

Baignade à la plage de la baie de Beauport

Développement d'un modèle de prévision de la qualité de l'eau

La plage de la baie de Beauport, au cœur de la ville de Québec, a été fermée à la baignade pendant plus de 50 ans pour cause de dépassement des normes de qualité de l'eau. Durant ces 50 années, près d'un demi-milliard de dollars ont été investis dans les infrastructures de la ville pour améliorer la qualité des cours d'eau. À la suite d'analyses réalisées en 2014 et 2015, un modèle prédictif de la qualité de l'eau a été développé, ce qui a permis d'ouvrir la plage à la baignade 39 jours durant l'été 2016, sans compromettre la santé des baigneurs.

PAR CYRIL GARNEAU

Ph. D, chercheur postdoctoral, modelEAU, Université Laval
cyril.garneau.1@ulaval.ca

ET PAR PETER VANROLLEGHEM

ing., Ph. D., professeur, modelEAU, Université Laval
peter.vanrolleghem@gci.ulaval.ca

La qualité de l'eau de la baie de Beauport est influencée, d'une part, par les apports en contaminants (surverses, pollution agricole, goélands, accidents) et, d'autre part, par la capacité du fleuve à diluer et à transporter ces polluants. Le fleuve dans la région de Québec est soumis à une marée qui provoque d'importantes inversions de courant et des patrons d'écoulement complexes et très changeants.

Au Québec, la norme de baignade est de 200 UFC/100 ml (unité d'*Escherichia coli* formant colonie par 100 ml). Des campagnes de mesure à la plage de la baie de Beauport ont montré que cette norme a été respectée 65 % du temps en 2014 et 2015. Cependant, ces mesures sont ponctuelles dans le temps (une mesure par jour au mieux) et ne permettent pas de suivre l'évolution temporelle de la contamination. De plus, le résultat de chaque mesure n'est disponible que le lendemain. Elles ne peuvent donc pas être utilisées pour décider de l'ouverture ou de la fermeture de la plage.

Dans ce contexte, le désir de la Ville de Québec d'ouvrir la plage à la baignade et la nécessité de ne pas compromettre la santé publique lui ont fait envisager le développement d'un modèle de prévision de la qualité de l'eau, permettant de prendre une décision chaque matin sur la base d'informations disponibles avant l'ouverture de la plage.

Le développement de ce modèle a été réalisé en cinq étapes :

1. Analyse des modèles de qualité de l'eau existants;
2. Collecte et mise en forme des données disponibles;
3. Choix des variables explicatives pour alimenter le modèle;
4. Analyse de l'incertitude des données et du modèle;
5. Développement d'un outil d'aide à la décision.

Le modèle développé a ainsi été intégré au cœur d'un outil d'aide à la décision utilisé pour permettre ou interdire la baignade pendant l'été 2016; une décision était ainsi prise tous les jours, à 8 h. Afin d'évaluer la performance du système décisionnel, cette décision était suivie d'un échantillonnage de la plage selon le protocole du programme Environnement-Plage du ministère du Développement durable, de l'Environnement et de la Lutte contre les changements climatiques.

ANALYSE DES MODÈLES DE QUALITÉ DE L'EAU EXISTANTS

Plusieurs types de modèles peuvent être utilisés pour prédire la contamination d'une plage : des modèles statistiques, empiriques, hydrodynamiques, etc. De Brauwere *et al.* (2014) offrent une excellente revue de littérature sur les modèles existants. Le choix du modèle a été basé sur le constat suivant : la contamination vient essentiellement de surverses ponctuelles. Cependant, de nombreux ouvrages de surverse ne fournissent pas cette information en continu. De plus, le fleuve Saint-Laurent possède un régime hydrodynamique très complexe. Ainsi, son très grand débit et ses marées de plus de quatre mètres au Port de Québec rendent difficiles le suivi des masses d'eau qui atteignent la plage et la dilution à laquelle est sujette la contamination.

Dans ce contexte, il a été déterminé que l'approche statistique offrait le meilleur potentiel de modélisation. De plus, le format journalier des données de qualité de l'eau à la plage a permis de considérer chaque mesure comme indépendante des autres, puisque près de deux cycles de marée les séparent. Le choix final



s'est donc arrêté sur un modèle de régression linéaire multiple :

$$Y = \beta_0 + \sum \beta_i X_i$$

où Y est la variable dépendante, la concentration en *E. coli*, β_0 est une constante, X_i sont les variables explicatives du modèle et β_i sont les paramètres du modèle. Les relations non linéaires peuvent éventuellement être exprimées en transformant les variables. Par exemple, la concentration d'*E. coli* varie exponentiellement avec la majorité des variables. La variable dépendante Y n'est donc pas la concentration d'*E. coli*, mais le logarithme en base 10 de cette concentration.

COLLECTE ET MISE EN FORME DES DONNÉES DISPONIBLES

La Ville de Québec mène des campagnes d'échantillonnage sur les principales plages du fleuve Saint-Laurent depuis plus de dix ans. Jusqu'à l'été 2015, de 3 à 5 mesures étaient réalisées chaque semaine selon le protocole du programme Environnement-Plage. Ce protocole, qui a évolué avec les années, consiste au prélèvement de plusieurs échantillons qui étaient analysés individuellement (jusqu'à 14 échantillons) ou sous forme de composites (jusqu'à 4 échantillons composites). La mesure pour une journée donnée consistait ensuite à faire la moyenne des résultats pour l'ensemble des échantillons.

Les variables disponibles pour expliquer la concentration d'indicateurs fécaux incluent les nombreux pluviomètres présents sur le territoire de la ville, les ouvrages en surverse du réseau unitaire ainsi que le temps de surverse (en minutes par jour), la prévision météorologique passée et présente, les cycles de marée et les observations réalisées sur la plage par le responsable de l'échantillonnage. Au total, quelques centaines de variables explicatives étaient disponibles pour expliquer la contamination.

La mise en forme des données a consisté à ramener toutes ces variables sur une base comparable. Ainsi, les variables ont été mises sur une base journalière, en accord avec les séries de données de contamination. Cette

étape a également servi à agréger différentes variables en une seule, par exemple en calculant la précipitation maximale sur des sous-ensembles de pluviomètres ou en calculant les surverses totales sur un secteur donné.

CHOIX DES VARIABLES EXPLICATIVES

Les variables explicatives sont les variables permettant de prédire la contamination en *E. coli*. Dans le cadre de la plage de la baie de Beauport, plusieurs centaines de combinaisons de variables ont été testées pour reproduire les concentrations d'*E. coli* observées. Le choix final s'est porté sur quatre variables (tableau 1).

Ces variables, en plus d'être fortement corrélées au logarithme de la mesure d'*E. coli*, peuvent être associées à des processus connus pour générer ou réduire la contamination. Leur influence potentielle peut être estimée, puisque le modèle de régression multilinéaire additionne les contributions des différentes variables au logarithme de la mesure d'*E. coli*, ce qui revient à multiplier cette mesure par un facteur équivalent.

La pluviométrie est reconnue comme étant le principal facteur de contamination, puisqu'une pluie trop intense surcharge le réseau unitaire et mène à des surverses dans les rivières et le fleuve.

La marée, quant à elle, vient inverser l'écoulement du fleuve. Ainsi, à marée descendante, la plage est plus sensible aux sources de pollution en amont, tandis qu'à marée montante, les sources de pollution en aval ont un impact supérieur. De plus, la marée haute entraîne une importante masse

TABLEAU 1

Variables indépendantes du modèle de prévision de la qualité de l'eau et l'influence potentielle entre leur valeur la plus favorable et la moins favorable.

VARIABLES INDÉPENDANTES	INFLUENCE POTENTIELLE
Pluviométrie des 24 dernières heures	4x
Hauteur de marée	3x
Goélands et grands oiseaux	2x
Direction du vent	1.5x

Baignade à la plage de la baie de Beauport

Développement d'un modèle de prévision de la qualité de l'eau



d'eau sur la plage possédant un très grand pouvoir de dilution de la pollution présente. La hauteur de marée constitue la seule variable à varier inversement avec la concentration d'*E. coli*, une augmentation du niveau menant à une diminution de la contamination.

Les oiseaux présents sur la plage constituent une source de contamination directe. Ainsi, un seul goéland peut rejeter jusqu'à 10^6 *E. coli* par jour (Alderisio et DeLuca, 1999). L'impact d'un seul oiseau sur la plage est négligeable. Cependant, lorsqu'une colonie de plusieurs dizaines d'oiseaux

s'installe, la contamination moyenne d'*E. coli* peut augmenter rapidement.

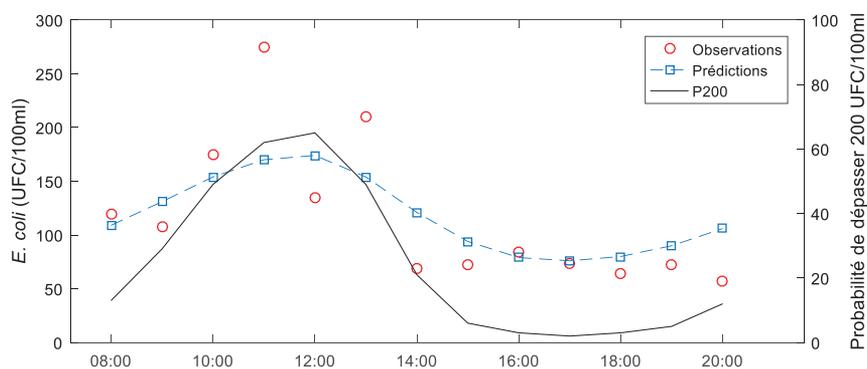
Enfin, le vent dominant sur la ville de Québec vient du sud-ouest. Cependant, un lien statistique significatif lie une contamination accrue à un vent du nord. Les raisons de cette contamination peuvent être multiples : courants de surface entraînant des contaminants de la côte de Beauport, événements météo plus violents, remise en suspension du sédiment sur la plage, etc.

L'influence potentielle des variables affichées au tableau 1 (p. 55) montre que la pluviométrie est sans conteste le principal contributeur à la contamination. Par contre, la marée, revenant deux fois par jour, est un facteur beaucoup plus présent. Les grands oiseaux peuvent influencer significativement la qualité de l'eau, mais il faut plus de 100 oiseaux sur la plage pour doubler la concentration d'*E. coli*. Dans ces conditions, le gestionnaire de la plage a décidé d'effaroucher les oiseaux en tout temps pour améliorer la qualité de l'eau. Enfin, la direction du vent a une influence limitée sur la prédiction de la qualité de l'eau. Son influence se fait donc sentir en conjonction avec une ou plusieurs variables additionnelles.

Durant le développement du modèle, l'effet de la marée s'est révélé beaucoup plus important que prévu initialement. Dans le but de valider son influence, une campagne de mesure spécifique a été menée pour prélever et analyser des échantillons toutes les heures, entre 8 h et 20 h. Les résultats de cette campagne sont présentés à la figure 1. On peut conclure que lors de cette campagne, le modèle est parvenu à prédire avec succès la période de vulnérabilité observée à marée basse (entre 10 h et 13 h) et la bonne qualité de l'eau à marée haute.

FIGURE 1

Résultats de la campagne d'échantillonnage du 16 juin 2016 (cercles rouges) et validation du modèle prévisionnel de qualité de l'eau (carrés bleus). La courbe noire indique la probabilité de dépasser la norme de qualité de l'eau de 200 UFC/100 ml au cours de la journée.



ANALYSE DE L'INCERTITUDE DES DONNÉES ET DU MODÈLE

Il est facile d'estimer les valeurs optimales des paramètres par la méthode des moindres carrés et, une fois ces paramètres calibrés, de prédire la concentration d'*E. coli* pour une valeur donnée des variables indépendantes. Une telle prédiction déterministe ne donne cependant pas d'information sur la probabilité de se tromper. Pour

FIGURE 2

Pseudocode de la technique de rééchantillonnage utilisée. La variable MC correspond au nombre de rééchantillonnages de Monte-Carlo, Y_{obs} correspond aux observations rééchantillonnées et Y_{pred} aux prédictions réalisées à l'aide des techniques de Monte-Carlo et du jackknife.

```
Pour i = 1 à MC
  Rééchantillonner  $Y_{obs}$  selon loi log-normale( $\mu_{obs}$ ,  $ET$ )
  Pour j = 1 à n
    Pour k = j à n
      Retirer les observations j et k
      Calculer  $\beta_{i,j,k}$  sur les observations restantes
      Calculer  $Y_{pred,i,j}$  et  $Y_{pred,i,k}$  à l'aide de  $\beta_{i,j,k}$ 
    Fin boucle k
  Fin boucle j
Fin boucle i
Calculer  $\mu_{pred}$  et  $\sigma_{pred}$  de chaque prédiction  $Y_{pred}$ 
```

pour pallier ce manque, deux méthodes statistiques ont été utilisées pour estimer l'incertitude du modèle due à l'incertitude sur les mesures et l'incertitude due à la structure du modèle.

Puisque chaque mesure d'*E. coli* est en fait la moyenne de 4 à 14 échantillons individuels, il est possible d'évaluer l'erreur-type sur la moyenne par l'équation suivante :

$$ET = \frac{\sigma}{\sqrt{n}}$$

où σ est l'écart-type des valeurs mesurées pour les échantillons individuels et n est le nombre d'échantillons.

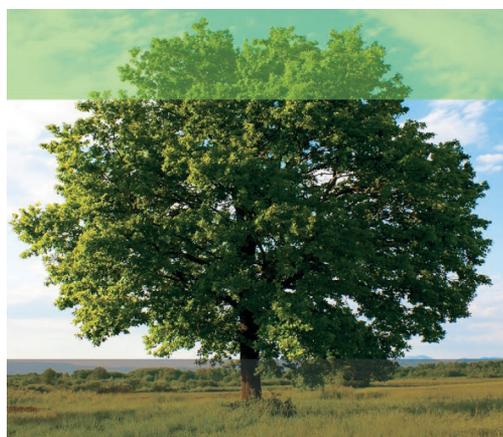
Afin d'évaluer l'incertitude sur la précision du modèle, une méthode de Monte-Carlo a été utilisée pour rééchantillonner les observations. Toutes les mesures étaient rééchantillonnées selon une loi log-normale de moyenne μ et écart-type ET .

Simultanément, l'incertitude due à la structure du modèle de régression linéaire multiple a été évaluée par la méthode du jackknife. Cette méthode consiste à retirer une ou deux

observations du jeu de données, recalculer les paramètres β_i sur toutes les autres données et valider les paramètres calculés sur les observations retirées. Ainsi, puisque ces observations sont totalement indépendantes des paramètres, il est possible de valider que ces paramètres permettront effectivement de faire des prédictions robustes dans le futur. L'algorithme est résumé à la figure 2 sous forme de pseudocode.

Sur la base de ces deux techniques, 200 rééchantillonnages des 101 observations ont été réalisés. Ainsi, plus d'un million de jeux de paramètres ont été calculés pour évaluer la capacité du modèle à reproduire toutes les conditions observées et conserver un caractère prévisible. La figure 3 (p. 58) compare les résultats de simulations (incertaines) avec les observations (incertaines) utilisées pour la construction du modèle. Le point noir représente la moyenne observée (en abscisse) et prédite (en ordonnée), et l'ellipse montre l'incertitude sur la mesure et sur la prédiction.

Ces résultats montrent que malgré la grande incertitude sur les mesures, les prédictions



VALORISEZ vos actions environnementales

ORGANISATIONS et ÉVÈNEMENTS

- ✓ Certification reconnue internationalement
- ✓ Services supplémentaires :
 - Réalisation d'inventaires de gaz à effet de serre
 - Production de déclarations GES
 - Rédaction de politiques environnementales



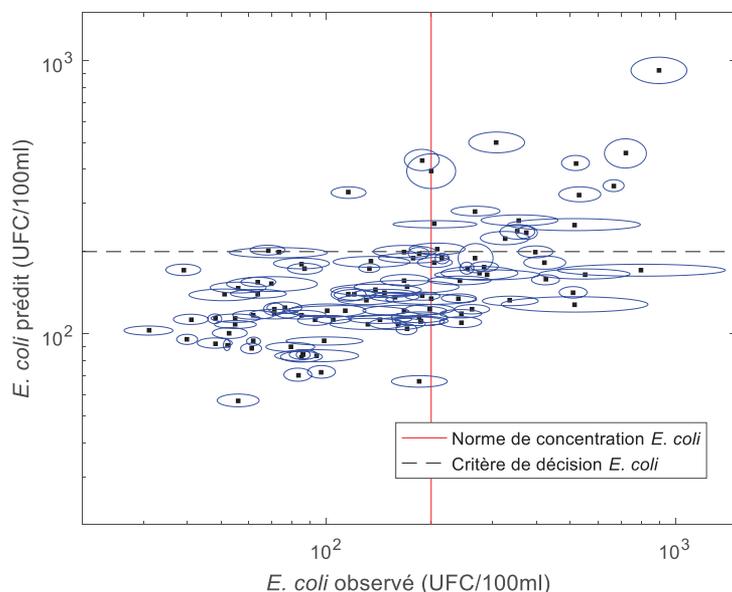
ECOCERT Canada – 510, rue Saint-Onésime, Lévis, Québec, Canada, G6V 5Z4
1-855-246-9383 – env.canada@ecocert.com – www.ecocertcanada.com

Baignade à la plage de la baie de Beauport

Développement d'un modèle de prévision de la qualité de l'eau

FIGURE 3

Valeurs observées et prédites d'*E. coli*. L'ellipse correspond à un écart-type d'incertitude sur l'observation (horizontal) et sur le modèle (vertical) selon une loi statistique log-normale. Le critère de décision a été fixé à 200 UFC/100 ml de façon arbitraire. Sa valeur doit être choisie par le gestionnaire, selon la tolérance au risque de dépassement de ce dernier.



restent peu dispersées, ce qui suggère que les variables utilisées par le modèle possèdent un bon pouvoir prédictif.

DÉVELOPPEMENT D'UN OUTIL D'AIDE À LA DÉCISION

Le modèle est capable de reproduire les observations passées raisonnablement. Un outil d'aide à la décision a donc été développé autour de ce modèle pour récupérer chaque jour les données requises, les fournir au modèle et calculer une prédiction de la qualité de l'eau et de son incertitude pour la journée. Le modèle a été compilé à partir du code source pour deux plateformes. La plateforme automatisée prend un fichier texte en entrée dans lequel les variables sont inscrites, et retourne en sortie un second fichier texte contenant les prévisions pour la journée. Cette plateforme est indépendante du programme qui l'exécute et peut donc être déployée très rapidement. La

version manuelle a été intégrée à un code Visual Basic for Applications (VBA) permettant d'entrer manuellement dans un tableur Excel les variables, de lancer la prédiction et de visualiser les résultats rapidement. Cet outil offre l'avantage de pouvoir lancer le modèle sous différentes conditions (par exemple avec des pluviométries différentes) pour bien visualiser l'effet de chaque variable.

Enfin, l'outil ne peut tenir compte des déversements en temps sec (causés par un bris d'équipement par exemple). Ces déversements, lorsqu'ils ont été observés, ont mené à des fermetures préventives.

RETOUR SUR UNE SAISON DE BAINNADE

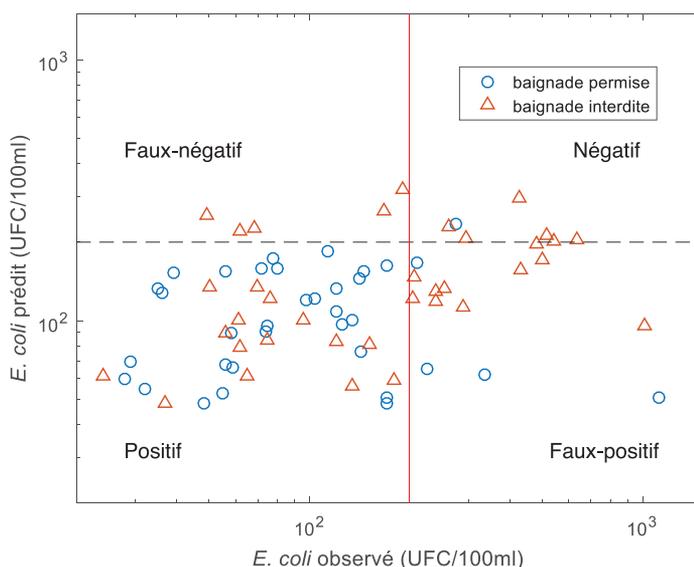
Le modèle prévisionnel a été lancé le 24 juin 2016 et a été utilisé tous les jours jusqu'au 9 septembre. Le protocole a consisté à calculer une prédiction avant 8 h chaque matin, à décider de l'ouverture ou de la fermeture de la plage, puis à prendre un échantillon d'eau à analyser au laboratoire d'environnement de la Ville de Québec. Les résultats d'analyse, connus le lendemain, venaient confirmer ou infirmer la prédiction. Ainsi, chaque jour de baignade a été validé par une ou plusieurs observations d'*E. coli*.

La figure 4 montre, pour chaque jour, la prédiction du modèle en fonction de l'observation réalisée au même moment. Les ronds bleus correspondent aux jours où la baignade était permise et les triangles rouges aux jours où elle était interdite. Les quatre quadrants délimités par la norme réglementaire et le critère de décision permettent d'identifier les zones suivantes : « positif » (le modèle et l'observation respectent la norme), « négatif » (le modèle et l'observation ne respectent pas la norme), « faux-positif » (le modèle respecte la norme, mais pas l'observation) et « faux-négatif » (le modèle ne respecte pas la norme tandis que l'observation la respecte). Le critère de décision a été fixé à 200 UFC/100 ml arbitrairement pour illustration seulement.

Au total, la plage a été ouverte à la baignade 39 jours et fermée 36 jours. Sur les 39 jours ouverts, seules cinq observations ont démontré une concentration en *E. coli* supérieure à la norme (faux-positif). Les raisons expliquant ces déviations entre les résultats du modèle et les

En ne suivant que quatre variables, soit la pluviométrie des 24 dernières heures, la marée, le nombre d'oiseaux présents sur la plage et la direction du vent, il a été possible d'ouvrir la plage à la baignade 39 jours pour la première fois depuis 50 ans.

FIGURE 4
Valeurs prédites et observées d'*E. coli* pour chaque jour de l'été 2016.



observations incluent la grande variabilité de la mesure d'*E. coli* et le fait que les prélèvements étaient réalisés à 8 h, tandis que la baignade était autorisée de 10 h à 18 h, ce que le modèle était capable de prédire. Une observation au-dessus de la norme à 8 h peut donc correspondre à une permission de baignade (voir figure 1, p. 56). Le même effet de la marée est responsable des fermetures préventives observées lorsque la prédiction et l'observation sont sous la limite des 200 UFC/100 ml, mais qu'une fermeture préventive a été ordonnée (faux-négatif).

CONCLUSION

Le développement d'un modèle statistique de prévision de la qualité de l'eau a été rendu possible par l'existence d'une base de données fiable, extensive et à jour. C'est en s'appuyant sur cette base de données que chaque variable explicative a pu être analysée et son influence quantifiée. Ainsi, en ne suivant que quatre variables, soit la pluviométrie des 24 dernières heures, la marée, le nombre d'oiseaux présents sur la plage et la direction du vent, il a été possible d'ouvrir la plage à la baignade 39 jours pour la première fois depuis 50 ans. Ces travaux

démontrent qu'il est possible d'exploiter de nouveaux usages des cours d'eau, lorsque les conditions sont favorables, et de protéger le public lorsque nécessaire.

REMERCIEMENTS

Nous voulons remercier le Service de l'environnement de la Ville de Québec, spécialement madame Sylvie Verreault et monsieur François Proulx. Nous remercions également les professeurs Sophie Duchesne et Yves Secrétan, ainsi que madame Marie-Ève Jean de l'Institut national de la recherche scientifique. ■

RÉFÉRENCES

Alderisio, K.A., DeLuca, N. (1999). « Seasonal enumeration of fecal coliform bacteria from the feces of ring-billed gulls (*Larus delawarensis*) and Canada geese (*Branta canadensis*) ». *Appl. Environ. Microbiol.*, vol. 65, n° 12, p. 5628-5630.

De Brauwere, A., Ouattara, N.K., Servais, P. (2014). « Modeling fecal indicator bacteria concentrations in natural surface waters: A review ». *Crit. Rev. Environ. Sci. Technol.*, vol. 44, n° 21, p. 2380-2453 (doi : 10.1080/10643389.2013.829978).

Assistez à la conférence
de Cyril Garneau sur le sujet
à AMERICANA!