

Cartographie des zéro-doses et des populations non vaccinées en Afrique de l'Ouest

Reach the Unreached – Digital technologies to map zero-dose and unreached children in West and Central Africa

Aperçu méthodologique de la cartographie de la population

WorldPop, Université de Southampton
30 novembre 2024

Des estimations précises des populations non et sous-immunisées dans diverses unités administratives sont essentielles pour une planification et une prestation efficaces de divers services basés sur les droits de l'homme dans les pays, en particulier les services de vaccination tels que la vaccination de routine et les campagnes, ainsi que les réponses humanitaires. Cependant, en raison de recensements obsolètes, de zones inaccessibles (en raison de conflits, de conditions météorologiques extrêmes, etc.), de mouvements de populations, de systèmes de gestion de l'information sanitaire médiocres et d'un manque de moyens, ces données sont souvent obsolètes, incomplètes ou indisponibles dans de nombreux pays. Pour combler cette lacune, l'initiative « Reach the Unreached » vise à développer ces ensembles de données essentiels et à les présenter dans des cartes numériques actuelles, fiables et exploitables. Ce rapport est axé sur la méthodologie de production de ces cartes.

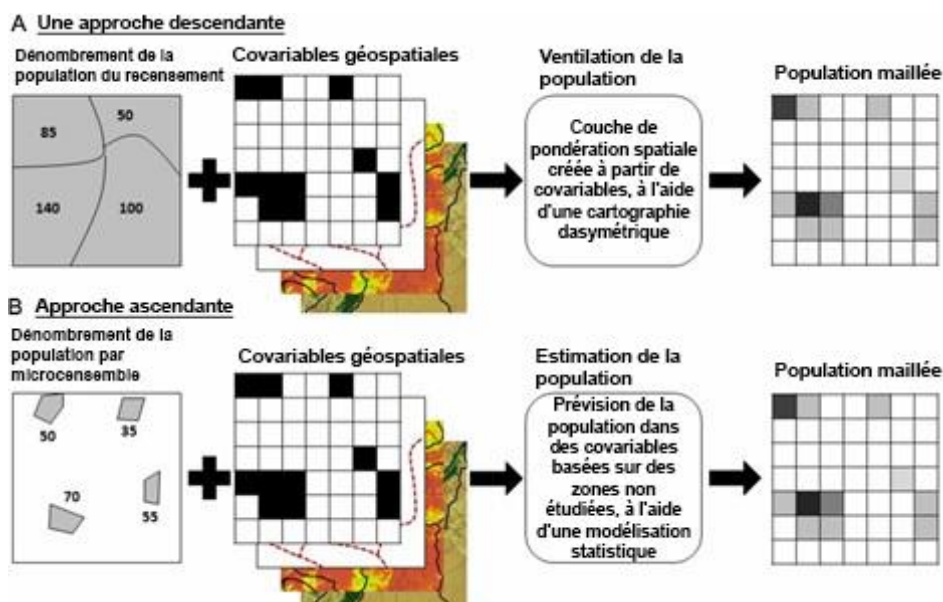


Figure 1 : Schéma de (A) approches de cartographie descendantes et (B) ascendantes. Les estimations des effectifs de populations sont supposées être limitées aux zones contenant des bâtiments résidentiels ; la densité de population dans ces zones est prédite en fonction d'ensembles de données auxiliaires tels que les réseaux routiers, la température ou la végétation verte.

Bref aperçu de la méthode

Pour produire des estimations de populations à haute résolution, le groupe de recherche WorldPop utilise deux principales méthodes : la méthode descendante et la méthode ascendante (Wardrop et al 2018, Figure 1). L'approche descendante utilise des données de population issues du recensement ou des projections démographiques subnationales, à couverture spatiale complète, et les désagrège pour produire des estimations de population à haute résolution. L'approche ascendante utilise des données de populations sur des zones échantillonnées (par exemple, les enquêtes de micro recensement ou les effectifs de population d'un échantillon de ménage enquêté) et prédit la population sur l'ensemble du territoire. Dans le cadre du projet Reach the Unreached, la méthode descendante est utilisée dans les pays où une estimation de la population était nécessaire.

La cartographie descendante est la technique de modélisation et de cartographie des populations la plus utilisée dans le domaine de la modélisation de la distribution de la population. Elle commence par des chiffres de recensement agrégés au niveau administratif et utilise une technique d'apprentissage automatique dite forêts aléatoires, ou « Random Forests » (RF) pour désagréger les chiffres de population au niveau du pixel. La méthode des RF ne produit pas des estimations de population de type uniforme, mais considère plutôt les variations à petite échelle des distributions de population, produisant une couche de population plus réaliste. L'algorithme RF intègre les relations spatiales entre la densité de population et une série de prédicteurs géospatiaux qui sont utilisées comme indicateurs de la présence humaine (par exemple, la topographie, l'utilisation des terres, la couverture terrestre, les lumières nocturnes, les lieux de conflit, etc.). Cette méthode évalue la qualité de l'ajustement, ainsi que l'ordre d'importance des différents prédicteurs dans la prédiction, et vérifie que la population désagrégée correspond à la population ayant servi de données d'entrée du modèle. La précision des estimations de population modélisées repose sur la résolution spatiale des prédicteurs géospatiaux, ainsi que sur le niveau de détail des données démographiques d'entrée. Ainsi, les variables géospatiales à haute résolution sont plus appropriées que les couches géospatiales à résolution moyenne ou grossière. De même, le fait de disposer de données démographiques plus détaillées (par exemple, au niveau de la commune plutôt qu'au niveau régional) peut améliorer de manière significative la précision spatiale des estimations de population carroyées.

Suivant l'approche de cartographie dasymétrique basée sur les RF (Stevens et al., 2015), le projet utilise le package popRF 'R' (Bondarenko et al., 2021) basé sur l'algorithme de Breiman (2001). Le package popRF produit comme résultat final une estimation à haute résolution de la population totale, avec une résolution spatiale de 100 x 100 m. Les caractéristiques démographiques, telles que la proportion de personnes appartenant à différents groupes d'âge et de sexe, sont estimées dans un deuxième temps, en multipliant les totaux de la population carroyée par les données de structures d'âge et de sexe existantes (Pezzulo et al., 2017).

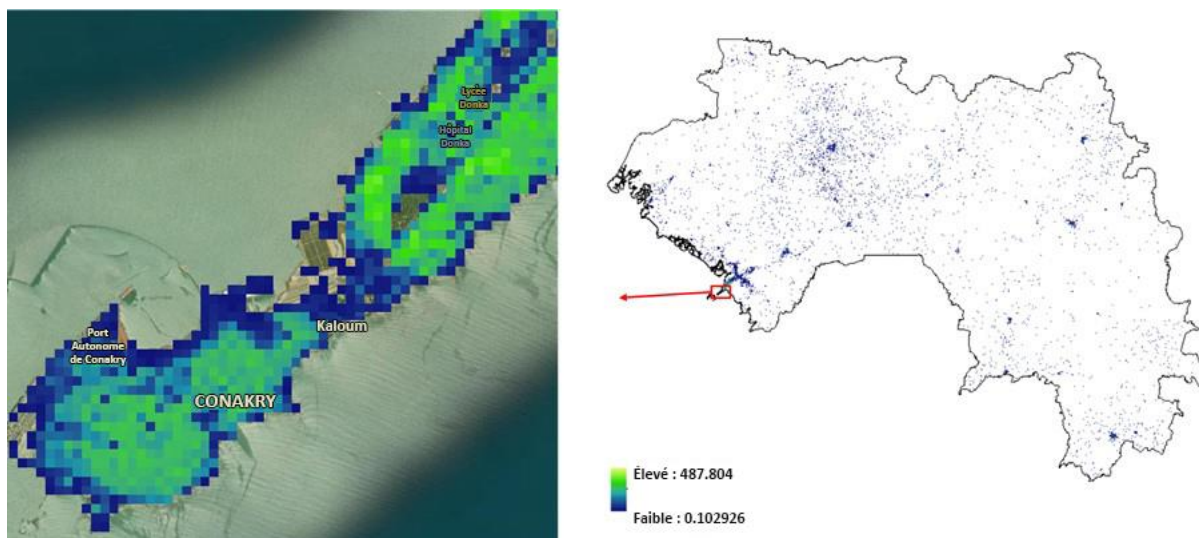


Figure 2 : La couche de population à haute résolution quadrillée pour la Guinée avec une vue agrandie de Conakry.

Spécification des résultats

Le format de sortie de la modélisation de la population est une représentation carroyée de la distribution de la population (c.-à-d. format matriciel, Figure 2). La trame géotiff contient des estimations de la taille totale de la population pour chaque cellule de grille d'environ 100 m. La valeur de chaque cellule de grille représente les estimations de la taille de la population totale, là où les établissements humains sont attendus, d'où la redistribution des effectifs de population aux empreintes des bâtiments seulement. En outre, trente-six trames désagrégées par âge/sexe sont fournies séparément en fonction des groupes d'âge/sexe (moins de 1 an, 1-4, 5-9, 10-14, 75-79, 80+). Quatre trames supplémentaires sont également fournies pour représenter des groupes démographiques souvent ciblés par les programmes et interventions. Il s'agit de (i) toutes les femmes et tous les hommes de moins d' 1 an, (ii) toutes les femmes et tous les hommes de moins de 5 ans, (iii) toutes les femmes et tous les hommes de moins de 15 ans et (iv) toutes les femmes âgées de 15 à 49 ans, inclusivement.

L'année de référence des estimations de population correspond aux données d'entrée utilisés. Ainsi, si les projections subnationales du recensement de 2021 sont utilisées pour la désagrégation descendante, la couche de population résultante aura comme année de référence 2021.

Hypothèses et limites

Les estimations de population à haute résolution sont basées sur des données du recensement, ou sur des projections démographiques calculées à partir du recensement. Les dénombrements des effectifs de population issus du recensement sont supposés être exhaustifs, sans omission de ménages ou de groupes de population. Mais les projections subnationales issues du recensement pourraient ne pas refléter la situation du moment, compte tenu du temps écoulé depuis le dernier recensement et des hypothèses de modélisation qui peuvent ne pas tenir compte des mouvements de population importants.

Les estimations de population carroyées sont contraintes aux zones habitées, qui sont acquises à partir d'images satellites. L'on suppose que les données relatives à l'empreinte des bâtiments, à partir desquelles ont été dérivées différentes métriques de caractéristiques du bâti, sont exactes et que chaque polygone de bâtiment correspond à une structure de bâtiment réelle. Dans les endroits où la couverture nuageuse est persistante, ou qui ont récemment connu des changements rapides de peuplement, par exemple l'établissement de nouvelles colonies, la croissance urbaine rapide ou l'abandon de colonies, les estimations de la population sont susceptibles d'être moins précises.

Disponibilité des logiciels et ressources d'apprentissage

Le site de téléchargement du package popRF : <https://github.com/wpdp/popRF>

Tutoriels sur la méthode appliquée :

Partie 1 -La méthode des

forêtsaléatoires :<https://data.worldpop.org/repo/docs/leasure2021small/>

Partie 2 - Le package popRF : <https://data.worldpop.org/repo/docs/lazar2021poprf>

Remerciements

The Reach the Unreached – Digital technologies to map zero-dose and unreached children in West and Central Africa project is funded by UNICEF — The United Nations Children’s Fund (contract No. 43387656). Le projet est dirigé par le Bureau Régional de l'UNICEF pour l'Afrique de l'Ouest, avec la participation de plusieurs partenaires, notamment les bureaux nationaux de l'UNICEF, WorldPop de l'Université de Southampton, MapAction et CartONG.

Ce rapport sur la méthodologie a été rédigé par Attila N Lazar et Assane Gadiaga. Les contributions de toute l'équipe WorldPop à la méthode développée et aux applications actuelles sont également reconnues.

Citation suggérée

WorldPop 2024. Aperçu méthodologique de la cartographie de la population. Mapping zero-dose and unreached populations in West Africa: Reach the Unreached – Digital technologies to map zero-dose and unreached children in West and Central Africa project. WorldPop, Université de Southampton, 30 novembre 2024.

Licence

Ce rapport peut être redistribué selon les termes d'une licence **Creative Commons Attribution 4.0 International (CC BY 4.0)**

Références

- Bondarenko M., Nieves J. J., Stevens F. R., Gaughan A. E., Tatem A. et Sorichetta A. 2020. wpgpRFPMS : Random Forests population modelling R scripts, version 0.1.0. Université de Southampton : Southampton, Royaume-Uni. <https://dx.doi.org/10.5258/SOTON/WP00665>
- Breiman, L. Random forests. *Mach. Learn.* 45, 5–32 (2001).
- Pezzulo C, Hornby GM, Sorichetta A, Gaughan AE, Linard C, Bird TJ, Kerr D, Lloyd CT, Tatem AJ. 2017. Sub-national mapping of population pyramids and dependency ratios in Africa and Asia. *Sci. Data* 4:170089. <https://dx.doi.org/10.1038/sdata.2017.89>
- Stevens, F. R., Gaughan, A. E., Linard, C. & Tatem, A. J. Disaggregating Census Data for Population Mapping Using Random Forests with Remotely-Sensed and Ancillary Data. *PLoS ONE* 10, e0107042 (2015). <https://doi.org/10.1371/journal.pone.0107042>
- Wardrop, N. A., W. C. Jochem, T. J. Bird, H. R. Chamberlain, D. Clarke, D. Kerr, L. Bengtsson, S. Juran, V. Seaman et A. J. Tatem (2018). "Spatially disaggregated population estimates in the absence of national population and housing census data." *Proceedings of the National Academy of Sciences* 115(14): 3529-3537.