

Évaluation de la modélisation prédictive pour la plage Parlee



Préparée pour :

Bureau du médecin-hygiéniste en chef



Collaborateurs :

Diane M.L. Mas, Ph. D., REHS/RS

Sage Hardesty, M. Sc.

Eau et ressources naturelles

Fuss & O'Neill, Inc.

www.fando.com

20190099.A10

Photo de couverture : Ministère de l'Environnement et des
Gouvernements locaux du Nouveau-Brunswick

ISBN: 978-1-4605-2092-2

Table des matières

Table des matières	2
Résumé	4
Section 1 – Données existantes sur la plage Parlee..	8
1.1. Données existantes	9
1.2. Données sur les bactéries indicatrices de contamination fécale	12
1.3. Données météorologiques	14
1.4. Données environnementales.....	15
1.5. Données manquantes	16
1.6. Conclusions et recommandations	17
Section 2 – Analyse exploratoire des données relatives à la plage Parlee	19
2.1 Objet et objectifs.....	20
2.2 Résumé des bactéries indicatrices de contamination fécale	20
2.3 Prétraitement des données	21
2.4 Analyse graphique	22
2.5 Analyse statistique.....	36
2.6 Conclusions et recommandations	41
Section 3 – Mise au point et évaluation d'un modèle préliminaire	44
3.1 Introduction.....	45
3. 2 Démarche de modélisation	45
3. 3 Évaluation du modèle préliminaire.....	48
3. 4 Modèles de précipitations antérieures..	52
3. 5 Conclusions et recommandations	59
Glossaire	63
Bibliographie	66

Cette page est laissée vide délibérément.

Résumé

Le rapport intitulé *Qualité de l'eau de la plage Parlee, Shediac (Nouveau-Brunswick) : rapport final* (avril 2018) récemment achevé comprenait une recommandation relative à l'exploration de la mise au point d'un outil visant à prédire la qualité de l'eau, en se fondant sur des données météorologiques et environnementales pertinentes, que pourrait utiliser le médecin-hygiéniste (MH) afin d'émettre des interdictions de baignade à la plage Parlee. Actuellement, les méthodes de culture normalisées servant à déterminer la qualité de l'eau de la plupart des plages au Canada et aux États-Unis au moyen de la mesure des concentrations de bactéries indicatrices de contamination fécale (entérocoques et *E. coli*) prennent au moins de 18 à 24 heures avant que les résultats ne soient disponibles (Francy et coll., 2013). Lorsque des méthodes de culture sont utilisées, les conditions de qualité de l'eau et tout affichage d'interdictions sont déterminés en fonction de la concentration de bactéries de la veille. Un outil, tel qu'un modèle de prévision, offrirait des possibilités de prédiction qui donneraient suite aux préoccupations relatives à l'exposition du public à des concentrations possiblement élevées de bactéries à cause du délai entre l'échantillonnage de la qualité de l'eau et la réception des résultats de laboratoire indiquant les concentrations d'entérocoques ou d'*E. coli*. La présente étude visait à évaluer la faisabilité d'un modèle de prévision pour la plage Parlee et les questions connexes relatives au programme actuel de surveillance de la qualité de l'eau ainsi qu'à l'utilisation du seuil de précipitation antérieure pour les interdictions.

Quelle est la qualité de l'eau de la plage Parlee?

En 2017 et 2018, des échantillons d'eau ont été prélevés à cinq emplacements le long de la plage Parlee et ont été analysés en vue de détecter la présence de bactéries indicatrices de contamination fécale (entérocoques et *E. coli*) à des fins de comparaison avec les Recommandations au sujet de la qualité des eaux utilisées à des fins récréatives au Canada (Santé Canada, 2012). Les dépassements de la moyenne géométrique recommandée étaient de 5 % pour les entérocoques et de < 1 % pour *E. coli* pour la période de relevé 2017-2018. De même, les dépassements de la concentration maximale dans un échantillon unique pour les entérocoques représentaient 3 % de toutes les observations aux cinq emplacements de surveillance de la plage pendant la période de relevé. Ces résultats représentent des conditions de bonne qualité de l'eau de la plage. Aux fins de comparaison, l'Environmental Protection Agency (EPA) des États-Unis a déterminé que, lorsque 16 % des échantillons dépassent la norme de qualité des eaux utilisées à des fins récréatives, cela révèle une « mauvaise qualité de l'eau » (Zepp et coll., 2010). De son côté, le Natural Resources Defense Council (NRDC) a, dans son rapport de 2014 intitulé *Testing the Waters* portant sur les plages américaines (Dorfman et coll., 2014), déterminé que les plages où 2 % ou moins des échantillons dépassent la norme ou la recommandation pertinente ont une qualité d'eau exceptionnelle (c'est-à-dire que ce sont des « superstars »). Par contre, lorsque 25 % ou plus des échantillons dépassent la norme ou la recommandation pertinente, une plage est considérée comme présentant des problèmes chroniques sur le plan de la qualité de l'eau (Dorfman et coll., 2014). De manière générale, 10 % de tous les échantillons prélevés par les 30 États qui transmettent des données à l'EPA des États-Unis et cités dans le rapport du NRDC excédaient la norme pertinente de qualité de l'eau. Compte tenu des données de référence de 2 %, de 10 % et de 25 %, la qualité de l'eau de la plage Parlee n'est pas loin d'être exceptionnelle, les dépassements sont légèrement inférieurs à ceux que l'on observe habituellement sur les plages américaines et la qualité est bien en deçà des niveaux qui la caractériseraient comme ayant une mauvaise qualité de l'eau ou des problèmes chroniques à cet égard. Il

est intéressant et important de noter que l'analyse exposée en détail à la **section 2** montre que la majorité des dépassements des limites recommandées de qualité des eaux utilisées à des fins récréatives à la plage Parlee ont eu lieu en juillet, en août et en septembre, sans égard aux précipitations.

Est-ce qu'un seuil de précipitation s'avère un critère utile en matière d'interdiction préventive?

L'utilisation d'un seuil de précipitation sur lequel fonder la délivrance d'une interdiction est une démarche empirique largement employée selon laquelle les concentrations élevées de bactéries indicatrices, et donc les éventuels agents pathogènes dans l'eau, sont habituellement associées à un certain seuil de précipitation. Ce modèle empirique s'appuie sur le modèle conceptuel selon lequel les précipitations et l'écoulement subséquent sont des mécanismes de transport des bactéries indicatrices vers les eaux littorales utilisées à des fins récréatives. Puisque des échantillons de la qualité de l'eau sont prélevés quotidiennement à la plage Parlee, il est possible de vérifier la précision du seuil de précipitation comme outil de prédiction.

Une analyse des données relatives aux entérocoques en ce qui a trait au seuil actuel de 10 mm de pluie au cours des 24 heures précédentes présentée à la **section 3** démontre que le seuil prédirait correctement environ 38 % des jours (5 des 13 jours en 2017-2018) de dépassement de la moyenne géométrique de 35 NPP/100 mL. Toutefois, 18 journées supplémentaires donneraient lieu à l'affichage d'interdictions avec ce seuil, alors que les concentrations de bactéries indicatrices ont par la suite fait état de valeurs inférieures aux limites recommandées (c'est-à-dire faux positifs). Si le seuil passait à 25 mm, seuls 2 des 13 jours (15 %) seraient affichés correctement, alors que 7 autres jours feraient l'objet d'interdictions injustifiées et que 11 jours n'auraient pas donné lieu à des interdictions même si les concentrations de bactéries indicatrices dépassaient la valeur moyenne géométrique recommandée. Il convient aussi de noter que, pendant les quatre jours de la période de relevé affichant les concentrations moyennes géométriques les plus élevées d'entérocoques, un seul aurait été correctement prédit comme étant une journée visée par une interdiction de baignade par le seuil de précipitation de 10 mm ou de 25 mm. De même, si l'ensemble des cinq stations de surveillance est pris en compte et comparé à la concentration maximale recommandée dans un échantillon unique de 70 NPP/100 mL d'entérocoques, le seuil de 10 mm prédirait environ 30 % des observations totales (journées d'interdiction de baignade) ayant dépassé la limite recommandée et le seuil de 25 mm en prédirait environ 19 %. En résumé, le seuil de précipitation est un prédicteur incertain de la qualité de l'eau à la plage Parlee. Alors que le seuil de précipitation vient appuyer jusqu'à un certain point la prise de décision, dans le cas des seuils de 10 mm et de 25 mm, il le fait au prix de presque trois fois plus d'interdictions de baignade que la qualité réelle de l'eau l'exigerait, tout en n'arrivant pas à déceler plus de la moitié des jours où la qualité de l'eau exigerait l'affichage d'interdictions.

Des modifications sont-elles recommandées concernant la surveillance actuelle de la plage?

La collecte quotidienne de données, qui offre un outil utile pour la surveillance et le suivi de la qualité des eaux utilisées à des fins récréatives, devrait se poursuivre pour le moment dans toute la mesure du possible. Alors qu'*E. coli* et les entérocoques peuvent servir d'organismes indicateurs, les entérocoques sont habituellement reconnus comme étant plus robustes à ce titre dans les eaux de mer ou d'estuaire (EPA des États-Unis,

2004). Par conséquent, si une bactérie indicatrice devait être utilisée aux fins de surveillance, les entérocoques seraient recommandés. La turbidité est recommandée comme paramètre supplémentaire de la qualité de l'eau sur place en raison de ses antécédents comme indicateur utile de la qualité de l'eau et comme variable éventuellement importante dans les modèles de prévision (Francy et Darner, 2006). Francy et Darner (2006) ont recommandé que la turbidité soit mesurée sur place à l'aide d'un turbidimètre de terrain ou *in situ* à l'aide d'un appareil de mesure de la qualité de l'eau. Si la turbidité est mesurée dans le cadre d'analyses de laboratoire, les échantillons devraient être conservés sur de la glace pendant le transport jusqu'au laboratoire. Anderson (2005) recommande d'utiliser le même type d'instrument pour les mesures de la turbidité, car ils pourraient ne pas produire de résultats équivalents si leur conception différait. La **section 2** du rapport présente plus d'information sur les turbidimètres portatifs. La collecte des données météorologiques de base – pluie et direction du vent – devrait continuer d'être surveillée. Toute personne recueillant des données sur la qualité de l'eau ou de l'environnement devrait recevoir une formation chaque année au début de la saison de fréquentation de la plage pour veiller à ce que des méthodes de collecte uniformes soient utilisées afin de permettre la comparaison des données d'une année à l'autre. L'utilisation de listes de vérification et la préparation de documents d'orientation peuvent être utiles pour la formation et pour maintenir l'uniformité entre les saisons et les différentes personnes qui recueillent les données.

La plage Parlee se prête-t-elle à l'utilisation de modèles de prévision?

L'étude a examiné la possibilité de mettre au point des modèles de régression linéaire multiple pour prédire les concentrations d'organismes indicateurs. L'analyse exploratoire des données décrite à la **section 2** a laissé entendre que la mise au point de modèles robustes et statistiquement significatifs pour la plage Parlee serait peu probable en raison du nombre relativement faible de dépassements des concentrations de bactéries indicatrices par rapport aux limites recommandées de qualité de l'eau. Habituellement, il faut que de 20 à 30 % des données dépassent la valeur d'intérêt de la limite conseillée de qualité de l'eau pour qu'un modèle ait une forte capacité de prédiction dans un large éventail de valeurs. Des modèles de régression linéaire multiple ont été mis au point dans la version 3.0.6 (VB3.0.6) du logiciel de modélisation Virtual Beach. Il convient de noter qu'un modèle de régression linéaire multiple peut être valide, c'est-à-dire respecter toutes les exigences en matière de RLM, sans offrir de fortes capacités de prédiction concernant la variable d'intérêt.

La mise au point du modèle s'est concentrée sur la prévision de concentrations d'entérocoques dans un échantillon unique et de moyennes géométriques à l'aide d'une série d'éventuelles variables météorologiques et environnementales indépendantes. Voici les variables indépendantes qui sont ressorties à titre de prédicteurs statistiquement significatifs : précipitations au cours des 24 heures précédentes, jours depuis la précipitation, augmentation du niveau d'eau, direction du vent et changement de la direction du vent. Les modèles mis au point pour prédire les concentrations de bactéries indicatrices durant la dernière partie de la saison de fréquentation de la plage (de juillet à septembre) étaient généralement de meilleurs prédicteurs selon la valeur R carré ajustée. Cependant, les valeurs R carré ajustées étaient généralement faibles et le modèle ayant eu le meilleur rendement a affiché une valeur R carré ajustée moyenne de 0,4954, indiquant qu'environ la moitié de la variabilité des données relatives aux bactéries indicatrices de contamination fécale pouvait être expliquée par les variables indépendantes du modèle. Les modèles de régression linéaire multiple se caractérisaient par une sensibilité

généralement faible, ce qui signifie qu'ils n'ont pas très bien réussi à déceler les vrais dépassements des limites recommandées de qualité des eaux utilisées à des fins récréatives. Lors de l'évaluation en fonction de la sensibilité, le meilleur modèle de régression linéaire multiple avait une valeur de sensibilité de 0,25, utilisant les précipitations au cours des 24 heures précédentes de juillet à septembre et le nombre de jours depuis les précipitations comme variables indépendantes. Toutefois, les modèles de régression linéaire multiple ont généralement produit moins de faux positifs, comme le démontre la plus grande spécificité de ces modèles (c'est-à-dire la capacité de prédire correctement les valeurs se situant en deçà de la valeur recommandée) par rapport au seuil de précipitation antérieure de 10 mm. De manière générale, les modèles de prévision mis au point dans le cadre de la présente étude n'affichent pas de solides rendements quant à la capacité de prédire le dépassement de la limite recommandée de qualité de l'eau sur le plan de la moyenne géométrique ou celui d'un échantillon unique. Si un projet pilote de modélisation devait être lancé à la plage Parlee, les modèles décrits à la **section 3** pour la période allant de juillet à septembre seraient recommandés.

En raison du faible nombre de dépassements des limites recommandées de qualité des eaux utilisées à des fins récréatives et des restrictions que cela comporte pour la modélisation prédictive, les autres méthodes d'évaluation rapide des conditions de la qualité de l'eau sur le plan microbien ne devraient pas être écartées. La réaction de polymérisation en chaîne quantitative, ou RPC quantitative, est une méthode rapide employée dans plusieurs plans d'eau utilisés à des fins récréatives aux États-Unis, qui a été acceptée par l'EPA des États-Unis pour l'évaluation des conditions de qualité de ces eaux pour les plages tant d'eau douce que de mer. Une combinaison du modèle de régression linéaire multiple ou du seuil de précipitation antérieure et de la RPC quantitative pourrait aussi être envisagée. Par exemple, le seuil de précipitation actuel pour les interdictions de baignade préventives a tendance à surestimer les jours où la qualité de l'eau dépasse réellement la limite recommandée, c'est-à-dire qu'il produit nombre de faux négatifs. Si une valeur de précipitation au cours des 24 heures précédentes, qui est aussi une variable météorologique prévisible, était utilisée comme outil de dépistage, alors, lorsque les précipitations dépassent le seuil, la RPC quantitative pourrait être utilisée avec les échantillons prélevés. Cela permettrait de créer un système de prévision de la qualité de l'eau qui pourrait orienter l'utilisation de l'analyse à l'aide de la RPC quantitative de manière ciblée et rentable, particulièrement pendant la période allant de juillet à septembre, et éventuellement réduire le nombre de « faux positifs » découlant de la seule utilisation du seuil de précipitation.

Section 1 – Données existantes sur la plage Parlee

1.1. Données existantes

Comme en fait mention le rapport *Qualité de l'eau de la plage Parlee : rapport final* (2018), nombre de facteurs peuvent avoir une incidence sur la qualité de l'eau à la plage Parlee mesurée en fonction des concentrations de bactéries indicatrices de contamination fécale (*E. coli* et entérocoques). Ces différents facteurs comprennent les sources de bactéries provenant des activités menées en amont ou dans le vent dans le bassin hydrologique de la plage, de la remise en suspension des sédiments et des activités des baigneurs, des plaisanciers et de la faune sur la plage, ainsi que de variables environnementales comme la pluie, le vent et les marées, qui influent sur le sort et le déplacement des bactéries.

La première étape permettant de comprendre la relation possible entre les concentrations de bactéries et les variables environnementales consiste à recueillir des données et à les résumer afin de connaître les renseignements disponibles, leur exhaustivité et les données temporelles manquantes pouvant avoir une incidence sur la possibilité de mettre au point un modèle de prévision ou par ailleurs d'évaluer les relations entre les bactéries indicatrices de contamination fécale et d'autres données environnementales plus facilement mesurées ou estimées.

Les données existantes sur la plage Parlee ont été déterminées dans le cadre de discussions avec le ministère de l'Environnement et des Gouvernements locaux et le Bureau du médecin-hygiéniste en chef, examinées pour en assurer l'exhaustivité et résumées dans une série de tableaux. Le **tableau 1-1** présente un condensé de toutes les données pouvant faire l'objet d'une collecte ou de mesures continues pendant la saison type de fréquentation de la plage (de mai à octobre). Les données recueillies dans le cadre d'une enquête autonome sur les sources potentielles de bactéries sont résumées séparément au **tableau 1-2**. Chacun des types d'ensembles de données est abordé plus en détail et la présente section se conclut par la détermination des périodes de chevauchement de collecte de variables environnementales au cours des saisons de fréquentation de la plage de 2017 et de 2018.

Tableau 1-1 Résumé des ensembles de données existants sur la plage Parlee

Ensemble de données	Paramètres	Fréquence	Période de relevé	Source	Lacunes
Station météorologique de la plage Parlee	Précipitations (mm)	Horaire	2017-06-05 au 2017-10-10 2018-06-18 au 2018-09-15	Ministère de l'Environnement et des Gouvernements locaux	2017-09-04 au 2017-09-09 2018-06-24 au 2018-06-28
	Durée des précipitations (10 s/compte)				
	Intensité de la pluie (mm/h)				
	Direction du vent (degrés)				
	Vitesse du vent (km/h)				
	Vitesse moyenne du vent (km/h)	Quotidienne	2017-06-05 au 2017-10-10		2017-09-04 au 2017-09-09
Observations des sauveteurs	Temp. de l'air (°C)	Deux fois par jour	2017-06-06 au 2017-08-19 2018-06-03 au 2018-08-25	Liste de vérification quotidienne des sauveteurs Ministère du Tourisme, du Patrimoine et de la Culture	2017-06-17
	Temp. de l'eau (°C)				
	Qualité de l'eau				
	Qualité du sable				
	Nbre d'utilisateurs de la plage				
	Nbre de baigneurs dans l'eau				
	Nbre d'oiseaux	Quotidienne			
	Direction du vent				
	Force du vent				
	Nbre de bateaux ancrés				
	Niveau de la marée				
	Nbre de bateaux passant devant la plage				
	Couverture nuageuse				
Rivière Coal Branch (01BS001)	Niveaux d'eau moyens (m)	Quotidienne	2017-01-01 au 2018-12-31	Division des relevés hydrologiques du Canada/Environnement et Changement climatique Canada	Aucune
Rivière Petitcodiac (01BU002)	Niveaux d'eau moyens (m)	Quotidienne	2017-01-01 au 2018-12-31	Division des relevés hydrologiques du Canada/Environnement et Changement climatique Canada	Aucune

Tableau 1-1 (suite) Résumé des ensembles de données existants sur la plage Parlee

Ensemble de données	Paramètres	Fréquence	Dates de couverture	Source	Lacunes
<i>Qualité de l'eau de la plage Parlee</i>	Précipitations dans les 24 heures précédentes (mm)	Quotidienne	2017-05-17 au 2018-10-10	Ministère de l'Environnement et des Gouvernements locaux	Aucune
	<i>E. coli</i> (NPP)		2017-05-15 au 2017-10-09		2017-10-02, 2017-05-27
	<i>E. coli</i> (comptes)		2018-05-22 au 2018-10-08		2018-06-02, 2018-06-03, 2018-06-14, 2018-08-18, 2018-09-19
	<i>Entérocoque</i>		2017-05-15 au 2017-10-10, 2018-05-22 au 2018-10-08		2017-05-27, 2017-10-02, 2017-10-10, 2018-06-02, 2018-06-03, 2018-06-14, 2018-08-18, 2018-09-19

Il convient de noter que les emplacements d'échantillonnage mentionnés au **tableau 1-2** se trouvent dans le rapport intitulé *Qualité de l'eau de la plage Parlee, Shediac (Nouveau-Brunswick) : rapport final* (2018), disponible à l'adresse https://www2.gnb.ca/content/dam/gnb/Departments/eco-bce/Promo/Parlee_Beach/pdfs/parlee_beach_water_quality_finalreport-f.pdf.

Tableau 1-2 Données sur les bactéries indicatrices de contamination fécale recueillies lors d'évènements distincts d'échantillonnage dans le bassin hydrographique

Indicateur	Unités	Emplacement d'échantillonnage	Dates de l'échantillonnage	Média
<i>Entérocoque</i>	NPP/100 mL	Lac des Boudreau Mudflats Zone humide entre Cap-Brûlé et The Bluff Étang au sud de la plage Parlee Sortie du ruisseau de la plage Parlee au nord de la rue St.John Estuaire South Cove	2017-09-06, 2017-09-07, 2017-10-30	Sédiment
<i>E. coli</i>	NPP/100 mL	Lac des Boudreau Mudflats Zone humide entre Cap-Brûlé et The Bluff Étang au sud de la plage Parlee Sortie du ruisseau de la plage Parlee au nord de la rue St. John Estuaire South Cove		Sédiment
<i>E. coli</i> (FFE)	NPP/100 mL	Shediac A, Shediac B, Shediac C, Shediac E Shediac G, Shediac H WQ1, WQ2 WQ3, WQ4 WQ5, WQ6 WQ7, WQ8 WQ9, WQ10 WQ11	2017-06-22, 2017-07-19, 2017-08-22, 2017-09-20, 2017-10-18	Eau
<i>E. coli</i> (PRBH)	NPP/100 mL	AG1, AG2, AG3, FW1, FW2, FW3, FW4, FW5, FW6, FW7, FW8, FW9, FW10, FW11, E1, E2, E2B, E3, E4, M1, M2, M3, M4, M6, SED1, SED2, SED3, SED4, SED5, SW1, SW2, SW3, SW4, SW5, SW6, SW7, SW8	2017-07-27, 2017-08-09, 2017-08-10, 2017-08-21, 2017-09-07, 2017-09-21, 2017-10-11, 2017-10-27	Eau
Entérocoque (PRBH)	NPP/100 mL	AG1, AG2, AG3, FW1, FW2, FW3, FW4, FW5, FW6, FW7, FW8, FW9, FW10, FW11, E1, E2, E2B, E3, E4, M1, M2, M3, M4, M6, SED1, SED2, SED3, SED4, SED5, SW1, SW2, SW3, SW4, SW5, SW6, SW7, SW8		Eau

1.2. Données sur les bactéries indicatrices de contamination fécale

Les bactéries indicatrices de contamination fécale (*E. coli* et entérocoques) ont été mesurées quotidiennement à la plage Parlee tout au long des saisons de fréquentation de

la plage de 2017 et de 2018 (**tableau 1-3**). Les emplacements et les détails de l'échantillonnage sont présentés dans le rapport *Qualité de l'eau de la plage Parlee : rapport final* (2018). Le changement d'unités dans les rapports sur *E. coli* révèle un changement de méthode de dénombrement employée par le laboratoire entre les saisons 2017 et 2018. Ce changement devrait avoir une incidence minimale sur la comparaison ou l'utilisation des données des deux années, qui peuvent être utilisées ensemble pour effectuer d'autres analyses et modélisations des données.

Les données manquantes sont minimes en 2017, alors qu'il manque seulement trois jours de données pour les entérocoques et deux jours pour *E. coli*. Même s'il manque cinq jours de données pour *E. coli* et les entérocoques en 2018, cela représente moins de 4 % des données recueillies en 2018.

Tableau 1-3 Résumé des données sur les bactéries indicatrices de contamination fécale

Indicateur	Unités	Fréquence	Période de relevé	Source	Lacunes
<i>Entérocoque</i>	NPP/100 mL	Quotidienne	2017-05-15 au 2017-10-10, 2018-05-22 au 2018-10-08	Ministère de l'Environnement et des Gouvernements locaux	2017-05-27, 2017-10-02, 2017-10-10, 2018-06-02, 2018-06-03, 2018-06-14, 2018-08-18, 2018-09-19
<i>E. coli</i> (NPP)	NPP/100 mL	Quotidienne	2017-05-15 au 2017-10-09		2017-10-02, 2017-05-27
<i>E. coli</i> (comptes)	UFC/100 mL	Quotidienne	2018-05-22 au 2018-10-08		2018-06-02, 2018-06-03, 2018-06-14, 2018-08-18, 2018-09-19

1.3. Données météorologiques

Les données météorologiques relatives à la plage Parlee proviennent de trois (3) ensembles de données (**tableau 1-4**). Les mesures quotidiennes des précipitations au cours des 24 heures précédentes ont été consignées et communiquées avec les mesures des concentrations de bactéries indicatrices de contamination fécale pendant les saisons de fréquentation de la plage de 2017 et de 2018. Les sauveteurs travaillant à la plage Parlee ont noté la température de l'air et de l'eau deux fois par jour pendant les mois de juin, de juillet et d'août pendant les saisons de fréquentation de la plage de 2017 et de 2018. La quantité, la durée et l'intensité des précipitations, ainsi que la direction et la vitesse du vent ont été notées toutes les heures par la station météorologique de la plage Parlee pendant les saisons de fréquentation de la plage de 2017 et de 2018, même si l'exactitude des données sur les précipitations n'a pas été confirmée¹. La couverture des deux ensembles de données est très bonne, alors qu'il ne manque que six jours ou moins de données dans chacun d'entre eux. La vitesse moyenne quotidienne du vent a aussi été communiquée en 2017 par la station météorologique.

Tableau 1-4 Données météorologiques

Variable	Fréquence	Période de relevé	Source	Lacunes
24 heures Précipitations antérieures (mm)	Quotidienne	2017-05-17 au 2018-10-10	Données relatives à la qualité de l'eau de la plage Parlee	Aucune
Temp. de l'air (°C)	Deux fois par jour	2017-06-06 au 2017-08-19	Liste de vérification quotidienne des sauveteurs – Ministère du Tourisme, du Patrimoine et de la Culture	2017-06-17
Temp. de l'eau (°C)		2018-06-03 au 2018-08-25		
Précipitations (mm)	Horaire	2017-06-05 au 2017-10-10 2018-06-18 au 2018-09-15	Station météorologique de la plage Parlee	2017-09-04 au 2017-09-09 2018-06-24 au 2018-06-28
Durée des précipitations (10 s/compte)				
Intensité de la pluie (mm/h)				
Direction du vent (degrés)				
Vitesse du vent (km/h)				
Vitesse moyenne du vent	Quotidienne			

¹ Même si les données sur les précipitations sont présentées ici, une communication avec le ministère de l'Environnement et des Gouvernements locaux a révélé qu'elles n'ont pas été confirmées et que les données relatives aux précipitations au cours des 24 heures précédentes présentées avec les concentrations de bactéries indicatrices de contamination fécale devraient être utilisées pour les analyses statistiques et de modélisation subséquentes.

1.4. Données environnementales

Les autres données environnementales (c'est-à-dire autres que des données météorologiques continues) recueillies pendant la surveillance de la qualité de l'eau de la plage sont de deux types – catégoriques et continues. Les données catégoriques sont des données pouvant être décrites par deux catégories ou plus, qui comprennent souvent une certaine part de subjectivité de la part de la personne qui les recueille. Dans le cas de la plage Parlee, les observations notées par les sauveteurs fournissent de l'information sur plusieurs variables catégoriques qui décrivent les conditions présentes à la plage (**tableau 1-5**). Même si la période de relevé se limite à juin, à juillet et à août, elle compte peu de lacunes et cet ensemble de données offre la seule information facilement accessible sur les marées, la couverture nuageuse et l'évaluation visuelle de la qualité de l'eau.

Tableau 1-5 Données environnementales catégoriques

Variable	Catégories	Fréquence	Période de relevé	Source	Lacunes
Qualité de l'eau	Mauvaise Faible Bonne S.O.	Deux fois par jour	2017-06-06 au 2017-08-19 2018-06-03 au 2018-08-25	Liste de vérification quotidienne des sauveteurs	2018-06-17
Qualité du sable	Bonne OK	Deux fois par jour			
Direction du vent	Directions	Quotidienne			
Force du vent	Fort Moyen Faible Nombres	Quotidienne			
Niveau de la marée	Élevé Faible S.O.	Quotidienne			
Couverture nuageuse	Dégagé Toute la journée Pluie Ensoleillé Nbre d'heures	Quotidienne			

D'autres données environnementales continues ont aussi été recueillies durant les saisons de fréquentation de la plage de 2017 et de 2018 (**tableau 1-6**). Les données sur le niveau d'eau sont consignées à deux stations limnimétriques avoisinantes, celle de la rivière Coal Branch (station 01BS001) et celle de la rivière Petitcodiac (station 01BU002), toutes deux entretenues par la Division des relevés hydrologiques du Canada/Environnement et Changement climatique Canada. Les données ne sont facilement accessibles qu'en ligne, en temps réel, pour les dates postérieures à 2017. Les données provisoires pour 2017 et 2018, qui ont été utilisées dans la présente analyse, ont été obtenues auprès de la Division des relevés hydrologiques du Canada/Environnement et Changement climatique Canada.

Les données quotidiennes consignées par les sauveteurs offrent certains renseignements supplémentaires sur les personnes fréquentant la plage et les baigneurs, qui sont une

source possible d'excrétion de matières fécales. Même si elles sont seulement disponibles pour juin, juillet et août, il y a peu de données manquantes.

Tableau 1-6 Données environnementales continues

Variable	Fréquence	Dates de couverture	Source	Lacunes
Rivière Coal Branch (01BS001) Niveau d'eau moyen (m)	Quotidienne	2017-01-01 au 2018-12-31	Division des relevés hydrologiques du Canada/Environnement et Changement climatique Canada	Aucune
Rivière Petitcodiac (01BU002) Niveau d'eau moyen (m)	Quotidienne	2017-01-01 au 2018-12-31	Division des relevés hydrologiques du Canada/Environnement et Changement climatique Canada	Aucune
Nbre de personnes fréquentant la plage	Deux fois par jour	2017-06-06 au 2017-08-19 2018-06-03 au 2018-08-05	Liste de vérification quotidienne des sauveteurs	Aucune
Nbre de baigneurs	Quotidienne	2017-06-06 au 2017-08-19 2018-06-03 au 2018-08-05	Liste de vérification quotidienne des sauveteurs	Aucune

1.5. Données manquantes

La **figure 1-1** résume la couverture de la période de relevé des ensembles de données décrits ci-dessus. Afin d'évaluer la relation entre les concentrations de bactéries indicatrices de contamination fécale à la plage Parlee et les données météorologiques et environnementales, les périodes de relevé des ensembles de données doivent se chevaucher. Les rectangles orange indiquent respectivement les zones de chevauchement en 2017 et en 2018. Par exemple, pour tous les types de données recueillies en 2018 pour servir à la création de modèles de prévision, la période de modélisation serait limitée aux mois de juin à août 2018. Ce sous-ensemble contient 84 jours de données, soit seulement 30 % du nombre de jours de fréquentation de la plage pour lesquels des données sur les bactéries indicatrices de contamination fécale ont été consignées en 2018. Les conditions sont semblables dans les données de 2017.

Si les données recueillies par les sauveteurs sont exclues, la période de relevé des données météorologiques de la station météorologique de la plage Parlee correspond à 87 % et 67 % de la période de relevé des données sur les bactéries indicatrices de contamination fécale de 2017 et de 2018 respectivement. Les trois paramètres dont les périodes de relevé coïncident avec les périodes de relevé des données sur les bactéries indicatrices de contamination fécale sont les données sur les précipitations au cours des 24 heures précédentes et celles relatives au niveau d'eau des rivières Coal Branch et Petitcodiac.

1.6. Conclusions et recommandations

Les données accessibles sur la qualité microbienne de l'eau, les conditions météorologiques et l'environnement ont été examinées pour pouvoir les résumer et en cerner les lacunes temporelles possibles. Voici les observations clés qui ont été faites :

- La couverture de données pendant la période de relevé des concentrations de bactéries indicatrices de contamination fécale, les précipitations au cours des 24 heures précédentes et le niveau d'eau des rivières avoisinantes est excellente.
- La méthode de laboratoire employée pour quantifier *E. coli* a été modifiée et, même si les résultats de chaque méthode devaient être comparables et utilisables ensemble, il faut garder ce changement à l'esprit au moment de l'examen des analyses subséquentes.
- Au sein de chaque ensemble de données, la couverture est généralement très bonne au cours de la période de relevé de l'ensemble, comptant peu de valeurs manquantes.
- Tant pour la saison de fréquentation de la plage de 2017 que pour celle de 2018, les données météorologiques disponibles couvrent une période plus courte que la période de relevé de l'échantillonnage des bactéries indicatrices de contamination fécale, ce qui limite la pleine utilisation des données sur ces bactéries aux fins de comparaisons avec les données météorologiques autres que les précipitations au cours des 24 heures précédentes.
- Les variables catégoriques introduisent la possibilité d'incohérence dans la description des conditions des données.
- Les échantillons d'eau et de sédiment dans le bassin hydrologique et les environs de la plage Parlee fournissent des « instantanés » de conditions. Alors qu'ils peuvent s'avérer utiles pour la mise au point de théories sur les sources de bactéries à la plage Parlee, la quantité limitée de données ne permet pas de les utiliser pour la réalisation d'analyses statistiques détaillées ou la création de modèles de prévision.

En fonction de cette analyse initiale, les recommandations suivantes sont aussi formulées :

- Dans la mesure du possible, les périodes de collecte de données par les sauveteurs et la station météorologique devraient coïncider avec le prélèvement d'échantillons d'eau aux fins de l'analyse des bactéries indicatrices de contamination fécale. Les analyses décrites aux **sections 2 et 3** indiquent que les données météorologiques sont plus déterminantes en ce qui a trait à la relation possible avec les concentrations de bactéries indicatrices de contamination fécale et un chevauchement assez large des périodes de relevé existe déjà entre ces variables et la collecte de données sur ces bactéries.
- Les sauveteurs devraient recevoir des directives détaillées et précises sur l'enregistrement des données catégoriques afin que les données consignées soient cohérentes.

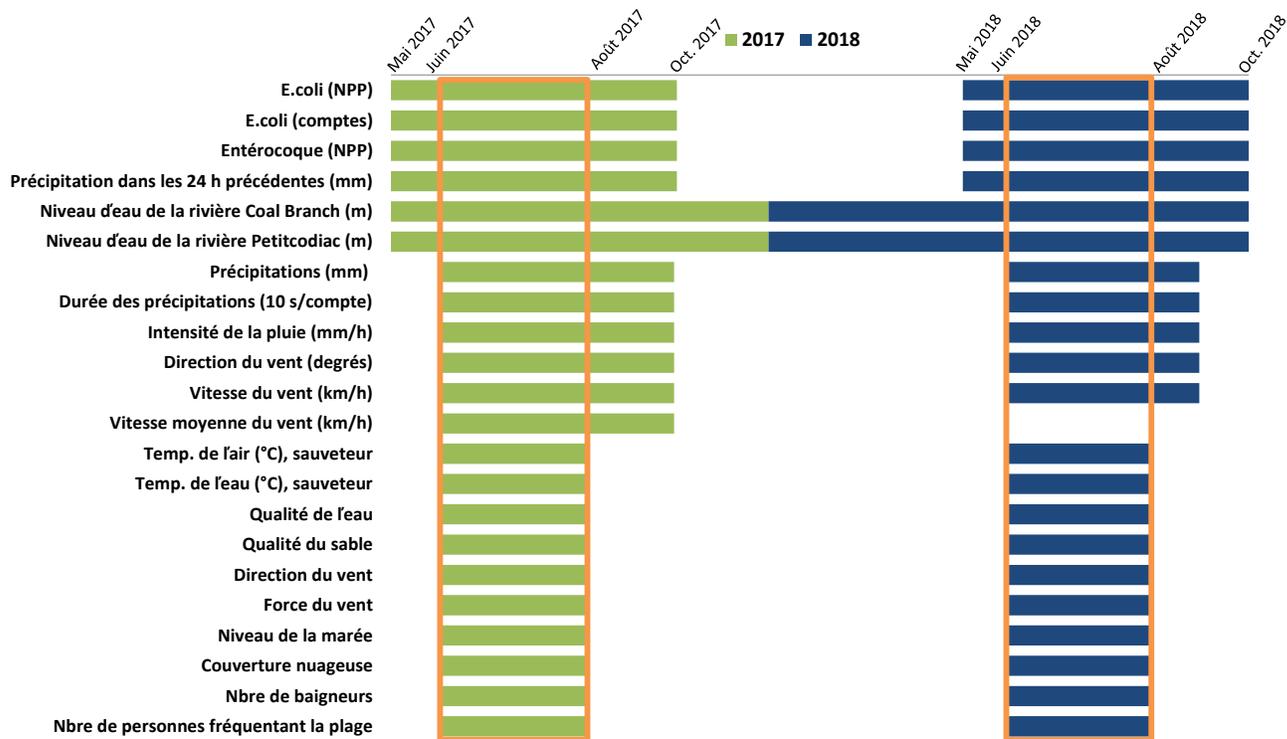


Figure 1-1. Couverture de la période de relevé des données relatives à la plage Parlee

Section 2 – Analyse exploratoire des données relatives à la plage Parlee

2.1 Objet et objectifs

L'analyse préliminaire des données est une première étape utile pour cerner les éventuelles relations importantes entre les bactéries indicatrices de contamination fécale comme les entérocoques et *E. coli* et les variables indépendantes qui caractérisent les conditions météorologiques ou autres conditions environnementales. La compréhension des relations éventuelles appuie le processus de sélection des variables pour la création d'un modèle de prévision tout en éclairant les efforts de collecte de données. Les variables météorologiques ou autres variables environnementales démontrant peu de rapport avec les bactéries indicatrices de contamination fécale ou des relations qui ne sont pas statistiquement significatives ont une valeur moindre en ce qui a trait à la détermination des conditions dans lesquelles les concentrations de ces bactéries sont élevées. Cette information peut s'avérer utile pour prioriser les efforts de collecte de données. L'analyse décrite dans le présent rapport, qui comprend une analyse à la fois graphique et statistique, utilise Microsoft Excel, Minitab 18 et la version 3.0.6 de Virtual Beach (VB3.0.6). Les données utilisées dans l'analyse sont résumées à la **section 1**.

2.2 Résumé des bactéries indicatrices de contamination fécale

La **figure 2-1** et le **tableau 2-1** présentent un résumé graphique et tabulaire des données recueillies sur la concentration d'entérocoques et d'*E. coli* en 2017 et 2018. Les données sur les entérocoques sont relativement cohérentes en 2017 et 2018 relativement aux valeurs moyennes et médianes ainsi qu'au pourcentage de jours où la moyenne géométrique a dépassé la limite recommandée de qualité des eaux utilisées à des fins récréatives de 35 NPP/100 mL. Le faible écart interquartile (c'est-à-dire la différence entre le 1^{er} et le 3^e quartile exprimée visuellement par la hauteur de la case à la **figure 2-1**) révèle que la majorité des données s'inscrivent dans une petite plage de valeurs. Compte tenu du fait que les données sont « censurées en amont » en raison d'un seuil de déclaration de 10 NPP/100 mL, la **figure 2-1** démontre que les données ont un biais positif, ce qui n'a rien d'étonnant.

Tableau 2-1. Résumé statistique des bactéries indicatrices de contamination fécale (2017-2018)

Bactéries indicatrices de contamination fécale		n	Moyenne	Écart-type	Min.	1 ^{er} quartile	Médiane	3 ^e quartile	Max.	% de dépassement ²
Entérocoques ¹ (NPP)	2017	146	13,54	14,94	10,00	10,00	10,00	11,49	144,52	5,5 %
	2018	135	17,57	33,77	10,00	10,00	10,00	12,54	331,36	4,4 %
	Toutes	281	15,48	25,80	10,00	10,00	10,00	11,49	331,36	5 %
Entérocoques ³ (NPP)	2017	726	17,90	72,92	10,00	10,00	10,00	10,00	1616,00	1,9 %
	2018	675	26,16	102,97	10,00	10,00	10,00	10,00	1850,00	4,2 %
	Toutes	1401	21,40	88,72	10,00	10,00	10,00	10,00	1850,00	3 %
<i>E. coli</i> ¹ (NPP ou comptes)	2017	146	16,00	36,59	1,00	1,42	3,20	11,44	234,33	1,4 %
	2018	135	7,09	18,01	2,00	2,00	2,64	4,70	187,41	0 %
	Toutes	281	11,72	29,47	1,00	2,00	2,64	7,52	234,33	0,7 %

Les données relatives à l'*E. coli* révèlent certaines différences entre 2017 et 2018, avec des valeurs moyennes, médianes et maximales et un écart interquartile supérieurs en 2017, comme l'indique la hauteur de la case des données de 2017 à la **figure 2-1**. Le pourcentage des valeurs dépassant la limite recommandée, même s'il est faible globalement, était également plus élevé en 2017 qu'en 2018. Même si les deux bactéries indicatrices de contamination fécale sont évaluées dans cette analyse, il est utile de garder à l'esprit que les entérocoques sont généralement considérés comme étant un indicateur préférable pour les eaux de mer ou d'estuaire puisqu'ils ont démontré une plus grande résistance à certains stress environnementaux dans les eaux utilisées à des fins récréatives, comme les conditions d'ensoleillement et de salinité.

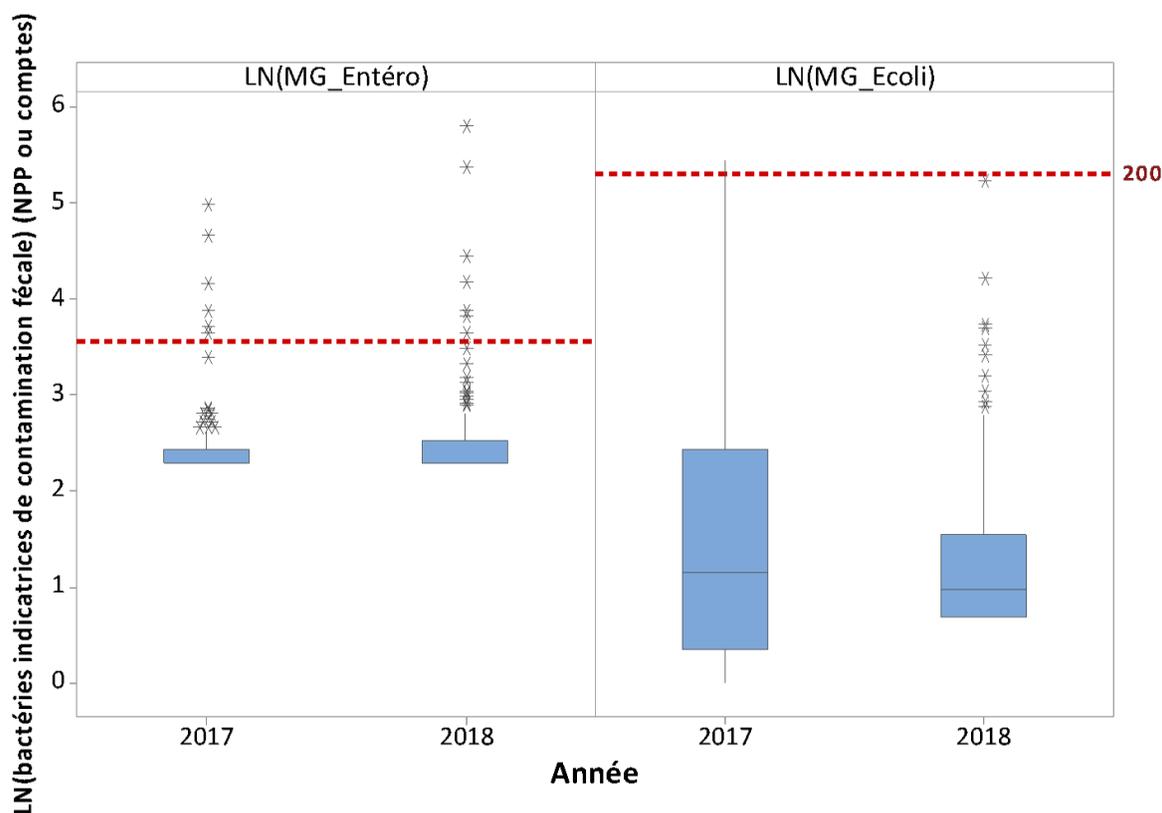


Figure 2-1. Diagrammes de quartiles de la moyenne géométrique des données logarithmiques naturelles transformées pour les entérocoques et *E. coli*

2.3 Prétraitement des données

Les données existantes sur la plage Parlee ont été prétraitées afin de les modifier selon un format facilitant l'analyse des relations entre elles et aussi de les préparer pour la création du modèle de prévision.

Comme le décrit le rapport *Qualité de l'eau de la plage Parlee : rapport final* (2018), les échantillons d'*E. coli* et d'entérocoques ont été prélevés à cinq emplacements d'échantillonnage distincts. Pour produire une variable unique de modélisation comparable pour la prévision de la moyenne géométrique (MG) des concentrations de bactéries

indicatrices de contamination fécale, la MG des cinq mesures a été calculée et enregistrée. La moyenne géométrique indiquant la valeur centrale d'un ensemble de nombres utilisant le produit de leur valeur, elle ne peut donc pas être calculée si un des nombres de l'ensemble a une valeur de 0. Compte tenu de cela, lorsqu'une station donnée dans l'ensemble des emplacements d'échantillonnage avait un compte de bactéries égal à 0, il était remplacé par la valeur 1 pour permettre le calcul de la moyenne géométrique. Les valeurs calculées de la MG ont été comparées à celles que l'on trouve à l'adresse <https://beaches.gnb.ca/en/SamplingLocation/Details/5882> et elles ont été confirmées. En outre, les concentrations d'entérocoques désignées comme étant d'une valeur inférieure (c'est-à-dire < 10) ont été consignées comme étant égales à la valeur du seuil de déclaration (c'est-à-dire < 10 a été fixé à 10 pour les entérocoques).

Les données météorologiques horaires ont été converties en moyennes quotidiennes pour en permettre la comparaison directe avec les données relatives aux bactéries indicatrices de contamination fécale en format quotidien. De plus, étant donné les observations du rapport *Qualité de l'eau de la plage Parlee : rapport final* (2018) relatives à l'influence possible de la direction du vent sur les concentrations de bactéries indicatrices de contamination fécale, la moyenne de la direction quotidienne du vent nouvellement calculée a été décalée d'un jour pour créer un nouvel ensemble de données appelé « direction du vent la veille ». Une variable décrivant le changement de la direction du vent par rapport à la veille a aussi été créée.

Les données catégoriques sont des données pouvant être décrites par deux catégories ou plus, qui comprennent souvent une certaine part de subjectivité de la part de la personne qui les recueille. Pour la plage Parlee, les observations consignées par les sauveteurs fournissent de l'information sur plusieurs variables catégoriques qui décrivent les conditions en vigueur à la plage. Les données catégoriques sur les marées, la qualité de l'eau et la qualité du sable ont reçu une valeur numérique (1, 2 ou 3) selon leur catégorie respective. La direction du vent consignée par les sauveteurs en point cardinal (N, E, S, O) a été convertie en degrés (N = 0° , S = 180° , O = 270° , etc.).

Les données quotidiennes consignées par les sauveteurs offrent certains renseignements supplémentaires sur les personnes fréquentant la plage et les baigneurs, qui sont une source possible d'excrétion de matières fécales. Pour de nombreuses journées au cours de la période de relevé, les sauveteurs ont présenté les données consignées sous forme de plage (par exemple 300 à 500 baigneurs). Dans ces cas, la limite supérieure de la plage a été utilisée dans l'analyse.

2.4 Analyse graphique

Une combinaison de graphiques chronologiques, de diagrammes de dispersion et de diagrammes de quartiles a été utilisée pour étudier les relations entre les données. L'analyse graphique est une étape utile de l'évaluation globale des données puisqu'elle peut révéler des relations non linéaires entre les variables qui ne sont pas apparentes dans l'analyse de corrélation linéaire décrite ci-dessous.

Graphiques chronologiques

Les **figures 2-2 et 2-3** présentent des graphiques chronologiques des concentrations quotidiennes d'entérocoques ou d'*E. coli* ainsi que la température moyenne quotidienne, la température de l'eau mesurée le matin par les sauveteurs et les précipitations au cours des 24 heures précédentes. Les **figures 2-4 et 2-5** présentent des graphiques chronologiques du niveau d'eau des rivières Coal Branch et Petitcodiac et des

concentrations d'entérocoques pendant les saisons de fréquentation de la plage de 2017 et 2018. Ces graphiques sont utiles pour évaluer les changements temporels de bactéries indicatrices de contamination fécale coïncidant avec les variables environnementales tracées.

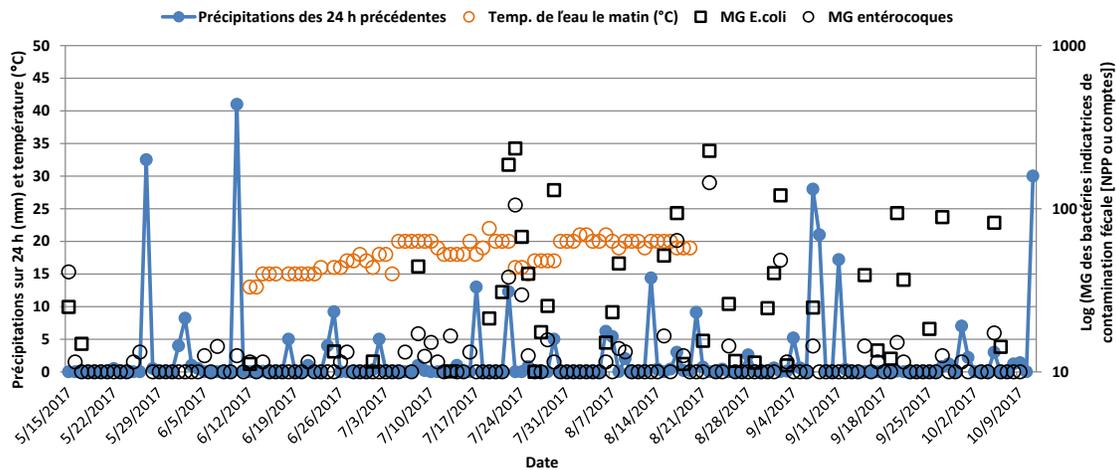


Figure 2-2. Graphique chronologique de 2017 relatif aux bactéries indicatrices de contamination fécale, à la température matinale et aux précipitations au cours des 24 heures précédentes

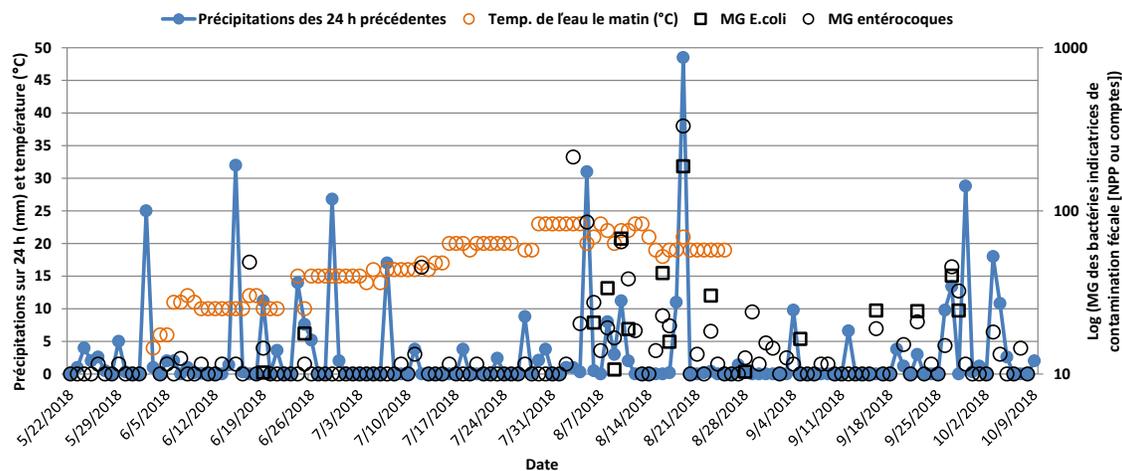


Figure 2-3. Graphique chronologique de 2018 relatif aux bactéries indicatrices de contamination fécale, à la température matinale et aux précipitations au cours des 24 heures précédentes

La lecture des graphiques chronologiques des figures 2-2 et 2-3 fait immédiatement ressortir plusieurs observations. Tout d'abord, les concentrations de bactéries indicatrices de contamination fécale les plus élevées sont habituellement observées de la fin juillet à la fin août, même si des concentrations supérieures d'*E. coli* ont été observées en octobre 2017. Cette période coïncidait avec une période d'augmentation de la température de l'eau en raison des hausses normales saisonnières de la température de l'air. Toutefois, les concentrations de bactéries indicatrices de contamination fécale ne semblent pas varier en fonction de la température. Par exemple, même lorsque les températures de l'eau

chutent après des précipitations (c'est-à-dire 22 juillet 2017, 8 août 2018), les concentrations de bactéries indicatrices de contamination fécale restent élevées.

Alors que les concentrations de bactéries indicatrices de contamination fécale semblent réagir dans une certaine mesure aux précipitations au cours des 24 heures précédentes, la relation ne semble pas entièrement cohérente entre l'importance des précipitations et celle des concentrations de bactéries indicatrices de contamination fécale, particulièrement au cours de la première partie de la saison de fréquentation de la plage. (Cette observation est abordée plus en détail à la **section 3.**) Par exemple, à la **figure 2-2** et à la **figure 2-3**, les précipitations du 27 mai 2017 et du 10 juin 2017 n'ont pas été suivies de hausses des bactéries indicatrices de contamination fécale. Les faibles précipitations de juillet et août 2017 ont coïncidé avec les concentrations de bactéries indicatrices de contamination fécale les plus élevées observées en 2017. De même, les pluies de 25 mm et de 26,8 mm du 2 juin 2018 et du 29 juin 2018 respectivement n'ont coïncidé avec aucune augmentation des concentrations de bactéries indicatrices de contamination fécale. Toutefois, les plus hautes moyennes géométriques des concentrations d'*E. coli* et d'entérocoques observées en 2018 coïncident avec les plus grandes précipitations au cours des 24 heures précédant le 19 août 2018. Ces observations laissent entendre que d'autres facteurs que la simple importance des précipitations influent sur les concentrations de bactéries indicatrices de contamination fécale.

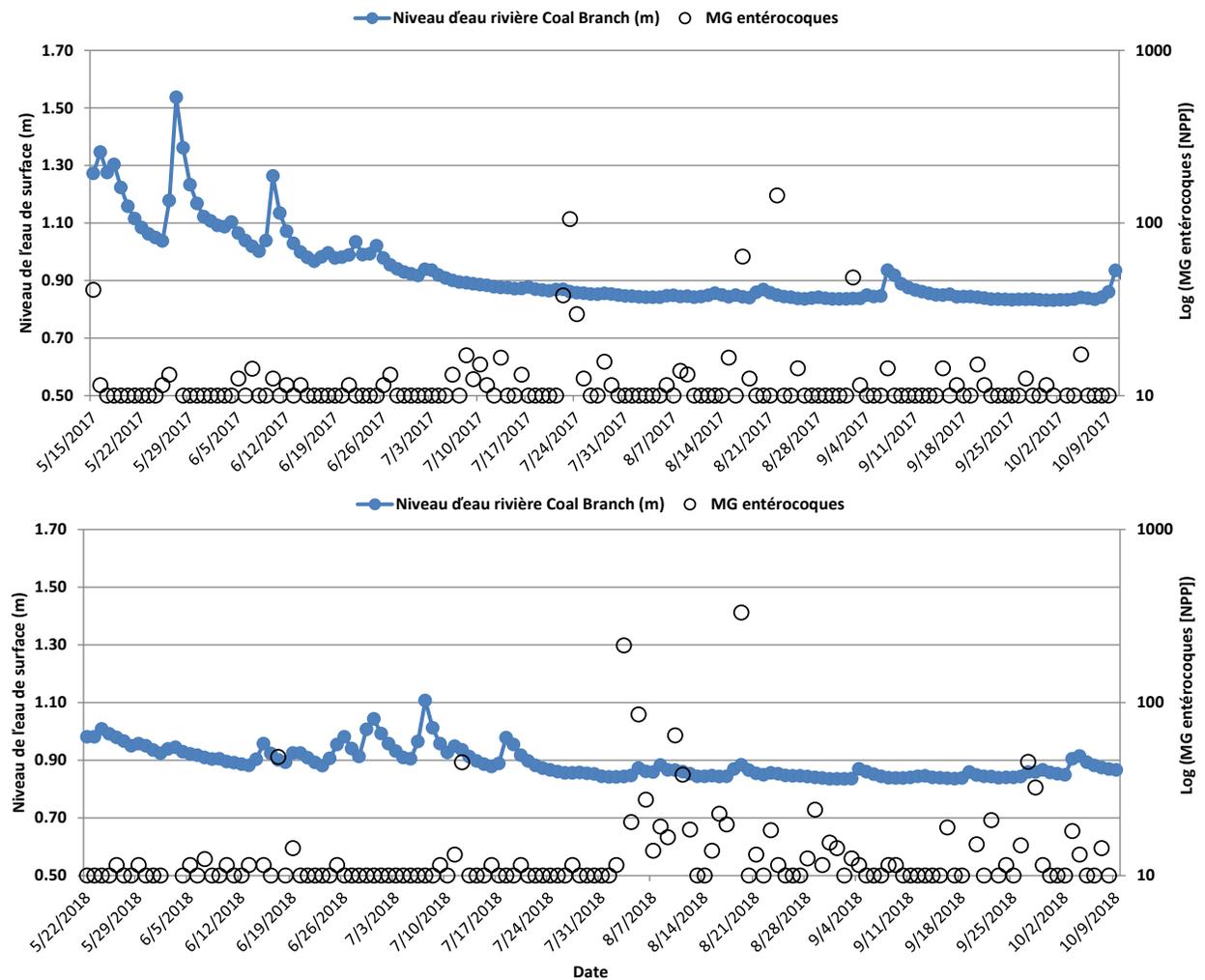


Figure 2-4. Graphique chronologique des entérocoques et du niveau d'eau de la rivière Coal Branch en 2017 (haut) et 2018 (bas)

Les entérocoques ont été utilisés aux fins de comparaison avec le niveau d'eau des rivières Coal Branch et Petitcodiac. La **figure 2-4** fait état de peu de changements coïncidents entre les concentrations d'entérocoques et le niveau d'eau de la rivière Coal Branch, les concentrations de ces bactéries s'étant élevées pendant une période de faible débit en été. La rivière Petitcodiac semble réagir davantage aux précipitations, mais, comme pour les précipitations présentées aux **figures 2-2 et 2-3**, il y a peu de coïncidence en début de saison entre la hausse des niveaux d'eau et les concentrations élevées d'entérocoques. Il convient de noter que les pluies de près de 50 mm tombées en août 2018, qui représentent près de la moitié des précipitations totales mensuelles moyennes pour ce mois, ont produit des valeurs élevées en ce qui concerne à la fois le niveau d'eau de surface et les concentrations d'entérocoques (**figure 2-5**), laissant entendre que des pluies très importantes et peu fréquentes en été pourraient coïncider avec des niveaux élevés d'entérocoques à la plage Parlee.

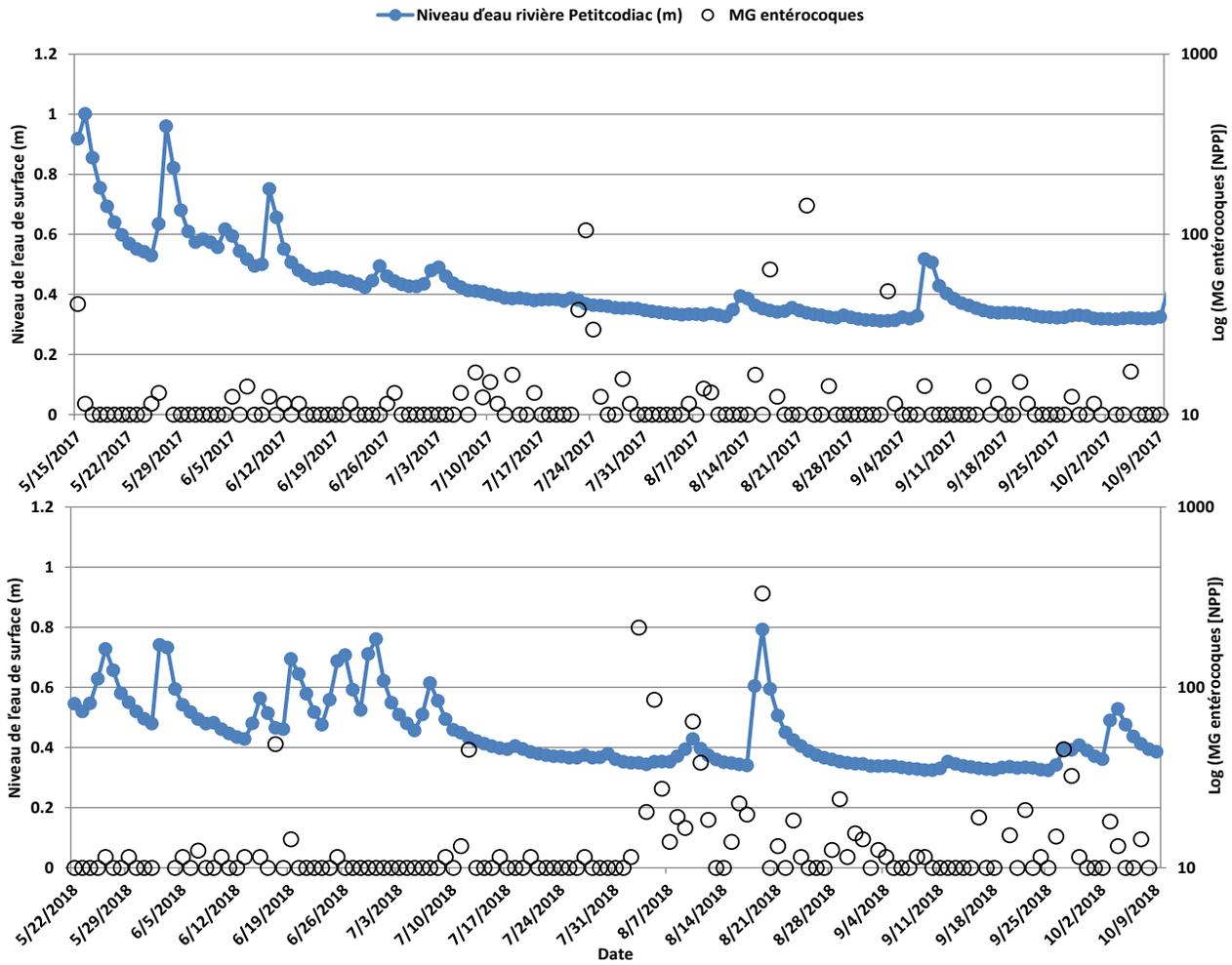


Figure 2-5. Graphique chronologique des entérocoques et du niveau d'eau de la rivière Petitcodiac en 2017 (haut) et 2018 (bas)

Diagrammes de dispersion

Les diagrammes de dispersion présentés sous forme matricielle permettent d'effectuer rapidement une évaluation visuelle de la relation possible entre des variables, ainsi que de détecter l'absence de toute relation perceptible visuellement. Une relation linéaire entre les bactéries indicatrices de contamination fécale et une variable indépendante présente le plus grand potentiel d'utilisation pour la modélisation prédictive. Dans plusieurs figures, les données ont subi des transformations logarithmiques, qui se révèlent utiles pour les données asymétriques en vue de faire ressortir davantage les schémas de données une fois qu'elles ont été tracées.

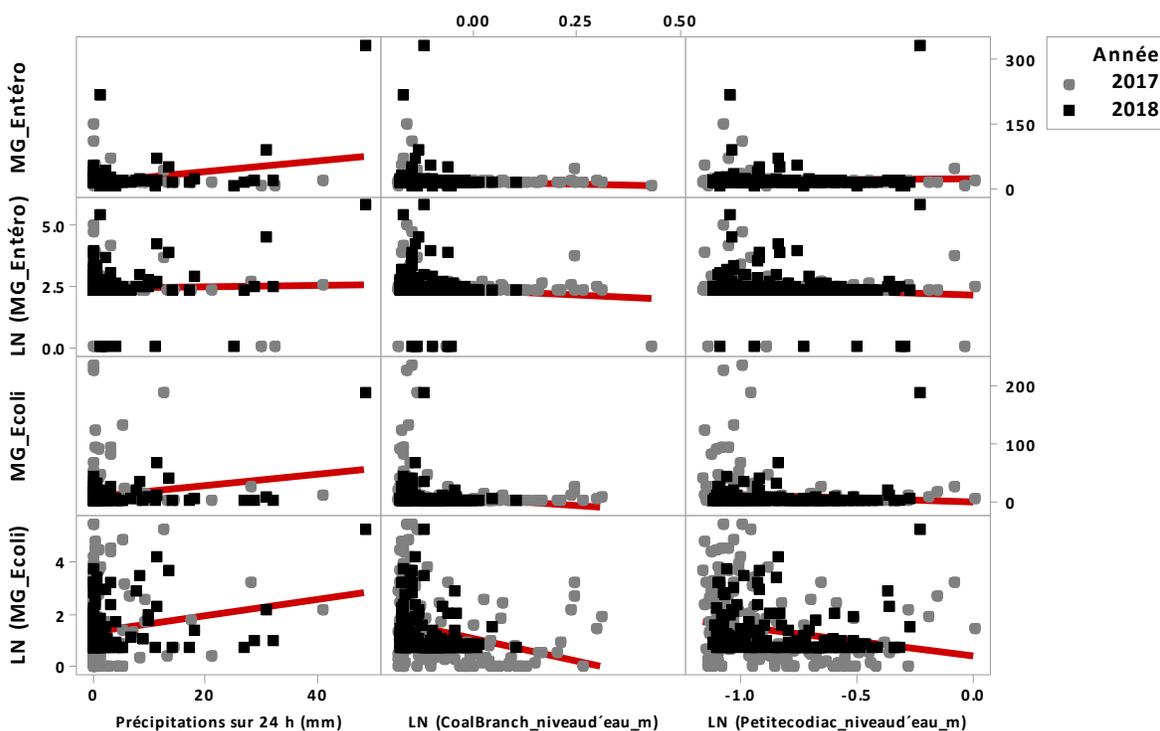


Figure 2-6. Diagrammes de dispersion matriciels avec courbe de régression – Précipitations et niveau d'eau

Les diagrammes de dispersion présentés à la **figure 2-6** montrent qu'il existe une légère relation positive avec les précipitations, les concentrations élevées de bactéries indicatrices de contamination fécale étant associées à des précipitations plus importantes au cours des 24 heures précédentes. Comme pour les graphiques chronologiques du niveau d'eau, les concentrations de bactéries indicatrices de contamination fécale montrent qu'il existe une légère relation négative avec le niveau d'eau, puisque l'augmentation du niveau d'eau est habituellement associée à la diminution des concentrations de ces bactéries, particulièrement dans le cas d'*E. coli* en 2018.

La **figure 2-7** indique que des concentrations supérieures de bactéries indicatrices de contamination fécale sont généralement observées lorsque la température de l'eau

augmente. Cette observation n'implique pas nécessairement qu'il existe une relation de cause à effet même si l'augmentation des températures de l'eau crée probablement de meilleures conditions de survie pour les bactéries indicatrices de contamination fécale et qu'elles coïncident avec des conditions d'eau plus attrayantes pour les personnes qui fréquentent la plage et les baigneurs, ce qui augmente la possibilité d'excrétion de matières fécales par ces personnes. Toutefois, les diagrammes de dispersion matriciels des concentrations de bactéries indicatrices de contamination fécale par rapport au nombre de personnes fréquentant la plage et de baigneurs (**figure 2-8**) ne font état d'aucune forte relation positive (c'est-à-dire que la hausse du nombre de baigneurs ne coïncide pas avec une augmentation des concentrations de bactéries indicatrices de contamination fécale).

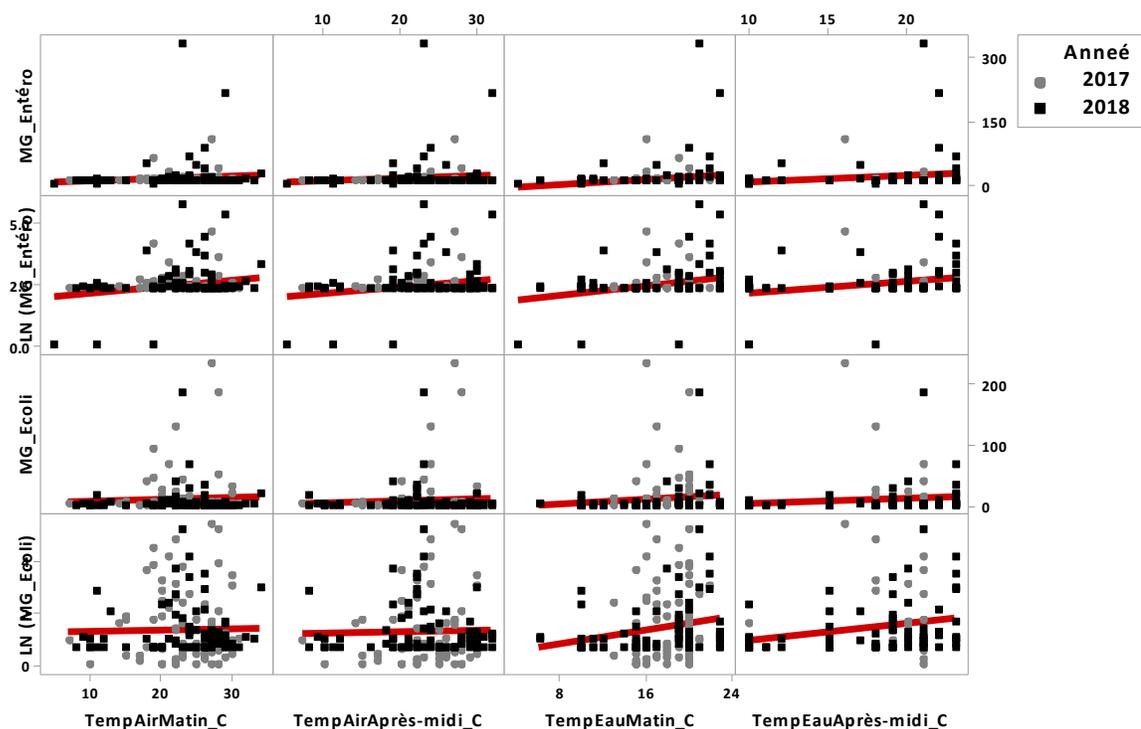


Figure 2-7. Diagrammes de dispersion matriciels avec courbe de régression – Températures de l'air et de l'eau

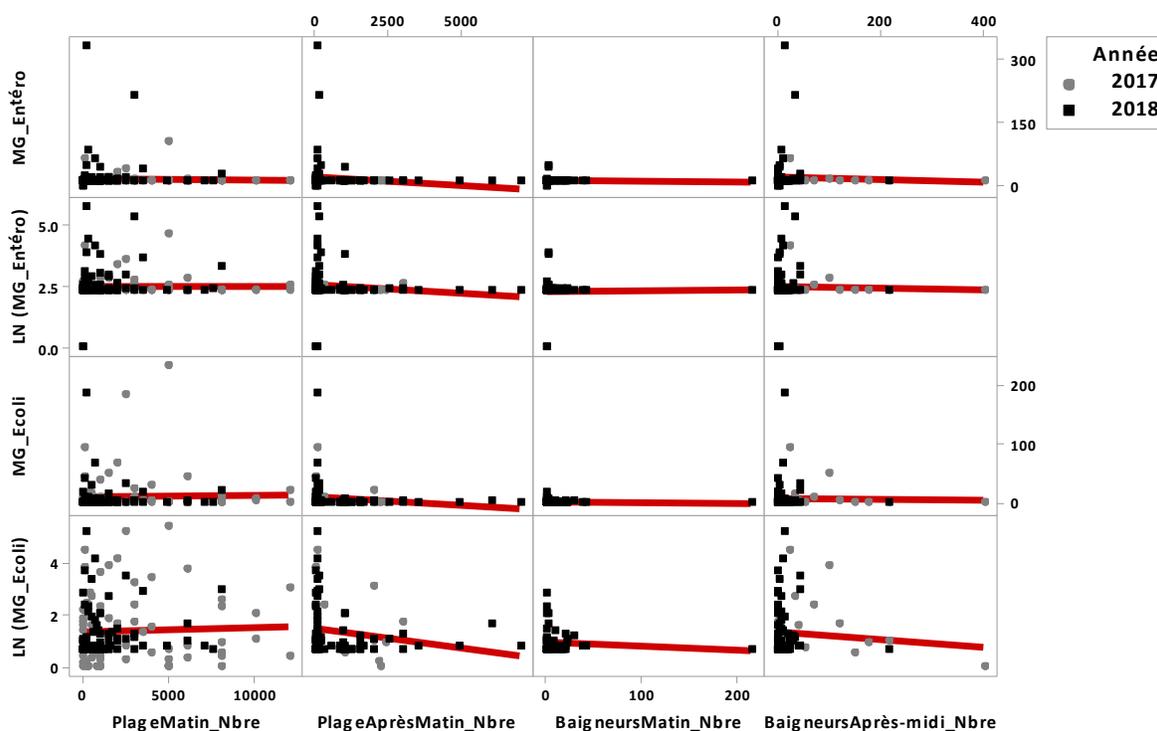


Figure 2-8. Diagrammes de dispersion matriciels avec courbe de régression – Personnes fréquentant la plage et baigneurs

Les diagrammes de dispersion peuvent aussi s'avérer utiles pour vérifier la concordance des mesures provenant de deux sources différentes. La **figure 2-9** montre la direction du vent relevée à la station météorologique de la plage Parlee par rapport aux observations de cette même variable par les sauveteurs. Il y a peu de concordance entre les deux ensembles de données (valeur R carré inférieure à 10 %). Bien que cette constatation puisse résulter de la variabilité de la direction du vent pendant la journée, elle illustre aussi la variation éventuelle d'une valeur pouvant découler des différentes sources de données et met en lumière le fait que la source utilisée pour la mise au point d'un modèle doit servir à sa mise en œuvre subséquente.

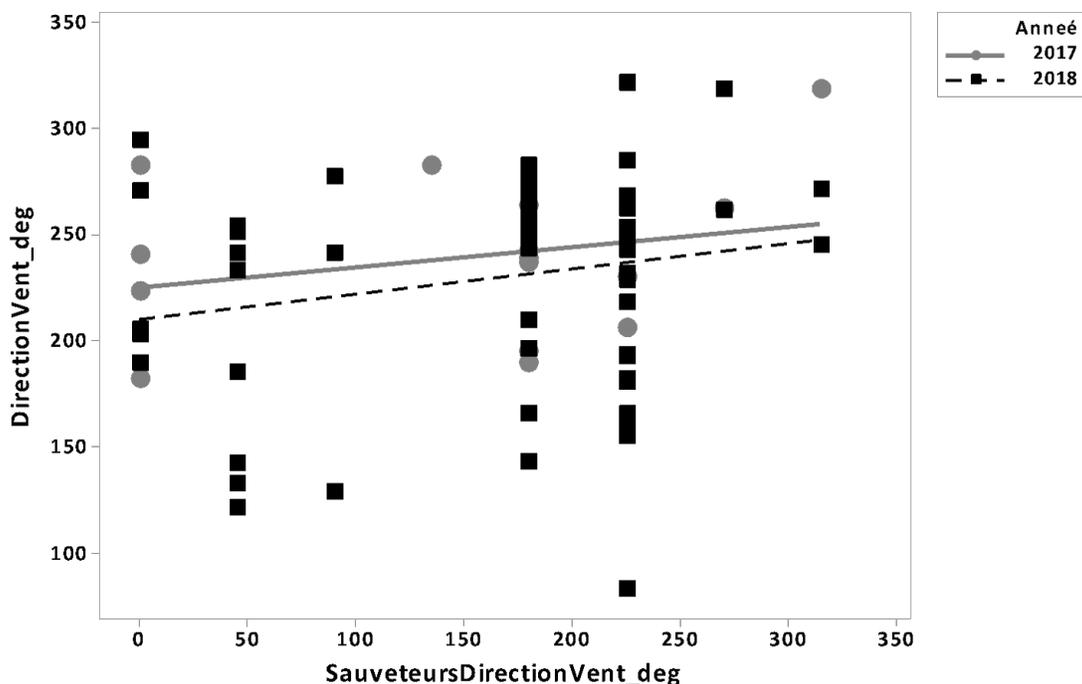


Figure 2-9. Diagramme de dispersion avec courbe de régression – Sauveteurs et station météorologique
Direction du vent en 2017 et 2018

Diagrammes de quartiles

Des diagrammes de quartiles (aussi appelés diagrammes en boîte à moustaches) ont été utilisés pour les variables catégoriques afin d'évaluer visuellement les différences entre les concentrations de bactéries indicatrices de contamination fécale pour différentes catégories de variables d'intérêt. Des courbes de référence pour les moyennes géométriques des critères de qualité de l'eau relativement aux entérocoques (35 NPP/100 mL) et *E. coli* (200 NPP/100 mL) sont aussi incluses dans chaque diagramme. Le niveau des marées (**figure 2-10**), la qualité de l'eau (**figure 2-11**), la qualité du sable (**figure 2-12**) et la direction du vent (**figure 2-14**) consignés par les sauveteurs à la plage Parlee sont présentés dans les diagrammes de quartiles ci-dessous. À l'exception des données sur les marées, qui ont été uniformément consignées en 2017 et 2018, il y a passablement de variabilité dans la manière dont l'information est signalée par les sauveteurs. Par exemple, alors que trois catégories ont été utilisées pour la qualité de l'eau en 2017, des catégories supplémentaires ont été employées en 2018. De même, les catégories utilisées varient d'une année à l'autre pour la collecte de données sur la qualité du sable.

La **figure 2-10** montre la concentration de bactéries indicatrices de contamination fécale par catégorie de marée pour les saisons de fréquentation de la plage 2017 et 2018. Alors que peu de dépassements du critère de qualité de l'eau ont été signalés au cours des deux saisons, les concentrations de bactéries indicatrices de contamination fécale affichent la plus forte variabilité à marée basse. On a aussi observé des dépassements du critère relatif aux entérocoques à marée haute, alors qu'aucun dépassement n'a été relevé et que

la plus faible variabilité a été observée pour les deux bactéries indicatrices de contamination fécale dans la catégorie « marée moyenne ». La variabilité accrue à marée basse laisse supposer qu'à marée basse, le bassin hydrologique de la plage a une incidence sur la qualité de l'eau et que le potentiel de dilution de la marée montante est inférieur. Comme pour les autres paramètres, les différences entre les concentrations d'*E. coli* en 2017 et 2018 sont notables.

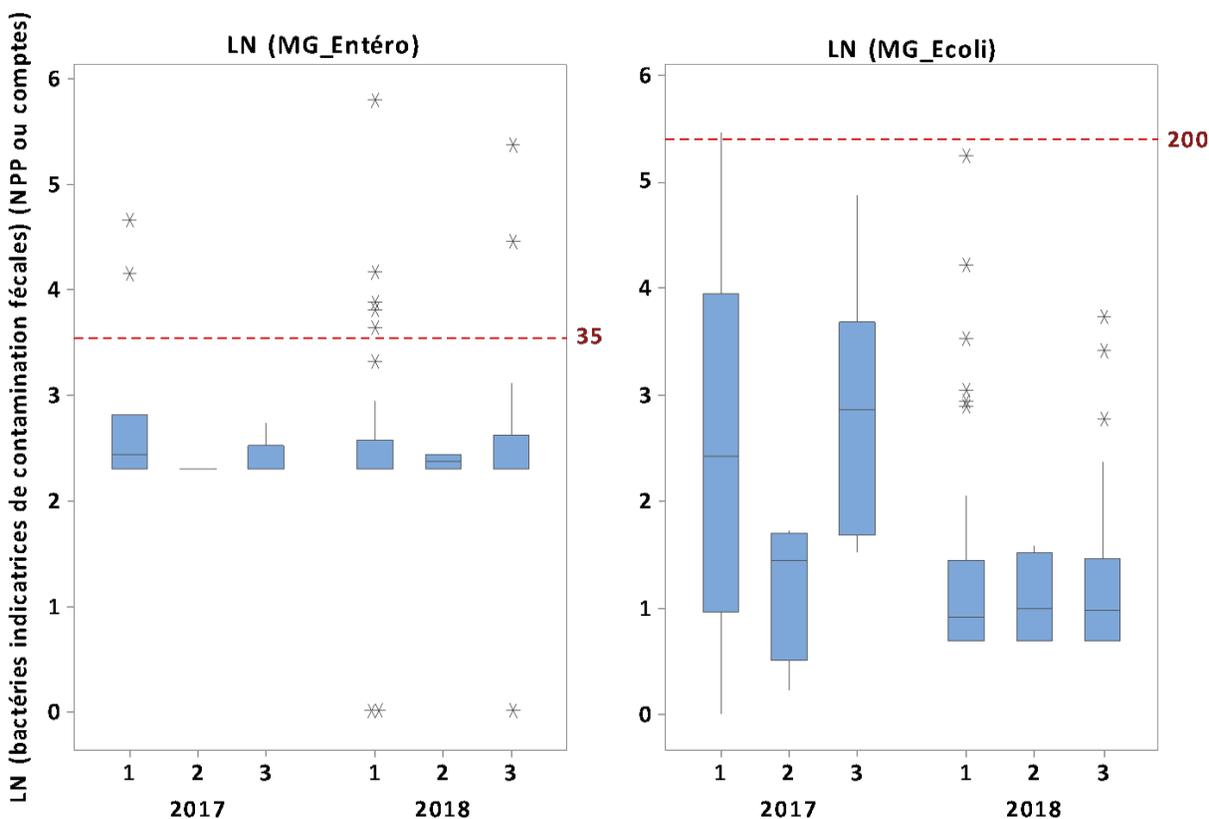


Figure 2-10. Diagramme de quartiles de la concentration de bactéries indicatrices de contamination fécale par catégorie de marée (1 = basse, 2 = moyenne, 3 = haute)

Même si le nombre de catégories varie d'une année à l'autre, lorsque les sauveteurs notaient des conditions de qualité de l'eau « faibles » ou « mauvaises », la concentration de bactéries indicatrices de contamination fécale était supérieure (**figure 2-11**). Cela démontre l'utilité de l'évaluation visuelle. Toutefois, il y a certains jours où les critères de qualité de l'eau ont été dépassés même si les sauveteurs avaient signalé de « bonnes » conditions de qualité de l'eau. Cela souligne le défi que présente l'évaluation de la qualité microbienne de l'eau puisque l'apparence visuelle ne fournit pas toujours un prédicteur fiable.

L'enregistrement par les sauveteurs de la qualité du sable présenté à la **figure 2-12** semble avoir peu de rapport avec les concentrations de bactéries indicatrices de contamination fécale, la valeur « bonne » étant attribuée presque tous les jours. L'utilité de la notation de

la qualité du sable semble limitée par rapport à la qualité de l'eau selon les données tracées à la **figure 2-12**.

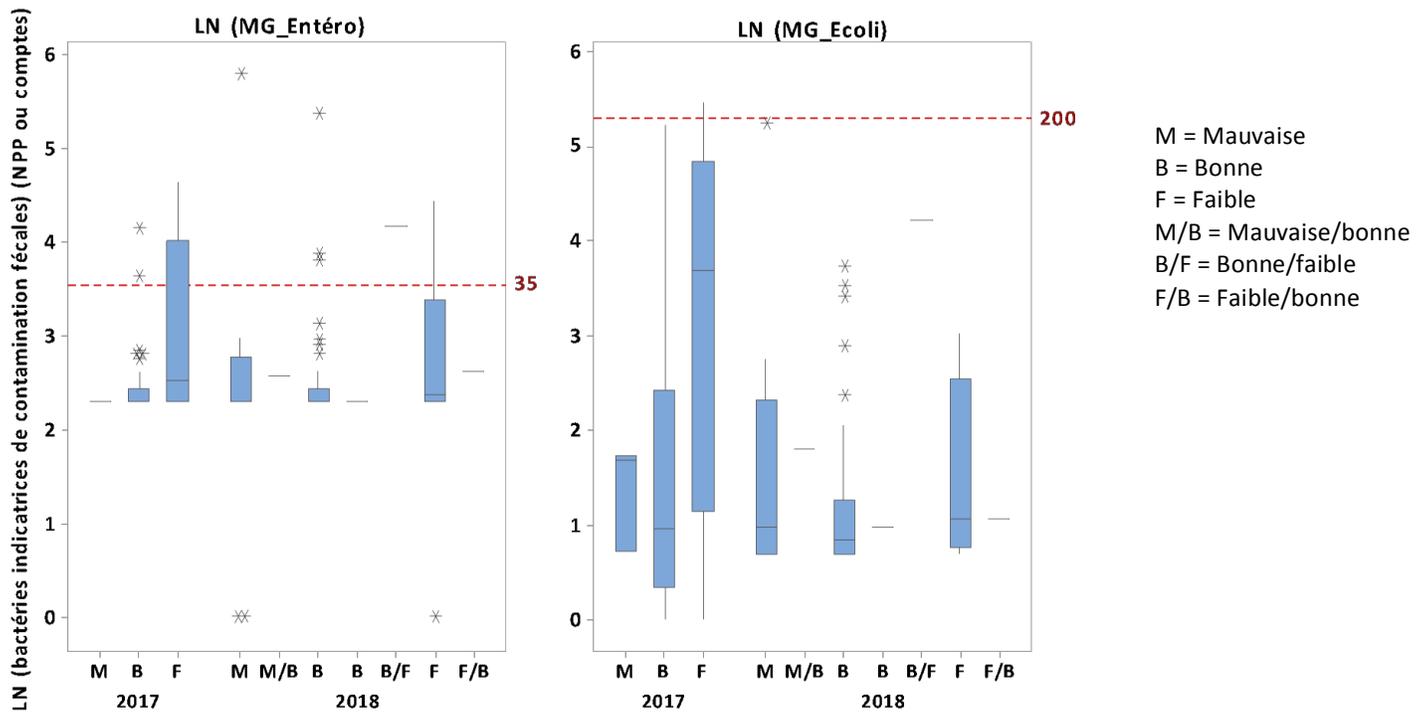


Figure 2-11. Diagramme de quartiles de la concentration de bactéries indicatrices de contamination fécale par catégorie de qualité d'eau

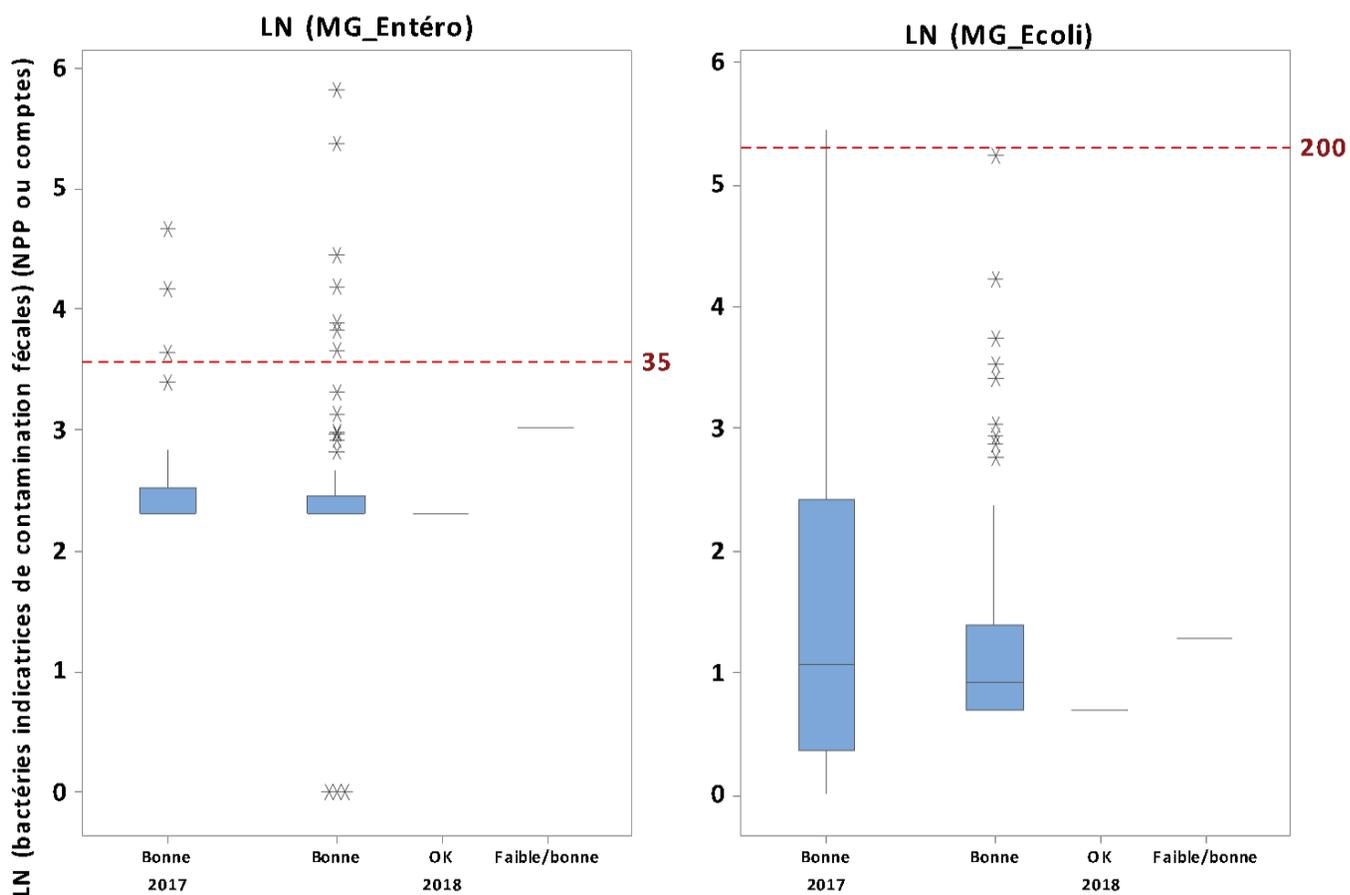


Figure 2-12. Diagramme de quartiles de la concentration de bactéries indicatrices de contamination fécale par catégorie de qualité du sable

Les observations suivantes peuvent être tirées des rapports des sauveteurs résumés dans les diagrammes de quartiles de la **figure 2-13**. La plus grande variabilité des concentrations d'entérocoques a été observée lorsque les vents soufflaient du nord et du sud-ouest, alors que les dépassements des critères de qualité des eaux utilisées à des fins récréatives coïncident avec les vents du nord, du nord-est et du sud-ouest. Les concentrations d'*E. coli* sont les plus élevées lorsque des vents du sud-ouest et du nord-nord-est sont signalés. Pour les deux bactéries indicatrices de contamination fécale, la variabilité la plus faible est observée avec des vents du sud et de l'ouest.

Ces résultats semblent contredire l'observation mentionnée dans le rapport *Qualité de l'eau de la plage Parlee : rapport final* (2018) qui énonce ce qui suit à propos de la direction du vent : « Ces données laissent entendre que, lorsque le vent souffle du sud-ouest (de 180 à 270 degrés), la qualité de l'eau de la plage Parlee est acceptable. Toutefois, lorsque le vent souffle du nord-ouest (de 270 à 360 degrés), on observe souvent une hausse des valeurs des tests. » Afin d'approfondir cette question, les données sur la direction du vent de la station météorologique de la plage Parlee ont été divisées en quatre quadrants (0-90 degrés = NE, 91-180 degrés = SE, 181-270 degrés = SE, 271-360 degrés = NO). Le **tableau 2-2** présente les comptes, les valeurs moyennes et maximales, ainsi que l'écart-type (ÉT) et l'écart interquartile (ÉI) des données pour décrire

la variabilité de chaque quadrant observée en 2017 et 2018. Les observations consignées sur la direction du vent indiquent qu'il souffle la plupart du temps du sud-ouest (vent de terre). Ces données ont également servi à générer des diagrammes de quartiles des bactéries indicatrices de contamination fécale par quadrant (**figure 2-14**). Alors que le **tableau 2-2** et la **figure 2-14** indiquent que les concentrations les plus élevées d'*E. coli* ont été observées lorsque le vent soufflait du nord-ouest, les dépassements du critère de qualité de l'eau relativement aux entérocoques ont été notés lorsque les vents provenaient de toutes les directions sauf du nord-est (qui a seulement été signalé deux fois au cours de la période de relevé). Pour mieux comprendre le pourcentage d'observations dépassant la norme, des courbes de distribution cumulative des concentrations d'entérocoques ont été tracées pour les vents soufflant du NO, du SE et du SO (**figure 2-15**). Bien que les trois quadrants affichent des dépassements, le quadrant NO présente le plus haut pourcentage de valeurs (~15 %) au-dessus du critère de qualité de l'eau, suivi de près par le quadrant SE. Les vents du sud-ouest sont les moins souvent associés aux dépassements du critère pour les entérocoques. Alors que les vents du sud-ouest sont des vents de terre, ceux du nord-ouest et du sud-ouest suivent le littoral, déplaçant l'eau parallèlement à la plage et agissant possiblement comme mécanisme de transport des bactéries vers la plage à partir d'autres emplacements situés au nord-ouest ou au sud-est.

Tableau 2-2. Caractéristiques des bactéries indicatrices de contamination fécale par quadrant du vent relevées à la station météorologique de la plage Parlee

Variable	Quadrant du vent	N	Pourcentage	Moyenne	ÉT	Maximum	ÉI
MG_Entéro	NE	2	0,69	10,00	0,00	10,00	*
	SE	26	8,93	25,30	62,90	331,40	2,2
	SO	147	50,52	13,80	19,16	214,15	1,49
	NO	32	11,00	20,79	28,99	144,52	6,24
MG_Ecoli	NE	2	0,69	6,410	0,72	6,92	*
	SE	26	8,93	13,40	36,42	187,41	6,34
	SO	143	49,14	11,80	25,07	186,07	6,52
	NO	32	11,00	27,00	58,30	234,30	15,0

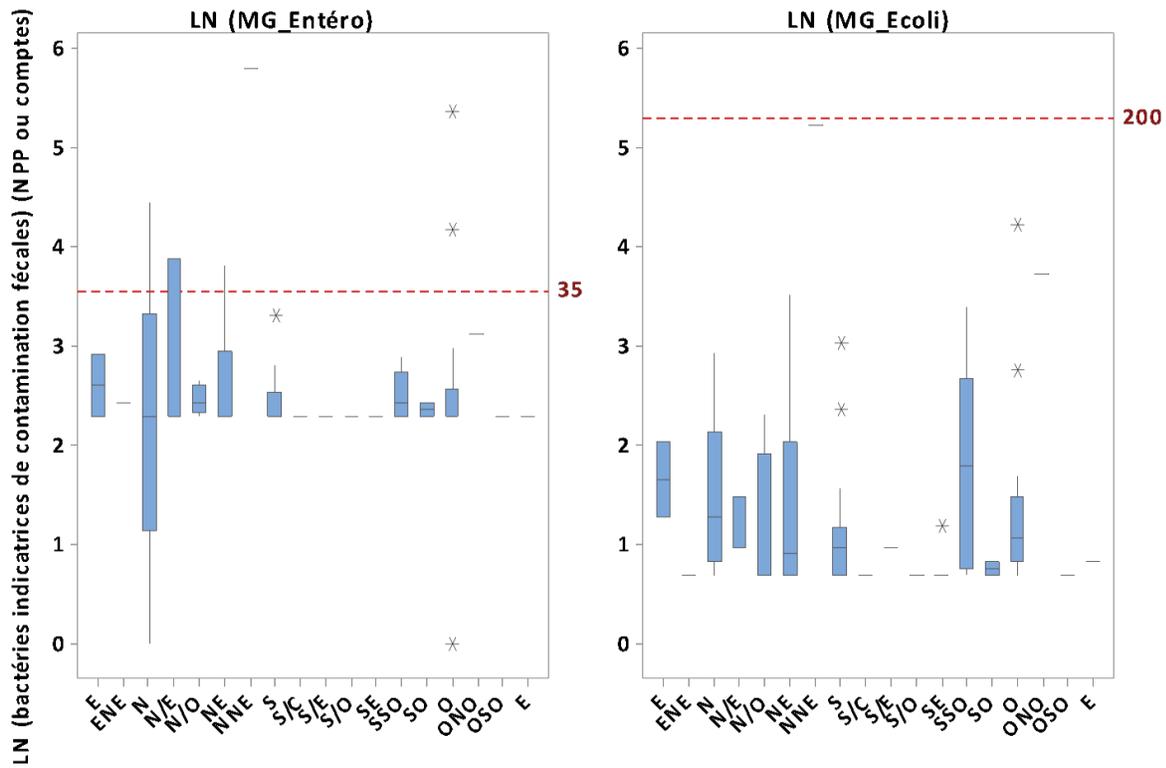


Figure 2-13. Diagramme de quartiles de la concentration de bactéries indicatrices de contamination fécale selon la direction du vent signalée par les sauveteurs (2018)

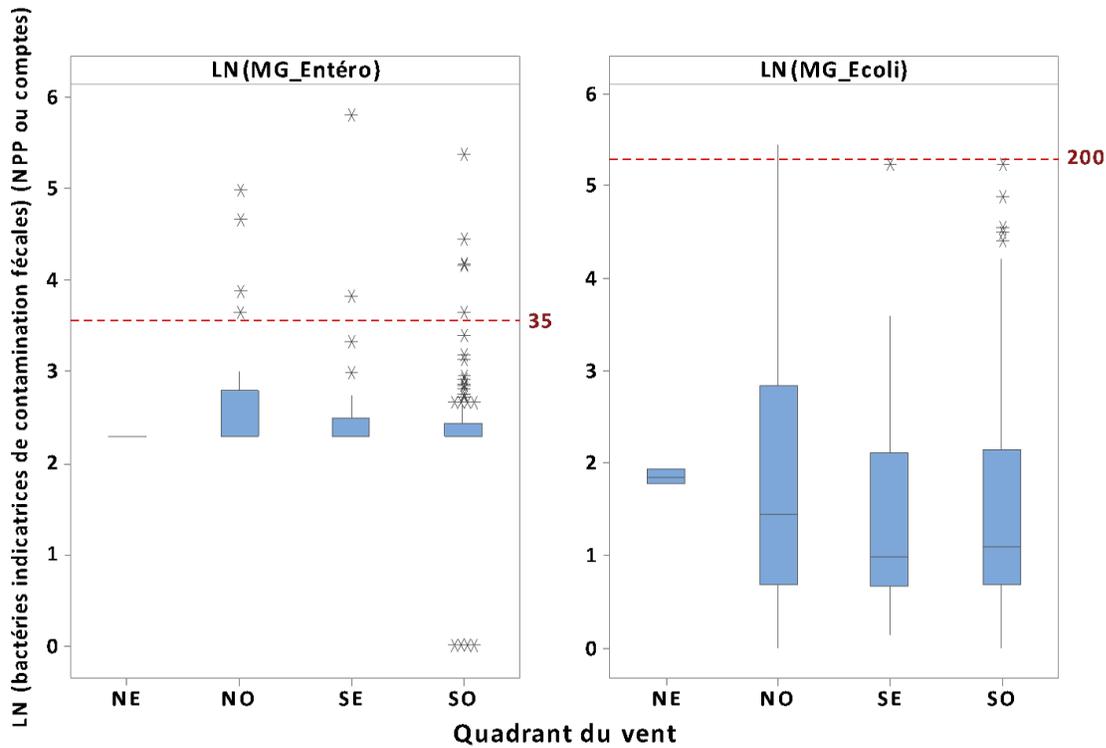


Figure 2-14. Diagramme de quartiles de la concentration de bactéries indicatrices de contamination fécale selon le quadrant de direction du vent à la station météorologique de la plage Parlee

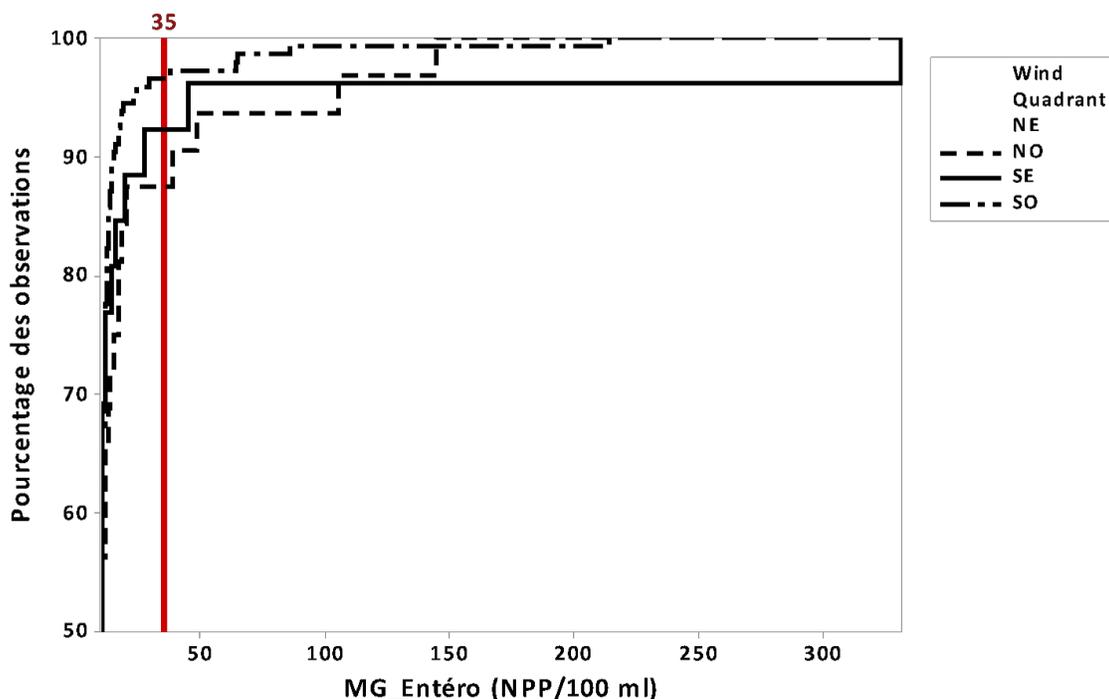


Figure 2-15. Distribution cumulative des concentrations d'entérocoques selon le quadrant de direction du vent (2017 et 2018)

2.5 Analyse statistique

L'analyse statistique a suivi les méthodes présentées dans le document intitulé *Guidance for Developing Predictive Models for Ontario Beaches* (Mas et Baker, 2011). Tout d'abord, une analyse statistique exploratoire des relations entre *E. coli* ou les entérocoques et les variables météorologiques ou environnementales indépendantes a été réalisée. À l'aide du logiciel VB3.0.6, les coefficients de corrélation linéaire de Pearson ont été calculés pour la moyenne géométrique (MG) quotidienne non transformée et logarithmique naturelle transformée pour *E. coli* et les entérocoques. Le coefficient de corrélation mesure la covariation de deux variables. Les variables météorologiques et environnementales indépendantes non catégoriques ont été analysées selon les conditions de transformation suivantes :

- non transformées,
- logarithme naturel,
- logarithme de base 10,
- carré,
- racine carrée.

Le logiciel VB3.0.6 possède un processus de sélection automatique utile pour l'analyse de corrélation. Les transformations ayant la corrélation statistiquement significative la plus forte ont été sélectionnées et consignées, sauf si la corrélation ne présentait pas une

amélioration de 20 %² par rapport à la corrélation non transformée. Dans ces cas, la corrélation non transformée a été sélectionnée et consignée. Les moyennes géométriques des bactéries indicatrices de contamination fécale ont été analysées comme un seul ensemble combiné de données (saisons de fréquentation de la plage 2017 et 2018) ainsi que pour chaque année afin d'explorer les différences sur le plan temporel ou celui de la méthodologie de mesure. Toutes les corrélations statistiquement significatives (valeur $p < 0,05$, seuil de confiance de 95 %) sont présentées dans les tableaux ci-dessous. Il convient de noter que, même si les données sur les précipitations relevées à la station météorologique de la plage Parlee ont été incluses dans l'analyse, la validité de ces données n'a pas été confirmée. Par conséquent, ces résultats sont présentés en gris pour indiquer la nature conditionnelle de la corrélation. De plus, comme cela est mentionné dans l'*Examen des données relatives à la plage Parlee*, la couverture des données et, par conséquent, le nombre de paires de données disponibles aux fins de corrélation varient selon le paramètre. Les données sur les précipitations au cours des 24 heures précédentes et le niveau d'eau possèdent la plus longue période de chevauchement. En comparaison, les données recueillies par les sauveteurs se limitent aux mois de juin à août, soit environ 60 % de la période de relevé des bactéries indicatrices de contamination fécale.

Tableau 2-3. Corrélations statistiquement significatives ($p < 0,05$) avec LN (MG entérocoques) ou MG entérocoques (Saisons de fréquentation de la plage 2017-2018)

Variable explicative	Corrélation (valeur p)
Carré (précipitations sur 24 heures au parc provincial de la plage Parlee)*	0,5260* (0,000)
Précipitations sur 24 heures au parc provincial de la plage Parlee	0,3127 (0,000)
Carré (précipitations à la station météorologique de la plage Parlee)*	0,2257* (0,001)
Intensité de la pluie à la station météorologique de la plage Parlee)*	0,1963* (0,0051)
LN (direction du vent la veille à la plage Parlee)*	-0,2009* (0,004)

*Corrélation avec la MG non transformée des entérocoques; Carré = Carré de la valeur entre parenthèses; LN = logarithme naturel de la valeur entre parenthèses.

Le **tableau 2-3** montre que les précipitations au cours des 24 heures précédentes sont la variable environnementale dont la corrélation est la plus constante et la plus forte avec les concentrations d'entérocoques pour l'ensemble de la période de relevé (2017 et 2018). La corrélation est plus forte ($r = 0,5260$) pour les entérocoques que pour *E. coli* ($r = 0,2377$). La corrélation positive est cohérente avec le concept voulant que les précipitations influent sur les concentrations de bactéries indicatrices de contamination fécale en transportant des bactéries vers la place pendant les événements produisant des écoulements. Toutefois, il importe de noter que la corrélation décrit le degré auquel les deux variables varient ensemble (c'est-à-dire

² Ce changement de 20 % de la force de corrélation est recommandé dans le logiciel VB3.0.6 comme étant le seuil à utiliser pour la sélection des variables indépendantes transformées par rapport aux variables indépendantes non transformées pour la mise au point du modèle.

qu'une variable augmente ou diminue alors que l'autre augmente). La corrélation entre deux variables ne prouve pas qu'il existe une relation de cause à effet entre ces deux variables (Helsel et Hirsch, 2002). La régression linéaire multiple (abordée à la **section 3**) cherche à déterminer et à quantifier la mesure dans laquelle une variable peut prédire la variation de l'autre variable.

Tableau 2-4. Corrélations statistiquement significatives ($p < 0,05$) avec LN (MG *E. coli*) ou MG *E. coli* (Saisons de fréquentation de la plage 2017-2018)

Variable explicative	Corrélation (valeur p)
Carré (précipitations sur 24 heures au parc provincial de la plage Parlee)*	0,2377* (0,0001)
Précipitations sur 24 heures au parc provincial de la plage Parlee	0,1659 (0,006)
Durée de la pluie à la station météorologique de la plage Parlee	-0,1438 (0,041)
Niveau d'eau de la rivière Coal Branch*	-0,1484* (0,013)
Niveau d'eau de la rivière Coal Branch	-0,2648 (0,000)
Niveau d'eau de la rivière Petitcodiac	-0,1947 (0,001)

*Corrélation avec la MG non transformée d'*E. coli*; Carré = carré de la valeur entre parenthèses; LN = logarithme naturel de la valeur entre parenthèses.

Il convient de noter qu'aucune corrélation statistiquement significative n'a été observée en 2017 entre les précipitations sur 24 heures et les concentrations d'*E. coli* (**tableau 2-5**). En plus, il existe d'autres différences notables entre les corrélations relatives à *E. coli* en 2017 et 2018. Alors que les entérocoques n'ont aucune corrélation importante avec l'une ou l'autre des variables observées par les sauveteurs au cours de la période de relevé de 2017 et 2018, les températures de l'eau en après-midi et de l'air le matin, le nombre de personnes à la plage l'après-midi, ainsi que les directions du vent la veille relevées par les sauveteurs ont toutes de fortes corrélations négatives importantes avec *E. coli* en 2018, mais aucune corrélation importante en 2017 (**tableau 2-6**). Il s'agit d'un résultat inattendu puisqu'il laisse entendre que la hausse des températures et l'augmentation du nombre de personnes à la plage sont corrélées avec des concentrations plus faibles de bactéries indicatrices de contamination fécale, alors qu'il s'agit de conditions qui vont à l'encontre des idées conceptuelles d'une augmentation possible de l'excrétion de matières fécales par les personnes fréquentant la plage et de la survie des bactéries indicatrices de contamination fécale dans des conditions de températures plus chaudes. Pour approfondir la question, les observations effectuées la veille par les sauveteurs ont aussi été vérifiées afin de déterminer si les conditions de la plage relevées la veille influent sur les concentrations de bactéries indicatrices de contamination fécale, mais aucune corrélation importante n'a été observée pour l'une ou l'autre des bactéries indicatrices de contamination fécale en 2017 ou 2018.

Cette variation de la relation entre 2017 et 2018 pour *E. coli* et les corrélations négatives inattendues pour *E. coli* en 2018 pourrait découler des différences également

notées dans les statistiques sommaires relatives à *E. coli*. L'analyse des données a révélé certaines différentes statistiques entre les concentrations des deux bactéries indicatrices de contamination fécale (voir le **tableau 2-1**) qui pourraient être liées à la méthode de laboratoire utilisée pour mesurer les concentrations d'*E. coli*. De toute façon, le faible pourcentage de dépassements qu'indiquent les données relatives à *E. coli* laisse supposer que son utilisation est limitée comme indicateur comparativement aux entérocoques.

Les niveaux d'eau des rivières Coal Branch et Petitcodiac ont des corrélations négatives moyennes avec les concentrations d'*E. coli* (tant avec l'ensemble combiné de données que pour chaque année) et aucune corrélation importante avec les entérocoques. La rivière Coal Branch, située au nord de la plage Parlee (**figure 2-16**), a une corrélation négative plus forte avec les concentrations d'*E. coli* que la rivière Petitcodiac. Les corrélations négatives avec le niveau d'eau indiquent que, pendant les périodes de fort débit, les concentrations d'*E. coli* sont plus faibles. Ce résultat est conforme aux graphiques chronologiques (**figures 2-4 et 2-5**). Puisque la saison de fréquentation de la plage coïncide avec l'été, qui est traditionnellement une saison plus sèche ayant un débit plus faible, cette corrélation pourrait ne pas être de nature causale, mais plutôt illustrer la covariation saisonnière qui existe.



Figure 2-16. Emplacement des jauges de niveau d'eau des rivières Coal Branch et Petitcodiac

Tableau 2-5. Corrélations statistiquement significatives ($p < 0,05$) avec LN (MG *E. coli*) ou MG *E. coli* (saison de fréquentation de la plage 2017)

Variable explicative	Corrélation (valeur p)
Durée de la pluie à la station météorologique de la plage Parlee*	-0,2011* (0,028)
Direction du vent à la station météorologique de la plage Parlee*	0,2479* (0,006)
Pluie à la station météorologique de la plage Parlee	-0,1913 (0,037)
Durée de la pluie à la station météorologique de la plage Parlee	-0,2436 (0,007)
Intensité de la pluie à la station météorologique de la plage Parlee	-0,1849 (0,044)
Niveau d'eau de la rivière Coal Branch*	-0,1981* (0,017)
Niveau d'eau de la rivière Petitcodiac*	-0,1779* (0,032)
Niveau d'eau de la rivière Coal Branch	-0,2997 (0,0002)
Niveau d'eau de la rivière Petitcodiac	-0,2461 (0,003)

*Corrélation avec la MG non transformée d'*E. coli*; Carré = carré de la valeur entre parenthèses; LN = logarithme naturel de la valeur entre parenthèses.

Tableau 2-6. Corrélations statistiquement significatives ($p < 0,05$) avec LN (MG *E. coli*) ou MG *E. coli* (saison de fréquentation de la plage 2018)

Variable explicative	Corrélation (valeur p)
Carré (précipitations au parc provincial de la plage Parlee)*	0,6949* (0,000)
Précipitations sur 24 heures au parc provincial de la plage Parlee)	0,3184 (0,0002)
Pluie à la station météorologique de la plage Parlee*	0,6407* (0,000)
Durée de la pluie à la station météorologique de la plage Parlee*	0,2384* (0,030)
Intensité de la pluie à la station météorologique de la plage Parlee*	0,7328* (0,000)
Pluie à la station météorologique de la plage Parlee	0,2449 (0,026)
Température de l'air notée le matin par les sauveteurs*	-0,5135* (0,003)
Température de l'air notée l'après-midi par les sauveteurs*	-0,3973* (0,0269)
LN (personnes à la plage notées le matin par les sauveteurs)*	-0,3574* (0,048)
Température de l'air notée le matin par les sauveteurs	-0,4649 (0,008)
Intensité de la pluie à la station météorologique de la plage Parlee	0,3248 (0,003)
Niveau d'eau de la rivière Petitcodiac*	0,2267* (0,008)
Niveau d'eau de la rivière Coal Branch	-0,2702 (0,002)

Variable explicative	Corrélation (valeur p)
LN (direction du vent la veille à la plage Parlee)*	-0,4136* (0,0001)
LN (direction du vent la veille à la plage Parlee)	-0,2421 (0,028)

*Corrélation avec la MG non transformée d'*E. coli*; Carré = carré de la valeur entre parenthèses; LN = logarithme naturel de la valeur entre parenthèses.

2.6 Conclusions et recommandations

L'analyse exploratoire des données a permis de cerner les principaux résultats suivants sur les caractéristiques des données relatives aux bactéries indicatrices de contamination fécale recueillies pendant les saisons de fréquentation de la plage 2017 et 2018 et les relations entre *E. coli* et les entérocoques et les variables environnementales indépendantes recueillies en même temps en 2017 et 2018 :

- Le pourcentage de concentrations de bactéries indicatrices de contamination fécale dépassant la moyenne géométrique de la limite recommandée de qualité des eaux utilisées à des fins récréatives (35 NPP/100 mL) était faible tant pour les entérocoques (~5 %) que pour *E. coli* (< 1 %). De même, le pourcentage de données relatives aux entérocoques dépassant la valeur maximale recommandée pour un échantillon unique de 70 NPP/100 mL était bas (~2-4 %). Cela indique de manière générale que la qualité de l'eau mesurée en fonction de la présence de bactéries indicatrices de contamination fécale est bonne et laisse aussi supposer que la mise au point de modèles de prévision robustes et statistiquement significatifs pourrait s'avérer improbable puisqu'il y a peu d'exemples de valeurs élevées de bactéries indicatrices de contamination fécale pour réussir à cartographier la relation empirique entre les concentrations de ces bactéries et les autres variables.
- Les valeurs de certaines variables proviennent de plusieurs sources (c'est-à-dire les données de station météorologique et celles recueillies par les sauveteurs), mais ces valeurs ne concordent pas nécessairement. Une comparaison des données sur la direction du vent provenant des observations des sauveteurs et de la station météorologique de la plage Parlee montre qu'il y a peu de concordance entre les deux ensembles de données (valeur R carré inférieure à 10 %). Bien que cette constatation puisse résulter de la variabilité de la direction du vent pendant la journée, elle illustre aussi la variation éventuelle d'une valeur pouvant découler des différentes sources de données et met en lumière le fait que la source utilisée pour la mise au point d'un modèle doit servir à sa mise en œuvre subséquente.
- En ce qui a trait à plusieurs des variables catégoriques, il y a pas mal de variabilité sur le plan la manière dont l'information est consignée par les sauveteurs. La normalisation des catégories et de leur signification est essentielle si les variables doivent servir à la comparaison au fil des saisons ou à la création et à la mise en œuvre d'un modèle.
- Les résultats de cette analyse ont renforcé la conclusion antérieure selon laquelle les vents du nord-ouest sont souvent associés à des concentrations élevées de bactéries indicatrices de contamination fécale. Toutefois, il importe de noter que tous les quadrants sauf celui du nord-est affichent des dépassements, même si le quadrant du nord-ouest a le pourcentage le plus élevé de valeurs (~15 %) qui dépassent le critère de qualité de l'eau, suivi de près par le sud-est. Les vents du sud-ouest sont les moins souvent associés aux dépassements du critère pour les

- entérocoques. Alors que les vents du sud-ouest sont des vents de terre, ceux du nord-ouest et du sud-ouest suivent le littoral, déplaçant l'eau parallèlement à la plage et agissant possiblement comme mécanisme de transport des bactéries vers la plage à partir d'autres emplacements situés au nord-ouest ou au sud-est.
- L'analyse de corrélation démontre que les variables suivantes conviennent probablement à la mise au point d'un modèle de prévision :
 - précipitations au cours des 24 heures précédentes;
 - direction du vent la veille;
 - direction du vent;
 - niveau d'eau à la jauge de la rivière Coal Branch.
 - Voici d'autres variables pouvant être utiles, mais qui, selon les observations, n'étaient liées qu'aux données relatives à *E. coli* de 2018 :
 - température de l'air;
 - température de l'eau;
 - personnes à la plage en après-midi.

En se fondant sur l'analyse exploratoire des données, les recommandations suivantes sont également formulées :

- Même si les deux bactéries indicatrices de contamination fécale ont été évaluées dans la présente analyse, les entérocoques devraient être utilisés pour la surveillance à venir puisqu'ils sont généralement considérés comme étant un indicateur préférable pour les eaux de mer et d'estuaire. Les entérocoques ont démontré une plus grande résistance à certains stress environnementaux dans les eaux utilisées à des fins récréatives, comme les conditions d'ensoleillement et de salinité.
- Les sauveteurs devraient recevoir des directives détaillées et précises sur l'enregistrement des données catégoriques afin que les données consignées soient cohérentes.
- Il ne sera probablement pas possible de remplacer les données provenant d'une source par celles d'une autre source. Par exemple, il est peu probable qu'un modèle de prévision mis au point en utilisant la direction du vent relevée à la station météorologique pourra utiliser avec succès les observations des sauveteurs sur cette variable; de même, une analyse statistique sur plusieurs saisons ne pourrait employer les données provenant de sources différentes.
- Les observations des sauveteurs sur la qualité du sable, qui semblent avoir peu de rapport avec la qualité de l'eau, pourraient être éliminées sans que cela ait d'incidence sur les prévisions en matière de qualité de l'eau.
- La turbidité devrait être ajoutée aux mesures effectuées sur le terrain dans le cadre du programme de collecte de données et de surveillance pour la plage Parlee en raison de ses antécédents à titre d'indicateur utile de la qualité de l'eau, mais aussi de variable potentiellement importante dans les modèles de prévision (Francy et Darnier, 2006). Il faudrait alors acheter ou louer un turbidimètre portatif qui répond aux critères de rendement précisés par la méthode de la norme ISO 7027 (DIN EN 27027) et qui est conforme à la norme ASTM D6855-03. Une plage de mesure allant de 0 à 1 000 uTN serait suffisante aux fins de la surveillance de la plage. Anderson (2005) recommande d'utiliser le même type d'instrument pour les mesures de la turbidité, car ils pourraient ne pas produire de résultats équivalents

si leur conception différait. Les turbidimètres convenables, qui coûtent habituellement 1 300 \$³ environ, sont étalonnés en usine, mais ils exigent des étalonnages mensuels avec des solutions de calibrage. Les solutions de calibrage doivent généralement être remplacées environ une fois par année et coûtent près de 70 \$. Il est également recommandé d'avoir des flacons d'échantillonnage supplémentaires pour les analyses. Ils coûtent environ 70 \$ par paquet de trois. Les turbidimètres sont relativement simples à utiliser et les lectures sont obtenues rapidement; ils exigeraient une formation minimale pour un échantillonneur de terrain. Les échantillons de turbidité pourraient être prélevés en même temps que les échantillons quotidiens de qualité de l'eau et devraient prendre un maximum de 10 à 15 minutes de plus par station pour le prélèvement des échantillons, l'analyse et la saisie des données. De nombreux turbidimètres offrent aussi l'enregistrement interne des données pour permettre le téléchargement des mesures après l'échantillonnage et éviter les erreurs possibles de transcription. Francy et Darner (2006) ont recommandé que la turbidité soit mesurée sur place à l'aide d'un turbidimètre de terrain ou *in situ* à l'aide d'un appareil de mesure de la qualité de l'eau. Si la turbidité est mesurée dans le cadre d'analyses de laboratoire, les échantillons devraient être conservés sur de la glace pendant le transport jusqu'au laboratoire. Bien que la turbidité puisse être mesurée en laboratoire, la mesure sur place est recommandée si les résultats doivent être utilisés dans le cadre d'un programme de modélisation prédictive ou si la turbidité doit servir d'indicateur « en temps réel » de la qualité de l'eau parce que la mesure sur place fournit des résultats immédiats comparativement au délai associé à l'analyse en laboratoire.

³ Dollars canadiens.

Section 3 – Mise au point et évaluation d'un modèle préliminaire

3.1 Introduction

L'un des principaux objectifs de la présente étude était de déterminer la faisabilité et les limites de l'utilisation de modèles empiriques ou fondés sur les données de prévision de la qualité de l'eau de la plage Parlee. Les modèles de prévision peuvent fournir un outil utile pour la prévision des conditions futures ou immédiates de qualité de l'eau de la plage. La régression linéaire multiple (RLM) est un outil très répandu pour la mise au point de modèles. Même si la régression linéaire multiple nécessite l'existence de relations linéaires entre une variable étudiée ou dépendante transformée ou non et des variables indépendantes ou explicatives, il s'agit d'une démarche de modélisation très transparente et facilement reproductible. L'analyse exploratoire préalable des données a révélé la présence d'une corrélation linéaire entre une bactérie indicatrice de contamination fécale et des variables environnementales (c'est-à-dire précipitations, niveau de la rivière, etc.), ce qui appuie la faisabilité d'une régression linéaire multiple. L'ensemble de données relatives à la plage Parlee comporte une limite, soit le nombre relativement faible de valeurs quotidiennes de bactéries indicatrices de contamination fécale qui dépassent les limites recommandées de qualité de l'eau des plages utilisées à des fins récréatives. Habituellement, pour que les modèles réussissent à prévoir les dépassements d'un seuil, on recommande qu'au moins 20 % des données disponibles pour la mise au point du modèle comprennent des valeurs dépassant le seuil. Comme l'ensemble combiné de données relatives à la plage Parlee pour 2017 et 2018 contient moins de 6 % de valeurs dépassant la limite recommandée pour les entérocoques et moins de 2 % dépassant la valeur recommandée pour *E. coli*, la mise au point de modèles qui réussissent à prévoir des valeurs supérieures aux limites recommandées de qualité des eaux utilisées à des fins récréatives d'intérêt devrait poser un défi.

À l'aide de la version 3.0.6 (VB3.0.6) du logiciel Virtual Beach, cette tâche visait à commencer par l'ensemble de variables éventuelles cernées à la **section 2** pour générer des modèles fondés sur la régression qui respectent les hypothèses sous-jacentes de régression linéaire présentées dans le document intitulé *Guidance for Developing Predictive Models for Ontario Beaches* (c'est-à-dire que les coefficients sont significatifs, les résidus semblent homoscedastiques, indépendants et normalement distribués, etc.). En plus d'étudier plusieurs variables explicatives, l'utilisation des seules données sur les précipitations a aussi été évaluée et l'utilité de la valeur d'interdiction préventive de 10 mm a été estimée.

3.2 Démarche de modélisation

En plus des variables indépendantes (explicatives) abordées à la **section 2 – Analyse exploratoire des données**, des variables indépendantes supplémentaires ont été créées et ajoutées aux ensembles de données servant à la mise au point du modèle, soit les précipitations au cours des 48 et 72 heures précédentes, la valeur numérique du jour et du mois de l'année, la valeur absolue du changement de direction du vent, le nombre de jours depuis les dernières précipitations et le nombre de jours depuis les dernières précipitations multiplié par la valeur absolue du changement de direction du vent.

Comme cela a été mentionné aux **sections 1 et 2**, les valeurs indépendantes utilisées pour la mise au point du modèle couvrent des plages variables de données. Afin de créer un modèle, le logiciel VB3.0.6 doit avoir un ensemble complet de données (c'est-à-dire un ensemble de données auquel il ne manque aucune valeur). Au cours du processus de validation des données du logiciel VB3.0.6, si pour une journée au sein de l'ensemble de données il manque des données ne serait-ce que pour une seule variable, la journée

complète sera automatiquement supprimée pour toutes les variables. Afin de tenir compte de cette suppression possible de données, avant d'utiliser le logiciel VB3.0.6, les valeurs indépendantes ont été classées en fonction de la longueur de la couverture de données et regroupées en ensembles de manière itérative par ordre décroissant de couverture des données. Par exemple, « l'ensemble de données A » correspondrait aux bactéries indicatrices de contamination fécale et à la valeur indépendante ayant la plus longue couverture de données. « L'ensemble de données B » serait « l'ensemble de données A » plus la variable indépendante ayant la deuxième plus longue couverture, etc. Cela a fait en sorte que le logiciel VB3.0.6 reçoive le plus grand nombre d'observations possible pour toute combinaison donnée de variables indépendantes.

En outre, on a créé un autre ensemble de données VB dont toutes les données ont été séparées en deux groupements temporels (de mai à juillet et d'août à septembre) avant d'entreprendre le processus itératif décrit ci-dessus afin d'étudier les différences de comportement survenant entre le début et la fin de la saison de fréquentation de la plage. Ces ensembles de données ont fait l'objet de l'analyse statistique décrite à la **section 2** afin d'étudier les relations linéaires entre les bactéries indicatrices de contamination fécale transformées ou non transformées et les variables indépendantes transformées ou non transformées.

Le logiciel VB3.0.6 compte plusieurs méthodes de modélisation statistique, dont la régression linéaire multiple, la régression partielle par les moindres carrés et la régression amplifiée (*boosted*) généralisée. Comme le mentionne le document intitulé *Guidance for Developing Predictive Models for Ontario Beaches* (Mas et Baker, 2011), les modèles de régression linéaire multiple sont une technique courante de modélisation de la qualité de l'eau utilisée pour diverses plages. La RLM vise à expliquer autant que possible la variation de la variable dépendante ou étudiée (les bactéries indicatrices de contamination fécale en l'occurrence) compte tenu d'un ensemble de variables explicatives ou indépendantes. La composante de régression linéaire multiple du logiciel VB3.0.6 trouve les meilleurs modèles de RLM possible en fonction des critères définis par l'utilisateur. Alors que le nombre de variables indépendantes incluses aux fins de modélisation augmente, le nombre de modèles possibles décrivant leur relation linéaire croît exponentiellement, de sorte qu'un examen minutieux de ces critères s'avère essentiel.

L'utilisateur doit d'abord établir le nombre maximal de variables d'un modèle. En règle générale, un modèle doit comporter à tout le moins de 10 à 20 observations par paramètre, sinon il peut être surajusté et, par conséquent, offrir une faible estimation. Conformément aux directives décrites dans le document intitulé *Virtual Beach 3.0.6: User's Guide*, une limite maximale permise de variables

indépendantes = $\frac{\text{nombre d'observations}}{5}$ a été fixée pour éviter le surajustement.

La composante RLM du logiciel VB3.0.6 traite automatiquement la multicolinéarité entre les variables indépendantes. La multicolinéarité survient lorsque deux variables explicatives d'un modèle de RLM sont étroitement reliées entre elles. Les variables explicatives devraient être indépendantes et cette corrélation pourrait causer des problèmes lorsque le modèle est ajusté ou interprété. Un facteur d'inflation de la variance (VIF) est utilisé dans le logiciel VB3.0.6 pour rejeter les modèles contenant des variables ayant un degré élevé de multicolinéarité. Un VIF de 1 est idéal. Si l'une des variables indépendantes d'un modèle a un VIF dépassant le seuil défini par l'utilisateur (VIF > 5 dans la présente étude), ce modèle est exclu.

De plus, la valeur R carré ajustée a été sélectionnée comme critère d'évaluation pour l'ajustement du modèle et ceux dont la valeur R carré ajustée était de 0,1 ou plus étaient retenus. La valeur R carré ajustée est une mesure de la variation de la concentration de bactéries indicatrices de contamination fécale qui est expliquée par les variables explicatives. Étant donné que les faibles valeurs R carré ajustées indiquent une faible capacité de produire des prévisions précises, le seuil a été fixé à 0,1 puisque les modèles ayant des valeurs R carré ajustées inférieures à ce seuil n'allaient probablement pas être utiles aux fins de la présente étude. La valeur R carré ajustée est une version modifiée de la valeur R carré qui tient compte du nombre de prédicteurs dans un modèle de régression et qui constitue ainsi un meilleur critère d'évaluation pour comparer divers modèles de régression linéaire contenant des nombres différents de prédicteurs. Si p est défini comme étant le nombre de paramètres dans un modèle, n comme étant le nombre d'observations dans l'ensemble de données, SCR comme étant la somme des carrés des résidus d'un modèle et STC comme étant la somme totale des carrés d'un modèle, le critère d'évaluation de tout modèle peut être défini comme suit :

$$\text{Valeur } R^2 = 1 - \frac{SCR}{STC} \text{ et Valeur } R^2 \text{ ajusté} = 1 - (1 - R^2) \frac{(n-1)}{(n-p-1)}$$

Les modèles dont les valeurs p des coefficients du point d'intersection et des variables indépendantes étaient inférieures à 0,05 ont été retenus comme étant statistiquement significatifs. Si un ou plusieurs des coefficients avaient une valeur $p > 0,05$, les modèles n'étaient plus évalués. Une valeur p inférieure ou égale à 0,05, qui indique de façon évidente que le coefficient est inférieur à zéro, est une valeur normalisée utilisée dans l'analyse de régression linéaire multiple.

Les modèles ont fait l'objet d'une évaluation plus approfondie en fonction de trois critères supplémentaires : la précision, la spécificité et la sensibilité. Ces critères comparent les résultats observés et prédits par le modèle à un critère (habituellement une norme ou une limite recommandée de qualité de l'eau) pour que les vrais/faux positifs et les vrais/faux négatifs puissent être définis. Le critère a été fixé à 35 pour les entérocoques et à 200 pour *E. coli* pour la modélisation des valeurs de la moyenne géométrique et à 70 NPP/100 mL pour la modélisation des concentrations d'entérocoques dans un échantillon unique. Les prévisions du modèle supérieures à ce seuil sont considérées comme étant des dépassements/positifs et celles qui sont inférieures à cette valeur sont considérées comme étant des non-dépassements/négatifs. Plus la valeur se rapproche de 1 pour ces critères, meilleur est le rendement du modèle.

$$\text{Précision} = \frac{(\text{vrais positifs} + \text{vrais négatifs})}{n}$$

$$\text{Spécificité} = \frac{\text{vrais négatifs}}{(\text{vrais négatifs} + \text{faux positifs})}$$

$$\text{Sensibilité} = \frac{\text{vrais positifs}}{(\text{vrais positifs} + \text{faux négatifs})}$$

De plus, même si la **section 2** a mis en lumière certaines préoccupations à propos des données relatives à *E. coli* en raison des caractéristiques de l'ensemble de données de 2017 comparativement à celui de 2018, tant les ensembles de données relatives aux entérocoques qu'à *E. coli* ont été utilisés pour la mise au point du modèle de calcul de la moyenne géométrique.

3.3 Évaluation du modèle préliminaire

Les **tableaux 3-1 et 3-2** résument les modèles affichant les meilleurs résultats qui ont été mis au point pour toute la période de relevé de 2017 et 2018 à l'aide de la méthodologie décrite ci-dessus. Lorsque les modèles incluait la même variable indépendante, mais avec une transformation des variables indépendantes ou des bactéries indicatrices de contamination fécale, le modèle le plus performant selon la valeur R carré ajustée, la spécificité et la sensibilité est présenté. La mise au point du modèle relatif à *E. coli* s'est avérée particulièrement ardue à cause du manque de dépassements du critère de qualité de l'eau (c'est-à-dire *E. coli* > 200 comptes ou NPP/100 mL). Par conséquent, parmi les quelques modèles qui ont réussi à respecter le critère de RLM, aucun n'a été en mesure de prédire les dépassements du critère de qualité de l'eau, entraînant une sensibilité de zéro. En dépit de cela, les précipitations antérieures, la direction du vent et le niveau d'eau de la rivière Coal Branch se sont révélés être les variables indépendantes des modèles, une constatation compatible avec le modèle conceptuel du transport des bactéries vers la plage Parlee.

Tableau 3-1. Modèles de RLM prédisant la moyenne géométrique d'*E. coli* à la plage Parlee (2017-2018)

Variable étudiée	Variables indépendantes	Valeur R carré ajustée	Faux positifs	Spécificité	Faux négatifs	Sensibilité	Précision
MG (<i>E. coli</i>)	Précipitations des 24 heures précédentes, niveau d'eau de la rivière Coal Branch, direction du vent	0,1223	0	1	2	0	0,9009
LN (MG [<i>E. coli</i>])	Précipitations des 24 heures précédentes, niveau d'eau de la rivière Coal Branch	0,1296	0	1	2	0	0,99277
LN (MG [<i>E. coli</i>])	Précipitations des 24 heures précédentes, niveau d'eau de la rivière Petitcodiac	0,1084	0	1	2	0	0,99277

Les modèles de RLM présentés au **tableau 3-2** font état d'une légère amélioration de la diminution des faux négatifs et affichent un nombre accru de faux positifs, mais ils ont globalement une capacité restreinte en matière de prévision. Le modèle qui concilie le mieux les faux positifs et les faux négatifs (**Éqn 1**) est celui dont les variables indépendantes sont composées des précipitations, du niveau d'eau, du jour de l'année et des jours depuis une précipitation. Même si ce modèle ne convenait pas à une mise en

œuvre opérationnelle, les variables indépendantes utilisées sont conformes aux observations empiriques selon lesquelles le moment de l'année et la durée depuis les dernières précipitations ont une certaine relation avec les concentrations de bactéries indicatrices de contamination fécale à la plage.

Tableau 3-2. Modèles de RLM prédisant la moyenne géométrique d'entérocoques à la plage Parlee (2017-2018)

Variable étudiée	Variabes indépendantes	Valeur R carré ajustée	Faux positifs	Spécificité	Faux négatifs	Sensibilité	Précision
MG (entérocoques)	Précipitations des 24 heures précédentes, niveau d'eau Coal Branch	0,1608	10	0,96212	10	0,23076	0,92779
MG (entérocoques)	Carré (précipitations des 24 heures précédentes), niveau d'eau Coal Branch, carré (niveau d'eau Petitcodiac)	0,3046	6	0,97727	11	0,15384	0,93862
MG (entérocoques)	Carré (précipitations des 24 heures précédentes), carré (précipitations des 48 heures précédentes), niveau d'eau Coal Branch, LOG10 (jour de l'année), racine carrée (jours depuis la précipitation)	0,4854	5	0,97297	9	0,18181	0,92857
LN (MG [entérocoques])	Précipitations des 24 heures précédentes, niveau d'eau Coal Branch	0,1262	0	1	12	0,07692	0,95667

GM_Entero_MPN =

$$494.2 + 0.07906 * (\text{SQUARE}(24\text{-hr AntecedentRainfall_mm})) + 0.01454 * (\text{SQUARE}(48\text{-hr Antecedent Rainfall_mm})) - 211.6 * (\text{CoalBranch_WSE_m}) - 123.2 * (\text{LOG10}(\text{DayOFYear})) - 5.915 * (\text{SQUAREROOT}(\text{DaysSinceRain}))$$

(Éqn 1)

La **figure 3-1** présente respectivement les graphiques chronologiques des saisons 2017 et 2018, calculés à l'aide de l'**équation 1**. Le rendement du modèle semble meilleur pour la saison 2018, particulièrement à la fin août en raison de fortes pluies (48,5 mm). Un diagramme de dispersion des concentrations d'entérocoques observées et prévues tirées de l'**équation 1** (**figure 3-2**) montre une courbe de correspondance parfaite (c'est-à-dire que les concentrations prévues et observées correspondent parfaitement). À la vue de ce diagramme, il est évident que la grande précision générale provient du nombre relativement élevé d'observations sous le seuil de 35 NPP/100 mL.

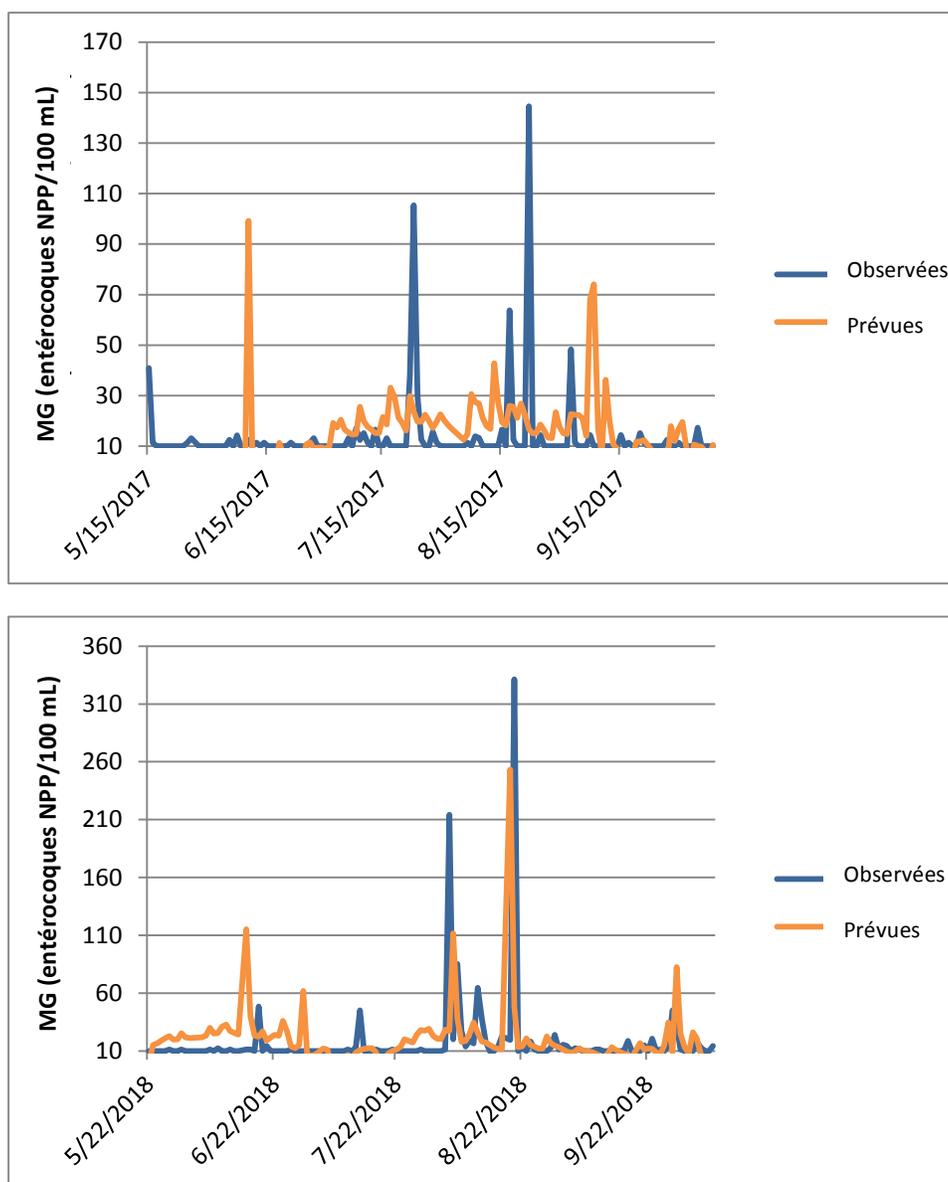


Figure 3-1. Graphiques chronologiques de la moyenne géométrique des concentrations prévues et observées d'entérocoques en 2017 (haut) et 2018 (bas)

À la suite de conversations avec le personnel du ministère de l'Environnement et des Gouvernements locaux et du Bureau du médecin-hygiéniste en chef du Nouveau-Brunswick, l'ensemble de données a aussi été divisé en deux périodes, soit de mai à juillet et d'août à octobre. Alors que les données de mai à juillet n'ont pas produit de modèles statistiquement significatifs, celles d'août à octobre ont produit des modèles répondant aux exigences de signification statistique en utilisant les précipitations au cours des 24 heures précédentes, les jours depuis les précipitations et le jour de l'année comme variables indépendantes. Malgré le fait que le rendement du modèle de prévision n'arrivait toujours pas à un niveau qui en justifierait la mise en œuvre opérationnelle, le résultat soutient deux conclusions découlant de la mise au point de ce modèle préliminaire et de l'analyse des seules données relatives aux précipitations qui sont décrites à la prochaine

section : les facteurs ayant une incidence sur les concentrations de bactéries indicatrices de contamination fécale à la plage Parlee semblent évoluer au cours de la saison récréative et la durée depuis la précipitation, en plus de la quantité de pluie, s'avère un facteur important.

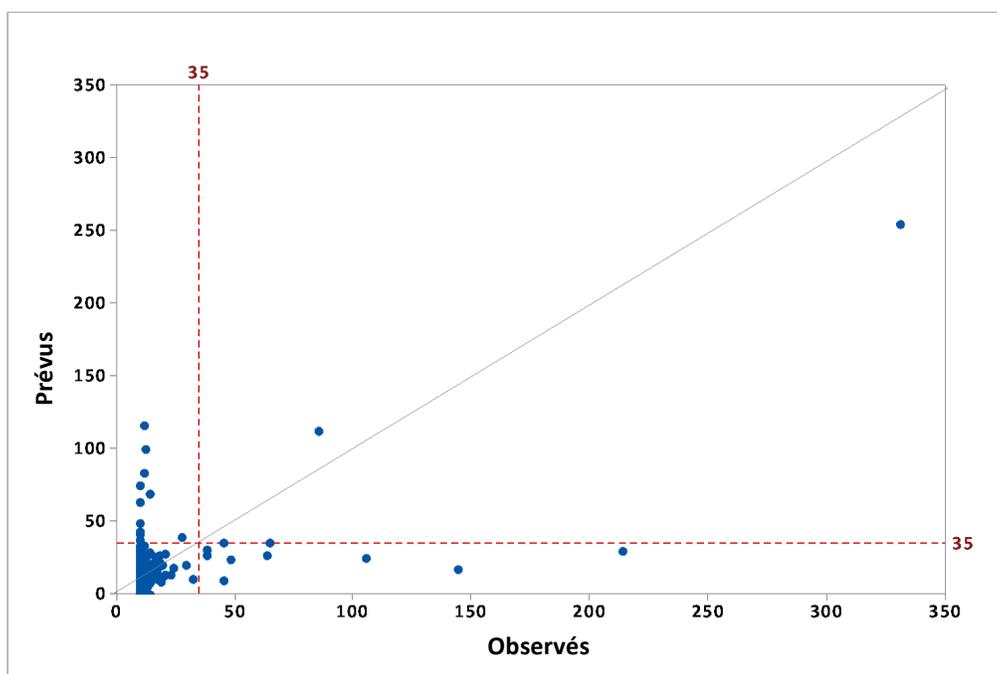


Figure 3-2. Diagramme de dispersion des entérocoques observés et prévus (Éqn 1) avec une courbe diagonale de correspondance parfaite et les courbes de référence du critère de qualité des eaux utilisées à des fins récréatives

3.4 Modèles de précipitations antérieures

Des précipitations de 10 mm ou plus observées au cours des 24 heures précédentes sont actuellement utilisées comme critère d'interdiction préventif pour la plage Parlee. Il s'agit d'une démarche empirique très répandue selon laquelle des concentrations élevées de bactéries indicatrices de contamination fécale sont habituellement associées à un certain seuil de précipitation et qui s'appuie sur le modèle conceptuel des précipitations et de l'écoulement subséquent comme mécanismes de transport de ces bactéries vers les eaux côtières utilisées à des fins récréatives. Dans le rapport *Qualité de l'eau de la plage Parlee : rapport final* (2018), certaines questions ont été soulevées à propos de l'utilité du seuil de précipitation de 10 mm comme indicateur utile de concentrations élevées de bactéries. L'un des objectifs de cette étude consistait à évaluer l'utilité du critère actuellement employé et à déterminer si des recommandations supplémentaires pouvaient être formulées concernant l'utilisation d'un seuil de précipitation pour les interdictions préventives. Les entérocoques ont été utilisés comme bactérie indicatrice de contamination fécale pour cette évaluation en raison du nombre élevé de valeurs dépassant le critère relatif à celles-ci et de la cohérence des données relatives aux entérocoques tout au long de la période de relevé.

La première étape de l'analyse a consisté à tracer les précipitations au cours des 24 heures précédentes par rapport aux concentrations d'entérocoques. Comme les **figures 3-3** et **3-4** le montrent, l'utilisation du seuil de 10 mm entraîne un nombre relativement important de faux positifs ainsi que plusieurs faux négatifs. Par exemple, à la **figure 3-3**, alors que l'emploi du seuil de 10 mm a produit cinq jours au cours desquels les concentrations d'entérocoques ont dépassé le critère de 35 NPP/100 mL et qu'une interdiction de baignade aurait été correctement affichée, il y a eu presque deux fois plus de jours durant lesquels l'utilisation des seules précipitations antérieures aurait laissé la plage ouverte à la baignade en dépit de concentrations d'entérocoques supérieures aux critères recommandés. Le nombre de jours au cours desquels les seules précipitations auraient justifié un affichage préventif à la plage, alors que les tests subséquents sur les bactéries indicatrices de contamination fécale ont indiqué des concentrations inférieures au niveau de préoccupation concernant l'utilisation des eaux à des fins récréatives est encore plus fréquent.

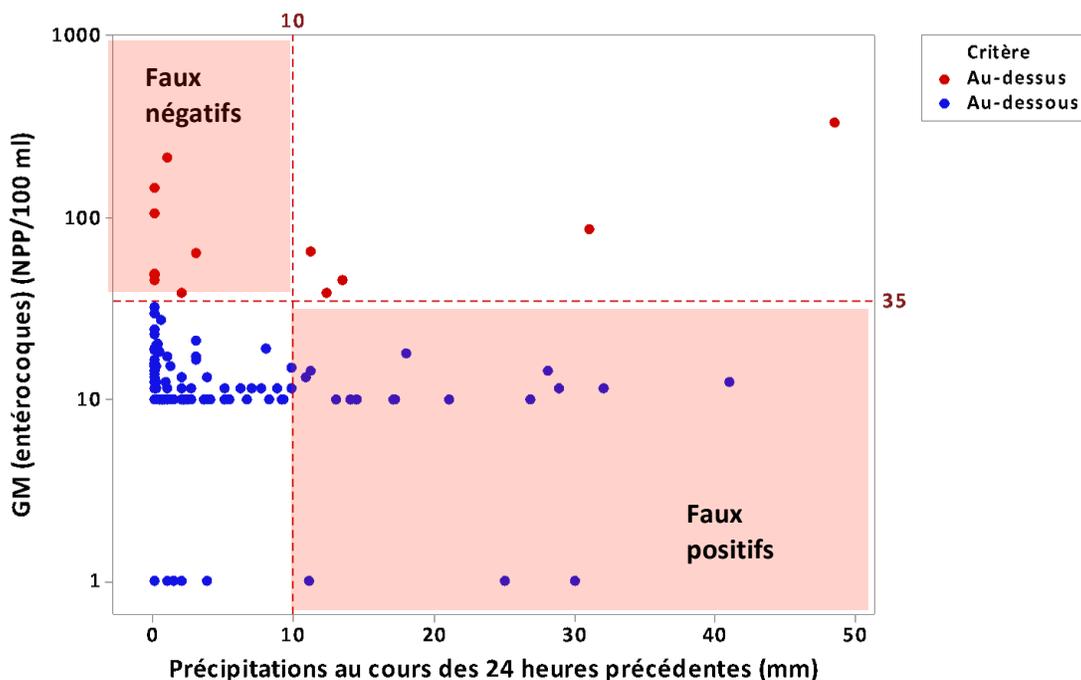


Figure 3-3. Diagramme de dispersion de la moyenne géométrique des entérocoques et des précipitations au cours des 24 heures précédentes (2017-2018)

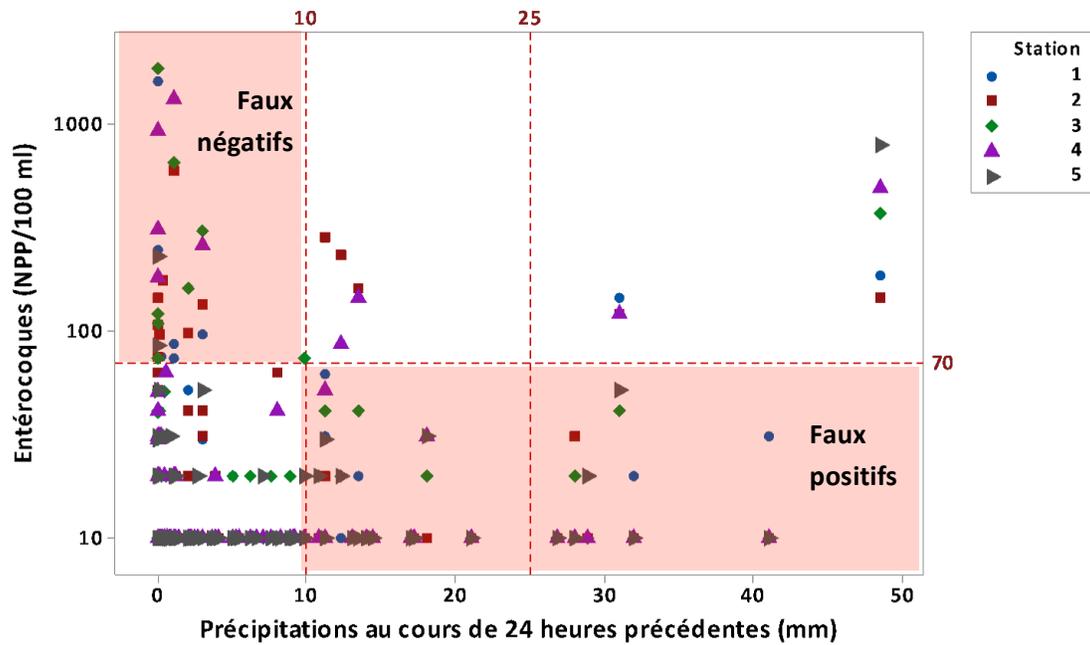


Figure 3-4. Diagramme de dispersion des entérocoques et des précipitations au cours des 24 heures précédentes (2017-2018)

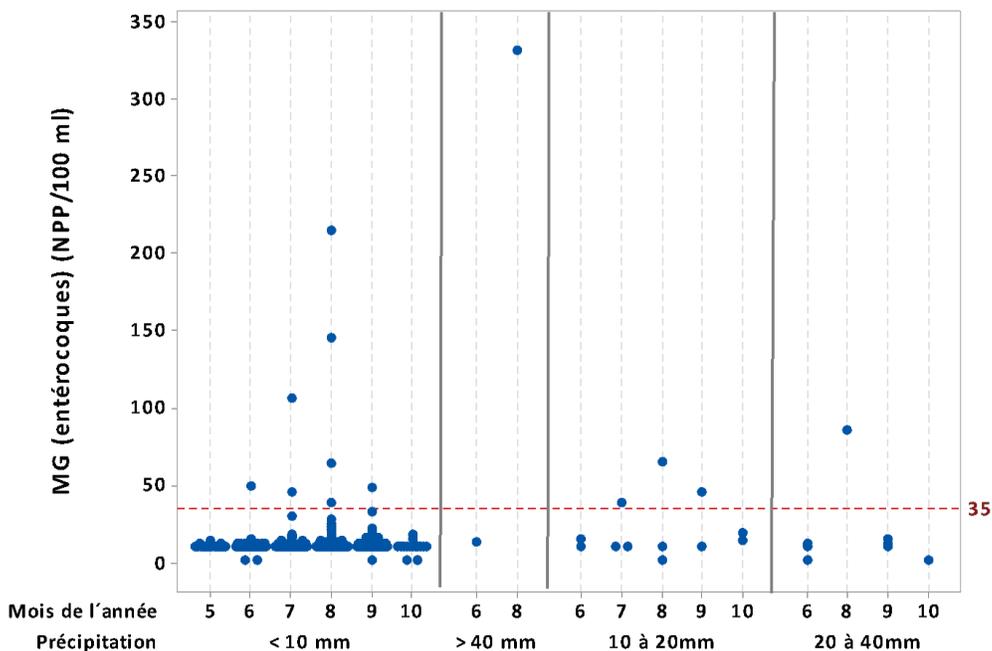


Figure 3-5. Diagramme des valeurs individuelles de la moyenne géométrique des concentrations d'entérocoques par mois et des précipitations associées au cours des 24 heures précédentes (2017-2018)

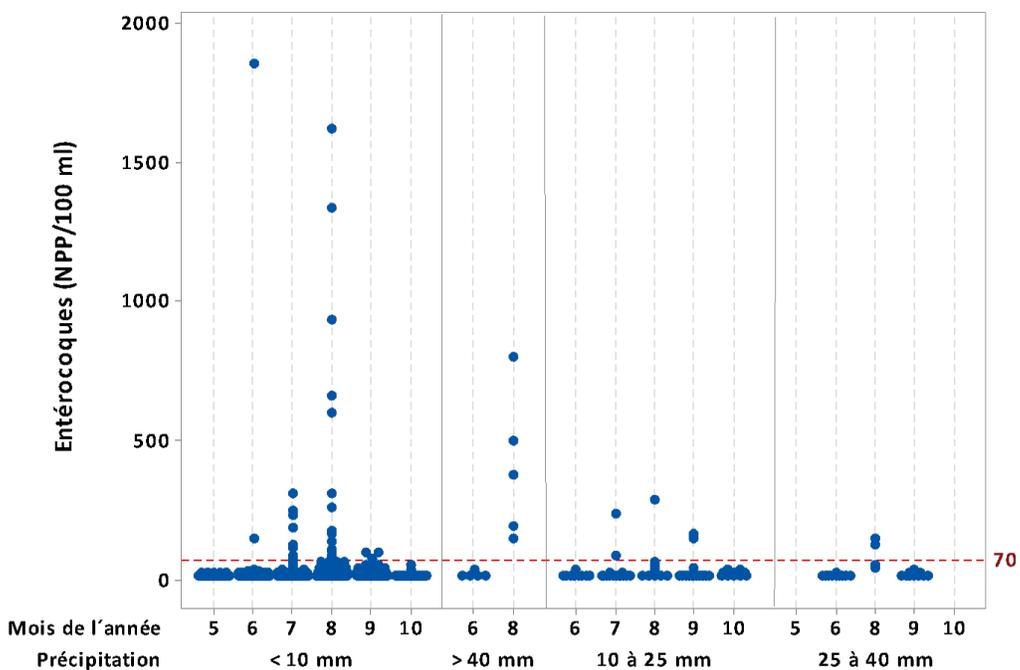


Figure 3-6. Diagramme des valeurs individuelles des concentrations d'entérocoques dans un échantillon unique par mois et des précipitations associées au cours des 24 heures précédentes (2017-2018)

Afin de mieux comprendre le comportement potentiel des précipitations antérieures comme prédicteur de la qualité microbienne de l'eau, les données relatives aux entérocoques ont été jumelées à une catégorie de précipitations (< 10 mm, 10 à 20 mm, 20 à 40 mm et > 40 mm) et ont été tracées par mois (**figure 3-5** et **figure 3-6**). Le schéma des données dans les deux figures donne lieu à des observations utiles. Tout d'abord, le nombre d'observations de concentrations d'entérocoques supérieures à 35 NPP/100 mL (**figure 3-5**) ou à 70 NPP/100 mL (**figure 3-6**) a augmenté au cours des saisons récréatives de 2017 et 2018 alors que les précipitations antérieures étaient en fait inférieures à 10 mm par rapport à toute autre valeur de précipitation. Même lorsque les valeurs de précipitation sont supérieures à 10 mm, il est plus probable qu'improbable que les concentrations d'entérocoques soient inférieures au critère de qualité des eaux utilisées à des fins récréatives. En outre, il convient de noter que de fortes pluies sont plus susceptibles d'agir comme prédicteur de juillet à septembre, même si c'est durant cette période que les dépassements du critère de qualité de l'eau risquent de se produire, sans égard à la quantité de pluie.

Le **tableau 3-3** démontre que le seuil prédirait correctement environ 38 % des jours (5 des 13 jours) durant lesquels la moyenne géométrique de 35 NPP/100 mL a été dépassée. Cependant, des interdictions seraient affichées au cours de 18 autres jours avec ce seuil, alors que les concentrations de bactéries indicatrices ont par la suite affiché des valeurs inférieures aux limites recommandées (c'est-à-dire faux positifs). Si le seuil passait à 25 mm (**tableau 3-4**), seuls 2 des 13 jours (15 %) seraient prédits correctement, alors que 7 autres jours feraient l'objet d'interdictions non justifiées et que 11 jours n'en feraient pas l'objet en dépit de concentrations de bactéries indicatrices supérieures à la moyenne géométrique de la limite recommandée. Il convient aussi de noter que, pendant les quatre jours de la période de relevé affichant les concentrations moyennes géométriques les plus élevées d'entérocoques, un seul aurait été correctement prédit comme étant une journée visée par une interdiction de baignade, soit par le seuil de précipitation de 10 mm, soit par celui de 25 mm.

De même, si les cinq stations de surveillance sont étudiées et comparées à la valeur maximale recommandée dans un échantillon unique de 70 NPP/100 mL, le seuil de 10 mm permettrait de prédire environ 30 % des observations totales (journées d'interdiction de fréquentation de la plage) ayant dépassé la limite recommandée et le seuil de 25 mm en prédirait environ 19 % (**tableau 3-4**). En l'absence d'autres données, le seuil de précipitation appuie dans une certaine mesure la prise de décision, mais, en ce qui a trait aux seuils de 10 mm et 25 mm, il le fait au prix de presque trois fois plus d'avis d'interdiction que la qualité réelle de l'eau ne l'exigerait, tout en n'arrivant pas à déterminer plus de la moitié des jours au cours desquels la qualité de l'eau nécessiterait l'affichage d'avis d'interdiction. Il convient de noter que le seuil de précipitation de 25 mm n'augmenterait pas beaucoup la spécificité, mais qu'il diminuerait considérablement la sensibilité.

Tableau 3-3. Résumé du rendement du seuil de précipitation pour la moyenne géométrique d'entérocoques (2017-2018)

Seuil de précipitation	n	Vrais positifs	Vrais négatifs	Faux positifs	Faux négatifs	Sensibilité	Spécificité	Précision
10 mm	285	5	254	18	8	0,3846	0,9338	0,9088
25 mm	285	2	265	7	11	0,1538	0,9743	0,9368

Tableau 3-4. Résumé du rendement du seuil de précipitation pour un échantillon unique d'entérocoques (2017-2018)

Seuil de précipitation	n	Vrais positifs	Vrais négatifs	Faux positifs	Faux négatifs	Sensibilité	Spécificité	Précision
10 mm	1 385	13	1 261	82	29	0,3095	0,9389	0,9199
25 mm	1 385	8	1 316	27	34	0,1905	0,9799	0,9559

Même si l'utilisation d'un seuil constitue fondamentalement un modèle empirique, comme il en a été question précédemment, la relation entre les précipitations et les concentrations d'entérocoques a été officialisée à l'aide de deux types d'analyse de régression. Tout d'abord, la régression linéaire multiple a été employée pour se pencher sur la relation entre la moyenne géométrique des entérocoques et des précipitations et les jours depuis les précipitations pour toute la période de relevé (équations A et B) et pour les mois de juillet à septembre (équations C et D). Les résultats de cette analyse avec le logiciel VB3.0.6 pour les entérocoques transformés et non transformés et les variables indépendantes indiquent que les modèles donnant les meilleurs résultats sont les suivants :

$$GM_Entero = 13.71 + 0.08533*(SQUARE(24-hr AntecedentRainfall_mm)) \quad (\text{Éqn A})$$

$$GM_Entero = 22.2087 + 0.0835*SQUARE[24-hr AntecedentRainfall_mm] - 6.8765*SQUAREROOT[DaysSinceRain] \quad (\text{Éqn B})$$

$$GM_Entero \text{ (July-Sept)} = 13.57 + 0.09939*(SQUARE(24-hr AntecedentRainfall_mm)) \quad (\text{Éqn C})$$

$$GM_Entero \text{ (July-Sept)} = 19.55 + 0.09631*(SQUARE(24-hr AntecedentRainfall_mm)) - 4.806*(SQUAREROOT(DaysSinceRain)) \quad (\text{Éqn D})$$

Tableau 3-5. Modèles de régression linéaire multiple utilisant seulement les précipitations pour prédire la moyenne géométrique des entérocoques à la plage Parlee

Équation	Valeur R carré ajustée	Faux positifs	Spécificité	Faux négatifs	Sensibilité	Précision	n
A	0,3697	3	0,98378	9	0,18182	0,93878	196
B	0,4002	3	0,98378	9	0,18182	0,93878	196
C	0,4298	5	0,97076	10	0,16666	0,91803	183
D	0,4442	6	0,96491	9	0,25	0,91803	183

Ces résultats (**tableau 3-5**) laissent supposer que de 30 à 40 % environ de la variabilité des bactéries indicatrices de contamination fécale peut s'expliquer par les précipitations

antérieures et les jours depuis la dernière pluie. Il est étonnant de constater que, même si le nombre de jours depuis les précipitations est effectivement apparu comme facteur dans deux des quatre modèles présentés au **tableau 3-5**, le coefficient correspondant aux jours depuis la dernière pluie est négatif. Ce résultat va à l'encontre de l'idée voulant qu'une plus longue période de temps sec entraîne un accroissement des bactéries qui seraient transportées vers la plage au cours des précipitations suivantes. Néanmoins, la régression linéaire multiple révèle que le nombre de jours, particulièrement au cours de la période allant de juillet à septembre, est une variable qui permet d'expliquer la variabilité observée sur le plan des concentrations de bactéries indicatrices de contamination fécale. Même si la capacité des modèles de prédire les faibles concentrations et le grand nombre de valeurs inférieures à 35 NPP/100 mL font en sorte que la précision des modèles soit élevée (> 90 %), le grand nombre de faux négatifs illustre le rendement relativement mauvais des modèles pour prédire les dépassements de la valeur du critère. Puisque le nombre d'observations utilisées pour élaborer les équations présentées au **tableau 3-5** est plus petit que le nombre total d'observations disponibles employées dans le cadre de l'analyse présentée au **tableau 3-3**, toute comparaison directe est limitée (c'est-à-dire que ce ne sont pas toutes les mêmes données qui sont incluses dans chaque calcul de spécificité et de sensibilité). Toutefois, en fonction des données utilisées pour chacun, comparativement aux modèles de RLM, il semble que le seuil de précipitation de 10 mm est plus susceptible de produire de faux positifs et que le seuil de 25 mm n'offre aucune amélioration par rapport aux modèles de RLM et qu'il entraîne un recul comparativement au seuil de 10 mm pour prédire, relativement aux entérocoques, les valeurs supérieures aux limites recommandées de qualité des eaux utilisées à des fins récréatives.

L'utilisation des valeurs dans un échantillon unique d'entérocoques accroît la taille de l'ensemble de données disponibles pour la mise au point des modèles, mais, à cause du faible pourcentage de valeurs dépassant la limite recommandée de qualité de l'eau de 70 NPP/100 mL, les RLM créées sont aussi caractérisées par une faible sensibilité (**tableau 3-6**), c'est-à-dire que la capacité de prédire les dépassements de la limite recommandée de qualité des eaux utilisées à des fins récréatives de 70 NPP/100 mL est limitée. Les valeurs R carré ajustées sont même inférieures aux RLM mises au point à l'aide des données relatives aux moyennes géométriques d'entérocoques. Selon le rendement de ces modèles, ils ne seraient pas recommandés pour utilisation dans un milieu opérationnel.

$$\begin{aligned} \text{LN(Enteroc)} = & 5.813 + 0.04242*(24\text{-hr AntecedentRainfall_mm}) + 0.01006*(\text{Rainfall_mm_48hr}) \\ & + 4.48e-06*(\text{SQUARE}(\text{WindDirection_deg})) + 1.708e-05*(\text{SQUARE}(\text{ChangeWindDirection_deg})) \\ & - 0.04143*(\text{DaysSinceRain}) - 3.62*(\text{CoalBranch_WSE_m}) - 8.507e-06*(\text{SQUARE}(\text{DayOFYear})) \end{aligned} \quad (\text{Éqn E})$$

$$\begin{aligned} \text{LN(Enteroc)} = & 5.754 + 0.05236*(24\text{-hr AntecedentRainfall_mm}) + 1.715e-05*(\text{SQUARE}(\text{ChangeWindDirection_deg})) \\ & - 0.05086*(\text{DaysSinceRain}) + 4.124e-06*(\text{SQUARE}(\text{WindDirection_deg})) - 8.31e-06*(\text{SQUARE}(\text{DayOFYear})) \\ & - 3.497*(\text{CoalBranch_WSE_m}) \end{aligned} \quad (\text{Éqn F})$$

$$\begin{aligned} \text{LN(Enteroc)} = & 10.37 + 0.03894*(24\text{-hr AntecedentRainfall_mm}) + 0.007361*(\text{Rainfall_mm_72hr}) \\ & + 2.584e-06*(\text{SQUARE}(\text{WindDirection_deg})) - 0.1038*(\text{SQUAREROOT}(\text{DaysSinceRain})) \\ & - 3.677*(\text{CoalBranch_WSE_m}) - 0.8801*(\text{LN}(\text{DayOFYear})) \end{aligned} \quad (\text{Éqn G})$$

Tableau 3-6. Modèles de RLM prédisant les concentrations d'entérocoques dans un échantillon unique à la plage Parlee

Équation	Valeur R carré ajustée	Faux positifs	Spécificité	Faux négatifs	Sensibilité	Précision	Période	n
E	0,1984	3	1	31	0,13888	0,96287	Juillet à septembre	835
F	0,1933	0	1	31	0,1388	0,9629	Juillet à septembre	835
G	0,1825	5	1	31	0,13888	0,96836	Mai à octobre	980

Comme alternative à la RLM, la régression logistique binomiale (RLB) a également été étudiée. La RLB modélise la relation entre un ensemble de prédicteurs et une réponse binaire, soit une réponse ayant deux résultats possibles. Dans ce cas, la valeur de la moyenne géométrique des bactéries indicatrices de contamination fécale est supérieure ou inférieure au critère de qualité des eaux utilisées à des fins récréatives. Le logiciel statistique Minitab 18 a été employé pour la création des équations de RLB. Pour la RLB, une probabilité de dépassement du seuil (c'est-à-dire probabilité de 50 %, de 25 %) est fixée, puis la prédiction pour l'une des deux réponses binaires est évaluée. Le modèle donnant les meilleurs résultats a été produit lorsque les précipitations antérieures et le nombre de jours depuis la pluie sont inclus comme prédicteurs. Deux niveaux de probabilité sont compris aux fins de comparaison, 25 % et 50 %. Une valeur de 50 % signifie que, si le modèle prédit qu'il y a une probabilité de dépasser le seuil de 35 NPP/100 mL de 50 % ou plus, le résultat binaire attribué veut que les entérocoques excèdent la norme. Même si la RLB peut parfois s'avérer utile pour prédire une valeur « supérieure/inférieure » comparativement à une valeur exacte, son utilisation semble ne pas beaucoup réduire le nombre de faux négatifs et pourrait ne pas offrir d'amélioration importante par rapport à l'utilisation du modèle de RLM pour les précipitations (tableau 3-7).

L'équation utilisée pour la régression logistique binomiale est la suivante :

$$\text{Probability (GM_Entero>35)} = \exp(Y') / (1 + \exp(Y'))$$

Où :

$$Y' = -2,004 + 0,0856 * \text{précipitations des 24 heures précédentes_mm} - 0,887 * \text{jours depuis la pluie (Éqn 2)}$$

Tableau 3-7. Modèles de RLB utilisant uniquement les précipitations pour prédire la moyenne géométrique des entérocoques à la plage Parlee

Seuil de probabilité	Faux positifs	Spécificité	Faux négatifs	Sensibilité	Précision	n
0,25	3	0,9769	9	0,3333	0,9402	183
0,50	0	0,9885	10	0,1666	0,9457	183

3.5 Conclusions et recommandations

La mise au point du modèle préliminaire et l'analyse des précipitations antérieures ont permis de tirer les principales conclusions suivantes à propos de l'utilisation des modèles

de prévision et des seuils de précipitation pour la gestion éventuelle de l'accès récréatif à la plage Parlee :

- Alors que plusieurs modèles de prévision valides ont été mis au point à partir des ensembles de données de 2017-2018, aucun n'a donné un rendement suffisant pour que l'on puisse en faire d'emblée la mise en œuvre opérationnelle à la plage Parlee. Les variables indépendantes utilisées dans les modèles comprenaient la plupart du temps les précipitations au cours des 24 heures précédentes, le niveau d'eau de la rivière Coal Branch et les jours depuis la précipitation. La fréquence des précipitations antérieures à titre de variable indépendante est conforme aux conclusions de l'analyse exploratoire antérieure des données.
- La direction du vent ou le changement de direction du vent n'ont pas semblé être des variables indépendantes importantes pour la modélisation de la moyenne géométrique des concentrations, seule la direction du vent ayant été utilisée dans l'un des modèles valides. Cela pourrait en partie découler de la prédominance des vents du sud-ouest et/ou du manque généralisé d'observations dépassant la limite recommandée de qualité des eaux utilisées à des fins récréatives. La direction du vent et le changement de direction du vent ont été utilisés comme variables indépendantes dans le cadre de la modélisation pour prédire la valeur de l'échantillon unique en ce qui a trait aux entérocoques.
- Alors que des modèles valides ont été confirmés, les valeurs R carré ajustées étaient généralement faibles, la modélisation la plus performante affichant une valeur R carré ajustée modérée de 0,4954.
- Les modèles de régression linéaire multiple se caractérisaient par une sensibilité généralement faible, ce qui signifie qu'ils n'ont pas très bien réussi à déceler les vrais dépassements des limites recommandées de qualité des eaux utilisées à des fins récréatives. Sur le plan de la sensibilité, le modèle de RLM donnant les meilleurs résultats avait une valeur de sensibilité de 0,25 et utilisait les précipitations au cours des 24 heures précédentes et le nombre de jours depuis les précipitations durant la période allant de juillet à septembre comme variables prédictives.
- La régression logistique binomiale (RLB) a également été mise à l'essai comme solution de rechange à la RLM. Le rendement a été semblable à celui de la RLM, avec une probabilité relativement faible de dépassement du critère requis pour atteindre une sensibilité conforme aux meilleurs modèles de RLM. Dans la pratique, l'utilisation d'un faible seuil de probabilité de dépassement réduit l'utilité du modèle.

À la suite de la mise au point du modèle préliminaire et de l'analyse des données relatives aux précipitations en vue de l'affichage d'interdictions préventives à la plage Parlee, les recommandations suivantes sont aussi formulées :

- Compte tenu du nombre relativement faible de dépassements des limites recommandées de qualité des eaux utilisées à des fins récréatives en ce qui a trait aux entérocoques et à *E. coli*, il est peu probable que l'on puisse mettre au point des modèles de prévision robustes convenant à une utilisation opérationnelle. Si la modélisation devait faire l'objet d'un projet pilote durant la saison 2019, les modèles présentés à l'**équation 1** ou l'**équation B** pour l'ensemble de la saison de fréquentation de la plage ou à l'**équation D** pour la période de juillet à septembre seraient recommandés selon les analyses réalisées. Même dans ce cas, il est peu probable que l'on obtienne des prévisions concluantes et les modèles ne devraient pas être utilisés comme seul outil de prise de décision concernant les avis d'interdiction.
- En raison du faible nombre de dépassements du critère de qualité des eaux utilisées à des fins récréatives et du défi que cela représente pour la modélisation prédictive, les autres méthodes d'évaluation rapide de la qualité microbienne de l'eau ne devraient pas être éliminées. La réaction de polymérisation en chaîne quantitative, ou RPC quantitative, est une méthode rapide employée dans plusieurs plans d'eau utilisés à des fins récréatives aux États-Unis, qui a été acceptée par l'EPA des États-Unis pour l'évaluation des conditions de qualité de ces eaux pour les plages tant d'eau douce que de mer.
- On n'a pas démontré que la variable des précipitations antérieures de 10 mm constituait un prédicteur fiable de la qualité microbienne de l'eau à la plage Parlee, selon les données recueillies pendant les saisons de fréquentation de la plage de 2017 et 2018.
- Nous recommandons d'ajouter la mesure de la turbidité sur le terrain comme paramètre de qualité de l'eau pour 2019 ainsi que l'évaluation subséquente du paramètre comme indicateur de substitution et/ou variable indépendante supplémentaire pour la modélisation prédictive. (Consulter la **section 2.6** pour obtenir des détails supplémentaires.)
- Même si les données de 2017-2018 indiquent que les dépassements du critère de qualité de l'eau sont plus nombreux à la fin de l'été, l'utilisation routinière de la RPC quantitative pourrait en fait être plus utile pour la protection de la santé publique au début de l'été (mai et juin), moment de la saison récréative où les précipitations semblent être un indicateur encore moins efficace de la qualité de l'eau.
- Une combinaison du modèle de régression linéaire multiple ou du seuil de précipitation antérieure et de la RPC quantitative pourrait aussi être envisagée. Par exemple, le seuil de précipitation actuel pour les avis d'interdiction préventive a tendance à surestimer les jours où la qualité de l'eau dépasse réellement la limite recommandée, c'est-à-dire qu'il produit de nombreux faux positifs. Si la valeur des précipitations sur 24 heures, qui est aussi une variable météorologique prévisible, était utilisée comme outil de dépistage, une analyse à l'aide de la RPC quantitative pourrait alors être effectuée sur les échantillons prélevés lorsque les précipitations dépassent le seuil. Cela permettrait de créer un système de prévision de la qualité de l'eau qui pourrait orienter l'utilisation de l'analyse à l'aide de la RPC quantitative de manière ciblée et rentable, particulièrement pendant la période

allant de juillet à septembre, et éventuellement réduire le nombre de « faux positifs » découlant de la seule utilisation du seuil de précipitation.

Glossaire

Glossaire

Précision – Mesure de la proximité des valeurs calculées par rapport à la valeur réelle. Dans la présente étude, la précision est mesurée par rapport à une valeur calculée (c'est-à-dire modélisée ou estimée) comme étant supérieure ou inférieure à un critère d'intérêt (c'est-à-dire une limite recommandée de qualité des eaux utilisées à des fins récréatives).

$$\text{Précision} = \frac{(\text{vrais positifs} + \text{vrais négatifs})}{n}$$

Valeur R carré ajustée – Mesure indiquant à quel point les données correspondent au modèle ou, autrement dit, le pourcentage de données expliquées par le modèle. Il s'agit d'une variante de la valeur R carré qui compense mieux les ensembles de données comptant de multiples variables (voir **Valeur R carré**). Si p est défini comme étant le nombre de paramètres d'un modèle et n , comme étant le nombre d'observations dans l'ensemble de données, le critère d'évaluation de tout modèle peut être défini comme suit :

$$\text{Valeur } R^2 \text{ ajusté} = 1 - (1 - R^2) \frac{(n - 1)}{(n - p - 1)}$$

Corrélation – Mesure de la force et de la direction de l'association entre deux variables. La corrélation décrit la mesure dans laquelle les deux variables varient de façon concomitante (c'est-à-dire qu'une variable augmente ou diminue alors que l'autre augmente). La corrélation entre deux variables ne prouve nullement l'existence d'une relation de cause à effet entre les deux variables (c'est-à-dire que cela ne signifie pas que la variation d'une variable est causée par l'autre variable).

Faux positif – Dans le cadre de la présente étude, valeur prédite de concentration de bactéries supérieure à la concentration recommandée en matière de qualité de l'eau lorsque la concentration mesurée était inférieure à la concentration recommandée.

Faux négatif – Dans le cadre de la présente étude, valeur prédite de concentration de bactéries inférieure à la concentration recommandée en matière de qualité de l'eau lorsque la concentration mesurée était supérieure à la concentration recommandée.

Variable indépendante ou explicative – La valeur d'une variable indépendante ne dépend pas d'une autre variable. Les variables indépendantes (dans le cadre de la présente étude, les précipitations, la température, etc.) servent à prédire la valeur d'une autre variable, la variable dépendante ou expliquée (dans le cadre de la présente étude, la concentration d'entérocoques ou d'*E. coli*).

Données censurées en amont – Ensemble de données ne comprenant aucune valeur en deçà d'un certain seuil. Dans le cadre de la présente étude, les données peuvent être rapportées sous forme de concentration bactérienne inférieure à une concentration donnée (c'est-à-dire < 10 NPP/100 mL), sans qu'il soit possible de déterminer exactement à quel point la concentration réelle se situe sous le seuil.

Régression linéaire multiple – Méthode statistique utilisée pour déterminer une relation mathématique entre de multiples variables indépendantes et une variable dépendante (dans le cadre de la présente étude, la concentration d'entérocoques ou d'*E. coli*).

Modèle surajusté – Il y a surajustement quand un modèle de régression s'ajuste de manière à correspondre aux caractéristiques d'un échantillon donné plutôt qu'à l'ensemble des données. Habituellement, un modèle surajusté peut bien correspondre aux données en question, mais être incapable de prédire précisément les données futures.

Valeur p – En régression linéaire multiple, la valeur p désigne la force de la preuve contre l'hypothèse selon laquelle le coefficient d'un modèle de régression linéaire n'est pas différent statistiquement de zéro (c'est-à-dire l'hypothèse nulle).

Valeur R carré – Mesure indiquant à quel point les données correspondent au modèle ou, autrement dit, pourcentage des données expliquées par le modèle (voir **Valeur R carré ajusté**). En utilisant SCR comme étant la somme des carrés des résidus d'un modèle et STC comme étant la somme totale des carrés d'un modèle, le critère d'évaluation d'un modèle peut être défini ainsi :

$$\text{Valeur } R^2 = 1 - \frac{SCR}{STC}$$

Sensibilité – Mesure du pourcentage de résultats positifs correctement déterminés comme étant des résultats positifs. Dans le cadre de la présente étude, la sensibilité est mesurée par rapport à une valeur calculée (c'est-à-dire modélisée ou estimée) comme étant supérieure ou inférieure à un critère d'intérêt (c'est-à-dire une limite recommandée de qualité des eaux utilisées à des fins récréatives).

$$\text{Sensibilité} = \frac{\text{vrais positifs}}{(\text{vrais positifs} + \text{faux négatifs})}$$

Spécificité – Mesure du pourcentage de résultats négatifs correctement déterminés comme étant des résultats négatifs. Dans le cadre de la présente étude, la spécificité est mesurée par rapport à une valeur calculée (c'est-à-dire modélisée ou estimée) comme étant supérieure ou inférieure à un critère d'intérêt (c'est-à-dire une limite recommandée de qualité des eaux utilisées à des fins récréatives).

$$\text{Spécificité} = \frac{\text{vrais négatifs}}{(\text{vrais négatifs} + \text{faux positifs})}$$

Statistiquement significatif – Expression utilisée pour décrire les données lorsque la probabilité de résultats survenant purement par hasard est inférieure à un seuil donné.

Bibliographie

ANDERSON, C. W. 2005. *Turbidity: U.S. Geological Survey Techniques of Water-Resources Investigations* (en ligne), Book 9, Chapter A6, Section 6.7, Washington (D.C.), U.S. Geological Survey. Sur Internet : <http://pubs.water.usgs.gov/twri9A6/>

DORFMAN, M., A. HAREN et J. DEVINE. 2014. *Testing the Waters. A Guide to Water Quality at Vacation Beaches*, s.l., NRDC.

FRANCY, D. S., A. M. G. BRADY, R. B. CARVIN, S. R. CORSI, L. M. FULLER, J. H. HARRISON, B. A. HAYHURST, J. LANT, M. B. NEVERS, P. J. TERRIO et T. M. ZIMMERMAN. 2013. *Developing and Implementing Predictive Models for Estimating Recreational Water Quality at Great Lakes Beaches* (en ligne), Reston (Va.), U.S. Geological Survey, « U.S. Geological Survey Scientific Investigations Report », n° 2013–5166, 68 p. Sur Internet : <http://dx.doi.org/10.3133/sir20135166/>

FRANCY, D., et R. DARNER. 2006. *Procedures for Developing Models to Predict Exceedances of Recreational Water Quality Standards at Coastal Beaches—Techniques and Methods 6-B5*, Reston (Va.), U.S. Geological Survey.

HELSEL, D. R., et R. M. HIRSCH. 2002. *Statistical Methods in Water Resources: Techniques of Water Resources Investigations* (en ligne), Book 4, Chapter A3, Washington (D.C.), U.S. Geological Survey. Sur Internet : <https://pubs.usgs.gov/twri/twri4a3/pdf/twri4a3-new.pdf>

MAS, D., et K. BAKER. 2011. *Guidance for Developing Predictive Models for Ontario Beaches*, Toronto, Ministère de l'Environnement de l'Ontario.

NOUVEAU-BRUNSWICK. COMITÉ DIRECTEUR DE LA QUALITÉ DE L'EAU DE LA PLAGE PARLEE. 2018. *Qualité de l'eau de la plage Parlee : rapport final* (en ligne), Fredericton, le Comité. Sur Internet : https://www2.gnb.ca/content/dam/gnb/Departments/eco-bce/Promo/Parlee_Beach/pdfs/parlee_beach_water_quality_finalreport-f.pdf

SANTÉ CANADA. 2012. *Recommandations au sujet de la qualité des eaux utilisées à des fins récréatives au Canada* (en ligne), 3^e éd., Ottawa, Santé Canada. Sur Internet : <https://www.canada.ca/fr/sante-canada/services/publications/vie-saine/recommandations-sujet-qualite-eaux-utilisees-fins-recreatives-canada-troisieme-edition.html>

USEPA. 2004. *Water Quality Standards for Coastal and Great Lakes Recreation Waters; Final Rule* (en ligne). Sur Internet : <https://www.govinfo.gov/content/pkg/FR-2004-11-16/pdf/04-25303.pdf>

ZEPP, R. G., M. CYTERSKI, R. PARMAR, K. WOLFE, E. M. WHITE et M. MOLINA. 2010. *Predictive Modeling at Beaches Volume II: Predictive Tools for Beach Notification*, Washington (D.C.), US EPA National Exposure Research Laboratory.

Fin du rapport