

La mise en place des enquêtes annuelles de recensement à Mayotte

Laurent Costa
Martin Chevalier



JMS 2022 30/03/2022

Lancement des Enquêtes Annuelles de Recensement (EAR) à Mayotte

- Dernier recensement général de la population à Mayotte en 2017
- Première EAR à Mayotte en 2021
- Première population légale publiée le 1^{er} janvier 2026 (millésimée 2023) à partir du premier **cycle RP** constitué (cumul de 5 EAR)

Mayotte est constitué de :

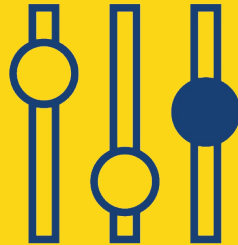
- 10 communes de 10 000 habitants ou plus, dites « grandes communes » (GC)
- 7 communes de moins de 10 000 habitants, dites « petites communes » (PC)

Cette présentation porte uniquement sur les grandes communes : réflexion sur la **déclinaison des méthodes de sondage en grande commune dans le contexte mahorais**.



01

Le recensement dans les DOM
et son application à Mayotte



02

Détermination d'un plan de
sondage alternatif par simulation

01

LE RECENSEMENT DANS LES DOM ET SON APPLICATION À MAYOTTE

À l'initialisation :

- au sein de chaque grande commune, des zones géographiques (appelées **îlots**) sont définies ;
- les îlots sont répartis en cinq **groupes de rotation** (GR) de la façon la plus équilibrée possible.

Chaque année :

- une **enquête cartographique exhaustive** est réalisée en N-1 dans les îlots du GR de l'année N ;
- un **tirage** est réalisé dans l'enquête cartographique de façon à constituer un **échantillon représentant 40 % des logements du GR de l'année N** ;
- les **adresses tirées au sort sont enquêtées**.

Importance de l'**habitat de fortune** (39 % des logements recensés en 2017) :

- Adresses **difficilement identifiables sur le terrain** : rien ne garantit que les logements *in fine* recensés sont bel et bien ceux qui ont été tirés au sort.
- **Evolution très rapide** : le décalage de quelques mois entre l'enquête cartographique servant au sondage et l'arrivée sur le terrain de l'enquête de recensement peut induire un biais de couverture.

→ Étude sur ce défaut de couverture et sur une adaptation de la méthode en conséquence

La disponibilité du RP 2017 a été **cruciale** pour l'étude méthodologique :

- base de données **exhaustive** ;
- information raisonnablement **fraîche**.

Une enquête cartographique elle-même exhaustive a précédé le RP 2017 :

- **géolocalisation fine** des unités de l'enquête cartographique : possibilité de localiser les adresses du RP dans **n'importe quelle partition du territoire** ;
- **décalage** entre enquête cartographique et RP de plusieurs mois : analogue (en plus long) à celui qui surviendra pour les EAR.

En règle générale, le défaut de couverture d'une base de sondage est à relier à un **problème d'actualité** : la population d'intérêt a beaucoup évolué entre la confection de la base de sondage et la date de validité de l'enquête.

Dans ce type de situation, on distingue trois types d'unités :

- les « présents-présents » : figurent dans la base de sondage et dans la population d'intérêt ;
- les « sortants » : figurent dans la base de sondage mais pas dans la population d'intérêt (destructions, cessations, décès, etc.) ;
- les « entrants » : figurent dans la population d'intérêt mais pas dans la base de sondage (constructions, créations, naissances, etc.).

La présence de **sortants** induit une augmentation de la **variance** des estimateurs, la présence d'**entrants** induit un **biais**.

Première évaluation du défaut de couverture par comparaison de l'enquête cartographique 2016 et du RP 2017

<i>Ensemble des grandes communes</i>	Présents	Entrants	Présents + entrants	Biais attendu du fait du défaut de couverture
Population	184 152	23 146	207 298	-11,2 %
Pourcentage d'étrangers	52,3 %	66,8 %	53,9 %	-3,0 %

Deux pistes ont été poursuivies pour limiter le biais de couverture ou affiner son évaluation :

- Extension de la base de sondage aux entités adressées en travaux : diminution du nombre d'entrants au prix d'une augmentation du nombre de sortants (arbitrage biais-variance très favorable ici) ;
- Hypothèses plus optimistes pour tenter de mesurer l'impact de « faux entrants » : adresses présentes dans la base de sondage et qui existent toujours au moment du RP 2017 mais qui n'ont pas été retrouvées sur le terrain :
 - scénario optimiste : appariement entre entités adressées sortantes/adresses entrantes d'un même îlot x type d'habitat (dur/non-dur) ;
 - scénario très optimiste : appariement entre entités adressées sortantes/adresses entrantes d'un même îlot (même quand le type d'habitat diffère).

<i>Ensemble des grandes communes</i>	Biais attendu du fait du défaut de couverture	
	Population	Pourcentage d'étrangers
En excluant les EA en travaux	- 11,2 %	- 3,0 %
En intégrant les EA en travaux	- 9,1 %	- 3,0 %
Hypothèse optimiste sur les « faux entrants »	- 3,8 %	- 1,5 %
Hypothèse très optimiste sur les « faux entrants »	- 3,1 %	- 1,2 %

Qualitativement, on retient que **le biais de l'estimateur de la population avec le plan de sondage classique du recensement dans les DOM est évalué entre -9,1 % et -3,8 % pour l'ensemble des GC de Mayotte.**

02

DÉTERMINATION D'UN PLAN DE SONDAGE ALTERNATIF PAR SIMULATION

Les performances des différents plans de sondage considérés sont **évaluées par simulation** :

- pour chaque « scénario » de tirage, un grand nombre de tirages (2 000) est réalisé dans le RP 2017 : **chacune de ces simulations produit un échantillon et un jeu de pondérations** ;
- pour chaque scénario, chaque simulation et chaque variable d'intérêt, **une estimation est calculée** à partir de l'échantillon et des pondérations ;
- pour chaque scénario et chaque variable d'intérêt, **la distribution des 2 000 estimations est comparée à la vraie valeur connue dans le RP 2017** pour **déterminer le biais et la variance de l'estimateur**.

62 scénarios ont été *in fine* comparés (23 plans de sondage avec à chaque fois plusieurs variantes : hypothèses ou estimateurs différents), ce qui représente de l'ordre de **460 000 simulations de tirage**.

Principe de la méthode

- Réalisation entre avril et août N-1 d'une enquête cartographique exhaustive sur le groupe de rotation dans lequel sera réalisé l'EAR N.
- Tirage dans l'enquête cartographique d'un échantillon d'adresses de sorte qu'il représente environ 40 % des logements de l'enquête cartographique.
- Recensement en février-mars N des seules adresses tirées.

Principe de la méthode

- Réalisation entre avril et août N-1 d'une enquête cartographique sur le groupe de rotation dans lequel sera réalisé l'EAR N.
- Tirage dans l'enquête cartographique d'un *échantillon d'îlots* de sorte que les logements de ces îlots représentent environ 40 % des logements de l'enquête cartographique.
- Recensement en février-mars N de l'ensemble des adresses figurant dans les limites des îlots tirés au moment de l'enquête (recensement exhaustif des îlots tirés).

Principe de la méthode

- Réalisation entre avril et août N-1 d'une enquête cartographique sur le groupe de rotation dans lequel sera réalisé l'EAR N.
- Tirage *parmi les logements « en dur »* de l'enquête cartographique d'un échantillon d'adresses de sorte qu'il représente environ 40 % des logements « *en dur* » de l'enquête cartographique.
- Recensement en février-mars N :
 - parmi les logements « en dur » : des seules adresses tirées ;
 - de tous les logements de fortune localisés au début de la collecte du recensement dans les îlots du groupe de rotation enquêté.

Principe de la méthode

- Réalisation entre avril et août N-1 d'une enquête cartographique sur le groupe de rotation dans lequel sera réalisé l'EAR N.
- Tirage *parmi les logements « en dur »* de l'enquête cartographique d'un échantillon d'adresses de sorte qu'il représente environ 40 % des logements « *en dur* » de l'enquête cartographique.
- Tirage dans l'enquête cartographique d'un *échantillon d'îlots* de sorte que les *logements de fortune* de ces îlots représentent environ 40 % des *logements de fortune* de l'enquête cartographique.
- Recensement en février-mars N :
 - parmi les logements « en dur » : des seules adresses tirées ;
 - parmi les logements de fortune : de tous les logements de fortune des *seuls îlots tirés*.

Ensemble des GC Minimum parmi les GC Maximum parmi les GC	Population		Pourcentage d'étrangers	
	Biais	CV	Biais	CV
Scénario 1 (méthode classique DOM)	- 9,1 % -18,3 % -2,6 %	0,5 % 0,9 % 2,5 %	- 2,9 % -10,6 % -0,6 %	0,5 % 0,8 % 4,2 %
Scénario 1 (avec hypothèse optimiste)	-3,8 % -7,7 % -0,5 %	0,5 % 0,9 % 2,4 %	- 1,5 % -5,6 % -0,3 %	0,5 % 0,8 % 4,0 %
Scénario 2 (tirage d'îlots)	0,0 % -0,5 % 0,3 %	1,2 % 2,2 % 4,9 %	- 0,2 % -2,0 % -0,1 %	1,4 % 2,4 % 7,4 %
Scénario 3 (recensement exhaustif des logements de fortune)	- 2,0 % -5,3 % -0,4 %	0,4 % 0,8 % 2,4 %	0,6 % 0,1 % 1,4 %	0,4 % 0,6 % 3,5 %
Scénario 3 (avec hypothèse optimiste)	- 0,3 % -0,9 % 0,0 %	0,4 % 0,7 % 2,2 %	0,1 % -0,1 % 0,6 %	0,4 % 0,6 % 3,6 %
Scénario 4 (tirage mixte)	- 2,2 % -5,6 % -0,9 %	0,9 % 1,4 % 3,9 %	0,4 % -0,6 % 1,4 %	0,8 % 1,3 % 4,1 %
Scénario 4 (avec hypothèse optimiste)	- 0,4 % -1,1 % 0,1 %	0,8 % 1,4 % 3,7 %	- 0,1 % -1,0 % 0,5 %	0,8 % 1,3 % 4,0 %

SYNTHÈSE

	Biais	Variance	Charge pour les communes	Évolutions du protocole et des traitements
Scénario 1 (méthode classique DOM)	Fort	Faible	Standard	Non
Scénario 2 (tirage d'îlots)	Nul	Forte	Standard	Oui
Scénario 3 (recensement exhaustif des logements de fortune)	Comparable aux autres DOM	Faible	Beaucoup plus importante	Oui
Scénario 4 (tirage mixte)	Comparable aux autres DOM	Moyenne	Standard	Oui

Moralité : pas de scénario idéal
 → **Arbitrage** : scénario 3 retenu

	<i>Logements en dur</i>	<i>Logements de fortune</i>	<i>Total</i>
Nombre de logements total	32 346 (60,7 %)	20 930 (39,3 %)	53 276
Nombre de logements dans le GR	6 809 (61,1 %)	4 341 (38,9 %)	11 150
Nombre de logements tirés	2 757 (38,8 %)	4 341 (61,2 %)	7 098
Nombre de logements collectés	2 790 (39,7 %)	4 235 (60,3 %)	7 025

- Collecte du 18 mars au 24 avril à cause du confinement.
- Très faible taux de non-réponse.
- Bonne implication des acteurs locaux et formation comprise pour les agents recenseurs.
- La collecte s'est globalement bien déroulée grâce à l'expérience du terrain des coordonnateurs communaux.
- Définition des *logements de fortune* à perfectionner lors de la cartographie.
- Les opérations de « *décasages massifs* », notamment dans les zones denses en *logements de fortune*, ont compliqué la collecte.
- Beaucoup de collecte par interview (difficultés de compréhension du questionnaire) et très peu par internet.

Retrouvez-nous sur

[insee.fr](https://www.insee.fr)



Laurent Costa
INSEE – DSDS – DMTR
Martin Chevalier
DREES