

Extrapolation des flux vélo



Programme

1. Pourquoi l'extrapolation ?
2. Les données utilisées
3. Méthodologie
 - a. Préparation et intégration des données
 - b. Résultats
4. Conclusion

Pourquoi l'extrapolation ?

Qu'est-ce que l'extrapolation ?

Méthode pour estimer le nombre de passages réels de cyclistes en tout point du réseau cyclable

Cela comprend :

- Un process d'analyse et de calibrage des données existantes,
- Un algorithme pour extrapoler les données à partir de données terrain et contextuelles,
- Des itérations pour calibrer le modèle.

Quels usages de l'extrapolation ?



Pilotage stratégique

- Identification des usages réels
- Analyse du maillage cyclable effectif
- Aide au dimensionnement (stationnement, services, etc.)



Valorisation & communication

- Visualisation cartographique
- Indicateurs pour bilans et élus
- Soutien aux demandes de subventions



Aide à l'aménagement

- Priorisation des axes à traiter
- Justification des investissements
- Évaluation des impacts d'aménagements



Suivi & évaluation

- Fréquentation par secteur/période
- Comparaison avant/après travaux
- Complément ou substitution aux capteurs fixes

La cible

Fréquentation des axes

Découvrez la répartition du trafic vélo sur votre territoire, pour mieux cibler les futurs aménagements cyclables.

< mars 2025 >

Sécurité des voies empruntées



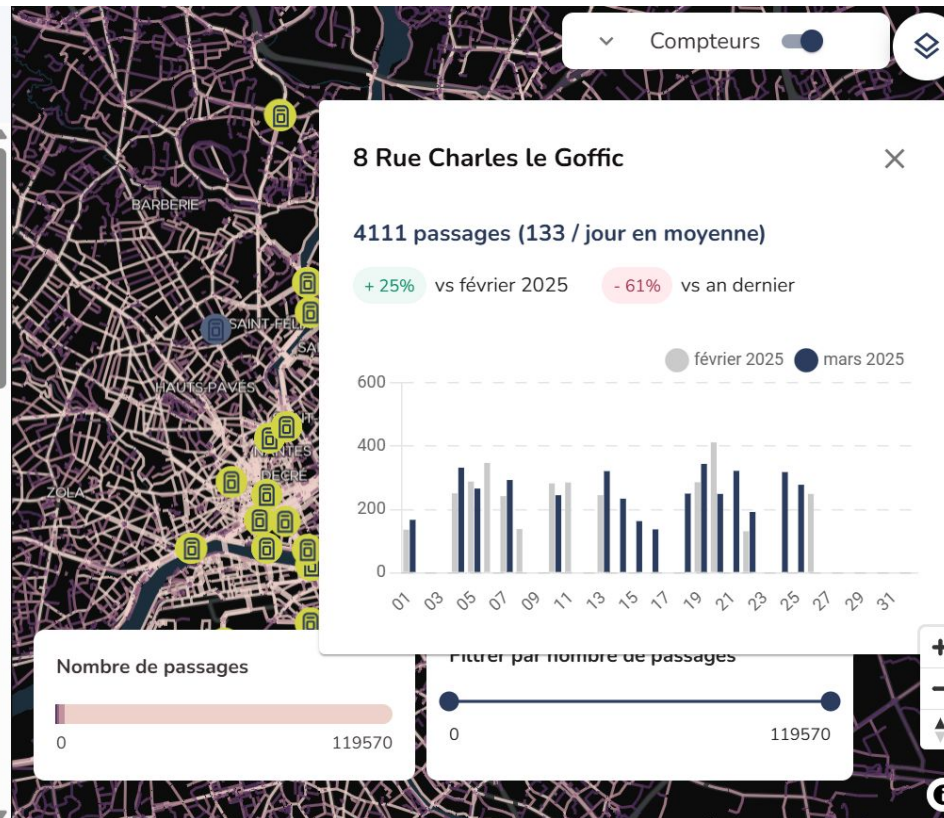
Aménagements cyclables

Autres aménagements
23.4 %

Voies non aménagées
54.5 %

+ 1 056 631 km vs mois précédent

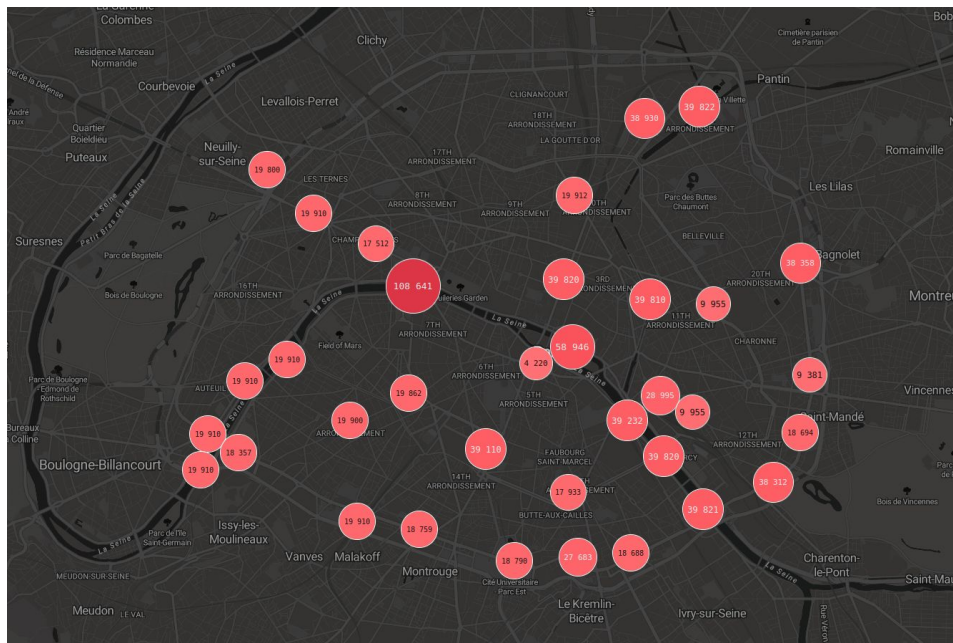
- 1 610 895 km vs an dernier



Les données utilisées

A quel type de données a-t-on accès ?

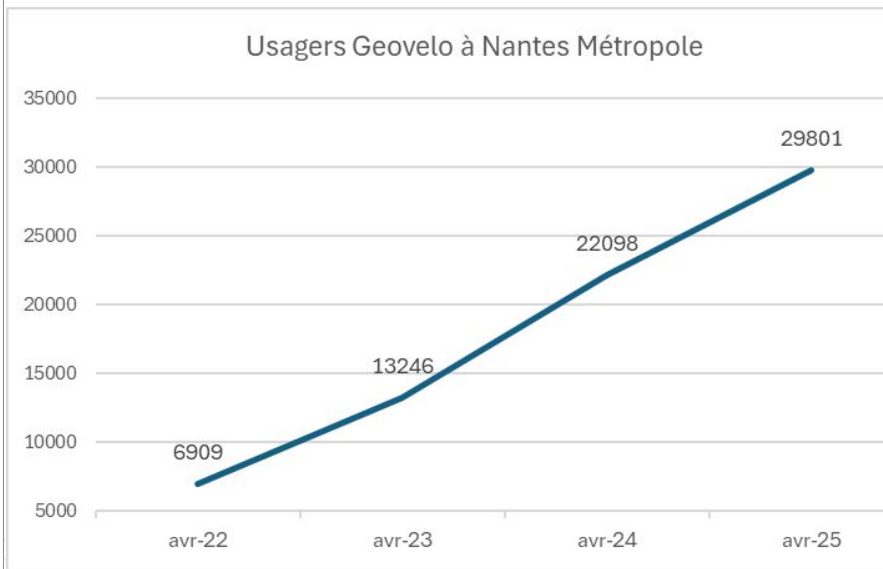
Données de comptage physique



- Comptage horaire
- Date et heure de comptage
- Coordonnées géographiques

A quel type de données a-t-on accès ?

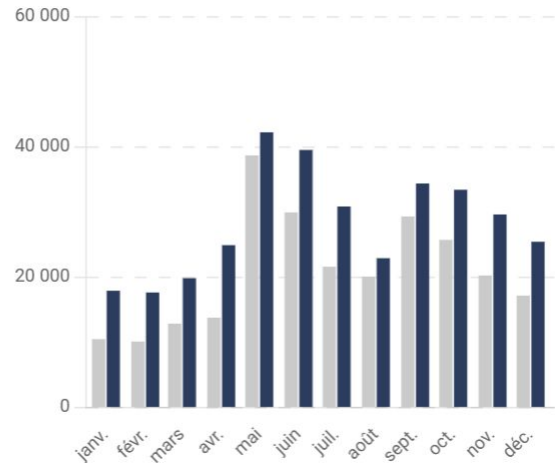
Traces GPS Geovelo



Trajets enregistrés (+35%)

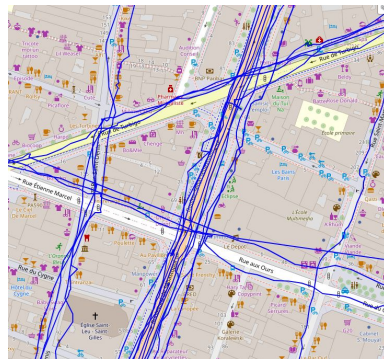
2024
338 904 trajets
enregistrés

2023
250 350 trajets
enregistrés



A quel type de données a-t-on accès ?

Traces GPS Geovelo mapmatchées



Trace GPS	Trace mapmatchée	Exemple de segment

A quel type de données a-t-on accès ?

Données externes :

- Insee (Population - Base infracommunale)
- Météo (température, précipitations)
- OpenStreetMap (type de voie, type d'aménagement...)



Méthodologie

Méthodologie

1. Préparation & intégration des données (Mois 1)

- Réunion de lancement avec les parties prenantes
- Collecte et qualification des données :
 - Données de comptage (Eco-Compteurs, Wintics, manuels...)
 - Données météo, socio-démo (IRIS), carto (OSM), traces Geovelo
- Recalage géographique des compteurs
 - Vérification de la position exacte sur le réseau (souvent imprécise ou erronée)
 - Corrélation avec la voie réelle empruntée par les cyclistes (OSM)
 - Correction du sens de comptage, si besoin
 - Essentiel pour assurer un bon match compteur ↔ flux Geovelo

2. Modélisation initiale & tests (Mois 2 à 3)

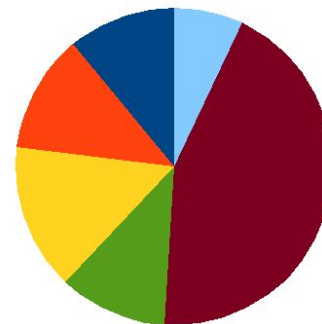
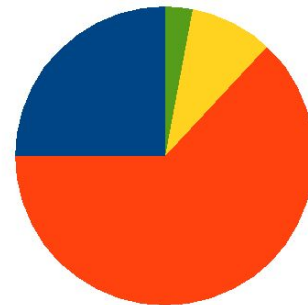
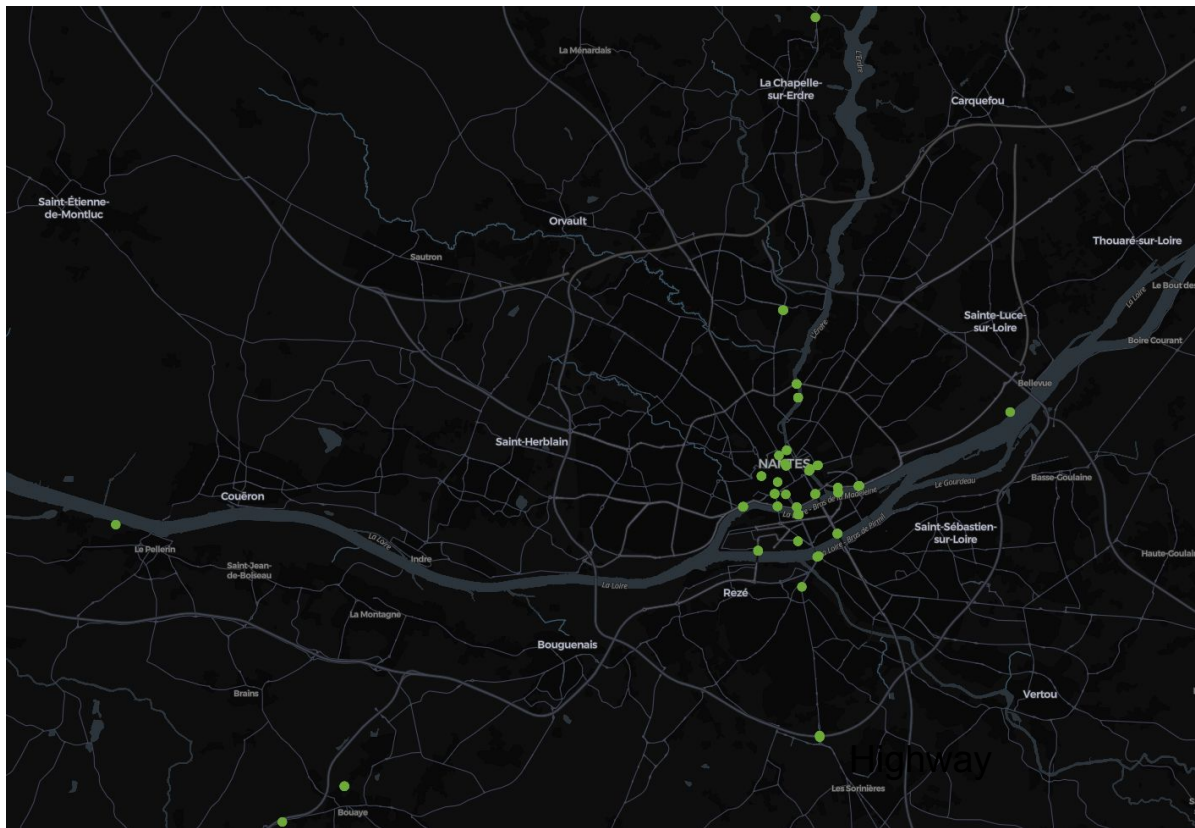
- Développement du modèle d'extrapolation
- Test de corrélation et de cohérence avec compteurs bien localisés
- Ajustement des facteurs d'échelle et identification des biais (flux faibles, zones périphériques...)

3. Itérations & optimisation (Mois 4 à 5)

- Amélioration continue via apprentissage sur données réelles (données de comptage de validation)
- Cartes de fiabilité des prédictions et étiquetage des zones à forte incertitude
- Réunions mensuelles de suivi avec livrables intermédiaires

Préparation et intégration des données

Compteurs Nantes ~ 59 compteurs



Localisation hyper centre, peu de données en périphérie

Correction de la localisation (10 corrigés)

Route des Sorinières CV0959 47.16, -1.54
Route des Sorinières CV0960

Libération CV0950 47.19, -1.54
Libération CV0949

Guy Mollet CV1042 47.25, -1.55
Guy Mollet CV10471 47.25, -1.55

Strasbourg CV9998 47.21, -1.55

Trémie CV0986 47.21, -1.54
Trémie CV0987 47.21, -1.54

Malakoff CV0679 47.21, -1.53

Correction du sens (16 corrigés)

Route des Sorinières CV0959 backward
Route des Sorinières CV0960 forward

Libération CV0950 forward
Libération CV0949 backward

La Chapelle-sur-Erdre X2H18014267 both

Trois Continents vers sud - CV0979 both
Trois Continents vers nord - CV0977 backward

Guy Mollet CV10471 forward
Guy Mollet CV1042 forward

50 Otages Nord - CV0785 et CV0786 both
50 Otages Sud - CV0787 et CV0788 both

Trémie CV0986 forward
Trémie CV0987 backward

Calvaire CV0744 forward
Calvaire CV0745 backward

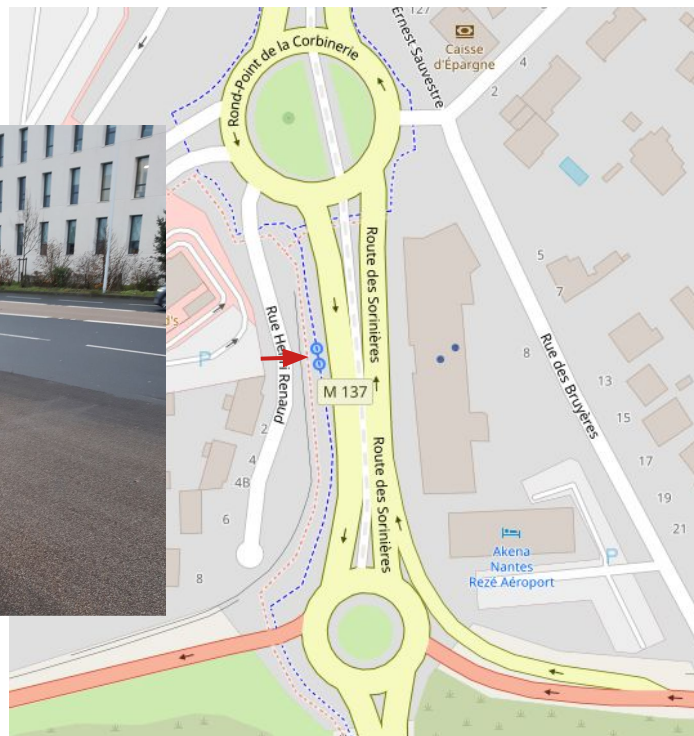
Malakoff CV0679 forward

Exemple de corrections

Avant



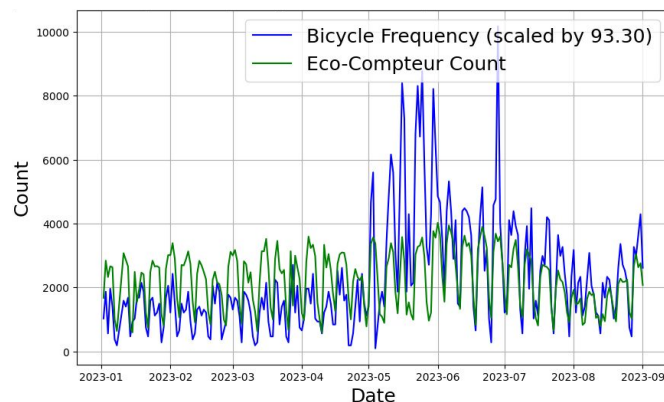
Après



Résultats

1. **Représentativité Geovelo** : calculée en comparant les **données réelles** aux données de **fréquentations Geovelo** sur une période.
2. **Représentativité Extrapolée** : Calculée en comparant les **données réelles** aux données de **fréquentations extrapolées** sur une période.

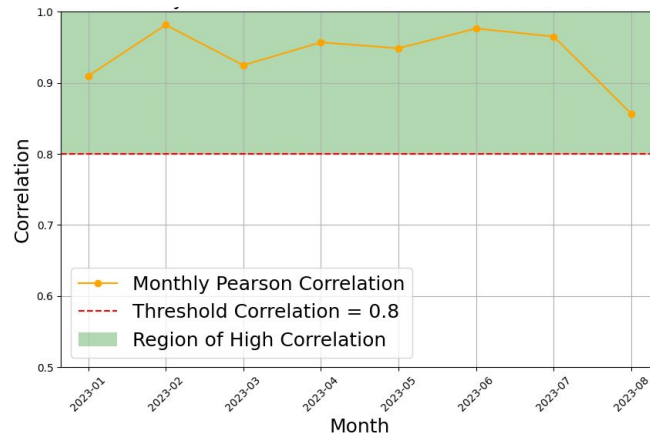
Facteur d'échelle



Pour les données Geovelo, plus le facteur d'échelle est faible, plus le nombre de cyclistes estimé par Geovelo se rapproche du nombre réel de cyclistes.

Pour les données extrapolées, la valeur idéale est 1.

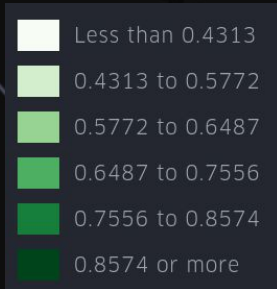
Corrélation de Pearson



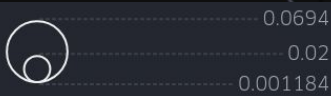
Mesure la force de la relation entre les deux séries temporelles.

Une valeur proche de 1 signifie qu'elles ont des tendances similaires, indépendamment des valeurs réelles.

Pearson Correlation

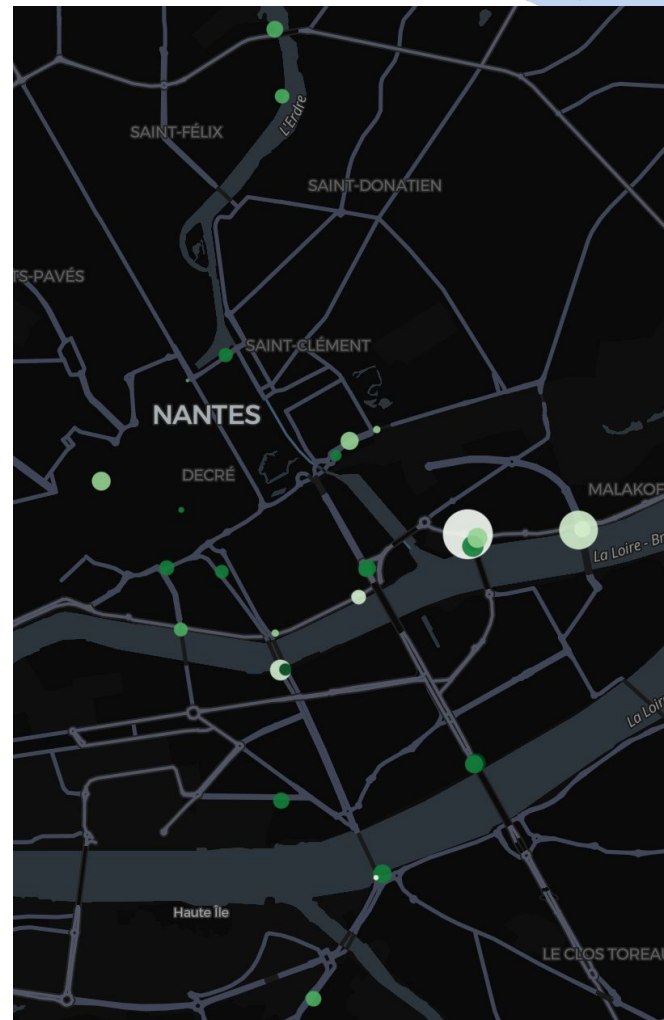


Inverse scaling factor

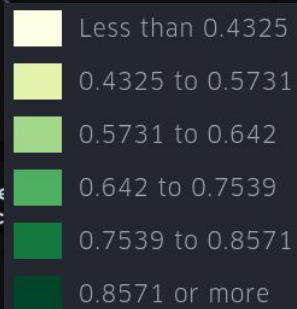


représentativité Geovelo

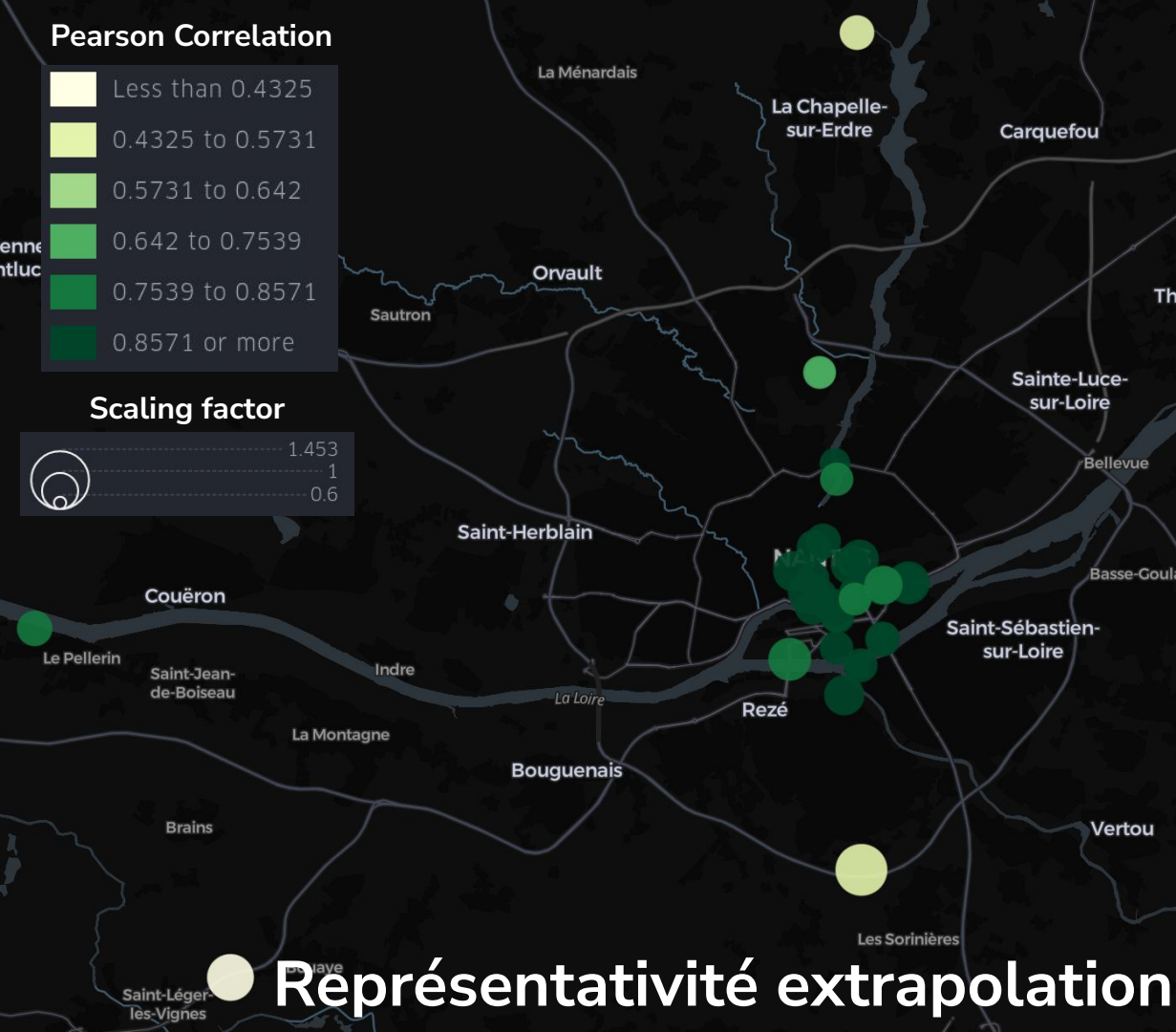
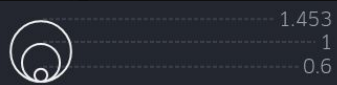
Une représentativité élevée n'est pas nécessaire pour avoir une corrélation forte



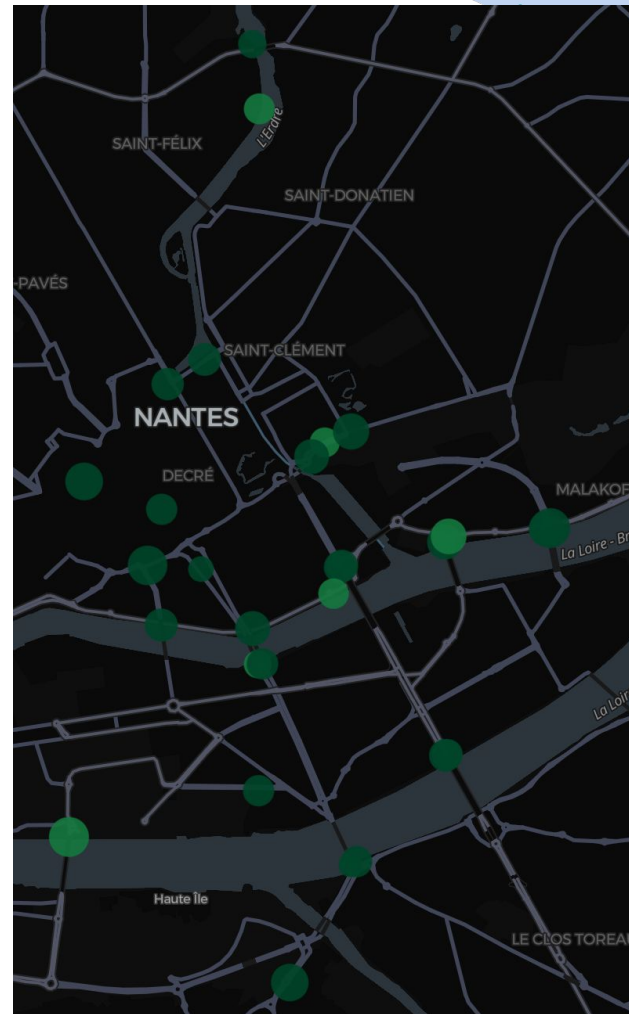
Pearson Correlation



Scaling factor



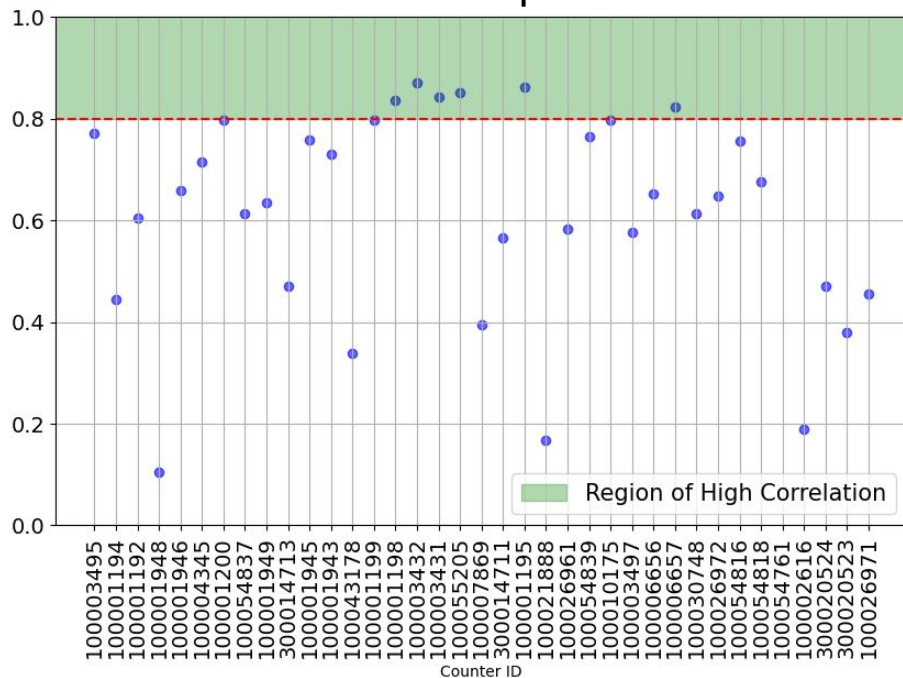
● Représentativité extrapolation



Pearson correlation

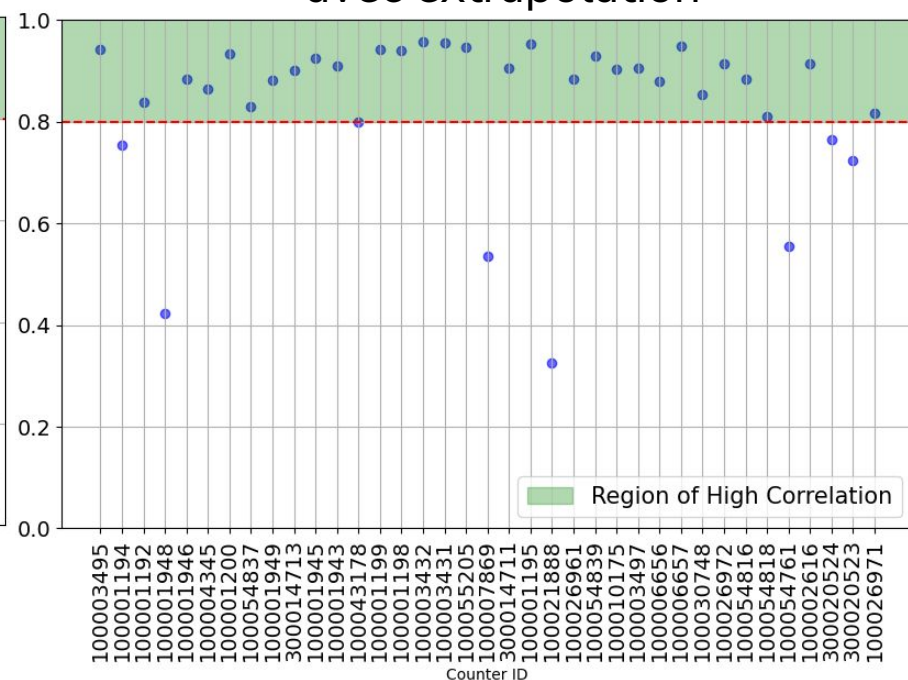
Avant les corrections 0 % sur la Region of High Correlation
Après les corrections 23 % sur la Region of High Correlation

sans extrapolation



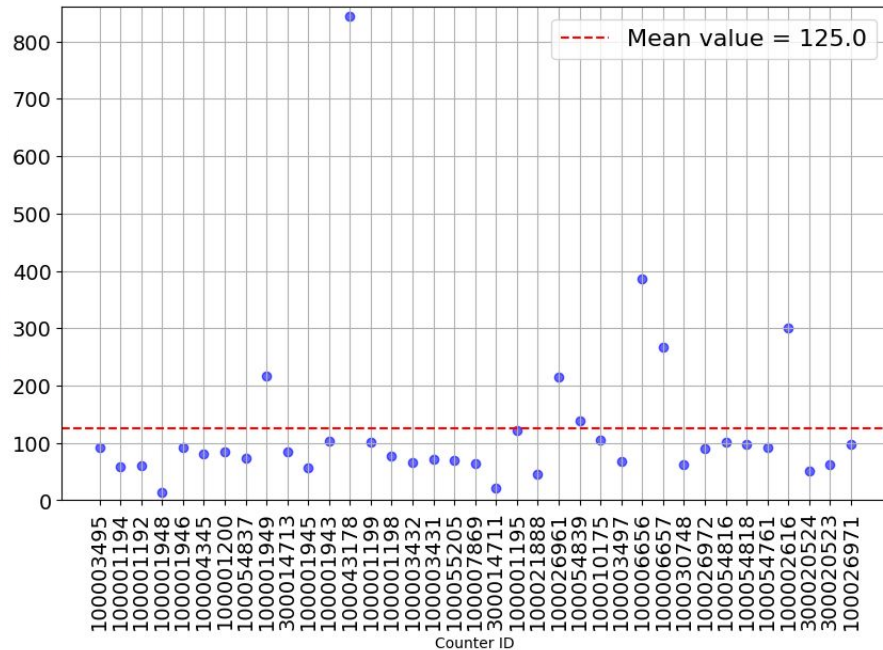
Avant les corrections 2.56% sur la Region of High Correlation
Après les corrections 81.08 % sur la Region of High Correlation

avec extrapolation

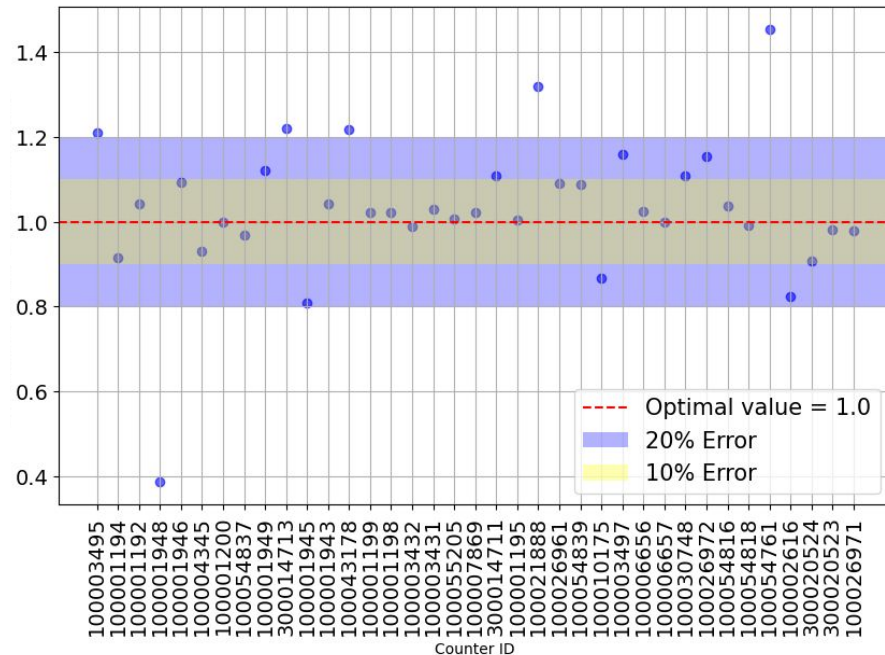


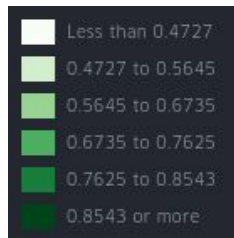
Scaling factor

sans extrapolation



avec extrapolation

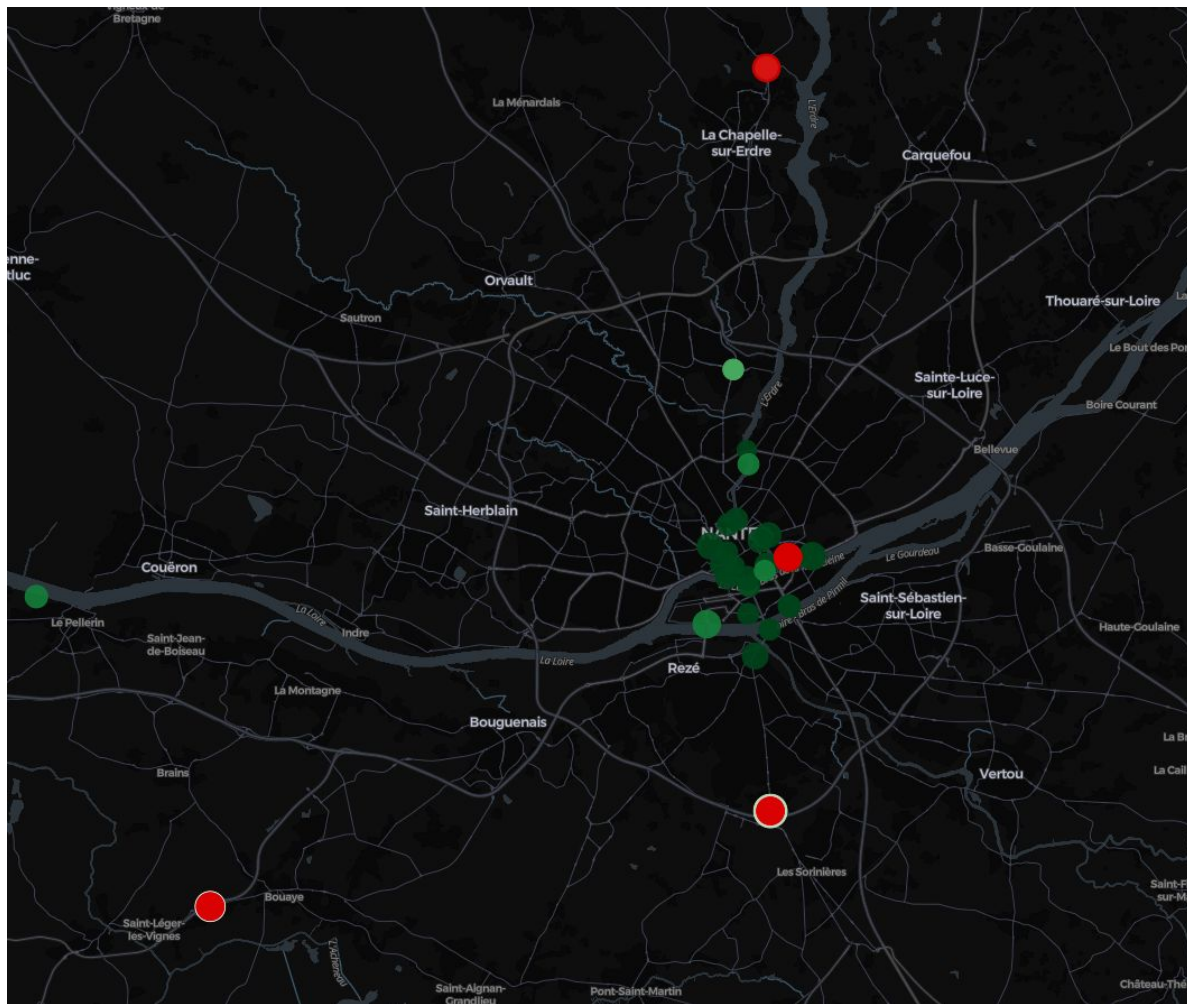




Points rouges (compteurs avec de mauvais résultats)

'scaling_factor_extrapolation' > 1.25
'scaling_factor_extrapolation' <= 0.7
'pearson_correlation_extrapolation' <= 0.6

La plupart des compteurs se
trouvent loin du centre-ville.



Pourquoi ?

counter_id	pearson_correlation extrapolation	scaling_factor extrapolation	pearson_correlation geovelo	scaling_factor geovelo
100007869	0.53	1.02	0.39	63.27
100021888	0.32	1.31	0.16	45.4
100001948	0.42	0.38	0.10	14.4
100054761	0.55	1.45	-0.05	91.78



<https://api.panoramax.xyz/#focus=pic&map=17.95/47.213226/-1.536722&pic=868e3856-6e9e-4b5e-930a-1e410b673cfb&speed=250&xyz=58.20/-49.38/0>

Classement des fréquences réelles, en mettant en avant les compteurs ayant les fréquences les plus basses

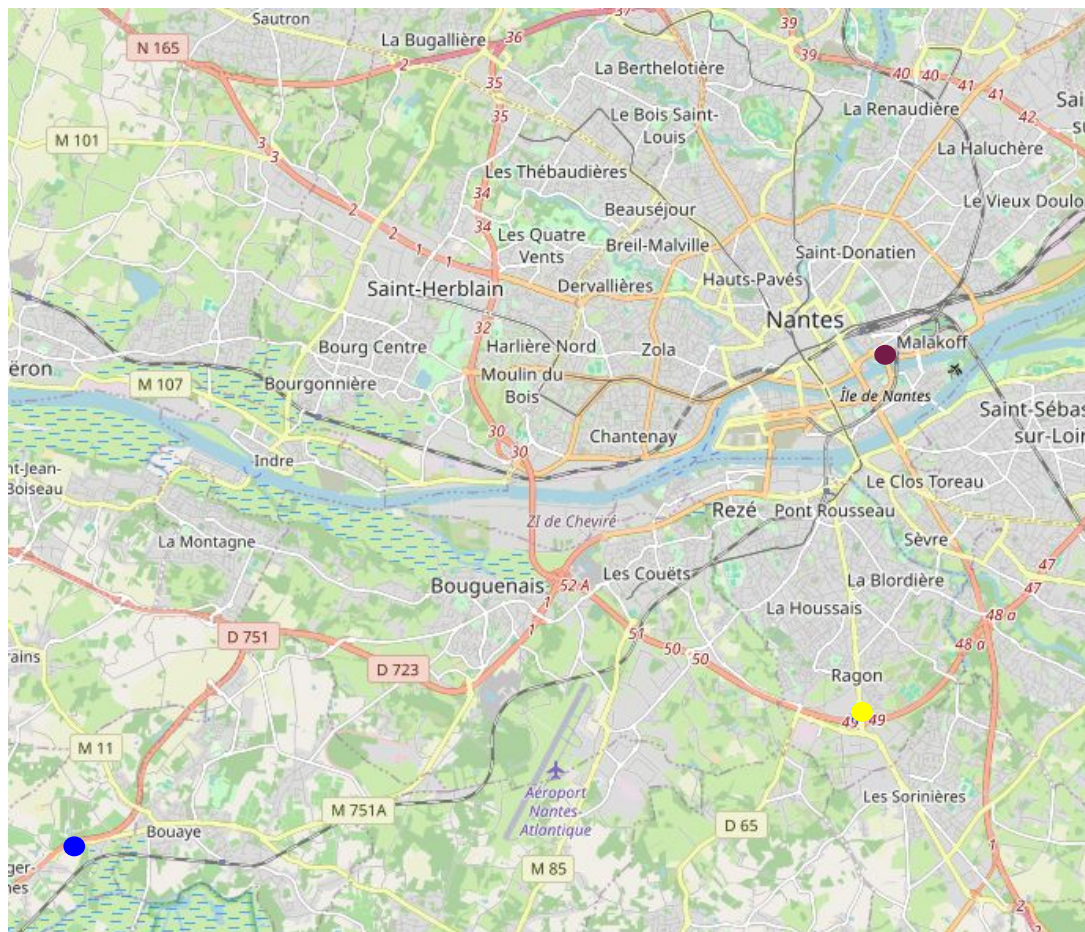
Counter ID

100001948

100021888

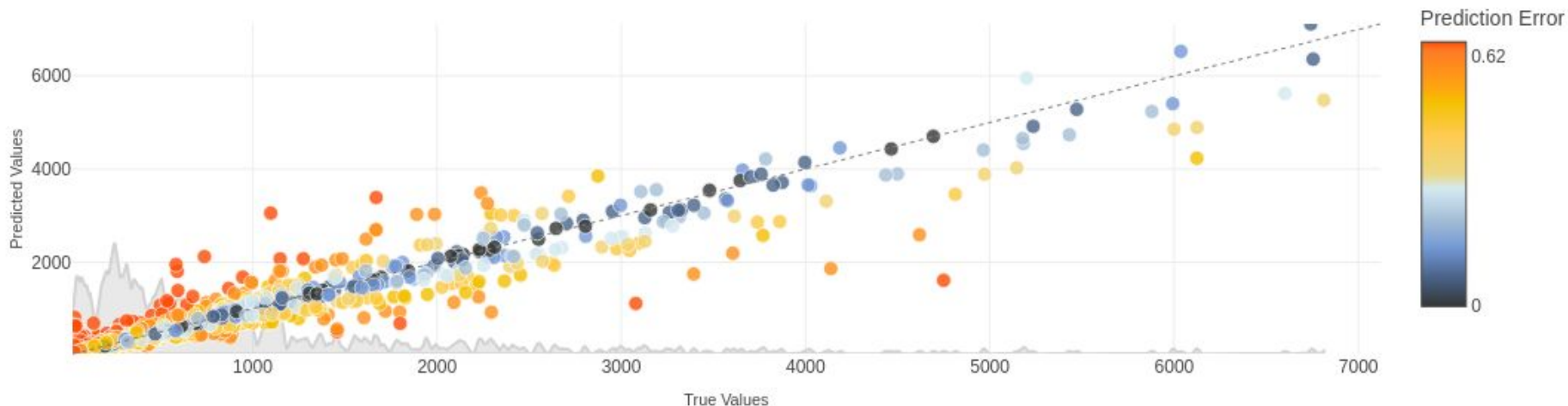
100054761

100054815



True Values vs Predicted Values

$$\text{Predicted Error} = \frac{\text{abs}(\text{True Values} - \text{Predicted Values})}{\text{True Values}}$$



- **Axe des X (Valeurs Réelles)** : Ce sont les vraies valeurs issues du jeu de données.
- **Axe des Y (Valeurs Prédites)** : Ce sont les valeurs estimées par le modèle.

Prédictions parfaites :

- Si le modèle était **parfait**, tous les points seraient situés **exactement sur la diagonale** ($y = x$).
- Cela signifie que les valeurs prédites correspondent **exactement** aux valeurs réelles.

Conclusion & prochaines étapes

Conclusion : très bons résultats de l'extrapolation

Le calibrage d'un modèle d'extrapolation a permis plusieurs choses pour la collectivité :

- **Analyse de la représentativité** des données Geovelo selon les compteurs,
- Identification de **compteurs mal positionnés** et repositionnement de ceux-ci dans les données,
- Accès à de **premiers résultats robustes** en zone urbaine (> 90%), permettant d'envisager un déploiement à plus grande échelle

Challenges :

- Avoir de meilleurs résultats hors agglomération,
- Passer à l'échelle

Un **modèle d'extrapolation innovant** et destiné aux applications terrain, présentant des résultats prometteurs tant sur les résultats techniques que sur les perspectives d'**efficacité de l'action publique** (besoin de moins de compteur, couverture d'une zone plus dense de connaissance terrain, ...).

Prochaines étapes

- Proposition de zones d'implantation de nouveaux compteurs
- Produire des cartes de fiabilisation des prédictions
- Améliorations techniques de l'algorithme

Contact

Ronan Bouquet, CEO

06 70 17 34 34

ronan.bouquet@geovelo.fr



