



**Chaire en  
Transformation du  
Transport**

# Modélisation de la motorisation : enjeux méthodologiques

Jérôme Laviolette, Candidat au doctorat, Génie Civil Transport,  
Polytechnique Montréal

Codirection: Catherine Morency, Owen Waygood et Konstadinos Goulias  
(UCSB)

# Plan de présentation

- ▶ Introduction
  - ▶ Mise en contexte
  - ▶ Questions de recherche et problématique
- ▶ Méthodologie
  - ▶ Approche de modélisation
  - ▶ Sources des données et préparation des variables
  - ▶ Procédure de modélisation
- ▶ Résultats
- ▶ Conclusions
  - ▶ Éléments clés
  - ▶ Limitations et étapes futures



# L'importance de comprendre les choix de motorisation des ménages

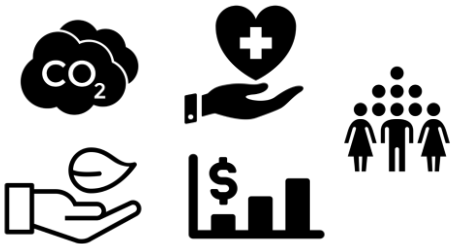


Au coeur du système  
*d'automobilité*

Impacts majeurs sur les  
comportements de mobilité  
et autres décisions

Compétition pour espaces  
urbains limités

Externalités  
négatives



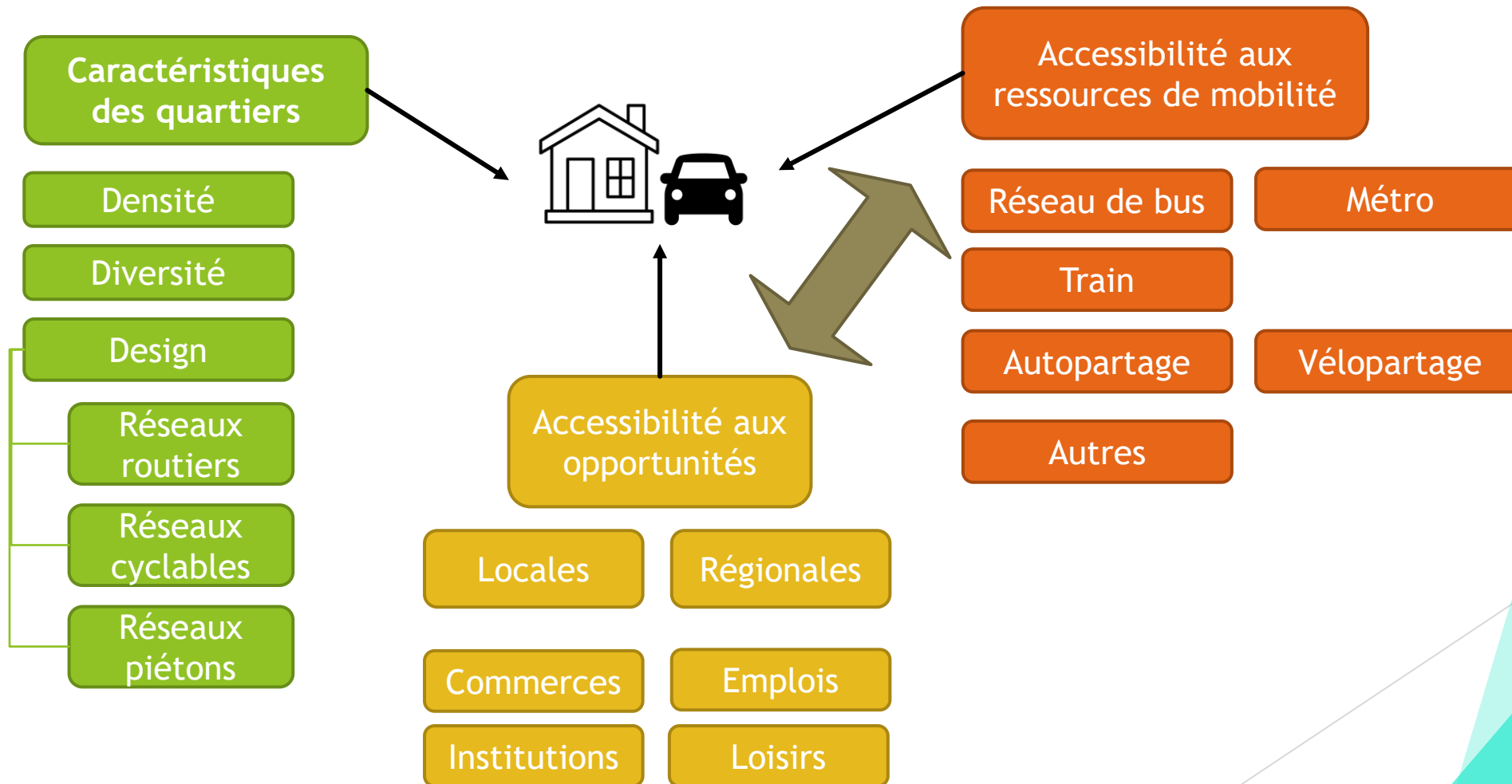
- Localisation, fréquence et durée des activités
- Choix modal, # de déplacements et distance parcourue
- Attitudes
- Choix de localisation résidentielle



Lavolette (2020),  
Gärling et Steg (2007)

Sioui, Morency, Trépanier (2012),  
von Behren et al. (2020)

# Lien entre l'environnement bâti et la motorisation des ménages

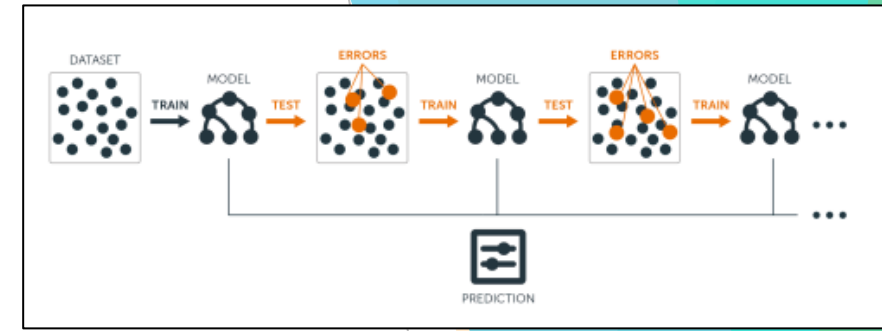


# Problématique et questions de recherche

- ▶ **Enjeux de modélisation** retrouvés dans la littérature sur la motorisation :
  - ▶ Distances / échelles spatiales de mesure souvent arbitraires, mais très rarement justifiée
  - ▶ Zones agrégées « fixes » souvent utilisées
  - ▶ Même distance / échelle spatiale utilisée pour toutes les variables
- ▶ **Questions de recherche:**
  - ▶ Quelle **distance / échelle spatiale** devrait être utilisée pour mesurer ces caractéristiques ?
    - ▶ Des distances différentes peuvent-elles mener à des conclusions et des recommandations de politiques publiques différentes ?
  - ▶ **Combien d'opportunités / quelles valeurs** (de densité, diversité, design) sont nécessaires pour avoir un impact sur les décisions de motorisation



# Gradient Boosting Machine (GBM) decision-trees



- ▶ Fonctionne par l'addition séquentielle de modèles simples (weak-learners) (ex. : petits arbres de décision à 1-6 embranchements)
- ▶ Chaque nouvel arbre est ajusté aux résidus de l'ensemble existant d'arbres
- ▶ L'algorithme fonctionne par *descente de gradient*: il peut être généralisé à n'importe quelle fonction de perte (loss function). Ceci permet de modéliser différents types de variables: continue, de comptage, classification binaire ou multinomial
- ▶ *Un gradient de descente stochastique* consiste à prendre une fraction aléatoire des données et permet d'éviter de trouver uniquement des minimums locaux.
- ▶ Modélisation faites avec *xgboost* dans R

1. Fit un arbre de décision aux données:  $F_1(x) = y$
2. Fit l'arbre suivant aux résidus du précédent:  $h_1(x) = y - F_1(x)$
3. Ajouter cet arbre à l'algorithm:  $F_2(x) = F_1(x) + h_1(x)$
4. ...etc. jusqu'à un mécanisme d'arrêt (ex.: validation-croisée)

Modèle séquentielle additif de  $b$  arbres de décisions individuels:

$$f(x) = \sum_{b=1}^B f^b(x)$$



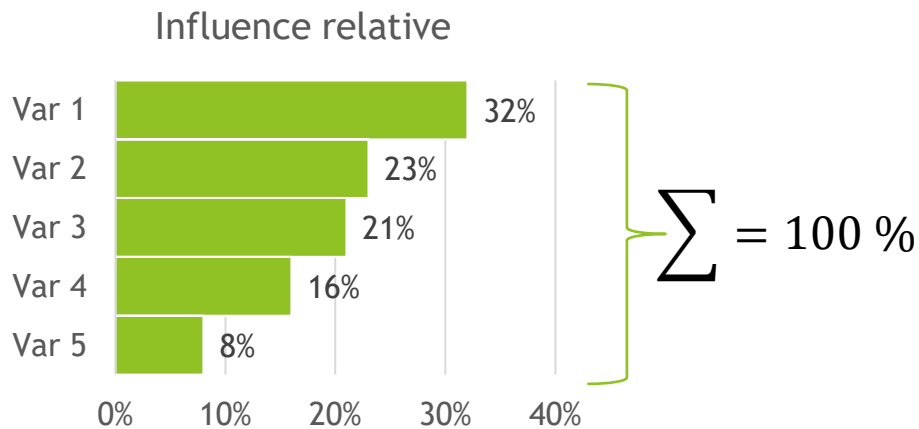
Hastie, Tibshirani, & Friedman (2009, ch. 10)  
Friedman (2001, 2002)  
Chen et He (2021)

# GBM : Outils d'interprétation des résultats

## 1. Influence relative des variables de prédiction

- Influence moyenne de la variable  $j$  sur l'ensemble d'arbre  $M$  :

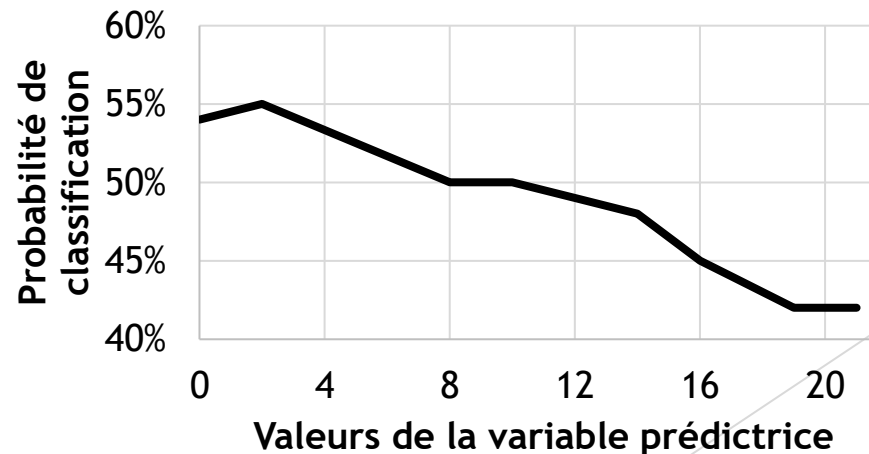
$$I_j^2 = \frac{1}{M} \sum_{m=1}^M \sum_{t=1}^{J-1} i_t^2 1(v_t = j)$$



## 2. Courbes de dépendance partielle (partial dependence plot - PDP)

- Changement moyen dans la variable prédit, en fonction des valeurs d'une variable prédictive  $j$ , en tenant les autres variables constantes à leur moyenne.

### Courbe de dépendance partielle



# Sources des données et territoire d'étude

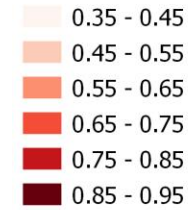
## Enquête OD 2018 de Montréal

- Variables sociodémographiques
- Variable dépendante modélisée (statut de motorisation)

## Variables de l'environnement bâti

- Fusion de plusieurs jeux de données spatiaux:
  - Recensement canadien 2016
  - GTFS (horaires et réseaux TC)
  - Utilisation des sols (CMM)
  - Stations Communauto
  - Stations BIXI
  - DMTI Spatial : Opp. Commerciales
  - OSM : Réseaux routiers et piétons

### Avg VEH / ADULT within HH



### Road Network

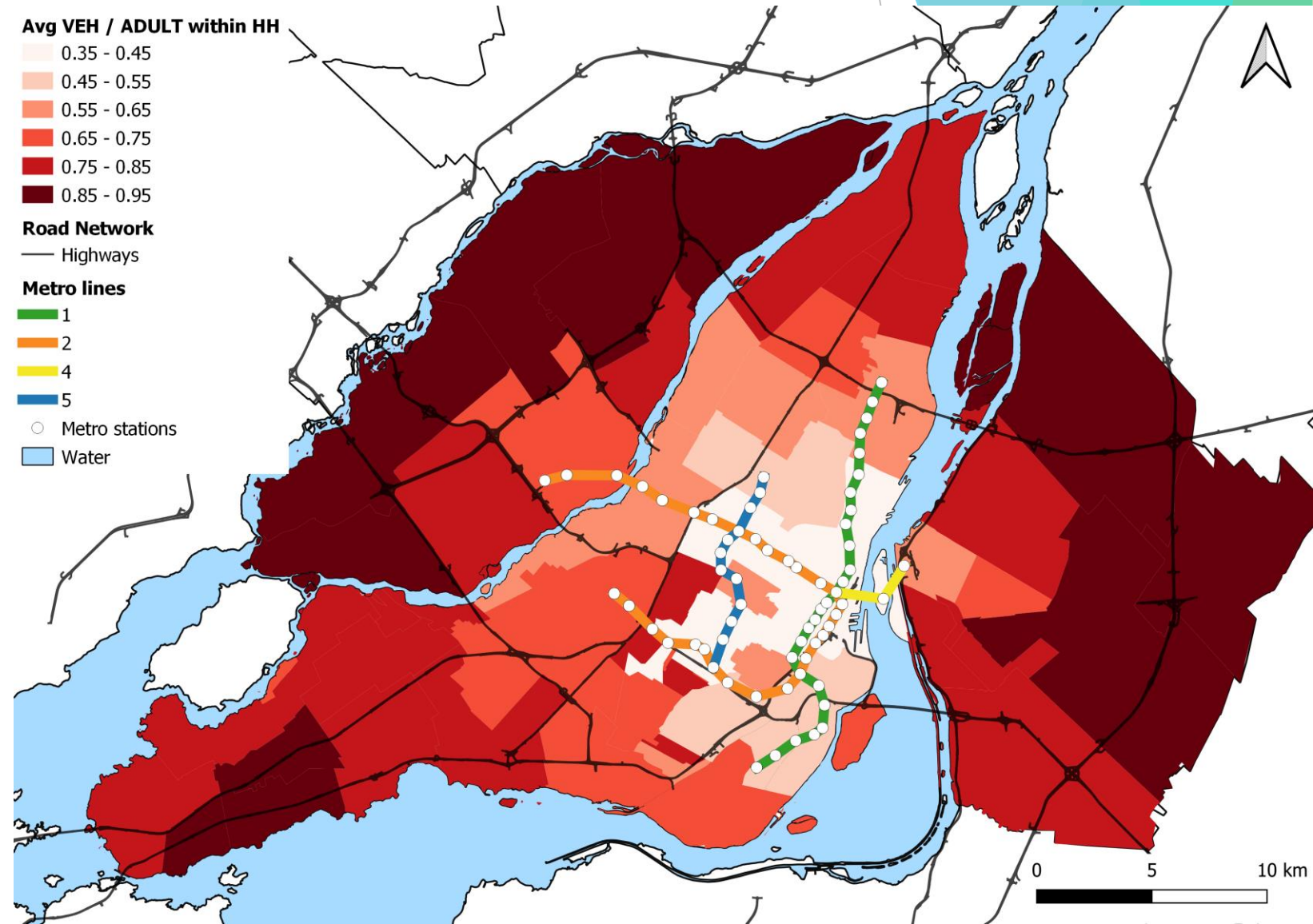
— Highways

### Metro lines



○ Metro stations

Water





# Calcul des zones tampons (« buffers »)

- ▶ Création de tampons de distance de marche autour des 46 517 ménages (unique à chaque ménage)
- ▶ Accessibilité réseau plus précise et réaliste que tampons à vol d'oiseau
- ▶ Réseau piéton OSM utilisé (incluant les chemins exclusifs piétons)
- ▶ 7 distances calculées : 400, 800, 1000, 1200, 1400 et 1600 m
- ▶ **OUTILS :**
  - ▶ « Analyse de réseau » de QGIS
  - ▶ Fonction « Concave Hull » de PostGIS

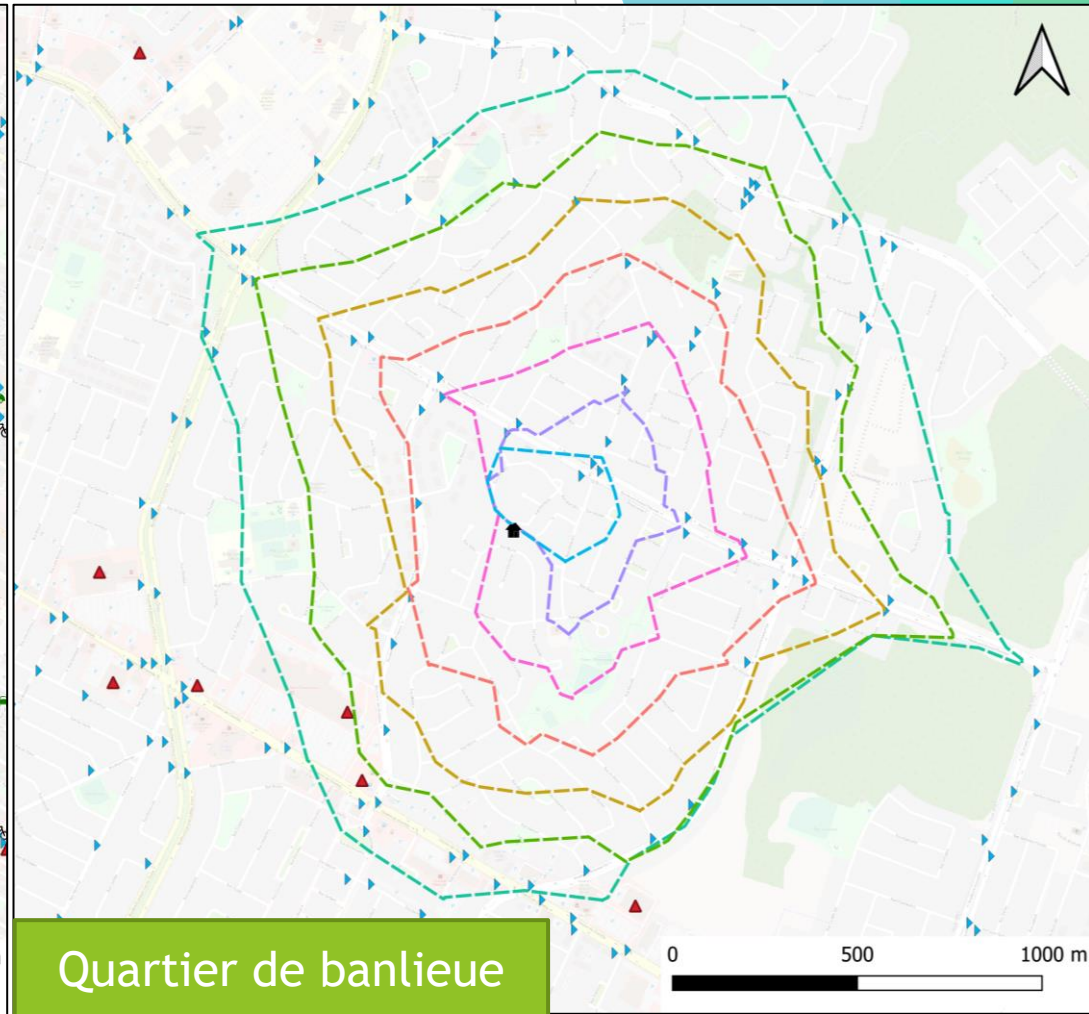
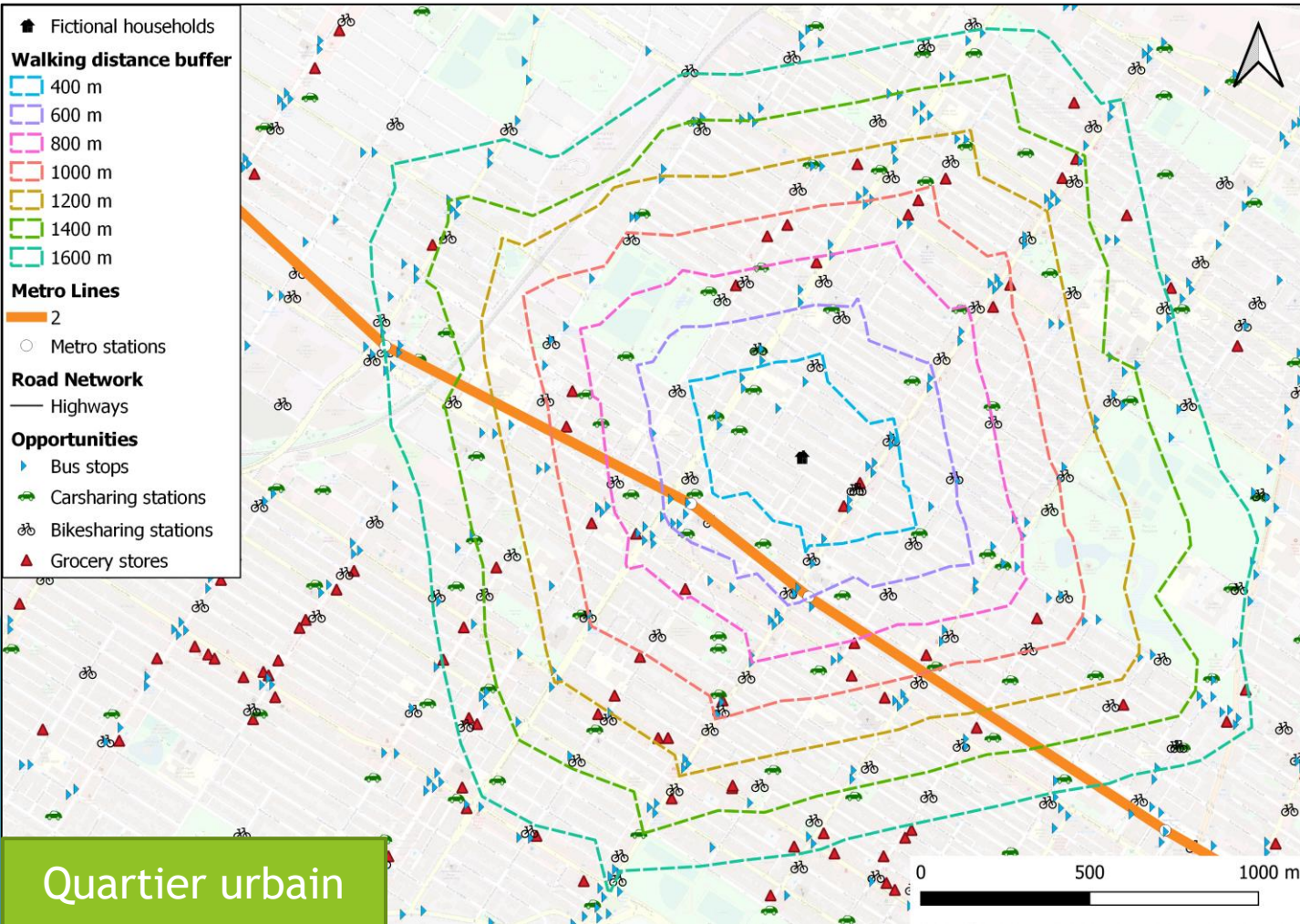


## Dernière étape:

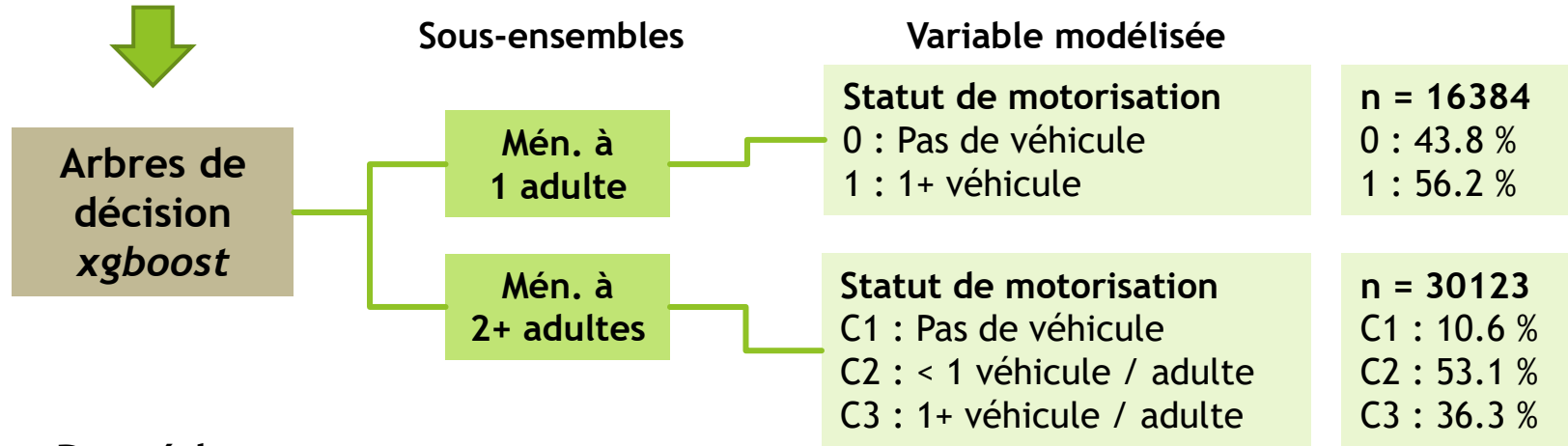
- ▶ Calculs des variables de l'environnement bâti à l'intérieur de chaque tampon pour chaque ménage.



# Calcul des zones tampons (« buffers »)



Quelle **distance-tampon** devrait être privilégiée pour chaque variable ?  
Quelle **valeur seuil** pour chaque variable a un effet structurel sur la motorisation ?



## Procédure

1. Construire les modèles de référence avec toutes les variables mesurées à 1 km de marche
2. Modèles entrainer avec 70 % des observations de chaque sous-ensemble
3. Pour chaque sous-ensemble, faire varier la taille de tampon **UNE** variable à la fois, par bond de 200 m [de 400 à 1600 m] en gardant les autres constantes à 1 km.
4. Analyser les courbes de dépendance partielle (PDP) et l'influence relative des variables
5. Comparer les effets entre sous-ensemble et selon chaque distance testée

# Modèles de référence : Ménages à 1 adulte

► Précision du modèle sur ensemble test (30 %) : **70.8 %**

► Classification correct:

0 véh. : **64.6 %**

1+ véh. : **75.7 %**

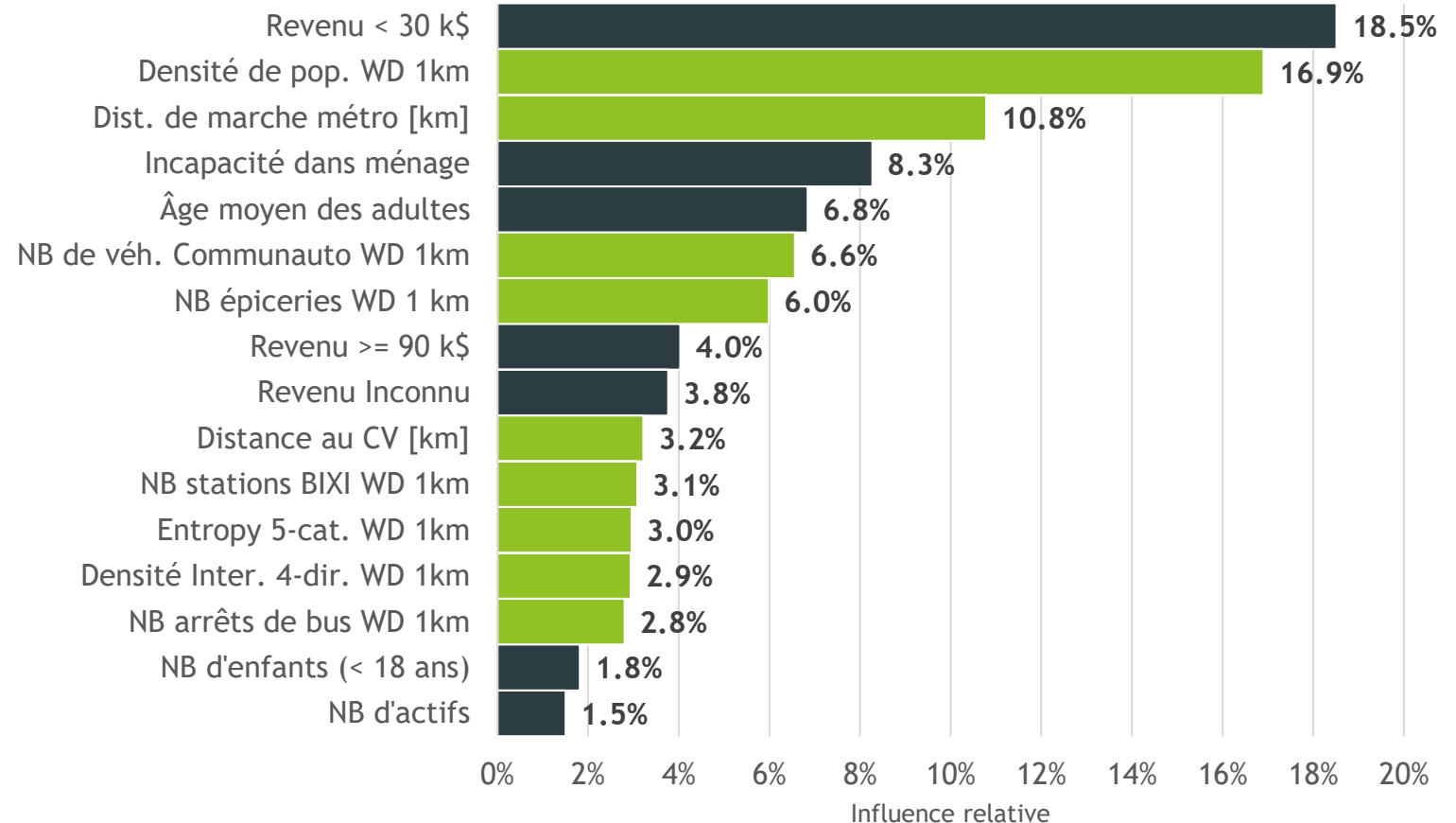
► Somme d'influence:

Sociodémo : **44.8 %**

Environnement bâti : **55.2 %**

## Influence relative des variables

■ Environnement bâti ■ Sociodémographique



\*WD: Walk-distance

# Modèles de référence : Ménages à 2+ adultes

► Précision du modèle sur ensemble test (30 %) : **64.8 %**

► Classification correct:

0 véh. : **17.6 %**

< 1 veh/adu. : **80.8 %**

>= 1 veh/adu. : **54.4 %**

► Somme d'influence:

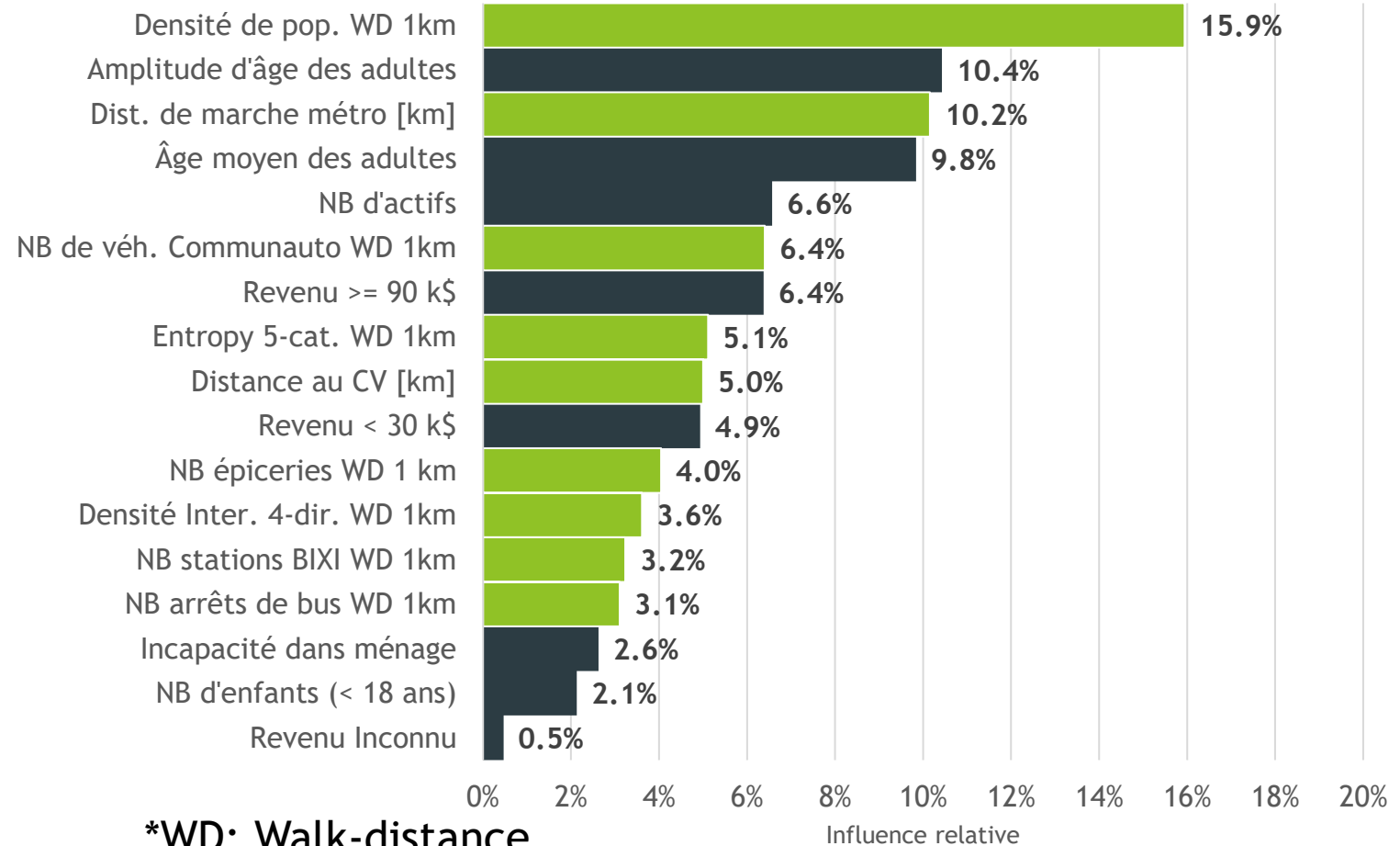
Sociodémo : **43.4 %**

**Environnement bâti :**

**56.6 %**

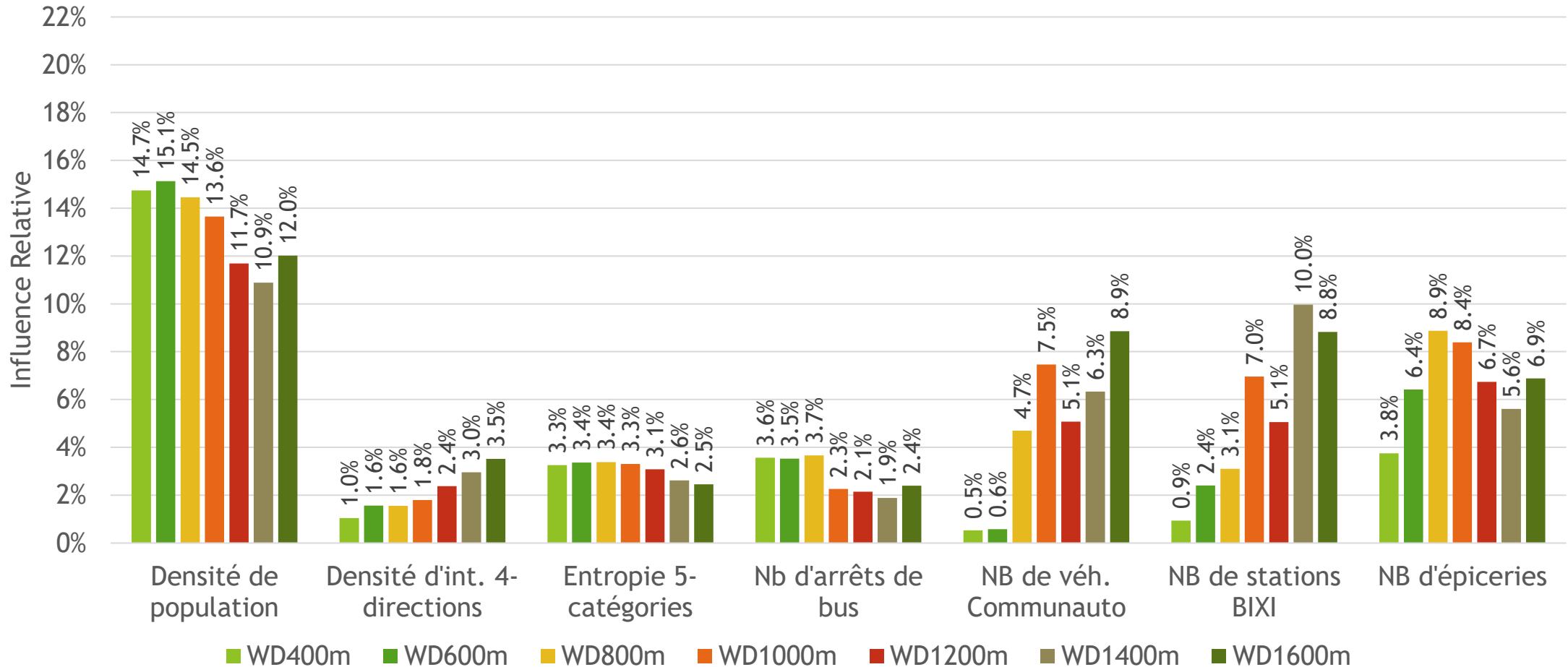
## Influence relative des variables

■ Environnement bâti ■ Sociodémographique



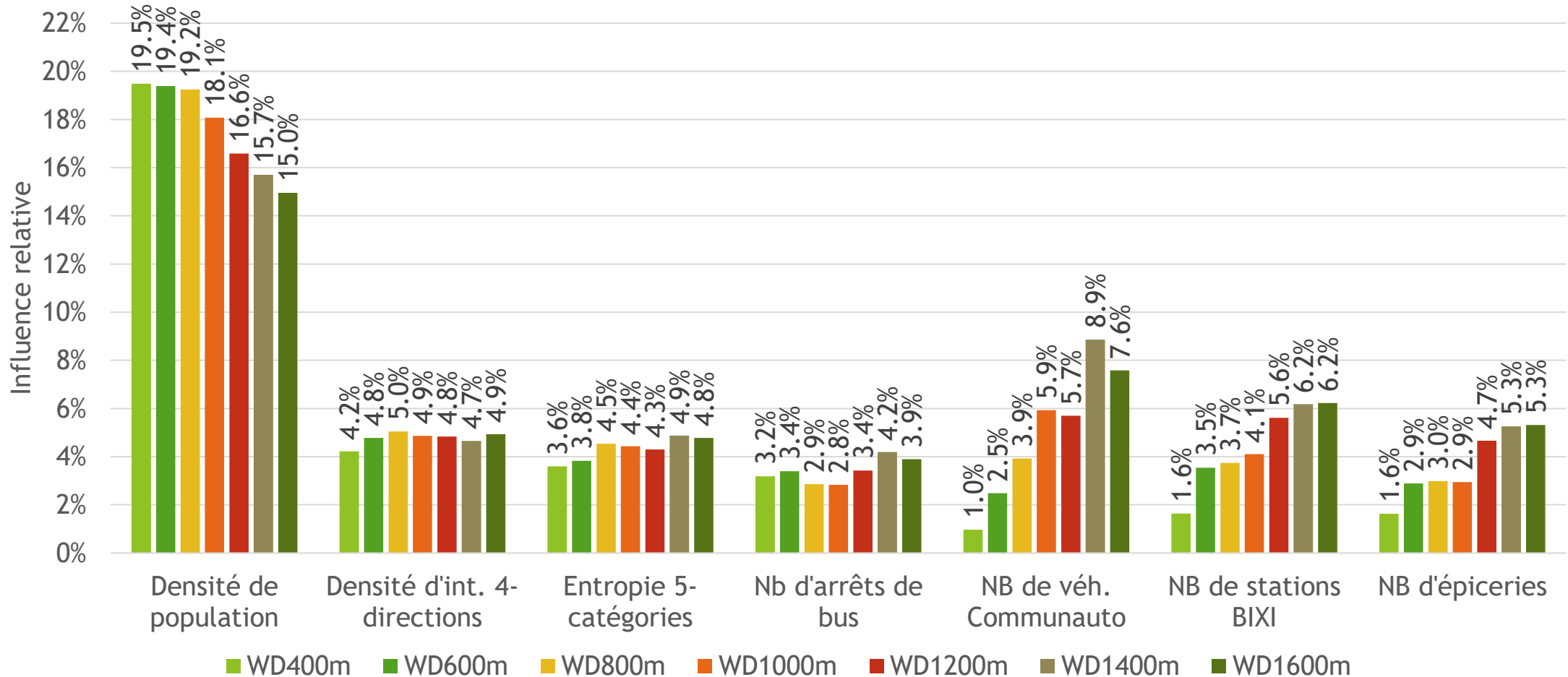
\*WD: Walk-distance

# Impacts de la variation de la taille des buffers : Ménages à 1 adulte



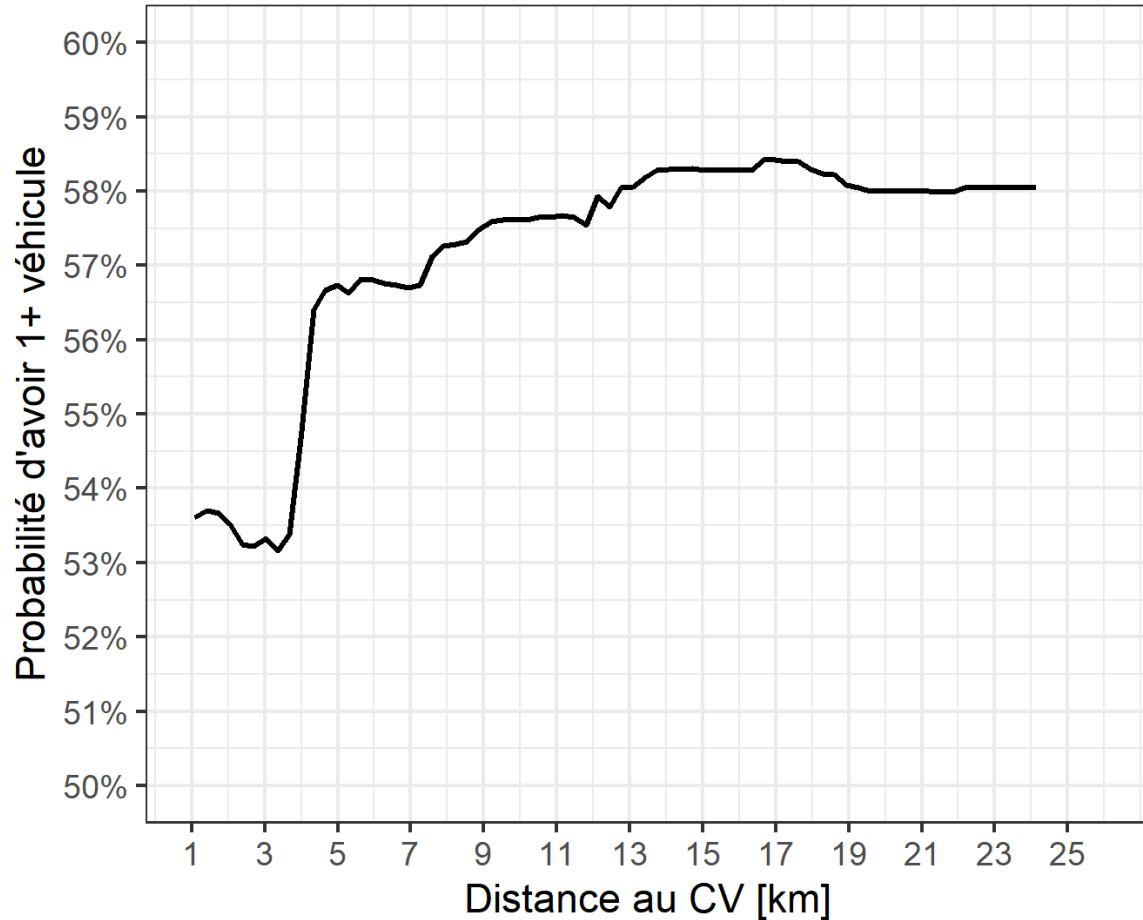
# Impacts de la variation de la taille des buffers :

## Ménages à 2+ adultes

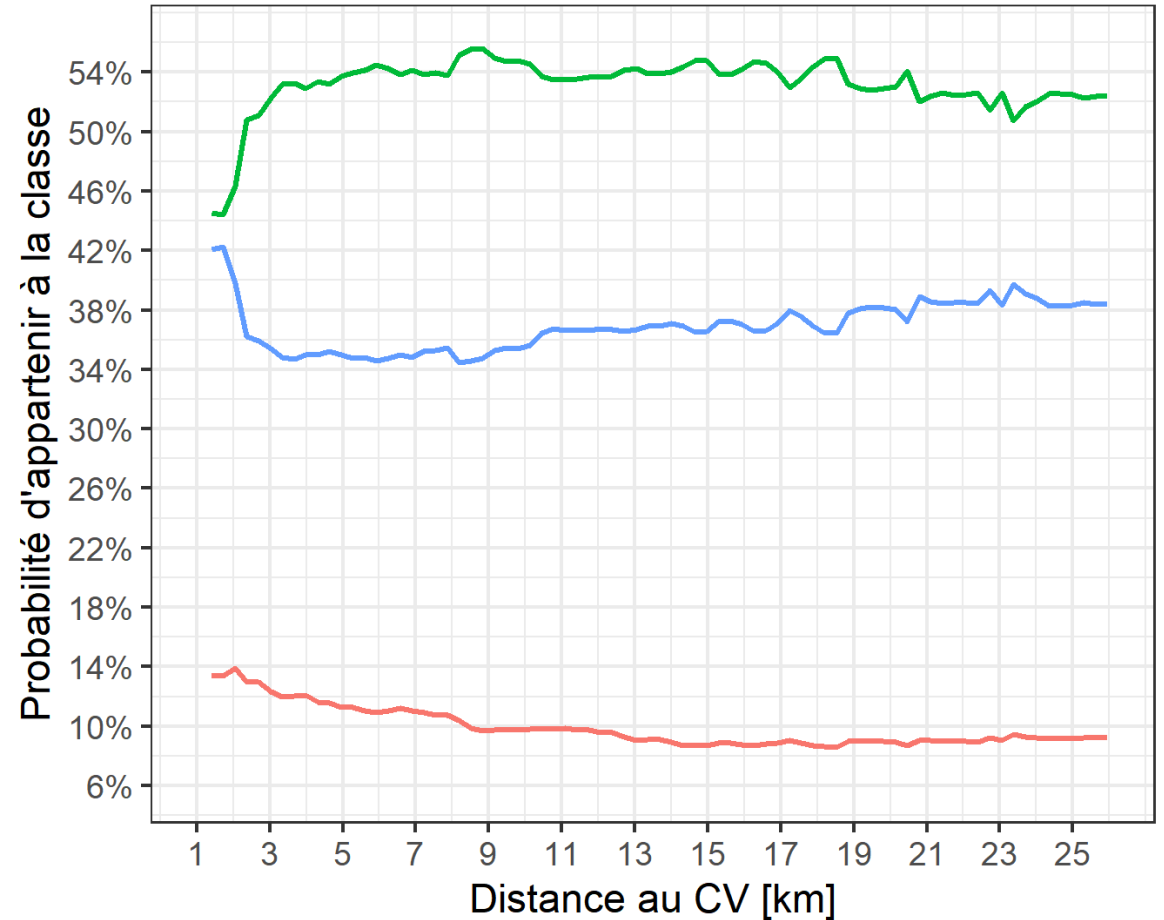


# PDP: Distance au centre-ville

**A** Mén. 1 adulte : 1+ car vs 0  
 Var : Distance au CV, IR = 3.22 %



**B** Mén. 2+ adultes : Véh. / adulte - 3 cat.  
 Var : Distance au CV, IR = 5 %

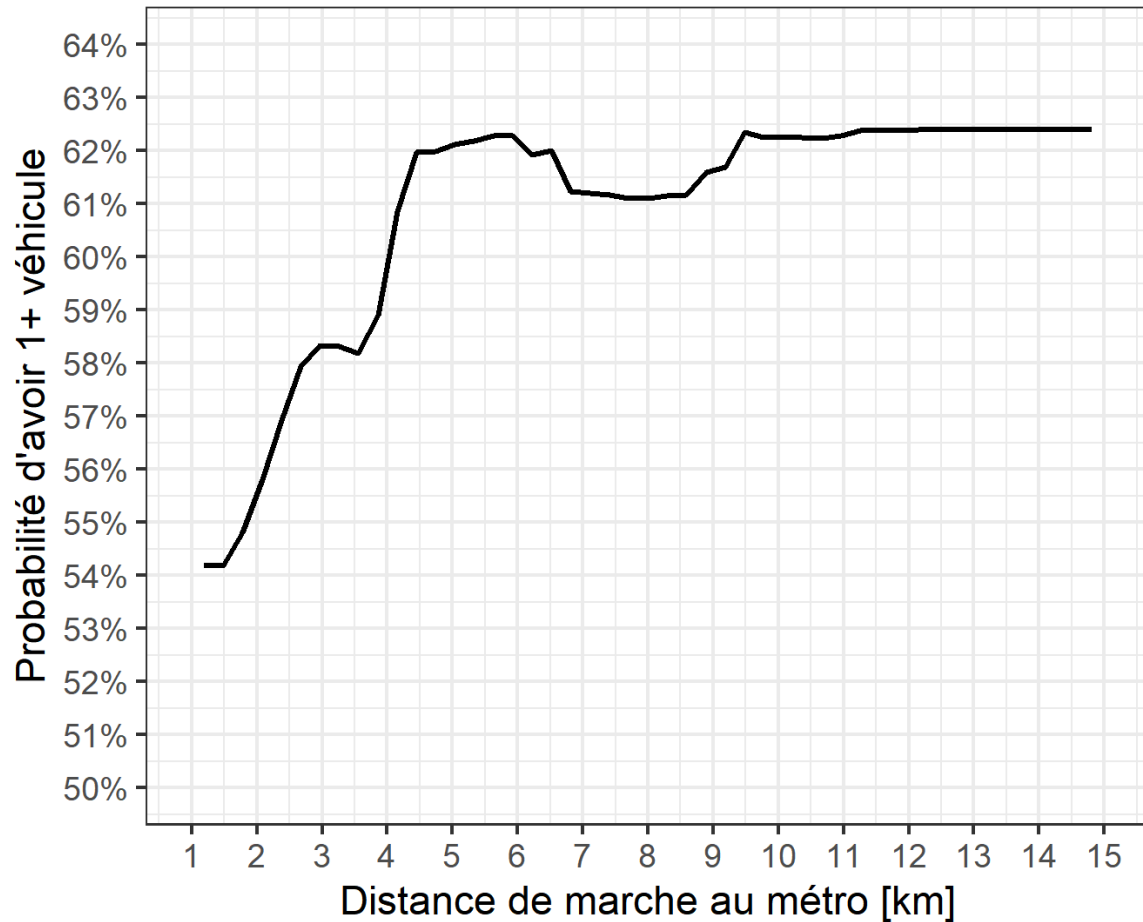


statut.motorisation — 1: 0 véh. — 2: < 1 véh. / adulte — 3: >= 1 véh. / adulte

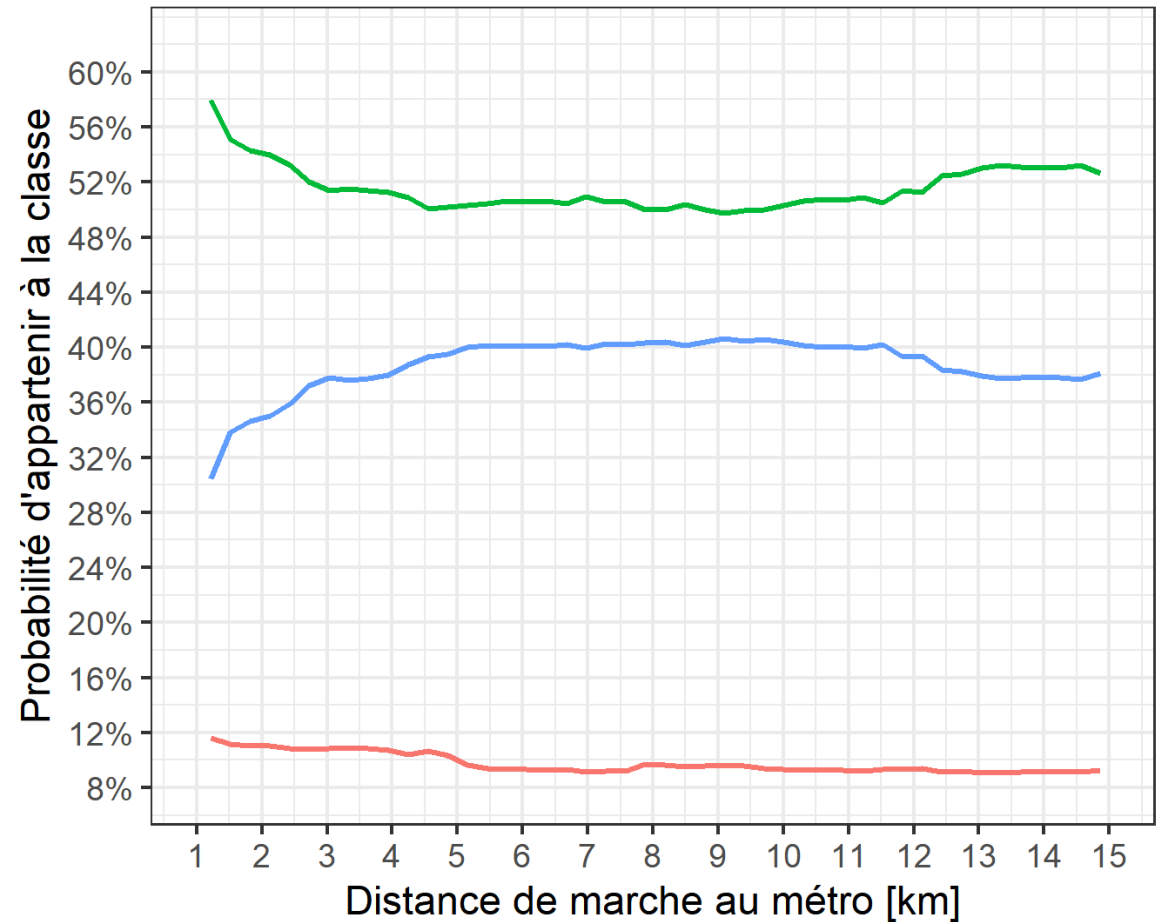


# PDP: Distance de marche à une station de métro

**A** Mén. 1 adulte : 1+ car vs 0  
 Var : Dist. de marche au métro, IR = 10.78 %



**B** Mén. 2+ adultes : Véh. / adulte - 3 cat.  
 Var : Dist. de marche au métro, IR = 10.16 %



statut.motorisation — 1: 0 véh. — 2: < 1 véh. / adulte — 3: >= 1 véh. / adulte

A : Modèle de Ménages à **1 adulte**



Probabilité d'être motorisé (1+ veh vs 0)  
en fonction de la variable analysée

B : Modèle de Ménages à **2+ adultes**

Probabilité d'être de la classe 1:  
**NON-motorisé: 0 véhicule**  
en fonction de la variable analysée

C : Modèle de Ménages à **2+ adultes**

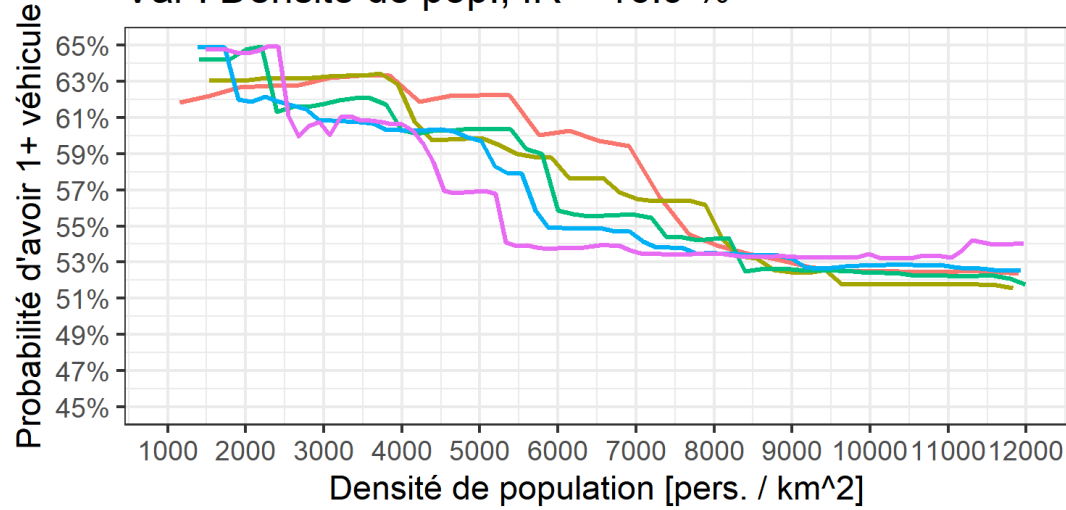
Probabilité d'être de la classe 2:  
**Motorisation partielle: < 1 véh. / adulte**  
en fonction de la variable analysée

D : Modèle de Ménages à **2+ adultes**

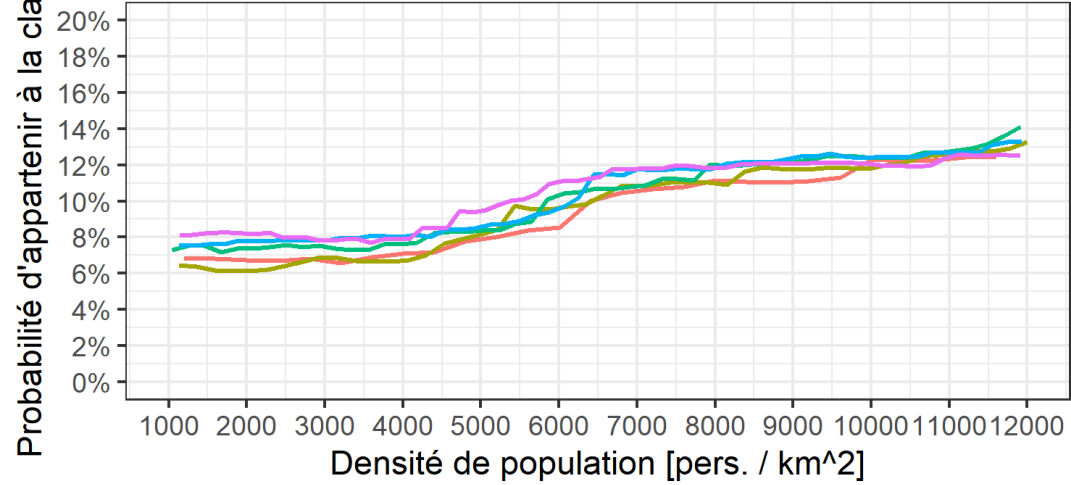
Probabilité d'être de la classe 3:  
**Motorisation complète: >= 1 véh. / adulte**  
en fonction de la variable analysée



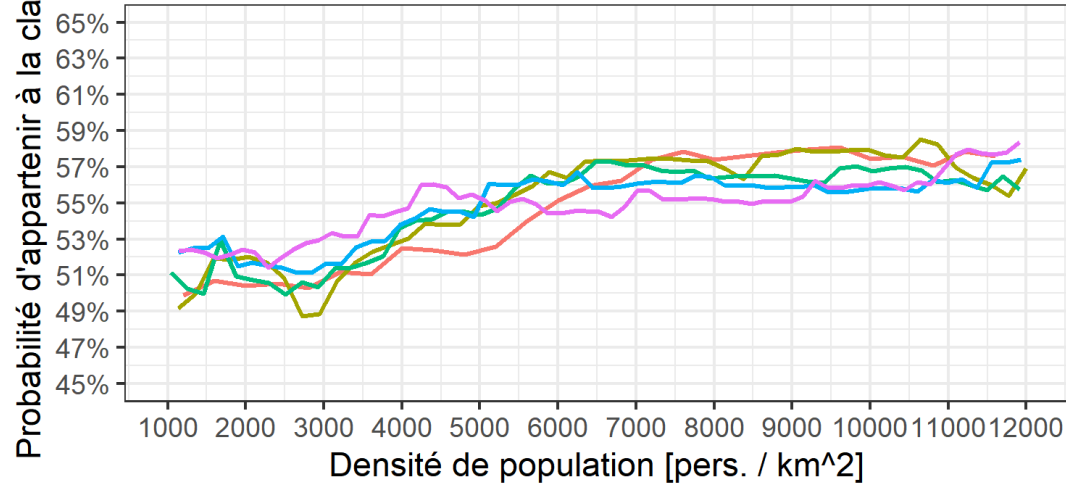
**A** Mén. 1 adulte : 1+ car vs 0  
Var : Densité de pop., IR = 16.9 %



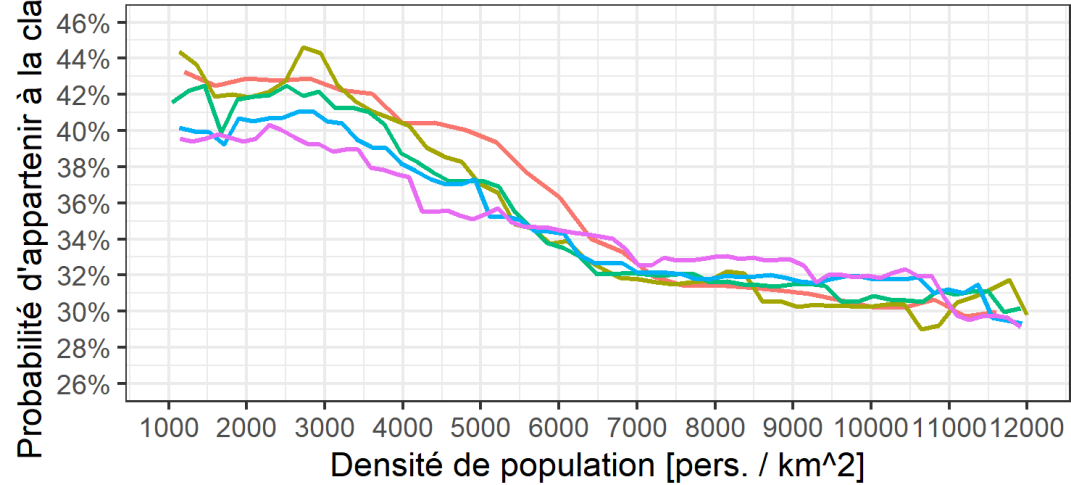
**B** Mén. 2+ adultes - C1: 0 véh.  
Var : Densité de pop., IR = 15.94 %



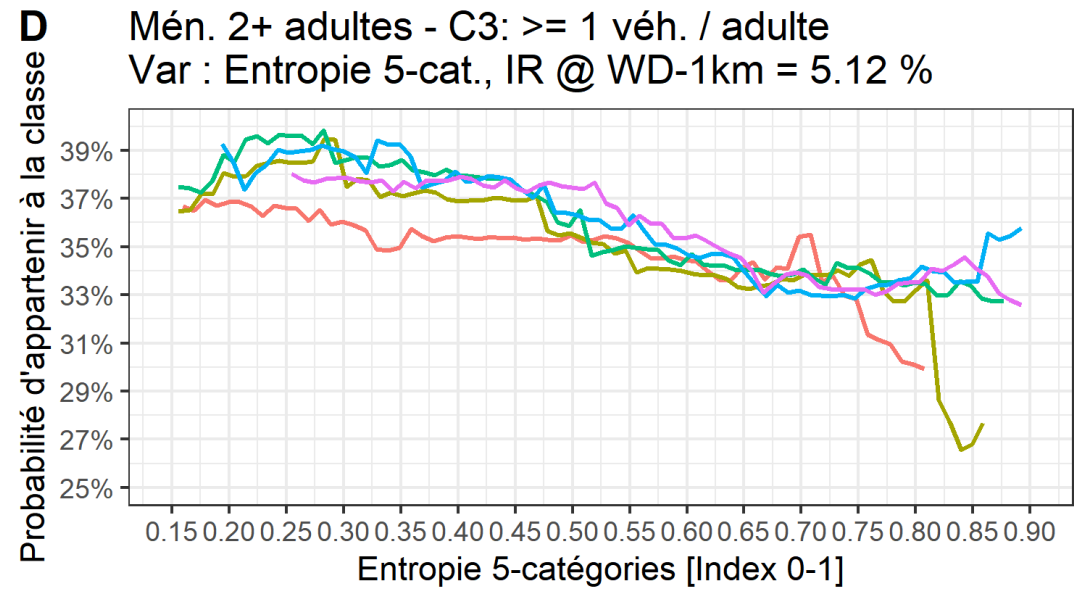
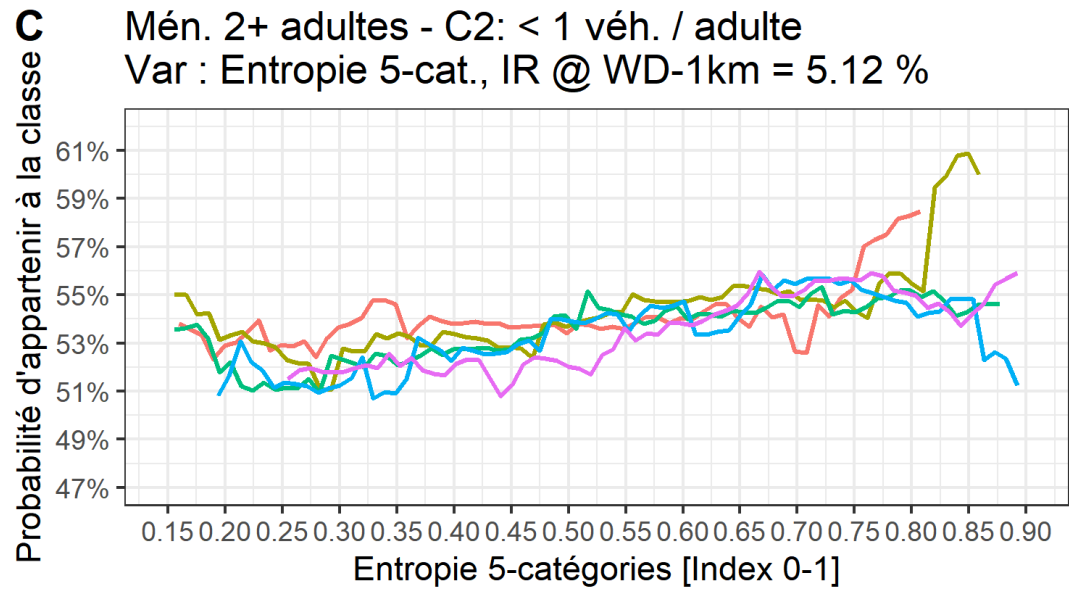
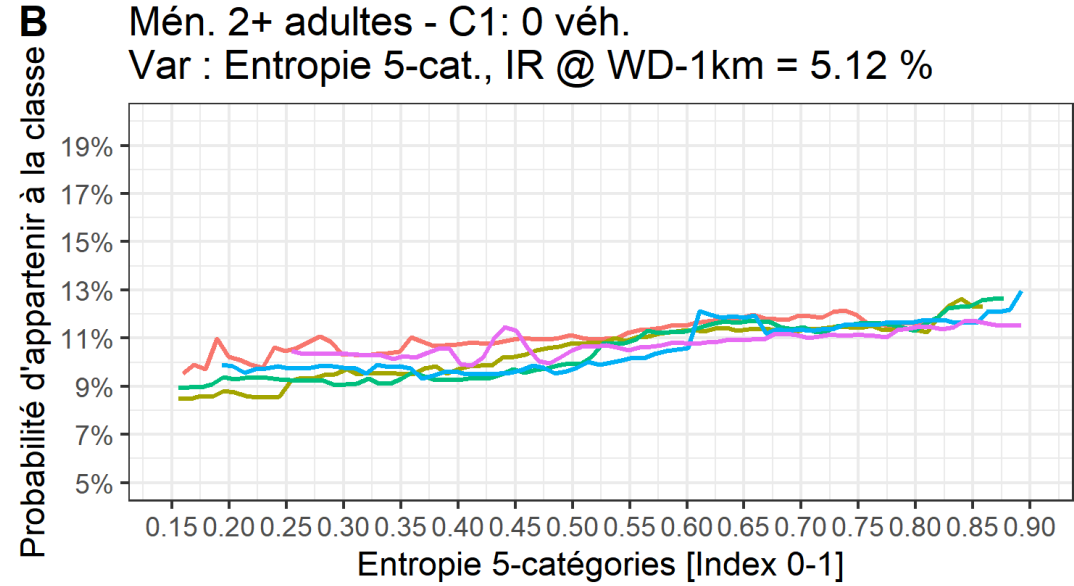
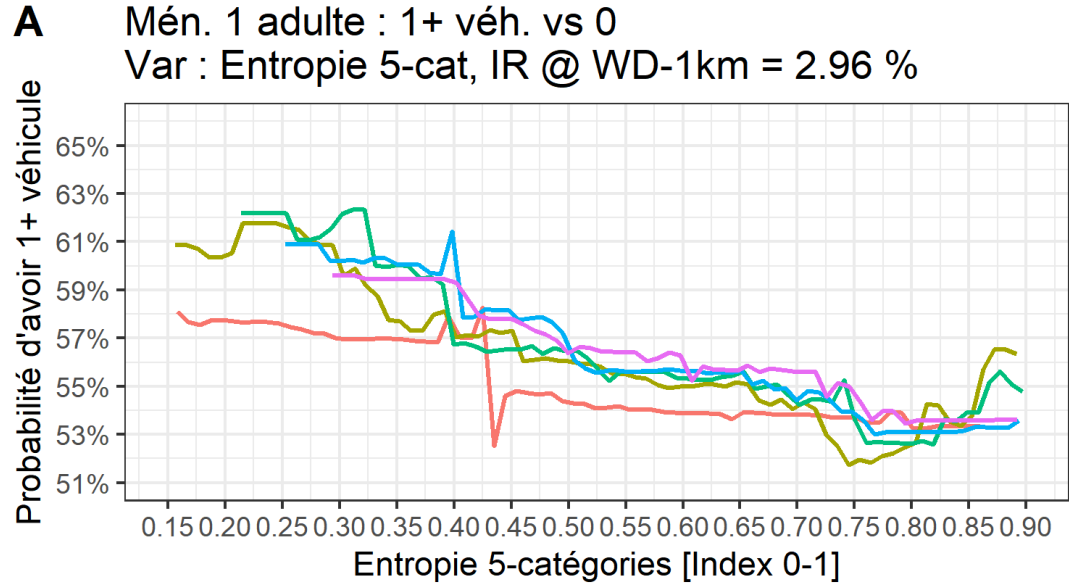
**C** Mén. 2+ adultes - C2: < 1 véh. / adulte  
Var : Densité de pop., IR @ WD-1km = 15.94 %



**D** Mén. 2+ adultes - C3: >= 1 véh. / adulte  
Var : Densité de pop., IR @ WD-1km = 15.94 %

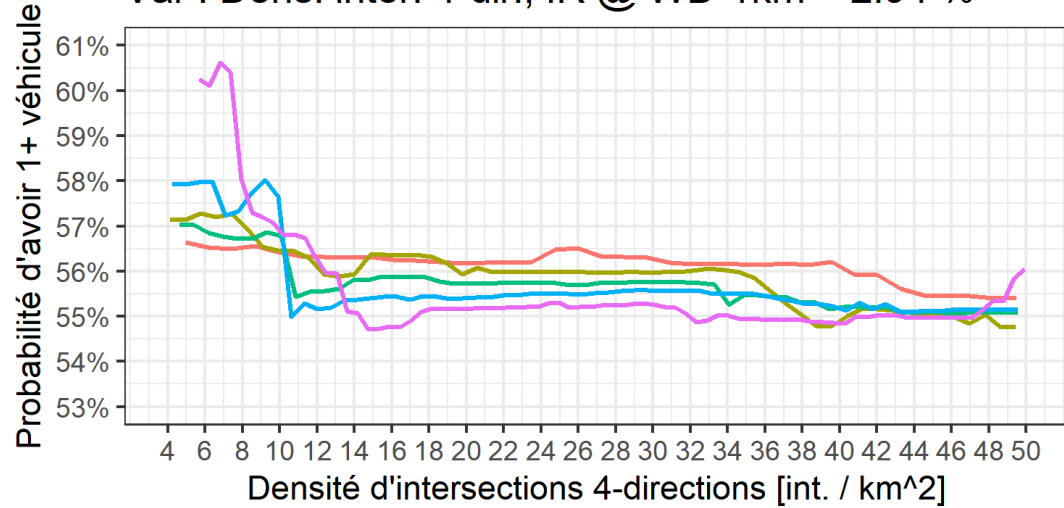


buffer size — wd400 — wd800 — wd1000 — wd1200 — wd1600

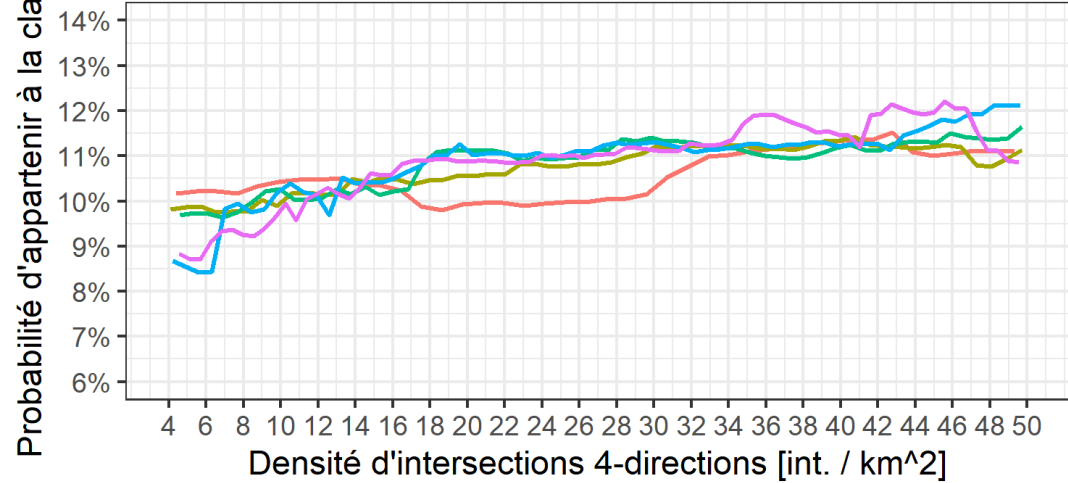


buffer size — wd400 — wd800 — wd1000 — wd1200 — wd1600

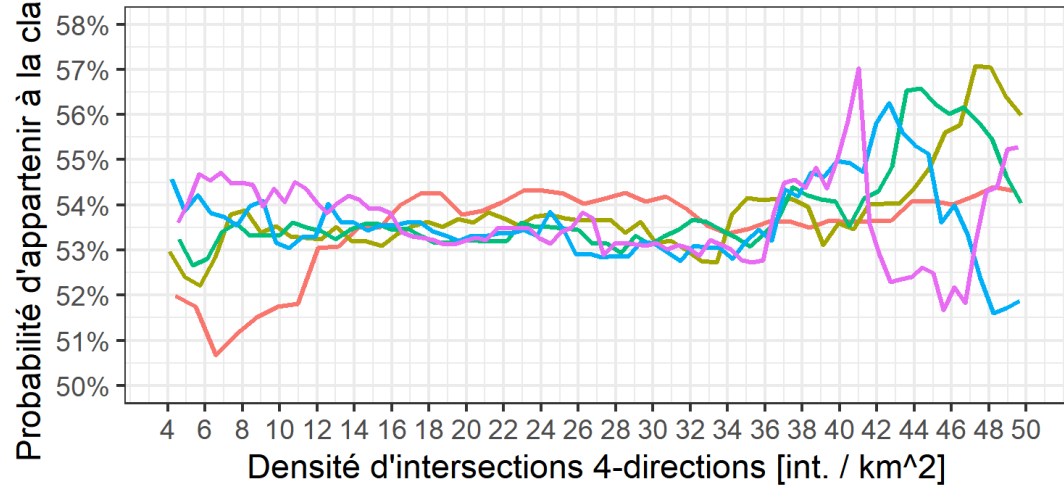
**A** Mén. 1 adulte : 1+ véh. vs 0  
Var : Dens. inter. 4-dir., IR @ WD-1km = 2.94 %



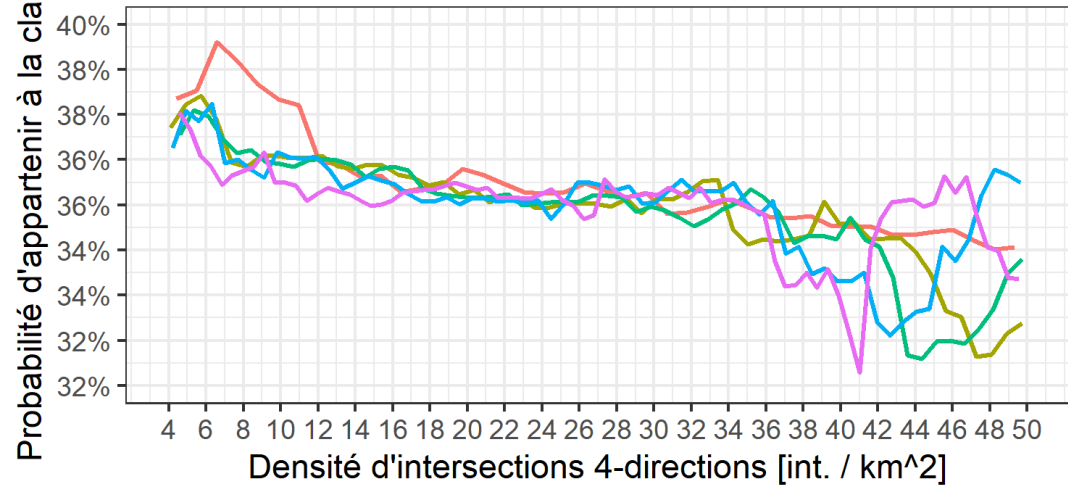
**B** Mén. 2+ adultes - C1: 0 véh.  
Var : Dens. inter. 4-dir., IR @ WD-1km = 3.61 %



**C** Mén. 2+ adultes - C2: < 1 véh. / adulte  
Var : Dens. inter. 4-dir., IR @ WD-1km = 3.61 %

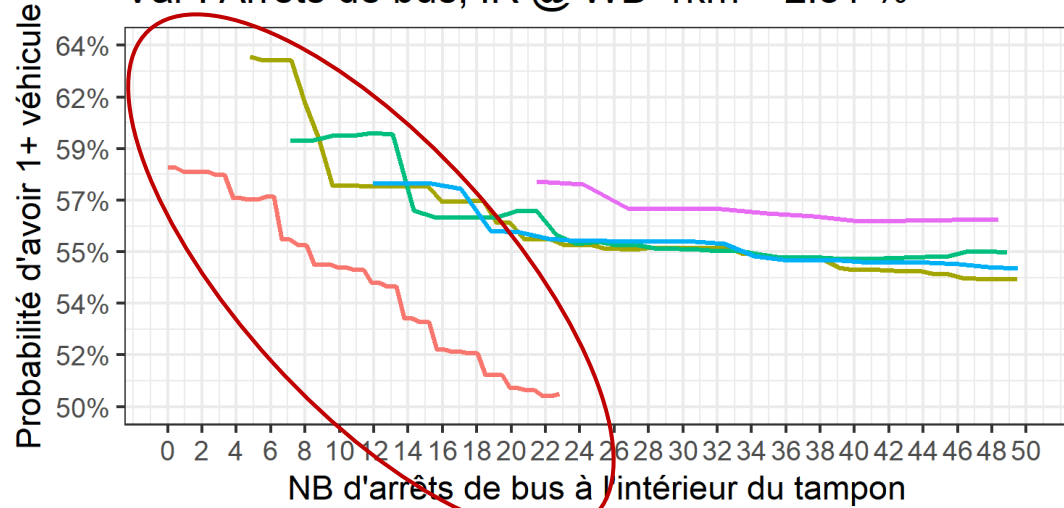


**D** Mén. 2+ adultes - C3: >= 1 véh. / adulte  
Var : Dens. inter. 4-dir., IR @ WD-1km = 3.61 %

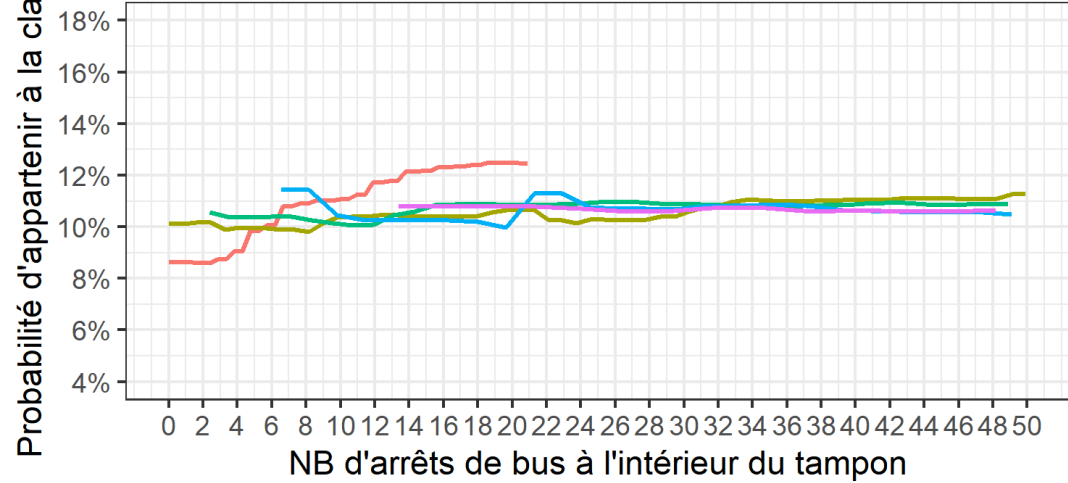


buffer size — wd400 — wd800 — wd1000 — wd1200 — wd1600

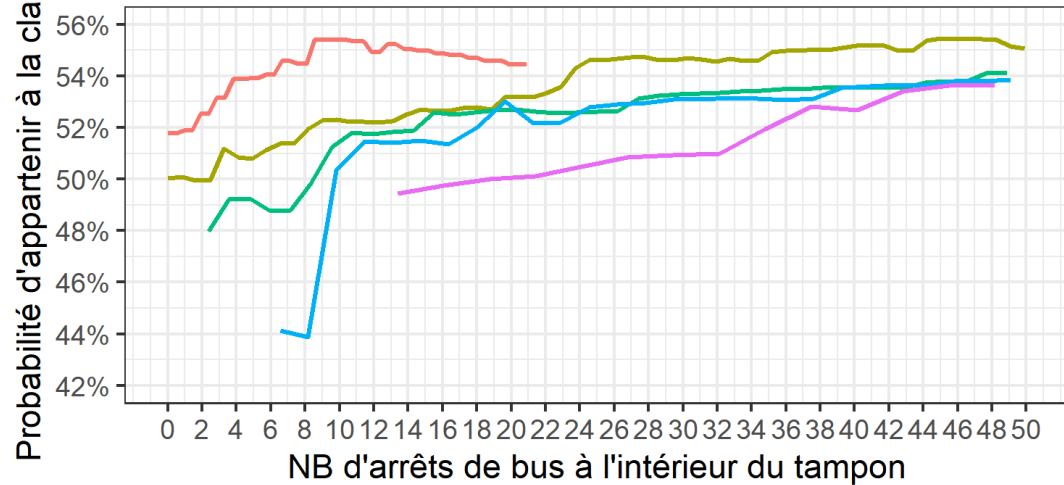
**A** Mén. 1 adulte : 1+ véh. vs 0  
Var : Arrêts de bus, IR @ WD-1km = 2.81 %



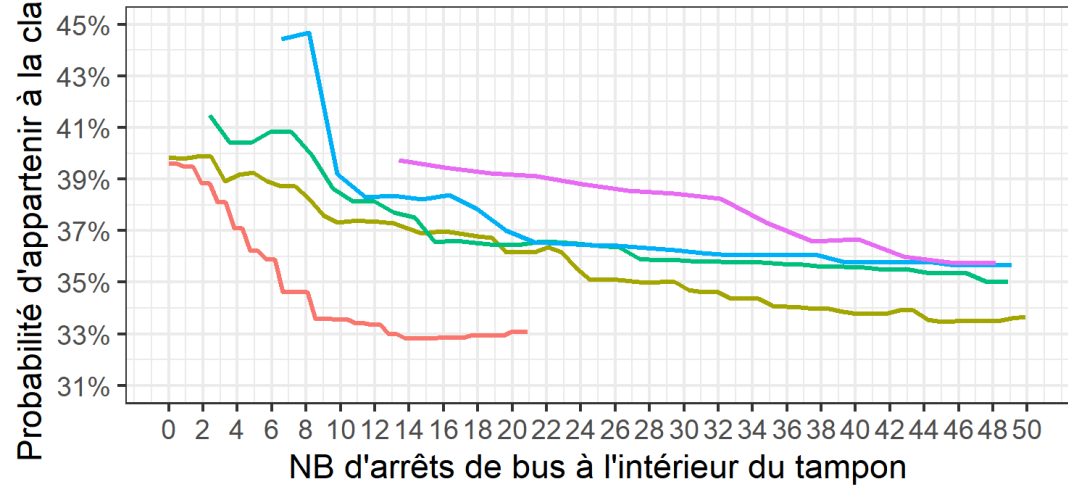
**B** Mén. 2+ adultes - C1: 0 véh.  
Var : Arrêts de bus, IR @ WD-1km = 3.11 %



**C** Mén. 2+ adultes - C2: < 1 véh. / adulte  
Var : Arrêts de bus, IR @ WD-1km = 3.11 %

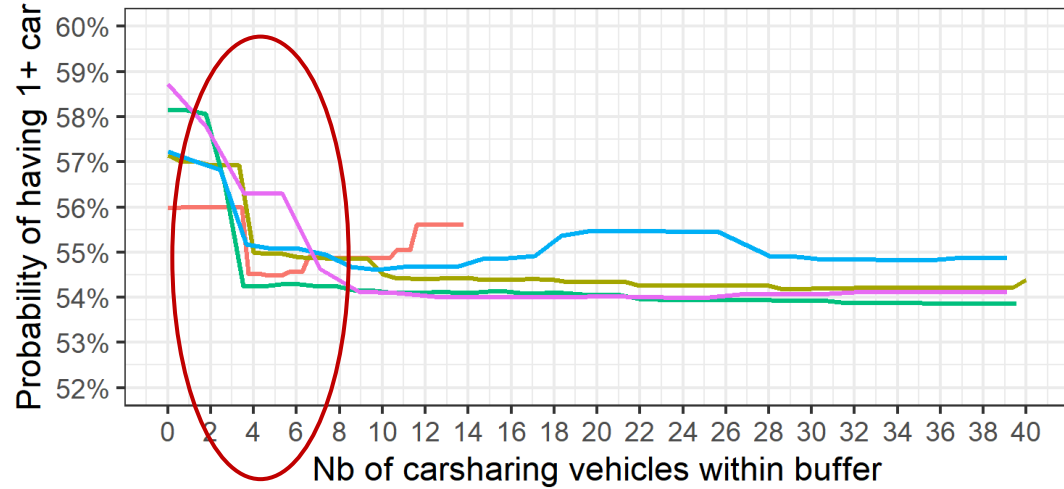


**D** Mén. 2+ adultes - C3: >= 1 véh. / adulte  
Var : Arrêts de bus, IR @ WD-1km = 3.11 %

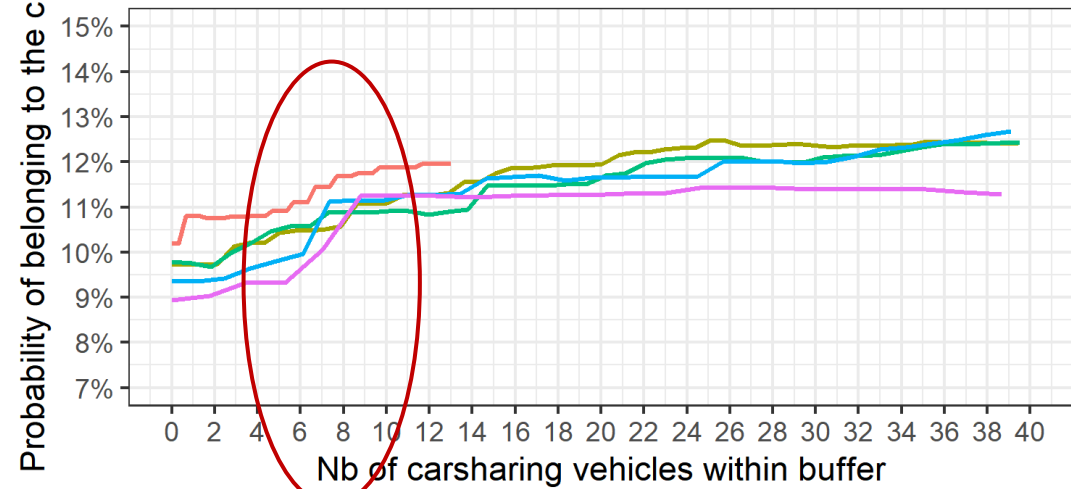


buffer size — wd400 — wd800 — wd1000 — wd1200 — wd1600

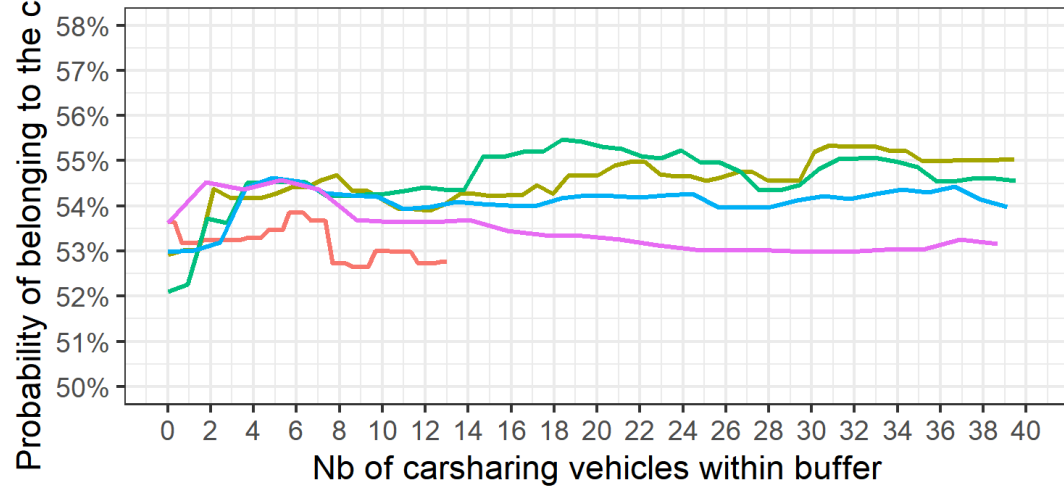
**A** 1adultHH - Outcome: 1+ car vs 0  
 Var : Nb of carsharing veh., RI = 6.57 %



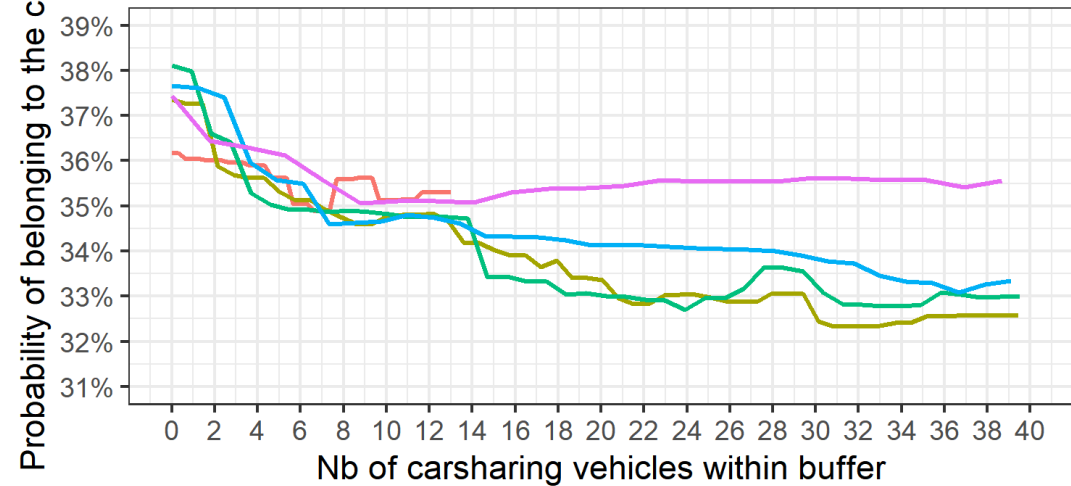
**B** Mén. 2+ adultes - C1: 0 véh.  
 Var : Nb de véh. Communauto., IR @ WD-1km = 6.4 %



**C** Mén. 2+ adultes - C2: < 1 véh. / adulte  
 Var : Nb de véh. Communauto., IR @ WD-1km = 6.4 %

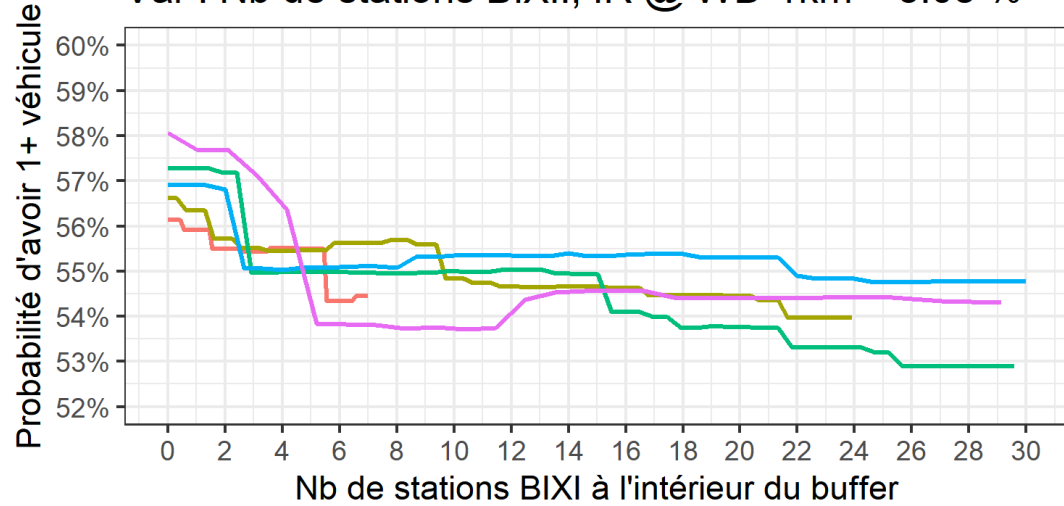


**D** Mén. 2+ adultes - C3: >= 1 véh. / adulte  
 Var : Nb de véh. Communauto., IR @ WD-1km = 6.4 %

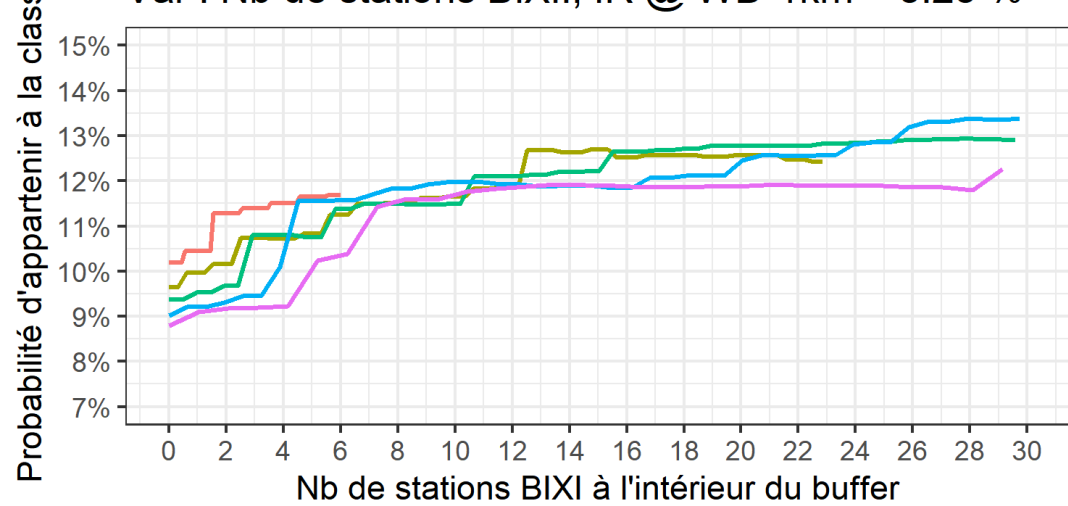


buffer size — wd400 — wd800 — wd1000 — wd1200 — wd1600

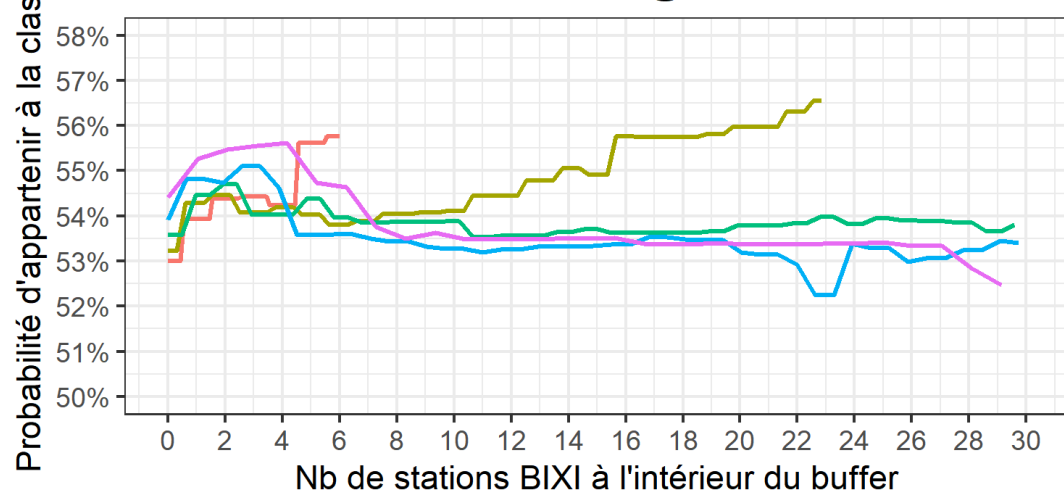
**A** Mén. 1 adulte : 1+ véh. vs 0  
 Var : Nb de stations BIXI., IR @ WD-1km = 3.08 %



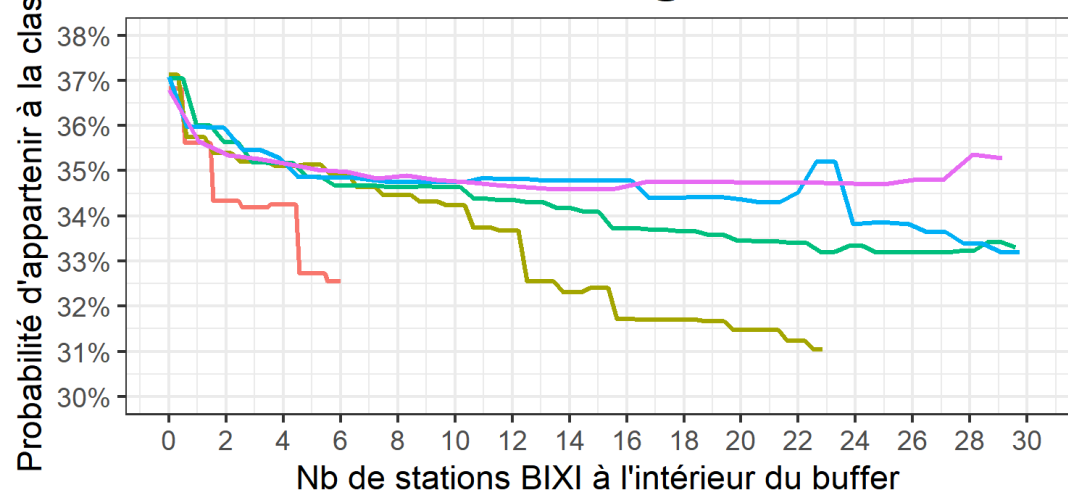
**B** Mén. 2+ adultes - C1: 0 véh.  
 Var : Nb de stations BIXI., IR @ WD-1km = 3.23 %



**C** Mén. 2+ adultes - C2: < 1 véh. / adulte  
 Var : Nb de stations BIXI., IR @ WD-1km = 3.23 %



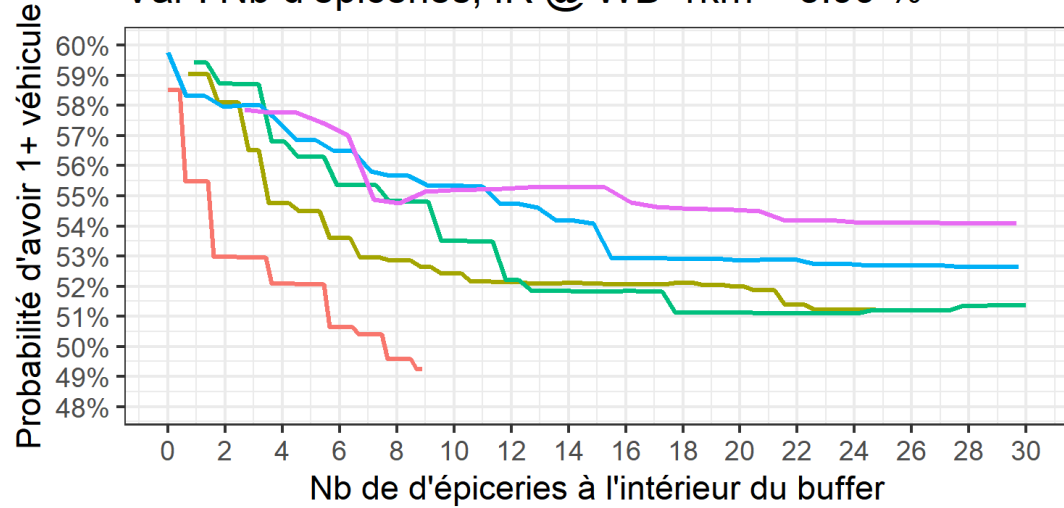
**D** Mén. 2+ adultes - C3: >= 1 véh. / adulte  
 Var : Nb de stations BIXI., IR @ WD-1km = 3.23 %



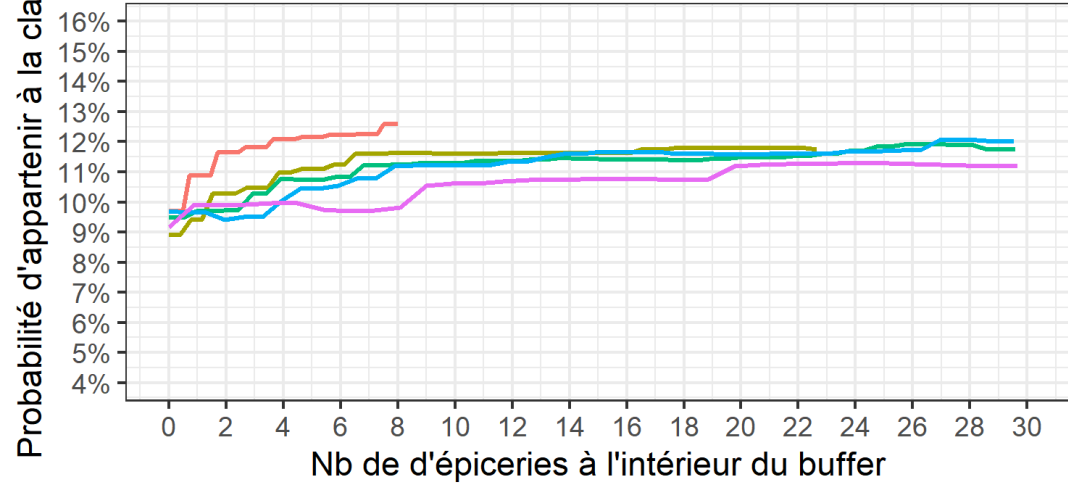
buffersize — wd400 — wd800 — wd1000 — wd1200 — wd1600



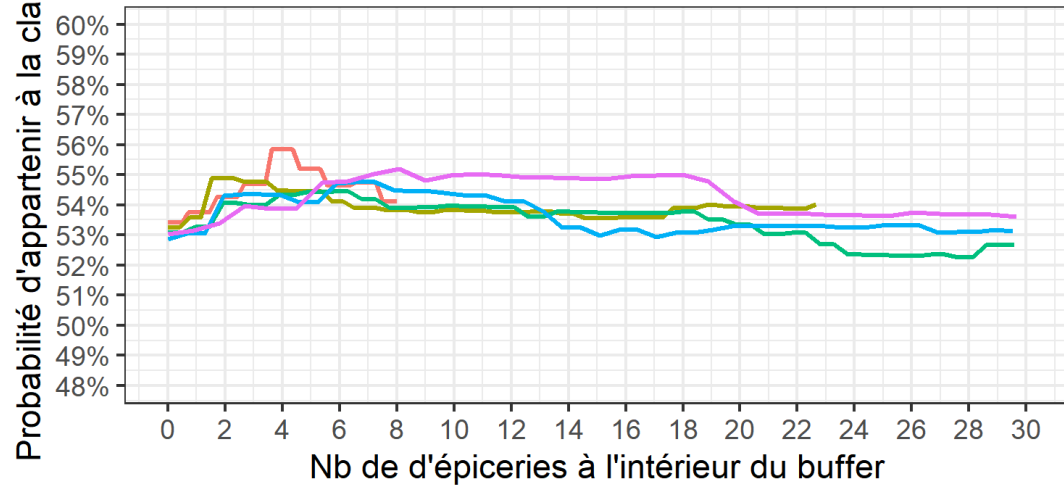
**A** Mén. 1 adulte : 1+ véh. vs 0  
Var : Nb d'épiceries, IR @ WD-1km = 5.99 %



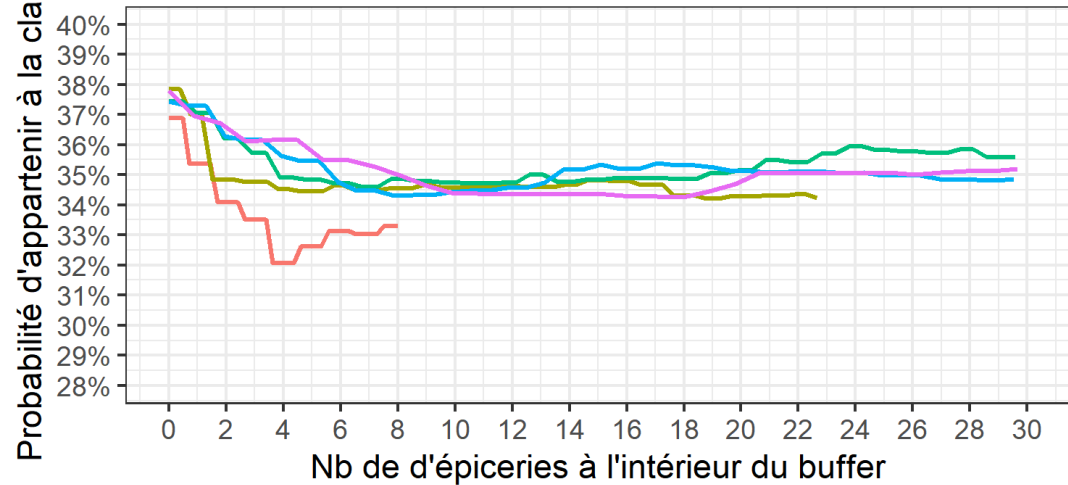
**B** Mén. 2+ adultes - C1: 0 véh.  
Var : Nb d'épiceries, IR @ WD-1km = 4.04 %



**C** Mén. 2+ adultes - C2: < 1 véh. / adulte  
Var : Nb d'épiceries, IR @ WD-1km = 4.04 %



**D** Mén. 2+ adultes - C3: >= 1 véh. / adulte  
Var : Nb d'épiceries, IR @ WD-1km = 4.04 %



buffer size — wd400 — wd800 — wd1000 — wd1200 — wd1600

## Sommaire des résultats : quoi retenir ?

### ▶ **Caractéristiques des quartiers:**

- ▶ **Densité de population**, **densité d'intersection 4-directions** (design) et **l'entropie** (mixité d'usage des sols)
- ▶ Effets relativement linéaires après ou avant une valeur seuil (ex: Densité, effet linéaire décroissant  $\leq 8500$  pers / km<sup>2</sup>)
- ▶ Ces variables sont moins sensibles à la taille des tampons utilisés

### ▶ **Accessibilité aux ressources de mobilité et opportunités commerciales**

- ▶ Variables très sensibles à la taille des tampons utilisées
- ▶ **Arrêts de bus** et **épicerie** : les plus petits tampons « performant » mieux et capturent mieux l'accès très locale à ces services.
- ▶ **Communauto** et **BIXI** : les tampons moyens (800-1000 m) affichent des changements de probabilité importants.  
Les plus gros tampons (1600 m) ont une influence relative + importante.  
MAIS semblent agir comme « proxies » pour les quartiers centraux.



## Sommaire des résultats : effets structurants

Variable	Résumé effet
Densité de population	Effet ~Linéaire jusqu'à 6500-8500 p/km <sup>2</sup> (tous les tampons) Pas d'effet additionnel après
Dens. Intersections 4-dir. (design)	Probabilité minimale atteinte à 12 int. / km <sup>2</sup> (tous les tampons)
Entropy (mixité d'usage)	Effet ~Linéaire entre 0.25 et 0.75
Distance au CV	↑ rapide de prob. motorisation entre 3 et 5 km. ↑ Lente jusqu'à de 5 à 14 km. Puis stable
Distance au Métro	↑ rapide de prob. motorisation entre 0 et 4.5 km.
Arrêts de bus	Tampon 400m : ↓ rapide de prob. motorisation entre 0 et 20 arrêts (1adulte) et entre 0 et 12 arrêts (2+ adultes)
Véhicules Communauto	1adulte : effet « seuil » (↓ rapide) à 4 véhicules 2+adultes : ↓ rapide de motorisation de 0 et 14 véhicules @ 800 m
Stations BIXI	1adulte : effet « seuil » (↓ rapide) entre 3-5 stations 2+adultes : ↓ rapide de motorisation de 0 et 12 stations @ 800 m
Épiceries	↓ rapide de prob. motorisation entre 0 et 4 épiceries @ 400 m

## Méthode, limites et recherches futures

- ▶ Illustre le potentiel des algos de Gradient Boosting Machine pour comprendre les comportements de mobilité
- ▶ **Limitations:**
  - ▶ Influence du revenu contrôlé, mais demanderait davantage d'analyse (effets de l'environnement bâti est très différents selon le revenu)
  - ▶ Les modèles ne font qu'analyser les corrélations. Pas possible d'évaluer l'effet causal.
  - ▶ Pas de contrôle pour l'auto-sélection résidentielle des ménages
  - ▶ Les corrélations entre les variables limitent l'interprétation des résultats, surtout dans les quartiers centraux où les opportunités sont fortement corrélées et se cumulent vite
- ▶ **Prochaines étapes:**
  - ▶ Valider les valeurs ayant des effets structurants
  - ▶ Valider les « distances seuils » pour atteindre certaines valeurs (ou nb d'opportunités)
  - ▶ Comprendre les effets combinés des variables
  - ▶ Scénarios d'impacts sur la motorisation

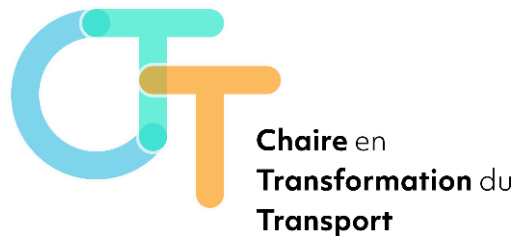


# Remerciements

L'article sera présenté le 10 janvier 2022 au Congrès Annuel du Transportation Research Board:

Lavolette, J., Morency, C., & Waygood, E. O. D. (2022). *A kilometer or a mile? Does buffer size matter when it comes to car ownership?* Communication présentée au 2022 Transportation Research Board Annual Meeting, Washington DC.

Article soumis pour publication au Journal of Transport Geography en octobre 2021.



**Fonds de recherche  
Nature et  
technologies**

**Québec** 

**POLYTECHNIQUE  
MONTREAL**

UNIVERSITÉ  
D'INGÉNIEURIE



**ARTM**

Autorité régionale  
de transport métropolitain

**CARGOM**

GRAPPE MÉTROPOLITAINE DE  
LOGISTIQUE ET TRANSPORT DE MONTRÉAL



**CIRRELT**

**COOP CARBONE**

**exo**

**IET** INSTITUT  
DE L'ÉNERGIE  
**TROTTIER**



**IVADO**

**Québec** 



Chaire en  
Transformation du  
Transport

# Références

- Lavolette, J. (2020). L'état de l'automobile au Québec : constats, tendances et conséquences. Fondation David Suzuki. Tiré de <https://fr.davidsuzuki.org/publication-scientifique/letat-de-lautomobile-au-quebec-constats-tendances-et-consequences/>
- Gärling, T., & Steg, L. (2007). Threats from car traffic to the quality of urban life: problems, causes and solutions. Bingley, West Yorkshire: Emerald Group Publishing Limited.
- Sioui, L., Morency, C., & Trépanier, M. (2012). How Carsharing Affects the Travel Behavior of Households: A Case Study of Montréal, Canada. *International Journal of Sustainable Transportation*, 7(1), 52-69. <https://doi.org/10.1080/15568318.2012.660109>
- von Behren, S., Bönisch, L., Niklas, U., & Chlond, B. (2020). Revealing Motives for Car Use in Modern Cities—A Case Study from Berlin and San Francisco. *Sustainability*, 12(13). doi: <https://doi.org/10.3390/su12135254>
- Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *The Annals of Statistics*, 29(5), 1189-1232. doi: <https://doi.org/10.1214/aos/1013203451>
- Friedman, J. H. (2002). Stochastic gradient boosting. *Computational Statistics & Data Analysis*, 38(4), 367-378. [https://doi.org/10.1016/S0167-9473\(01\)00065-2](https://doi.org/10.1016/S0167-9473(01)00065-2)
- Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: data mining, inference, and prediction*: Springer Science & Business Media.
- Chen, T., & He, T. (2021). xgboost: eXtreme Gradient Boosting. Tiré de <https://cran.r-project.org/web/packages/xgboost/xgboost.pdf>





Questions ?

[jerome.laviolette@polymtl.ca](mailto:jerome.laviolette@polymtl.ca)



Chaire en  
Transformation du  
Transport