FLOWMINDER.ORG

Estimations de la population et de la mobilité pour la République démocratique du Congo

Rapport méthodologique et description des ensembles de données

Version 2.1 des estimations

Copyright © 2026 Flowminder Foundation - All Rights Reserved

Estimations de la population et de la mobilité pour la République démocratique du Congo

Rapport méthodologique et description des ensembles de données

Version 2.1 des estimations

Dernière mise à jour : novembre 2026

Liste du contenu

1. Résumé de la méthodologie 3
- 1.1. Comprendre les métadonnées de téléphonie mobile (données CDR) 3
- 1.2. Comprendre nos données 3
- 1.3. Protection des données et de la vie privée 5

2. Estimations du nombre de résidents par zone de santé 6
- 2.1. Informations générales 6
- 2.2. Estimations du nombre de résidents par zone de santé 7
- 2.3. Estimations des flux entrants par zone de santé 8
- 2.4. Estimations des flux sortants par zone de santé 8

3. Estimations des changements de lieu de résidence entre zones de santé 10
- 3.1. Informations générales 10
- 3.2. Estimations des changements de lieu de résidence entre zones de santé 10

4. Sources de données 13
- 4.1. Métadonnées de téléphonie (données CDR), Vodacom RDC 13
- 4.2. Microrecensement de la RDC, 2021, Flowminder/WorldPop 13
- 4.3. Enquête téléphonique en RDC, 2021, Flowminder 13
- 4.4. Estimations démographiques de l'UN OCHA, 2020 13
- 4.5. Annuaires statistiques de l'INS, 2017 et 2020 13
- 4.6. Fichiers de forme (shapefiles) des zones de santé, DHIS2 14

5. Limitations 15
- 5.1. Biais “de mesure” 15
- 5.2. Représentativité et biais 15

Copyright © 2026 Flowminder Foundation - All Rights Reserved

1. Résumé de la méthodologie

1.1. Comprendre les métadonnées de téléphonie mobile (données CDR)

Les métadonnées de téléphonie mobile (comptes rendus d'appels (CRA) ou données CDR, de l'anglais Call Detail Records) sont un type d'informations régulièrement enregistrées par les opérateurs de réseaux mobiles sur l'utilisation du réseau par leurs abonnés, à des fins de facturation. Chaque fois qu'un abonné est impliqué dans un événement de réseau - qu'il s'agisse de passer ou de recevoir un appel, d'envoyer ou de recevoir un message SMS ou d'utiliser des données mobiles - l'opérateur enregistre le type d'événement de réseau, l'heure de l'événement, l'identifiant du numéro de téléphone et l'identifiant de la tour de téléphonie mobile qui a acheminé l'événement. Sur la base de ces informations et de l'emplacement des tours de téléphonie cellulaire, nous pouvons évaluer la façon dont les gens se déplacent à l'intérieur d'un pays. L'analyse de l'addition des mobilités des abonnés donnent un aperçu de la mobilité de la population, tout en protégeant la vie privée de chaque abonné (les données sont, entre autres, anonymisées (pseudonymisées) et agrégées) (voir Protection des données et de la vie privée). Pour des explications plus détaillées sur les CDR, veuillez consulter FlowGeek, notre centre de connaissances sur l'analyse des données CDR ici (en anglais uniquement).

1.2. Comprendre nos données

Afin d'y extraire des informations utilisables sur la mobilité, qui protègent la vie privée des individus et sont pertinentes pour les secteurs du développement et de l'aide humanitaire, les données CDR doivent passer par plusieurs étapes de traitement et être combinées avec des données d'enquête et des données démographiques.

Les indicateurs et mesures (statistiques) relatifs aux résidents et aux changements de lieu de résidence présentés dans nos rapports sont le résultat d'années de recherche et de développement de méthodes, et ont été corrigés pour tenir compte des évolutions de la population :

- **Être représentatif de l'ensemble de la population :** à l'aide de données d'enquête et d'estimations de la population existantes, nous mettons à l'échelle nos estimations afin qu'elles ne représentent pas uniquement le nombre d'identifiants ou d'abonnés de Vodacom RDC. Nos estimations représentent la population dans son ensemble, indépendamment de l'opérateur utilisé ou du fait qu'une personne utilise ou non un téléphone mobile.

et

- **Réduire l'influence du comportement en matière d'utilisation du téléphone :** nous estimons le nombre de résidents à partir de la mobilité observée des abonnés (et des estimations de population existantes) plutôt qu'en comptant le nombre d'abonnés résidant dans chaque lieu, afin de garantir que les données représentent davantage la mobilité réelle que l'utilisation variable du téléphone.

Copyright © 2026 Flowminder Foundation - All Rights Reserved

Les indicateurs et les mesures relatifs aux résidents et aux changements de lieu de résidence présentés dans nos rapports sont des estimations de la population globale de la République démocratique du Congo (RDC), et pas seulement du nombre d'abonnés de Vodacom RDC. **Ils ont été mis à l'échelle et ajustés pour être représentatifs de la population dans son ensemble - pour mesurer la mobilité réelle et la population actuelle, en réduisant la sensibilité à l'utilisation variable du téléphone.**

Il est important de noter que pour estimer le nombre de résidents, il faut d'abord estimer les changements de lieu de résidence. Les **résidents** estimés **par zone de santé et par mois** sont conceptualisés comme la population qui a passé la majorité du mois dans cette zone de santé (la population de fait). Les étapes du calcul sont les suivantes :

- Tout d'abord, nous attribuons à chaque identifiant un **lieu "de résidence" (localisation résidentielle)** par mois, c'est-à-dire la section communale qui contient les antennes de téléphonie cellulaire près desquelles l'identifiant se trouvait la majeure partie du mois (si aucune section communale n'est majoritaire, aucun lieu n'est attribué).
- Nous définissons ensuite les **changements de lieu de résidence** comme une modification de cette localisation résidentielle d'un mois à l'autre.
- Nous **pondérons** (ajustons et mettons à l'échelle) les changements de lieu de résidence entre zones de santé en utilisant un paramètre SIM-utilisateur et une combinaison de la couverture de la population des données CDR dans la zone de santé d'origine et la zone de santé de destination. Il s'agit de notre méthode actuelle pour corriger les biais de représentativité.
- Nous additionnons tous les changements de lieu de résidence pondérés vers chaque zone de santé par mois (Total des entrées, flux entrants) et tous les changements de lieu de résidence pondérés depuis chaque zone de santé par mois (Total des sorties, flux sortants) pour calculer les **flux nets (changements de lieu de résidence nets)** estimés pour chaque zone de santé. Cela correspond à la différence entre les résidents qui se sont installés et les résidents qui sont partis, du fait de la mobilité mensuelle uniquement (entrées moins sorties).
- Nous calculons ensuite les **estimations démographiques de base/référence** pour 2020 à partir des **estimations de l'OCHA pour 2020 par zone de santé** et des **estimations au niveau provincial tirées de l'Annuaire statistique 2020 de l'INS**.
- À ces estimations de base, nous **ajoutons les changements de lieu de résidence nets** (flux nets, ou "flux entrants moins flux sortants") pour chaque zone de santé entre ce mois et le mois suivant. Seule la mobilité détectée dans les CDR est utilisée, en combinaison avec des pondérations, pour estimer les résidents - et non le nombre de lieux d'habitation dérivé des CDR. **C'est notre méthode pour réduire l'influence des changements de comportement en matière d'utilisation du téléphone dans nos estimations de la mobilité et de la population dérivées des CDR.**
- Nous multiplions ces estimations mensuelles par des **taux de variation mensuels de la population** (également tirés des estimations démographiques publiées dans l'Annuaire statistique 2020 de l'INS) pour tenir compte des composantes démographiques non couvertes par les données CDR (naissances, décès, immigration, émigration).

Copyright © 2026 Flowminder Foundation - All Rights Reserved

- Nous **répétons** ce processus chaque mois jusqu'au dernier mois disponible dans les données CDR, afin d'estimer le nombre de résidents à partir de la population de référence existante, de la mobilité interne et d'autres changements démographiques survenus depuis lors.

1.3. Protection des données et de la vie privée

Aucune donnée personnelle, telle que l'identité, les données démographiques, la localisation, les contacts ou les déplacements d'un individu, n'est à aucun moment mise à la disposition du gouvernement ou de toute autre tierce partie. Tous les résultats que nous produisons et publions sont agrégés (par exemple, le nombre d'habitants d'une municipalité donnée), ce qui signifie qu'ils ne contiennent aucune information sur les abonnés individuels. Les données sont entièrement anonymes. Cette approche est conforme au règlement général sur la protection des données de l'Union européenne (UE GDPR 2016/679). Les données sont traitées sur un serveur installé derrière le pare-feu de Vodacom RDC, en République démocratique du Congo, et seules les données agrégées quittent les locaux de l'opérateur.

Copyright © 2026 Flowminder Foundation - All Rights Reserved

2. Estimations du nombre de résidents par zone de santé

2.1. Informations générales

2.1.1. Version

Version 2.0

2.1.2. Description

Un ensemble de données contenant des estimations mensuelles de la population de fait des zones de santé de la RDC, **depuis mars 2020.**

2.1.3. Unités temporelles et couverture

Mois civils.

2.1.4. Unités géographiques et couverture

Les estimations couvrent l'ensemble des zones de santé de la RDC où les métadonnées de téléphonie mobile (données CDR) sont disponibles. Les limites des zones de santé sont celles des données DHIS2 2024.

2.1.5. Rédaction

Les zones de santé pour lesquelles la part des résidents issus des CDR par rapport à la population générale n'a jamais dépassé 1 % dans la série chronologique ont été supprimées. Toutes les valeurs inférieures à 15 ont été supprimées (et marquées « supprimées (nombre < 15) ») pour des raisons de protection des données.

2.1.6. Estimations de base de la population

Les **estimations de la population infranationale compilées par le Bureau de la coordination des affaires humanitaires des Nations unies (OCHA) pour 2020** ont été utilisées comme estimations de référence, en combinaison avec les **estimations au niveau provincial fournies par l'Institut national de la statistique (INS) pour 2019** et le **taux de croissance annuel dérivé des estimations chronologiques de l'INS (1,033).**

$est_pop_2020_03_a$ = $est_pop_ocha_2020_a * ((est_pop_ins_2019_adm1_a * ann_growth_ins_adm0) / est_pop_ocha_2020_adm1_a)$

$est_pop_2020_03_{an}$	est l'estimation du nombre d'habitants pour la zone de santé *a* pour mars 2020
$est_pop_ins_2019_adm1_a$	est l'estimation du nombre d'habitants de l'INS pour la province (adm1) de la zone de santé *a* pour 2019
$est_pop_ocha_2020_adm1_a$	est l'estimation du nombre d'habitants de l'OCHA pour la province (adm1) de la zone de santé *a* pour 2020
ann_growth_ins_adm0	est le taux de croissance annuel de la population totale du pays (adm0) dérivé des estimations chronologiques de l'INS

Copyright © 2026 Flowminder Foundation - All Rights Reserved

2.2. Estimations du nombre de résidents par zone de santé

L'estimation des résidents de fait dans la zone de santé *"a"* pour le mois *"n"* (est_pop_{an}) est calculée comme la somme de la population de cette zone de santé au cours du mois précédent *m* (est_pop_{am}) et des changements de lieu de résidence nets pour cette zone de santé entre les deux mois ($est_netflows_{amn}$), multipliée par un facteur de croissance de la population ($1 + change_rate_{a}$). Le mois de référence est mars 2020 (*m=0*) pour lequel l'estimation des résidents est basée sur les estimations de population existantes.

2.1.1. Calcul

L'estimation des résidents peut être exprimée sous la forme d'un système d'équations récursives :

$$est_pop_{an} = (est_pop_{am} + est_netflows_{amn}) * (1+change_rate)_{a}$$

$$est_pop_{a0} = est_pop_{abase}$$

Où :

est_pop_{an}	est l'estimation du nombre de résidents dans la zone de santé *"a"* pour le mois en cours *"n"*
est_pop_{am}	est l'estimation du nombre de résidents de la zone de santé *"a"* pour le mois précédent *"m"*
$est_netflows_{amn}$	est le nombre total estimé de changements de lieu de résidence nets pour la zone de santé *"a"* entre les mois *"m" et "n"*.
change_rate	est le taux de variation mensuel moyen estimé de la populationen pourcentage
est_pop_{a0}	est l'estimation des résidents de la zone de santé *a* pour *m=0* (mars 21020), *la* base de référence *de l'estimation* de la population (est_pop_{abase})

L'estimation des changements de lieu de résidence nets pour la zone de santé *"a"* entre les mois *"m" et "n"* ($est_netflows_{amn}$) est la somme de tous les changements de lieu de résidence estimés vers cette zone de santé ($est_inflows_{amn}$) moins la somme de tous les changements de lieu de résidence estimés à partir de cette zone de santé ($est_outflows_{amn}$) :

$$est_netflows_{amn} = est_inflows_{amn} - est_outflows_{amn}$$

Voir aussi **flux entrants estimés** et **flux sortants estimés**.

2.2.2. Estimations des limites supérieures et inférieures

Copyright © 2026 Flowminder Foundation - All Rights Reserved

$est_pop_LB_{an}$	est l'estimation du nombre de résidents de la zone de santé *"a"* pour le mois "n", limite inférieure
$est_pop_UB_{an}$	est l'estimation du nombre de résidents de la zone de santé *"a"* pour le mois "n", limite supérieure

2.3. Estimations des flux entrants par zone de santé

Nombre estimé de personnes qui sont entrées (c-a-d qui se sont installées) dans une zone de santé (à partir de toutes les autres zones de santé) entre le mois précédent et le mois en cours. En d'autres termes, il s'agit du **total de tous les flux entrants dans une zone de santé.**

2.3.1. Calcul

La somme des entrées dans la zone de santé *"a"* au cours du mois *"n"* est calculée comme étant la somme des changements de lieu de résidence estimés vers la zone de santé *"a"* depuis toutes les autres zones de santé *"b"* entre les mois *"m"* et" *n"* :

$$est_inflows_{amn} = \sum_{b=1}^{k} est_flows_{bamn}$$

Où :

est_flows_{bamn}	est le nombre de changements de lieu de résidence estimés vers la zone de santé *"a"* à partir de toutes les zones de santé *b* (b≠a) entre les mois *"m" et "n"*

Pour le calcul des changements de lieu de résidence bilatéraux estimés (est_flow_{bamn}), voir les **estimations de changements de lieu de résidence**.

2.3.2. Filtres et expurgations

Les valeurs inférieures à 15 ont été expurgées pour des raisons de protection des données.

2.4. Estimations des flux sortants par zone de santé

Nombre estimé de personnes ayant quitté une zone de santé (vers toutes les autres zones de santé) entre le mois précédent et le mois en cours. En d'autres termes, il s'agit du **total de tous les flux sortants d'une zone de santé.**

2.4.1. Calcul

La somme des flux sortants de la zone de santé *"a"* au cours du mois *"n"* est calculée comme étant la somme des changements de lieu de résidence estimés de la zone de santé *"a"* vers toutes les autres zones de santé *"b"* entre les mois *"m" et "n"* :

$$est_outflows_{amn} = \sum_{b=1}^{k} est_flows_{abmn}$$

Où :

Copyright © 2026 Flowminder Foundation - All Rights Reserved

est_flows_{abmn}	est le nombre de changements de lieu de résidence estimées de la zone de santé “a” vers toutes les zones de santé “b” (b≠a), entre les mois “m” et “n”

Pour le calcul des changements de lieu de résidence bilatéraux estimés (est_flows_{abmn}), voir les **estimations des changements de lieu de résidence**.

2.4.2. Filtres et expurgations

Les valeurs inférieures à 15 ont été expurgées pour des raisons de protection des données.

Copyright © 2026 Flowminder Foundation - All Rights Reserved

3. Estimations des changements de lieu de résidence entre zones de santé

3.1. Informations générales

3.1.1. Version
Version 2.0

3.1.2. Description
Ensemble de données contenant des estimations des changements de lieu de résidence de population mois par mois entre les zones de santé de la RDC, depuis juin-juillet 2023.

3.1.3. Unités temporelles et couverture
Chaque mois, les estimations des résidents couvrent un mois civil.

Les estimations des changements de lieu de résidence, d'un mois à l'autre, sont opérationnelles en tant que changements des lieux de séjour estimés ("lieux d'origine") entre le premier mois et le second mois. Par exemple, les estimations des changements de lieu de résidence pour mars 2024 couvrent les changements dans les lieux de séjour entre février et mars 2024.

3.1.4. Unités géographiques et couverture
Les estimations couvrent **les flux de population directionnels** entre les zones de santé de la RDC dans les 26 provinces du pays pour les zones de santé pour lesquelles des données CDR sont disponibles. Les limites des zones de santé issues des données DHIS2 sont utilisées.

3.1.5. Rédaction
Les valeurs inférieures à 15 ont été expurgées pour des raisons de protection des données.

3.2. Estimations des changements de lieu de résidence entre zones de santé

Le nombre estimé de personnes changeant de lieu de résidence d'une zone de santé à une autre zone de santé entre le mois en cours et le mois précédent.

3.2.1. Calcul
Les changements de lieu de résidence de la zone de santé *a* vers la zone de santé *b* entre les mois *m et n* sont estimés sur la base des agrégats CDR des changements de lieu de résidence, c'est-à-dire le nombre d'ID qui changent de domicile de la zone de santé *a* vers la zone de santé *b* entre ces mois.

L'emplacement du domicile est déterminé comme la zone de santé contenant les tours cellulaires qui ont le plus souvent (et au cours d'au moins trois semaines différentes) acheminé le dernier événement de la journée d'un identifiant au cours d'un mois calendaire. Pour chaque identifiant, les

Copyright © 2026 Flowminder Foundation - All Rights Reserved

changements de lieu de résidence sont alors détectés comme un changement dans la section communale du lieu de résidence d'un mois à l'autre. **Pour les mois pour lesquels les agrégats CDR font défaut, les déménagements sont calculés à partir de l'évolution du lieu de résidence entre le mois précédant le début de la période de lacune et le premier mois suivant cette période.**

Ensuite, les agrégats CDR des changements de lieu de résidence (cdr_flow_{abmn}) de la section communale *a* vers la section communale *b* entre les mois *m* et *n* sont ajustés, en tenant compte du ratio estimé SIM/utilisateur ($sims_a^{-1}$) et de la couverture de la population par les Données CDR ($median_12m(geom(pop_coverage_{am}, pop_coverage_{bm})^{-1})$).

$$est_flows_{abmn} = cdr_flows_{abmn} * sims_a^{-1} * median_12m(geom(pop_coverage_{am}, pop_coverage_{bm})^{-1})^{att}$$

Où :

est_flows_{abmn}	est l'estimation des changements de lieu de résidence de la zone de santé *"a"* vers la zone de santé *"b"* entre les mois *"m" et "n"*
cdr_flows_{abmn}	sont des changements de lieu de résidence de la zone de santé *"a"* vers la zone de santé *"b"*, dérivés du CDR, entre les mois *"m" et "n"*
geom()	*est la moyenne géométrique*
$sims_a$	*est le nombre de SIM par utilisateur dans la province d'origine de la zone de santé a*
median_12m()	*est la médiane sur 12 mois pour les mois m, m-1, …, m-12*
$pop_coverage_{am}$	*est la couverture de population des agrégats CDR dans la zone de santé a, mois m*
att	*est le facteur d'atténuation pour les pondérations (réduction de puissance)*

Le paramètre $sims_a$ est calculé comme une estimation à double cadre basée sur les données du microrecensement de 2021 et les données de l'enquête téléphonique de 2021, et saisit le nombre moyen de SIM de Vodacom par utilisateur de Vodacom par province d'origine. Le terme $sims_per_user_origin^{-1}$ réduit les agrégats CDR de changements de lieu de résidence pour tenir compte des enregistrements CDR multiples des mêmes utilisateurs. Ce paramètre est constant dans le temps.

Afin d'éviter d'importantes fluctuations des pondérations dues au faible nombre de localisations de logements certains mois, on utilise la médiane sur 12 mois de la moyenne géométrique. Cette médiane est ensuite réduite à l'aide d'un exposant (*att* = 0,5) afin de réduire la variance des pondérations. **Lorsque le calcul de la médiane sur 12 mois coïncide avec des lacunes dans les données CDR (juin-juillet 2021 et mai-juillet 2026), il utilise la valeur médiane d'une période plus courte, en excluant les mois pour lesquels les données CDR font défaut. Bien que cela signifie qu'une telle valeur sera calculée à partir d'une période inférieure à 12 mois, nous ne pensons pas que ces durées réduisent l'effet modérateur de la pondération par la médiane sur 12 mois.**

Copyright © 2026 Flowminder Foundation - All Rights Reserved

Pour les **estimations de la limite supérieure**, cette réduction d'échelle est mise en œuvre en utilisant l'exposant 0,7

$$est_flows_UB_{abmn} = cdr_flows_{abmn} * sims_{a}^{-1} * median_12m(geom(pop_coverage_{am}, pop_coverage_{bm})^{-1})^{0.7}$$

Pour les **estimations de la limite inférieure**, cette réduction d'échelle est mise en œuvre en utilisant l'exposant 0,3

$$est_flows_LB_{abmn} = cdr_flows_{abmn} * sims_{a}^{-1} * median_12m(geom(pop_coverage_{am}, pop_coverage_{bm})^{-1})^{0.3}$$

NOTE : Les changements de lieu de résidence font référence aux changements de lieu de résidence bilatéraux directionnels, de la zone de santé *"a"* vers la zone de santé *"b"*. Ils ne sont généralement pas égaux au nombre de changements de lieu de résidence de la *zone "b"* vers la *zone "a"*.

3.2.2. Filtres et expurgations

Les valeurs inférieures à 15 ont été expurgées pour des raisons de protection des données.

Copyright © 2026 Flowminder Foundation - All Rights Reserved

4. Sources de données

4.1. Métadonnées de téléphonie (données CDR), Vodacom RDC

Pour la version actuelle des estimations, on utilise les lieux de séjour mensuels et les changements de lieu de résidence d'un mois à l'autre dérivés des métadonnées de téléphonie mobile (comptes rendus d'appels (CRA) ou Call Detail Records en anglais (données CDR)), de Vodacom RDC.

À des fins de facturation, les opérateurs conservent un enregistrement des activités d'un identifiant (ID) dans une base de données. Ces enregistrements sont générés chaque fois qu'un ID passe ou reçoit un appel, envoie ou reçoit un SMS, ou utilise des données mobiles. C'est ce qu'on appelle les CDR (Call Detail Records). Les CDR contiennent des informations sur l'identifiant d'envoi et de réception d'un appel ou d'un SMS, la durée d'un appel ou le volume d'une session de données, ainsi que sur l'identifiant de la tour de téléphonie mobile qui achemine l'appel.

4.2. Microrecensement de la RDC, 2021, Flowminder/WorldPop

En coopération avec l'Institut National de la Statistique (INS) de la RDC, le groupe WorldPop de l'Université de Southampton et l'École de Santé Publique de Kinshasa (ESPK), Flowminder a commissionné et coordonné, en 2021, un microrecensement dans sept provinces de la RDC : Haut-Katanga, Haut-Lomami, Ituri, Kasaï, Kasaï-Oriental, Lomami et Sud-Kivu. Les données ont été collectées par le biais d'entretiens en face à face entre le 14 mars et le 27 avril 2021. L'ensemble de données final comprend des données sur 85 982 ménages et 367 792 individus.

4.3. Enquête téléphonique en RDC, 2021, Flowminder

Flowminder a commissionné et coordonné une enquête téléphonique auprès des utilisateurs de téléphone en RDC afin d'obtenir des informations empiriques sur l'utilisation du téléphone et sur la mobilité en RDC. L'objectif de l'enquête téléphonique était également de valider les méthodes de Flowminder pour la production d'estimations de la mobilité basées sur les données CDR, en particulier des concepts clés (localisation du domicile, mobilité, migration), et de connaître la structure socio démographique des utilisateurs de téléphone en RDC. La collecte des données s'est déroulée du 15 octobre au 8 novembre 2021. Des entretiens CATI ont été réalisés auprès de 7 523 répondants.

4.4. Estimations démographiques de l'UN OCHA, 2020

Les estimations de population 2020 par zone de santé publiées par UN OCHA sur la plateforme HDX.

4.5. Annuaires statistiques de l'INS, 2017 et 2020

L'Institut national de la statistique (INS) de la RDC a publié des annuaires statistiques en 2017 et 2020 qui contenaient des projections démographiques par province pour les années 2012 à 2019.

Copyright © 2026 Flowminder Foundation - All Rights Reserved

Les estimations démographiques pour 2019 et le taux de croissance annuel moyen (+3,3 %) dérivé de la série chronologique des projections sont utilisés pour calibrer les totaux par province.

4.6. Fichiers de forme (shapefiles) des zones de santé, DHIS2

Pour s'aligner sur les unités administratives utilisées dans le logiciel DHIS2 (District Health Information System 2.0), les agrégats et les estimations sont produits sur la base d'une correspondance entre les tours de téléphonie mobile cartographiées et les fichiers shapefiles des zones de santé du DHIS2. Le fichier shapefile pertinent, daté de 2024, couvre 519 zones de santé distinctes de la RDC.

Copyright © 2026 Flowminder Foundation - All Rights Reserved

5. Limitations

5.1. Biais “de mesure”

L'exactitude, la précision et la validité des statistiques dérivées des données CDR dépendent à la fois des erreurs ou des biais de "mesure" et de représentation.

En général, la **densité des tours de la téléphonie mobile** dans une zone donnée influe sur la **précision** des estimations de **localisation.** La densité des tours de la téléphonie mobile est généralement plus élevée dans les zones urbaines et plus faible dans les zones rurales.

La mesure dans laquelle les statistiques dérivées des données CDR constituent un indicateur fiable de la mobilité de la population dépend de la correspondance entre les **activités observables** à partir des données CDR disponibles et les activités réelles des utilisateurs. Les épisodes de voyage non ponctués par des communications (par exemple, lorsque les cartes SIM ne sont pas utilisées ou en dehors des zones couvertes) ne sont pas récupérables.

Inversement, une augmentation de la **fréquence d'utilisation du téléphone** au niveau individuel ou global peut correspondre à une augmentation de la mobilité mesurée, étant donné que les mouvements qui se produisaient auparavant entre les communications sont alors saisis. En outre, les données CDR enregistrées pour chaque identifiant ne sont pas distribuées uniformément dans le temps et dépendent de la fréquence à laquelle les utilisateurs lancent et/ou reçoivent des appels, envoient des SMS et/ou utilisent des sessions de données pour la connectivité internet sur leurs téléphones. Nous ne "voyons" un abonné que lorsqu'il utilise son téléphone. S'il n'utilise pas son téléphone un jour donné, nous ne pouvons pas dire avec certitude où il se trouve ce jour-là. Les abonnés dont la fréquence d'utilisation est faible ont tendance à être exclus de l'analyse - car les épisodes manquants des trajectoires individuelles ne peuvent pas être déduits, ou parce qu'ils apportent une quantité négligeable d'informations pour les agrégats de groupe. Une fréquence d'utilisation plus faible pourrait toutefois être associée à des schémas de mobilité différents et, en général, à une mobilité plus faible.

Une autre source de biais est la mesure dans laquelle les données CDR correspondant à un identifiant individuel correspondent réellement à une **seule personne.** Les individus peuvent utiliser plusieurs cartes SIM simultanément, ou changer fréquemment la carte SIM qu'ils utilisent, et une carte SIM peut être partagée entre plusieurs individus. La propriété peut également être transférée à un utilisateur différent ou à des groupes d'utilisateurs.

Les biais de mesure compliquent l'extraction d'informations à partir des CDR (sur la mobilité et sur la localisation de lieux significatifs tels que le domicile) et compliquent l'inférence sur la base des traces comportementales dans les données CDR.

5.2. Représentativité et biais

Les biais en termes de représentation peuvent être classés en trois catégories principales. Les principales **erreurs de couverture et de sélection** des statistiques dérivées des données CDR sont liées à :

Copyright © 2026 Flowminder Foundation - All Rights Reserved

- la **part des utilisateurs de téléphones mobiles** dans la population d'une région donnée
- la **part de marché de certains opérateurs** dans la zone, et
- le **sous-ensemble d'utilisateurs suffisamment actifs** dans la zone qui est utilisé pour les analyses.

Tout le monde n'a pas accès à un téléphone portable. Dans les pays à revenu faible et intermédiaire, la possession d'un téléphone portable et (dans une moindre mesure) l'accès à ce type d'appareil se sont avérés disproportionnés en faveur des hommes, des personnes instruites, des citadins, des riches et des personnes en âge de travailler, dont la mobilité peut différer de celle de la population en général.

Lorsque les données CDR ne sont disponibles qu'auprès d'un seul opérateur, cela peut ajouter une erreur de sous-couverture, car il n'est pas rare que les caractéristiques de la base d'abonnés d'un opérateur particulier soient biaisées en faveur d'individus relativement plus riches ou pauvres.

En d'autres termes, les données CDR ne saisissent qu'un **sous-ensemble non aléatoire de la population,** qui peut également changer au fil du temps en raison de l'évolution des prix des téléphones et des cartes SIM, de l'extension de la couverture des signaux, des stratégies de marketing et de tarification des opérateurs de réseau mobile, ainsi que des changements démographiques et sociétaux. Cela peut se traduire par une augmentation de l'utilisation du téléphone pour les anciens groupes à faible utilisation, et conduire à des changements dans la représentation de la population. Ces erreurs de sous-couverture et leurs variations temporelles peuvent fausser l'interprétation et l'utilisation des observations dérivées des CDR.

Les parts de marché et les zones de couverture des opérateurs évoluent dans le temps, ce qui aura un impact sur l'analyse des données CDR. En outre, le taux d'erreur dans la gestion des données des opérateurs et les problèmes de réseau est généralement élevé, ce qui rend l'analyse de leurs données plus difficile en raison des lacunes dans les données dans le temps et l'espace, ainsi que des incohérences. Ces problèmes sont suivis par Flowminder avant l'étape de l'analyse.

Contactez nous

Pour toute question ou information sur nos estimations sur la population et la mobilité en RDC, sur les méthodes présentées dans ce document ou sur l'analyse des données de téléphonie mobile, veuillez nous contacter à l'adresse **rdc@flowminder.org.**