

Élaboration de scénarios prospectifs de la mobilité

Louise Ligonniere, Amina Manseur, Lila Mekki

ENSAE Paris

23 mai 2024



Plan de la présentation

1 Résumé non-technique

- Objectifs et problématique
- Les données
- Les scénarios
- L'approche par les pondérations
- Les résultats principaux

2 Problématique technique : les pondérations

- La théorie des pondérations et le calage sur marges
- Règle de trois, sur une seule variable
- L'approche retenue de calage sur marges
- Le calcul des pondérations jours

3 Retour sur les résultats

- Le scénario démographique : interprétation des résultats
- Le scénario réchauffement climatique : limite de notre approche

4 Conclusion

Objectifs du projet

Transition énergétique = défi de taille pour la société ! L'utilisation croissante de véhicules électriques nécessite d'anticiper et d'organiser les besoins en énergie.

Notre objectif a été de mettre en œuvre des scénarios de mobilité aux horizons 2040 à 2070.



Problématique pour EDF

► Un outil développé par EDF : CharME.

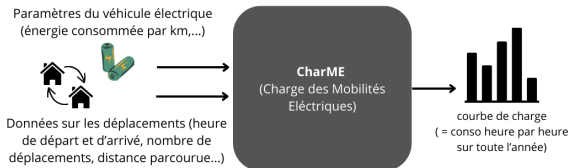


Figure: Schéma explicatif du fonctionnement de CharME

► Deux hypothèses :

- 1 Le type d'énergie de la voiture n'influe pas sur les déplacements des individus.
- 2 La mobilité ne dépend pas de l'horizon de prévision.

Les données

- ▶ **Enquête de Mobilité des Personnes (EMP), Insee, 2019**
 - 13 825 ménages,
 - 45 169 déplacements,
 - habitudes de déplacements quotidiens et occasionnels,
- ▶ **Variables cibles** → représentatives des résultats de sortie de CharME.

Les scénarios

Les deux scénarios sélectionnés :

- ❶ **Scénario d'évolution démographique** : selon l'âge, le sexe et le statut d'activité (Insee)
- ❷ **Scénario de réchauffement climatique** : évolution de la température moyenne et des précipitations totales sur une journée (Météo France, Centre National de Recherche Météorologique)

L'approche par les pondérations

Approche retenue =

- 1 Définir des variables dont on prévoit l'évolution d'ici l'horizon souhaité.
- 2 Modifier les pondérations de l'enquête pour obtenir les bons effectifs projetés sur ces variables.
- 3 Calculer les variables cibles (représentatives des sorties de CharME) et étudier leur évolution.

Point important : on ne modifie pas la façon dont les individus, à caractéristiques contrôlées, se déplacent.

Les résultats du scénario démographique

Pour rappel, nous avons pris en compte l'évolution des trois variables suivantes dans le recalcul des pondérations :

- L'âge ;
- Le sexe ;
- Le statut d'activité (actif ou non actif).

Variable d'intérêt	2019	2040	2070
Pour rappel : population totale (6 ans et plus)	59,5 millions	64,9 millions	64,2 millions
Nombre total de déplacements en une journée	174 millions	178,1 millions	176,2 millions
Distance totale parcourue en une journée (km)	2,22 milliards	2,26 milliards	2,25 milliards
Nombre moyen de déplacements par jour par individu	2,93	2,74	2,75
Distance moyenne par jour par individu (km)	37,58	34,90	35,02
Nombre total de déplacements EN VOITURE en une journée	110,6 millions	116,2 millions	117,8 millions
Distance totale parcourue EN VOITURE en une journée (km)	1,66 milliards	1,72 milliards	1,72 milliards
Nombre moyen de déplacements EN VOITURE par jour par individu	1,86	1,79	1,84
Distance moyenne EN VOITURE par jour par individu (km)	27,94	26,58	26,87

Figure: Résultats du scénario démographique à horizons 2040 et 2070

Les pondérations dans l'EMP

- ▶ **Les pondérations individus (pond_ind) :**
 - Représentativité au niveau de la population française
 - Correction de la non-réponse totale
 - Calage sur marges, simultané au niveau ménage et individu

Le calage sur marges : théorie

- **But** : Corriger le biais de sélection : certaines catégories de personnes sur-représentées dans l'échantillon
- **Les variables de calage** : Variables $X_1, \dots, X_j, \dots, X_J$ pour lesquelles on connaît les effectifs sur la population
- **Cadre formel** :
 - Population U , échantillon s
 - Pour tout individu $k \in U$, probabilité d'inclusion dans s notée π_k
 - Poids d'estimation d_k , initialement $d_k = \frac{1}{\pi_k}$
 - Estimateur pondéré pour Y : $\hat{Y} = \sum_{k \in s} d_k y_k$

Le calage sur marges : théorie

► Les équations de calage :

$$\forall j = 1, \dots, J : \sum_{k \in s} w_k x_{j,k} = X_j \quad (1)$$

► Objectif : Pour une fonction de distance G fixée :

$$\begin{aligned} \min_{(w_k)_k} \quad & \sum_{k \in s} d_k G\left(\frac{w_k}{d_k}\right) \\ \text{s.c.} \quad & \forall j = 1, \dots, J : \sum_{k \in s} w_k x_{j,k} = X_j \end{aligned} \quad (2)$$

► Quatre méthodes usuelles, correspondant à quatre fonctions de calage :

- Méthode linéaire
- Méthode raking ratio
- Méthode logit
- Méthode linéaire tronquée

Règle de trois, sur une seule variable - exemple : l'âge

- ▶ Tranches d'âge $i \in 1, \dots, 9$
- ▶ Ratio au sein de chaque catégorie d'âge, en notant $s_i = \{k \in s \mid \text{age}_k = i\}$:

$$\forall k \in s_i : \text{pond_ind}_{k,2040} = \text{pond_ind}_{k,2019} \frac{\text{tot}_{i,2040}}{\text{tot}_{i,2019}}$$

- ▶ Limite de cette méthode : l'enquête n'est plus représentative sur les autres variables de calage
- ▶ Exemple : pas assez d'actifs en 2040 !

L'approche retenue de calage sur marges : poids initiaux

► Enquête par sondage aléatoire simple stratifié :

- Population découpée en strates selon 3 variables (statut QPV, type de communes urbaines et rurales, motorisation)
- Tirage non-uniforme entre les strates (selon la variance)

► Pour notre méthode, poids uniformes sur l'ensemble de l'échantillon :

- Probabilité d'inclusion $\pi_k = \frac{\#s}{\#U}$
- Poids d'estimation $d_k = \frac{1}{\pi_k} = \frac{\#U}{\#s}$

► Limites :

- Ne prend pas en compte le plan de sondage spécifique
- Ne corrige pas la non-réponse totale

L'approche retenue de calage sur marges : marges de calage

- ▶ Projections à horizon h lorsqu'elles sont disponibles
- ▶ Pour les autres variables :
 - *Hypothèse 1* : Leur distribution n'évolue pas
→ Utiliser les effectifs calculés en 2019
 - *Hypothèse 2* : Leur distribution n'évolue pas **conditionnellement aux variables de calage**
→ Ne pas mettre de contrainte sur les effectifs de ces variables

L'approche retenue de calage sur marges : méthode de calage

- Essai de différentes méthodes, et choix selon les rapports de poids

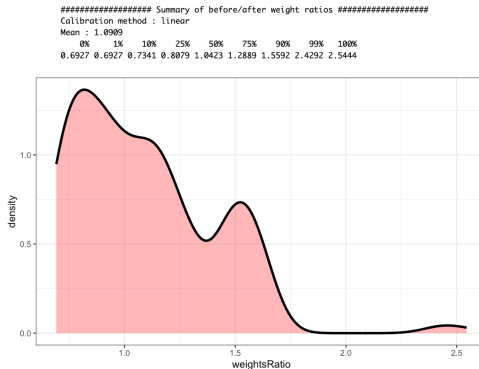


Figure: Rapports des poids individus après application de la méthode linéaire, scénario démographie 2040

Le calcul des pondérations jours

► Les pondérations déplacements (pond_jour) :

- Jour d'interrogation : tiré parmi les jours avec déplacements dans les 7 jours précédant l'enquête
- Pas réparti de façon uniforme dans la base : plus d'enquêtés interrogés sur leurs déplacements un jour de semaine qu'un samedi ou un dimanche
- Déplacements fortement corrélés au type de jour de la semaine
- Objectif : représenter la mobilité des Français au cours d'une semaine type

Le calcul des pondérations jours

- **Objectif** : Donner un poids à chaque ligne selon le type de jour d'interrogation, pour obtenir une semaine complète
- Les pondérations jours représentent des individus-jours de déplacement.
- On note N_j^* le nombre d'individus de la population qui se sont déplacés le jour j .
- En notant jour_k le jour d'interrogation de l'individu k :

$$\forall j = 1, \dots, 7 : \sum_{k \in s, \text{jour}_k = j} d_{2,k} = N_j^*$$

- Sur une semaine complète :

$$\sum_{k \in s} d_{2,k} = \sum_{j=1 \dots 7} N_j^*$$

Le calcul des pondérations jour

- On calcule les N_j^* pour chaque type de jour $j = 1, \dots, 7$, en appliquant les pondérations individu $d_{1,k}$ projetées à l'étape précédente.
- Pour chaque type de jour j , on estime ensuite le nombre d'individus dans la population dont on dispose des déplacements le jour j :

$$n_j^* = \sum_{k \in s, \text{jour}_k = j} d_{1,k}$$

- On applique ensuite le ratio suivant pour obtenir les pondérations jours :

$$d_{2,k} = d_{1,k} \sum_{j=1 \dots 7} 1(\text{jour}_k = j) \frac{N_j^*}{n_j^*}$$

- Limite de cette méthode : le ratio $d_{2,k}/d_{1,k}$ est le même pour tous les individus ayant été interrogés un même jour j (pas le cas dans l'EMP 2019)

Le scénario démographique : les résultats

Variable d'intérêt	2019	2040	2070
Pour rappel : population totale (6 ans et plus)	59,5 millions	64,9 millions	64,2 millions
Nombre total de déplacements en une journée	174 millions	178,1 millions	176,2 millions
Distance totale parcourue en une journée (km)	2,22 milliards	2,26 milliards	2,25 milliards
Nombre moyen de déplacements par jour par individu	2,93	2,74	2,75
Distance moyenne par jour par individu (km)	37,38	34,90	35,02
Nombre total de déplacements EN VOITURE en une journée	110,6 millions	116,2 millions	117,8 millions
Distance totale parcourue EN VOITURE en une journée (km)	1,66 milliards	1,72 milliards	1,72 milliards
Nombre moyen de déplacements EN VOITURE par jour par individu	1,86	1,79	1,84
Distance moyenne EN VOITURE par jour par individu (km)	27,94	26,58	26,87

Figure: Résultats du scénario démographique à horizons 2040 et 2070

Le scénario démographique : les résultats

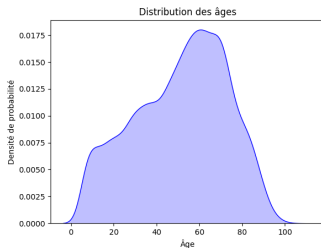


Figure: Distribution des âges en 2019

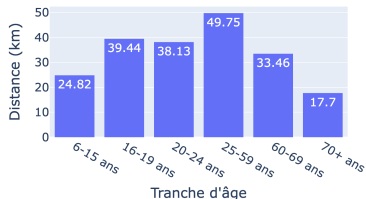


Figure: Distance parcourue selon l'âge

Le scénario démographique : les résultats

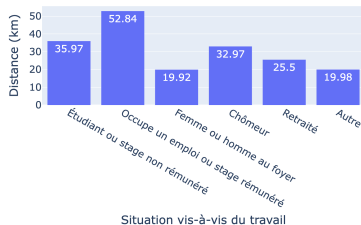


Figure: Nombre de déplacement selon la situation vis-à-vis du travail

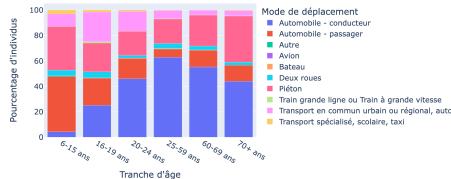


Figure: Mode de transport selon l'âge

Le scénario démographique : les résultats

► Qu'en retient-on ? :

- 2040 : Baisse du nombre de déplacements par individus mais ne compense pas l'augmentation de population. **Au total, le nombre de déplacements augmente.**
- 2070 : La population diminue légèrement : baisse du nombre total de déplacements et de la distance totale.
- 2070 : Le nombre moyen de déplacements et la **distance moyenne par individu réaugmentent légèrement** (contrairement à 2040-2070). Plus de déplacements en voiture.

Le scénario climatique : les projections

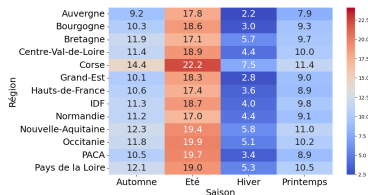


Figure: Mai 2018 - Avril 2019

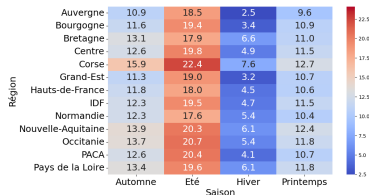


Figure: 2021 - 2050

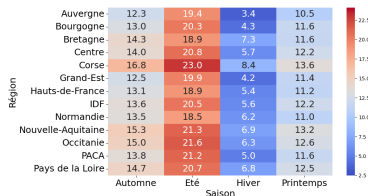


Figure: 2041 - 2070

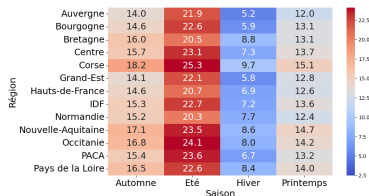


Figure: 2071 - 2100

Le scénario climatique : les projections

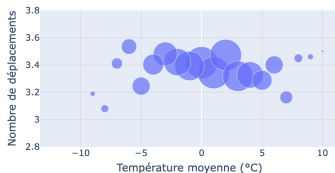


Figure: Déplacements et température

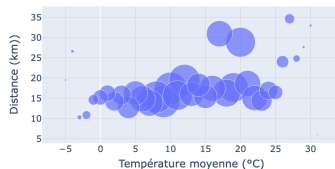


Figure: Distance et température

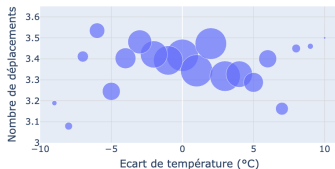


Figure: Déplacements et écarts de température

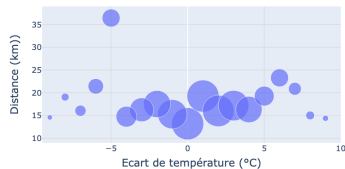


Figure: Distance et écarts de température

Le scénario climatique : approche et limites

► Notre approche :

- Utilisation du calage sur marges
- Aux températures journalières ?
- À des tranches de températures ?

► Ses limites :

- Manque de données dans l'EMP
- Pas assez de températures ou précipitations extrêmes
- Difficulté de modélisation de l'impact des températures sur les déplacements

Conclusion

- ▶ **Base solide pour l'élaboration de nouveaux scénarios**
- ▶ **Deux limites principales :**
 - Méthode adaptée pour des évolutions de variables socio-démographiques, mais peu concluante pour d'autres variables.
 - Hypothèse assez forte de non-évolution des pratiques de déplacements à variables de calages égales.