



L'apprentissage automatique pour prédire la dengue en Colombie

KATE ZINSZER PhD MSc  
Professeure adjointe  
Département de médecine sociale et préventive  
Université de Montréal

JASP  
4 décembre, 2018

## Arbovirus

- ◆ Dengue, chikungunya et Zika en croissance

introduction

WWII      2013      2014

- ◆ *Aedes aegypti* et *Aedes albopictus* des moustiques opportunistes
- ◆ Infections graves et des séquelles chroniques
- ◆ Coûts directs et indirects sont substantiels

2

## Colombie



- ◆ L'un des pays les plus touchés par dengue, chikungunya, et Zika
- ◆ 49.5 million d'habitants dans diverses conditions socioéconomiques, géographiques et climatiques
- ◆ Maladies à déclaration obligatoire (SIVIGILA)

surveillance et réponse



cas de dengue sévère



interventions

3

## Prévision



- ◆ Mécanisme d'alerte pour les interventions, services de santé et gestion de la chaîne d'approvisionnement
- ◆ Méthodes différentes : déterministes, statistiques et d'apprentissage automatique
  - ◆ Capacité à inclure très grand nombre de prédicteurs
- AA
  - ◆ Robustesse face à la multicollinéarité
  - ◆ Capacité à analyser un grand volume de données
  - ◆ Gérer les données manquantes

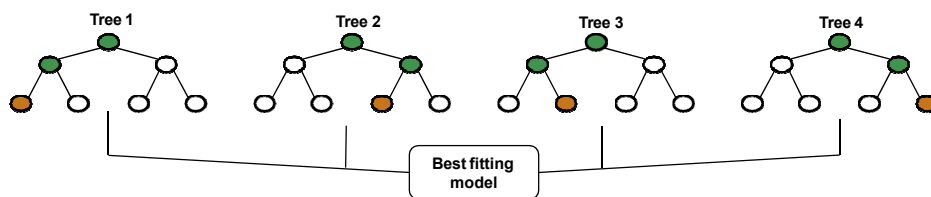
4

# Objectif

Développer et évaluer des modèles de prédiction individuels et conjoints pour la dengue, le chikungunya et le Zika

# Méthode

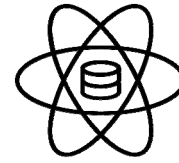
## Forêts aléatoires



- ♦ Échantillons bootstrap sont tirés et des arbres de décision sont construits
- ♦ Résultat final = moyenne sur tous les arbres de l'ensemble
- ♦ Identifie les prédicteurs en fonction de leur importance relative dans la prédiction

6

# Données



## ◆ SIVIGILA

◆ Cas suspects et confirmés de dengue, de chikungunya et de Zika

## ◆ Enquêtes nationales et recensement

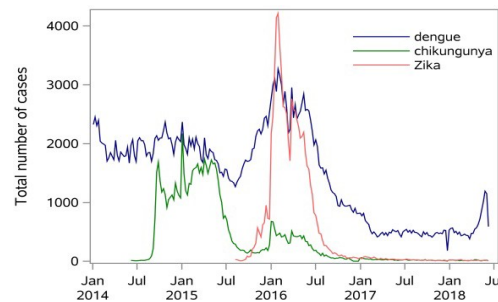
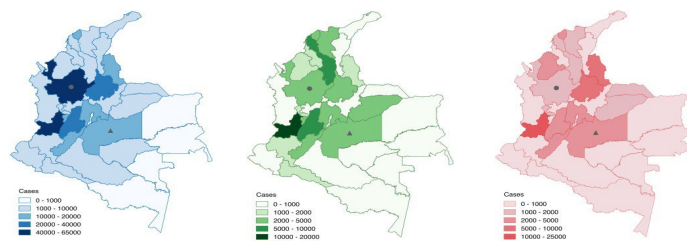
◆ Taille de la population, éducation, pauvreté, accès aux soins de santé

## ◆ Données de télédétection

◆ Précipitations, température, végétation, utilisation des sols (par exemple, urbanisation)

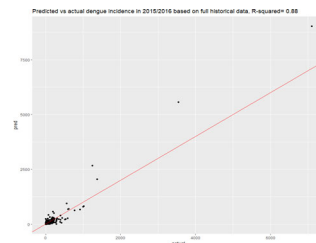
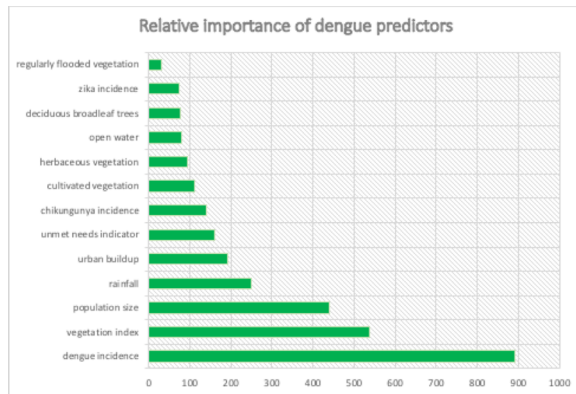
7

# Données descriptives



8

# Prévisions préliminaires



1 month-ahead

Historical chikungunya and dengue cases  
 $R^2 = 0.89$

2 month-ahead

Historical chikungunya and dengue cases  
 $R^2 = 0.70$

Year-ahead

Historical dengue and chikungunya cases  
 $R^2 = 0.66$   
Environment, land cover, SES, historical dengue and chikungunya  
 $R^2 = 0.88$

9

## Discussion

- ◆ Certains prédicteurs sont beaucoup plus importants que d'autres
- ◆ Les résultats sont prometteurs
- ◆ Besoin d'examiner les valeurs aberrantes
- ◆ Calculer d'autres indicateurs de précision
- ◆ Quel niveau de précision est nécessaire?

10

## Prochaines étapes

- ◆ Explorer l'échelle hebdomadaire à une résolution spatiale moins précise
- ◆ Construire des modèles pour le chikungunya et le Zika
- ◆ Comprendre comment construire des modèles conjoints
- ◆ Tester d'autres approches
  - ◆ Machines à vecteurs de support
  - ◆ Réseaux de neurones profonds

11

## Remerciements

Université de Montréal  
Katia Charland  
Audrey Smargiassi  
Marc-Antoine Tutt-Guërette



Universidad Cooperativa de Colombia  
Gloria Jaramillo Ramirez  
Cesar Balaguera



McGill University  
Naizhuo Zhao  
Mabel Carabali  
Mathieu Maheu-Giroux  
Alexandra Schmidt  
Erin Liu



Noun project - icons

12