

STAGE DE MASTER :

Fiabiliser et enrichir les fichiers SIRENE pour la simulation du transport de marchandises en ville

Le LAET propose depuis 2000 un modèle de simulation des flux de marchandises en ville conçu sur la base de vastes enquêtes quantitatives réalisées auprès des activités économiques. En entrée, ce modèle (Freturb, à terme remplacé par Silogues) fonctionne avec deux sources de données :

- Le répertoire SIRENE, élaboré par l'INSEE, qui recense tous les établissements économiques implantés en France. C'est un fichier qui est mis à jour quotidiennement mais qui a subi au fil du temps un certain nombre d'évolutions (ajout / suppression / modification de variables)
- Un fichier de zonage qui présente les caractéristiques du territoire sur lequel porte la simulation (densité, étendue géographique, etc.)

La qualité de la simulation obtenue via Freturb-Silogues est en lien direct avec celle des données d'entrée utilisées par le modèle. Le stage proposé vise donc d'abord à appréhender la fiabilité du fichier SIRENE concernant les types d'activités et les effectifs salariés déclarés. Au-delà de cette mission il est également attendu du stagiaire de proposer des méthodes visant à enrichir le fichier afin d'y introduire des variables importantes pour la modélisation Freturb-Silogues mais aujourd'hui disparues du fichier SIRENE (notamment les tranches d'effectifs salariés, la nature de l'établissement) ou mal renseignées. Ce stage s'inscrit dans le cadre du projet de réalisation de la plateforme SILOGUES de simulation du transport de marchandises en ville sous différents scénarios.

Mission

Le stagiaire devra tout d'abord établir un diagnostic critique des données proposées par SIRENE en les comparant avec diverses sources de données qu'il aura pu identifier comme pertinentes pour conduire son analyse et qui comprendront :

- Les Enquêtes Transport de Marchandises en Ville, conduites par le LAET et dont une partie du questionnaire vise à contrôler que l'information de SIRENE est bien exacte,
- Les autres sources de données relatives aux activités produites par l'INSEE : Recensement des Entreprises et des Etablissements, Connaissance Locale de l'Appareil Productif, etc.
- Les données des services fiscaux : URSSAF notamment
- Les données disponibles dans les bases d'information géographique : données IGN

Pour effectuer son travail le stagiaire aura à disposition les bases de données issues des Enquêtes Transport de Marchandises en Ville réalisées sous la direction scientifique du LAET à Paris et Bordeaux en 2011 et 2013. Les sources de données complémentaires non disponibles en accès libre seront acquises par le laboratoire pour permettre leur confrontation à SIRENE.

Une fois identifiés les points faibles du répertoire, le stagiaire proposera une méthode d'enrichissement et de correction de SIRENE qui permettra à Freturb-Silogues d'améliorer la classification des établissements dans les typologies qu'il utilise.

Le stagiaire devra mobiliser de solides compétences en analyse de données, statistique et modélisation afin de construire une procédure de préparation du fichier SIRENE qui pourra être intégrée à la plateforme SILOGUES.

Pour mener à bien sa mission, le stagiaire pourra s'appuyer sur un travail exploratoire conduit sur cette question à la Métropole de Rouen.

Modalités pratiques

Ce stage d'une durée de 6 mois sera accueilli au sein du LAET.

Gratification : 27,3 euros par jour.

Contact au LAET : Mathieu Gardrat (mathieu.gardrat@entpe.fr)

Pour en savoir plus :

Sur la base SIRENE : <https://www.data.gouv.fr/fr/datasets/base-sirene-des-entreprises-et-de-leurs-etablissements-siren-siret/>

Sur le modèle Freturb : <http://tmv.laet.science/modele/index.html>

Sur le modèle SILOGUES : <http://tmv.laet.science/projets/SILOGUES/index.htm>

Sur l'Enquête Transport de Marchandises en Ville :

<http://tmv.laet.science/projets/Enquetes/index.htm>