

# Questions à se poser avant d'utiliser les résultats d'un modèle / d'une source de données pour analyser les comportements de mobilité

# 1. Pertinence

- **Indicateur 1.1 : L'objectif, les applications ou débouchés de l'analyse sont-ils clairement identifiés ?**
- **Indicateur 1.2 : Les sources de données utilisées sont-elles clairement identifiées ? La méthode de traitement est-elle expliquée ? Les sources de données et la méthode retenues sont-elles pertinentes par rapport aux objectifs fixés ?**
- **Indicateur 1.3 : La méthode a-t-elle été positionnée par rapport à un état de l'art ?**
- **Indicateur 1.4 : Si un modèle est utilisé (implicitement ou explicitement), les segmentations retenues pour les variables explicatives sont-elles pertinentes par rapport aux objectifs assignés ?**
  - Attention : Un redressement est déjà un modèle, avec des hypothèses, même si elles ne sont qu'implicites. Avant utilisation de données redressées, il faut rendre ce modèle explicite et vérifier sa pertinence. Il est également nécessaire de comprendre quelle analyse statistique a été réalisée pour définir les segmentations retenues, afin de savoir si elles sont pertinentes par rapport aux objectifs fixés.

# 1. Pertinence

- **Indicateur 1.5 : Si un besoin de comparabilité (spatiale entre territoires ou temporelle) a été identifié, les données et la méthode les autoriseront-elles ?**
- **Indicateur 1.6 : Si un besoin de tester des scénarios d'évolution de l'offre de transport et de la demande de déplacements projetés dans le temps a été identifié, les données et la méthode l'autoriseront-elles ?**
- **Indicateur 1.7 : Les coûts d'acquisition des données et de mise en œuvre de la méthode ont-ils été évalués et sont-ils compétitifs par rapport à des approches équivalentes ?**

## 2. Fiabilité

- **Indicateur 2.1 : Les biais systématiques des sources de données utilisés ont-ils été clairement identifiés et documentés ? L'échantillonnage des sources de données non-exhaustives utilisées est-il compatible avec les objectifs de l'analyse ?**
  - Exemples de biais systématiques : non-réponse plus forte des hyper-mobiles, mauvaise couverture des trajets courts, non-couverture des résidences collectives, couverture exclusive des utilisateurs d'un service / d'un fournisseur, sous-représentation des personnes âgées, sur-représentation des professions et catégories socio-professionnelles (PCS) supérieures, etc.
  - Par exemple, utiliser une source issue d'un équipement des véhicules haut-de-gamme sur-représentant les PCS supérieures serait inadaptée pour analyser la mobilité dans un quartier majoritairement habité par des foyers à faibles revenus.
- **Indicateur 2.2 : Les résultats intermédiaires et finaux ont-ils fait l'objet de validations par rapport à des données de référence et ces validations sont-elles documentées ? Les écarts aux valeurs de référence sont-ils compatibles avec les objectifs assignés à la méthode ? Les intervalles de confiance des données ont-ils été calculés et présentés ?**
  - Les notions de précision ou d'intervalle de confiance doivent dans toutes les situations être prises en compte. Toutefois, si des données présentant des biais d'échantillonnage systématiques sont utilisées, ces vérifications sont d'autant plus cruciales.

## 3. Actualité

- **Indicateur 3.1 : Le degré d'actualité des données utilisées est-il compatible avec les objectifs assignés ?**
- **Indicateur 3.2 : La périodicité de mise à jour des données utilisées est-elle compatible avec les objectifs assignés ?**

## 4. Cohérence

- **Indicateur 4.1 : Les incohérences (temporelles, en termes de zonages, de population cible, etc.) entre les sources de données utilisées sont-elles clairement identifiées ?**
- **Indicateur 4.2 : Des méthodes ont-elles été mises en place pour limiter ou compenser les incohérences entre les sources de données utilisées ? La validation de la méthode a-t-elle porté sur l'impact de ces incohérences ?**
  - Si les données utilisées sont en décalage spatial par rapport à la cible de l'analyse, la validation de la méthode a-t-elle permis de mettre en évidence ses bonnes capacités de transfert spatial ?
  - Si les données utilisées sont en décalage temporel par rapport à la cible de l'analyse, la validation de la méthode a-t-elle permis de mettre en évidence ses bonnes capacités de transfert temporel ?
  - La méthode fournit-elle des résultats sur un zonage cohérent avec les zonages administratifs ou les zonages d'études de l'INSEE, permettant leur comparaison avec d'autres données ?

## 5. Accessibilité et clarté

- **Indicateur 5.1** : Pour chaque source de données utilisée, l'origine, le propriétaire, la date et le mode de collecte, la précision spatiale et temporelle, la définition des variables, ainsi que la qualité de l'échantillon sont-ils documentés au sein de métadonnées qui en facilitent une correcte interprétation ?
- **Indicateur 5.2** : Les principes statistiques et algorithmiques de la méthode utilisée sont-ils clairement décrits ? Les chaînes de traitement itératifs et les éventuels bouclages sont-ils décrits ? Les entrées et sorties de chaque traitement sont-ils décrits ?
- **Indicateur 5.3** : Les éventuelles méthodes permettant de corriger ou redresser les résultats sont-elles décrites (par ex. les méthodes de correction de matrices OD sur des comptages ponctuels) ?
- **Indicateur 5.4** : La méthode permet-elle de mettre facilement en avant des corrélations entre les variables d'entrée et les variables de sortie ? Quel est le niveau d'effet "boîte noire" de la méthode ?

# Références utilisées

- **La construction de cette grille de questions s'est basée sur les travaux suivants :**
  - **Code de bonnes pratiques des statistiques européennes\* (principes 11 à 15) :**
    - Pertinence
    - Exactitude et fiabilité
    - Actualité et ponctualité
    - Cohérence et comparabilité
    - Accessibilité et clarté
  - **Bonnafous Alain (1989), Le siècle des ténèbres de l'économie, Economica, Paris.**
    - Conditions d'opérationnalité d'un modèle :
      - Pertinence : capacité du modèle à être conforme à ce que l'on croit savoir de la réalité.
      - Cohérence interne : condition de cohérence logico-mathématique.
      - Cohérence d'objectifs
      - Mesurabilité des variables en jeu dans le modèle, des relations causales et des paramètres des équations.

\*

<https://ec.europa.eu/eurostat/documents/4031688/9332274/KS-02-18-142-FR-N.pdf/130905e7-45a7-4475-b37c-8f699b5e33e1>





## Contacts :

- [modelisation-deplacements@cerema.fr](mailto:modelisation-deplacements@cerema.fr)
- [observation.mobilite@cerema.fr](mailto:observation.mobilite@cerema.fr)