L'apprentissage machine: un outil pour une gestion plus efficace des eaux récréatives

Émile Sylvestre[1,2], Joseph De Soissan[1], Naysan Saran[1], Nicolas Fortin St-Gelais[1,3]

[1] CANN Forecast
[2] Polytechnique Montréal
[3] Université du Québec à Montréal

23 novembre 2016
Congrès INFRA 2016
Palais des congrès de Montréal

Projet de recherche multidisciplinaire Safe Ruhr (2012-2015)
www.sciencedirect.com/science/journal/14384639/219/7/supp/PB

Plan

- Défis en gestion des eaux récréatives
- Utiliser l’apprentissage machine pour mieux gérer les risques
- Étude de cas en région de Montréal

Risque microbiologique en eaux récréatives

- Les pathogènes sont la principale cause des maladies et des problèmes de santé hydrique en eaux récréatives.

- Profil de risque : fluctuations à court terme (évènements intenses)

- Effets aigus sur la santé:
  - Gastro-entérites
  - Infections (yeux, nez, oreilles)
  - Affection de la peau

Limites du programme environnement-plage

Classification et qualité des eaux de baignade		Moyenne arithmétique en UFC/100 ml	
		Plages en milieu d'eau douce	Plage en milieu marin*
Cote	Qualité	*Escherichia coli*	Entérocoques
A	Excellente	≤ à 20	≤ à 5
B	Bonne	de 21 à 100	de 6 à 20
C	Passable	de 101 à 200	de 21 à 35
D	Polluée	≥ à 201	≥ à 36

http://www4.gouv.qc.ca/fr/Portail/citoyens/programme-service/Pages/Info.aspx?sqctype=sujet&sqcid=622

Échantillonage prévu pour l'été prochain :
Cote A : 2 échantillons
Cote B : 3 échantillons
Cote C : 5 échantillons

Logique à première vue, mais qu'un point de départ en pratique.

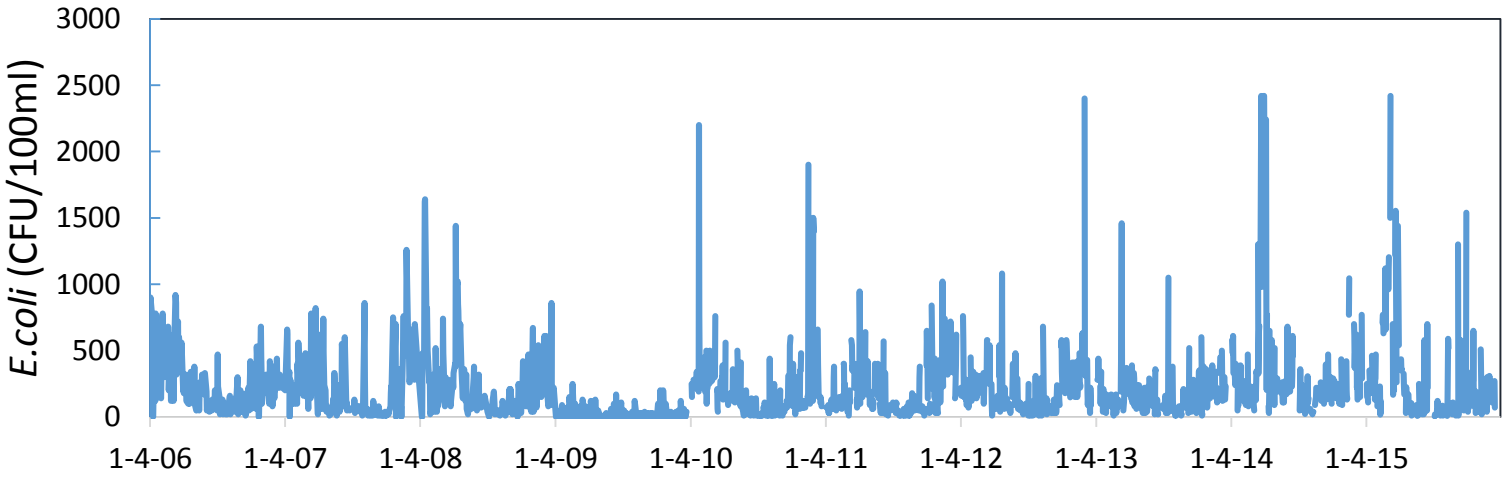
Inapplicabilité en milieu urbain

L'échantillonage routinier est basé sur une hypothèse de stationnarité à la source d'eau.

Cette hypotèse statistique n'est pas valide en milieu urbain en raison de l'effet combiné:

- de l'augmentation des surfaces de ruisselement sur le territoire
- des effets des changements climatiques

Milly, P. C. D. et al. (2007)



Suivi long terme de la qualité à une prise d'eau potable sur rivière en milieu urbain en région de Montréal

Nouvelles lignes directrices fondées sur le risque microbien à court terme

Statu quo :

- Les fluctuations à court terme ne sont captées par la moyenne arithmétique.
- Le résultat des échantillonnages par culture prend minimum 24 heures.

Opportunités :

- Mettre en place une stratégie d'évaluation du risque adaptée aux conditions locales.
- Améliorer la communication du risque avec le public.

Stratégie de gestion du risque par étapes

1- Évaluation provisoire du statut du cours d'eau et identification des périodes critiques de contamination

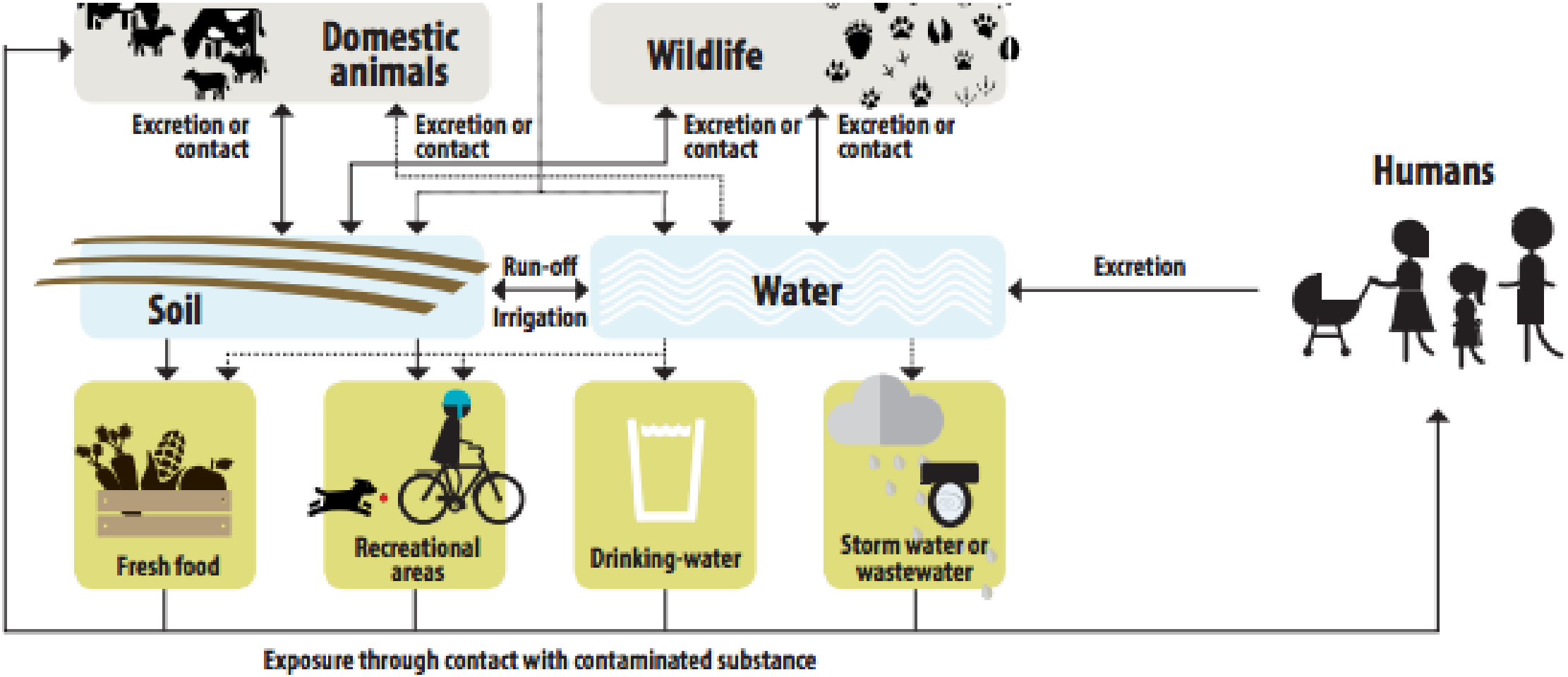
2- Caractérisation préliminaire des périodes de risque

- Échantillonnage des évènement critiques (indicateurs fécaux et pathogènes)

3- Caractérisation détaillée des périodes de risque (ampleur, durée, fréquence, prédicteurs/indicateurs, traceurs)

- Analyse quantitative du risque microbien - risque d'infection/personne/année (WHO, 2016)

1.1 Quels sont les principaux risques ?



WHO. (2016)

1.2 Identification des périodes critiques

Utilisation des données historiques

- Source de pollution
  - Échantillonage *E.coli*/coliformes fécaux
  - Durées/volumes des évènements de surverses

Municipalités

© VWM Solutions

- Transport et évolution dans l'environnement
  - Débit/niveau du cours d'eau récepteur
  - Pluviométrie
  - Température

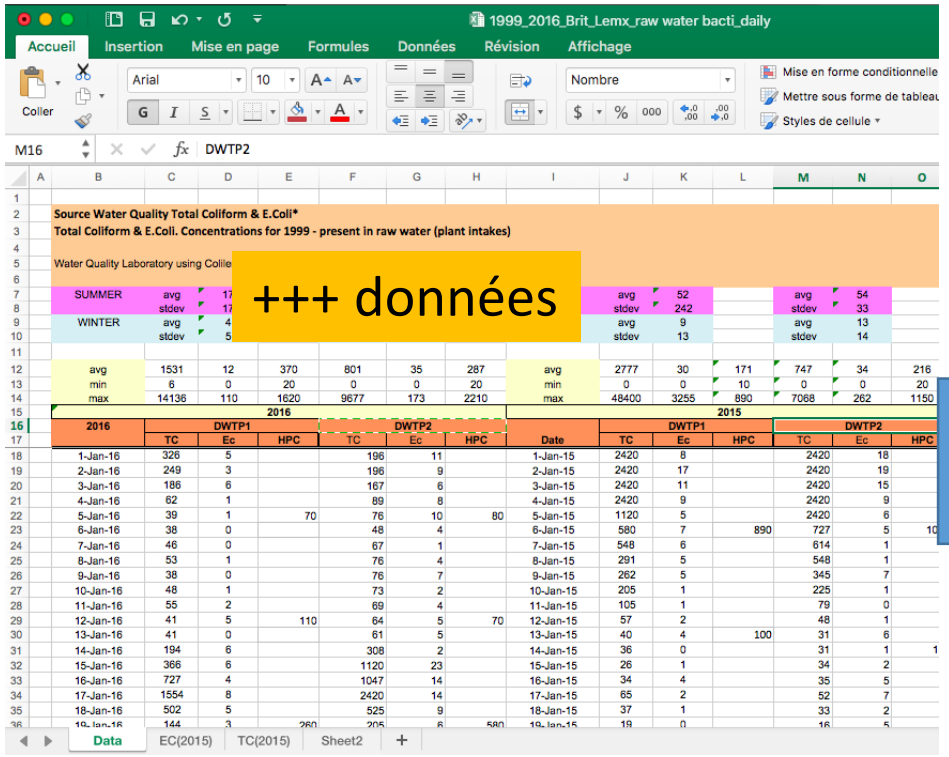
Environnement Canada (données ouvertes)

- Indices
  - Pluviométrie antécédente
  - Période sèche antécédante
  - First-flush

2.1 Construction de modèles prédictifs à optimisation continue

<u>Actuellement :</u>

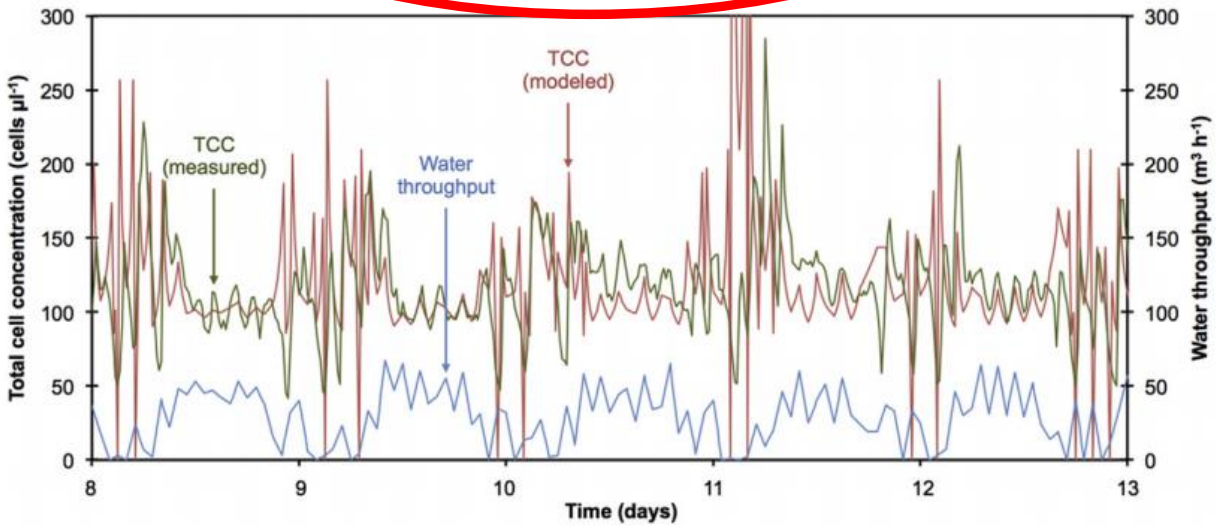
Par tendance centrale et percentiles (déterministe)



<u>À venir :</u>

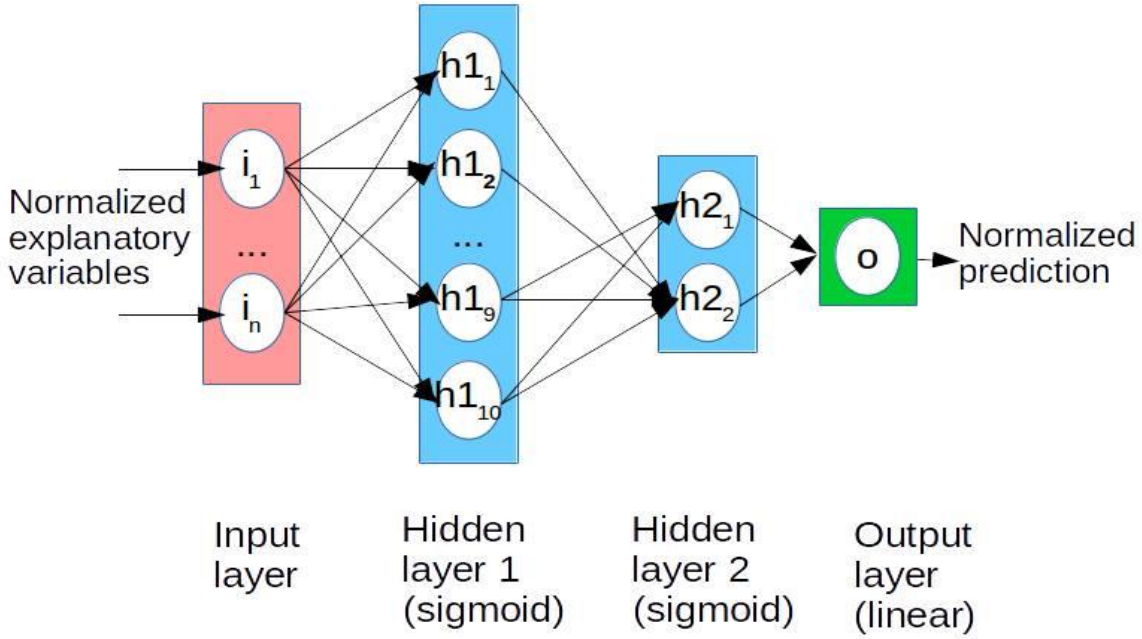
Modèle d'apprentissage en temps réel (stochastique)

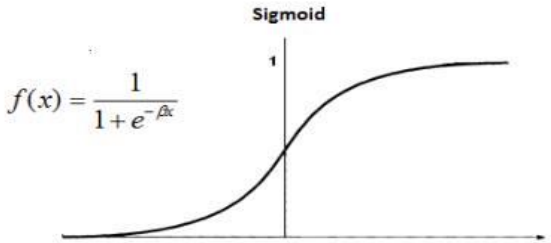
- Prévisions immédiates
- Amélioration continue

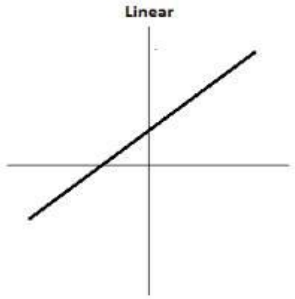


Messner, M. J. et al. (2016)

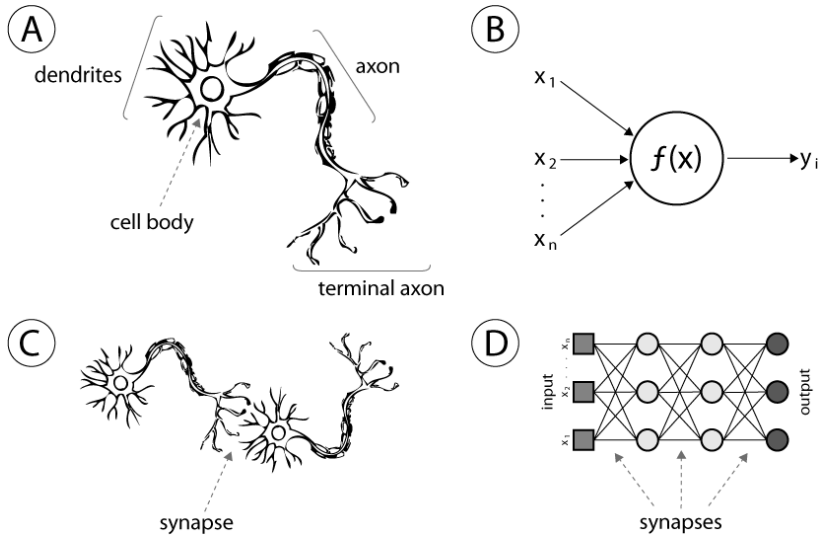
Structure d'un réseau neuronal artificiel (RNA)







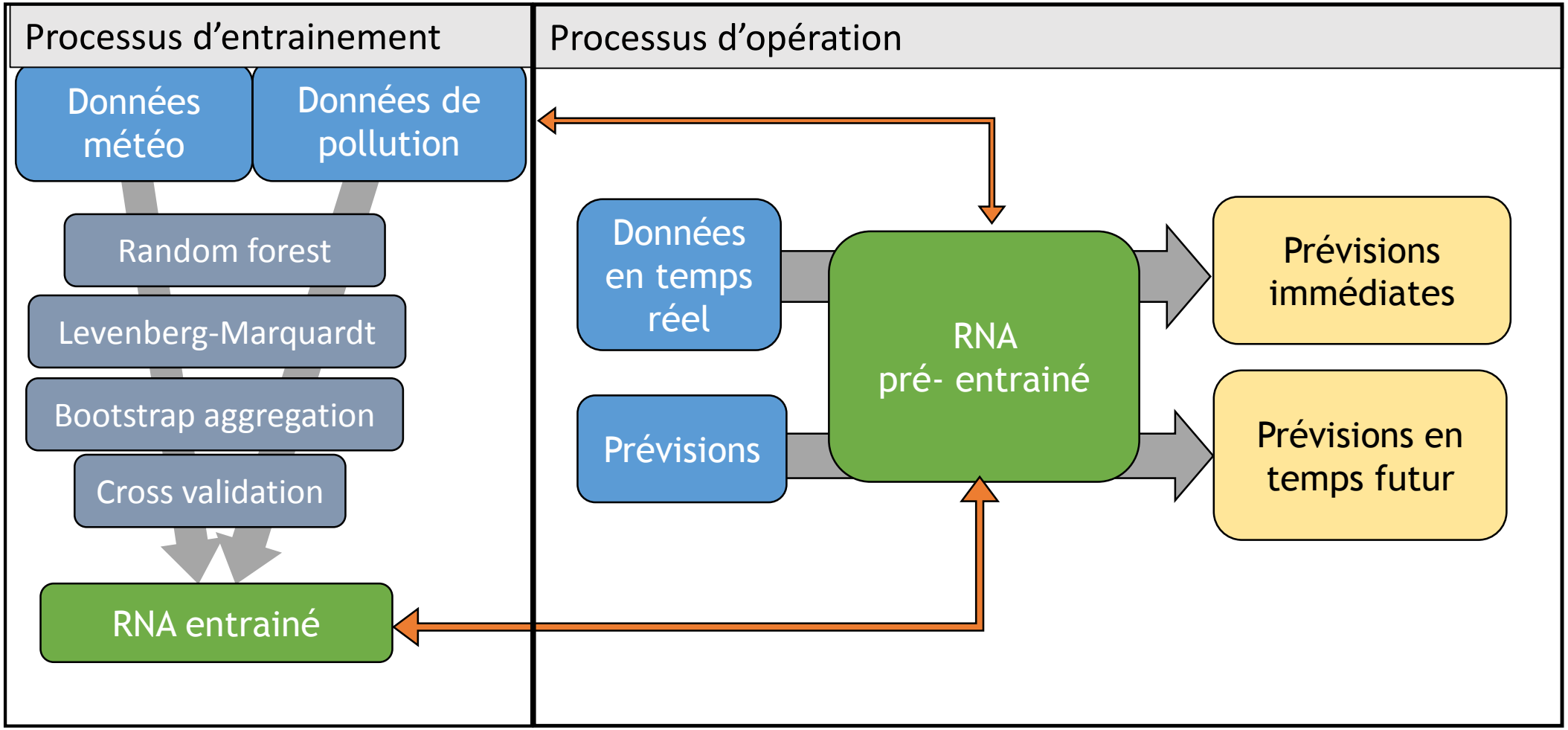
RNA et fonction d'activation de stimuli-response



A) Neurone humain B) Neurone artificiel
C) Synapse biologique D) Synapse d'un RNA

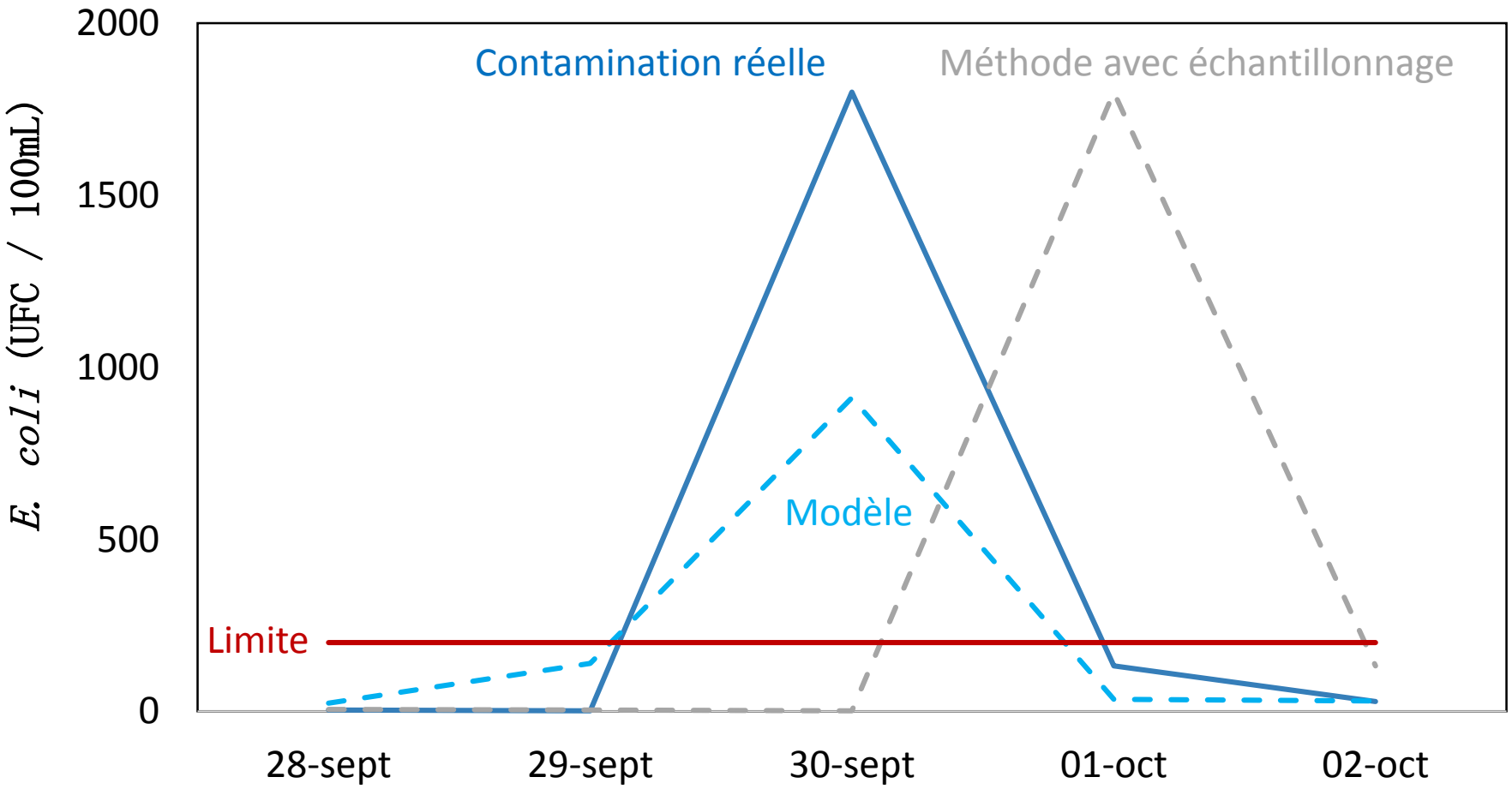
From: http://www.intechopen.com/books/artificial-neural-networks-architectures-and-applications/applications-of-artificial-neural-networks-in-chemical-problems

Processus d'entrainement et d'opération



Permet de faire l'étude de sensibilité de multiples sources de pollution pour prioriser les actions correctives.

Améliorer la communication du risque



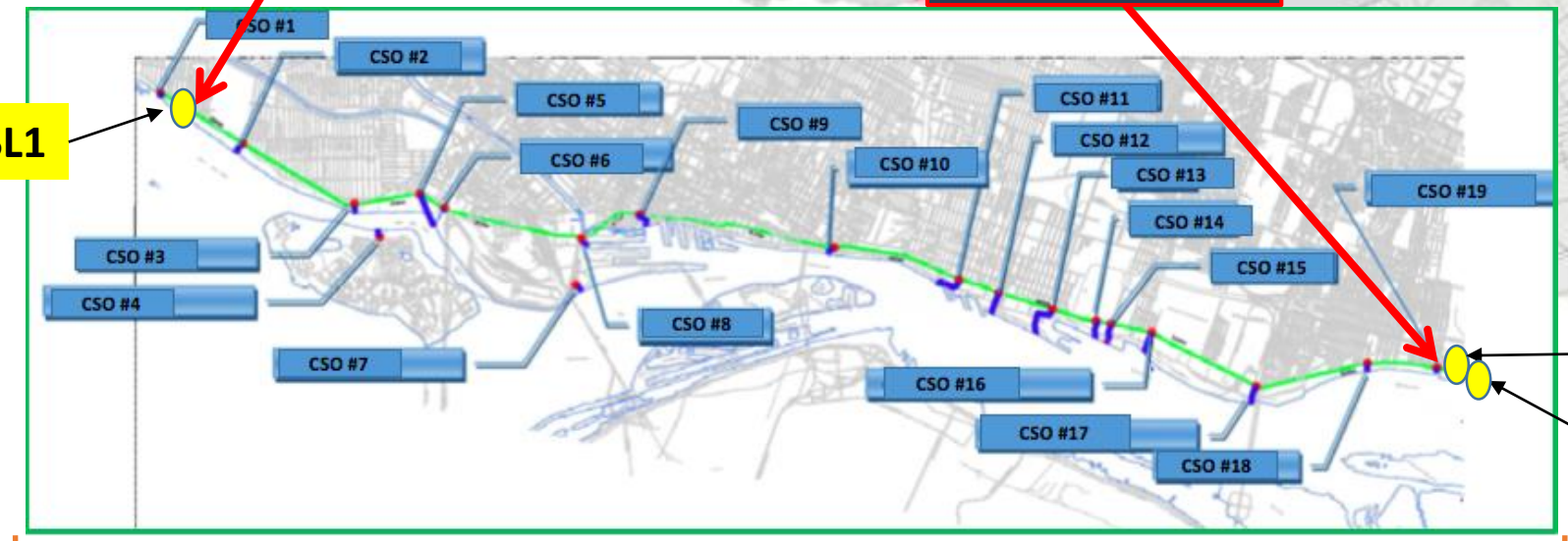
Plan

- Défis en gestion des eaux récréatives
- Utiliser l’apprentissage machine pour mieux gérer les risques
- Étude de cas en région de Montréal

Étude de cas en région de Montréal

Région 1

Région 2

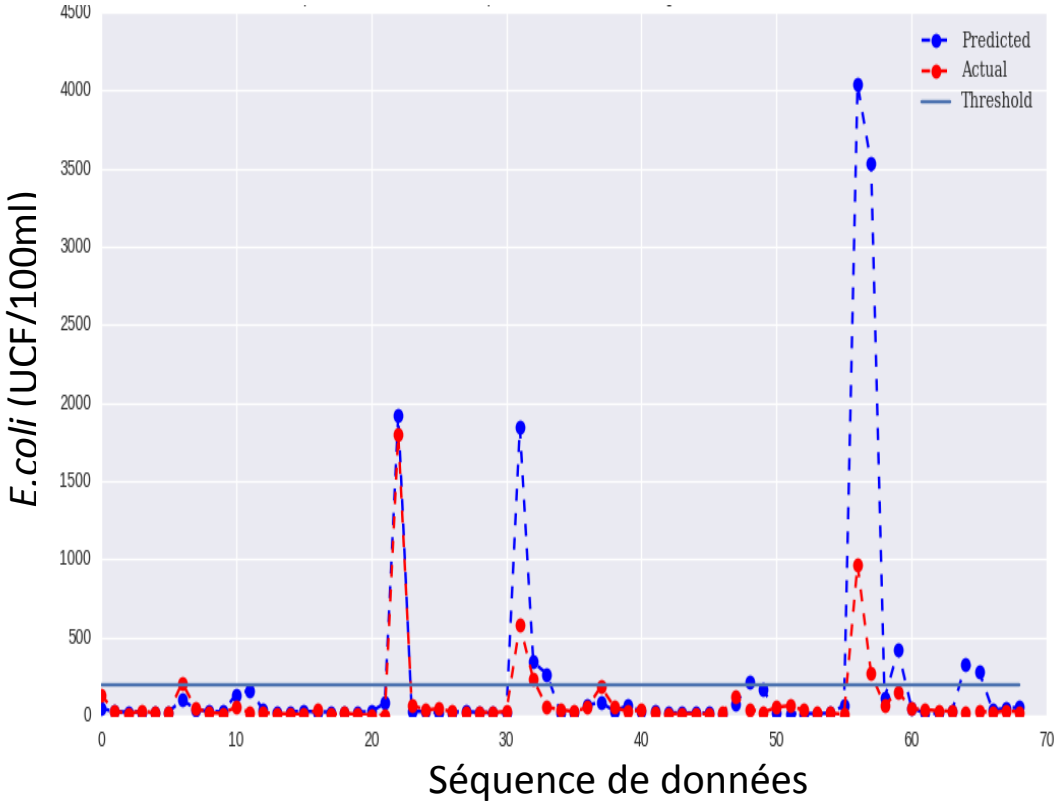


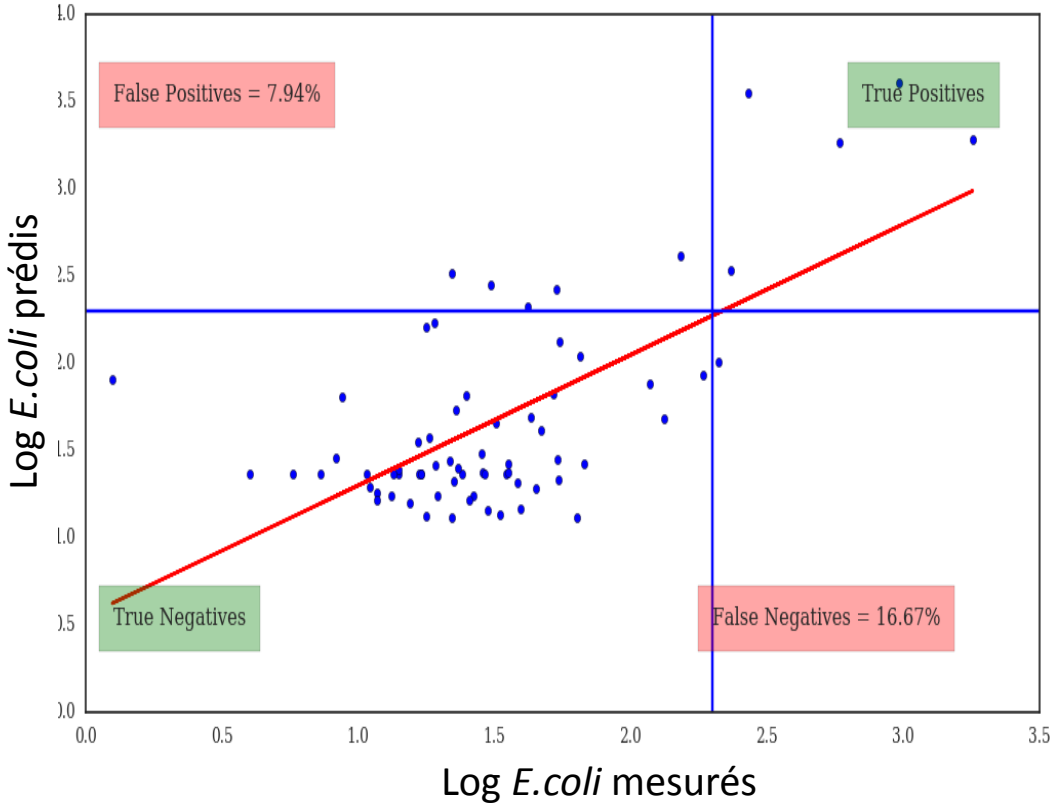
Lieux d'échantillonnage

19 ouvrages de surverse

Résultats : Test 1

Région 1 pour 2015:
- Échantillonnage CF journalier à SL1 (n=85)
- Durée cumulative (2 jours) des rejets au CSO #1





L'échantillonnage haute-fréquence (journalier) sur une saison fournie de l'information valable pour une modélisation RNA préléminaire.

Temps de rétablissement

Région 1

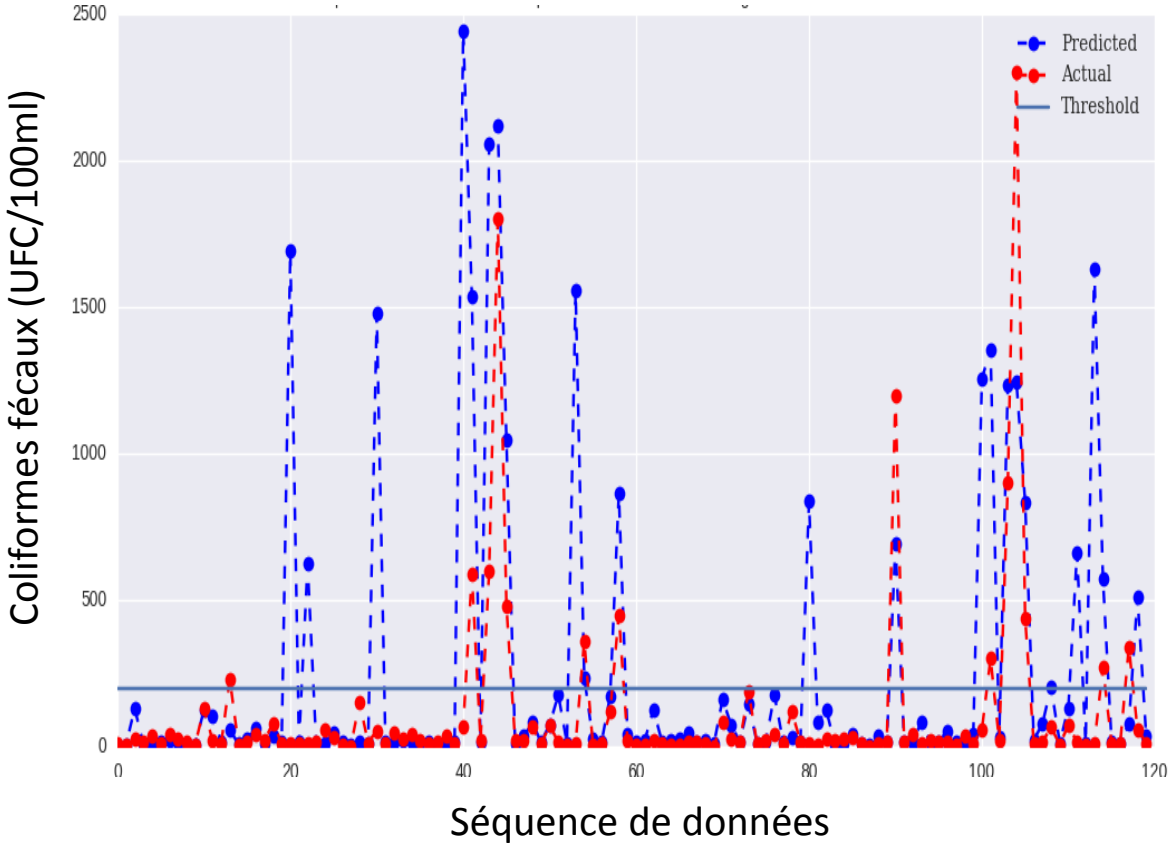
Conditions	Moyenne *E.coli* (UFC/100ml)	
Sans surverses (baseline)	36	
Jour d'évènement (période critique)	1117	> 200 (critère enviro-plage)
Jour 1 suivant l'évènement (période critique)	211	
Jour 2 suivant l'évènement (baseline)	49	

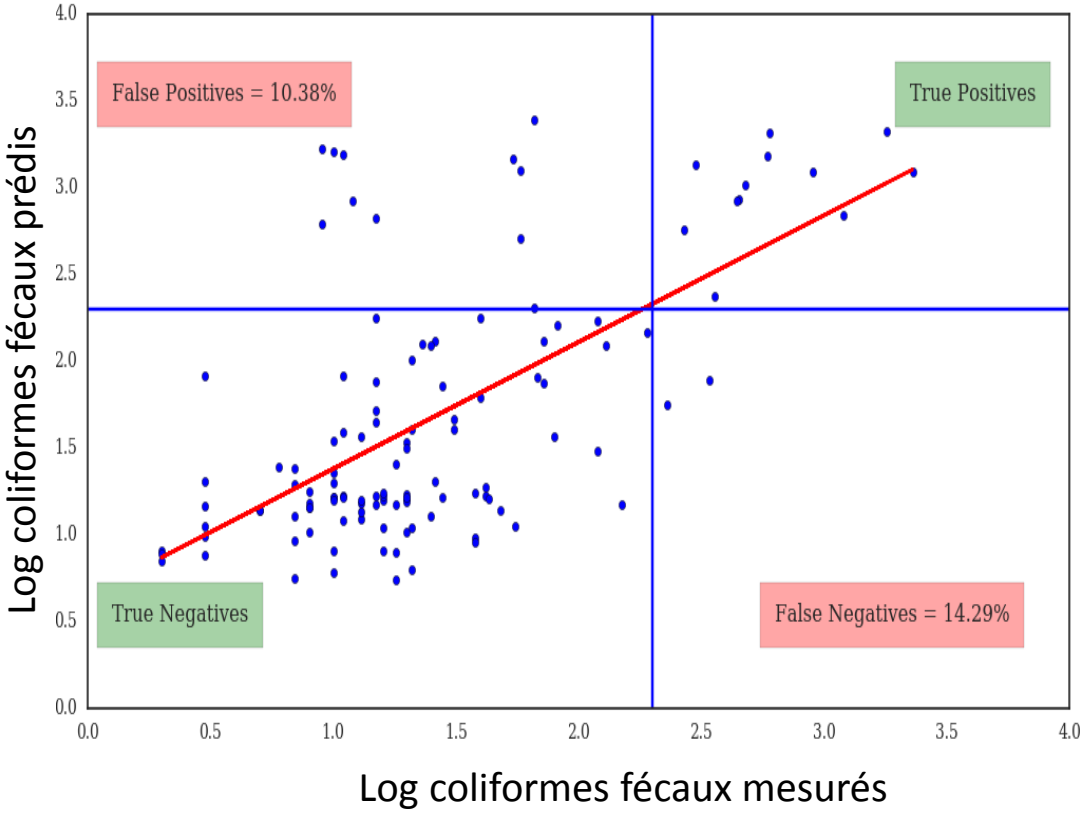
Taux de fermeture :

- 7% du temps (juin à septembre)
- Légère surestimation potentielle (+20% de pluie que moyenne sur période 1981-2010)

Résultats : Test 2

Région 2 pour 2013 à 2015 :
- Échantillonnage CF hebdomadaire à SL2 et SL3
- Pluviométrie





Conclusions

- Le programme environnement-plage doit être amélioré afin de répondre aux contraintes actuelles en milieu urbain.
- Environnement Canada et les municipalités possèdent une quantité importante de données pour mettre en place des modèles prédictifs. **(étape 1)**
- L'échantillonnage de la contamination en période critique (e.g. forte pluie) améliore la performance des modèles prédictifs. **(étape 2)**
- La modélisation par réseau neuronal artificiel améliore considérablement le potentiel de communication du risque pour la santé publique et la mise en place de mesures de protection de la source d'eau.
- L'indicateur *E.coli* est utile pour comprendre la dynamique microbienne à la source d'eau, mais l'échantillonage de pathogènes et requis pour faire une évaluation quantitative du risque microbien. **(étape 2 et 3)**

Remerciements

- Fondation de Gaspé-Beaubien
- Sarah Dorner (Polytechnique Montréal)
- Michèle Prévost (Polytechnique Montréal)
- Diane Mas (Fuss & O'Neill)
- Alain Charron (Ville de Montréal)
- Bertrand Denis (Environnement Canada)

Pour plus d'information :
facebook.com/infobaignademtl/
emile.sylvestre@cannforecast.com

Références

Besmer, M. D., & Hammes, F. (2016). Short-term microbial dynamics in a drinking water plant treating groundwater with occasional high microbial loads.*Water Research*, *107*, 11-18.

Milly, P. C. D., Julio, B., Malin, F., Robert, M., Zbigniew, W., Dennis, P., & Ronald, J. (2007). Stationarity is dead. *Ground Water News & Views*, *4*(1), 6-8.

Raschka, S. (2015). *Python Machine Learning*. Packt Publishing Ltd.

Roser, D. J., Davies, C. M., Ashbolt, N. J., & Morison, P. (2006). Microbial exposure assessment of an urban recreational lake: a case study of the application of new risk-based guidelines. *Water Science and Technology*, *54*(3), 245-252.

World Health Organization. (2016). Protecting surface water for health. Identifying, assessing and managing drinking-water quality risks in surface-water catchments.

World Health Organization. (2016). Quantitative microbial risk assessment. Application for water safety management