

NE RAYER PERSONNE DE LA CARTE

GUIDE DES DONNÉES DÉMOGRAPHIQUES
MAILLÉES POUR LE DÉVELOPPEMENT DURABLE

2020





NE RAYER PERSONNE DE LA CARTE

GUIDE DES DONNÉES DÉMOGRAPHIQUES MAILLÉES POUR LE DÉVELOPPEMENT DURABLE

Rapport du Thematic Research Network on Data and Statistics (TReNDS) du Réseau des solutions pour le développement durable (SDSN) des Nations unies, en soutien au POPGRID Data Collaborative

2020

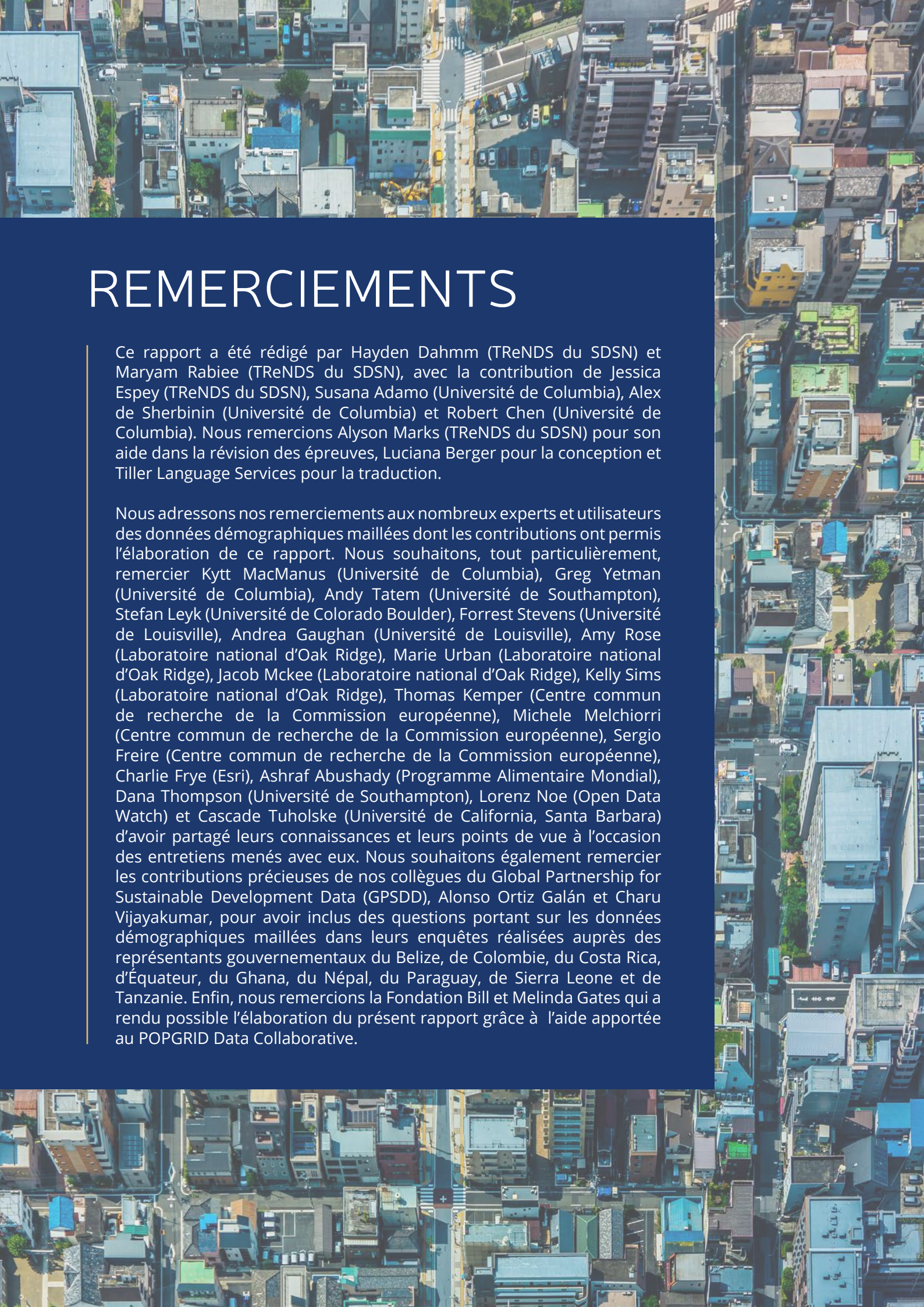
AVERTISSEMENT

Ce rapport a été élaboré par le TReNDS du SDSN en soutien au POPGRID Data Collaborative. Les résultats, avis et analyses présentés dans le présent rapport ne reflètent pas nécessairement le point de vue des gouvernements, des organisations ou des organismes cités. Ce rapport est disponible en espagnol et en français.



Le présent rapport, dans son contenu, est protégé par la licence émise par Creative Commons Attribution 4.0 International.

Conception : Luciana Berger



REMERCIEMENTS

Ce rapport a été rédigé par Hayden Dahmm (TReNDS du SDSN) et Maryam Rabiee (TReNDS du SDSN), avec la contribution de Jessica Espey (TReNDS du SDSN), Susana Adamo (Université de Columbia), Alex de Sherbinin (Université de Columbia) et Robert Chen (Université de Columbia). Nous remercions Alyson Marks (TReNDS du SDSN) pour son aide dans la révision des épreuves, Luciana Berger pour la conception et Tiller Language Services pour la traduction.

Nous adressons nos remerciements aux nombreux experts et utilisateurs des données démographiques maillées dont les contributions ont permis l'élaboration de ce rapport. Nous souhaitons, tout particulièrement, remercier Kytt MacManus (Université de Columbia), Greg Yetman (Université de Columbia), Andy Tatem (Université de Southampton), Stefan Leyk (Université de Colorado Boulder), Forrest Stevens (Université de Louisville), Andrea Gaughan (Université de Louisville), Amy Rose (Laboratoire national d'Oak Ridge), Marie Urban (Laboratoire national d'Oak Ridge), Jacob Mckee (Laboratoire national d'Oak Ridge), Kelly Sims (Laboratoire national d'Oak Ridge), Thomas Kemper (Centre commun de recherche de la Commission européenne), Michele Melchiorri (Centre commun de recherche de la Commission européenne), Sergio Freire (Centre commun de recherche de la Commission européenne), Charlie Frye (Esri), Ashraf Abushady (Programme Alimentaire Mondial), Dana Thompson (Université de Southampton), Lorenz Noe (Open Data Watch) et Cascade Tuholske (Université de California, Santa Barbara) d'avoir partagé leurs connaissances et leurs points de vue à l'occasion des entretiens menés avec eux. Nous souhaitons également remercier les contributions précieuses de nos collègues du Global Partnership for Sustainable Development Data (GPSDD), Alonso Ortiz Galán et Charu Vijayakumar, pour avoir inclus des questions portant sur les données démographiques maillées dans leurs enquêtes réalisées auprès des représentants gouvernementaux du Belize, de Colombie, du Costa Rica, d'Équateur, du Ghana, du Népal, du Paraguay, de Sierra Leone et de Tanzanie. Enfin, nous remercions la Fondation Bill et Melinda Gates qui a rendu possible l'élaboration du présent rapport grâce à l'aide apportée au POPGRID Data Collaborative.

ACRONYMES

ADAM	Automatic Disaster Analysis and Mapping (Cartographie et analyse automatique des catastrophes)	IFPRI	International Food Policy Research Institute (Institut international de recherche sur la politique alimentaire)
CCR	Centre commun de recherche de la Commission européenne	LSMS	Living Standards Measurement Study (Enquête sur la mesure des niveaux de vie)
CIAT	Centre international d'agriculture tropicale	ODD	Objectifs de développement durable
CIESIN	Center for International Earth Science Information Network (Centre pour un réseau d'information international en sciences de la Terre)	OMS	Organisation Mondiale de la Santé
CSNU	Commission de statistique des Nations unies	ONU	Organisation des Nations unies
DMSP-OLS	Defense Meteorological Satellite Program Optical Linescan System (Programme de satellites météorologiques de la défense - Système de caméra optique)	ONS	National Statistics Office (Office national des statistiques)
DPNU	Division de la population des Nations unies	ORNL	Laboratoire national d'Oak Ridge
DTC	Diphtérie, tétanos, coqueluche	PAGER	Prompt Assessment of Global Earthquakes for Response (Évaluation immédiate des tremblements dans le monde pour mobilisation de la réponse)
ETM	Landsat Enhanced Thematic Mapper (Instrument de cartographie thématique amélioré de Landsat)	PAM	Programme Alimentaire Mondial
GHSL	Global Human Settlement Layer (Coupe de l'établissement humain dans le monde)	PAR	Population à risque
GHS-BUILT	Global Human Settlement Layer Built Up Extent (Coupe de l'établissement humain dans le monde – extension urbanisation)	POPGRID	POPGRID Data Collaborative
GHS-POP	Global Human Settlement Layer – Population (Coupe de l'établissement humain dans le monde – population)	SEDAC	NASA Socioeconomic Data and Applications Center (Centre d'applications et de données socio-économiques de la NASA)
GPSDD	Partenariat mondial pour des données sur le développement durable	SIG	Systèmes d'information géographique
GPW	Gridded Population of the World (Population maillée dans le monde)	TReNDS	Thematic Research Network on Data and Statistics (Réseau de recherche thématique sur les données et les statistiques)
GRID3	Geo-Referenced Infrastructure and Demographic Data for Development Initiative (Infrastructure géoréférencée et données démographiques pour l'Initiative pour le développement)	USGS	U.S. Geological Survey (Bureau des études géologiques des États-Unis)
GRUMP	Global Rural Urban Mapping Project (Projet mondial de cartographie urbaine et rurale)	VIIRS	Visible Infrared Imaging Radiometer Suite (Suite radiométrique d'imagerie visible à infrarouge)
HRSL	High Resolution Settlement Layer (Coupe à haute résolution de l'établissement humain)	WPE	World Population Estimate (Estimations de la population mondiale)

TABLE DES MATIÈRES

ACRONYMES	4
RÉSUMÉ	7
INTRODUCTION	10
RECENSEMENTS : DIFFICULTÉS ET OPPORTUNITÉS	11
LE RECENSEMENT RESTE PRIMORDIAL	12
CHAPITRE 1 : UTILISATION DES DONNÉES DÉMOGRAPHIQUES MAILLÉES DANS LE MONDE	13
1.1 APPLICATIONS IMMÉDIATES	13
ÉVALUATION DES RISQUES DE CATASTROPHE ET INTERVENTION	13
SANTÉ ET MALADIES INFECTIEUSES	15
PLANIFICATION D'ENQUÊTE	16
1.2 APPLICATIONS À MOYEN TERME	16
EXAMEN DES DÉCISIONS POLITIQUES	17
LES OBJECTIFS DE DÉVELOPPEMENT DURABLE	18
TABLEAU 1 - ODD EXIGEANT DES INFORMATIONS DÉMOGRAPHIQUES	20
1.3 APPLICATIONS À LONG TERME	21
HAUSSE DU NIVEAU DE LA MER	21
DISPONIBILITÉ DES RESSOURCES EN EAU	22
1.4 ÉLARGIR LE CHAMP D'APPLICATION ET COMPRENDRE LES OPTIONS DISPONIBLES	22
CHAPITRE 2 : SE FAMILIARISER AVEC LES JEUX DE DONNÉES DÉMOGRAPHIQUES MAILLÉES	23
2.1 DIFFÉRENCES DANS LES PARAMÈTRES DE MODÉLISATION	23
DONNÉES SAISIES	24
AJUSTEMENT DES DONNÉES	26
MÉTHODE DE MODÉLISATION	26
2.2 CARACTÉRISTIQUE DES SORTIES	27
2.3 VUE D'ENSEMBLE DES JEUX DE DONNÉES DÉMOGRAPHIQUES MAILLÉES À L'ÉCHELLE MONDIALE	28
GRIDDED POPULATION OF THE WORLD (GPW) V4	29
GLOBAL RURAL URBAN MAPPING PROJECT (GRUMP)	30
GLOBAL HUMAN SETTLEMENT LAYER - POPULATION (GHS-POP)	31
WORLD POPULATION ESTIMATE (WPE)	32
HIGH RESOLUTION SETTLEMENT LAYER (HRSL)	33
LANDSCAN	34
WORLDPOP	35
TABLEAU 2 PARAMÈTRES DES JEUX DE DONNÉES DÉMOGRAPHIQUES MAILLÉES POPGRID	36

TABLE DES MATIÈRES

2.4 AUTRES APPROCHES DEMODÉLISATION	37
UTILISATION DENOUVELLES SOURCES DE DONNÉES	37
2.5 CHOIX DE LA SOLUTIONADAPTÉE	38
.....	
CHAPITRE 3 :COMPARAISON CROISÉE ET VALIDATION DES JEUX DE DONNÉES MAILLÉES	39
3.1 COMPARAISON ENTRE LESJEUX DE DONNÉES	39
DESCRIPTION DESDIFFÉRENCES ENTRE LES ESTIMATIONS	40
INCIDENCES SUR LESRÉSULTATS	40
ESTIMATION DES EFFETS DUCHANGEMENT CLIMATIQUE	41
3.2 MÉTHODES DEVALIDATION DES ESTIMATIONS DÉMOGRAPHIQUES	41
COMPARAISON AVEC LESDONNÉES DE RECENSEMENT	41
COMPARAISON DES MÉTHODES ÉCARTS RÉGIONAUX	42
SUPÉRIORITÉ DES DONNÉES ÀHAUTE RÉOLUTION	43
VARIATION DES ERREURS EN FONCTION DE LA GÉOGRAPHIE	44
3.3 AUTRES CATÉGORIES DE DONNÉES DE VALIDATION	45
3.4 NÉCESSITÉ D'UNTRAVAIL DE VALIDATION COMPLÉMENTAIRE	45
.....	
CHAPITRE 4 :DIRECTIVES POUR LA PRISE DE DÉCISIONS	46
4.1 IMPORTANCE DEL'EMPLACEMENT DES POPULATIONS À UN MOMENT DONN	47
4.2 CONNAÎTRE L'APPROCHEDE MODÉLISATION UTILISÉE POUR RÉPARTIR LA POPULATION	47
4.3 IMPORTANCE DUREGROUPEMENT DÉMOGRAPHIQUE, Y COMPRIS PAR ÂGE ET PAR SEXE	48
4.4 LES ÉTUDES PAR PAYSSONT DÉLICATES	48
4.5 IDENTIFIER LESPÉRIODES DE TEMPS PERTINENTES	49
4.6 PRENDRE EN COMPTE LESDIFFÉRENCES ENTRE LES ESTIMATIONS URBAINES ET RURALES	49
4.7 LES FACTEURSENVIRONNEMENTAUX ONT UNE INCIDENCE SUR LA RÉPARTITION DE LA POPULATION	49
4.8 ENVISAGER LESCOVARIABLES	50
4.9 SE RENSEIGNER SUR LESCOÛTS ET DROITS DE RÉUTILISATION	50
MESSAGES POUR ÉVITER LES MÉPRISES	51
.....	
CONCLUSIONS	52
.....	
RÉFÉRENCES	54
.....	
ANNEXE A	61

RÉSUMÉ

Chaque année, près de 160 millions de personnes sont victimes de catastrophes naturelles. Après une catastrophe naturelle, chaque seconde compte pour sauver des vies et garantir que les fournitures indispensables parviennent aux personnes sinistrées. Pour identifier rapidement et avec exactitude les communautés touchées par la catastrophe, les organismes de secours, tels que le Programme Alimentaire Mondial (PAM) des Nations unies, utilisent les données démographiques maillées pour calculer, presque en temps réel, l'impact des tremblements de terre et des tempêtes tropicales sur les populations et les infrastructures. L'utilisation des données démographiques maillées s'avère essentielle pour définir les secours à mobiliser dans les situations d'urgence et minimiser les souffrances.

Disposer promptement de données fiables sur la population peut faire la différence entre la vie et la mort pour les personnes plongées dans des situations de crise ou vivant en zones de conflits. Ces données sont cruciales pour faire face à ces défis, mais également pour la prise de décision et la planification des interventions. Nous devons savoir où sont situées les populations, dans quelle situation elles se trouvent, quelles sont les infrastructures disponibles et à quels services de base elles ont accès. Pour calculer le nombre de personnes situées dans des zones difficiles d'accès ou prisonnières de zones de conflits, il n'y a pas de solution miraculeuse. Aussi, les données démographiques maillées offrent-elles un moyen prometteur de disposer de données exploitables dans des circonstances difficiles.

En 2015, lorsque 193 dirigeants mondiaux se sont mis d'accord sur les 17 Objectifs du développement durable (ODD), ils ont également promis de « ne laisser personne pour compte ». Or, en l'absence de données démographiques fiables et actualisées portant sur une aire géographique donnée, il est difficile de garantir que tout le monde sera pris en compte et que personne ne sera laissé pour compte. La plupart des gouvernements et décideurs politiques s'en remettent aux sources de données conventionnelles, telles que les enquêtes sur les ménages et les recensements de la population. Cependant, le développement des politiques et des programmes nécessaires pour éradiquer la pauvreté et améliorer la santé, l'éducation et les services de base ne peuvent se contenter des sources conventionnelles dont les limites, tant géographiques, temporelles que logistiques, sont nombreuses. Ainsi, à l'heure où nous rédigeons le présent rapport, près de 60 pays prévoient un retard confirmé ou prévisible de leurs activités de recensement, en raison de la pandémie du COVID-19 (FNUAP, n.d.). Il est heureux que les récents progrès en matière de capacité d'observation de la terre et de méthodes statistiques permettent désormais d'obtenir des estimations plus fréquentes et plus fines de la population dans le monde, grâce aux jeux de données démographiques maillées.

La cartographie de population maillée (ou matricielle), composée d'un quadrillage de rangées et de colonnes, illustre la répartition de la population, référencée à l'aide des coordonnées de latitude et de longitude. De plus en plus de fournisseurs de données associent les informations de recensement aux géoréférencement par satellite pour répartir la population et produire des données démographiques maillées. En dépit de ces progrès, il règne encore un certain degré de confusion, ou peut-être un manque de connaissance, au sujet des données démographiques maillées. Le grand nombre de jeux de données désormais disponible peut intimider les utilisateurs, en particulier ceux dont le temps et l'expertise technique sont insuffisants pour comprendre la différence entre les diverses options et en évaluer les avantages et les inconvénients selon la finalité recherchée.

Le POPGRID Data Collaborative, mis en place en 2018, entend répondre à ces nombreuses difficultés, en mettant en lien les utilisateurs de données, les prestataires et les parties prenantes, tant du secteur public que privé, qui travaillent à partir de données géoréférencées sur la population, les regroupements humains et les infrastructures.

Dans le présent rapport, Ne rayer personne de la carte : guide des données démographiques maillées pour un développement durable, notre objectif est de combler le manque connaissances en essayant d'améliorer la compréhension et l'accès aux fichiers de données démographiques maillées pour les décideurs politiques et autres utilisateurs. L'élaboration du présent rapport obéit à deux grands axes :

- Dans quelle mesure les données démographiques maillées peuvent-elles compléter les sources de données sur la population actuellement disponibles et permettre aux utilisateurs de la communauté du développement durable de prendre des décisions rapides et éclairées ?
- Quels sont les jeux de données démographiques maillées les plus adaptés à la finalité recherchée ?

S'inspirant de l'abondante littérature disponible et des entretiens menés auprès des principaux fournisseurs et utilisateurs du POPGRID Data Collaborative, le présent rapport propose un aperçu, une analyse et des recommandations sur l'utilisation des fichiers démographiques maillés dans divers champs d'application, tels que la réponse aux situations de catastrophe, les interventions sanitaires et la planification d'enquêtes. Le présent rapport compare, notamment, sept jeux de données démographiques maillées du POPGRID Data Collaborative, c'est-à-dire l'analyse des données de base, des méthodes et des hypothèses de base, ainsi que les avantages et les inconvénients de chaque fichier de données, dans une terminologie simple. Il propose également une analyse comparative des divers fichiers de données et des résultats obtenus. Enfin, il répond aux nombreuses idées fausses concernant les données démographiques maillées et conclut en présentant neuf principes dont l'objectif est d'orienter les utilisateurs dans leur processus de sélection.



LES MESSAGES CLÉS DU RAPPORT SONT :

LES DONNÉES DE RECENSEMENT RESTENT IMPORTANTES ; LES DONNÉES DÉMOGRAPHIQUES MAILLÉES NE REMPLACENT PAS LES DONNÉES DE RECENSEMENT.

Les jeux de données présentés dans le présent rapport viennent étayer les recensements nationaux en y apportant des éléments supplémentaires portant sur les calculs relatifs à la population, lesquels peuvent être mis à jour plus fréquemment ou calculés selon un degré de résolution spatiale plus élevé.

LES DONNÉES DÉMOGRAPHIQUES MAILLÉES SONT IMPARFAITES.

Bien qu'elles permettent de compenser certaines des limites liées aux sources conventionnelles, les méthodes utilisées pour produire ces estimations comportent leurs propres facteurs d'incertitude. Les utilisateurs doivent en être informés et être conscients de ces incertitudes possibles.

LES UTILISATEURS DE DONNÉES DOIVENT TENIR COMPTE DE PLUSIEURS FACTEURS DANS LE CHOIX DU JEU DE DONNÉES SELON LA FINALITÉ RECHERCHÉE POUR RÉPONDRE À LEURS BESOINS.

Chaque jeu de données est établi sur la base de diverses données sous-jacentes et d'hypothèses variées. Les décideurs politiques et les chercheurs doivent ainsi examiner avec rigueur les caractéristiques de chaque jeu de données, à savoir les caractéristiques démographiques, la résolution spatiale, les intervalles de temps, le coût des données et les droits d'utilisation, entre autres éléments.

DES TRAVAUX DE VALIDATION SUPPLÉMENTAIRES SONT NÉCESSAIRES POUR POUVOIR COMPARER LES CALCULS OBTENUS À PARTIR DES DONNÉES DÉMOGRAPHIQUES MAILLÉES ET CEUX OBTENUS À PARTIR DES DONNÉES QUI FONT AUTORITÉ EN MATIÈRE DE RÉPARTITION DE LA POPULATION.

Il est urgent d'entamer une analyse systématique de ces produits et un processus de validation objectif pour mieux affiner les méthodes et améliorer le degré de précision et la finalité des modèles. Ce travail est actuellement mené grâce au POPGRID Data Collaborative.

Il ne nous reste plus que 10 ans pour atteindre les ODD. Nous sommes donc à un croisement. Les données démographiques maillées participent d'ores et déjà à la poursuite de ces objectifs ambitieux en améliorant l'accès, la cohérence et la désagrégation spatiale des indicateurs ODD, permettant ainsi aux initiatives nationales et internationales de mieux cibler leurs efforts en vue des ODD et de toucher les populations qui, sans elles, seraient laissées pour compte. Ces données ne se révéleront utiles que pour autant que les décideurs politiques en comprennent les limites, les applications et la portée. De plus, elles n'atteindront leur plein potentiel que lorsqu'elles auront été intégralement validées au regard des données concrètes se rapportant au monde réel. Il est donc indispensable d'accélérer ces travaux de recherche majeurs et la généralisation de l'application des données démographiques maillées à l'échelle mondiale, si l'on veut atteindre les ODD et ne rayer personne de la carte.

INTRODUCTION

En 2019, la population mondiale a atteint le chiffre de 7,7 milliards d'habitants (Nations unies, n.d.). Bien que la population ait presque quintuplée depuis le début du XXe siècle, nous ne connaissons pas, ou peu, les conditions de vie d'une bonne partie de la population. C'est d'autant plus vrai pour nombre de ces communautés parmi les plus vulnérables et difficiles d'accès. En 2015, quand les Objectifs de développement durable (ODD) ont été adoptés, 193 pays se sont engagés à « ne laisser personne pour compte » (Nations unies, 2016). Cependant, en l'absence de données géoréférencées précises sur la localisation de ces populations, il nous est impossible d'atteindre ces objectifs et de veiller à comptabiliser l'ensemble de la population.

En règle générale, les connaissances que nous avons de la population proviennent des recensements, c'est-à-dire du décompte officiel de tous les habitants vivant dans un pays donné. Les Nations unies (ONU) définissent cet exercice comme étant « l'ensemble des opérations qui consistent à recueillir, grouper, analyser et publier les données démographiques, économiques et sociales se rapportant, à un moment donné, à tous les habitants d'un pays ou d'une partie bien déterminée d'un pays » (Nations unies, n.d.). Ce décompte fournit des renseignements précieux sur les communautés, permet d'orienter les services, peut déterminer la représentation politique et bien plus encore (Dahmm, 2018). Au total, 214 pays ou régions ont organisé un recensement au cours de l'année 2010 (DSNU, n.d.). De nombreux pays ne disposent, toutefois, que de données anciennes qui remontent à plus de 10 ans, en particulier les pays à faible revenu ou les pays plongés dans le conflit. Ainsi, le dernier recensement réalisé à Madagascar date de 2003, celui de la République démocratique du Congo remonte à 1984, en Afghanistan, il date de 1979, en Somalie de 1975 et au Liban de 1932 (Wardrop et al., 2018).¹ En outre, même lorsque les recensements sont réalisés à intervalles réguliers, des difficultés peuvent surgir et empêcher un comptage exhaustif. Les recenseurs se heurtent parfois à la barrière linguistique dans certaines communautés indigènes et groupes minoritaires et les zones touchées par les catastrophes ou les situations de guerre peuvent se révéler difficiles d'accès (pays partenaires GPSDD, 2019). De plus, dans certains pays, les conflits ethniques peuvent également et systématiquement se traduire par un calcul à la baisse des minorités, les communautés vulnérables étant susceptibles d'être toujours sous représentées. Par ailleurs, les conditions sur le terrain peuvent avoir considérablement changé en l'espace de dix ans, au gré des naissances, des décès et des déménagements, ce qui peut avoir une incidence considérable sur l'économie des pays. Au Salvador, par exemple, l'utilisation d'un recensement aux données caduques pour allouer des fonds municipaux se serait soldée par un détournement de près de 92 millions de dollars américains (valeur réelle du dollar américain en 2018), entre 2000 et 2007 (Roseth et al., 2019). Les recensements fournissent des données irremplaçables mais, depuis une vingtaine d'années, l'utilisation d'autres outils complémentaires, tels que les données démographiques maillées, permettent de contourner ces difficultés, entre autres.

Les produits de données démographiques maillées proposent des estimations chiffrées de la population selon des cellules de grille homogènes sur l'ensemble de la planète (Leyk et al., 2019). Cette catégorie de données permet de mieux comprendre l'emplacement des zones d'implantation des populations. Les grilles de population à modélisation minimum répartissent les données de recensement de façon équitable dans les cellules de grille, selon l'hypothèse de la répartition homogène des habitants. D'autres jeux de données incluent l'imagerie satellitaire ainsi que de nombreuses autres sources de données pour affiner les estimations de répartition, ce qui permet une mise à jour régulière des estimations.

Ces fichiers de données matricielles comportent un certain nombre d'avantages :

- Ils fournissent des estimations documentées et actualisées sur la répartition de la population.
- Ils facilitent l'utilisation directe de ces données en parallèle avec d'autres sources de données scientifiques disponibles sous format matriciel.
- Dans les zones n'ayant pas fait l'objet d'un recensement récent, ils fournissent un point de départ pour les études et l'évaluation des besoins.

Fort heureusement, ces jeux de données innovants tendent à être intégrés aux « processus de décision » de nombreux intervenants. Cependant, un sondage récemment mené par le Global Partnership for Sustainable Development Data (GPSDD) auprès des bureaux nationaux des statistiques a révélé que de nombreux agents ignorent encore l'existence de ces ressources (pays partenaires GPSDD, 2019). En novembre 2019, le GPSDD a réalisé une enquête auprès des pays partenaires afin d'évaluer leurs connaissances sur les fichiers démographiques maillés et l'utilisation qu'ils en font (voir annexe A pour l'enquête). Cette enquête comprend les entretiens menés auprès des représentants officiels des gouvernements du Belize, de Colombie, du Costa Rica, d'Équateur, du Ghana, du Népal, du Paraguay, de Sierra Leone et de Tanzanie. L'analyse des résultats de l'enquête par pays est reprise dans les diverses sections du présent rapport.²

Ce rapport se propose de dresser l'état des connaissances en matière de fichiers démographiques maillés afin que le potentiel, autant que les limites, de ce type de données soient plus largement reconnus et que les forces dont ils sont porteurs soient plus accessibles aux décideurs politiques décidés à œuvrer en faveur des ODD.

ENCADRÉ 1 | RECENSEMENTS : DIFFICULTÉS ET OPPORTUNITÉS

Pour organiser un recensement, les pays doivent surmonter plusieurs difficultés, lesquelles ont été discutées avec les représentants des bureaux nationaux de statistique qui ont été interviewés. Ainsi, les représentants de la Colombie, du Costa Rica et de l'Équateur évoquent le manque de ressources financières parmi les éléments qui empêchent l'aboutissement des opérations de recensement, dans les délais établis. Par ailleurs, en Colombie, accéder aux communautés qui vivent dans les zones reculées ou gangrénées par la violence s'avère difficile et les recenseurs se heurtent à la barrière linguistique et à la résistance des habitants. Au même moment, au Belize, l'Office national des statistiques rencontrait des difficultés de recrutement de recenseurs en nombre suffisant et des complications logistiques accrues causées par les modifications intervenues dans la répartition des habitants. Le Népal, quant à lui, a été confronté à des erreurs commises dans l'identification des ethnies et des castes. Le recensement peut donc poser de graves problèmes d'inexactitude. Ainsi, le dernier recensement organisé au Paraguay, en 2012, semble être passé à côté d'un quart de la population (pays partenaires GPSDD, 2019).

Fort heureusement, il existe de nouvelles méthodes et technologies qui permettent d'améliorer le processus en augmentant le niveau d'efficacité et le degré d'exactitude. Par exemple, en 2018, le Département administratif national de la statistique (DANE) a intégré des covariables spatiales dans le calcul estimatif des populations situées dans les zones reculées ou inaccessibles. En outre, en Tanzanie, le recensement prévu pour 2022 s'appuiera sur une nouvelle approche dans l'élaboration des cartes, grâce au système d'information géographique (SIG). Le gouvernement entend également passer du support papier au support numérique sur tablette pour recueillir les données. Le Costa Rica et la Sierra Leone envisagent également l'utilisation des tablettes pour le recueil des données. Par ailleurs, au Belize, le gouvernement se lance dans l'élaboration d'une cartographie à partir d'images satellites pour évaluer les changements intervenus depuis le dernier recensement, tandis que le Népal utilise Google Maps pour mettre à jour sa carte des zones de dénombrement.

(Pays partenaires GPSDD, 2019)

Il existe désormais plusieurs alternatives aux fichiers démographiques maillés à l'échelle mondiale, chacun ayant ses particularités, des avantages spécifiques et des inconvénients. La communauté internationale d'experts en données sur la population reconnaît qu'il faut davantage de coopération et que le POPGRID Data Collaborative (POPGRID) rassemble les principaux producteurs de données ainsi que les utilisateurs et les promoteurs de données géoréférencées en matière de population, d'établissements humains et d'infrastructures. Ensemble, les membres de POPGRID cherchent à améliorer l'accès et la cohérence des données, à soutenir les utilisateurs en répondant aux besoins les plus urgents mais sans créer de risque de confusion et à encourager l'innovation.

Dans le présent rapport, nous décrivons sept des plus grands fichiers de population maillée dans le monde, produits par des organismes qui participent au POPGRID : Gridded Population of the World (GPW), Global Rural Urban Mapping Project (GRUMP), Global Human Settlement Layer (GHSL), LandScan, WorldPop, World Population Estimate (WPE) et le High Resolution Settlement Layer (HRSL). Ces fichiers de données mondiales utilisent une méthode descendante pour désagréger et répartir dans des cellules de grille le nombre d'habitants dénombrés par les recensements. Le résultat dépend, par conséquent, des données de recensement à partir desquelles est estimé le calcul de la population. Le présent rapport met en exergue les fichiers de données produits à partir des méthodes descendantes, présentés au chapitre 2. Cependant, des méthodes alternatives sont en cours de développement pour obtenir des estimations qui ne dépendent pas de l'existence de données de recensement, telles que les méthodes de recensement ascendantes ou hybrides qui recourent aux enquêtes micro-censitaires et à l'imagerie satellitaire détaillée pour estimer le nombre d'habitants lorsque les habituelles données de recensement national ne sont pas disponibles (Wardrop et al., 2018, Weber et al., 2018, FNUAP, 2017).

Dans le chapitre suivant, nous passons en revue les fichiers de données démographiques maillées qui offrent une couverture mondiale ou quasi mondiale et leurs applications possibles. Nous examinons ensuite les mesures d'ores et déjà entreprises pour valider ces données et nous expliquons pourquoi il est primordial d'assurer la poursuite des travaux de recherche. Sur la base de cet ensemble d'informations, nous énumérons quelques principes d'orientation sur l'utilisation des données démographiques maillées. Employées de façon responsable et éclairée, les données démographiques maillées peuvent représenter un atout majeur pour que personne ne soit rayé de la carte.

ENCADRÉ 2 | LE RECENSEMENT RESTE PRIMORDIAL

Les données démographiques maillées ne remplacent en rien les données du recensement. Au contraire, tous les fichiers démographiques maillés à l'échelle mondiale sont élaborés à partir des données de recensement. La fiabilité des données d'entrée est l'élément le plus important pour garantir l'exactitude de l'ensemble des estimations qui en résultent (Wardrop et al., 2018). Les données maillées doivent être envisagées comme un outil complémentaire pour mieux exploiter les données de recensement. Y compris lorsque les méthodes reposent sur l'imagerie satellitaire, le recensement reste une variable nécessaire à la plupart des estimations car l'analyse de l'image ne peut remplacer un décompte fiable effectué sur le terrain. L'analyse automatisée des images (également appelée classification non supervisée) peut être utilisée pour contourner certaines difficultés mais elle comprend de nombreuses limites. Ainsi, par exemple, un couvert végétal fourni peut obscurcir les habitations et laisser croire que la zone est inhabitée ou encore, des constructions peuvent ressembler à des bâtiments résidentiels et n'être que des bâtiments commerciaux ou industriels. S'il est vrai que les données démographiques maillées permettent de documenter les changements intervenus entre deux recensements, d'apprécier la répartition de la population selon un degré de résolution supérieur et de proposer une estimation du nombre d'habitants dans les zones où aucun recensement n'a été effectué, leur utilité générale se trouve renforcée par des données de recensement plus fines.

UTILISATION DES DONNÉES DÉMOGRAPHIQUES MAILLÉES DANS LE MONDE

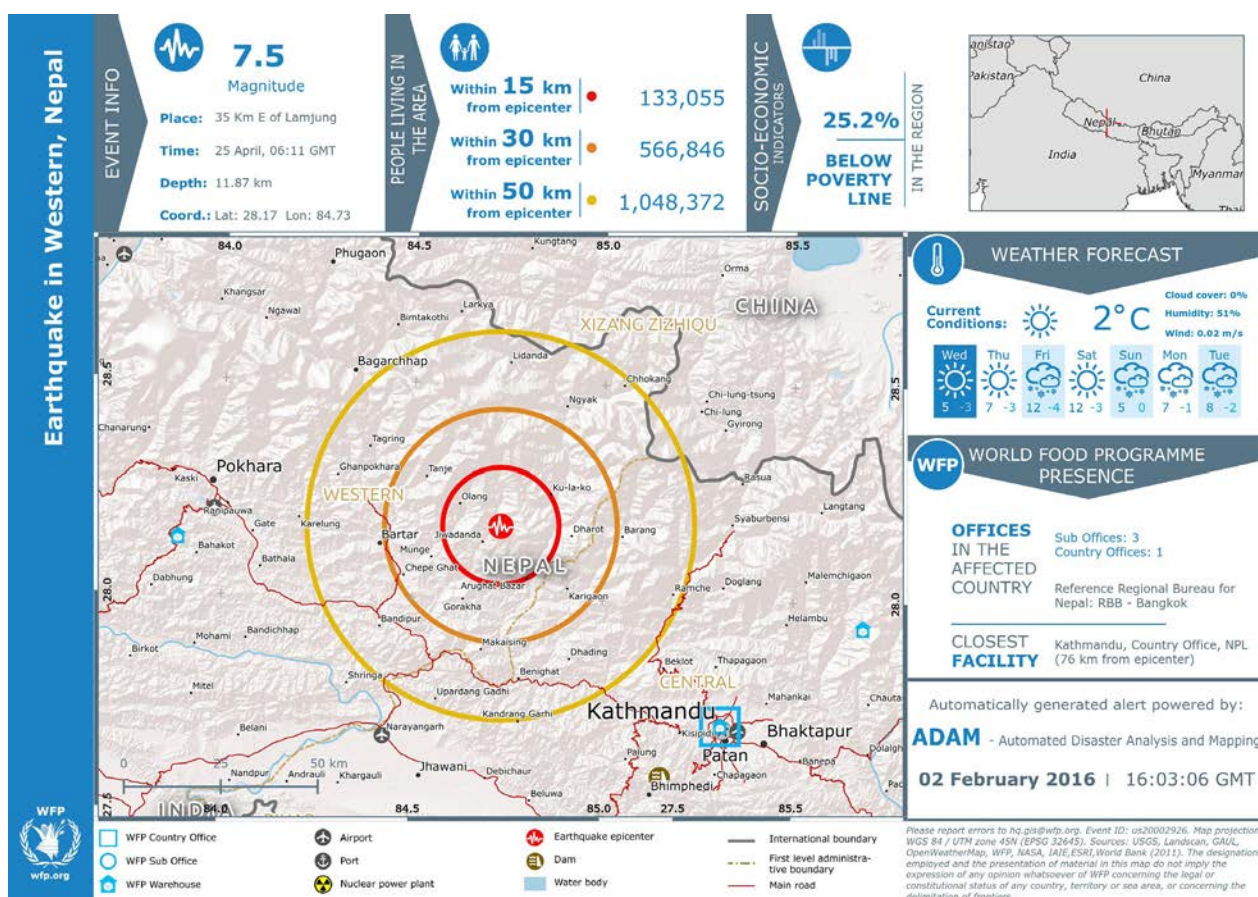
Au sein de la communauté internationale, se dégage un mouvement grandissant qui exige que les décisions soient prises sur la base de faits (Cairney, 2016). Les Nations unies réclament la production de données et de statistiques de développement durable plus fiables, grâce aux moyens de la révolution numérique, de sorte que les décisions des gouvernements reposent sur des faits, qu'elles renforcent la responsabilisation et permettent le suivi des progrès réalisés en direction des objectifs mondiaux (IEAG, 2014). Il se trouve que le recensement est l'un des outils majeurs dont nous disposons en matière de données factuelles. Comprendre où vivent les populations et savoir identifier les communautés les plus vulnérables est un facteur indispensable pour honorer les engagements qui sont les nôtres, tant à l'échelle nationale que mondiale. Dans certains cas, cependant, le recensement affiche ses limites et les utilisateurs de données, un peu partout dans le monde, se tournent vers les données démographiques maillées pour faire face à des situations d'urgence très variées, allant de l'intervention d'urgence dans le cadre d'une crise humanitaire immédiate au suivi des engagements politiques en matière de changement climatique. La pluralité des applications possibles souligne le potentiel que recèlent ces outils, non seulement pour caractériser la population de notre planète, mais également pour améliorer les conditions de vie de chaque individu.

1.1 APPLICATIONS IMMÉDIATES

Chaque année, les catastrophes naturelles frappent près de 160 millions de personnes dans le monde (OMS, n.d.), provoquent le déplacement de près de 17 millions de personnes (IDMC, 2019) et les maladies infectieuses tuent plus de 6 millions de personnes (Baylor College of Medicine, n.d.). Si nous voulons être en mesure de porter secours aux victimes, il est indispensable que nous sachions immédiatement et avec certitude où elles se trouvent. Les données démographiques maillées sont utilisées pour accélérer et cibler la mobilisation des ressources de l'intervention humanitaire et pour renseigner le choix des méthodes d'intervention.

ÉVALUATION DES RISQUES DE CATASTROPHE ET INTERVENTION

L'évaluation du risque de catastrophe et la planification de l'intervention sont parmi les applications des fichiers de données démographiques maillées les plus communément utilisées (MacManus, 2019). Elles servent ainsi à anticiper l'impact potentiel de scénarios de catastrophe les plus divers. Ainsi, suite à la fusion de l'installation Fukushima Daiichi au Japon, les chercheurs ont calculé, à partir des fichiers du Projet Global Rural-Urban Mapping (GRUMP), que plus d'un million de personnes dans le monde vit dans un rayon de 30 km autour des 21 centrales nucléaires considérées dans l'étude (Butler, 2011), ces étant utilisées pour anticiper la gestion des risques provoqués par d'autres incidents semblables. Basée sur les données du Global Human Settlement Layer (GHSL) de 2015, une autre étude est parvenue à un calcul indiquant que plus de 35 % de la population mondiale est potentiellement exposée à un risque de tremblement de terre et qu'un risque d'inondation menace probablement 1 milliard de personnes (Ehrlich et al., 2018). Lorsqu'une catastrophe se déclare, il



Exemple de tableau de bord de l'ADAM"

Source : [Programme Alimentaire Mondial](http://ProgrammeAlimentaireMondial)

est indispensable que les secours soient acheminés le plus rapidement possible aux populations touchées. C'est pourquoi des méthodes sont élaborées pour calculer « à la volée » le nombre de personnes exposées aux catastrophes naturelles ou aux catastrophes d'origine humaine (Garb et al., 2007). À titre d'exemple, une étude préliminaire avait révélé le potentiel de ce type d'approche en utilisant les données démographiques pour calculer le nombre de morts provoqués par le tsunami de 2004, en Indonésie. Il avait ainsi été démontré qu'il était possible d'obtenir une estimation raisonnable à partir de ces données, laquelle s'est avérée utile à la réponse humanitaire (Doocy et al., 2007).

Le Programme Alimentaire Mondial des Nations unies (PAM) s'appuie désormais activement sur les données démographiques maillées pour planifier ses interventions (Abushady, 2019). Ne disposant que de très peu de temps après qu'une catastrophe se déclare pour évaluer le volume des ressources à déployer, le PAM accélère ses calculs en utilisant l'[Automatic Disaster Analysis & Mapping \(ADAM\)](#), un outil d'analyse automatique et de cartographie des catastrophes qui s'inspire à la fois des fichiers du LandScan et de ceux du WorldPop pour calculer l'impact des tremblements de terre et des tempêtes tropicales, quasiment en temps réel. L'outil élabore alors des cartes et des tableaux de bord sur les populations touchées, qui sont partagés avec le reste de la communauté humanitaire. Ces estimations ont une incidence directe sur les décisions du PAM relatives au premier acheminement de l'aide alimentaire. Le calcul du nombre d'habitants est ensuite rapidement ajusté grâce aux données qui remontent des opérations sur le terrain mais les premières estimations obtenues grâce à ADAM restent primordiales pour déclencher le niveau d'intervention approprié. En l'absence de données démographiques maillées, il serait impossible de déclencher une réaction aussi rapide et efficace et il y aurait des retards dans l'acheminement des secours.

Le gouvernement des États-Unis s'appuie, lui aussi, sur les données démographiques maillées pour ses interventions dans les catastrophes internationales. Il a ainsi utilisé ces données dans plusieurs cas de catastrophes naturelles allant du tremblement de terre de 2010, en Haïti, à celui de 2015 au Népal et aux innombrables ouragans qui s'abattent régulièrement sur les Caraïbes (Rose, 2019). En outre, le Bureau des statistiques des États-Unis a développé un fichier de données démographiques spécialisé portant sur le Pakistan, qui tient compte de la spécificité et de la diversité du pays (Azar et al., 2013), lequel a été utilisé par la section humanitaire du Département d'État américain pour estimer le nombre de victimes lors des inondations de 2010. Le Bureau d'études géologiques des États-Unis (USGS) utilise également les données démographiques de LandScan dans son système [PAGER](#) (Prompt Assessment of Global Earthquakes for Response), un outil d'évaluation rapide de l'impact des tremblements de terre dans le monde.

SANTÉ ET MALADIES INFECTIEUSES

Lorsque l'épidémie d'Ébola s'est déclarée en Afrique occidentale, en 2013–2016, une des difficultés majeures rencontrée par les services d'urgence sanitaire était le manque d'informations fiables sur la population locale (Cori et al., 2017). Privés de données démographiques détaillées et récentes, les intervenants ont eu les pires difficultés pour localiser les communautés et calculer avec exactitude le taux d'infection. Les données démographiques maillées peuvent aider à contourner ces difficultés, à anticiper la propagation des maladies infectieuses et à définir les stratégies d'endiguement. Une étude démontre ainsi qu'il aurait été possible de faire une projection de la propagation du virus Ébola grâce aux données démographiques du LandScan, utilisées en complément du nombre de cas déclarés et du calcul des distances entre les districts touchés et les districts non touchés, en Guinée, en Sierra Leone et au Libéria (Rainisch et al., 2014). Aujourd'hui, alors que le nouveau virus COVID-19 atteint des proportions de pandémie mondiale, à l'heure même où nous rédigeons ce rapport, les données démographiques maillées jouent un rôle important dans les calculs prévisionnels de la propagation du virus en Chine et à travers le monde, le travail du groupe WorldPop ayant démontré l'impact majeur des stratégies de confinement sur le ralentissement de la propagation (Lai et al., 2020a; Lai et al., 2020b). Par ailleurs, les producteurs du GPW ont développé un outil interactif de visualisation mondiale du COVID-19 qui affiche les statistiques du nombre de cas déclarés avec les estimations démographiques par tranche d'âge et sexe, une information essentielle pour comprendre les risques de propagation et la gravité du virus (SEDAC. n.d.).

Les études de santé publique menées sur les maladies infectieuses se fondent elles aussi, de plus en plus, sur les données démographiques maillées, qui sont notamment essentielles dans le calcul de la « population à risque » (PAR), c'est-à-dire du pourcentage de la population exposée au risque de maladie (Linard and Tatem, 2012). Par exemple, l'Organisation mondiale de la santé (OMS) utilise le fichier WorldPop pour calculer le nombre de cas de paludisme dans les zones pour lesquelles il n'y a pas de statistiques fiables, en particulier sur le continent africain, où sont concentrés plus de 90 % des cas de paludisme (OMS, 2019). En superposant les différents modèles de prévalence des maladies du projet Atlas Malaria et les données démographiques maillées, il est possible de prévoir la charge de morbidité. Grâce à ces calculs, l'OMS estime à 228 millions le nombre de cas de paludisme dans le monde pour l'année 2018, soit une diminution du nombre de cas par rapport aux 251 millions comptabilisés en 2010, mais en-deçà des objectifs internationaux de lutte contre le paludisme. De même, les estimations des modèles à haute résolution sur le nombre d'adultes atteints du virus de l'immunodéficience humaine (VIH) en Afrique ont été obtenues en associant les données du WorldPop et les études et renseignements de terrain indiquant l'emplacement des centres de santé et concluent à des variations considérables du taux de transmission d'une région à l'autre (Dwyer-Lindgren et al., 2019).

Les données démographiques sont également essentielles pour vérifier l'existence de services de santé primaire. Une étude a cartographié la couverture vaccinale des enfants contre la diphtérie-tétanos-coqueluche (DTC) en Afrique, à partir du WorldPop et des enquêtes de santé. Elle conclut que si la couverture s'est améliorée dans près de trois quarts des unités administratives secondaires en Afrique, de 2000 à 2016, seuls le Maroc et le Rwanda ont atteint l'objectif de 80 % de couverture infantile (Mosser et al., 2019). Les données démographiques maillées ont, par ailleurs, été utilisées pour soutenir les mesures de lutte contre la malnutrition et le retard de croissance des enfants, pour développer des indicateurs de prise de décision sanitaire en milieu urbain et pour faire un calcul estimatif de la répartition géographique des grossesses et des naissances (Local Burden of Disease Child Growth Failure Collaborators, 2020; Thomson et al., 2019; James et al., 2018).

PLANIFICATION D'ENQUÊTE

Les enquêtes auprès des ménages sont également un outil important pour planifier les interventions sanitaires et autres dispositions d'urgence. Pour constituer un échantillonnage local représentatif, les enquêteurs doivent comprendre la façon dont la population est répartie. Lorsque les données de recensement sont caduques, imprécises ou réservées à la haute administration, planifier une enquête peut s'avérer compliqué. Or, les données démographiques maillées fournissent aux enquêteurs des estimations actualisées car, dans certains cas, les cellules de grille sont plus représentatives des zones unitaires réelles qui font l'objet de l'étude (Thomson, 2019). Par ailleurs, l'homogénéité du quadrillage facilite le travail d'échantillonnage spatial, de sorte que certains groupes d'habitants, plus petits et souvent sous représentés, peuvent être étudiés plus en détail (Thomson et al., 2017). L'outil GridSample a été conçu par Flowminder pour aider les enquêteurs à générer automatiquement des unités d'échantillons sur la base des données démographiques maillées (Thomson et al., 2017). Dans l'élaboration de la planification des enquêtes, les données démographiques maillées sont parfois encore jugées comme expérimentales, pourtant, elles ont déjà été utilisées un peu partout dans le monde pour concevoir des enquêtes dans le domaine de la planification sanitaire, notamment au Rwanda et en Haïti (Thomson, 2019).

1.2 APPLICATIONS À MOYEN TERME

La population mondiale devrait atteindre 8,5 milliards d'habitants d'ici 2030 (DPNU). Grâce aux données démographiques maillées, le suivi des conditions économiques et environnementales en parallèle de la croissance de la population peut être renforcé, tout comme les efforts déployés en vue des ODD et autres engagements à moyen terme.



EXAMEN DES DÉCISION POLITIQUES

Une meilleure compréhension de la façon dont les populations sont réparties peut renforcer la planification et l'évaluation des politiques. Du point de vue le plus général, les données démographiques maillées sont un outil qui permet d'évaluer les résultats des diverses stratégies économiques mises en œuvre. Bien que dans une majorité de pays, un processus de décentralisation soit en cours ces dernières décennies, les retombées économiques qui en découleraient font encore débat. Une récente étude économique s'est appuyée sur les mesures de fragmentation géographique à partir des données démographiques pour conclure que l'augmentation des dépenses régionales s'accompagne d'une légère hausse du PIB, cette relation n'étant pertinente que dans les pays développés (Canavire et al., 2020). Une étude de la Banque mondiale s'est penchée sur les implications de la Nouvelle route de la soie de la Chine sur les économies d'Asie centrale, en créant un modèle d'équilibre spatial qui inclut les données démographiques du GPW (Bird et al., 2019). Les résultats suggèrent que les investissements de cette initiative pourraient avoir des retombées positives dans certaines zones où les prévisions pointent vers une hausse des revenus et où la population pourrait plus que doubler, mais qui en désavantageraient d'autres. Ainsi, les données démographiques maillées permettent également de faire la lumière sur le coût humain que représentent tant le commerce mondial que la consommation mondiale. À titre d'exemple, une étude, réalisée sur la base des données du LandScan pour examiner l'impact des particules fines dans la pollution de l'air des pays asiatiques résultant de la consommation des cinq plus grands pays consommateurs, conclut que cette pollution a contribué à la mort prématurée de plus d'un million de personnes rien qu'en 2010 (Nansai et al., 2020).

Comprendre la façon dont la population est répartie se révèle également être fondamental pour la planification des infrastructures et la prestation de services, ne fût-ce que parce que les données démographiques maillées peuvent mettre la lumière sur les changements intervenus dans l'accès aux services de base. Les données du GPW ont été utilisées par l'un des modèles créés pour documenter le taux d'électrification en Afrique. Selon ce modèle, le pourcentage de la population du continent ayant accès à l'électricité est passé de 26,8 % à 35,5 % entre 2000 et 2013 (Andrade-Pacheco et al., 2019). De même, une étude réalisée à partir des données du WorldPop, en association avec d'autres études géoréférencées, conclut que le pourcentage d'Africains ayant accès à de meilleures conditions de logement a plus que doublé de 2000 à 2015, passant de 11 % à 23 % (Tusting et al., 2019). Au-delà d'aider à mieux appréhender les conditions actuelles, les données démographiques maillées peuvent contribuer à l'élaboration de nouvelles infrastructures. Ainsi, les données du GPWv4 ont été utilisées dans le cadre d'une récente évaluation des micro-réseaux solaires au Malawi, où seuls 12 % de la population, sur un total de 18 millions d'habitants, ont accès au réseau électrique (Eales et al., 2020). En intégrant la taille et l'emplacement des communautés dans les paramètres, l'étude a conclu que les micro-réseaux solaires pourraient être le moyen le plus rentable pour apporter l'électricité à 37 % de la population du Malawi, que l'expansion du réseau électrique existant pourrait être la méthode la plus rentable pour arriver à 42 % de la population, les 21 % restants bénéficiant de systèmes solaires individuels. Dans le même ordre d'idées, une analyse sur l'hypothèse d'une élimination progressive du secteur énergétique du charbon en Chine a tenu compte des tendances en matière de population et de pollution dans ses recommandations sur la faisabilité de stratégies techniques qui seraient également équitables (Cui et al., 2020). D'après ses producteurs, le LandScan est utilisé par le secteur privé dans la conception des réseaux de téléphonie mobile et l'étude de nouveaux axes de transport (Rose, 2019).

LES OBJECTIFS DE DÉVELOPPEMENT DURABLE

En prenant l'engagement des ODD, les 193 pays signataires ont également accepté de surveiller 232 indicateurs portant sur divers aspects de la durabilité. Ces indicateurs sont essentiels pour atteindre les objectifs mondiaux, en ceci qu'ils offrent un mécanisme de mesure des progrès accomplis et permettent de développer et de tester les politiques mises en place. Cependant, de nombreux pays sont encore privés des méthodes ou des données nécessaires pour assurer le suivi régulier de ces indicateurs. Une étude des Nations unies datant de 2018 conclut qu'en Afrique et en Asie, seuls 20 % des indicateurs ODD, en moyenne, sont actuellement disponibles (Nations unies, 2018).

Les données démographiques maillées peuvent fournir des informations essentielles sur la population et s'appliquent à ces nombreux indicateurs ODD qui décrivent l'état de la population ou le degré de variation des caractéristiques démographiques, comme par exemple l'indicateur ODD 1.1.1. « proportion de la population vivant au-dessous du seuil de pauvreté fixé au niveau international, par sexe, âge, situation dans l'emploi et lieu de résidence » (Tatem, 2019). Elles peuvent également être utilisées pour mesurer la distance qui sépare les populations des services et infrastructures et identifier ainsi les communautés reculées ou difficiles d'accès, qui sont parfois oubliées par les méthodes conventionnelles de recueil de données (Tatem, 2019). De plus, les données démographiques modélisées offrent l'avantage d'assurer un suivi plus régulier des changements intervenus puisque le décompte du nombre d'habitants par le biais des recensements n'est, en général, disponible qu'une fois tous les dix ans (MacManus, 2019).

Plusieurs méthodes sont en cours de développement qui utilisent les données démographiques maillées rapportées à plusieurs des indicateurs ODD. À titre d'exemple, les connexions aux réseaux de transports sont essentielles pour porter assistance aux pauvres des zones rurales et l'indice d'accès rural (ODD 9.1.1) mesure le pourcentage d'habitants en milieu rural vivant non loin d'une route. Cette mesure a toujours été obtenue avec les données sur les ménages et n'est disponible que pour l'échelon national, ce qui est moins pertinent dans le cadre de la planification du développement (Iimi et al., 2016). La Banque mondiale a récemment mis au point une méthode qui combine les données démographiques maillées du WorldPop et les sources de données nationales pour calculer l'indice d'accès rural (Iimi et al., 2016) à l'échelon régional. Cette nouvelle méthode permet, non seulement d'obtenir des mesures à intervalle régulier, mais elle offre également une lecture plus détaillée. Ainsi, la Banque mondiale a pu décrire des différences notoires de conditions routières et de densité entre l'Afrique et l'Asie (Banque mondiale, 2016). Depuis 2018, la Banque mondiale a calculé les indices de 20 pays et cette méthode a obtenu l'aval officiel de la Division des statistiques des Nations unies (Vincent, 2018 ; IAEG-ODD ONU, 2019).

Tout indicateur relatif aux difficultés du développement urbain peut également tirer parti des avantages que représentent les données démographiques maillées. Une méthode a été développée pour mesurer le niveau moyen pondéré annuel de particules fines en fonction du nombre d'habitants (indicateur ODD 11.6.2) en associant le GPW aux mesures effectuées sur le terrain et aux censeurs satellites à distance (Shaddick et al., 2017). L'OMS a officiellement adopté cette méthodologie et cet indicateur est désormais disponible dans 178 pays (DPNU, 2017). De même, le taux d'occupation des sols (indicateur ODD 11.3.1) calcule le ratio entre le taux de consommation d'espace et le taux d'accroissement de la population. Les chercheurs du Centre commun de recherche de la Commission européenne avancent une méthode qui permet de mesurer cet indicateur en utilisant les données démographiques maillées du GHSL (Melchiorri et al., 2019). Ils ont ainsi mesuré les changements du taux d'occupation des sols de 10 000 centres urbains entre 1990 et 2015, en tenant compte de l'augmentation mondiale de la densité de la population urbaine au cours de cette même période. Il a également été suggéré que le GHSL pourrait mesurer le pourcentage de la

population qui a aisément accès aux transports publics (indicateur ODD 11.2.1), ainsi que 11 autres lacunes relatives aux indicateurs urbains (ONU – GGIM Europe, 2019 ; Melchiorri et al., 2019).

Bien que divers autres indicateurs ODD se basent sur le calcul du nombre d'habitants ou des zones peuplées, la plupart d'entre eux ne sont pas confirmés à l'échelon local (Gaughan, 2019). Selon une étude réalisée par TREnds, au moins 73 des indicateurs spécifiques exigent des données démographiques sous une forme ou sous une autre et 34 disposent toujours de données inadéquates ou de méthodes indéfinies (voir [Tableau 1](#)). Parmi les indicateurs les plus pertinents, plusieurs exigent d'avoir accès à des renseignements sur la répartition de la population en fonction des zones d'impact ou des zones desservies. Il s'agit notamment du pourcentage de la population touchée par les pertes économiques directement imputables aux catastrophes (indicateurs ODD 1.5.1 et 1.5.2), l'accès aux services d'assainissement (indicateur ODD 6.2.1) et à l'électricité (indicateur ODD 7.1.1). À l'instar du taux d'occupation des sols, les données démographiques maillées peuvent s'avérer utiles pour le traçage de plusieurs indicateurs urbains, dont le pourcentage de population urbaine vivant dans des quartiers de taudis ou des logements inadéquats (indicateur ODD 11.1.1). Ainsi, en élargissant l'application des données démographiques maillées, il est possible d'accroître le degré de responsabilité et la sensibilisation à l'égard des ODD.

La Commission de statistique des Nations unies (CSNU) a entériné l'utilisation d'une méthode de calcul permettant de distinguer les zones urbaines des zones rurales aux fins des comparaisons de statistiques internationales, à partir d'une grille démographique (CSNU, 2020). Cette méthode facilite le recueil d'informations et l'emploi qui en est fait dans le cadre des indicateurs ODD qui imposent de distinguer les données urbaines des données rurales (1.1.1, 2.4.1, 3.3.1, 4.5.1, 9.1.1, 11.1.1), et permet de comparer les résultats des indicateurs ODD urbains sensibles à la délimitation des frontières urbaines (11.2.1, 11.3.1, 11.6.2, 11.7.1) (Centre commun de recherche de la Commission européenne, 2018).

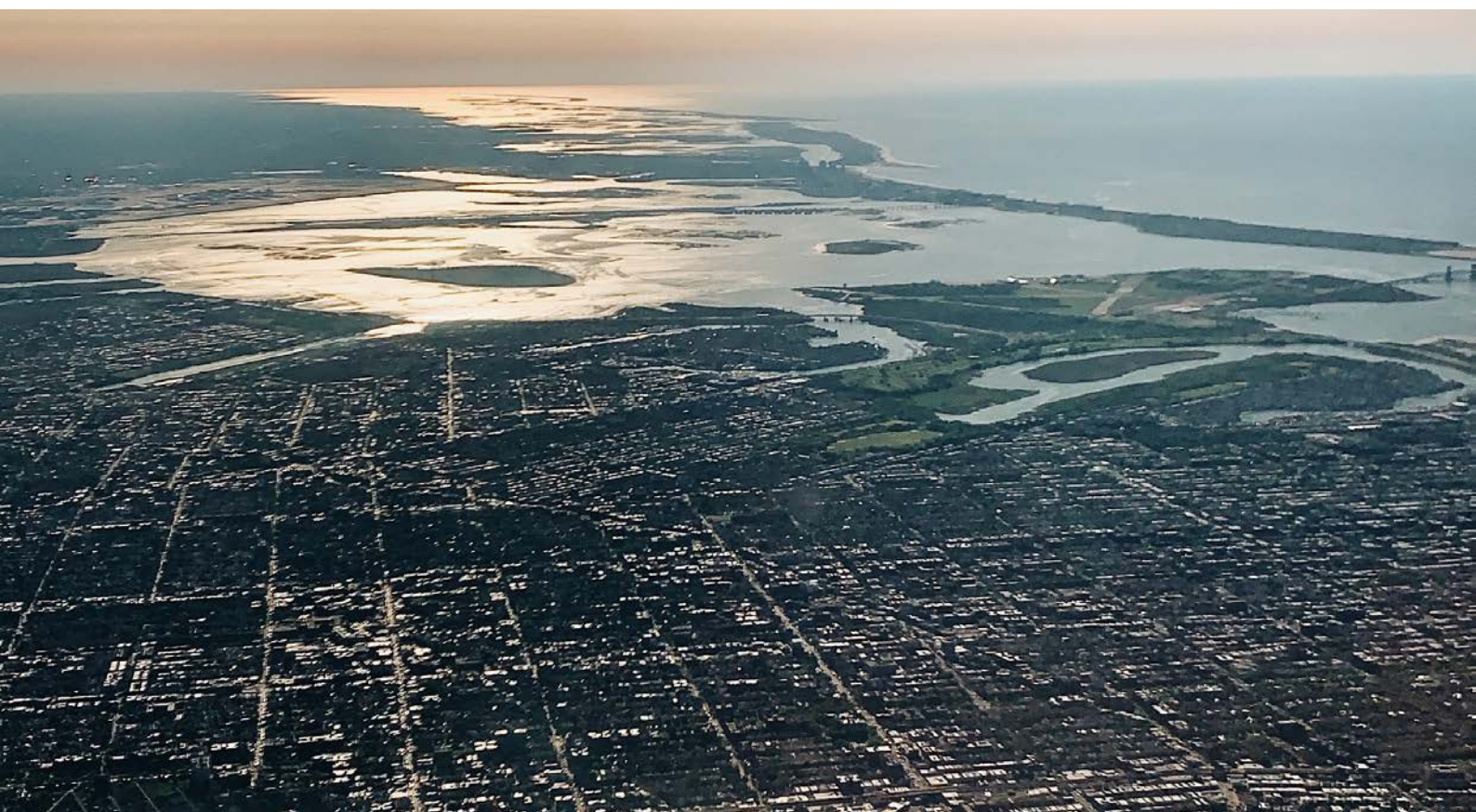


TABLEAU 1 - ODD EXIGEANT DES INFORMATIONS DÉMOGRAPHIQUES

ODD	INDICATEUR	ODD	INDICATEUR
1 PAS DE PAUVRETÉ 	1.1.1, 1.2.1, 1.2.2, 1.4.1, 1.4.2, 1.5.1, 1.5.3	8 TRAVAIL DÉCENT ET CROISSANCE ÉCONOMIQUE 	8.1.1, 8.4.2, 8.5.2, 8.6.1, 8.7.1, 8.10.1, 8.10.2
2 FAIM «ZÉRO» 	2.1.1, 2.1.2, 2.2.1, 2.2.2	9 INDUSTRIE, INNOVATION ET INFRASTRUCTURE 	9.1.1, 9.2.1, 9.5.2, 9.C.1
3 BONNE SANTÉ ET BIEN-ÊTRE 	3.1.1, 3.2.1, 3.2.2, 3.3.1, 3.3.2, 3.3.3, 3.3.4, 3.4.1, 3.4.2, 3.5.2, 3.6.1, 3.7.1, 3.7.2, 3.8.2, 3.9.1, 3.9.2, 3.9.3, 3.A.1, 3.B.1	10 INÉGALITÉS RÉDUITES 	10.2.1, 10.3.1
4 ÉDUCATION DE QUALITÉ 	4.2.1, 4.2.2, 4.3.1, 4.6.1	11 VILLES ET COMMUNAUTÉS DURABLES 	11.1.1, 11.2.1, 11.3.1, 11.5.1, 11.6.2, 11.7.2, 11.A.1
5 ÉGALITÉ ENTRE LES SEXES 	5.2.1, 5.2.2, 5.3.1, 5.3.2, 5.5.2, 5.6.1, 5.B.1	12 CONSOMMATION ET PRODUCTION RESPONSABLES 	12.2.2, 12.4.2
6 EAU PROPRE ET ASSAINISSEMENT 	6.1.1, 6.2.1	13 MESURES RELATIVES À LA LUTTE CONTRE LES CHANGEMENTS CLIMATIQUES 	13.1.1
7 ÉNERGIE PROPRE ET D'UN COÛT ABORDABLE 	7.1.1, 7.1.2	16 PAIX, JUSTICE ET INSTITUTIONS EFFICACES 	16.1.1, 16.1.3, 16.1.4, 16.2.1, 16.2.2, 16.2.3, 16.6.1, 16.6.2, 16.7.2, 16.9.1

1.3 APPLICATIONS À LONG TERME

Les données démographiques maillées font partie intégrante d'une évaluation à long terme des risques de changements environnementaux à venir, parmi lesquels la hausse du niveau de la mer provoquée par le changement climatique, les tempêtes et les sécheresses. Le résultat de ces travaux d'évaluation est primordial pour agencer une préparation efficace face à la catastrophe et planifier la résilience. Les frontières administratives utilisées pour enregistrer les données de recensement étant géographiquement irrégulières, il peut être difficile d'associer ces données à l'état brut avec d'autres types d'observations (Bai et al., 2018). Les données démographiques répondent à cette difficulté en fournissant aux chercheurs des informations de base qui les aident à mieux appréhender les risques futurs en matière de développement durable (MacManus, 2019). De fait, parce que les données maillées représentent la répartition actuelle de la population, elles sont utilisées dans de très nombreux scénarios de développement pour élaborer des projections spatiales de la population future selon divers niveaux de résolution, afin de mieux anticiper la vulnérabilité des populations aux aléas à venir (Gao, 2017 ; Jones et O'Neill, 2020).

HAUSSE DU NIVEAU DE LA MER

Les études scientifiques qui calculent le degré de vulnérabilité des populations côtières aux changements climatiques et aux inondations recourent régulièrement aux jeux de données démographiques à haute résolution pour saisir l'incroyable variabilité géographique du littoral (Mondal et Tatem, 2012). Les projections disponibles suggèrent ainsi que la hausse du niveau de la mer pourrait avoir un effet désastreux sur l'humanité. Une analyse réalisée à partir du projet GRUMP suggère que 411 millions de personnes dans le monde pourraient être exposées à des risques extrêmes d'inondation d'ici 2060, par rapport aux 189 millions estimés en 2000 (Neumann et al., 2015). La superposition des données démographiques maillées avec le calcul du produit intérieur brut (PIB) et le taux d'occupation des terres agricoles indique qu'une augmentation du niveau de la mer de 1 mètre met les pays en développement dans une situation de risque accru (Dasgupta et al., 2011). Par ailleurs, la modélisation de la hausse du niveau de la mer et d'autres phénomènes environnementaux suggère que le nombre de personnes en situation de risque pourrait être supérieur. Un rapport de la Banque mondiale réalisé à partir de LandScan, en association avec d'autres outils, pour évaluer le risque accru d'inondations résultant de la disparition des palétuviers, conclut que le risque d'inondations pourrait augmenter d'un tiers dans 42 pays en développement, exposant ainsi deux fois plus d'habitants (Blankespoor et al., 2016). Des travaux plus récents, qui ont également utilisé le LandScan en association avec CoastalDEM, la coupe d'élévation côtière améliorée, estiment que les populations en situation de risque dans les zones côtières pourraient être beaucoup plus nombreuses (Kulp and Strauss, 2019). Des études comparatives sont donc en cours pour mieux comprendre les zones d'incertitudes (CIESIN et CUNY CIDR, 2020 prochainement), sans compter que l'impact économique de la hausse du niveau de la mer pourrait être désastreux. Une modélisation récente, ayant utilisé les données du GRUMP, suggère que si des mesures d'adaptation plus poussées ne sont pas mises en œuvre, l'hypothèse d'une importante fonte des glaces pourrait entraîner une hausse du niveau de la mer qui provoquerait une baisse du PIB pouvant aller jusqu'à 4 % (Schinko et al., 2020).

DISPONIBILITÉ DES RESSOURCES EN EAU

Pour aussi significatif que puisse être le changement climatique à l'avenir, il semble que l'accès aux ressources hydriques soit beaucoup plus sensible aux changements démographiques. Selon une étude réalisée avec des données démographiques maillées, si les pires projections climatiques se concrétisent et s'accompagnent simultanément d'une croissance démographique inédite à ce jour, des milliards de personnes pourraient se trouver en situation de stress hydrique grave d'ici 2100. Cette étude a associé les données de LandScan et la projection sur les variations du niveau de précipitations pour analyser la complexité des interactions entre la population et le climat (Parish et al., 2012). Des scénarios socio-économiques ont été dressés pour estimer la croissance de la population jusqu'à la fin du siècle. Les chercheurs ont ensuite fait une modélisation de la disponibilité en eau par tête d'habitant pour chacun des 26 929 bassins versants que compte la planète. Les ressources hydriques étant une question transnationale, les données maillées ont permis d'élaborer une modélisation à partir de la géographie naturelle plutôt que des frontières artificielles. S'il est vrai que des variations radicales pourraient affecter les précipitations au cours du XXI^e siècle, certaines régions s'asséchant et d'autres devenant plus humides, les simulations effectuées à partir de différentes combinaisons de projections démographiques et climatiques suggèrent que les ressources hydriques par tête d'habitant seraient surtout sensibles à la croissance de la population. Une autre étude visant à élaborer un modèle économique plus avancé, également réalisée à partir des données de LandScan, estime que 8,6 milliards de personnes, soit deux tiers de la population projetée, sont susceptibles de connaître une situation de stress hydrique (Kiguchi et al., 2014). Bien que le degré de certitude de ces approches ne soit pas absolu, l'examen simultané des projections sur la répartition démographique et sur le climat nous permet d'identifier, en amont, les zones les plus vulnérables et de prendre les mesures qui s'imposent.

1.4 ÉLARGIR LE CHAMP D'APPLICATION ET COMPRENDRE LES OPTIONS DISPONIBLES

Comme nous l'avons démontré, les données démographiques maillées peuvent constituer un outil puissant pour comprendre et améliorer les conditions de vie des populations. Non seulement elles peuvent être une aide inestimable dans le cadre d'une réponse à la catastrophe, mais elles peuvent aider à comprendre les risques de maladie et prévoir l'impact futur de la dégradation de l'environnement. De plus, elles permettent d'assurer le suivi des indicateurs identifiés par la communauté mondiale comme étant nécessaires pour parvenir au développement durable. Cependant, le potentiel de ces données démographiques maillées est rendu complexe par la diversité de l'offre et par les subtilités d'applications des jeux de données. Il convient que les utilisateurs saisissent bien les caractéristiques propres à chaque produit afin de porter leur choix sur l'option la plus adaptée à leurs besoins. Ces questions sont abordées dans le chapitre suivant.

SE FAMILIARISER AVEC LES JEUX DE DONNÉES DÉMOGRAPHIQUES MAILLÉES

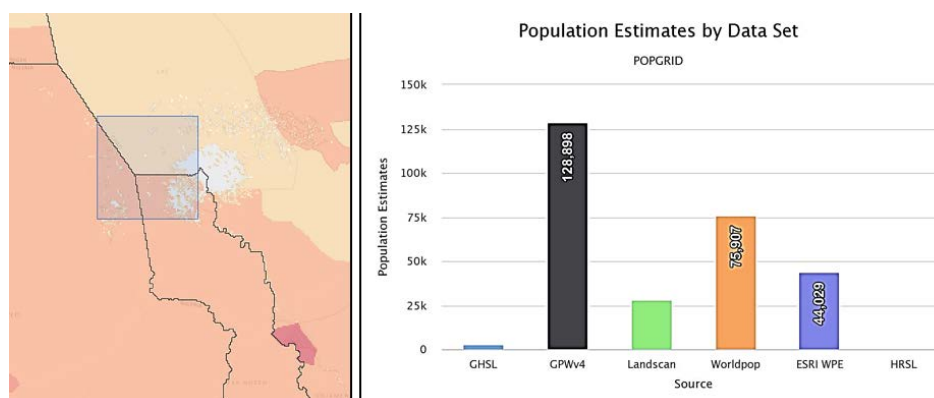
Comme nous l'avons expliqué au chapitre 1, le PAM utilise les données démographiques maillées pour mobiliser ses ressources dans les situations d'urgence. Actuellement, la plate-forme ADAM utilise cependant deux jeux de données démographiques maillés différents. Après avoir testé et comparé ces fichiers, le choix du PAM s'est porté sur le WorldPop, qui est plus adapté à la plupart des applications du PAM, tandis que le LandScan recueille la préférence pour les contextes insulaires. Si toutes ces solutions permettent de quantifier la population, ce choix souligne combien ces produits sont fondamentalement différents dans les méthodologies qu'ils appliquent. Il convient que les utilisateurs prennent le temps de comprendre les caractéristiques de chaque produit pour répondre correctement à leurs besoins en données.

Si certains fichiers de données démographiques maillées sont désormais régulièrement utilisés par diverses organisations, les résultats obtenus sont très distincts. Notre propos n'est en aucun cas de vanter un produit plus qu'un autre. La pertinence de chaque produit dépend du contexte et du type de questions auquel il faut apporter une réponse. Cependant, jeter son dévolu sur le premier produit pourrait aboutir à des informations trompeuses (voir l'exemple ci-dessous, dans le schéma 1, concernant une zone située à la frontière du Tchad, du Nigéria et du Cameroun). Le choix d'un produit exige de procéder à un examen rigoureux et de commencer par comprendre les méthodes de calcul employées par les différents fichiers de données et ce qui les distinguent les uns des autres. Dans ce chapitre, nous expliquons quelles sont les différences majeures entre les divers modèles de données démographiques. Nous passons ensuite en revue les sept principaux produits, en décrivant les données sous-jacentes, les méthodes (descendante-ascendante) et les hypothèses de base ainsi que les forces et faiblesses respectives de chaque produit.

2.1 DIFFÉRENCES DANS LES PARAMÈTRES DE MODÉLISATION

Les divers jeux de données maillées présentés dans cette section répondent généralement à une méthode de production similaire reposant sur la géorépartition des données démographiques. Les estimations sont cependant élaborées à partir de différents types de données et recourent à des méthodes de calcul différentes. Les principaux paramètres qui décrivent le mode d'élaboration des fichiers démographiques maillés sont au nombre de trois : 1) données saisies, 2) ajustement des données et 3) méthode de modélisation.

Schéma 1 : estimations démographiques divergentes concernant une zone située à la frontière du Tchad, du Nigéria et du Cameroun, à l'ouest du lac Tchad.



Source : <https://sedac.ciesin.columbia.edu/mapping/popgrid/comparison-view/>

DONNÉES SAISIES

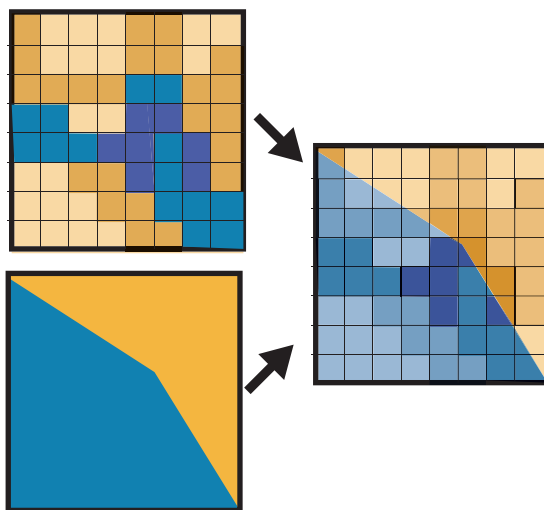
Les données de recensement, les données sur les limites administratives et enfin leur corrélation géospatiale sont les trois principales données de départ utilisées dans l'élaboration des jeux de données démographiques maillés.

DONNÉES DE RECENSEMENT – Il s'agit du décompte officiel du nombre d'habitants dans une unité administrative et/ou unité de recensement donnée (zone géographique représentant une entité régionale, c'est-à-dire un état, une province, un district, une agglomération, une ville, un secteur de recensement) ainsi que l'information démographique relative aux habitants qui y vivent. En théorie, les recensements doivent être menés selon une périodicité minimum de 10 ans. Certains pays n'y parviennent pas, souvent en raison du coût élevé que représente cet exercice, de la difficulté d'accès à certaines populations ou encore de situations de conflits et d'insécurité. Les pays n'ont alors à leur disposition que des données caduques ou lacunaires (DPNU, 2017). Par ailleurs, il est fréquent que les unités administratives changent d'un recensement à l'autre, au gré des fusions ou scissions, que les frontières ou les noms changent avec la croissance des villes ou suite aux mutations politiques ou économiques, entre autres facteurs. Bien que les données de recensement soient habituellement recueillies au niveau des ménages, les bureaux de statistiques nationaux ne diffusent généralement que les données agrégées correspondant à un niveau administratif spécifique, pour des raisons de confidentialité et de respect de la vie privée. Il arrive que certaines caractéristiques relatives aux individus ou aux foyers, spécifiques et détaillées, soient divulguées dans le cadre d'un échantillonnage anonyme de foyers (microdonnées de recensement).³

DONNÉES SUR LES LIMITES ADMINISTRATIVES – Il s'agit de renseignements géographiques relatifs à une parcelle de la surface terrestre correspondant à une unité administrative du recensement. Ces données sont essentielles pour géolocaliser le comptage du recensement et permettre la répartition de ces données et d'autres variables sur une grille géographique classique composée de cellules ou « pixels », une unité de forme plus ou moins carrée, généralement définie par des coordonnées de latitude et de longitude. La taille d'une cellule peut varier, allant de dizaines ou de centaines de mètres carrés à des dizaines ou centaines de kilomètres carrés, et varie selon la latitude (une cellule de 1° située au niveau de l'équateur compte 12 364 km² tandis qu'à 45° de latitude, elle compte 8 743 km²). Lorsque plusieurs unités administratives se superposent dans une même cellule, une formule de calcul doit être introduite pour chaque cellule afin d'y attribuer

une valeur unique (ex. population totale), de sorte que la moyenne des valeurs soit pondérée par les zones de chevauchement. De même, la population totale ou toute autre caractéristique attribuée à une unité administrative doit être affectée à toutes les cellules qui se superposent à elle, de façon proportionnelle à la zone de chevauchement ou selon une autre formule à définir. Des complications peuvent surgir selon les ambiguïtés, incertitudes, modifications et erreurs de données relatives aux limites administratives, telles que traits de côte fluctuants, territoires contestés, frontières définies par des cours d'eau ou toute autre caractéristique dynamique, et selon l'exactitude spatiale et les questions d'échelle.

Schéma 2 : Représentation visuelle de la répartition des chiffres du recensement sur une grille de cellules géographiques ou « pixels »



CORRÉLATION GÉOSPATIALE (intrants auxiliaires) – Il s'agit de données géographiques connues pour être (vraisemblablement) corrélées à la répartition de la population et qui peuvent ainsi être utilisées pour améliorer l'exactitude prédictive (Lloyd et al., 2017). Plutôt que d'employer de simples formules statistiques qui partent de l'hypothèse que les populations sont réparties de façon homogène dans toutes les unités administratives, les chercheurs ont développé un ensemble de modèles qui utilisent ces corrélations géospatiales pour identifier les zones où la densité de population est supérieure ou inférieure à la moyenne. Les jeux de données examinés dans cette section utilisent différentes approches de modélisation pour réaffecter la population correctement, tout en conservant les chiffres de la population totale pour chaque unité administrative. Les variables suivantes ont été utilisées ou testées dans le développement des fichiers de données démographiques maillées :

- ▶ structures bâties et surfaces imperméables,
- ▶ villes et établissements connus (ex. zones peuplées),
- ▶ données topographiques relatives à l'élévation, la pente et le trait de côte,
- ▶ données relatives aux infrastructures de différents services (ex. énergie, communication, santé, assainissement, éducation),
- ▶ réseaux routier, ferroviaire, de transit et autres réseaux de transport,
- ▶ données sur la couverture terrestre et l'utilisation des terres dans le domaine agricole, industriel, forestier, etc.,
- ▶ pollution lumineuse captée par satellite,
- ▶ zones protégées (ex. parcs nationaux, réserves sauvages),
- ▶ plans d'eau, dont rivières, lacs, mers et océans.

AJUSTEMENT DES DONNÉES

Les données de saisie pour l'élaboration des jeux de données démographiques maillées doivent être ajustées pour correspondre à un laps de temps spécifique (ex. un jour précis ou une année particulière) et au chiffre total de la population nationale correspondante.

Les recensements n'étant généralement pratiqués que tous les 10 ans et pas toujours la même année selon les pays, les estimations démographiques doivent être ajustées sur la même période de temps pour permettre un maillage croisé à l'échelle nationale. Ces ajustements reposent généralement sur les calculs et projections de la Division de la population des Nations unies (DPNU) pour assurer, à tout le moins, une cohérence générale avec les normes internationales. Les premières estimations de la DPNU commencent avec le décompte de la population au 1er juillet 1950 et continuent avec des estimations démographiques, par tranche d'âge et par sexe, sur des intervalles de cinq ans. Ces estimations s'appuient sur une méthode de cohorte, soit la projection des taux de fécondité, de mortalité et de migration de chaque groupe de la population né au cours d'une année déterminée (DPNU, 1917). Ces estimations sont évaluées selon diverses sources de données qui comprennent les recensements, les enquêtes post-censitaires, les enquêtes de ménages, les statistiques de l'état civil et les registres de la population (DPNU, 1917). La DPNU procède généralement à une révision de ses Perspectives de la population mondiale tous les deux ans, la dernière édition datant de 2019 (DPNU, 1919). Les ajustements à l'échelon régional sont généralement tributaires des changements intervenus à l'échelon national, sur l'ensemble de la population.

MÉTHODE DE MODÉLISATION

Les méthodologies utilisées pour répartir le décompte des habitants par cellule varie selon les jeux de données. Les méthodologies utilisent des corrélations empiriques à partir de plusieurs variables et sélectionnent divers critères pour attribuer à chaque cellule un nombre d'habitants donné.

PONDÉRATION ARÉOLAIRE – Il s'agit de la plus ancienne méthode de grille de population. Cette méthode consiste à répartir, de façon homogène, le nombre d'habitants d'une unité de recensement sur l'ensemble des cellules de grille de cette unité. Lorsqu'il y a chevauchement de cellules entre unités, on leur attribue un pourcentage de population inférieur, proportionnel à la surface terrestre de la cellule au sein de cette unité. La somme des habitants de toutes les cellules qui composent l'unité sera égale à la population totale de l'unité en question (Hallisey et al., 2017). Ainsi, la précision du décompte d'habitants au sein d'une cellule dépend essentiellement de la taille de l'unité de recensement. De façon générale, plus les unités sont petites, plus le décompte de la population est précis dans chaque cellule de grille.

À l'inverse, la plupart des calculs réalisés par les partenaires du POPGRID adopte une approche descendante qui répartit les données démographiques dans des cellules plus petites que celles habituellement utilisées dans les recensements (Lloyd et al., 2019).⁴ S'il est vrai que la plupart des modèles obéissent à une approche descendante, ils se distinguent cependant par le niveau et la méthode de fractionnement utilisés. Dans ce chapitre, les méthodes de modélisation utilisées pour constituer les jeux de données comprennent les éléments suivants :

CARTOGRAPHIE DASYMÉTRIQUE – Répartition homogène de la population en unités spatiales. La mention « pas de données » est attribuée aux cellules où il est peu probable qu'il y ait des habitants ; un décompte de la population est attribué aux cellules où il est probable qu'il y ait des habitants. Cette méthode peut être appliquée de différentes façons :

- ▶ **Cartographie dasymétrique binaire** – Porte la méthode de cartographie dasymétrique à l'échelon supérieur puisqu'elle introduit une variable supplémentaire, un jeu de données auxiliaires, qui permet la désagrégation spatiale des données de recensement des grandes unités administratives vers de plus petites cellules pour développer des fichiers de données à haute résolution. Contrairement à la pondération aréolaire, cette méthode ne repose pas sur une hypothèse de densité de population homogène pour toutes les cellules relevant d'une même unité administrative. Cette méthode exclut les zones inhabitées dans la répartition des habitants par zone habitée, et les zones inhabitées sont définies individuellement, par cellule, sur une base binaire (Eicher et Brewer, 2001).
- ▶ **Cartographie dasymétrique à variables multiple** – Utilise le même concept que la cartographie dasymétrique binaire mais s'appuie sur de multiples corrélations géospatiales (jeux de données auxiliaires) pour le calcul des estimations à haute résolution.
- ▶ **Forêt d'arbres décisionnels** – Les variables utilisées dans ce type de modélisation représentent le champ des variables possibles à partir desquelles opérer un choix (ex. les innombrables arbres d'une forêt, qui donne son nom à la méthode). Cette diversité permet à un algorithme d'identifier la variable la plus pertinente, parmi les variables « aléatoires », au regard de la densité de population. La méthode dite de forêt d'arbres décisionnels utilise des variables et des techniques de validation croisée pour évaluer l'effet pondéré à l'échelle d'un pays. À partir de la validation croisée et des données d'apprentissage, l'algorithme « apprend » quelle est la covariable ou la variable à laquelle attribuer la plus haute valeur pondérée pour un lieu donné. Les covariables sont, par conséquent, choisies de façon « aléatoire » pour chaque emplacement. L'algorithme calcule la densité de population en divisant la somme du décompte du recensement par zone d'unité agrégée. Dans cette méthode, les zones inhabitées (ex. aires protégées, plans d'eau) ne sont pas prises en compte.⁵ Cette méthode offre un degré de corrélation plus élevé entre la densité de population et les covariables géographiques de distance (ex. proximité de routes) (Stevens et al., 2015).

2.2 CARACTÉRISTIQUE DE SORTIE

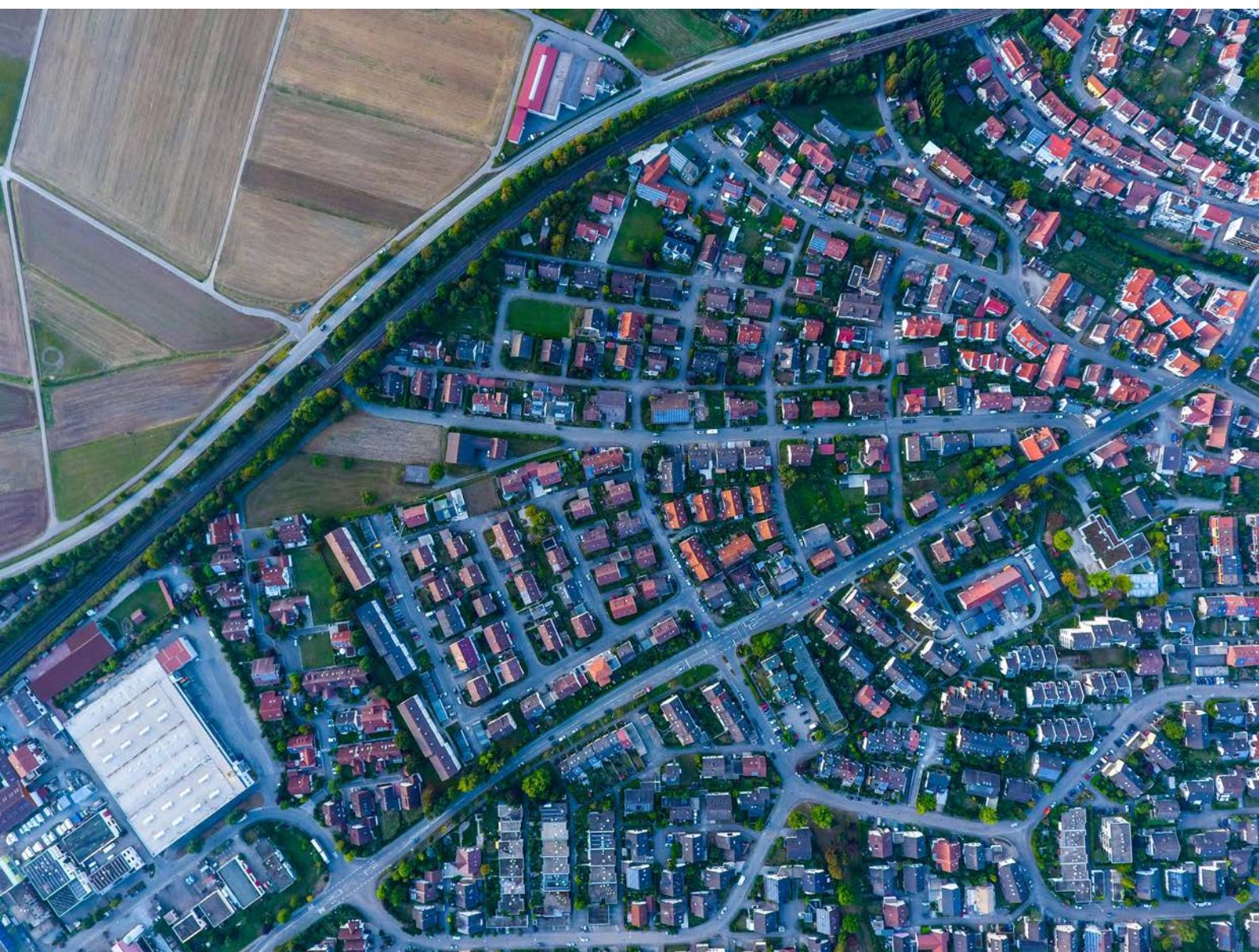
Les fichiers de données démographiques maillées utilisent diverses résolutions spatiales. Ce que l'on appelle la résolution spatiale est la taille de la cellule de grille. Toutefois, comme nous l'avons évoqué par ailleurs, cette résolution spatiale est généralement beaucoup plus fine que la résolution d'entrée, laquelle comprend des unités de recensement de tailles et de formes diverses. Cette résolution est exprimée en mètres ou en secondes d'arc,⁶ qui sont des unités de mesure communes au système de référencement géographique. Ainsi, 30 secondes d'arc au niveau de l'équateur est égal à environ 1 km. Les grilles ne sont généralement élaborées que pour les continents et les îles et excluent les régions de l'Arctique et de l'Antarctique.

⁵ Ce modèle exclut les zones inhabitées, ce qui ne signifie pas que personne ne vit dans les parcs ou autres lieux généralement admis comme étant inhabités.

27 | ⁶ La mesure par seconde d'arc découpe le globe en 360 radians de taille égale (degrés) où chaque radian est subdivisé en 60 minutes d'arc, elles-mêmes divisées en 60 secondes d'arc. Une seconde d'arc mesure la latitude ou la longitude parcourue à la surface de la terre en une seconde et représente la taille de la cellule (Esri, n.d.).

2.3 VUE D'ENSEMBLE DES JEUX DE DONNÉES DÉMOGRAPHIQUES MAILLÉES À L'ÉCHELLE MONDIALE

Les jeux de données démographiques maillées utilisent un large éventail de données d'entrée et de méthodes de modélisation pour produire les calculs de population. Pour obtenir un aperçu plus technique et une évaluation des jeux de données du POPGRID selon une grille « adaptée à l'emploi », les lecteurs sont invités à consulter les travaux de Leyk et al. (2019) « The spatial allocation of population ». La section suivante propose un bref aperçu de sept types de jeux de données développés par les organisations participant au POPGRID ([Tableau 2](#)). Ces jeux de données sont classés selon un degré de modélisation rapproché. Veuillez noter que cet aperçu n'est autre qu'un instantané ; aussi les lecteurs sont-ils invités à consulter les sites internet des fournisseurs de données et à se renseigner sur les dernières versions disponibles pour chaque fichier de données.



GRIDDED POPULATION OF THE WORLD (GPW) V4

En 1995, la première version du GPW a été développée par le géographe Waldo Tobler à l'Université de Californie-Santa Barbara, avec l'aide du Consortium for International Earth Science Information Network, une organisation à but non lucratif établie dans le Michigan. D'autres versions du GPW ont été développées par la suite et diffusées par le biais du Socioeconomic Data and Applications Center de la NASA (SEDAC), exploité par le Center for International Earth Science Information Network (CIESIN), désormais l'un des centres de l'Institut de la Terre de l'Université de Columbia, basé à New-York. La version 4 est sortie en 2017 et son extension 4.11 en 2018. Élaboré avec un minimum de variables et de modélisations, le GPWv4 est la source d'entrée de nombreux autres jeux de données, dont le Global Human Settlement Layer-Population et le WorldPop.

DONNÉES DÉMOGRAPHIQUES :

recensements de population et d'habitat de 2010, réalisés entre 2005 et 2014

CORRÉLATIONS GÉOSPATIALES :

zones protégées et plans d'eau

IMAGES SATELLITAIRES :

aucune

SOURCE D'AJUSTEMENT DES DONNÉES :

chiffres officiels des pays pour les recensements et estimations de la DPNU

MÉTHODE :

pondération aréolaire

ANNÉES PRISES EN COMPTE DANS LES ESTIMATIONS :

2000, 2005, 2010, 2015 et 2020.

RÉSOLUTION :

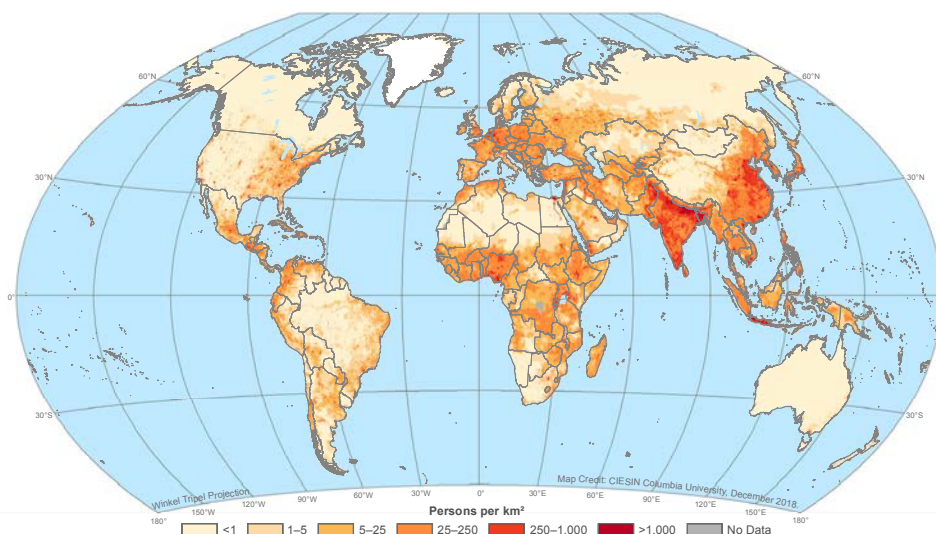
30 secondes d'arc (1 km)

POLITIQUE DE DISTRIBUTION :

accès libre (licence internationale émise par Creative Commons Attribution 4.0)

FORCES ET FAIBLESSES :

le GPWv4 est l'un des deux fichiers de données qui propose un calcul estimatif des principales caractéristiques démographiques (âge et sexe⁷) ainsi que du nombre d'habitants et de la densité de population (le second est le WorldPop). Au fil des années, les données de recensement régionales sont devenues de plus en plus accessibles, permettant ainsi d'améliorer considérablement le degré de résolution des résultats (Aubrecht et al., 2016; Doxsey-Whitfield et al., 2015). Le GPW est un fichier faiblement modélisé qui reste fidèle aux données de recensement et représente, par conséquent, un choix approprié pour toute analyse à grande échelle couvrant plusieurs régions et pour les analyses exigeant un décompte de population indépendant, qui ne soit pas influencé par les covariables utilisées dans la plupart des autres modes de modélisation. Toutefois, une telle approche à faible degré de modélisation conclut tendanciellement à un nombre d'habitants beaucoup plus élevé, notamment dans les zones rurales des pays où les données de recensement ne portent que sur les unités géographiques de plus grande taille (voir schéma 1, ci-dessus, concernant la frontière Nigéria-Tchad-Cameroun).



Source : [CIESIN](https://www.ciesin.columbia.edu/data/Products/GPW411)

⁷ Il convient de noter que, dans certains cas, les estimations sont établies à partir des données de recensement correspondant à des unités administratives plus importantes en taille, rapportées ensuite à des unités plus petites. Par exemple, les répartitions par âge et la proportion des sexes s'appliquent à tous les districts d'une même province. Les données portant sur l'âge et le sexe dans des unités de recensement plus petites (ex. zones de dénombrement) peuvent ne pas être disponibles pour des raisons de respect des règles de confidentialité.

GLOBAL RURAL URBAN MAPPING PROJECT (GRUMP)

Le CIESIN, l'Institut international de recherche sur les politiques alimentaires (IFPRI), la Banque mondiale et le Centro Internacional de Agricultura Tropical (CIAT) ont développé un projet mondial de cartographie urbaine et rurale (GRUMP) et ont publié le recueil de données GRUMP en 2007. Cet ensemble de données opère une répartition des données démographiques proportionnelle aux fichiers auxiliaires sur le logement et les implantations.

DONNÉES DÉMOGRAPHIQUES :

données de recensement régional du GPWv3 pour le recensement de 2000 (1995-2004)

CORRÉLATIONS GÉOSPATIALES :

éclairage nocturne pour les zones d'implantation les plus étendues, zones-tampon pour les implantations plus petites, zones protégées et plans d'eau

IMAGES SATELLITAIRES :

imagerie de l'éclairage nocturne produite par le Defense Meteorological Satellite Program Optical Linescan System (DMSP-OLS) pour l'année 1995

SOURCE D'AJUSTEMENT DES DONNÉES :

estimations de la DPNU

MÉTHODE :

cartographie dasymétrique binaire

ANNÉES PRISES EN COMPTE DANS LES ESTIMATIONS :

1990, 1995 et 2000.

RÉSOLUTION :

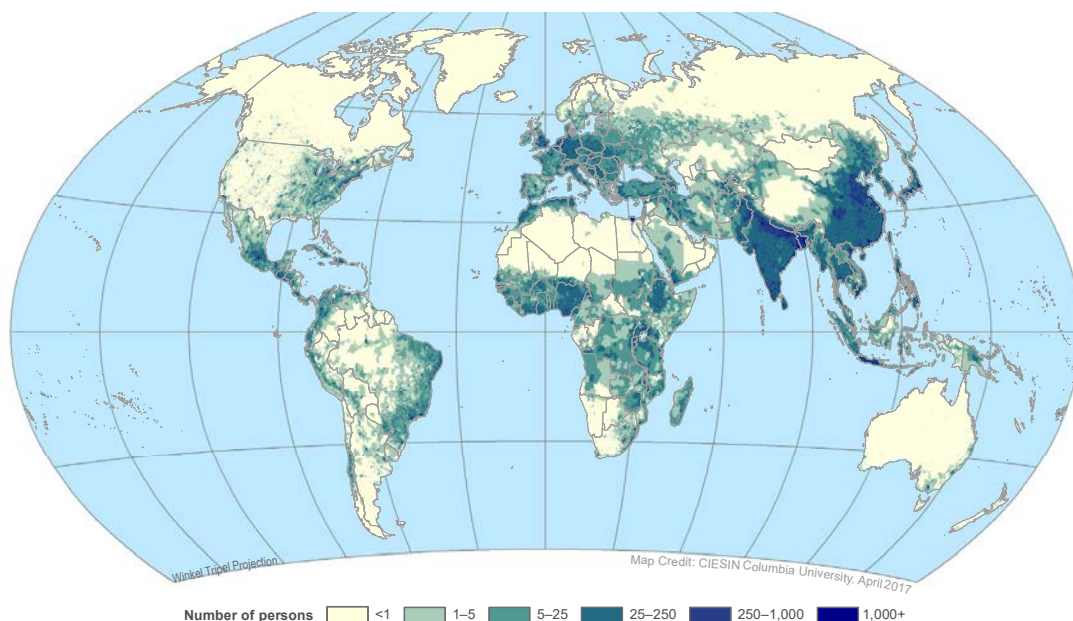
30 secondes d'arc (1 km)

POLITIQUE DE DISTRIBUTION :

accès libre (utilisation non commerciale)

FORCES ET FAIBLESSES :

Ile GRUMP est l'un des premiers jeux de données maillées à l'échelle mondiale proposant une vraie modélisation des populations urbaines et rurales, ses fichiers auxiliaires (extensions urbaines et zones peuplées) ayant été abondamment utilisés. L'une des faiblesses de ce système est le degré de résolution grossier et « l'efflorescence » de l'éclairage nocturne DMSP-OLS qui tend à surestimer les extensions urbaines dans certaines zones (Aubrecht et al., 2016, Gunasekera et al., 2015, Schneider et al., 2010). Le GRUMP est un ensemble utile pour faire une estimation des populations à risque et des populations des zones côtières à faible élévation car il propose une résolution spatiale plus fine ; il est également utile pour les nombreuses études qui doivent distinguer les zones rurales des zones urbaines (Mondal et Tatem, 2012). Le GRUMP a désormais été dépassé par d'autres fichiers de données dont la liste figure ci-dessous.



Source : [CIESIN](#)

GLOBAL HUMAN SETTLEMENT LAYER - POPULATION (GHS-POP)

En 2015, le Centre commun de recherche de la Commission européenne et le CIESIN ont publié la première version du jeu de données du Global Human Settlement Layer - Population (GHS-POP). La dernière édition date de 2019. Ce jeu associe les données de recensement de la population et celles de la surface bâtie par observation satellitaire et répartit, dans une grille, le nombre d'habitants par unité administrative, en fonction de l'existence ou de l'absence de surfaces bâties dans la cellule de grille. Les renseignements ci-dessous se rapportent à la dernière édition.

DONNÉES DÉMOGRAPHIQUES :

données du GPWv4.10

CORRÉLATIONS GÉOSPATIALES :

cartographie des surfaces bâties du Global Human Settlement Layer Built Up Extent (GHS-BUILT)

IMAGES SATELLITAIRES :

imagerie Landsat pour 1975, 1990, 2000 et 2014

SOURCE D'AJUSTEMENT DES DONNÉES :

estimations de la DPNU

MÉTHODE :

dasymétrique binaire

ANNÉES PRISES EN COMPTE DANS LES ESTIMATIONS :

1975, 1990, 2000 et 2015

RÉSOLUTION :

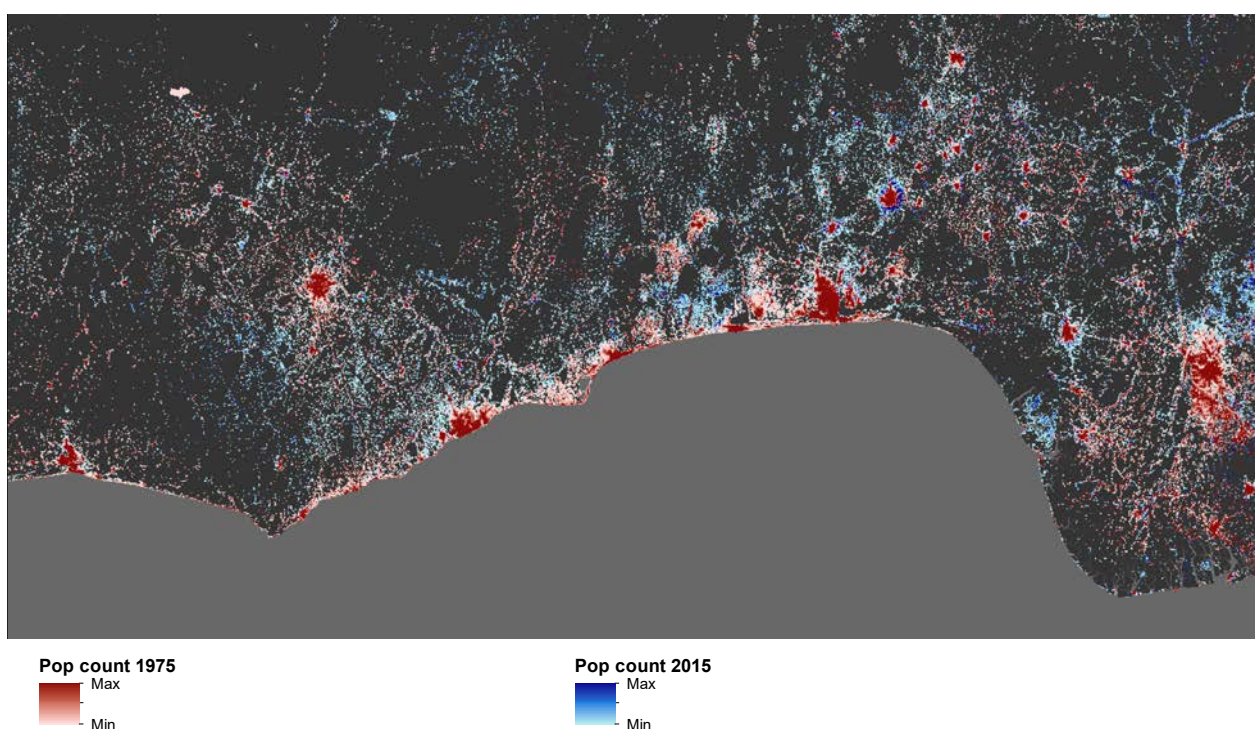
250 m et 1 km en projection de Mollweide à l'échelle mondiale, 9 secondes d'arc et 30 secondes d'arc dans le système géodésique mondial de 1984 (WGS84)

POLITIQUE DE DISTRIBUTION :

accès libre (licence émise par Creative Commons Attribution 4.0)

FORCES ET FAIBLESSES :

le GHS-POP a été développé à partir d'une hypothèse de cellules de grille comprenant un plus grand nombre de surfaces bâties et ayant une plus forte densité de population. Ce jeu de données utilise une méthode simple et reproductible dont la principale force est d'offrir une série chronologique cohérente à l'échelle mondiale, ce qui en fait une méthode adaptée aux analyses étalées sur plusieurs années (Freire et al., 2018). Sa faiblesse réside dans le fait que ce jeu n'emploie que la cartographie des surfaces bâties détectées uniquement par observation satellitaire. Ainsi, aux zones qui ne sont pas détectées comme étant des surfaces bâties n'est affectée que la densité de population déclarée pour l'unité de recensement (GPWv4), même si dans certaines zones rurales, elles comportent des unités d'habitation trop petites pour être repérées par un capteur à distance de résolution moyenne. Aussi, le GHS-POP peut-il indiquer une absence de population dans certaines zones rurales, alors qu'il existe peut-être de petites habitations éparses.



Source : [JRC](#)

WORLD POPULATION ESTIMATE (WPE)

Le World Population Estimate (ou WPE) développé par Esri rassemble des jeux de données démographiques à l'échelle mondiale qui ont été publiés entre 2014 et 2017. L'outil WPE indique l'emplacement où les populations résident (désignées comme « populations nocturnes ») et non l'endroit où elles se trouvent pendant la journée. Ce rassemblement de jeux de données a été conçu dans une optique d'analyse économique de la consommation car les habitudes des consommateurs varient en fonction de leur lieu de résidence.

DONNÉES DÉMOGRAPHIQUES :

estimations commerciales de l'année écoulée basées sur les recensements de différents pays et les données des Nations unies.⁸

CORRÉLATIONS GÉOSPATIALES :

carrefours et points de lieux habités fournis par GeoNames⁹

IMAGES SATELLITAIRES :

images Landsat 8

SOURCE D'AJUSTEMENT DES DONNÉES :

estimations officielles par pays dont 138 ont été traitées par le Michael Bauer Research GmbH

MÉTHODE :

modélisation dasymétrique

ANNÉES PRISES EN COMPTE DANS LES ESTIMATIONS :

2013, 2015 et 2016.

RÉSOLUTION :

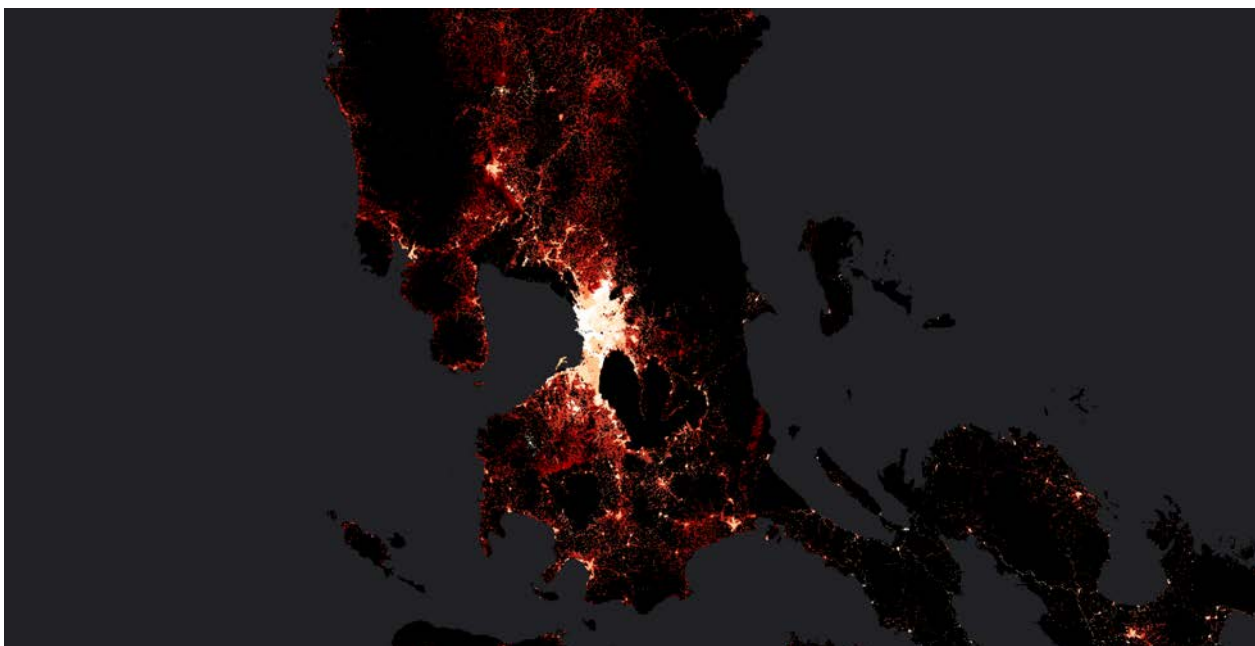
150 m en 2016 et 250 m au cours des années précédentes

POLITIQUE DE DISTRIBUTION :

fin commerciale (gratuit pour les personnes qui possèdent un compte d'utilisateur nominatif dans les organisations ArcGIS)

FORCES ET FAIBLESSES :

les estimations changent d'une année à l'autre, c'est pourquoi les résultats ne peuvent pas être comparés dans le temps. L'outil WPE est adapté aux analyses à grande échelle, ainsi qu'aux analyses économique, commerciale et d'entreprise. Le WPE est conçu pour estimer la répartition des affiliations aux partis politiques, l'impact humain sur l'environnement, la localisation des populations touchées par les catastrophes et les épidémies ainsi que pour analyser l'accès aux infrastructures (Frye, 2017). Une documentation détaillée sur les méthodes et les métadonnées au niveau national est disponible pour tous les jeux de données.



Source : [Esri](#)

⁸ Les estimations commerciales sont fournies par le Michael Bauer Research GmbH : elles commencent par le recensement le plus récent pour produire une estimation pour l'année en cours. L'outil WPE utilise cette méthode dans 138 pays et utilise les données des Nations unies pour les autres pays.

⁹ Il est possible de recueillir des points de lieux habités auprès de fournisseurs comme [GeoNames](#): les « lieux habités » sont extraits d'une base de données qui contient 25 millions de noms géographiques et se compose de millions de caractéristiques uniques. Les lieux de peuplement indiquent une plus grande possibilité d'établissements humains dans un faible rayon.

HIGH RESOLUTION SETTLEMENT LAYER (HRSL)

Dans le cadre d'une initiative conjointe de Facebook, du CIESIN et de la Banque mondiale, la première version du High Resolution Settlement Layer (HRSL), sortie en 2017, couvrait 30 pays. Facebook utilise des techniques de vision par ordinateur pour identifier les bâtiments à partir d'images satellitaires en vente libre. Dans un deuxième temps, le CIESIN utilise des données sur les bâtiments et celles issues des recensements pour produire des estimations démographiques. La méthodologie, validée en collaboration avec la Banque mondiale, s'appuie sur le programme LSMS (Étude de mesure des niveaux de vie) et le programme d'enquêtes anonymes auprès des ménages (en situation réelle et sur le terrain) et les utilise comme bases de comparaison avec le jeu de données démographiques maillées. La version actualisée, produite en 2018-2019, comprend environ 160 pays et propose une ventilation par âge et par sexe.

DONNÉES DÉMOGRAPHIQUES :

GPWv4

CORRÉLATIONS GÉOSPATIALES :

structures bâties

IMAGES SATELLITAIRES :

DigitalGlobe (désormais Maxar)

SOURCE D'AJUSTEMENT DES DONNÉES :

voir GPWv4

MÉTHODE :

modélisation dasymétrique binaire

ANNÉES PRISES EN COMPTE DANS LES ESTIMATIONS :

2015

RÉSOLUTION :

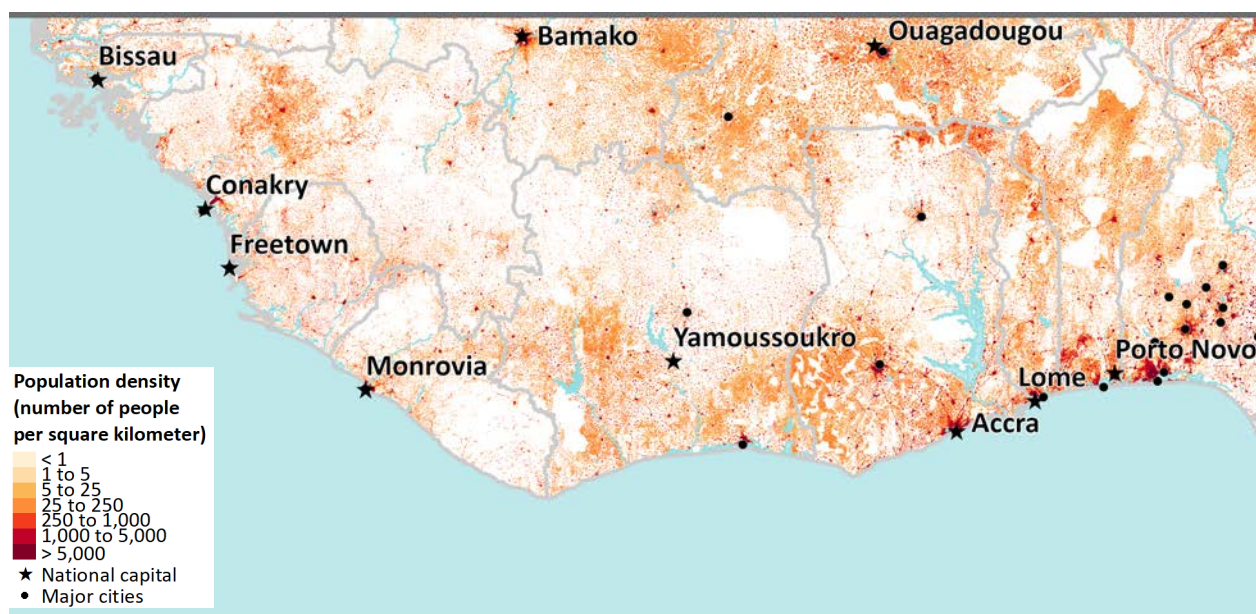
1 seconde d'arc (environ 30 m)

POLITIQUE DE DISTRIBUTION :

accès libre

FORCES ET FAIBLESSES :

les grilles démographiques de ce jeu de données fournissent une présentation détaillée des établissements humains dans les zones urbaines et rurales, ce qui peut étayer les travaux de recherches, notamment en matière d'intervention humanitaire et réaction en cas de catastrophe. Rappelons que l'outil HRSL ne peut établir de distinction entre les caractéristiques des bâtiments résidentiels et commerciaux. Toutefois, lors d'une évaluation de jeux de données démographiques maillées dans six pays, les corrélations du HRSL relatives aux zones construites se sont révélées être les plus significatives dans tous les pays (Reed et al., 2018).



Source : [CIESIN](#)

LANDSCAN

Le Laboratoire National d'Oak Ridge (ORNL) a sorti son premier jeu de données LandScan en 1998 et a publié des mises à jour annuelles depuis lors. La dernière version des jeux de données LandScan utilise la technique « d'interpolation intelligente » qui consiste à intégrer les données de recensement et les jeux de données géospatiales ou auxiliaires pour estimer la répartition démographique. À ce jour, le jeu de données de 2018 est la version disponible la plus récente.

DONNÉES DÉMOGRAPHIQUES :

recensement infranational et données administratives

CORRÉLATIONS GÉOSPATIALES :

routes, couverture terrestre, structures bâties, municipalités ou zones urbaines, infrastructures, données environnementales, zones protégées et plans d'eau

IMAGE SATELLITE :

Maxar (autrefois DigitalGlobe)

SOURCE D'AJUSTEMENT DES DONNÉES :

estimations du Centre des programmes internationaux du Bureau du recensement des États-Unis

MÉTHODE :

modélisation dasymétrique à variables multiples mais la méthode employée varie selon les pays et les périodes

ANNÉES PRISES EN COMPTE DANS LES ESTIMATIONS :

jeux de données annuels depuis 1998

RÉSOLUTION :

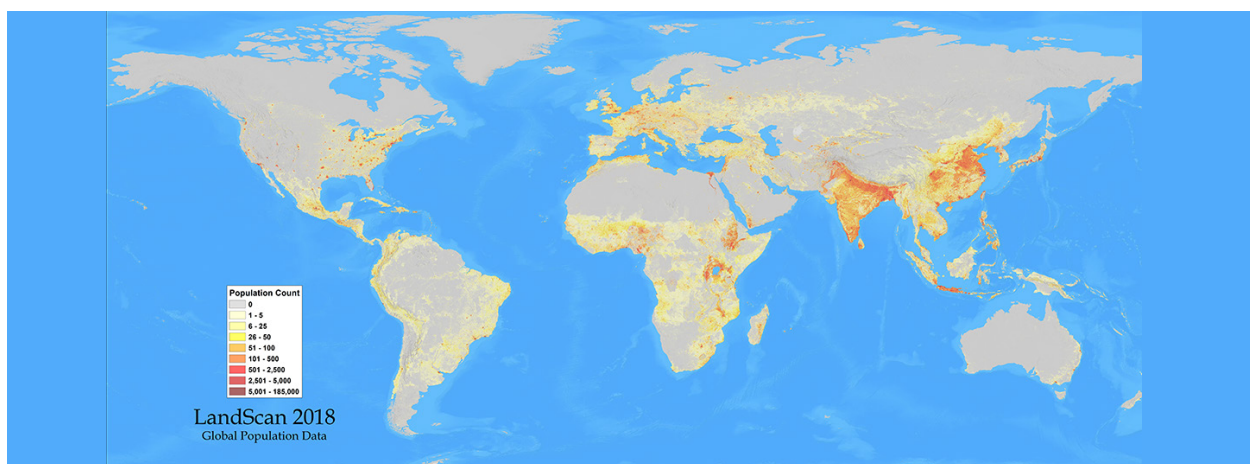
30 secondes d'arc (1 km)

POLITIQUE DE DISTRIBUTION :

accès libre dans le cadre de la recherche universitaire, à des fins humanitaires et pour les organismes publics ou les fonctionnaires des États-Unis et leurs partenaires de mission. Licence commerciale disponible pour les autres acteurs.

FORCES ET FAIBLESSES :

une des particularités du LandScan est qu'il modélise la « population ambiante », ce qui signifie que pour les régions où les déplacements domicile-travail sont fréquents, la grille démographique se situe quelque part entre une population résidentielle de nuit (ce que mesurent les recensements) et une population de jour (où les gens travaillent). Le LandScan utilise une modélisation complexe adaptée à chaque pays de façon à pondérer les cellules en fonction de la présence éventuelle de population tout au long de la journée. Cela peut être utile pour l'évaluation des risques et la réaction aux catastrophes en cas de menaces diurnes. En raison de la diversité des données saisies et des méthodes utilisées, il convient de comparer avec soin les estimations du LandScan avec d'autres modèles, en particulier pour les régions plus petites ou les zones urbaines où les populations dites « ambiantes » et nocturnes peuvent être très différentes. Le LandScan inclut des données démographiques par âge et par sexe, sur une durée de 5 ans, au niveau 1 de l'unité administrative. Le LandScan 2012 est également un outil adapté pour détecter la répartition de la population urbaine, mais il est moins fiable dans les zones rurales (Gunasekera et al., 2015). Cela est dû en partie aux insuffisances des images satellitaires dans certains environnements qui rendent difficile la détection des établissements humains (par exemple, les maisons peuvent se fondre dans le paysage ou être cachées par des arbres). En outre, les données et les méthodes employées n'étant que peu divulguées, les données ne sont pas entièrement accessibles à tous les utilisateurs (Gunasekera et al., 2015). Toutefois, le LandScan prévoit de passer à un accès totalement libre à l'avenir. Le LandScan développe de nouveaux modèles pour les estimations de chaque année, ce qui peut exclure l'utilisation des données comme série chronologique homogène pour des zones peu étendues.



Source : [ORNL](#)

WORLDPOP

Lancé en 2012, le WorldPop rassemble les projets de cartographie démographique AfriPop, AsiaPop et AmeriPop afin d'englober l'ensemble de la population mondiale. Le WorldPop produit des données démographiques maillées au moyen d'algorithmes de l'apprentissage machine.

DONNÉES DÉMOGRAPHIQUES :

données de recensement GPWv4

CORRÉLATIONS GÉOSPATIALES :

routes, couverture terrestre, structures bâties, municipalités ou zones urbaines, éclairage nocturne, infrastructures, données environnementales, zones protégées et plans d'eau

IMAGE SATELLITE :

Landsat Enhanced Thematic Mapper (ETM), Landsat, TerraSAR-X, TanDEM-X, DMSP et Visible Infrared Imaging Radiometer Suite (VIIRS)

SOURCE D'AJUSTEMENT DES DONNÉES :

estimations officielles par pays et estimations du DPNU

MÉTHODE :

forêts aléatoires mais la méthode employée varie selon les pays et les périodes

ANNÉES PRISES EN COMPTE DANS LES ESTIMATIONS :

chaque année depuis 2000-2020

RÉSOLUTION :

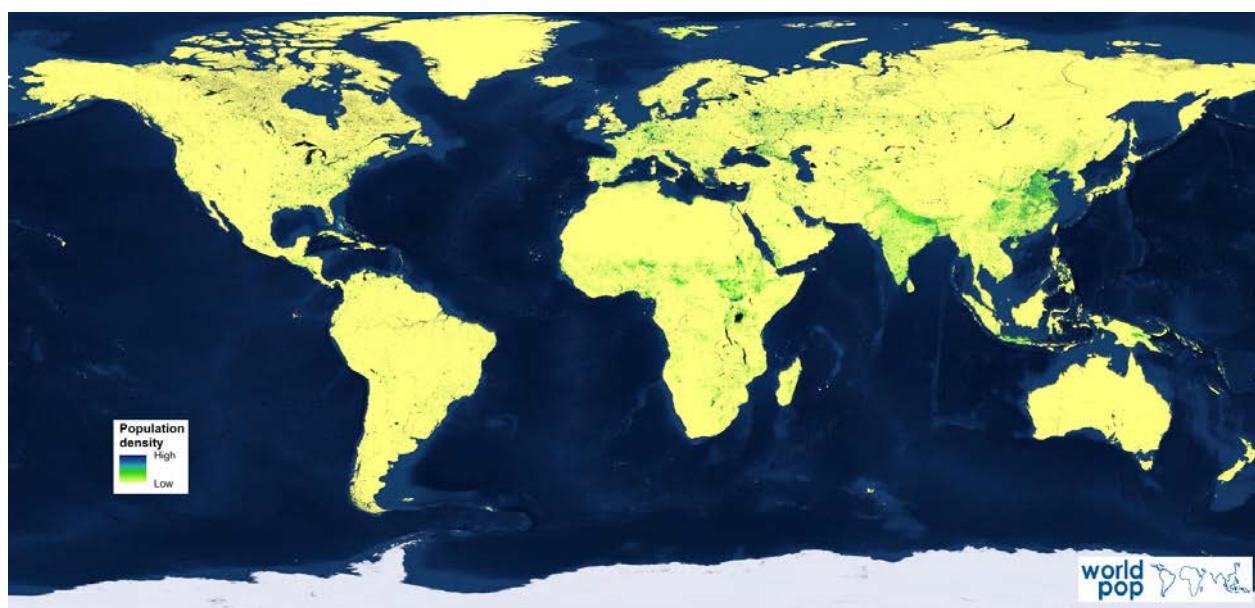
3 secondes d'arc (100 m)

POLITIQUE DE DISTRIBUTION :

libre accès

FORCES ET FAIBLESSES :

un des points forts du modèle WorldPop est sa capacité à utiliser l'apprentissage machine pour identifier les relations significatives à partir des données de recensement collectées et à omettre une partie de la population rurale dans les zones qui ne semblent pas avoir de zones bâties dérivées des satellites ; cela est dû au fait que les satellites manqueront inévitablement certains endroits peuplés. Le WorldPop publie également l'ensemble du code source et fait preuve de transparence quant aux méthodes employées, afin que les utilisateurs puissent développer leurs propres grilles en utilisant différents ensembles d'hypothèses. Le WorldPop intègre également diverses données d'entrée et de données auxiliaires de sorte que le modèle redistribue les chiffres de population en utilisant différentes pondérations entre les unités de recensement ou unités administratives (Stevens et al., 2015). Comme le GPWv4, le WorldPop inclut des ventilations démographiques par âge et par sexe, ce qui en fait un candidat de choix pour les indicateurs ODD liés à la population, et fournit également un jeu de données avec une résolution de 100 m pour chaque année de 2000 à 2020. Étant donné que ce modèle ne masque pas les zones non construites, il peut produire des estimations de population non nulles dans les déserts et les forêts. Cependant, parce qu'il regroupe les populations, le modèle a été critiqué pour avoir exclu plus de surfaces inhabitées que les autres ensembles de données (Aubrecht et al., 2016). Il est important de souligner que le jeu de données mondiales décrit ici est conçu pour être comparable dans le temps.



Source : [WorldPop](https://worldpop.org/)

TABLEAU 2 PARAMÈTRES DES JEUX DE DONNÉES DÉMOGRAPHIQUES MAILLÉES POPGRID

	GPWV4	GRUMP	GHS-POP	WPE	HRSL	LANDSCAN	WORLDPOP
DONNÉES DÉMOGRAPHIQUES	Recensements de population et d'habitat de 2010	Données de recensement régional du GPWv3 pour le recensement de 2000	Données du GPWv4.10	Recensements de plusieurs pays, données commerciales et de l'ONU	GPWv4	Données de recensement infranational et administratives	Données de recensement GPWv4
CORRÉLATS GÉOSPATIAUX	Zones protégées et plans d'eau	Éclairage nocturne pour les zones d'implantation les plus étendues, zones-tampon pour les implantations plus petites, zones protégées et plans d'eau	Cartographie des surfaces bâties du Global Human Settlement Layer Built Up Extent (GHS-BUILT)	Carrefours et points de lieux habités fournis par GeoNames	Structures bâties	Routes, couverture terrestre, structures bâties, municipalités ou zones urbaines, infrastructures, données environnementales, zones protégées et plans d'eau	Routes, couverture terrestre, structures bâties, municipalités ou zones urbaines, éclairage nocturne, infrastructures, données environnementales, zones protégées et plans d'eau
IMAGES SATELLITAIRES	Néant	Imagerie de l'éclairage nocturne produite par le Defense Meteorological Satellite Program Optical Linescan System (DMSP-OLS) pour l'année 1995	Imagerie Landsat pour 1975, 1990, 2000 et 2014	Landsat8	DigitalGlobe (désormais Maxar)	Maxar (autrefois DigitalGlobe)	Landsat Enhanced Thematic Mapper (ETM), Landsat, TerraSAR-X, TanDEM-X, DMSP et VIIRS
SOURCE DE L'AJUSTEMENT DES DONNÉES	Chiffres officiels des pays pour les recensements et estimations de la DPNU	Estimations de la DPNU	Estimations de la DPNU	Estimations officielles par pays dont 138 ont été traitées par Michael Bauer Research GmbH	GPWv4	Estimations du Centre des programmes internationaux du Bureau du recensement des États-Unis	Estimations officielles par pays et estimations du DPNU
MÉTHODE	Pondération areolaire	Modélisation dasymétrique binaire	Modélisation dasymétrique binaire	Modélisation dasymétrique	Modélisation dasymétrique binaire	Modélisation dasymétrique à variables multiples	Forêts aléatoires
ANNÉES COMPRISES	2000, 2005, 2010, 2015 et 2020	1990, 1995 et 2000	1975, 1990, 2000 et 2015	2013, 2015 et 2016	2015	Chaque année depuis 1998	2000-2020
RÉSOLUTION	30 secondes d'arc (1 km)	30 secondes d'arc (1 km)	250 m et 1 km en projection de Mollweide à l'échelle mondiale, 9 secondes d'arc et 30 secondes d'arc (WGS84)	150 m en 2016 et 250 m pour les années précédentes	1 seconde d'arc (environ 30 m)	30 secondes d'arc (1 km)	3 secondes d'arc (100 m)
POLITIQUE DE DISTRIBUTION	Accès libre (licence internationale émise par Creative Commons Attribution 4.0)	Accès libre (utilisation non commerciale)	Accès libre (licence internationale émise par Creative Commons Attribution 4.0)	Fin commerciale (gratuit pour les personnes qui possèdent un compte d'utilisateur nominatif dans les organisations ArcGIS)	Accès libre	Accès libre dans le cadre de la recherche universitaire, à des fins humanitaires et pour les organismes publics ou les fonctionnaires des États-Unis et leurs partenaires de mission. Licence commerciale disponible pour les autres acteurs.	Accès libre

2.4 AUTRES APPROCHES DE MODÉLISATION

L'approche descendante de la modélisation appliquée aux jeux de données décrits dans ce document repose sur la disponibilité des données de recensement, malgré les différences entre celles-ci sur plusieurs plans : échelle, qualité et chronologie. Cependant, certains fournisseurs de données ont étudié des méthodes d'estimation qui adoptent une approche ascendante pour compenser le déficit de données de recensement fiables (Wardrop et al., 2018 ; Weber et al., 2018). Avec l'approche ascendante, les données auxiliaires et le recueil de données à petite échelle sur le terrain (Reed et al., 2018) permettent de calculer le dénombrement et la densité démographiques. Par exemple, le WorldPop utilise l'approche ascendante avec des caractéristiques issues des satellites et des enquêtes sur les ménages lorsque les données de recensement sont obsolètes, peu fiables ou non disponibles. Dans le cadre de l'initiative [GRID3](#) (Infrastructure géo-référencée et données démographiques pour une initiative de développement), le but du WorldPop est d'associer les données provenant du dénombrement de population localisé à des images satellitaires pour estimer les populations des zones dont les données de recensement sont obsolètes. À ce jour, l'approche ascendante a été appliquée aux estimations de population au Nigeria, en Afghanistan, en Colombie et en République démocratique du Congo. Le LandScan HD (3 secondes d'arc) utilise également une approche ascendante pour produire des données démographiques à une résolution de 3 secondes d'arc. Bien que les données à cette résolution ne soient pas accessibles aux utilisateurs, elles sont agrégées à une résolution de 30 secondes d'arc et utilisées dans LandScan Global.

ENCADRÉ 3 | UTILISATION DE NOUVELLES SOURCES DE DONNÉES

Dans la grande majorité des entretiens du GPSDD, les représentants des bureaux nationaux des statistiques ont indiqué qu'ils méconnaissaient les données démographiques maillées décrites dans ce rapport. La seule exception était la Colombie, qui a coopéré avec le Fonds des Nations unies pour la population et l'Université de Southampton pour développer des modèles d'estimation de la population au niveau des pâtés de maisons (en utilisant une approche ascendante) ; ces modèles sont susceptibles de réduire le degré d'incertitude des estimations pour les zones difficiles d'accès.

Bien que les données démographiques maillées soient moins répandues, de nombreux pays commencent à recourir à d'autres sources de données. La Tanzanie aurait entamé un dialogue avec la Banque mondiale en vue d'augmenter l'utilisation des données géospatiales dans le Bureau national des statistiques, quoique cette initiative ne s'étende pas actuellement aux estimations de la population. Par ailleurs, plusieurs pays envisagent de nouvelles utilisations pour les données administratives. Il s'agit de données collectées dans le cadre de la tenue générale de registres et qui peuvent ensuite être réutilisées pour les calculs statistiques. Par exemple, le Paraguay estime, actuellement, les évolutions démographiques sur la base des registres des naissances, des décès et des migrations, et ce pays prévoit de procéder à un recensement basé sur les données administratives. L'Équateur prévoit également de comparer les résultats de son recensement de 2020 avec des données administratives. De plus, la Colombie, le Costa Rica et le Ghana utilisent des données administratives pour estimer l'évolution de la population. En outre, le service statistique du Ghana s'est associé à la société de télécommunications Vodafone pour accéder à des données agrégées sur les appels qui seront utilisées pour comprendre les évolutions en matière de mobilité et d'accès aux services publics. Grâce aux nouvelles sources de données, actuellement étudiées ou déjà exploitées, les données démographiques maillées peuvent s'ajouter à la gamme croissante d'outils mis à la disposition des décideurs.

(Pays partenaires du GPSDD, 2019)

2.5 CHOIX DE LA SOLUTION ADAPTÉE

Il existe actuellement différentes façons d'estimer la répartition de la population, au moyen de diverses sources de données, méthodes de base et résolutions de sortie. Les parties prenantes, dont les décideurs politiques et les universitaires, ont désormais le choix entre plusieurs produits. La comparaison des approches de modélisation est la première étape de cette sélection. Cependant, pour déterminer si un jeu de données est adapté à l'application souhaitée, il est important de vérifier comment les résultats de ces modèles soutiennent la comparaison, autant entre eux que par rapport aux conditions sur le terrain. Dans le chapitre suivant, nous passons en revue les informations disponibles sur la comparaison et la validation de ces produits de données démographiques. En outre, nous justifions la nécessité d'un travail de validation supplémentaire pour améliorer la fiabilité des résultats de ces différents fichiers de données et leur adéquation à différentes applications.



COMPARAISON CROISÉE ET VALIDATION DES JEUX DE DONNÉES MAILLÉES

Comme décrit au premier chapitre, les données démographiques maillées peuvent s'avérer utiles dans le cadre de certaines initiatives importantes comme la réaction en cas de catastrophe et la planification des enquêtes auprès de ménage. Cependant, on peut se rendre rapidement compte des inexactitudes et des erreurs une fois sur le terrain (Thomson, 2019). Dans certains cas, les calculs étaient sensiblement erronés. Des enquêteurs au Népal, par exemple, ont découvert que le WorldPop avait placé de manière incorrecte des habitants dans un bâtiment du Parlement, un aéroport et même une usine de Coca-Cola (Thomson, 2019). De même, en Namibie, de grandes formations rocheuses ont été identifiées, à tort, comme des maisons, lesquelles ne figurent généralement pas sur les résultats des régions du monde au couvert végétal fourni (Thomson, 2019). Des expériences comme celles-ci permettront d'affiner les prévisions à l'avenir mais illustrent également que les prédictions démographiques basées sur des modèles sont loin d'être parfaites.

Différents jeux de données démographiques étant désormais disponibles, il est important de bien comprendre la manière dont la sélection des données peut influencer les études et les décisions. Il est également essentiel de mettre à l'épreuve la précision de ces données en comparant leurs calculs aux conditions connues sur le terrain. Les estimations démographiques comportent toujours un certain degré d'incertitude. Toutefois, il n'existe pour l'heure aucun moyen reconnu de mesurer ou de communiquer les degrés d'incertitude associés aux produits de données actuellement disponibles. Les chercheurs ont commencé à s'atteler à cette problématique et leurs premières conclusions indiquent plusieurs pistes. Toutefois, il est nécessaire de mener des études plus systématiques et plus approfondies pour mettre au point des moyens objectifs de mesure de la précision et de l'incertitude qui soient en mesure d'aider les utilisateurs à exploiter les estimations démographiques de manière efficace. Dans la section ci-dessous, nous décrivons la manière dont les chercheurs testent, valident et améliorent leurs modèles.

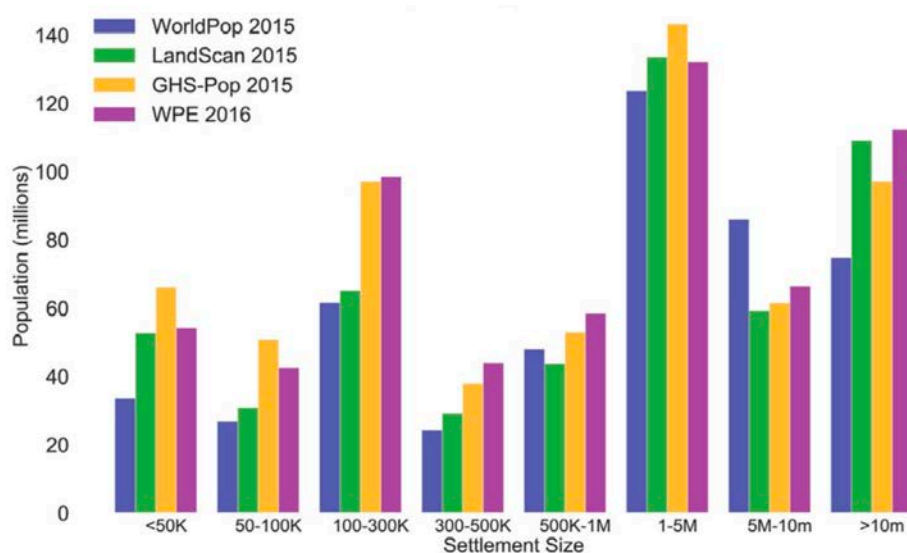
3.1 COMPARAISON ENTRE LES JEUX DE DONNÉES

Étant donné que les différentes grilles démographiques sont élaborées à partir d'hypothèses distinctes sur la répartition des populations, il n'est pas surprenant que les estimations qui en résultent soient également différentes. Une comparaison objective des grilles démographiques peut nous aider à mieux comprendre les différences et les limites des jeux de données et la nature de ces différences. En l'absence de données de validation reconnues, il n'est pas possible de se prononcer sur l'estimation la plus proche de la population réelle, mais les comparaisons entre les modèles peuvent faire la lumière sur les conséquences que peut avoir un choix de modèle plutôt qu'un autre.

DESCRIPTION DES DIFFÉRENCES ENTRE LES ESTIMATIONS

Les écarts entre les dénombrements de population pour la plupart des pays sont négligeables, mais peuvent être significatifs pour les zones plus petites (Calka et Bielecka, 2019). Une étude de 2007 a comparé le LandScan Global 2004 au GPWv3 dans plusieurs régions, dont les villes de Washington, Los Angeles, Houston, Londres et en Iran (Sabesan et al., 2007). Les chercheurs ont constaté des écarts significatifs dans les centres-villes où le LandScan donne généralement des estimations démographiques plus importantes, en raison de son procédé de mesure de la population « ambiante ». En outre, le LandScan indique généralement une plus forte densité de la population, alors que le GPWv3 indique une répartition plus diffuse dans l'espace. Les auteurs ont conclu que le LandScan semblait saisir une plus grande variabilité géospatiale, ce qui suggère qu'il est plus adapté à la planification des catastrophes. Dans une autre étude, les chercheurs ont comparé les estimations démographiques des établissements urbains en Afrique selon cinq méthodes différentes (Tuholske et al., 2019). Les estimations de la population urbaine totale du continent vont de 479,15 millions pour le WorldPop 2015 à 608,89 millions pour le WPE 2016. Cet écart souligne l'incertitude qui entoure l'identification et le décompte des populations urbaines et rurales, en particulier dans les pays en développement.

Schéma 3 : population urbaine totale par taille d'établissement pour l'Afrique



Source : [Tuholske et al., 2019](#)

INCIDENCES SUR LES RÉSULTATS

Des écarts de calcul peuvent affecter l'ensemble d'un modèle et avoir une incidence sur les prévisions issues de variables essentielles. Le fichier de données AfriPop, particulier à l'Afrique de l'Est, était un prédécesseur du WorldPop. Les développeurs de cet outil ont effectué une série de comparaisons informelles avec le LandScan 2008 et 2000 et le GRUMP Beta (Linard et al., 2010). Ils ont testé les répercussions que la sélection d'un modèle pouvait avoir sur les prévisions d'exposition au paludisme en Somalie : les modèles comparés ont montré que le GRUMP estimait qu'une proportion significativement plus élevée de la population serait exposée. Même si l'AfriPop et le LandScan ont obtenu des répartitions géographiques similaires, de grands écarts étaient à déplorer dans les chiffres absolus de la population. De plus, le GRUMP a estimé un nombre beaucoup plus important de personnes localisées dans les zones d'épidémie. Ces écarts peuvent avoir un impact considérable sur des décisions capitales, y compris les décisions en matière de santé publique, et peuvent influencer les prévisions des taux de morbidité et de mortalité utilisées dans la planification des épidémies.

Le choix d'un modèle influence aussi les prévisions relatives aux effets du changement climatique et de l'élévation du niveau de la mer sur les populations vulnérables. Selon Mondal et Tatem (2012), les études qui estiment les populations côtières vulnérables au changement climatique et aux inondations utilisent généralement des jeux de données démographiques à haute résolution mais ne tiennent pas compte de l'incidence du choix des données. Pour évaluer cette incidence, les chercheurs ont adopté des techniques existantes de prédiction des inondations côtières, puis ont utilisé en alternance le LandScan 2008 et le GRUMP 2000 v1. Les chercheurs ont constaté que la variation de la population totale à risque entre le GRUMP et le LandScan était faible au niveau continental mais qu'il existait des différences significatives entre les différents pays. Les pays développés disposant de données d'entrée à haute résolution présentaient peu d'écarts dans les estimations, tandis que les pays à faible revenu affichaient des écarts plus importants. En outre, la majorité des pays pour lesquels la sélection du modèle était particulièrement lourde de conséquences étaient de petits pays insulaires (plus de 25 % d'écart entre les estimations). En somme, les degrés d'incertitude sont les plus élevés dans le cas des populations les plus exposées au risque d'élévation du niveau de la mer. Une étude plus récente a examiné le risque d'inondation côtière dans 18 pays d'Asie, d'Amérique latine et d'Afrique et a comparé les conséquences associées au choix d'un des modèles suivants : HRSL 2014, WorldPop (version inconnue) et LandScan 2015 ; chacun de ces modèles était disponible à différentes résolutions spatiales (30 m, 90 m et 900 m respectivement) (Smith et al., 2019). Dans les 18 pays, l'utilisation du HRSL à haute résolution a permis d'obtenir des estimations plus faibles de la population exposée à une crue centennale, tandis que le WorldPop et le LandScan ont montré des réductions allant jusqu'à 60 %. Les auteurs ont conclu que les populations se concentrent naturellement dans les zones moins sujettes aux inondations et que les connaissances issues des analyses sont nécessaires à la planification de l'adaptation au climat. Les sensibilités des modèles peuvent introduire une incertitude considérable dans les recherches sur la population mais cette incertitude est souvent sous-estimée. Les chercheurs devraient évaluer leur choix de modèle de manière rigoureuse et faire preuve de transparence quant aux incertitudes qui y sont associées.

3.2 MÉTHODES DE VALIDATION DES ESTIMATIONS DÉMOGRAPHIQUES

Les grilles démographiques doivent être impérativement validées au regard de la population sur le terrain pour garantir des estimations aussi précises que possible. Même si les calculs concordent bien entre eux, cela ne garantit pas qu'ils reflètent également le domicile et le lieu de travail réels des personnes, en particulier dans des environnements en évolution rapide, comme les zones périurbaines des pays en développement. La validation doit donc passer par une comparaison des calculs aux chiffres de population officiels. Une telle comparaison est cependant difficile à réaliser dans de nombreuses régions du globe. Si plusieurs études de validation au niveau national et local ont été menées dans le passé, la validation à l'échelle mondiale reste difficile, en partie parce que les données de validation ne sont pas actuellement disponibles ou accessibles pour de nombreuses régions du monde (Calka et Bielecka, 2019 ; Engstrom, 2019).

COMPARAISON AVEC LES DONNÉES DE RECENSEMENT

Selon la littérature actuelle, la méthode de validation la plus courante consiste à procéder à des comparaisons avec les données de recensement disponibles (Bai et al., 2018 ; Biljecki et al., 2016 ; Calka et Bielecka, 2019 ; Engstrom et al., 2019 ; Gaughan et al., 2013 ; Hall et al., 2012 ; Hay et al., 2005 ; Linard et al., 2010). Cette approche est particulièrement utile lorsque les données de recensement peuvent être mises à disposition à des résolutions élevées. Bien que les modèles démographiques

fonctionnent sur la base de données de recensement agrégées, leurs résultats peuvent être comparés à des données plus fines pour vérifier la précision de leur répartition démographique. Les données de recensement des ménages géoréférencées sont idéales pour les travaux de validation (Hall et al., 2012) mais les organismes de recensement ne publient généralement que des données agrégées (Chen, 2019). En outre, la validation est particulièrement difficile lorsqu'un modèle est produit à une résolution plus élevée que le niveau administratif le plus petit selon les données de recensement disponibles (Bai et al., 2018).

COMPARAISON DES MÉTHODES

Qu'il s'agisse de grilles faiblement ou hautement modélisées, toutes les méthodes comportent des erreurs de calculs. Cependant, on ne peut connaître la nature de ces erreurs sans effectuer un travail de validation. La pondération aréolaire, employée pour créer les bases de données GPW et GRUMP, présente l'avantage de ne formuler que très peu d'hypothèses. Toutefois, cette même particularité peut entraîner d'autres types d'erreurs de calculs. Dans de bonnes conditions, les méthodes de cartographie dasymétrique peuvent fournir des résultats plus réalistes (Calka et Bielecka, 2019). Une étude réalisée en Pologne a permis d'en faire la démonstration : les chercheurs ont testé la précision des données du GRUMP faiblement modélisées et des données plus finement modélisées du LandScan 2012 par rapport à une grille démographique de référence fournie par les pouvoirs publics, qui était basée sur les données du recensement de 2011 et établie conformément au cahier des charges de la directive INSPIRE de l'UE (Da Costa et al., 2017 ; Calka et Bielecka, 2019 ; Backer et Bloch, 2011). Dans les deux cas, les chercheurs ont d'abord identifié les écarts entre les calculs obtenus et les données de référence, puis ont analysé les caractéristiques spatiales de ces écarts. Cette méthode a fait apparaître d'importantes divergences dans le GRUMP par rapport à la grille de référence, seules 1,5 % des cellules de grille n'affichant aucune différence entre les deux (Da Costa et al., 2017). En revanche, le LandScan était généralement conforme à la grille de référence : 10,5 % des cellules de grille donnaient des estimations démographiques équivalentes à la référence et des estimations extrêmement fiables ou très fiables s'appliquaient à 72 % du pays (Calka et Bielecka, 2019). De même, un autre groupe de chercheurs a conjugué des sources de données pour localiser avec précision le million d'habitants de la Scanie, en Suède, puis ont transposé ces données sur une grille standard pour les comparer directement avec le GPWv3, le LandScan, le GRUMP et la grille de densité de population de l'UE-27+ (Hall et al., 2012). Comme en Pologne, les calculs du LandScan étaient dans l'ensemble plus proches des chiffres de la population de référence que les autres méthodes.

Ce travail de validation met en évidence les tendances et les lacunes des différentes méthodes d'estimation. Il ressort de ce travail que le GRUMP, dont les calculs reposent sur les données de l'éclairage nocturne, a tendance à surestimer les extensions urbaines (Gaughan et al., 2013), comme le démontrent les études comparatives ci-dessus. En Pologne, le GRUMP, dont la tendance est de sous-estimer la densité de population dans les centres villes et de la surestimer dans les banlieues, a évalué à la baisse la densité de population moyenne, selon un écart de 50 % (Da Costa et al., 2017). Le LandScan présente également des lacunes : par exemple, en Pologne, il surestime les centres villes, mais sous-estime d'autres zones à forte densité démographique et les banlieues (Calka et Bielecka, 2019). Il faut cependant souligner que le LandScan estime la population « ambiante » (moyenne sur 24 heures) ; il faut par conséquent s'attendre à ce qu'il propose une répartition de la population différente de celle du recensement traditionnel. Toutefois, ces tendances sont renversées à Scanie, en Suède, où le GPWv3 et le GRUMP ont tous deux conclu à une surestimation dans les villes et à une sous-estimation dans les banlieues, tandis que le degré de précision du LandScan variait en fonction de la géographie et de facteurs tels que l'utilisation des terres agricoles (Hall et al., 2012). Ces études de validation montrent, non seulement qu'un modèle peut être mieux adapté

à une zone ou un pays donné et que la fiabilité des calculs peut varier au sein d'un même pays, mais aussi la nécessité d'améliorer les méthodes et les données pour garantir la pertinence des travaux de validation. Il est également important d'utiliser les versions les plus récentes des données disponibles, car les méthodes et les données sont régulièrement mises à jour (par exemple, le GRUMP et le GPWv3 ont été remplacés par des versions plus récentes ou des jeux de données différents).

ÉCARTS RÉGIONAUX

Les modèles de population qui sont bien adaptés aux pays très développés ne conviennent pas nécessairement aux pays en développement, notamment lorsque les caractéristiques de la répartition de la population sont différentes ou que les données sous-jacentes sont plus difficiles à obtenir. Les modèles AsiaPop et AfriPop, tous deux prédécesseurs du modèle WorldPop, se sont avérés plus précis dans les régions pour lesquelles ils ont été créés. Par exemple, une mise à l'essai d'AsiaPop au Cambodge et au Vietnam a montré qu'il était généralement adapté aux besoins et présentait moins de variabilité dans les erreurs par rapport aux données de référence que le GPWv3 ou le GRUMPv1 (Gaughan et al., 2013). De même, les calculs d'AfriPop étaient très proches des données de référence de la DPNU en Somalie, alors que le LandScan 2008 avait tendance à donner des surestimations et le GRUMP bêta 2000 à donner des sous-estimations dans la plupart des districts, les écarts étant particulièrement marqués dans les zones périurbaines (Linard et al., 2010). En Chine, l'équipe Bai et al. (2018) a mis à l'essai le GPWv3, le GRUMPv1, le WorldPop, ainsi qu'une grille de population spécifique à la Chine (CnPop), en calculant les erreurs par rapport aux données de recensement au niveau des villes liées à un SIG à partir de l'année 2000. Ils ont constaté que le modèle WorldPop était le plus précis des quatre. La méthode de pondération aréolaire, telle qu'utilisée par le GPWv3 et le GRUMP, pose des problèmes d'estimation en Chine, alors que les méthodes analysant la couverture terrestre, comme le WorldPop et le CnPop, semblent plus précises. Il convient toutefois de noter que le GPWv3 n'est disponible qu'à une résolution plus grossière que le GPWv4 (la version la plus récente).

Dans le cadre d'une comparaison au Sri Lanka entre plusieurs grilles de population et les données de référence du recensement de 2011 pour le Gram Niladhari (GN), l'étude menée par Engstrom et al. (2019) a montré que le LandScan était l'outil le moins précis, tandis que le HRSL de Facebook suivi du GHS-POP étaient les plus précis (valeurs du coefficient de détermination de 0,84 et 0,67, respectivement). Le GPWv4 a obtenu des résultats assez satisfaisants avec une valeur du coefficient de détermination de 0,4. Pour les échelons administratifs moins précis, la concordance entre les outils était, dans l'ensemble, supérieure. Il semble que l'approche de modélisation « ambiante » du LandScan n'ait pas permis de rapprochement avec les résultats du recensement des ménages, bien que les créateurs reconnaissent également que les données démographiques sous-jacentes puissent être imparfaites (Bhaduri, communication personnelle, 2018).

SUPÉRIORITÉ DES DONNÉES À HAUTE RÉOLUTION

Divers travaux de validation ont démontré que des données d'entrée à plus haute résolution peuvent fournir des estimations plus fiables. Les développeurs du HRSL ont testé la précision de leurs calculs en les comparant à un jeu de données référencées par l'homme et à des données d'enquêtes de ménage dans 18 pays d'Asie, d'Amérique latine et d'Afrique et ont conclu que les erreurs semblaient être négligeables (Tiecke et al., 2017). En effet, le HRSL repère souvent des bâtiments que d'autres modèles ne détectent pas et ses calculs sont particulièrement supérieurs dans les zones rurales.

Engstrom et al. (2019) ont piloté une méthode ascendante au Sri Lanka qui combinait des données d'enquêtes auprès des ménages et des données satellites ; ils ont mis à l'épreuve leur nouvelle méthode en parallèle avec les méthodes existantes au regard des données de recensement. Ils ont constaté que les estimations ascendantes étaient plus précises que les estimations descendantes des modèles GPWv4 (2010 et 2015), GHSL 2014, LandScan 2010 et WorldPop 2010, bien que les calculs du HRSL 2015 et du WorldPop 2015 soient encore plus précis car ils s'appuient sur les données de recensement les plus récentes et intègrent des images satellites à plus haute résolution. Des approches novatrices et ciblées géographiquement peuvent aussi permettre d'améliorer les calculs. Par exemple, une méthode testée en Chine a montré que la prise en compte de données de « points d'intérêt » produisait des résultats plus précis que ceux du WorldPop (Ye et al., 2019). Dans une étude sur le Pakistan, le Bureau américain du recensement a constaté que sa nouvelle méthode était plus proche des valeurs du recensement que le LandScan, le GPW ou le GRUMP, et qu'elle excellait dans l'identification des grands groupes de population agricole (Azar et al., 2013). Par conséquent, les chercheurs devraient étudier les nouvelles méthodes et sources de données et les utilisateurs devraient envisager de nouveaux jeux de données qui pourraient être mieux adaptés aux conditions locales.

VARIATION DES ERREURS EN FONCTION DE LA GÉOGRAPHIE

Le fait qu'une grille de population puisse fournir des estimations à haute résolution ne signifie pas que les estimations sont plus précises au niveau concerné car l'incertitude et les erreurs augmentent si l'on étudie les résultats de près. Par exemple, lors de la validation de différentes grilles de population au regard des données de recensement au Kenya, Hay et al. (2004) ont constaté que les modèles étaient plus précis au niveau provincial, mais que la capacité de prédiction diminuait aux échelons administratifs inférieurs. Cependant, comme l'on pouvait s'y attendre, ils ont observé une forte variation de la précision prédictive lorsque les calculs étaient en dessous du niveau des données de recensement utilisées pour générer un modèle donné. En effet, la résolution des données de recensement peut souvent être un obstacle dans l'estimation de la population. D'autres travaux ont noté que l'utilisation de données auxiliaires n'améliore la précision que si elles sont plus détaillées que les données de recensement d'entrée (Linard et al., 2012). Toutefois, la relation entre la résolution et la précision peut être dynamique, comme le montrent les travaux de validation d'une nouvelle méthode d'estimation proposée aux Pays-Bas (Biljecki et al., 2016). Les chercheurs n'ont trouvé qu'une erreur prédictive de 0,5 % au niveau national, mais l'erreur est passée à 9,3 % au niveau administratif le plus élevé, puis est tombée à 3 % au niveau administratif le plus bas. Ces résultats soulignent la nécessité de continuer à tester les jeux de données démographiques à différentes résolutions géographiques. Il est légitime que les utilisateurs des données souhaitent disposer de données démographiques de résolution aussi fine que possible, mais ils doivent examiner attentivement les options disponibles pour déterminer si leur choix répond réellement à leurs besoins.

3.3 AUTRES CATÉGORIES DE DONNÉES DE VALIDATION

En dehors des données de recensement produites par les États, il existe d'autres sources de données qui peuvent être utilisées pour valider les grilles de population. Le HRSL a été validé en partie en utilisant les données de l'enquête intégrée auprès des ménages du Malawi (IHS3) (Tiecke et al., 2017). De même, les chercheurs ont utilisé les données géospatiales de l'enquête de ménage conduite à Bo City, en Sierra Leone, pour valider une nouvelle méthode ascendante d'estimation de la population (Hillson, 2014). De plus, des enquêtes micro-censitaires menées dans le nord du Nigeria ont été validées à l'aide de données de référence recueillies lors d'une campagne de vaccination contre la polio (Weber et al., 2018), ce qui permet d'envisager une validation de la grille de population. Bien que l'on ne puisse pas valider une grille de population dans son ensemble au moyen d'autres types de données, ce procédé peut renforcer la confiance que l'on peut accorder aux résultats, et les approches novatrices et les sources de données non conventionnelles peuvent ouvrir la voie à un processus de validation plus poussé.

3.4 NÉCESSITÉ D'UN TRAVAIL DE VALIDATION COMPLÉMENTAIRE

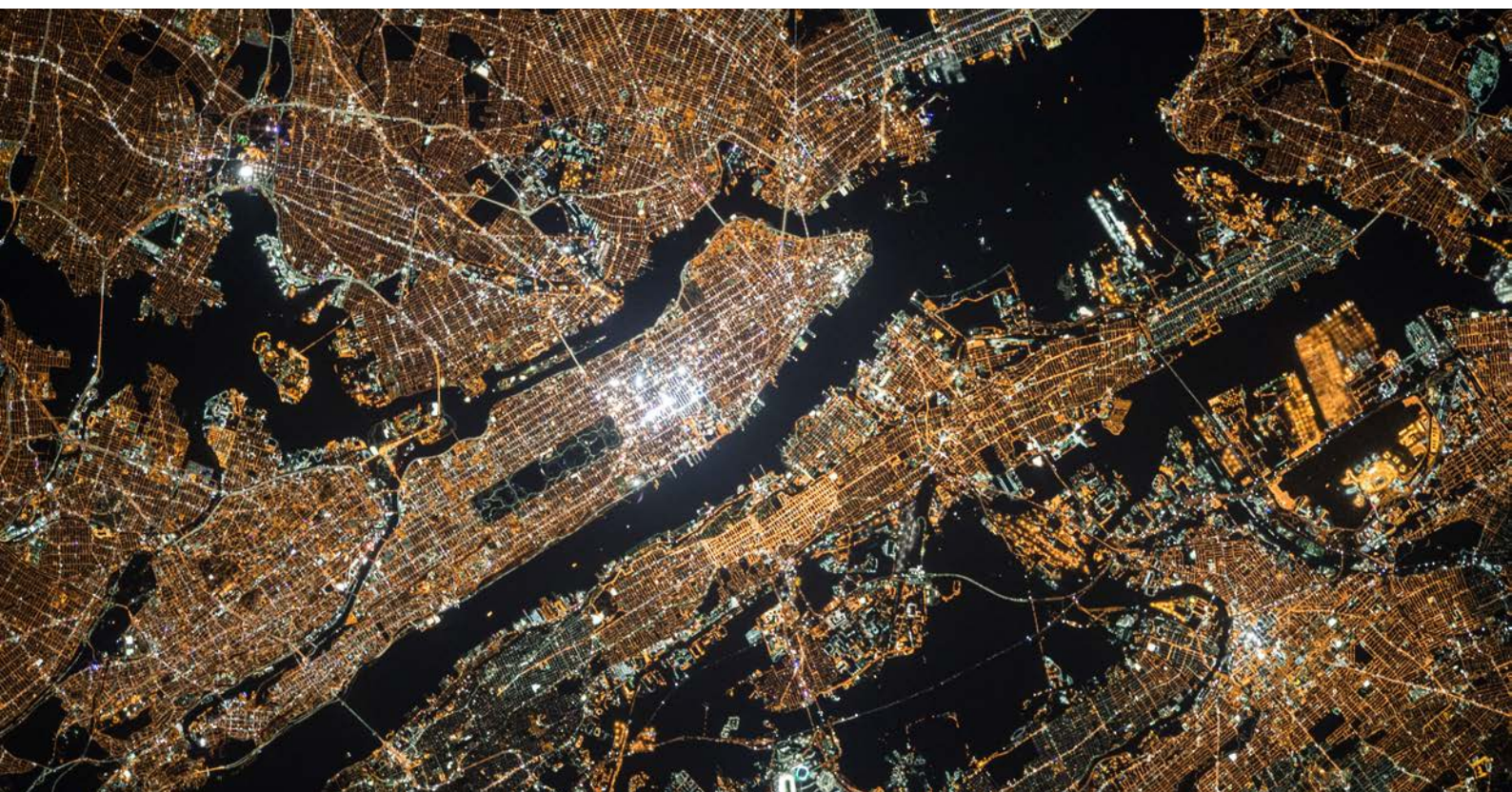
Grâce à la multiplication des comparaisons de modèles et des travaux de validation, on constate que les caractéristiques des différentes grilles de population commencent à être répertoriées, l'utilité des estimations proposées étant ainsi mise à l'épreuve. Cependant, il ne suffit pas d'avoir une connaissance parfaite des subtilités de chaque jeu de données ou de savoir comment leurs performances peuvent varier en fonction du pays et du milieu, ou bien de connaître les modèles parfaitement adaptés à des besoins spécifiques en matière de politique. Par ailleurs, les recherches menées par le passé sont encore trop limitées et inadéquates (Bai et al., 2018). Il est donc impératif de réaliser des comparaisons systématiques, de portée mondiale et qui ne visent pas que les pays séparément. Toutefois, le déficit de données nécessaires à la validation demeure une contrainte majeure. Aussi, les experts ont-ils proposé l'emploi des données de microrecensement ou des données d'enquêtes géoréférencées sur les ménages, à l'instar des travaux susmentionnés de Tiecke et al. (2017) et de Webber et al. (2018). Cette approche de validation permettrait d'éviter une comparaison de chaque cellule des jeux de données maillées et pourrait s'appuyer sur les sources de données existantes (Chen, 2019). Le POPGRID Data Collaborative est à la recherche de mesures analytiques clés permettant de valider ses différents jeux de données maillées afin de vérifier la précision des méthodes actuelles et de comprendre comment les améliorer.

Les utilisateurs des données ont toujours la possibilité de prendre des dispositions pour tenir compte du degré d'incertitudes des estimations. Comme décrit dans le présent document, il peut y avoir des écarts significatifs entre les diverses grilles de population. Ainsi les utilisateurs peuvent chercher à connaître la nature de l'incertitude des calculs en utilisant plusieurs grilles, au lieu de partir du principe qu'un jeu de données reflète nécessairement la réalité. De plus, les utilisateurs devraient tenir compte des caractéristiques géographiques lorsqu'ils appliquent une grille de population car des travaux de validation en cours peuvent contribuer à justifier la sélection d'un jeu plutôt qu'un autre pour une région donnée ou une catégorie d'occupation des sols précise. Par ailleurs, comme indiqué dans le chapitre suivant, les utilisateurs devraient prendre en compte certains facteurs dans leur choix de modèle. Une utilisation plus éclairée des grilles de population renforcera la qualité de l'analyse et de la prise de décision.

DIRECTIVES POUR LA PRISE DE DÉCISIONS

Comme nous l'avons expliqué dans les chapitres précédents, les jeux de données démographiques maillées actuelles sont tous différents. Il est donc indispensable de disposer des connaissances générales nécessaires pour analyser leurs forces et leurs faiblesses. De plus, les entretiens avec les producteurs et les utilisateurs de données nous ont montré que les attentes des différents acteurs étaient disparates. D'une part, les producteurs de données s'inquiétaient de ce que l'exploitation efficace des informations pertinentes des jeux de données démographiques maillées soit freinée par le manque de connaissances et de ressources des décideurs politiques. D'autre part, les utilisateurs de données signalaient qu'en période de crise, le besoin urgent d'extraire une donnée utile à la prise de décision avait priorité sur la compréhension générale des jeux de données (Neumann et al., 2015). Dans certains cas, ce sont les questions logistiques, les problèmes d'accès ou les facteurs politiques qui détermineront le choix du jeu de données. Bien qu'il soit impossible de surmonter tous ces obstacles, quelques considérations majeures devraient permettre de guider les différents acteurs dans leur sélection de données, en dépit de leur complexité.

Un article récent a adopté une approche « adaptée à l'emploi » dans la formulation de recommandations générales sur l'utilisation des données démographiques maillées (Leyk et al., 2019). Sur la base de cette étude technique, nous exposons neuf des principaux critères de décision permettant un processus de sélection plus ouvert et mieux informé. Par ailleurs, ces critères permettent d'identifier les jeux de données susceptibles de ne pas convenir à une application donnée et d'expliquer les raisons pour lesquelles d'autres jeux pourraient fournir des estimations plus adaptées aux besoins.



1. IMPORTANCE DE L'EMPLACEMENT DES POPULATIONS À UN MOMENT DONNÉ

Étant donné que la population effectue ses déplacements sur différentes échelles temporelles (quotidienne, hebdomadaire et saisonnière), de nombreux jeux de données démographiques se fondent sur des concepts différents pour calculer la répartition temporelle de la population. En majorité, les fichiers se fondent sur des données de recensement qui tiennent compte du lieu de résidence des personnes (population résidentielle). Le lieu de résidence est généralement associé à la population nocturne, puisque la grande majorité des personnes passent la nuit dans leur résidence principale et travaillent dans un autre endroit pendant la journée. La population diurne correspond à l'endroit où les gens se trouvent pendant la journée, tandis que la population ambiante (concept temporel utilisé par le Landscan) correspond à l'endroit où les gens sont susceptibles de se trouver tout au long de la journée (par exemple, pendant plus de 24 heures). L'analyse de l'évolution de la population à long terme est généralement basée sur la population nocturne car les recensements sont fondés sur cette population et qu'il existe peu d'estimations fiables de la population diurne (Leyk et al., 2019). Pourtant, pour certaines applications dont les interventions d'urgences, les utilisateurs peuvent avoir besoin d'informations sur la population diurne afin de mieux repérer l'emplacement réel des personnes si une catastrophe se produit pendant la journée (DeVillie et al., 2014). Les calculs des fichiers GPWv4, GHS-POP, WorldPop et WPE portent sur la population résidentielle, tandis que les calculs du LandScan font la moyenne des chiffres de la population diurne et nocturne.¹⁰ Ainsi, les utilisateurs devraient tout d'abord savoir si leurs travaux doivent se baser sur des données de population diurne ou nocturne, puis étudier les solutions correspondantes.

2. CONNAÎTRE L'APPROCHE DE MODÉLISATION UTILISÉE POUR RÉPARTIR LA POPULATION

S'il est tentant d'opter pour un fichier de données maillées de la plus haute résolution disponible, comme le HRSL, distribué à une résolution de 3 secondes d'arc (10 m), chaque fichier utilise une définition différente de ce qui constitue un établissement humain, avec les limites des capteurs satellites utilisés pour les définir. Le HRSL utilise un modèle où les surfaces bâties de cinq mètres sont produites par un algorithme d'apprentissage automatique appliqué à l'imagerie Maxar. En théorie, l'empreinte de chaque bâtiment peut être identifiée. Le GHS-POP, en revanche, utilise un produit de 30 m dérivé des fichiers du Landsat qui définit le degré d'urbanisation en termes de fractions. Ces deux fichiers de données utilisent une logique qui déplace automatiquement les personnes vers les surfaces bâties, tandis que le WorldPop utilise un algorithme qui laisse systématiquement un nombre résiduel d'habitants dans les zones non identifiées comme bâties (en supposant que les capteurs satellites n'identifient pas toujours toutes les zones peuplées). Comme d'autres facteurs déterminent si les différents fichiers sont adaptés aux besoins (par exemple, la disponibilité de séries chronologiques cohérentes, le recours à la population résidentielle ou à la population ambiante, etc.), une haute résolution ne saurait garantir leur pertinence même si le modèle avec la plus haute résolution pourrait sembler un choix intéressant. Une application au niveau local peut justifier le choix d'un produit relativement modélisé. En revanche, s'il s'agit d'un enjeu régional, comme le changement climatique, des fichiers à résolution plus grossière pourraient suffire (MacManus, 2019). Comme décrit au chapitre 3, les erreurs de prévision peuvent augmenter au fur et à mesure que la résolution augmente. La résolution des données d'entrée issues du recensement est un facteur important

¹⁰ Le CCR a développé les premières grilles démographiques continentales qui rendent compte des variations quotidiennes et mensuelles de la population en considérant la présence de résidents, de la population active des différents secteurs, des étudiants et des touristes, ainsi que les lieux de résidence et d'activité (Schiavina et al., 2020). Ces grilles ont été produites dans le cadre du projet ENACT pour tous les pays de l'UE-28 en conjuguant des données statistiques officielles au niveau régional à des données géospatiales provenant de sources de données conventionnelles et non conventionnelles pour l'année de référence 2011 (dernière série de recensements).

dans cette réflexion (le GPWv4 documente soigneusement cette question par pays), et il est important de se rappeler que la qualité du recensement peut varier considérablement d'un pays à l'autre. Enfin, si les données démographiques doivent être utilisées en complément d'autres jeux de données, telles que les données d'élévation ou de pente, les produits modélisés qui exploitent ces facteurs en tant que données d'entrée de leurs modèles ne conviendront pas (voir également le point 8 ci-dessous).

3. IMPORTANCE DU REGROUPEMENT DÉMOGRAPHIQUE, Y COMPRIS PAR ÂGE ET PAR SEXE

Certaines catégories démographiques peuvent être particulièrement vulnérables à divers aléas socio-économiques et environnementaux. Par exemple, les catastrophes ont généralement des retombées socio-économiques plus importantes sur les femmes, les enfants et les personnes âgées, et ces groupes sont particulièrement exposés à des problèmes de santé, à l'insécurité alimentaire et à la pauvreté (McGlade et al., 2019). Les données démographiques ventilées par âge et par sexe peuvent être utilisées pour évaluer s'il existe des proportions exceptionnellement élevées ou faibles de groupes vulnérables dans des zones particulières afin d'aider à orienter et à améliorer les interventions. À l'heure actuelle, le GPWv4 et le WorldPop sont les deux seuls jeux de données démographiques maillées à l'échelle mondiale qui fournissent des estimations de population maillées par sexe et par âge. De plus, le LandScan inclut des données démographiques par âge et par sexe sur 5 ans dans l'unité administrative de niveau 1. Dans le cas du GPWv4, seuls 20 % des pays disposent de données par âge et par sexe pour l'ensemble du pays, y compris de nombreux petits États insulaires. Basées sur des méthodes d'apprentissage automatique, des estimations à haute résolution par âge et par sexe ont récemment été publiées pour être exploitées par le HRSL, mais elles doivent être considérées comme expérimentales.

4. LES ÉTUDES PAR PAYS SONT DÉLICATES

L'évolution de la population au niveau mondial dépend du nombre annuel de naissances et de décès. Cependant, l'évolution de la population à l'échelle nationale est déterminée non seulement par les naissances et les décès mais aussi par les migrations. Pour comprendre la dynamique de la population dans les différents pays et régions, les utilisateurs de données ont tout intérêt à utiliser des jeux de données ajustés en fonction d'un ensemble d'estimations démographiques cohérentes à l'échelle mondiale. Par exemple, les grilles de population mondiale du GPWv4 sont basées sur les totaux et les grilles de population du recensement initial de chaque pays et sont ajustées aux estimations par pays du rapport « Perspectives de la population dans le monde » des Nations unies, de façon à ce que la somme de toutes les cellules de grille corresponde aux estimations des Nations unies. Il est important de rappeler qu'en se basant sur le recensement de départ, les populations de migrants, de réfugiés et de personnes déplacées à l'intérieur de leur pays ne sont pas toujours représentées. De même, le LandScan ajuste les estimations infranationales en fonction des totaux par pays des programmes internationaux du Bureau du recensement des États-Unis pour une année donnée. Pour les analyses transnationales, il est préférable d'utiliser les jeux de données ajustés.

5. IDENTIFIER LES PÉRIODES DE TEMPS PERTINENTES

Le taux d'accroissement de la population a un impact considérable sur la production alimentaire, la gestion de l'eau, la mise en valeur des terres et la politique climatique. Un accroissement rapide de la population, en particulier dans les centres urbains, exige des politiques éclairées et des moyens suffisants pour améliorer la qualité de vie des populations et protéger l'environnement. Les utilisateurs ont généralement besoin des données les plus récentes qui correspondent à une finalité précise et il est de leur intérêt d'examiner les tendances récentes ou les projections pour l'avenir. En dépit du fait que certains jeux de données démographiques maillées fournissent des estimations chaque année, ces estimations ne sauraient être automatiquement utilisées comme série chronologique cohérente à une résolution fine. Par exemple, le WPE et le LandScan utilisent des méthodes différentes chaque année et ne sont généralement pas indiqués pour les comparaisons dans le temps. D'autres jeux de données fournissent des estimations à des intervalles de 5 ans (ou autres) qui sont conçues spécifiquement pour être cohérentes sur ces périodes. Le GPWv4, par exemple, utilise les recensements de 2010 comme référence et comprend une simple projection jusqu'en 2020. Dans la mesure du possible, les utilisateurs devraient vérifier l'origine des données de recensement issues du pays concerné afin d'évaluer la qualité des estimations.

6. PRENDRE EN COMPTE LES DIFFÉRENCES ENTRE LES ESTIMATIONS URBAINES ET RURALES

Dans leur majorité, les jeux de données démographiques actuels, à l'échelle mondiale, se focalisent sur la caractérisation des tendances et de la croissance des établissements urbains. En effet, la majorité de la population réside dans les zones d'établissements urbains dont les données de recensement sont généralement publiées avec une résolution plus fine et les infrastructures sont plus faciles à détecter. De nombreuses zones rurales, comme les unités administratives de recensement, ayant une superficie assez importante, les villages et les infrastructures rurales peuvent être plus difficiles à repérer à l'aide de données de télédétection en raison de la végétation, du terrain ou des types et matériaux de construction (Leyk et al., 2019). Pour les utilisateurs qui se focalisent principalement sur la densité ou la répartition de la population urbaine, les jeux de données comportant des variables relatives à l'établissement humain ou à l'extension urbaine, dont le GHS-POP, le LandScan et le WorldPop, sont sans doute plus adaptés pour évaluer la répartition de la population urbaine. Pour les travaux qui nécessitent une représentation plus détaillée de la répartition de la population rurale, le HRSL constitue une solution intéressante, s'il est disponible pour le ou les pays concernés, car il utilise des données de télédétection à haute résolution qui détectent de manière satisfaisante les bâtiments dans les régions rurales.

7. LES FACTEURS ENVIRONNEMENTAUX ONT UNE INCIDENCE SUR LA RÉPARTITION DE LA POPULATION

Les utilisateurs devraient connaître le milieu particulier pour lequel ils cherchent à obtenir une estimation de la population. Les déserts, les zones à forte déclivité ou à végétation dense et les régions à forte couverture nuageuse présentent des défis particuliers. En effet, dans ces cas-ci, les images de télédétection utilisables sont peu nombreuses et la détection des bâtiments sur ces images peut être difficile. Même les données de recensement sous-jacentes peuvent

être peu fiables dans certains cas, en particulier dans le cas des populations nomades, qui pratiquent l'agriculture de subsistance ou qui sont très éloignées. Dans certains cas, les jeux de données sont fondés sur l'hypothèse que personne ne vit dans des plans d'eau ou des zones protégées, mais cela peut conduire à sous-estimer les populations qui peuvent être présentes dans ces zones à certaines périodes de l'année. Dans sa dernière version (R2019), le SGH-POP, qui ne dévie pas de ces hypothèses, a néanmoins revu en profondeur les zones considérées inhabitées dans ses données d'entrée de recensement (Freire et al., 2018).

8. ENVISAGER LES COVARIABLES

Les utilisateurs devraient tenir compte des covariables utilisées dans chaque jeu de données. Les covariables sont utilisées pour répartir les données de population entre les unités administratives de manière plus précise. La corrélation entre une covariable et la densité de population a une incidence sur les estimations de la population. Par exemple, le LandScan, le WorldPop et le WPE utilisent les routes comme covariables dans leur modélisation. Si les utilisateurs souhaitent analyser le rapport entre les routes et la densité de population, les jeux de données qui utilisent les routes dans leur modèle démographique ne seront sans doute pas le meilleur choix, car ils sont élaborés à partir d'une hypothèse préétablie sur la façon dont les populations se regroupent à proximité des routes. Il est particulièrement important de comprendre cela dans le contexte des ODD qui évaluent l'accès aux services et aux installations exploités comme covariables pour estimer les répartitions de la population maillée. Il est important de noter que le nombre de covariables utilisées pour produire des estimations de population varie selon les jeux de données et affecte le degré de modélisation utilisé pour produire les estimations. Le WorldPop et le LandScan utilisent un plus grand nombre de covariables dans leur modélisation, tandis que le GHS-POP et le HRSL utilisent moins de covariables et sont faiblement modélisés.

9. SE RENSEIGNER SUR LES COÛTS ET DROITS DE RÉUTILISATION

Les organismes publics et à but non lucratif ne disposent pas toujours des structures et des moyens nécessaires pour financer l'accès des décideurs politiques et autres utilisateurs aux données. Dans certains cas, les utilisateurs peuvent avoir besoin de données pour des analyses et des prises de décision internes et disposent des moyens nécessaires pour acheter les données qui ne sont pas en accès libre. Cependant, dans de nombreux cas, les droits de réutilisation et les coûts sont un facteur important à prendre en compte, notamment lorsque les utilisateurs souhaitent transmettre des données à des partenaires ou à des utilisateurs finaux. Un certain nombre de jeux de données sont entièrement libres d'accès et publics, à des fins d'utilisation et de réutilisation commerciale et non commerciale (généralement désignés par la licence internationale Creative Commons Attribution 4.0). Certains jeux de données peuvent limiter l'utilisation commerciale, même dans le cas où une entité commerciale contribue à un projet d'intérêt public. Il convient de noter que le concept d'« utilisation équitable » diffère d'un pays à l'autre, si bien que les utilisateurs devront connaître les lois applicables en matière de bases de données et de droits d'auteur dans leur propre juridiction ou à quelle finalité se destinent les données qui y sont acheminées ou utilisées. Par exemple, le GPW4, le GHS-POP, le GRUMP et le WorldPop donnent accès à leurs données gratuitement, et le WPE est disponible gratuitement pour les personnes qui ont un compte d'utilisateur nominatif dans les organisations de l'ArcGIS. Le LandScan est gratuit et en accès libre uniquement à des fins humanitaires, de recherche universitaire et pour le gouvernement fédéral américain et ses partenaires de mission. Les entités commerciales doivent obtenir une licence payante pour l'utiliser. Cependant, le jeu de données LandScan USA est une ressource gratuite et en accès libre.

Ce chapitre propose des critères de décision afin d'orienter les utilisateurs dans leur choix du jeu de données le mieux adapté à la finalité recherchée. Toutefois, les directives susceptibles d'orienter ce choix ne se limitent pas aux neuf critères mentionnés ci-dessus. Nous encourageons les utilisateurs à évaluer également les facteurs contextuels (socio-économiques, environnementaux, etc.) relatifs au champ d'application lors de la sélection d'un jeu de données. Par exemple, lors de l'élaboration d'indicateurs pour les ODD, il peut être souhaitable qu'un pays se limite à un seul jeu de données démographiques maillées qui servira de base commune pour les indicateurs axés sur l'accès aux services, ce qui permettra d'harmoniser les indicateurs entre les objectifs et les secteurs.

ENCADRÉ 5 | MESSAGES POUR ÉVITER LES MÉPRISES

TReNDS a interrogé les fournisseurs de données du POPGRID présentés dans ce rapport afin de comprendre les messages qu'ils souhaitent transmettre aux utilisateurs en vue d'encourager une approche « adaptée à l'emploi » concernant les données démographiques maillées. Les fournisseurs de données souhaitent communiquer les points suivants afin d'éviter toute méprise sur les estimations démographiques et les jeux de données :

- ▶ Les données démographiques maillées sont une désagrégation ou une redistribution des données de recensement dans des modèles descendants, et non une solution de substitution aux données de recensement.
- ▶ De nombreux jeux de données démographiques maillées utilisent les mêmes données de recensement d'entrée, mais ne sont pas homogènes. Les fournisseurs de données emploient des méthodes et des hypothèses différentes qui produisent des résultats différents.
- ▶ Une meilleure communication entre les fournisseurs de données et les utilisateurs permettra aux fournisseurs de données de recommander des jeux de données adaptés à un champ d'application précis.
- ▶ La publication d'une documentation complète sur un jeu de données peut encourager la transparence et garantir que les utilisateurs choisiront le jeu de données le plus adapté à la finalité recherchée.
- ▶ Le [POPGRID Viewer](#) est un outil interactif qui permet aux utilisateurs de comparer visuellement les résultats de différents jeux de données afin d'identifier et de comprendre leurs différences.

CONCLUSIONS

Pour relever les plus grands défis de l'humanité et atteindre nos objectifs communs, il est vital d'avoir une connaissance précise de l'emplacement des populations. Les données démographiques maillées constituent un outil important qui permet d'approfondir ces connaissances. Dans les chapitres précédents, nous avons présenté cette catégorie de données et indiqué un certain nombre de critères que les utilisateurs doivent garder à l'esprit. Actuellement, les données démographiques maillées viennent appuyer les opérations de secours du PAM dans les cas de catastrophe, facilitent la réponse en cas d'épidémie, fournissent de nouveaux étalons de mesure du développement à la Banque mondiale et contribuent à des travaux de recherche scientifique déterminants. Cependant, selon une enquête menée par les bureaux nationaux des statistiques, de nombreux décideurs qui pourraient bénéficier de cette ressource ne la connaissent pas encore. Nous espérons que cet aperçu des jeux de données disponibles et de leurs particularités permettra de mieux les faire connaître et de multiplier leur champs d'applications.

Bien que les sept jeux de données ici décrits mesurent des éléments similaires, ils divergent sur de nombreux points importants. Il convient que les utilisateurs comprennent comment un jeu de données est élaboré, en particulier les types de données d'entrée et de données auxiliaires, la méthode de modélisation sous-jacente et la manière dont les données sont ajustées pour correspondre aux chiffres de la population nationale. Les caractéristiques des jeux de données obtenus sont également différentes, notamment la résolution spatiale, le type de désagrégation de la population disponible, les années représentées et la cohérence des méthodes entre les versions. Ces différences ne se limitent pas aux caractéristiques générales des jeux de données. Des études ont montré qu'il peut y avoir des écarts significatifs dans les chiffres de population entre les jeux de données. Par ailleurs, l'incertitude des modèles démographiques influence les prévisions qui en découlent, telles que l'exposition aux maladies et la vulnérabilité au changement climatique. Les utilisateurs de données doivent orienter leur choix de modèle de manière délibérée et faire preuve de transparence quant aux incertitudes qui y sont associées.

Alors que les premiers efforts de validation des données de population maillées fournissent déjà des indications précieuses, il semble que certaines méthodes soient en mesure de produire des calculs réellement fiables. Les utilisateurs doivent cependant être conscients des erreurs systématiques des produits disponibles car la qualité des estimations peut varier en fonction de la géographie ou du type d'établissement humain. Les études de validation soulignent également que ce n'est pas parce qu'un modèle fournit des estimations à une résolution plus élevée qu'il est plus précis. La communauté POPGRID reconnaît la nécessité d'effectuer des recherches de validation complémentaires qui permettraient de tester les performances des jeux de données dans une variété de contextes et de s'assurer que les outils sont utilisés de manière adéquate. La nouvelle série de recensements mondiaux actuellement en cours ouvre la voie à des possibilités uniques d'accéder aux chiffres de population actuels et d'effectuer des études de validation détaillées. Les multiples sources de données des enquêtes auprès des ménages pourraient également fournir des informations utiles. Les travaux de validation nécessiteront une planification et un financement stratégiques, et la communauté internationale devrait s'engager dans une stratégie visant à exploiter pleinement le potentiel des données démographiques maillées.

Si l'investissement dans les données démographiques maillées est assurément nécessaire, il ne doit pas remplacer l'amélioration des recensements. L'une des étapes les plus importantes pour renforcer les estimations démographiques maillées est l'amélioration de la précision, de la

résolution et de l'exhaustivité des données d'entrée. Des données de recensement de qualité sont essentielles pour les nombreux jeux de données de type descendant dont nous avons dressé le profil. Bien que les nouvelles techniques de données démographiques maillées ascendantes puissent être utiles dans les domaines où les données de recensement sont manquantes ou caduques, elles ne sauraient les remplacer.¹¹

En effet, les données de recensement sous-jacentes peuvent être le principal facteur contraignant lors de l'élaboration d'une estimation descendante de la population. Ce problème est particulièrement marqué dans les milieux à faibles revenus où les données à haute résolution sont souvent indisponibles, dans les zones de conflit ou dans les zones en pleine mutation. Les États et le système des Nations unies dans son ensemble doivent veiller à maintenir la fiabilité des recensements et autres sources de données, même si de nouvelles sources de données continuent d'apparaître.

Indépendamment des questions liées à un éventuel investissement, les utilisateurs de données doivent répondre à des questions essentielles pour évaluer un modèle démographique en fonction de critères précis avant d'en retenir un. Veulent-ils calculer la population pendant la journée ou pendant la nuit ? S'intéressent-ils uniquement au nombre total d'habitants ou à la répartition démographique ? Veulent-ils étudier les caractéristiques de la population en fonction des pays ou selon une échelle temporelle ? S'intéressent-ils aux populations des zones urbaines ou rurales ? Leurs réponses sont susceptibles de les orienter vers un choix logique et adapté à leurs besoins. Un travail de validation complémentaire permettra de préciser d'autres subtilités à prendre en compte lors du choix et de l'évaluation des jeux de données. Par exemple, il serait utile de savoir si les modèles sont performants dans les régions côtières, dans les zones forestières ou dans d'autres endroits difficiles à caractériser. Le POPGRID Data Collaborative s'est engagé à étudier ces questions et à promouvoir un dialogue sur les données démographiques maillées. A mesure que nos connaissances évoluent, des facteurs pertinents pourraient être présentés dans un arbre de décision interactif conçu pour aider les utilisateurs et leur fournir des conseils personnalisés. Les travaux de recherche et la sensibilisation croissante des utilisateurs permettront aux données démographiques de contribuer non seulement à documenter la qualité de vie dans le monde entier mais aussi à l'améliorer.



¹¹ Des initiatives sont en cours pour cartographier les millions de réfugiés et de populations déplacées, dits les « millions disparus » grâce à la télédétection par satellite ; de plus, la Matrice de suivi des déplacements de l'Organisation internationale pour les migrations (OIM-DTM) a réalisé des progrès dans la géolocalisation des populations déplacées (Van Den Hoek, 2018).

RÉFÉRENCES

- Abushady, A. (2020, 11 octobre). Entretien avec H. Dahmm et M. Rabiee.
- Andrade-Pacheco, R., Savory, D. J., Midekisa, A., Gething, P. W., Sturrock, H. J., & Bennett, A. (2019). Household Electricity Access in Africa (2000–2013): Closing Information Gaps with Model-based Geostatistics. *PloS One*, 14(5). [doi:10.1371/journal.pone.0214635](https://doi.org/10.1371/journal.pone.0214635)
- Aubrecht, C., Gunasekera, R., Ungar, J., & Ishizawa, O. (2016). Consistent Yet Adaptive Global Geospatial Identification of Urban–Rural Patterns: The iURBAN model. *Remote Sensing of Environment*, 187, 230–240. [doi:10.1016/j.rse.2016.10.031](https://doi.org/10.1016/j.rse.2016.10.031)
- Azar, D., Engstrom, R., Graesser, J., & Comenetz, J. (2013). Generation of Fine-Scale Population Layers Using Multi-resolution Satellite Imagery and Geospatial Data. *Remote Sensing of Environment*, 130, 219–232. [doi:10.1016/j.rse.2012.11.022](https://doi.org/10.1016/j.rse.2012.11.022)
- Backer, L., & Holt Bloch, V. V. (2011). GEOSTAT 1A—Representing Census Data in a European Population Grid. Informe final. Extrait de <http://www.efgs.info/wp-content/uploads/geostat/1a/GEOSTAT1A-final-report.pdf>
- Bai, Z., Wang, J., Wang, M., Gao, M., & Sun, J. (2018). Accuracy Assessment of Multi-Source Gridded Population Distribution Datasets in China. *Sustainability*, 10(5), 1363. [doi:10.3390/su10051363](https://doi.org/10.3390/su10051363)
- Banque mondiale. (2016). SDG Indicator 9.1.1. Troisième réunion de IAEG-SDG, Mexico. Extrait de : <https://unstats.un.org/sdgs/files/meetings/iaeg-sdgs-meeting-03/3rd-IAEG-SDGs-presentation-WorldBank--9.1.1.pdf>
- Baylor College of Medicine. (s. f.) Introduction to Infectious Diseases. Extrait de <https://www.bcm.edu/departments/molecular-virology-and-microbiology/emerging-infections-and-biodefense/introduction-to-infectious-diseases>
- Bhaduri, B. (2018). Comunicación personal con R. Chen y A. de Sherbinin.
- Biljecki, F., Ohori, K. A., Ledoux, H., Peters, R., & Stoter, J. (2016). Population Estimation Using a 3D City Model: A Multi-scale Country-wide Study in the Netherlands. *PloS One*, 11(6). [doi:10.1371/journal.pone.0156808](https://doi.org/10.1371/journal.pone.0156808)
- Bird, J. H., Lebrand, M. S. M., & Venables, A. J. (2019). The Belt and Road Initiative: Reshaping Economic Geography in Central Asia. The World Bank. Extrait de <http://documents.worldbank.org/curated/en/836301554729486900/The-Belt-and-Road-Initiative-Reshaping-Economic-Geography-in-Central-Asia>
- Blankespoor, B., Dasgupta, S., & Lange, G. (2016). Mangroves as a Protection from Storm Surges in a Changing Climate. *Ambio*, 46(4), 478–491. [doi:10.1007/s13280-016-0838-x](https://doi.org/10.1007/s13280-016-0838-x)
- Butler, D. (2011). Reactors, Residents and Risk. *Nature*, 474, 36. [doi: 10.1038/472400a](https://doi.org/10.1038/472400a)
- Cairney, P. (2016). The Politics of Evidence-based Policy Making. Palgrave Springer.
- Calka, B., & Bielecka, E. (2019). Reliability Analysis of LandScan Gridded Population Data. The Case Study of Poland. *ISPRS International Journal of Geo-Information*, 8(5), 222. [doi:10.3390/ijgi8050222](https://doi.org/10.3390/ijgi8050222)
- Canavire-Bacarreza, G., Martinez-Vazquez, J., & Yedgenov, B. (2020). Identifying and Disentangling the Impact of Fiscal Decentralization on Economic Growth. *World Development*, 127, 104742. [doi:10.1016/j.worlddev.2019.104742](https://doi.org/10.1016/j.worlddev.2019.104742)
- Chen, R. (2019, 19 novembre). Entretien avec H. Dahmm et M. Rabiee.
- CIESIN & CUNY CIDR. (2020 forthcoming). Low Elevation Coastal Zone (LECZ) Urban-Rural Population and Land Area Estimates, Version 3. NASA Socioeconomic Data and Applications Center (SEDAC) [Dataset].

Cori, A., Donnelly, C. A., Dorigatti, I., Ferguson, N. M., Fraser, C., Garske, T., & Van Kerkhove, M. D. (2017). Key Data for Outbreak Evaluation: Building on the Ebola Experience. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 372 (1721), 20160371. [doi:10.1098/rstb.2016.0371](https://doi.org/10.1098/rstb.2016.0371)

Cui, R., Hultman, N., Jiang, K., McJeon, H., Yu, S., Cui, D., & Clarke, L. (2020). A High Ambition Coal Phaseout in China: Feasible Strategies Through a Comprehensive Plant-by-Plant Assessment. Center for Global Sustainability. Extrait de https://cgs.umd.edu/sites/default/files/2020-01/AHighAmbitionCoalPhaseoutInChina_EN_Full.pdf

Da Costa, J. N., Bielecka, E., & Calka, B. (2017). Uncertainty Quantification of the Global Rural-Urban Mapping Project over Polish Census Data. *Proceedings of 10th International Conference "Environmental Engineering"* (Vol. 10, pp. 1-7). [doi:10.3846/enviro.2017.221](https://doi.org/10.3846/enviro.2017.221)

Dahmm, H. (2018). Valuation of the New Zealand Census. SDSN TREnds. <https://www.sdsntrends.org/research/2018/11/16/case-study-new-zealand-census>

Dasgupta, S., Laplante, B., Murray, S., & Wheeler, D. (2011). Exposure of Developing Countries to Sea-Level Rise and Storm Surges. *Climatic Change*, 106(4), 567-579. [doi:10.1007/s10584-010-9959-6](https://doi.org/10.1007/s10584-010-9959-6)

Déville, P., Linard, C., Martin, S., Gilbert, M., Stevens, F. R., Gaughan, A. E., & Tatem, A. J. (2014). Dynamic Population Mapping Using Mobile Phone Data. *Proceedings of the National Academy of Sciences*, 111(45), 15888-15893. [doi:10.1073/pnas.1408439111](https://doi.org/10.1073/pnas.1408439111)

Doocy, S., Gorokhovich, Y., Burnham, G., Balk, D., & Robinson, C. (2007). Tsunami Mortality Estimates and Vulnerability Mapping in Aceh, Indonesia. *American Journal of Public Health*, 97(Supplement_1), S146-S151. [doi:10.2105/AJPH.2006.095240](https://doi.org/10.2105/AJPH.2006.095240)

Doxsey-Whitfield, E., MacManus, K., Adamo, S. B., Pistolesi, L., Squires, J., Borkovska, O., & Baptista, S. R. (2015). Taking Advantage of the Improved Availability of Census Data: A First Look at the Gridded Population of the World, Version 4. *Papers in Applied Geography*, 1(3), 226-234. [doi:10.1080/23754931.2015.1014272](https://doi.org/10.1080/23754931.2015.1014272)

DPNU. (2017). World Population Prospects The 2017 Revision: Methodology of the United Nations Population Estimates and Projections. Division de la population des Nations unies (DPNU). Extrait de : https://population.un.org/wpp/Publications/Files/WPP2017_Methodology.pdf

DPNU. (2019). World Population Prospects Highlights. Division de la population des Nations unies (DPNU). Extrait de : https://population.un.org/wpp/Publications/Files/WPP2019_Highlights.pdf

DPNU. (2020). A Recommendation on the Methodology to Delineate Cities Urban and Rural Areas. Items for Consideration and Decision: Demographic Statistics. Commission de statistique des Nations unies (UNSC). Extrait de <https://unstats.un.org/unsd/statcom/51st-session/documents/BG-Item3j-Recommendation-E.pdf>

DSNU. (2017a). Indicator 11.6.2: Annual Mean Levels of Fine Particulate Matter (e.g. PM2.5 and PM10) in Cities (Population Weighted). Division de statistique des Nations Unies (DSNU). Extrait de : <https://unstats.un.org/sdgs/metadata/files/Metadata-11-06-02.pdf>

DSNU. (2017b). Division de statistique des Nations Unies. Principles and Recommendations for Population and Housing Censuses. Division de statistique des Nations Unies (DSNU). Extrait de : https://unstats.un.org/unsd/demographic-social/Standards-and-Methods/files/Principles_and_Recommendations/Population-and-Housing-Censuses/Series_M67rev3-E.pdf

Dwyer-Lindgren, L., Cork, M. A., Sligar, A., Steuben, K. M., Wilson, K. F., Provost, N. R., & Biehl, M. H. (2019). Mapping HIV Prevalence in Sub-Saharan Africa Between 2000 and 2017. *Nature*, 570(7760), 189-193. [doi:10.1038/s41586-019-1200-9](https://doi.org/10.1038/s41586-019-1200-9)

Eales, A., Alsop, A., Frame, D., Strachan, S., & Galloway, S. (2020). Assessing the Market for Solar Photovoltaic (PV) Microgrids in Malawi. *Hapres Journal of Sustainability Research*, 2(1). [doi:10.20900/jsr20200008](https://doi.org/10.20900/jsr20200008)

Ehrlich, D., Melchiorri, M., Florczyk, A. J., Pesaresi, M., Kemper, T., Corbane, C., & Siragusa, A. (2018). Remote Sensing Derived Built-up Area and Population Density to Quantify Global Exposure to Five Natural Hazards Over Time. *Remote Sensing*, 10(9), 1378. [doi:10.3390/rs10091378](https://doi.org/10.3390/rs10091378)

Eicher, C. L., & Brewer, C. A. (2001). Dasymetric Mapping and Areal Interpolation: Implementation and Evaluation. *Cartography and Geographic Information Science*, 28(2), 125-138.
[doi:10.1559/152304001782173727](https://doi.org/10.1559/152304001782173727)

Engstrom, R., Newhouse, D. L., & Soundararajan, V. (2019). Estimating Small Area Population Density Using Survey Data and Satellite Imagery: An Application to Sri Lanka. The World Bank. Extrait de <http://documents.worldbank.org/curated/en/256241552483977593/Estimating-Small-Area-Population-Density-Using-Survey-Data-and-Satellite-Imagery-An-Application-to-Sri-Lanka>

Esri. (n.d.) Measuring in Arc-Seconds. Extrait de <https://www.esri.com/news/arcuser/0400/wdside.html> Accessed 16 April 2020

European Commission Joint Research Centre. (2018). Atlas of the Human Planet 2018. EU Science Hub. Extrait de <https://ec.europa.eu/jrc/en/publication/eur-scientific-and-technical-research-reports/atlas-human-planet-2018>

FNUAP. (n.d.). Impact of COVID-19 on Census. Fonds des Nations Unies pour la population (FNUAP). Extrait de : https://arcgis-pilot.unfpa.org/open_data/census_cal/census_delay_list.php?pagesize=-1 Consulté le 16 avril 2020

FNUAP (2017). New Methodology: A Hybrid Census to Generate Spatially-Disaggregated Population Estimates. Fonds des Nations Unies pour la population (FNUAP). Extrait de : <https://www.unfpa.org/resources/new-methodology-hybrid-census-generate-spatially-disaggregated-population-estimates>

Freire, S., Schiavina, M., Florczyk, A. J., MacManus, K., Pesaresi, M., Corbane, C., & Sliuzas, R. (2018). Enhanced Data and Methods for Improving Open and Free Global Population Grids: Putting 'Leaving No One Behind' Into Practice. *International Journal of Digital Earth*, 13(1), 61-77. [doi:10.1080/17538947.2018.1548656](https://doi.org/10.1080/17538947.2018.1548656)

Frye, C. (2017). Esri's 2016 World Population Estimate Methodology. Esri. Extrait de https://www.esri.com/content/dam/esrisites/en-us/media/whitepaper/Esri_2016_WPE_Methodology_WhitePaper.pdf

Garb, J. L., Cromley, R. G., & Wait, R. B. (2007). Estimating Populations At Risk for Disaster Preparedness and Response. *Journal of Homeland Security and Emergency Management*, 4(1).
[doi:10.2202/1547-7355.1280](https://doi.org/10.2202/1547-7355.1280)

Gao, J. (2017). Downscaling Global Spatial Population Projections From 1/8-Degree to 1-km Grid Cells. Technical Notes NCAR, National Center for Atmospheric Research, Boulder, CO., USA. Extrait de <https://opensky.ucar.edu/islandora/object/technotes%3A553/datastream/PDF/view>

Gaughan, A. E., Stevens, F. R., Linard, C., Jia, P., & Tatem, A. J. (2013). High Resolution Population Distribution Maps for Southeast Asia in 2010 and 2015. *PloS One*, 8(2). [doi:10.1371/journal.pone.0055882](https://doi.org/10.1371/journal.pone.0055882)

Gaughan, A. (2019, September 24). Entretien avec H. Dahmm et M. Rabiee Gunasekera, R., Ishizawa, O., Aubrecht, C., Blankespoor, B., Murray, S., Pomonis, A., & Daniell, J. (2015). Developing an Adaptive Global Exposure Model to Support the Generation of Country Disaster Risk Profiles. *Earth-Science Reviews*, 150, 594-608.
[doi:10.1016/j.earscirev.2015.08.012](https://doi.org/10.1016/j.earscirev.2015.08.012)

Hall, O., Stroh, E., & Paya, F. (2012). From Census to Grids: Comparing Gridded Population of the World with Swedish Census Records. *The Open Geography Journal*, 5(1). Extrait de <https://pdfs.semanticscholar.org/aa54/496eaa46a532ba3c4353843c4a56b7712db3.pdf>

Hallisey, E., Tai, E., Berens, A., Wilt, G., Peipins, L., Lewis, B., Graham, S., Flanagan, B., Lunsford, N. (2017). Transforming Geographic Scale: A Comparison of Combined Population and Areal Weighting to Other Interpolation Methods. *International Journal of Health Geographics*, 16:29.
[doi: 10.1186/s12942-0170102-z](https://doi.org/10.1186/s12942-0170102-z)

Hay, S. I., Guerra, C. A., Tatem, A. J., Noor, A. M., & Snow, R. W. (2004). The Global Distribution and Population at Risk of Malaria: Past, Present, and Future. *The Lancet Infectious Diseases*, 4(6), 327-336. [doi:10.1016/S1473-3099\(04\)01043-6](https://doi.org/10.1016/S1473-3099(04)01043-6)

Hay, S. I., Noor, A. M., Nelson, A., & Tatem, A. J. (2005). The Accuracy of Human Population Maps for Public Health Application. *Tropical Medicine and International Health*, 10(10), 1073-1086.
[doi:10.1111/j.1365-3156.2005.01487.x](https://doi.org/10.1111/j.1365-3156.2005.01487.x)

Hillson, R., Alexandre, J. D., Jacobsen, K. H., Ansumana, R., Bockarie, A. S., Bangura, U., ... & Stenger, D. A. (2014). Methods for Determining the Uncertainty of Population Estimates Derived from Satellite Imagery and Limited Survey Data: A Case Study of Bo City, Sierra Leone. *PloS one*, 9(11). [doi:10.1371/journal.pone.0112241](https://doi.org/10.1371/journal.pone.0112241)

IDMC. (2019). Global Report on Internal Displacement 2019. Internal Displacement Monitoring Centre. Extrait de <https://www.internal-displacement.org/global-report/grid2019/>

IEAG. (2014). A World That Counts: Mobilising the Data Revolution for Sustainable Development. United Nations Secretary-General's Independent Expert Advisory Group on a Data Revolution for Sustainable Development (IEAG). Extrait de <https://www.undatarevolution.org/wp-content/uploads/2014/12/A-World-That-Counts2.pdf>

Imi, A., Ahmed, F., Anderson, E. C., Diehl, A. S., Maiyo, L., Peralta-Quirós, T., & Rao, K. S. (2016). New Rural Access Index: Main Determinants and Correlation to Poverty. The World Bank. Retrieved from <https://openknowledge.worldbank.org/handle/10986/25676>

James, W. H., Tejedor-Garavito, N., Hanspal, S. E., Campbell-Sutton, A., Hornby, G. M., Pezzulo, C., &

Kerr, D. (2018). Gridded Birth and Pregnancy Datasets for Africa, Latin America and the Caribbean. *Nature*, 5, 180090. [doi:10.1038/sdata.2018.90](https://doi.org/10.1038/sdata.2018.90)

Johnson, A. M. (2019). Liberia: US\$18.9 Million Census Set for March 2021. *FrontPageAfrica*. Extrait de <https://frontpageafricaonline.com/news/liberia-us18-9-million-census-set-for-march-2021/>

Jones, B., & O'Neill, B. C. (2020). Global One-Eighth Degree Population Base Year and Projection Grids Based on the Shared Socioeconomic Pathways, Revision 01. NASA Socioeconomic Data and Applications Center (SEDAC) [Dataset]. [doi:10.7927/m30p-j498](https://doi.org/10.7927/m30p-j498). Accessed 16 April 2020

Kiguchi, M., Shen, Y., Kanae, S., & Oki, T. (2014). Re-evaluation of Future Water Stress Due to Socio-Economic and Climate Factors Under a Warming Climate. *Hydrological Sciences Journal*, 60(1), 14-29. [doi:10.1080/02626667.2014.888067](https://doi.org/10.1080/02626667.2014.888067)

Kulp, S. A., & Strauss, B. H. (2019). New Elevation Data Triple Estimates of Global Vulnerability to Sea-Level Rise and Coastal Flooding. *Nature Communications*, 10(1), 1-12. [doi:10.1038/s41467-019-12808-z](https://doi.org/10.1038/s41467-019-12808-z)

Lai, S., Bogoch, I. I., Ruktanonchai, N., Watts, A. G., Li, Y., Yu, J., & Li, Z. (2020a). Assessing Spread Risk of Wuhan Novel Coronavirus Within and Beyond China, January-April 2020: A Travel Network-Based Modelling Study. *Health Sciences*. [doi: 10.1101/2020.02.04.20020479](https://doi.org/10.1101/2020.02.04.20020479)

Lai, S., Ruktanonchai, N. W., Zhou, L., Prosper, O., Luo, W., Floyd, J. R., & Tatem, A. J. (2020b). Effect of Non-Pharmaceutical Interventions for Containing the COVID-19 Outbreak: An Observational and Modelling Study. *MedRxiv*. [doi:10.1101/2020.03.03.20029843](https://doi.org/10.1101/2020.03.03.20029843)

Leyk, S., Gaughan, A. E., Adamo, S. B., de Sherbinin, A., Balk, D., Freire, S., & Comenetz, J. (2019). The Spatial Allocation of Population: A Review of Large-scale Gridded Population Data Products and Their Fitness for Use. *Earth System Science Data*, 11(3). [doi:10.5194/essd-11-1385-2019](https://doi.org/10.5194/essd-11-1385-2019)

Linard, C., Alegana, V. A., Noor, A. M., Snow, R. W., & Tatem, A. J. (2010). A High Resolution Spatial Population Database of Somalia for Disease Risk Mapping. *International Journal of Health Geographics*, 9(1), 45. [doi:10.1186/1476-072X-9-45](https://doi.org/10.1186/1476-072X-9-45)

Linard, C., & Tatem, A. J. (2012). Large-scale Spatial Population Databases in Infectious Disease Research. *International Journal of Health Geographics*, 11(1), 7. [doi:10.1186/1476-072X-11-7](https://doi.org/10.1186/1476-072X-11-7)

Lloyd, C. T., Sorichetta, A., & Tatem, A. J. (2017). High Resolution Global Gridded Data for Use in Population Studies. *Scientific Data*, 4(1), 1-17. [doi:10.1038/sdata.2017.1](https://doi.org/10.1038/sdata.2017.1)

Lloyd, C. T., Chamberlain, H., Kerr, D., Yetman, G., Pistolesi, L., Stevens, F. R., & Sinha, P. (2019). Global Spatio-Temporally Harmonised Datasets for Producing High-resolution Gridded Population Distribution Datasets. *Big Earth Data*, 3(2), 108-139. [doi:10.1080/20964471.2019.1625151](https://doi.org/10.1080/20964471.2019.1625151)

Local Burden of Disease Child Growth Failure Collaborators. (2020). Mapping Child Growth Failure Across Low-and Middle-Income Countries. *Nature*, 577(7789), 231. [doi:10.1038/s41586-019-1878-8](https://doi.org/10.1038/s41586-019-1878-8)

MacManus, K. (2019, 9 octobre). Entretien avec H. Dahmm et M. Rabiee.

McGlade, J., Bankoff, G., Abrahams, J., Cooper-Knock, S. J., Cotecchia, F., Desanker, P., & Hirsch, F. (2019). Global Assessment Report on Disaster Risk Reduction 2019. Extrait de <https://www.undrr.org/publication/global-assessment-report-disaster-risk-reduction-2019>

Melchiorri, M., Pesaresi, M., Florczyk, A. J., Corbane, C., & Kemper, T. (2019). Principles and Applications of the Global Human Settlement Layer as Baseline for the Land Use Efficiency Indicator—SDG 11.3. 1. *ISPRS International Journal of Geo-Information*, 8(2), 96. [doi:10.3390/ijgi8020096](https://doi.org/10.3390/ijgi8020096)

Mondal, P., & Tatem, A. J. (2012). Uncertainties in Measuring Populations Potentially Impacted by Sea Level Rise and Coastal Flooding. *PloS One*, 7(10). [doi:10.1371/journal.pone.0048191](https://doi.org/10.1371/journal.pone.0048191)

Mosser, J. F., Gagne-Maynard, W., Rao, P. C., Osgood-Zimmerman, A., Fullman, N., Graetz, N., & Earl, L. (2019). Mapping Diphtheria-Pertussis-Tetanus Vaccine Coverage in Africa, 2000–2016: A Spatial and Temporal Modelling Study. *The Lancet*, 393(10183), 1843-1855. [doi:10.1016/S0140-6736\(19\)30226-0](https://doi.org/10.1016/S0140-6736(19)30226-0)

Nansai, K., Tohno, S., Chatani, S., Kanemoto, K., Kurogi, M., Fujii, Y., & Lenzen, M. (2020). Affluent Countries Inflict Inequitable Mortality and Economic Loss on Asia via PM2. 5 Emissions. *Environment International*, 134, 105238. [doi:10.1016/j.envint.2019.105238](https://doi.org/10.1016/j.envint.2019.105238)

Nations unies. (2016). The Sustainable Development Goals Report 2016. United Nations. Extrait de <https://unstats.un.org/sdgs/report/2016/The%20Sustainable%20Development%20Goals%20Report%202016.pdf>

Nations unies. (2018). The Sustainable Development Goals Report 2018. United Nations Department of Economic and Social Affairs. Extrait de <https://www.un.org/development/desa/publications/thesustainable-development-goals-report-2018.html>

Nations unies. (n.d.). Population. Extrait de : <https://www.un.org/en/sections/issues-depth/population/index.html> Consulté le 16 avril 2020

Neumann, B., Vafeidis, A. T., Zimmermann, J., & Nicholls, R. J. (2015). Future Coastal Population Growth and Exposure to Sea-level rise and Coastal Flooding - A Global Assessment. *PloS One*, 10(3), e0118571. [doi:10.1371/journal.pone.0118571](https://doi.org/10.1371/journal.pone.0118571)

OMS. (n.d.). Natural Events. Organisation mondiale de la Santé (OMS). Extrait de : https://www.who.int/environmental_health_emergencies/natural_events/en/ Consulté le 16 avril 2020

OMS. (2019). World Malaria Report 2019. Organisation mondiale de la Santé (OMS). Extrait de <https://www.who.int/publications-detail/world-malaria-report-2019>

Parish, E. S., Kodra, E., Steinhäuser, K., & Ganguly, A. R. (2012). Estimating Future Global Per Capita Water Availability Based on Changes in Climate and Population. *Computers and Geosciences*, 42, 79-86. [doi:10.1016/j.cageo.2012.01.019](https://doi.org/10.1016/j.cageo.2012.01.019)

Pays partenaires du GPSDD. (2019). Entretien avec Alonso Ortiz Galan et Charu Vijayakumar.

Rainisch, G., Shankar, M., Wellman, M., Merlin, T., & Meltzer, M. I. (2014). Regional Spread of Ebola Virus, West Africa, 2014. *Emerging Infectious Diseases*, 21(3), 444. doi.org/10.3201/eid2103.141845

Reed, F. J., Gaughan, A. E., Stevens, F. R., Yetman, G., Sorichetta, A., & Tatem, A. J. (2018). Gridded Population Maps Informed by Different Built Settlement Products. *Data*, 3(3), 33. [doi:10.3390/data3030033](https://doi.org/10.3390/data3030033)

Rose, A. (2019, 14 novembre). Entretien avec H. Dahmm et M. Rabiee.

Roseth, B., Reyes, A. Yee Amezcaga, K. (2019). The Value of Official Statistics: Lessons from Intergovernmental Transfers. *Inter-American Development Bank (IDB)*. [doi:10.18235/0001883](https://doi.org/10.18235/0001883)

Sabesan, A., Abercrombie, K., Ganguly, A. R., Bhaduri, B., Bright, E. A., & Coleman, P. R. (2007). Metrics for the Comparative Analysis of Geospatial Datasets with Applications to High-resolution Grid-based Population Data. *GeoJournal*, 69(1-2), 81-91. [doi:10.1007/s10708-007-9103-y](https://doi.org/10.1007/s10708-007-9103-y)

Schiavina, M., Freire, S., Rosina, K., Ziemba, L., Marin Herrera, M., Craglia, M., Lavallo, C., Kemper, T., Batista, F. (2020). ENACT-POP R2020A - ENACT 2011 Population Grid. European Commission, Joint Research Centre (JRC) [Dataset]. [doi:10.2905/BE02937C-5A08-4732-A24A-03E0A48BDCDA](https://doi.org/10.2905/BE02937C-5A08-4732-A24A-03E0A48BDCDA). Accessed 15 April 2020.

Schinko, T., Drouet, L., Vrontisi, Z., Hof, A., Hinkel, J., Mochizuki, J., & Lincke, D. (2020). Economy-Wide Effects of Coastal Flooding Due to Sea level rise: A Multi-model Simultaneous Treatment of Mitigation, Adaptation, and Residual impacts. *Environmental Research Communications*, 2(1), 015002. [doi: 10.1088/2515-7620/ab6368](https://doi.org/10.1088/2515-7620/ab6368)

Schneider, A., Friedl, M. A., & Potere, D. (2010). Mapping Global Urban Areas Using MODIS 500-m Data: New Methods and Datasets Based on 'Urban Ecoregions'. *Remote Sensing of Environment*, 114(8), 1733-1746. [doi:10.1016/j.rse.2010.03.003](https://doi.org/10.1016/j.rse.2010.03.003)

SEDAC. (n.d.). SEDAC Global COVID-19 Viewer: Population Estimates by Age Group and Sex. Socioeconomic Data and Applications Center (SEDAC). Extrait de <https://sedac.ciesin.columbia.edu/mapping/popest/covid-19/> Accessed 16 April 2020

Shaddick, G., Thomas, M. L., Green, A., Brauer, M., van Donkelaar, A., Burnett, R., ... & Gurney, S. (2017). Data Integration Model for Air Quality: A Hierarchical Approach to the Global Estimation of Exposures to Ambient Air Pollution. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 67(1), 231-253. [doi:10.1111/rssc.12227](https://doi.org/10.1111/rssc.12227)

Smith, A., Bates, P. D., Wing, O., Sampson, C., Quinn, N., & Neal, J. (2019). New Estimates of Flood Exposure in Developing Countries Using High-Resolution Population Data. *Nature Communications*, 10(1), 1-7. [doi:10.1038/s41467-019-09282-y](https://doi.org/10.1038/s41467-019-09282-y)

Stevens, F. R., Gaughan, A. E., Linard, C., & Tatem, A. J. (2015). Disaggregating Census Data for Population Mapping Using Random Forests with Remotely-sensed and Ancillary Data. *PloS One*, 10(2). [doi:10.1371/journal.pone.0107042](https://doi.org/10.1371/journal.pone.0107042)

Tatem, A. (2019 September 4). Entretien avec H. Dahmm et M. Rabiee.

Thomson, D. (2019 November 12). Entretien avec H. Dahmm et M. Rabiee.

Thomson, D. R., Linard, C., Vanhuysse, S., Steele, J. E., Shimoni, M., Siri, J., ... & Georganos, S. (2019). Extending Data for Urban Health Decision-Making: A Menu of New and Potential Neighborhood-level Health Determinants Datasets in LMICs. *Journal of Urban Health*, 96(4), 514-536. [doi:10.1007/s11524-019-00363-3](https://doi.org/10.1007/s11524-019-00363-3)

Thomson, D. R., Stevens, F. R., Ruktanonchai, N. W., Tatem, A. J., & Castro, M. C. (2017). GridSample: An R Package to Generate Household Survey Primary Sampling Units (PSUs) From Gridded Population Data. *International Journal of Health Geographics*, 16(1), 25. [doi:10.1186/s12942-017-0098-4](https://doi.org/10.1186/s12942-017-0098-4)

Tiecke, T.G., Liu, X., Zhang, A., Gros, A., Li, N., Yetman, G., Kilic, T., Murray, S., Blankespoor, B., Prydz, E.B., & Dang, H.H. (2017). Mapping the world population one building at a time. ArXiv, abs/1712.05839. Extrait de <https://www.semanticscholar.org/paper/Mapping-the-world-population-one-building-at-a-time-Tiecke-Liu/b17d3ded924a92b31cce96d11c345a9b82367457>

Tuholske, C., Caylor, K., Evans, T., & Avery, R. (2019). Variability in Urban Population Distributions across Africa. *Environ. Res. Lett.* 14 085009. [doi:10.1088/1748-9326/ab2432](https://doi.org/10.1088/1748-9326/ab2432)

Tusting, L. S., Bisanzio, D., Alabaster, G., Cameron, E., Cibulskis, R., Davies, M., & Okumu, F. O. (2019). Mapping Changes in Housing in Sub-Saharan Africa from 2000 to 2015. *Nature*, 568(7752), 391-394. [doi:10.1038/s41586-019-1050-5](https://doi.org/10.1038/s41586-019-1050-5)

ONU. IAEG-SDG. (2019). Tier Classification for Global SDG Indicators 11 décembre 2019. Nations unies Groupe d'experts des Nations Unies et de l'extérieur chargé des indicateurs relatifs aux objectifs de développement durable. Extrait de : <https://unstats.un.org/sdgs/files/Tier-Classification-of-SDG-Indicators-11-December-2019-web.pdf>

UN-GGIM Europe. (2019). The Territorial Dimension in SDG Indicators: Geospatial Analysis and its Integration with Statistical Data. Comité d'experts sur la gestion de l'information géospatiale à l'échelle mondiale des Nations unies. Extrait de https://un-ggim-europe.org/wp-content/uploads/2019/05/UN_GGIM_08_05_2019-The-territorial-dimension-in-SDG-indicators-Final.pdf

UDSNU. (n.d.). Demographic and Social Statistics. Division de statistique des Nations Unies (DSNU). Extrait de : <http://mdgs.un.org/unsd/demographic/sources/census/wphc/censusclockmore.htm>
Consulté le 16 avril 2020

Van Den Hoek, J. (2018). Mapping the Missing Millions: A Global Analysis of Informal Settlements and the Geography of Exclusion. In AGU Fall Meeting Abstracts. Extrait de <https://ui.adsabs.harvard.edu/abs/2018AGUFMPA43E1387V/abstract>

Vincent, S. (2018). Status Review of the Updated Rural Access Index (RAI). ReCAP for DFID. Extrait de http://research4cap.org/Library/Vincent-CDS-2018-StatusReviewUpdatedRAI-FinalReport_GEN2033C-180529.pdf

Wardrop, N. A., Jochem, W. C., Bird, T. J., Chamberlain, H. R., Clarke, D., Kerr, D., & Tatem, A. J. (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.
[doi:10.1073/pnas.1715305115](https://doi.org/10.1073/pnas.1715305115)

Weber, E. M., Seaman, V. Y., Stewart, R. N., Bird, T. J., Tatem, A. J., McKee, J. J., & Reith, A. E. (2018). Census-independent Population Mapping in Northern Nigeria. *Remote Sensing of Environment*, 204, 786-798. [doi:10.1016/j.rse.2017.09.024](https://doi.org/10.1016/j.rse.2017.09.024)

Ye, T., Zhao, N., Yang, X., Ouyang, Z., Liu, X., Chen, Q., & Jia, P. (2019). Improved Population Mapping for China Using Remotely Sensed and Points-of-Interest Data Within a Random Forests model. *Science of the Total Environment*, 658, 936-946.
[doi:10.1016/j.scitotenv.2018.12.276](https://doi.org/10.1016/j.scitotenv.2018.12.276)

ANNEXE A

1. À quelle date le dernier recensement a-t-il été mené dans votre pays et à quelle fréquence sont-ils effectués ?
2. Quelles sont les difficultés dans les données associées aux recensements dans votre pays ?
3. Avez-vous envisagé d'utiliser d'autres sources de données démographiques provenant de données d'OT ou du secteur des télécommunications qui complèteraient les sources démographiques conventionnelles ? Quelles sources avez-vous envisagées et comment obtenez-vous des informations à leur sujet ?
4. Y a-t-il des situations où les pouvoirs publics ou d'autres acteurs de votre pays utilisent des estimations démographiques maillées (estimations spatiales désagrégées ou matricielles provenant de plusieurs sources primaires comme les données de recensements menés dans le passé, les données satellites, etc.) pour obtenir des statistiques ou éclairer les décisions ? Si oui, pouvez-vous donner des exemples ?
5. Quels sont les avantages des données démographiques maillées par rapport aux données de l'état civil ou administratives, enquêtes par sondage, etc. ? Pourriez-vous expliquer la façon dont vous utilisez les différentes sources de données ?
6. Quelles sources de données démographiques maillées utilisez-vous (par exemple, GPW, LandScan, etc.) ? Utilisez-vous différentes sources pour différentes applications ou régions géographiques ? Pourriez-vous donner des exemples ?
7. Le choix de certaines sources de données démographiques maillées a-t-il été motivé par des raisons méthodologiques, politiques ou de disponibilité particulières ? A-t-on comparé plusieurs sources de données avant de procéder à une sélection ?
8. Quelles sont les lacunes que vous remarquez en matière de disponibilité et d'accessibilité des sources de données démographiques maillées ? De quelles ressources et de quels outils supplémentaires avez-vous besoin pour combler ces lacunes ?

