

# Couper des arbres avec R

Combiner l'apprentissage automatique et les capacités géospatiales de R pour mesurer le risque d'inondation en assurance.

Par : Frédérick Guillot et Etienne Larrivée-Hardy

R à Québec, 26 mai 2017



# Qui sommes-nous?

## Frédéric Guillot

- M. Sc. en sciences actuarielles
- Directeur Recherche et Innovation chez Co-operators
- Pas un expert en géomatique / hydrologie / ...

## Etienne Larrivée-Hardy

- M. Sc. en sciences actuarielles
- Analyste Recherche et Innovation chez Co-operators
- Pas un expert en géomatique / hydrologie / ...

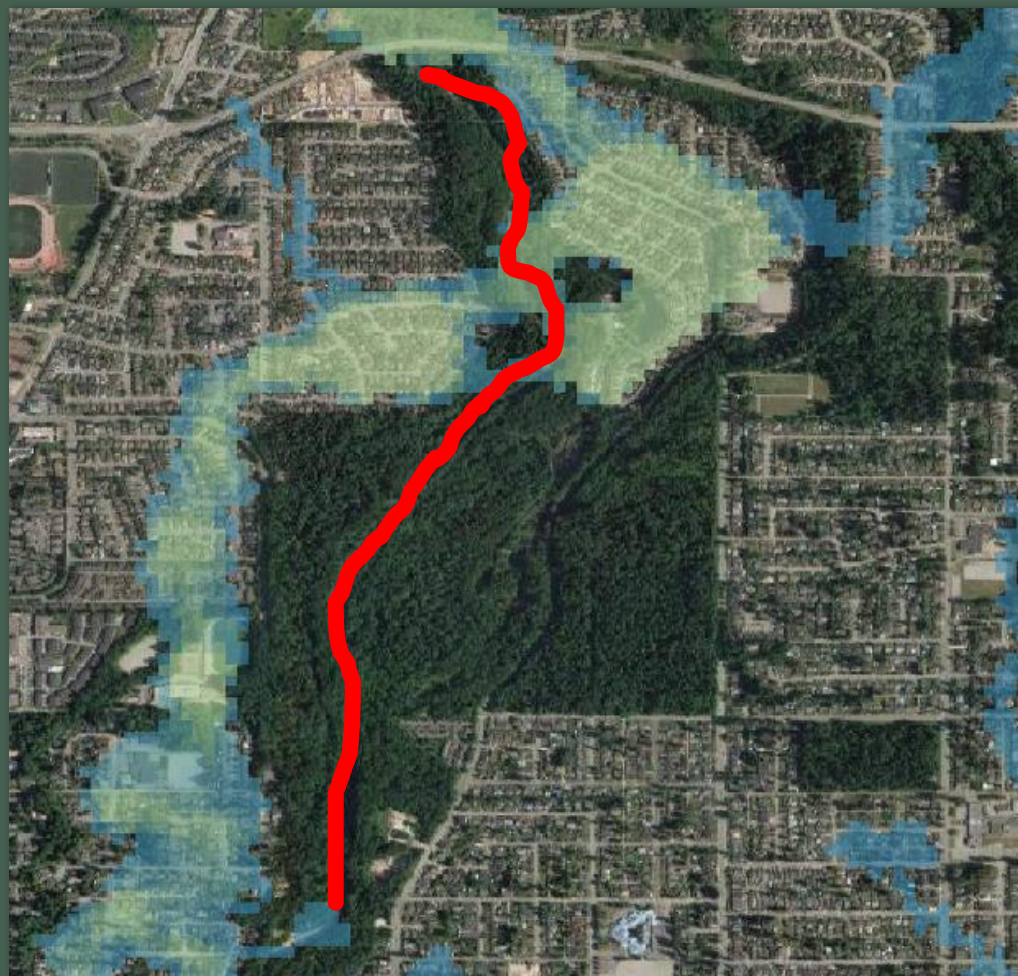


# Plan de la présentation

- Le problème
- Quelques itérations
- Notre solution
- Conclusion



# Le problème



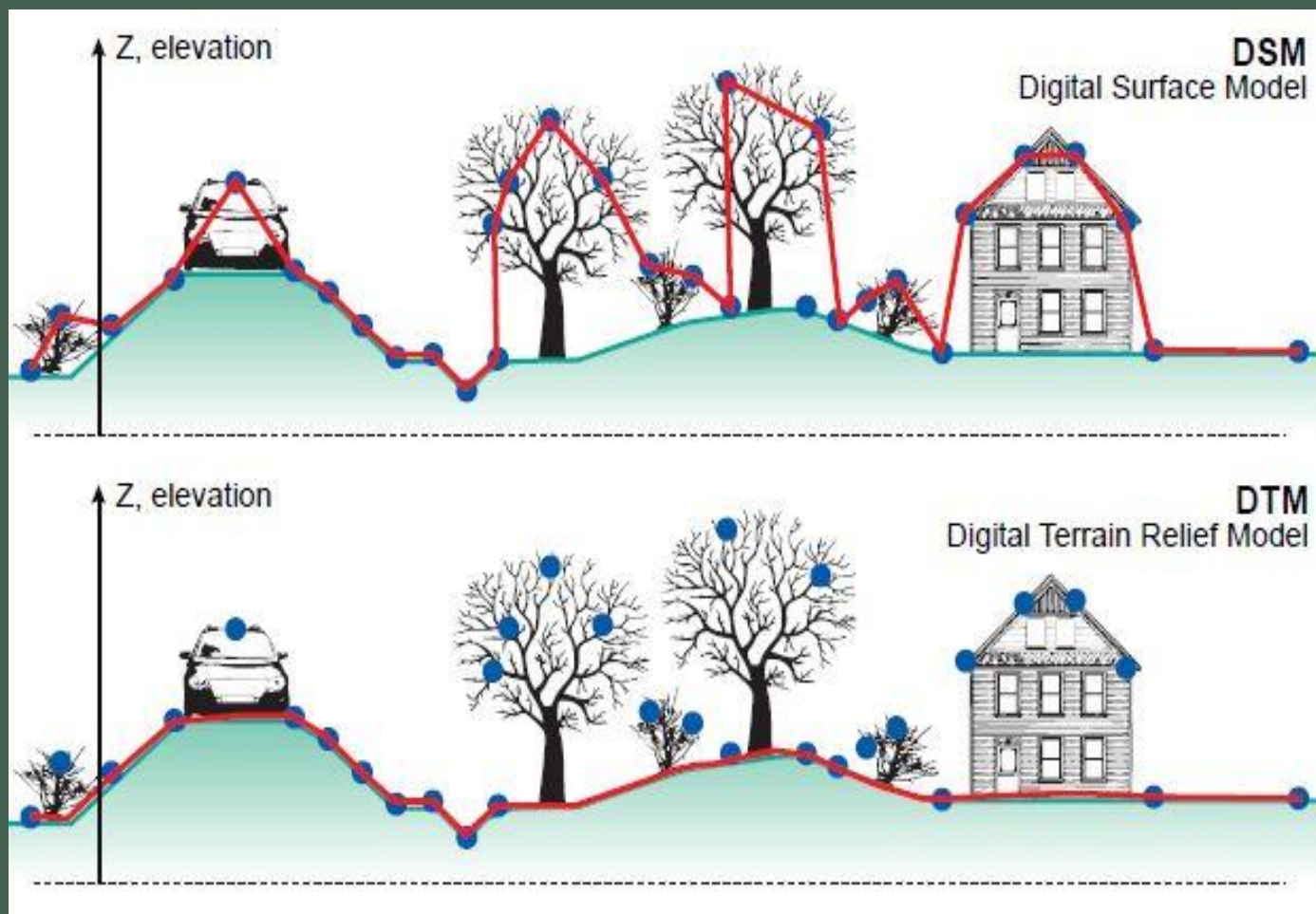


# Le problème





# Le problème



# Itération #1

:: Prédire les BORDURES ::



Gallant, J. C., Read, A. M., & Dowling, T. I. (2012).  
*Removal of tree offsets from SRTM and other digital surface models.*

1. Identifier les bordures de la forêt
2. Estimer la hauteur des arbres
3. Corriger l'élévation

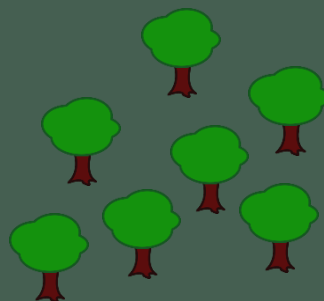
# Itération #1

:: Prédire les BORDURES ::



Gallant, J. C., Read, A. M., & Dowling, T. I. (2012).  
*Removal of tree offsets from SRTM and other digital surface models.*

1. Identifier les bordures de la forêt
2. Estimer la hauteur des arbres
3. Corriger l'élévation







# Itération #1

:: Prédire les BORDURES ::

Gallant, J. C., Read, A. M., & Dowling, T. I. (2012).  
*Removal of tree offsets from SRTM and other digital surface models.*

1. Identifier les bordures de la forêt
2. Estimer la hauteur des arbres
3. Corriger l'élévation



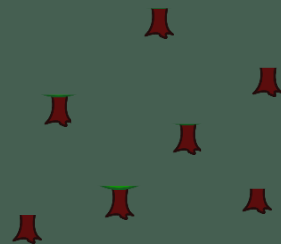


# Itération #1

:: Prédire les BORDURES ::

Gallant, J. C., Read, A. M., & Dowling, T. I. (2012).  
*Removal of tree offsets from SRTM and other digital surface models.*

1. Identifier les bordures de la forêt
2. Estimer la hauteur des arbres
3. Corriger l'élévation





# Itération #1

:: Prédire les BORDURES ::

Gallant, J. C., Read, A. M., & Dowling, T. I. (2012).  
*Removal of tree offsets from SRTM and other digital surface models.*

1. Identifier les bordures de la forêt
2. Estimer la hauteur des arbres
3. Corriger l'élévation



## Données

- SRTM
- Hansen

## Librairies

- raster

## Fonctions

- raster()
- projectRaster()
- lm()



# Itération #1

:: Prédire les BORDURES ::

## Avantages

- Plusieurs possibilités pour la détection des lisières et pour l'interpolation

## Désavantages

- Difficile de trouver une recette qui fonctionne dans des cas différents (petite région vs grande région)
- Ne fonctionne pas très bien en pratique
- Ignore l'information contenue dans les mesures effectuées dans les forêts
- Lent à appliquer et nécessite du travail manuel

# Itération #2

:: Prédire les PENTES ::



Pour faire une bonne modélisation hydrologique, l'important c'est que les pentes soient bonnes. L'élévation est moins importante.

1. Modèle qui prédit la pente en l'absence d'arbres
2. Recréer l'élévation selon la pente prédite



# Itération #2

:: Prédire les PENTES ::

Pour faire une bonne modélisation hydrologique, l'important c'est que les pentes soient bonnes. L'élévation est moins importante.

1. Modèle qui prédit la pente en l'absence d'arbres
2. Recréer l'élévation selon la pente prédite



## Données

- CDEM

## Librairies

- smoothie

## Fonctions

- kernel2dmeitsjer()
- Kernel2dsmooth()

# Itération #2

:: Prédire les PENTES ::



## Avantages

- Très bons résultats

## Désavantages

- Difficile de calibrer le modèle
- Processus itératif qui peut être lent / ne pas converger
- Manque de justification théorique / mathématique



# Notre solution

:: Prédire l'ÉLEVATION ::

Revenir à la base et prédire l'élévation

1. Ajout de données externes (ex.: météo) et dérivées
2. Modèle qui prédit la l'élévation en l'absence d'arbres
3. Correction de l'élévation



## Données

- Daymet
- ICESat
- Points de contrôle

## Librairies

- biglm

## Fonctions

- biglm()





# Notre solution

:: Prédire l'ÉLEVATION ::

## Avantages

- Très bons résultats
- Très rapide à appliquer
- Fonctionne bien dans un large éventail de conditions

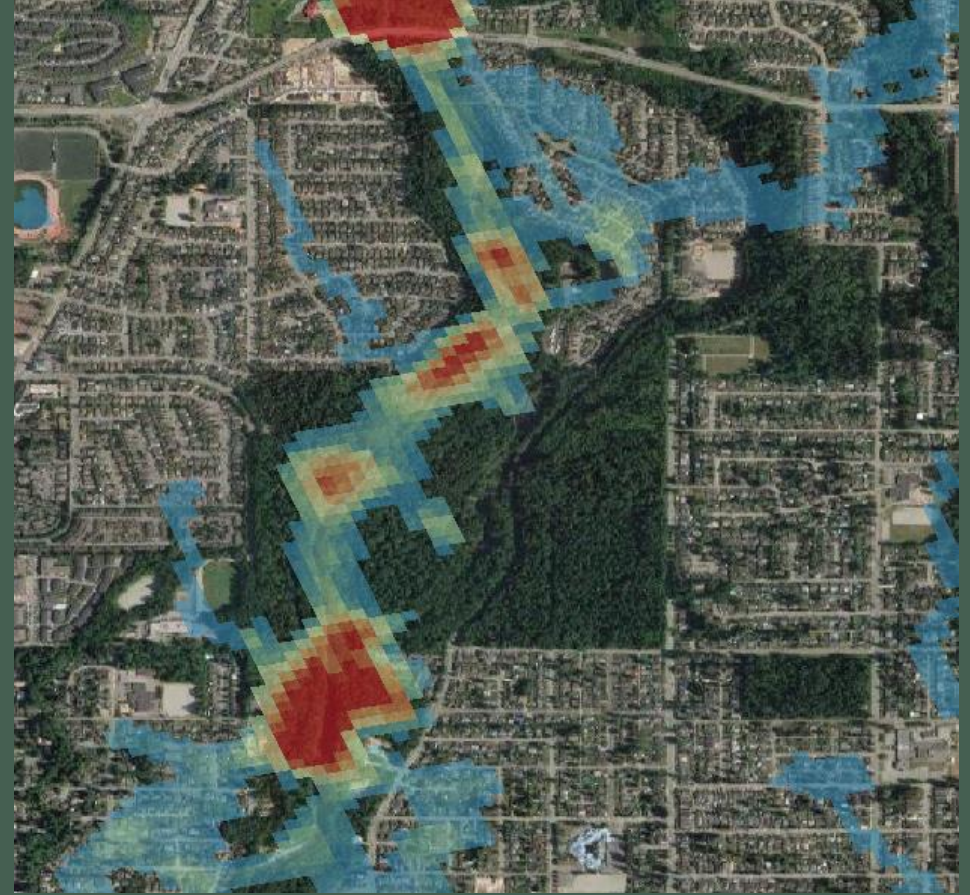
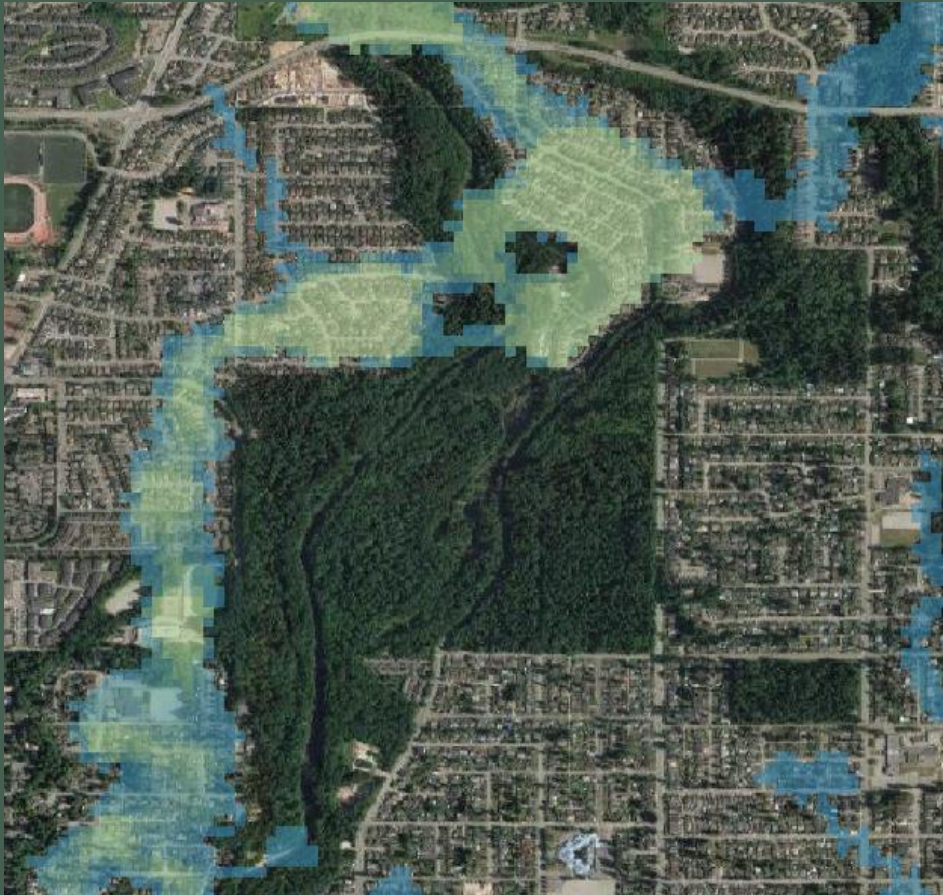
## Désavantages

- Pourrait encore être amélioré
- Modèle relativement simple



# Notre solution

:: Prédire l'ÉLÉVATION ::





# À retenir !

- R est un excellent outil pour combiner l'analyse géospatiale à des techniques d'apprentissage automatique!



# Les outils que nous avons utilisés

- R avec les librairies suivantes :
  - Rgeos, rgdal, sp, raster, smoothie, biglm.
  - Pour construire et appliquer le modèle.
- QGIS
  - Pour la visualisation des entrées et des sorties des modèles.



# Les données que nous avons utilisés

- Données d'élévation
  - MNEC
  - MNSC
- Imagerie satellite
  - Landsat
  - Sentinel
- Données météorologiques
  - Précipitation
  - Température
- Données historiques d'inondations;
- Zones inondables produites par des ingénieurs.



# Autres idées intéressantes

- Inclure les données des pixels voisins dans le modèle :
  - Moyenne pondérée,
  - Gradient.
- Inclure des interactions spécifiques ayant une interprétation physique.
- Empêcher le modèle de faire des prédictions négatives.
  - Normalement, on s'attend à ce que le MNS soit toujours plus haut ou égal au MNT.
- Utiliser une méthode de régression en ligne.
  - C'est-à-dire qui peut être entraînée itérativement.
- Utiliser des données de ICESat.