

Modéliser le parc automobile francilien et ses émissions de polluants : redressement d'une population synthétique à partir de données marginales

Organismes d'accueil : laboratoires [SPLOTT](#) et [LVMT](#)

Profil recherché : étudiant·e en école d'ingénieur ou master, spécialisé·e en mathématiques appliquées, informatique ou statistiques

Localisation : le stagiaire pourra être localisé dans un des deux laboratoires

- Université Gustave Eiffel (Campus de Nantes, allée des ponts et chaussées, Bouguenais)
- Laboratoire Ville Mobilité Transport (Marne-la-Vallée, RER A Noisy-Champs)

Quelle que soit la localisation principale, le stagiaire pourra être amené à se rendre ponctuellement sur l'autre site pour échanger avec les équipes sur place.

Période : début à partir de mars 2023 (flexible), durée ~ 5/6 mois à temps plein

Rémunération : environ 600 € par mois (variable selon le nombre de jours travaillés dans le mois)

Contexte

Alors que les États européens multiplient les mesures de régulation du parc automobile afin d'améliorer leur qualité de l'air en ville, il est essentiel de comprendre les déterminants de la possession de voitures et du choix du type de véhicule par les ménages pour évaluer les effets sociaux et environnementaux de ces mesures. En effet, l'Union Européenne a adopté des normes d'émissions de polluants pour les constructeurs automobiles et de nombreuses métropoles européennes ont également mis en place des zones à faibles émissions (ZFE) limitant le trafic routier dans une zone restreinte, principalement en fonction des émissions des véhicules. Les modèles de mobilité multi-agents sont ainsi largement utilisés aujourd'hui pour évaluer les ZFE et leurs impacts. La représentation d'un parc de véhicules désagrégé au niveau des ménages fournirait donc de meilleures entrées pour les modèles de mobilité, avec l'objectif de calculer les émissions quotidiennes de polluants liées au trafic routier sur la base d'une population et d'un parc de véhicules synthétiques.

Le contexte de ce travail de recherche est l'utilisation et la production de données ouvertes pour aider le décideur à aboutir à des choix éclairés en matière de politiques environnementales. En effet, les chercheurs et praticiens ont, depuis peu, accès facilement à plusieurs sources de données permettant une meilleure connaissance des ménages : recensement de la population, enquête ménage déplacement, base de données fiscales, bases de données véhicules (Crit'air).

Méthode et objectifs

L'objectif final est de participer à la construction d'indicateurs pertinents pour la prise de décision. Ces indicateurs sont construits à partir d'une population synthétique composée d'agents avec des attributs permettant de calculer les indicateurs importants pour le décideur. L'ajout d'attributs pertinents à cette population en fonction des politiques visées est crucial

pour ensuite construire des indicateurs. Cette étape est le cœur du stage proposé : générer une population synthétique au niveau d'une zone administrative et lui ajouter des attributs environnementaux sur ses véhicules.

En pratique, des modèles de machine learning visant à attribuer à chaque ménage un véhicule avec ses caractéristiques (carburant, norme euro d'émission de polluants) pour une population donnée ont déjà été produits par l'équipe encadrante du LVMT à partir d'enquêtes ménage déplacement. L'objectif était d'attribuer les voitures aux ménages en fonction de leurs caractéristiques socio-économiques et de mobilité. Il s'agit à présent pour ce stage de mobiliser des méthodes statistiques afin d'ajuster ce parc automobile à l'échelle régionale et communale, afin qu'il soit également cohérent avec les statistiques agrégées de type de Crit'air possédés par exemple.

Pour ce faire, des méthodes statistiques telles que l'iterative proportional fitting procedure (IPF) (Yameogo et al, 2021) et le Penalized Maximum Entropy Dasymetric Modeling (P-MEDM) (Ieyk and al, 2013) seront mobilisées. Les méthodes et outils mobilisés dans le cadre de ce stage ont été déjà utilisés (ou développés) par l'équipe encadrante de Splott dans le cadre de travaux scientifiques sur la génération de population synthétique à partir de diverses données.

Déroulement du stage :

Les différentes tâches du stage sont les suivantes :

- Étude bibliographique sur les méthodes de redressement des données à partir de données marginales et prise en main des outils existants : pipeline Eqasim [Horl et al] et la librairie python Bhepop2
- Etudes des données ouvertes disponibles : Enquête Ménage Déplacement, base de données Crit'air du point de vue de leurs cohérences, de leurs incertitudes, etc.
- Développement d'un algorithme pour redresser l'association des véhicules aux ménages
- Etudes des performances de l'algorithme
- Insertion de l'algorithme dans le Pipeline Eqasim

Écosystème du stage :

Le stage se déroulera au sein du laboratoire [AME-Splott](#) de l'université Gustave Eiffel (UGE), situé sur le campus de Nantes, en collaboration avec le [LVMT](#), laboratoire de l'UGE et de l'École des Ponts ParisTech situé à Marne-la-Vallée en Île-de-France. Le stage impliquera des déplacements au LVMT qui seront financés (transport et hébergement inclus).

Le projet présenté s'intègre notamment dans le projet scientifique de Splott dont trois thèses sont sur des thématiques connexes. Les études proposées dans ce stage bénéficieront des synergies avec ces trois thèses.

Merci d'adresser vos candidatures (cv et lettre de motivation) à :

pierre-olivier.vandanjon@univ-eiffel.fr ; marjolaine.lannes@enpc.fr

Références :

Hörl, Sebastian, and Milos Balac. "Synthetic population and travel demand for Paris and Île-de-France based on open and publicly available data." *Transportation Research Part C: Emerging Technologies* 130 (2021): 103291.

Lannes Marjolaine, Coulombel Nicolas, and Roustan Yelva. "Car fleet synthesis for agent-based mobility models: a comparison of machine learning and discrete choice methods." (2022) <https://www.ivt.ethz.ch/en/info/news/2022/05/31-matsim-user-meeting.html>

S. Leyk, N. N. Nagle, and B. P. Battenfield. Maximum entropy asymmetric modeling for demographic small area estimation. *Geographical Analysis*, 45(3) :285–306, 2013

Yaméogo, Boyam Fabrice, Vandanjon, Pierre Olivier, Gastineau, Pascal, Hankach, Pierre, 2021, Generating a two-layered synthetic population for French municipalities: Results and evaluation of four synthetic reconstruction methods, *JASSS-Journal of Artificial Societies and Social Simulation*, 24, 27p, DOI: 10.18564/jasss.4482

Yaméogo, Boyam Fabrice, Vandanjon, Pierre-Olivier, Pierre Hankach, and Pascal Gastineau. "Methodology for Adding a Variable to a Synthetic Population from Aggregate Data: Example of the Income Variable." (2021)

https://hal.science/hal-03282111/file/Article_methodo_allocation_revenu.pdf