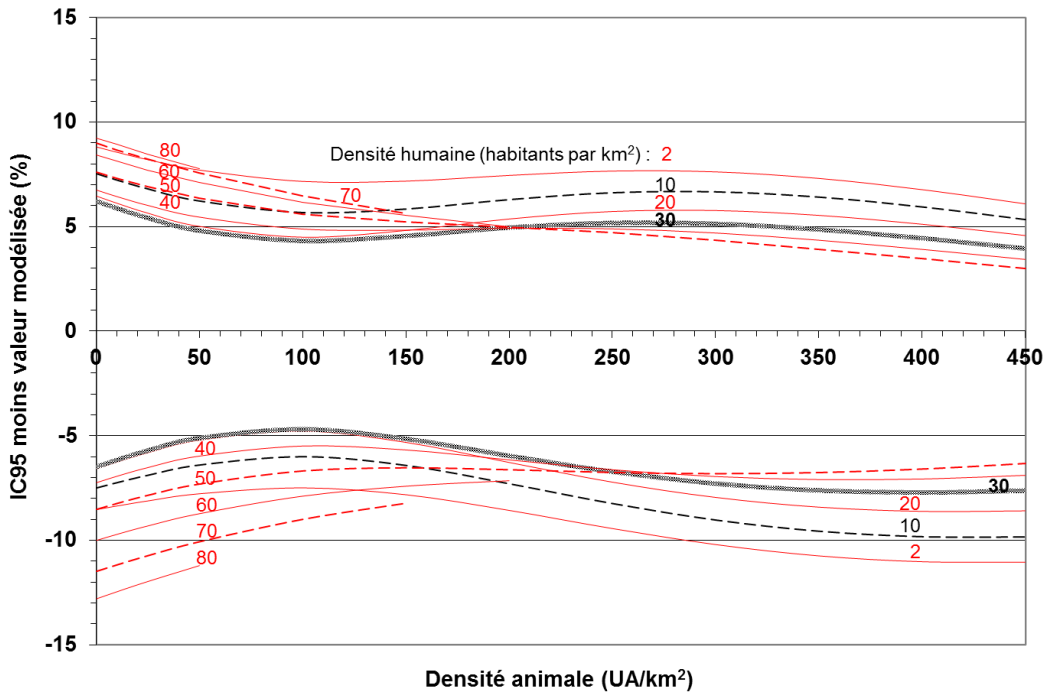


a) Critère pour les activités de contact indirect (1 000 UFC/100 ml)

Figure A2.4 Intervalle de confiance à 95 % (IC95) du pourcentage de dépassement des critères de CF modélisé selon la DA pour des DH de 2 à 80 habitants par km² – modèles 2006RL

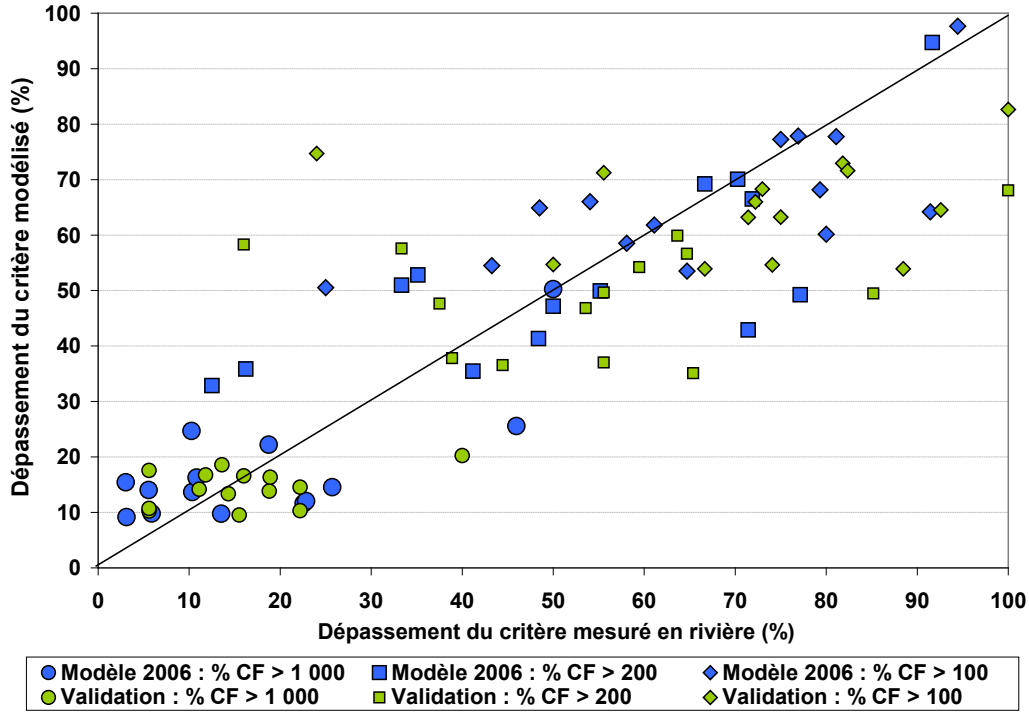


b) Critère pour les activités de contact direct (200 UFC/100 ml)

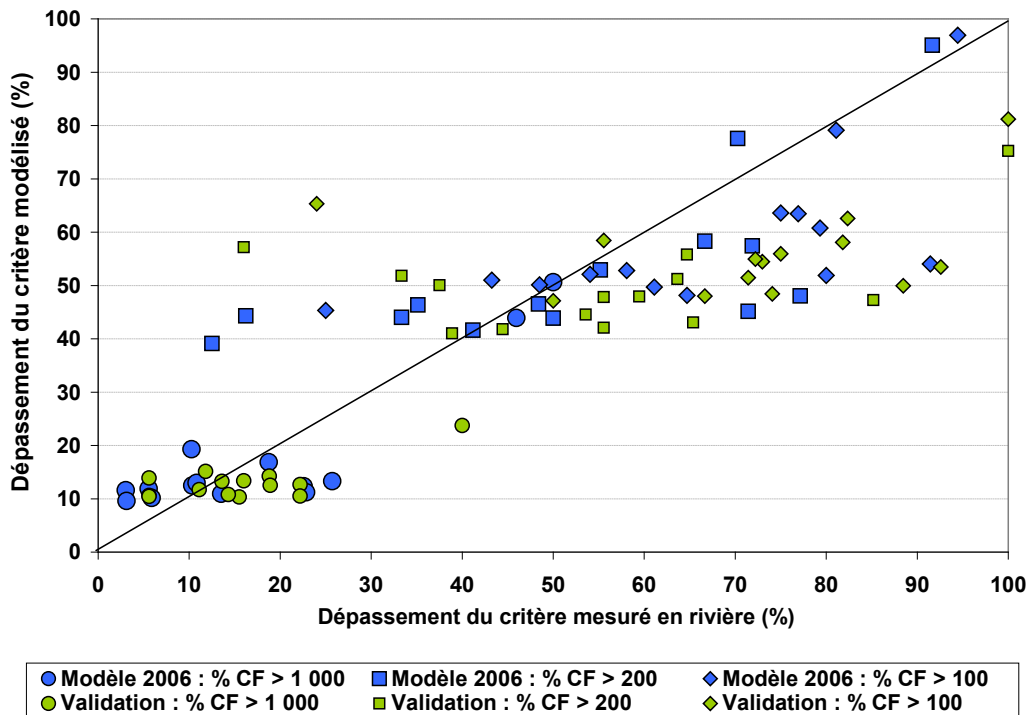


c) Critère pour l'irrigation des cultures (100 UFC/100 ml)

Figure A2.4 Intervalle de confiance à 95 % (IC95) du pourcentage de dépassement des critères de CF modélisé selon la DA pour des DH de 2 à 80 habitants par km² – modèles 2006RL (suite)

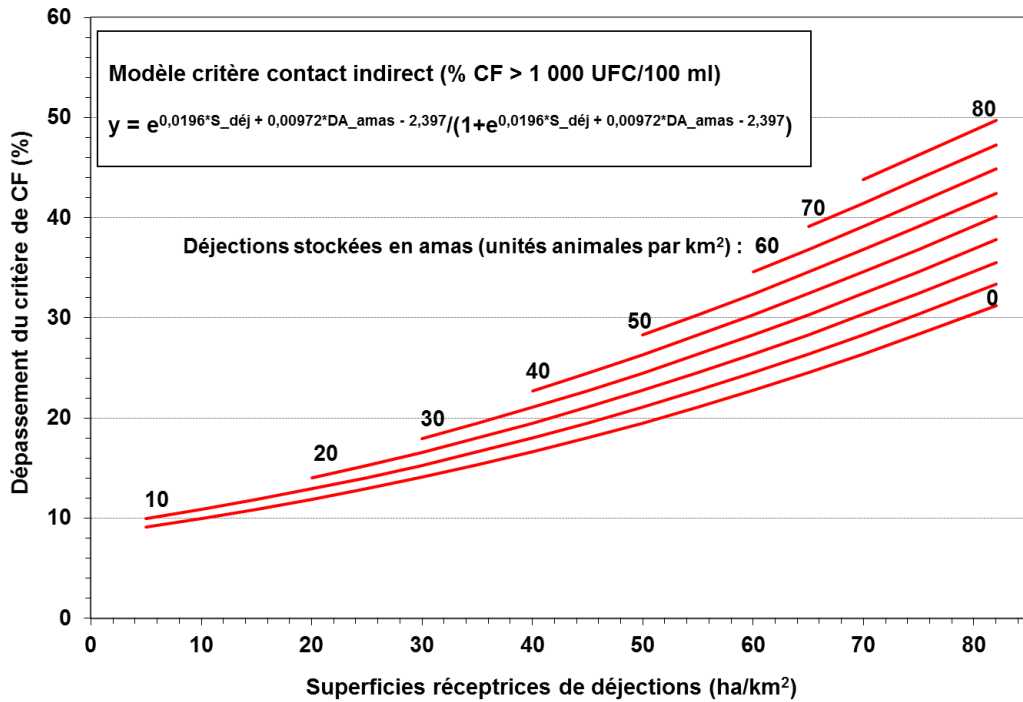


a) Équation 2006RL' – prédiction à l'aide de la S_déj et de la DA_amas

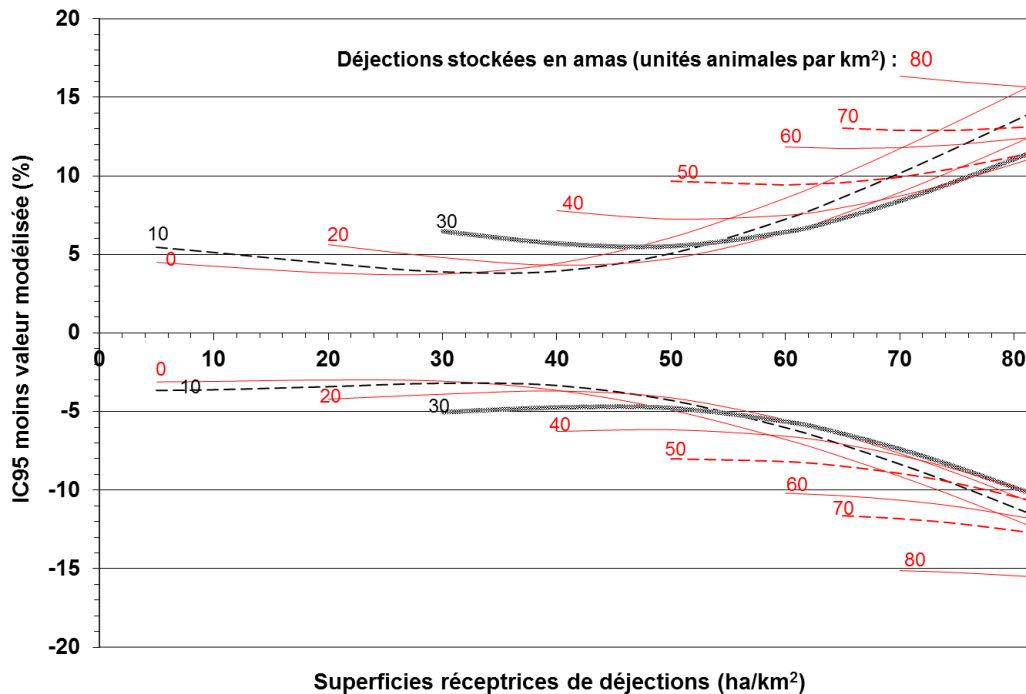


b) Équation 2006RL'' – prédiction à l'aide de la DA_entr et de la DA_amas

Figure A2.5 Dépassements estivaux des critères de coliformes fécaux mesurés et prédits à l'aide de la DA_amas et de la S_déj ou de la DA_entr – modèles 2006RL' et 2006RL''

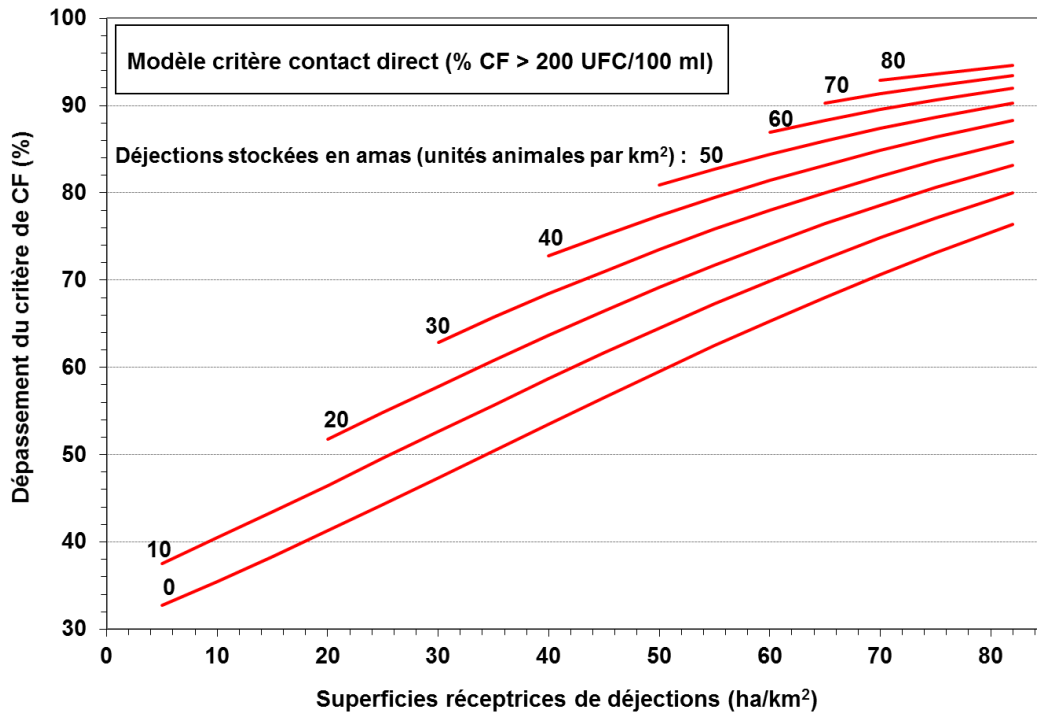


a) Représentation de l'équation 2006RL' – critère contact indirect

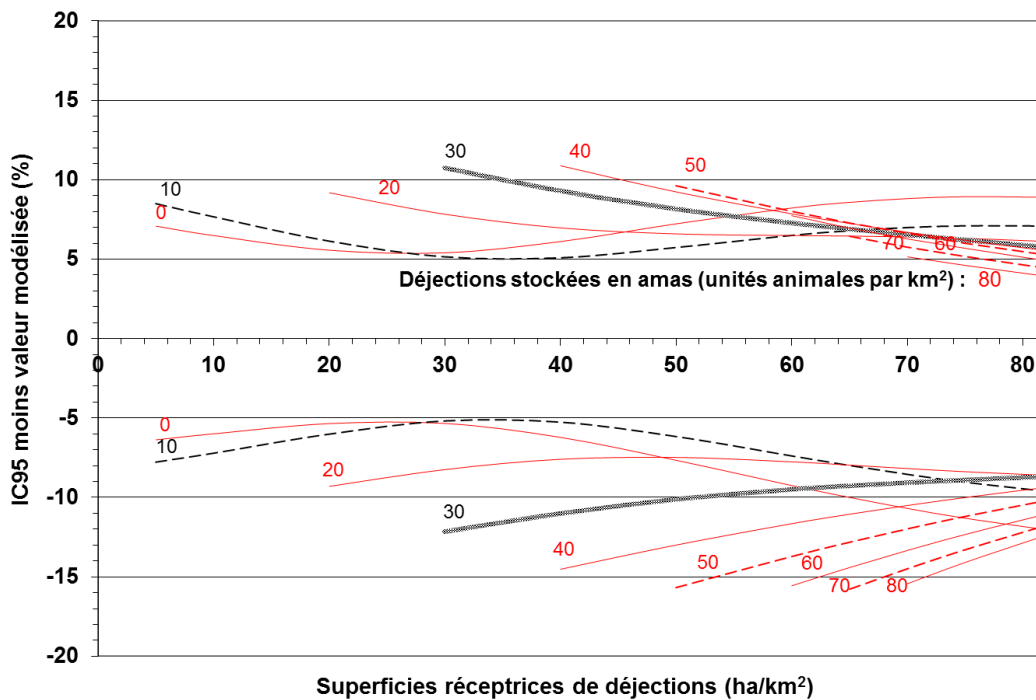


b) Intervalle de confiance à 95 % autour de l'équation 2006RL' – critère contact indirect

Figure A2.6 Prédiction du dépassement du critère de contact indirect selon la $S_{\text{dég}}$ pour des DA_{amas} de 0 à 80 UA par km² – modèle 2006RL'

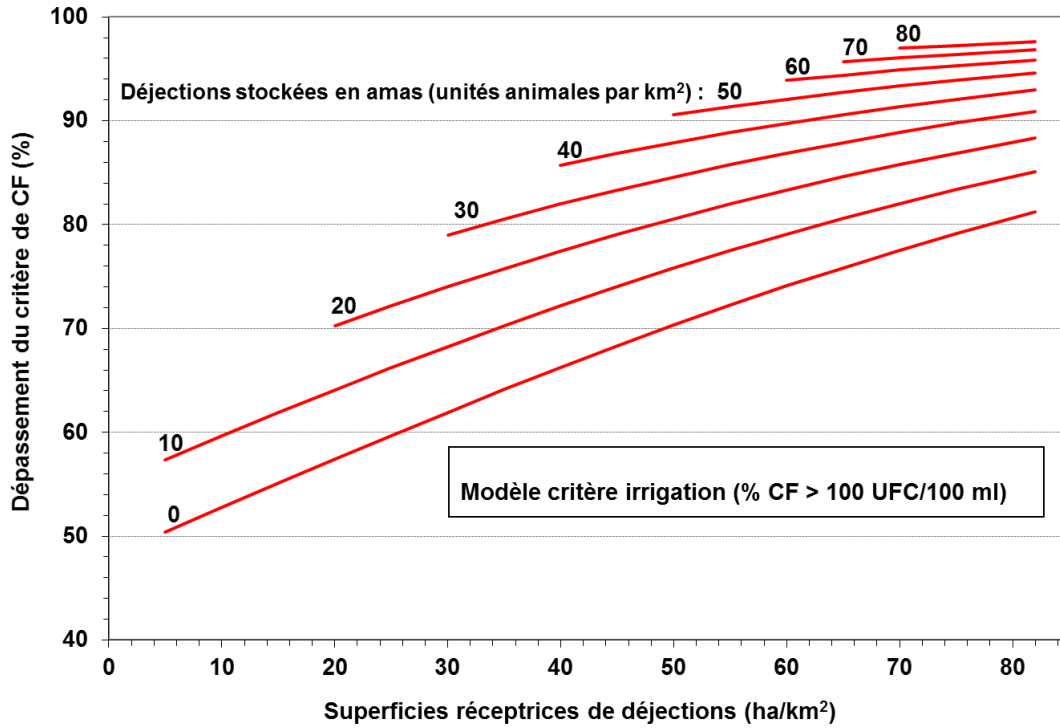


a) Représentation de l'équation 2006RL' – critère contact direct

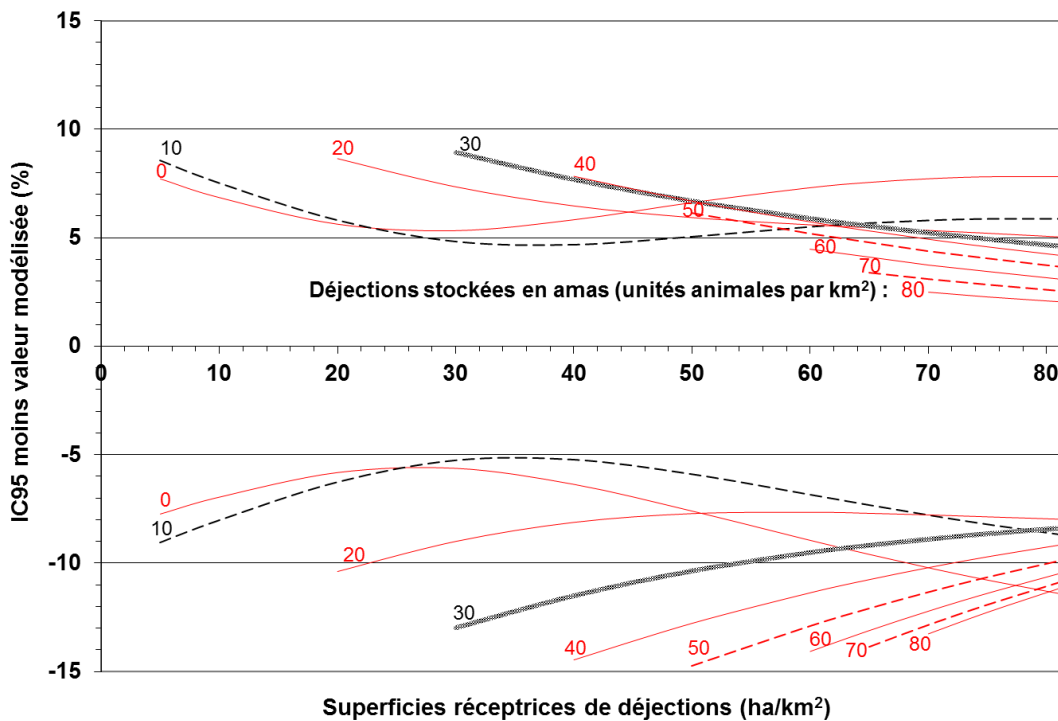


b) Intervalle de confiance à 95 % autour de l'équation 2006RL' – critère contact direct

Figure A2.7 Prédiction du dépassement du critère de contact direct selon la $S_{d\acute{e}j}$ pour des DA_{amas} de 0 à 80 UA par km² – modèle 2006RL'

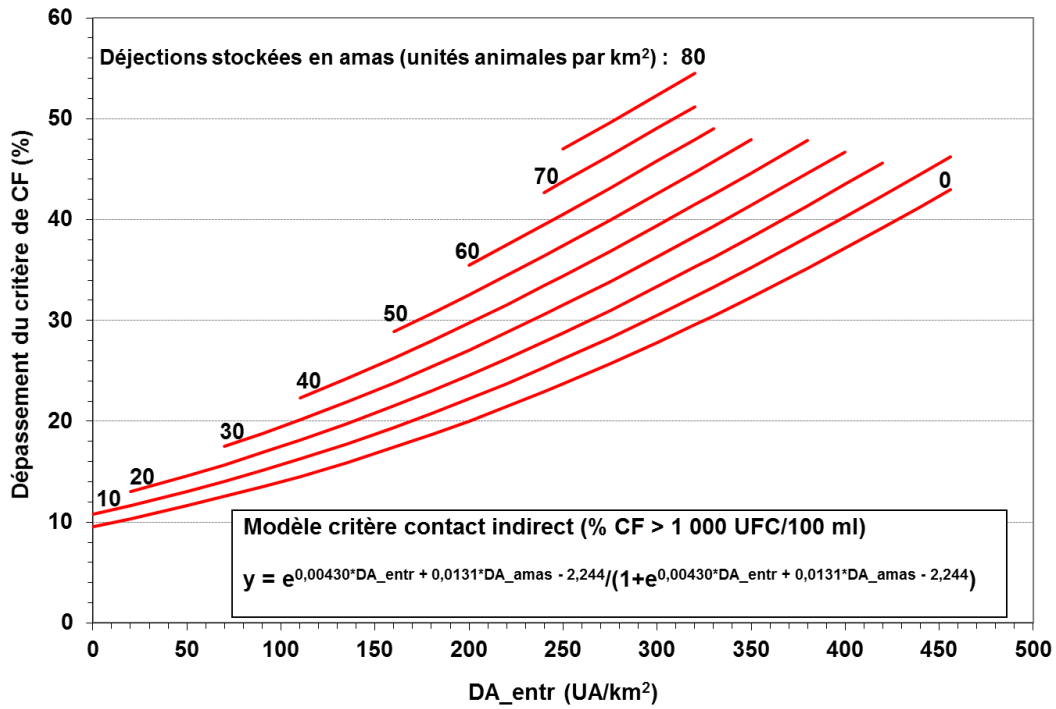


a) Représentation de l'équation 2006RL' – critère irrigation

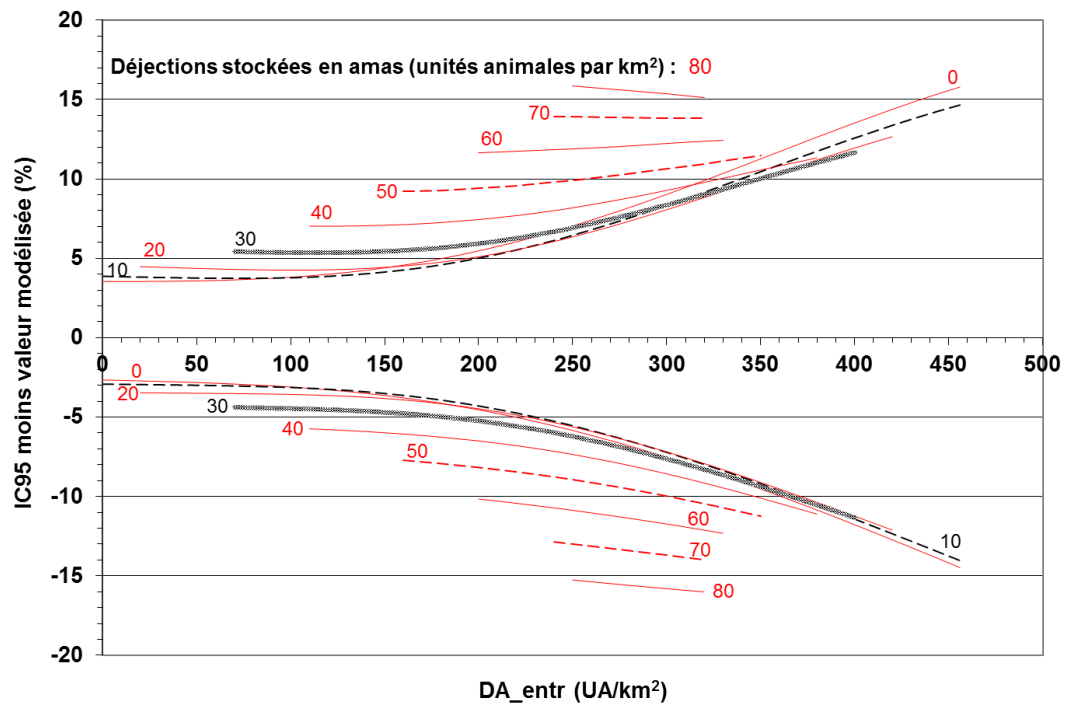


b) Intervalle de confiance à 95 % autour de l'équation 2006RL' – critère irrigation

Figure A2.8 Prédiction du dépassement du critère pour l'irrigation des cultures selon la S_{déj} pour des DA_{amas} de 0 à 80 UA par km² – modèle 2006RL'

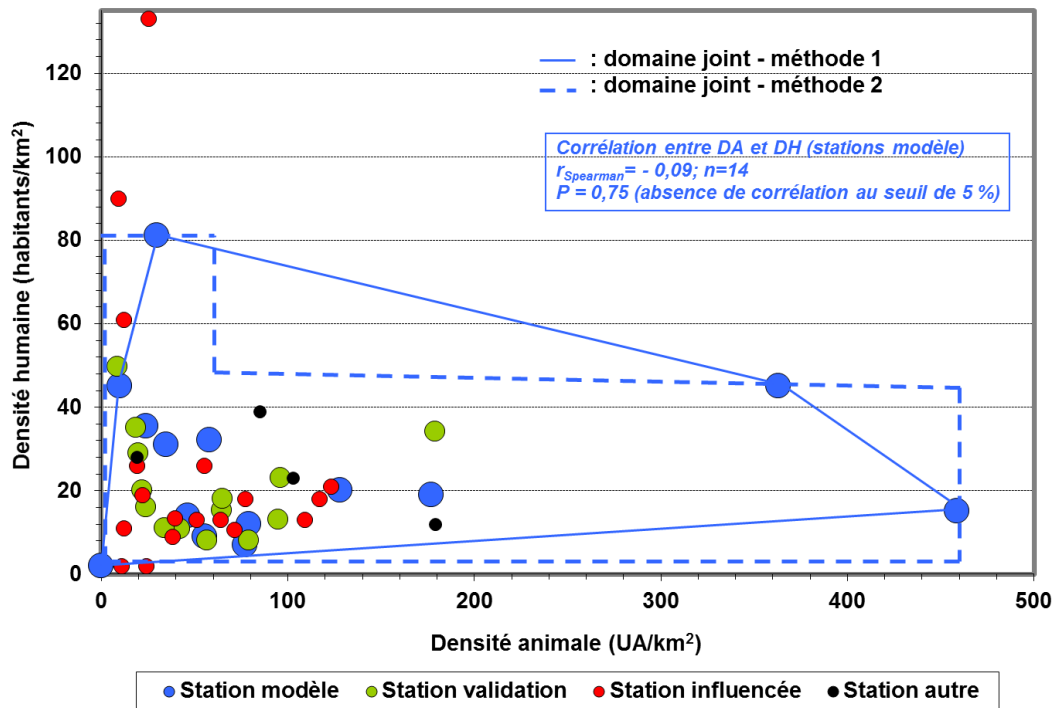


a) Représentation de l'équation 2006RL'' – critère contact indirect

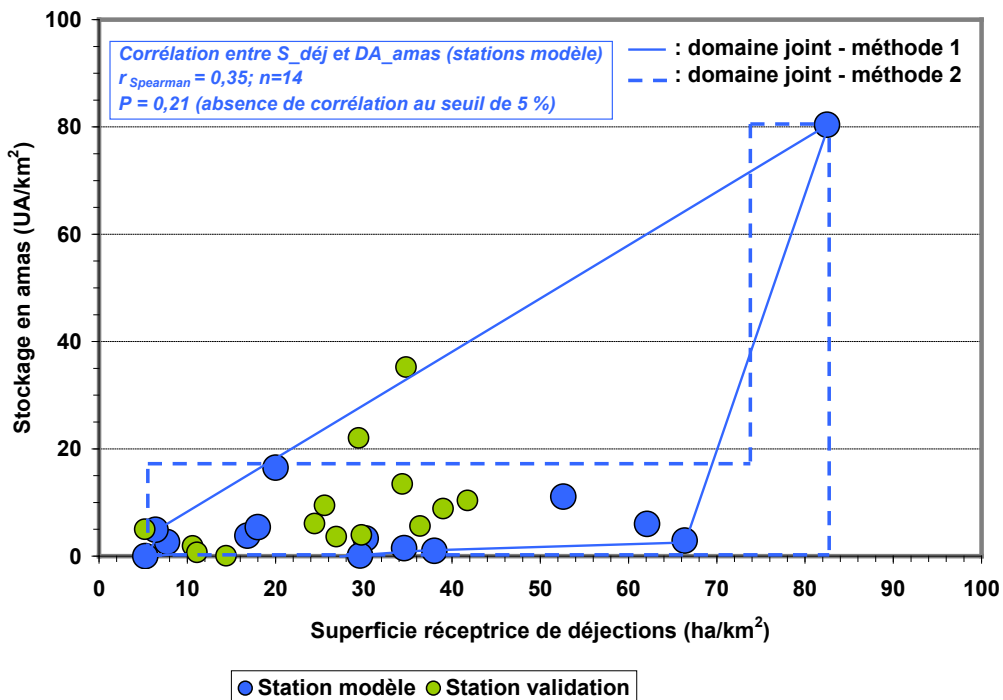


b) Intervalle de confiance à 95 % autour de l'équation 2006RL'' – critère contact indirect

Figure A2.9 Prédiction du dépassement du critère de contact indirect selon la DA_entr pour des DA_amas de 0 à 80 UA par km² – modèle 2006RL''

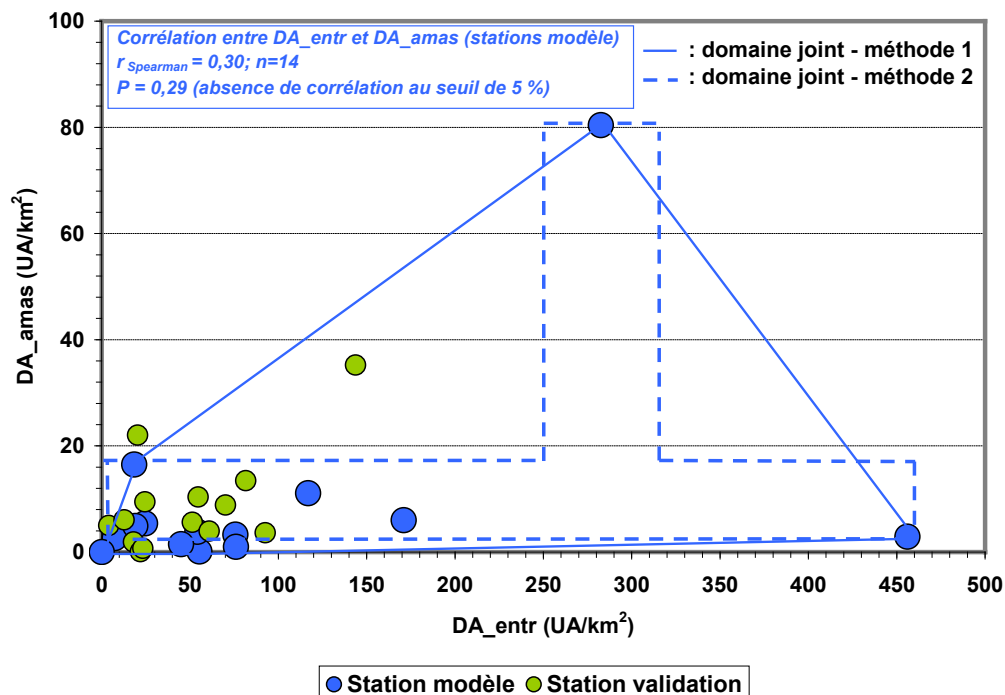


a) Modèles basés sur la densité animale (DA) et la densité humaine (DH)



b) Modèles basés sur la superficie réceptrice de déjections (S_déj) et le stockage en amas (DA_amas)

Figure A2.10 Domaine joint des modèles 2006 bivariés et corrélation entre les variables explicatives



c) Modèles basés sur la DA avec stockage en amas (DA_amas) et autre qu'en amas (DA_entr)

Figure A2.10 Domaine joint des modèles 2006 bivariés et corrélation entre les variables explicatives (suite)

Tableau A2.3 Données de densités animales et humaines, de superficies réceptrices de déjections, d'entreposage en amas et de qualité d'eau utilisées pour la construction des modèles 2006^a

Station (figure 1)	Superficie (km ²)	DA (UA/km ²)	DH (hab./km ²)	S_déj (ha/km ²)	DA_amas (UA/km ²)	Coliformes fécaux (mai à octobre)					
						Période	Médiane	%>100	%>200	%>1 000	N
1	47	459	15	66	3	2005-2007	800	81	70	46	37
2	22	363	45	82	80	2000-2002	1 150	94	92	50	36
3	36	128	20	53	11	2004-2006	345	75	72	19	32
5	98	79	12	30	3	2004-2006	455	91	77	26	35
6	18	55	9	30	0,1	2000-2002	205	61	50	6	36
7	28	35	31	20	17	2004-2006	250	79	55	10	29
8	15	58	32	17	4	2000-2002	170	58	48	23	31
9	30	46	14	35	2	2000-2002	100	48	33	3	33
10	94	10	45	8	3	2000-2002	175	65	41	6	34
11	28	30	81	18	5	2005-2007	340	80	71	23	35
17	65	177	19	62	6	2005-2007	300	77	67	10	39
26	27	77	7	38	1	2006-2008	120	54	35	11	37
35	14	0,1	2	5	0,0	2006-2008	30	25	13	3	32
40	23	24	35	6	5	2002/2003, 2008-2009	80	43	16	14	37

a : Les dépassements de critères en « % > » sont présentés à titre indicatif, les régressions logistiques utilisant toutes les données. DA = Densité animale (une unité animale (UA) équivaut à 1 bovin de plus d'un an, 5 veaux, 5 porcs, 4 truies, 75 dindons, 250 volailles autres; données municipales de cheptel de Statistique Canada (2006 et 2002) transformées en UA et réparties par bassin versant à l'aide des données des lieux d'élevage dans SAGO); DH = Densité humaine; S_déj = Superficies réceptrices de déjections non traitées de Statistique Canada (2006 et 2002) réparties au prorata de la superficie; DA_amas = Densité des élevages dont les déjections sont entreposées en amas selon SAGO; N = Nombre d'échantillons. Les coliformes fécaux sont exprimés en UFC/100 ml.

Annexe 3 Validation externe des équations de prédiction de la concentration médiane estivale de coliformes fécaux

Tableau A3 Coefficient de validation externe des modèles estivaux 2006 et 2001

Station (figure 1)	Superficie (km ²)	DA (UA/km ²)	DH (hab./km ²)	S_déj (ha/km ²)	DA_amas (UA/km ²)	Qualité de l'eau		Q ² _{EXT} des modèles 2006 et 2001					
						Période	CF	N	2006A	2006A'	2006A''	Biv2001	Univ2001
Stations non influencées et spatialement indépendantes retenues pour la validation externe des modèles													
45	17	9	50	5	5	2008-2009	425	26	0,50	-0,19	0,10	0,38	-0,16
48	62	19	35	24	6	2008-2009	215	28	0,98	1,00	0,97	0,95	0,87
44	96	20	29	11	2	2008-2009	240	27	0,92	0,82	0,89	0,83	0,79
34	12	22	20	14	0,1	2006-2008	120	18	1,00	1,00	1,00	0,97	0,95
33	14	24	16	11	1	2006-2008	135	18	0,99	0,99	1,00	1,00	0,99
49	55	34	11	26	9	2005-2007	265	18	0,73	1,00	0,97	0,55	0,78
42	35	42	11	29	22	2008-2009	50	25	0,93	-0,25	0,26	-0,12	-0,40
27	13	57	8	36	6	2006-2008	360	37	0,47	0,97	0,77	0,23	0,60
47	67	65	15	30	4	2008-2009	470	27	0,10	0,47	0,27	0,00	0,29
14	17	65	18	42	10	2007-2008	495	22	0,01	0,85	0,37	-0,03	0,21
31	14	79	8	39	9	2006-2008	165	18	0,99	0,61	0,87	1,00	0,94
50	29	95	13	34	13	2005-2007	220	17	1,00	0,80	0,88	0,99	1,00
41	53	96	23	27	4	2008	150	16	0,80	0,92	0,86	0,80	0,80
43	70	179	34	35	35	2008-2009	910	25	-0,99	-0,83	-0,17	0,51	0,35
Coefficient de validation externe moyen (Q²_{EXT}) des 14 stations									0,60	0,58	0,65	0,58	0,57
Stations non influencées hors du domaine de superficie des modèles													
20	3	179	12	--	--	2000-2002	110	37	0,03	--	--	-0,27	-0,51
21	4	85	39	--	--	2000-2002	600	39	0,13	--	--	0,36	0,08
39	106	19	28	--	--	2006-2008	465	38	-0,14	--	--	-0,21	-0,22
18	202	112	23	--	--	2000-2002	180	35	0,80	--	--	0,87	0,87
18	202	103	23	--	--	2005-2007	215	18	0,93	--	--	0,98	0,98

DA = Densité animale; DH = Densité humaine; S_déj = Superficies recevant des déjections animales non traitées; DA_amas = Densité animale avec entreposage en amas; CF = Médiane estivale de coliformes fécaux pour la période; N = Nombre de valeurs de CF; 2006A, 2006A' et 2006A'' = Modèles 2006 estivaux correspondant à l'équation A et à l'équation A' présentées à la figure A2a et au modèle 2006A'' respectivement; Biv2001 et Univ2001 = Modèles 2001 estivaux bivarié et univarié de Patoine (2011).

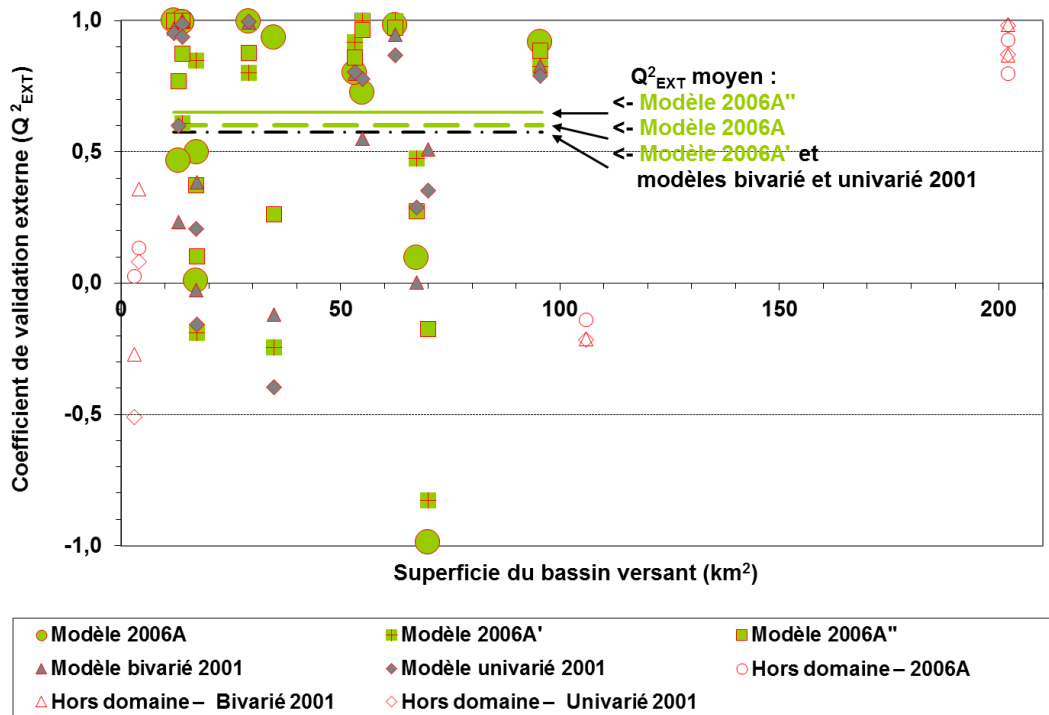


Figure A3 Coefficient de validation des modèles de prédiction de la valeur estivale de CF

Annexe 4 Résidus des modèles 2006

Tableau A4.1 Résidus des modèles 2006 aux stations influencées

Station (figure 1)	Station BQMA ^a	Coliformes fécaux (mai-octobre)			Cours d'eau	DA ^b (UA/km ²) (hab./km ²)	DH ^b	Résidus ^c Modèles 2006				Influence particulière ^d	
		Période	Médiane (UFC/100 ml)	> 1 000 (%)				N	A	B	C		RL
Influence causée par une source répertoriée ou indéterminée autre que la DA et la DH													
14	02300004	2000-2002	4 200	85,7	35	Du Portage	91	17	3 953	2,97	3,20	73	RR, TA, ACE
37	05050098	2006-2008	3 050	85,0	40	Le Petit Bras	19	26	2 909	3,15	3,11	75	RR, écoulement d'amas de fumier
12	05220240	2000-2002	2 800	80,6	36	Saint-Pierre	15	86	2 426	2,39	2,08	57	ACE, SU
12	05220240	2005-2008	815	44,7	38	Saint-Pierre	28	90	399	0,93	0,71	19	SU
25	03010086	2006-2009	1 650	62,5	32	Saint-Zéphirin amont	51	13	1 498	2,44	2,54	53	SU, écoulement d'amas de fumier
28	03040195	2006-2008	1 400	62,2	37	À l'Ours	22	19	1 282	2,46	2,43	53	SU
51	03040073	2005-2007	1 300	58,8	17	Ewing	71	11	1 117	2,09	2,23	49	Indéterminée
46	06200004	2008-2009	1 200	50,0	26	Bibeau	39	13	1 069	2,22	2,27	41	Indéterminée
22	03030253	2006-2008	1 300	61,1	36	Des Aulnages	109	13	1 033	1,77	2,00	49	Indéterminée
16	03030237	2000-2002	1 350	61,8	34	Du Sud-Ouest	135	18	1 011	1,54	1,82	47	RNT 835 (Sainte- Brigide-d'Iberville)
16	03030237	2005-2008	1 300	56,3	16	Du Sud-Ouest	117	18	997	1,61	1,87	42	Indéterminée
15	02330029	2002-2005	1 200	55,9	34	Saint-Patrice	77	18	976	1,80	2,00	44	Groupe de chalets construits avant 1981
38	05050099	2006-2008	1 500	65,6	32	Du Moulin	25	133	917	1,21	0,64	24	SU
13	05220063	2000-2002	900	40,5	37	Du Point du Jour	12	71	593	1,46	1,20	22	RR
29	12020001	2006-2008	435	22,2	18	Dauphine	24	2	381	1,79	1,53	15	Indéterminée
32	02340119	2006-2008	590	44,4	18	Laflamme	55	26	378	1,11	1,28	33	Indéterminée
19	02340091	2000-2002	700	42,4	33	Turmel	123	21	373	0,89	1,18	27	Indéterminée
36	05050097	2006-2008	355	16,7	36	Le Grand Bras	11	2	327	1,80	1,39	10	Déjections du bassin versant voisin
24	02400048	2006-2008	415	26,3	38	Bourbon	38	9	305	1,30	1,29	18	RNT 211 (Sainte- Sophie-d'Halifax)
30	03010046	2010-2012	305	16,7	18	Saint-Zéphirin	64	13	127	0,64	0,78	7	RNT 280 (Saint- Zéphirin-de-Courval) RNT 74 (La Visitation- de-Yamaska)
Influence causée par une atténuation naturelle													
23	02190010	2006-2008	9	0,0	26	Petite rivière Neigette	12	11	-57	-2,24	-2,43	-8	Lac en amont de la station
Seuil selon le test de Grubb (standard : bilatéral 5 %)									± 464 ± 1,52 ± 1,97 ± 26				
Seuil selon le test de Grubb (bilatéral 10 %)									± 417 ± 1,37 ± 1,77 ± 23				
Unités des résidus : UFC/100 ml, non transformées pour le modèle A et transformées en Ln pour les modèles B et C; pourcentage des analyses de CF > 1 000 UFC/100 ml pour le modèle de régression logistique (RL).													

a : Numéro de station, Banque de données sur la qualité du milieu aquatique.

b : Les densités animales et humaines utilisées pour le calcul des résidus proviennent du tableau 1 pour les périodes de 2005 à 2009 et de l'annexe 1 pour les périodes de 2000 à 2005, sauf pour le ruisseau Turmel (non présenté à l'annexe 1 mais utilisation de la même méthode d'attribution au bassin versant des élevages des municipalités, selon les lieux d'élevage de SAGO). Dans le cas de la rivière Saint-Zéphirin (période 2010-2012), les données du tableau 1 ont aussi été utilisées puisque les données de Statistique Canada de 2011 à l'échelle municipale ne sont pas comparables à celles des années antérieures.

c : Les résidus correspondent à la différence entre les valeurs médianes estivales de coliformes fécaux mesurées dans le cours d'eau et celles calculées à l'aide des modèles 2006A, B, C et RL (% CF > 1 000) présentés à l'annexe 2. Les résidus anormaux selon le test de GRUBB bilatéral selon un seuil standard de 5 % sont indiqués en caractère gras.

d : RR = Rejet de résidence isolée; TA = Traverse d'animaux; ACE = Accès du bétail au cours d'eau; SU = Secteur urbain en amont rapproché de la station; RNT XXX = Rejet non traité d'un réseau d'égout municipal desservant XXX personnes.

N = Nombre de données utilisées pour calculer la valeur estivale de coliformes fécaux.

Tableau A4.2 Résidus des modèles 2006 aux stations non influencées

Station (figure 1)	Station BQMA ^a	Coliformes fécaux (mai-octobre)			Cours d'eau	DA ^b (UA/km ²)	DH ^b (hab./km ²)	Résidus ^c Modèles 2006				Influence particulière ^d	
		Période	Médiane (UFC/100 ml)	> 1 000 (%)				N	A	B	C		RL
Stations construction modèles 2006													
1	02340099	2005-2007	800	45,9	37	Bras d'Henri	459	15	-172	-0,20	-0,36	-2	--
2	02340086	2000-2002	1 150	50,0	36	Des Îles Brûlées	363	45	249	-0,03	0,02	2	--
3	02300003	2004-2006	345	18,8	32	Boyer Nord	128	20	12	0,18	0,46	4	--
5	03030008	2004-2006	455	25,7	35	Runnels	79	12	252	0,95	1,12	15	--
6	02340098	2000-2002	205	5,6	36	Morency	55	9	60	0,42	0,50	-4	--
7	05220241	2004-2006	250	10,3	29	Saint-Esprit	35	31	58	0,37	0,45	-1	--
8	05220249	2000-2002	170	22,6	31	Desroches	58	32	-72	-0,26	-0,08	9	--
9	02340093	2000-2002	100	3,0	33	Belair	46	14	-47	-0,34	-0,26	-6	--
10	05080053	2000-2002	175	5,9	34	Aux Pommes	10	45	-24	0,17	-0,01	-7	--
11	05220239	2005-2007	340	22,9	35	Vacher	30	81	-43	0,12	-0,03	-0,3	--
17	02300002	2005-2007	300	10,3	39	Boyer Sud	177	19	-127	-0,20	0,08	-8	--
26	03020388	2006-2008	120	10,8	37	Veillette	77	7	-59	-0,22	-0,12	1	--
35	05050096	2006-2008	30	3,1	32	Le Grand Bras amont	0,1	2	24	-0,28	-1,02	-3	--
40	02200022	2002/2003, 2008-2009	80	13,5	37	Levasseur	24	35	-108	-0,71	-0,71	2	--
Station construction modèle (2001 seulement)													
4	02300005	2000-2002	145	18	38	Honfleur	151	19	-230	-0,80	-0,51	2	--
Stations validation													
14	02300004	2007-2008	495	13,6	22	Du Portage	65	18	295	1,00	1,18	3	--
27	03020389	2006-2008	360	18,9	37	Bradley	57	8	216	1,00	1,07	10	--
31	02330048	2006-2008	165	5,6	18	Du Moulin Bernier	79	8	-22	0,05	0,17	-5	--
33	02350002	2006-2008	135	5,6	18	Bourret	24	16	25	0,16	0,13	-3	--
34	02500001	2006-2008	120	5,6	18	Des Moulanges	22	20	-2	-0,01	-0,04	-4	--
41	02330050	2008	150	18,8	16	Le Bras	96	23	-132	-0,52	-0,26	5	--
42	03020065	2008-2009	50	16	25	Brook	42	11	-76	-0,92	-0,88	7	--
43	03030110	2008-2009	910	40,0	25	À la Barbue	179	34	418	0,63	0,93	17	--
44	03090018	2008-2009	240	22,2	27	Esturgeon	20	29	86	0,55	0,52	12	--
45	04310084	2008-2009	425	15,5	26	Rousse	9	50	210	1,02	0,81	2	--
47	05280053	2008-2009	470	22,2	27	Chacoura	65	15	282	1,01	1,18	12	--
48	06200004	2008-2009	215	14,3	28	Rouge	19	35	39	0,36	0,32	3	--
49	02E90002	2005-2007	265	11,1	18	Fouquette	34	11	155	0,84	0,83	3	--
50	03040066	2005-2007	220	11,8	17	Walbridge	95	13	-19	0,09	0,29	0,1	--
Stations autre (non indépendantes ou hors du domaine de 10 à 100 km²)													
18	02300001	2000-2002	180	11,4	35	Boyer	112	23	-133	-0,44	-0,16	-3	--
18	02300001	2005-2007	215	27,8	18	Boyer	103	23	-80	-0,20	0,07	14	--
21	02300013	1998-2000	600	41	39	Grillade	85	39	276	0,69	0,93	25	--
20	02340092	2000-2002	110	5,4	37	Binet	179	12	-292	-1,06	-0,82	-11	--
39	05050100	2006-2008	465	23,7	38	La Chevrotière	19	28	317	1,24	1,20	13	--
Seuil selon le test de Grubb (standard : bilatéral 5 %)								± 464 ± 1,52 ± 1,97 ± 26					
<i>Seuil selon le test de Grubb (bilatéral 10 %)</i>								<i>± 417 ± 1,37 ± 1,77 ± 23</i>					
Unités des résidus : UFC/100 ml, non transformées pour le modèle A et transformées en Ln pour les modèles B et C; pourcentage des analyses de CF > 1 000 UFC/100 ml pour le modèle de régression logistique (RL).													

a : Numéro de station, Banque de données sur la qualité du milieu aquatique.

b : Les densités animales et humaines utilisées pour le calcul des résidus proviennent du tableau A2.3 pour les stations « construction », sauf pour le ruisseau Honfleur (tableau A1.2), du tableau 1 pour les stations « validation » et « autre » pour les périodes de 2005 à 2008 et d'une répartition selon les lieux d'élevage de SAGO pour les stations « autre » pour les périodes de 2000 à 2002.

c : Les résidus correspondent à la différence entre les valeurs médianes estivales de coliformes fécaux mesurées dans le cours d'eau et celles calculées à l'aide des modèles 2006A, B et C et 2006RL (% CF > 1 000) présentés à la figure A2.1 et à la figure 5 respectivement. Les résidus anormaux selon le test de GRUBB bilatéral selon un seuil standard de 5 % sont indiqués en caractère gras.

d : RR = Rejet de résidence isolée; TA = Traverse d'animaux; ACE = Accès du bétail au cours d'eau; SU = Secteur urbain en amont rapproché de la station; RNT XXX= Rejet non traité d'un réseau d'égout municipal desservant XXX personnes.

N = Nombre de données utilisées pour calculer la valeur estivale.

Annexe 5 Données de pression microbienne du territoire

Tableau A5 Sources de données disponibles en lien avec les processus de contamination microbienne de l'eau par des matières fécales dans les petits bassins versants

Indicateur potentiel	Type de données	Source des données	Limites géographiques	Disponibilité temporelle ^a
Indicateur général				
Densité animale	Nombre d'animaux par type d'élevage	Recensement de l'agriculture : Statistique Canada	Subdivision de recensement unifiée (SRU) ^b	Années 2001, 2006, 2016 et 2011 ^b
		Système d'aide à la gestion des opérations (SAGO) : MELCC	Lieu d'élevage	Années 2003 et suivantes ^c
		Fiche d'enregistrement des exploitations agricoles : MAPAQ	Position centrale de l'exploitation agricole (centroïde)	Années 2000, 2003, 2006 et 2009, 2012 ^d
Densité humaine	Population	Répertoire des municipalités du Québec : MAMH	Municipalité	Annuelle ^e
Densité faunique	Faune à sang chaud	Aucune base de données à l'échelle des petits bassins versants		
Indicateur de certaines pratiques agricoles				
Densité animale avec déjections stockées en amas	Amas ou accès au cours d'eau ^f et nombre d'animaux par type d'élevage	Système d'aide à la gestion des opérations (SAGO): MELCC	Lieu d'élevage	Années 2003 et suivantes
Densité animale avec accès au cours d'eau				
Superficie réceptrice de déjections		Recensement de l'agriculture : Statistique Canada	Subdivision de recensement unifiée (SRU) ^b	Années 2001, 2006, 2016 et 2011 ^b

a : Disponibilité temporelle à l'intérieur de la période de suivi de qualité de l'eau considérée dans la présente étude, soit de 1998 à 2012. Des données plus précises sont requises aux périodes retenues pour le développement des modèles, soit de 2000 à 2002 et de 2004 à 2009 (voir la note ci-dessous).

b : Les données par SRU sont disponibles en ligne sur le site de Statistique Canada à compter de 2006. Le MELCC a compilé par municipalité les données par SRU des recensements de 2001 et de 2006, mais non celles du recensement de 2011 parce qu'elles ne sont pas comparables à celles des années 2001 et 2006

c : Le Système d'aide à la gestion des opérations (SAGO) est une base de données à l'usage des employés du MELCC.

d : Les données des fiches d'enregistrement des exploitations agricoles sont colligées dans une base de données du ministère de l'Agriculture, des Pêcheries et de l'Alimentation du Québec (MAPAQ) non accessible en ligne aux autres ministères ou organismes.

e : Données disponibles en ligne sur le site du ministère des Affaires municipales et de l'Habitation (MAMH).

f : Le système SAGO indique la présence de stockage en amas contigu à un bâtiment d'élevage ou au champ et l'accès d'animaux au cours d'eau, sans toutefois indiquer le cheptel qui s'y rattache parmi le nombre d'animaux par type d'élevage indiqué pour le lieu d'élevage.

NOTE : Les données de cheptel par municipalité de Statistique Canada 2001 et 2006 ont été retenues pour le calcul des densités animales en raison de leur concordance temporelle aux données de suivi de qualité de l'eau et de leur accessibilité. Pour la construction des modèles, l'attribution de ces données par municipalité au bassin versant a été réalisée à l'aide des données par lieu d'élevage du système SAGO en raison d'un besoin de précision accru, et dans les autres cas, une attribution au prorata des superficies de la municipalité dans le bassin versant a été retenue.

Annexe 6 Fréquence cumulée des densités animales et humaines des petits bassins versants

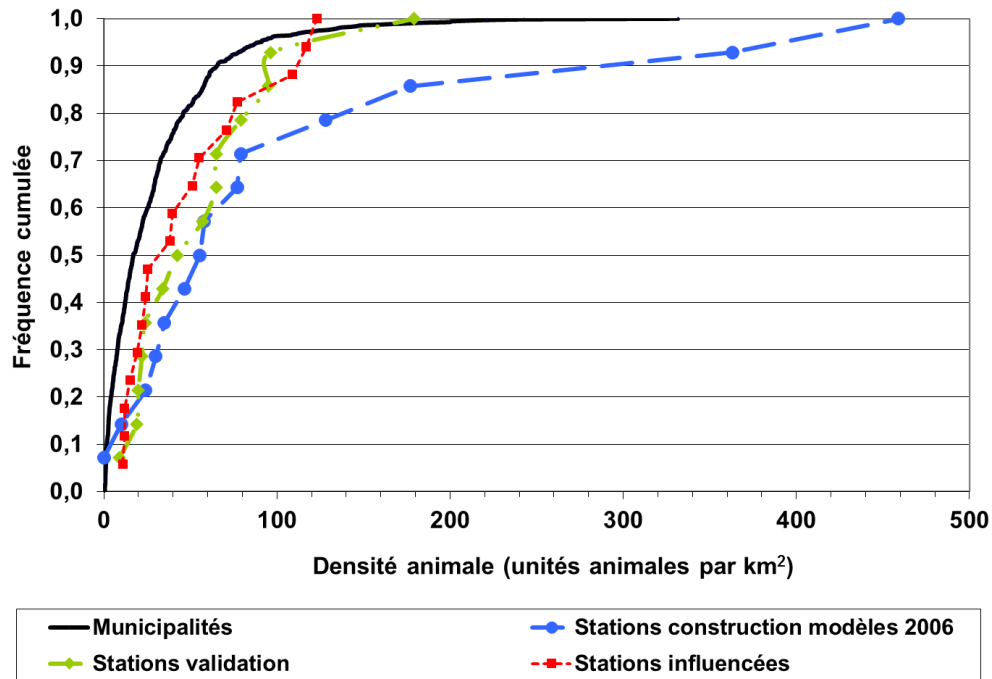


Figure A6.1 Fréquence cumulée des densités animales^a

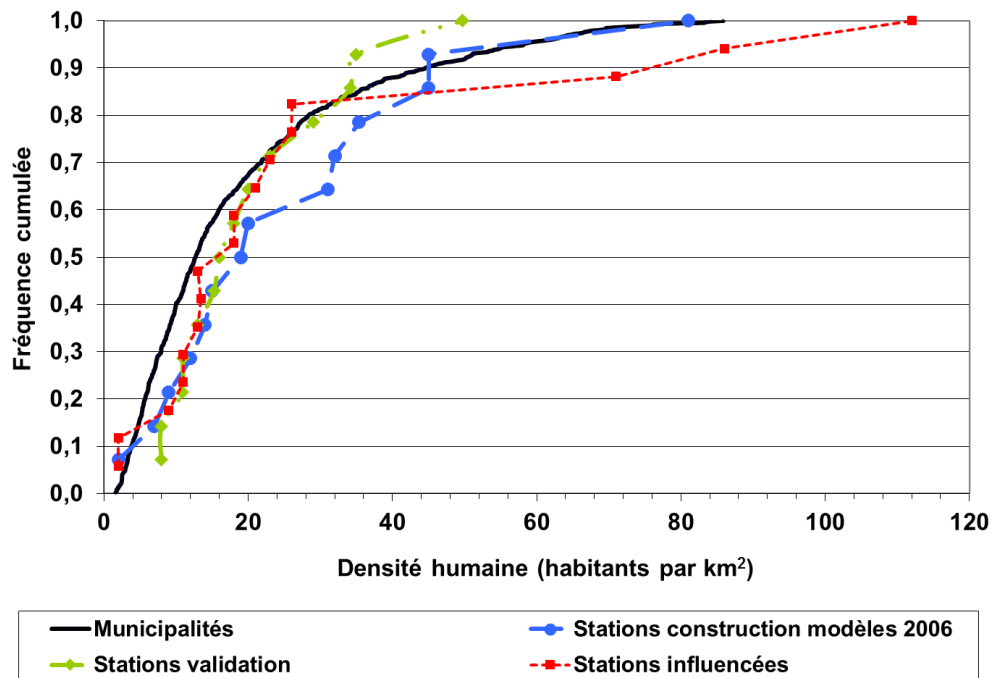


Figure A6.2 Fréquence cumulée des densités humaines^a

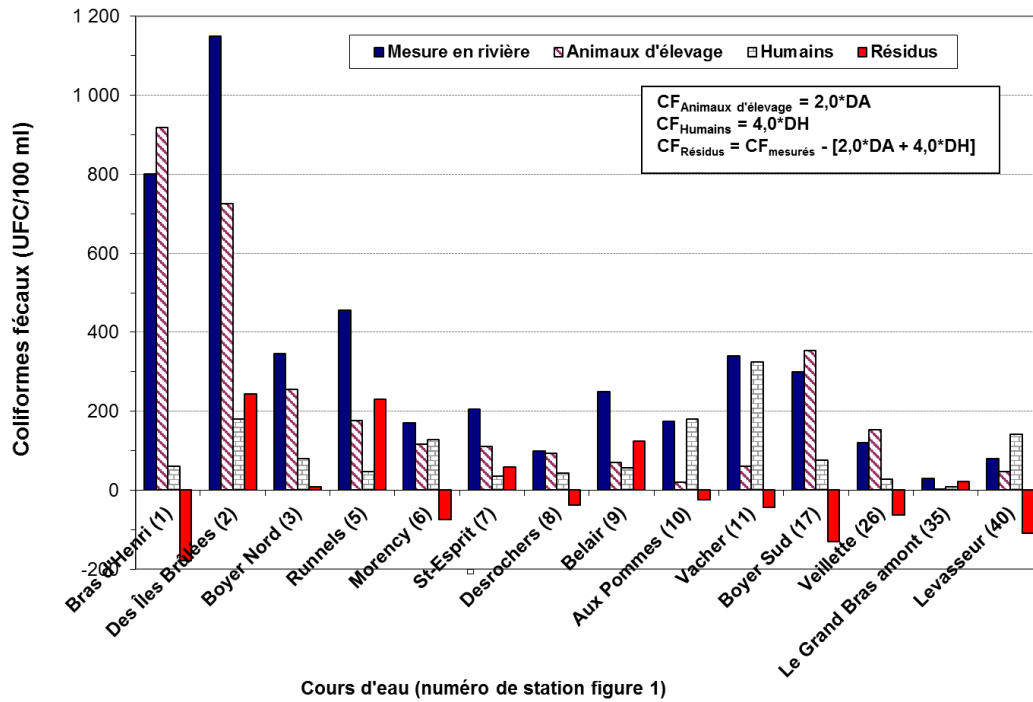
a : Fréquence estimée pour les petits bassins versants du Québec méridional à l'aide de la distribution des municipalités situées dans les régions administratives 01 à 05, 07, 12, 14, 16 et 17 et comportant une densité animale et humaine dans le domaine joint des modèles 2001 et 2006 (données de Statistique Canada et du Répertoire des municipalités du Québec pour l'année 2006).

Annexe 7 Utilisations possibles des relations entre les CF et le territoire

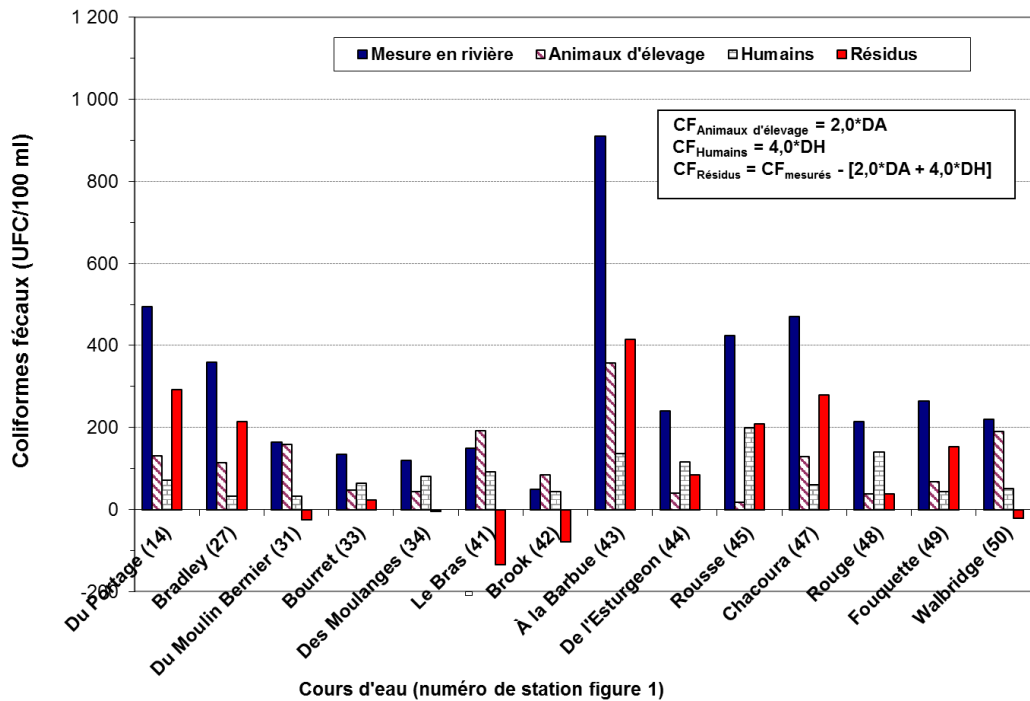
Modèle	Utilisation possible à des fins de prédiction (P) ou de diagnostic (D)	Prérequis	Utilisateur potentiel ^a
2006A (fig. 3 et A2.2) 2006A' ou 2006A'' (fig. 4 et A2.3)	(P) Médianes estivales de CF – conditions 2006	DA, DH ou S_déj, DA_amas ou DA, DA_amas	MELCC MAPAQ OBV MAMH MRC Municipalité
2001 univarié estival (fig. A1.1a)	(P) Médianes estivales de CF – conditions 2001 (DH 9 à 51 hab./km ²)	DA	
2001 bivarié estival (fig. A1.3 et A1.4)	(P) Médianes estivales de CF – conditions 2001 (DH > 51 hab./km ²)	DA, DH	
2006RL (fig. 6, 7 et A2.4)	(P) Dépassement du critère de contact indirect, direct et irrigation – conditions 2006	DA, DH	
2006RL' ou 2006RL'' (fig. A2.6 ou A2.9)	(P) Dépassement du critère de contact indirect – conditions 2006	S_déj, DA_amas ou DA, DA_amas	
2001RL (fig. A1.2)	(P) Dépassement du critère de contact indirect, direct et irrigation – conditions 2001 (DH 9 à 51 hab./km ²)	DA	
2006A (fig. 3) 2006RL (fig. 6)	(D) Identification de stations anormalement influencées par d'autres facteurs que DA et DH	CF mesurés et DA, DH	
2006A simplifié 2001A' ou 2006A'' (fig. 4)	(D) Contribution relative des sources animales (élevage) et humaines de CF	CF mesurés et DA, DH ou S_déj, DA_amas ou DA, DA_amas	
Équations ci-haut	(P) CF dans un bassin versant sans suivi de la qualité de l'eau (situation actuelle ou projetée après assainissement ou développement)	DA ou DA, DH ou S_déj, DA_amas ou DA, DA_amas	
Équations ci-haut	(P) Effet d'une intervention d'assainissement ou d'un projet de développement dans un bassin versant	DA ou DA, DH ou S_déj, DA_amas ou DA, DA_amas	

a : MELCC = Ministère de l'Environnement et de la Lutte contre les changements climatiques; MAPAQ = Ministère de l'Agriculture, des Pêcheries et de l'Alimentation; OBV = Organisme de bassin versant; MAMH = Ministère des Affaires municipales et de l'Habitation; MRC = Municipalité régionale de comté.

Annexe 8 Contribution des variables à la médiane de coliformes fécaux et résidus

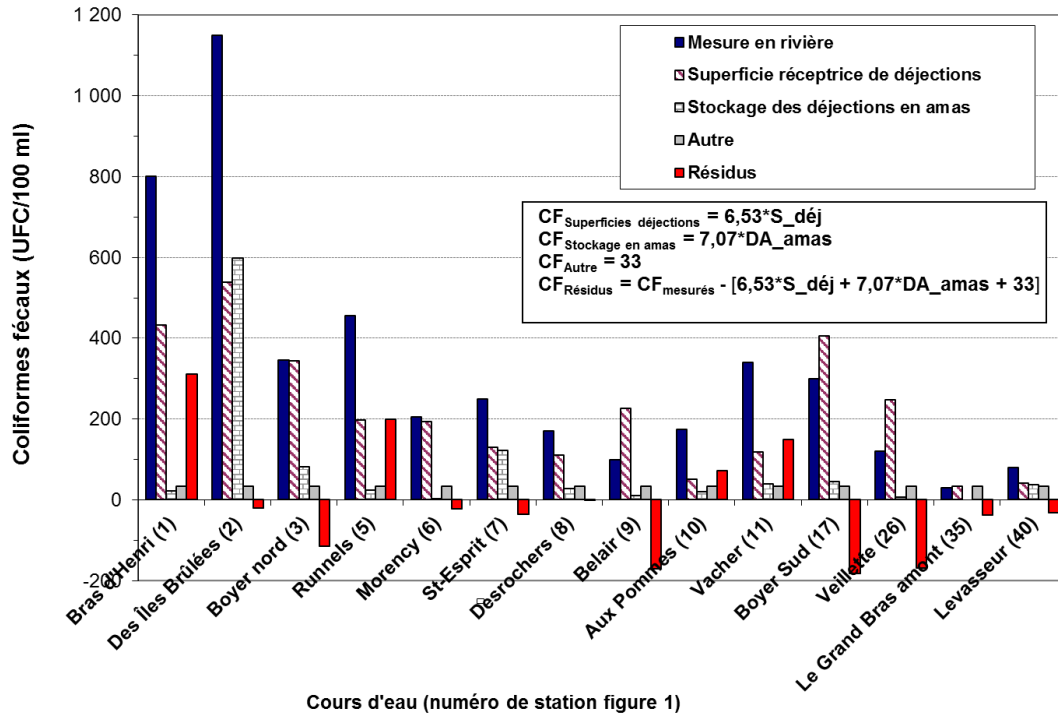


a) Stations utilisées pour la construction des modèles 2006

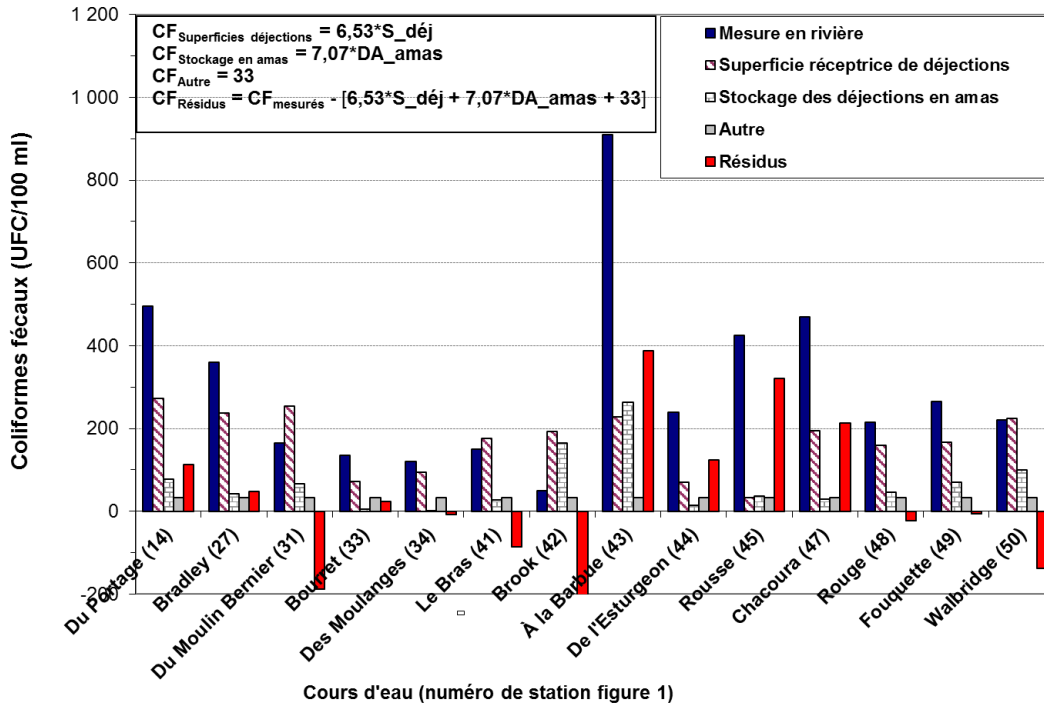


b) Stations utilisées pour la validation des modèles 2006

Figure A8.1 Concentrations médianes estivales de CF mesurées à 28 stations et contribution estimée des animaux d'élevage et des humains

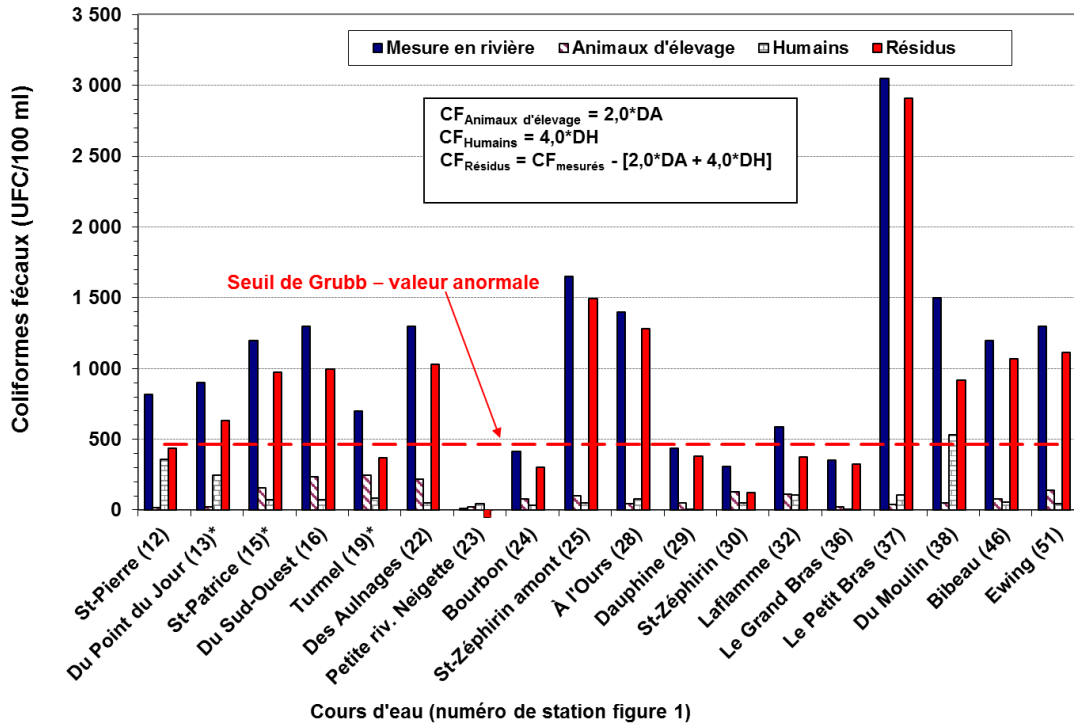


a) Stations utilisées pour la construction des modèles 2006



b) Stations utilisées pour la validation des modèles 2006

Figure A8.2 Concentrations médianes estivales de CF mesurées à 28 stations et contribution estimée des superficies réceptrices de déjections et du stockage en amas



Note : La période de mesure des CF considérée se situe entre 2005 et 2009, sauf aux stations suivies d'un astérisque, où elle est de 2000 à 2002 et à la station de la rivière Saint-Zéphirin (numérotée 30) où elle est de 2010 à 2012.

Figure A8.3 Concentrations médianes estivales de CF mesurées aux 18 stations influencées et contribution estimée des animaux d'élevage et des humains

Annexe 9 Exemples d'utilisations des relations pour prédire le risque de contamination

Les trois exemples suivants présentent des utilisations potentielles des relations développées pour prédire le risque de contamination. Ils permettent également d'illustrer la procédure de calcul et de comparer le résultat obtenu à l'aide des différents modèles.

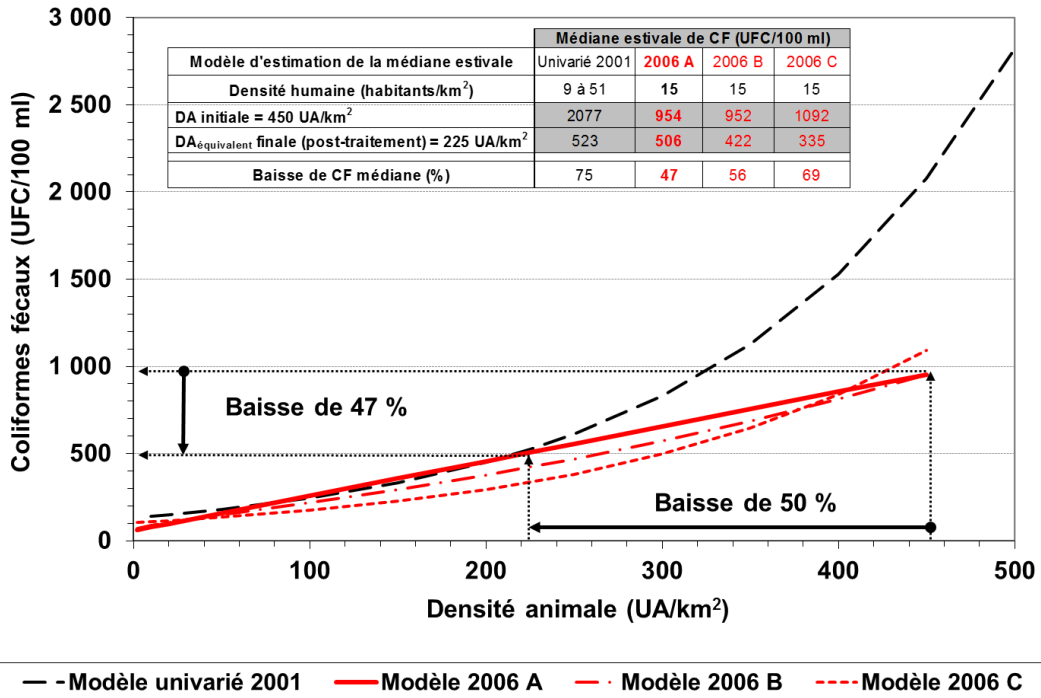
Le premier exemple illustre le cas d'un bassin versant sans station de mesure de la qualité de l'eau comportant 100 UA/km² et une densité humaine de 50 habitants/km². La médiane estivale des CF estimée à l'embouchure du bassin versant à l'aide du modèle 2001 univarié (figure A1.1a) serait de 643 UFC/100 ml (intervalle de prédiction à 95 % : de 85 à 698 UFC/100 ml), alors qu'elle serait de 398 UFC/100 ml (intervalle de prédiction à 95 % : de 74 à 722 UFC/100 ml) si le modèle 2006A était utilisé (figure 3 et figure A2.2). Le modèle 2006A forcé à LD/2 donnerait la même médiane estivale (398 UFC/100 ml), mais un intervalle de prédiction (figure A2.2) plus faible (de 92 à 704 UFC/100 ml). Le dépassement du critère de qualité de l'eau pour les activités de contact indirect comme la pêche, estimé à l'aide de la figure A1.2, serait de 20 % (intervalle de confiance à 95 % : de 15 à 27 %). Cet exemple montre la possibilité d'utiliser divers modèles selon que les pratiques agricoles s'apparentent à celles des années 2001 ou 2006 et selon le type de prédiction des valeurs de CF désiré (valeur médiane ou dépassement de critère de qualité d'eau). Ce type de résultat pourrait servir, par exemple, à déterminer les bassins versants les plus à risque.

La figure A9 illustre deux autres exemples d'utilisation des relations présentées dans le tableau de l'annexe 7. Le premier exemple (figure A9a) estime l'effet du traitement des déjections de la moitié des UA d'un bassin versant comportant une agriculture très intensive et une faible densité de population (DA = 450 UA/km²; DH = 15 UA/km²; S_déj = 66 ha/km²; DA_amas = 3 UA/km²), comme celui de la rivière Bras d'Henri. Dans le cas d'un traitement qui élimine les bactéries des déjections, une baisse de l'ordre de 47 % de la médiane estivale des CF serait attendue selon le modèle 2006A, en considérant que le traitement équivaut à une baisse de 50 % de la DA.

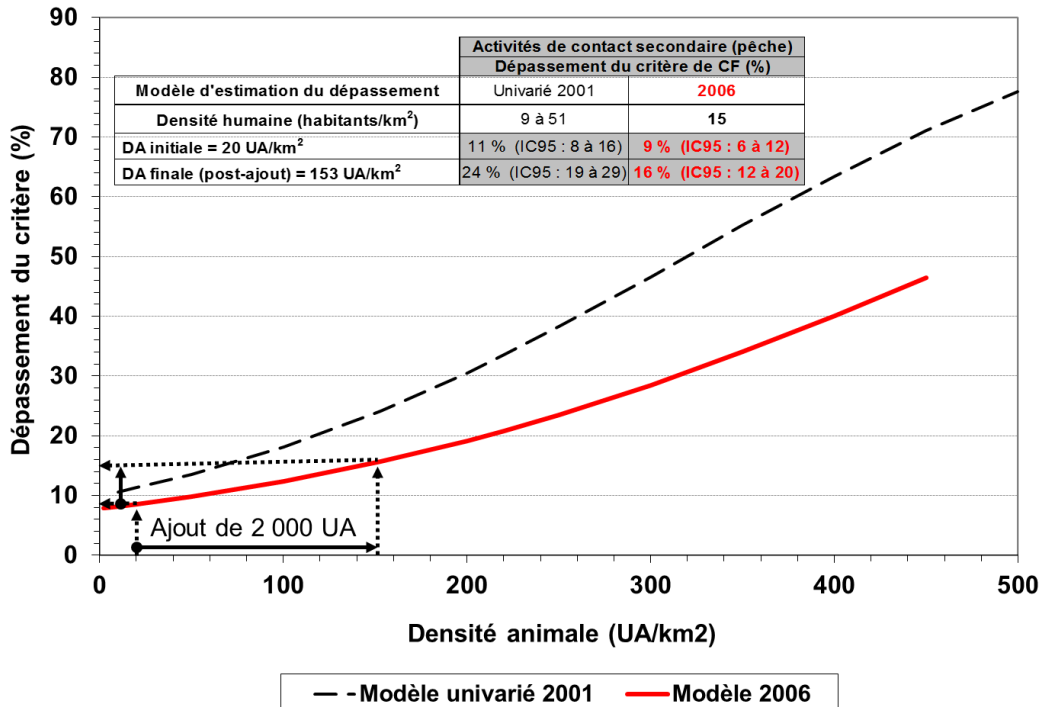
La baisse de CF attendue selon le modèle 2006A' en considérant une diminution de moitié des S_déj serait du même ordre (44 % ou 47 % selon que la DA_amas est maintenue à 3 UA/km² ou diminuée aussi de moitié). Ce type de résultat pourrait aider notamment à estimer l'effet de certaines pratiques comme le traitement des déjections animales à l'échelle de certains petits bassins versants ou encore à encourager l'adoption de ces pratiques.

Le second exemple (figure A9b) estime l'effet de l'ajout de 2 000 UA dans un petit bassin versant de 15 km² comportant une agriculture extensive et une faible densité de population (DA = 20 UA/km²; DH = 15 habitants/km²) et présentant un intérêt pour la pêche sportive. Le modèle 2006RL permet d'estimer que la fréquence de dépassement du critère de CF pour les activités de contact indirect comme la pêche passerait de 9 % (IC95 : de 6 % à 12 %) à 16 % (IC95 : de 12 % à 20 %) à la suite de cet ajout d'unités animales. Le modèle 2001RL, plus simple d'utilisation, mais considérant des pratiques comparables à celles qui prévalaient vers l'an 2000, donnerait plutôt un dépassement estimé du critère de CF qui passerait de 11 % (IC95 : de 8 % à 16 %) à 24 % (IC95 : de 19 % à 29 %).

La figure 9, qui illustre la contribution des amas (DA_amas) aux médianes estivales de CF dans les bassins versants ayant servi à la construction et à la validation des modèles, permet aussi de visualiser la baisse prévisible associée à l'entreposage étanche dans les bassins versants en enlevant la contribution des amas dans le modèle 2006A' (DA_amas = 0). Le même exercice peut aussi être réalisé à l'aide du modèle 2006A". Ce modèle permet également d'estimer à environ 6,6 UFC/100 ml la baisse des CF dans les cours d'eau pour chaque UA par kilomètre carré de bassin versant dont les déjections stockées en amas seraient entreposées dans un ouvrage étanche, en considérant que le coefficient multipliant la densité animale dans l'équation 2006A" passerait de 8,05 à 1,44 à la suite de l'entreposage étanche des déjections.



a) Effet du traitement des déjections de la moitié des unités animales sur la médiane estivale de CF



b) Effet de l'ajout de 2 000 UA dans un bassin versant de 15 km² avec une DA de 20 UA/km² sur le dépassement du critère pour la protection des activités de contact indirect comme la pêche

Figure A9 Exemples d'utilisation des relations

Annexe 10 Références bibliographiques des annexes

- Ministère des Affaires municipales et des Régions (MAMR), 2008. Système de suivi des ouvrages municipaux d'assainissement des eaux (SOMAE). Rapport annuel de performance des ouvrages de surverse.
- PATOINE, M., 2011. « Influence de la densité animale sur la concentration de coliformes fécaux dans les cours d'eau du Québec méridional, Canada », *Revue des sciences de l'eau*, vol. 24, n° 4, p. 421-435, [<http://www.erudit.org/revue/rseau/2011/v24/n4/1007628ar.html?vue=resume>].
- PATOINE. M. et F. D'Auteuil-Potvin, 2015. *Contamination bactériologique des petits cours d'eau en milieu agricole : état et tendances*, Québec, ministère du Développement durable, de l'Environnement et de la Lutte contre les changements climatiques, Direction du suivi de l'état de l'environnement, ISBN 978-2-550-72699-9 (PDF), 35 p. + 8 annexes.
- SAWYER, R., 1982. "Sample size and the accuracy of predictions made from multiple regression equations", *Journal of Educational and Behavioral Statistics*, vol. 7, no 2, p. 91-104.
- Statistique Canada, 2006. *Recensement de l'agriculture de 2006*. Disponible à <http://www.statcan.gc.ca/pub/95-629-x/2007000/4123856-fra.htm> (consulté le 3 janvier 2013).
- Statistique Canada, 2002. *Recensement de l'agriculture de 2001*, Données sur les exploitations agricoles diffusion complète. CD-ROM numéro 95F0304XCB.



**Environnement
et Lutte contre
les changements
climatiques**

Québec 