



GRID³

GEO-REFERENCED INFRASTRUCTURE AND
DEMOGRAPHIC DATA FOR DEVELOPMENT

Cartographie et classification des localités

Juin 2021



Table des matières

2

Glossaire

4

Résumé

8

Introduction

11

Créer des couches de localités

16

Approches de classification

19

Applications

25

Conclusion

26

Reconnaissance et attribution

26

Travaux cités

26

Historique des versions
papier

Glossaire

Classification au niveau de la zone

Une approche de classification des localités qui utilise des ensembles de données d'empreintes de bâtiment à haute résolution pour identifier les schémas de construction qui peuvent être observés.

Empreintes de bâtiments

Contours numérisés des bâtiments dérivés d'imagerie satellite.

Classification des localités en fonction des types de bâtiment

Le niveau le plus élémentaire de classification des localités, qui catégorise chaque bâtiment d'une localité en 'résidentiel ou non résidentiel.

Couche de localité exhaustive

Un ensemble de données qui combine l'étendue des localités avec les attributs de la couche de points (par exemple, les noms de lieux ou les unités administratives). Une couche de localité exhaustive peut se présenter sous la forme d'une couche de points ou d'une couche de polygones.

Schéma de base de données

Structure du tableau des lignes et des colonnes utilisées dans un fichier de données.

Ensemble

Une combinaison de prédictions entre les modèles utilisés, pour améliorer la prédiction globale et pour éviter l'enchevêtrement de tout modèle individuel dans l'ensemble.

Étendue

La couverture géographique d'un groupe de localités.

Catégorisation des quartiers au sein d'une localité

L'application de méthodes pour identifier et classer les zones locales et quartiers dans une localité en différents types en fonction de la taille, de la forme et de la disposition des structures, ainsi que de leur relation avec les infrastructures et d'autres caractéristiques du paysage.

Apprentissage automatique

Un sous-ensemble d'intelligence artificielle dans lequel un ordinateur utilise un modèle mathématique statistique (ou un ensemble) afin d'analyser les schémas non-apparents et les inférences dans un ensemble de données. Le modèle permet à l'analyse d'être appliquée (c'est-à-dire de prédire) efficacement au-delà des données d'apprentissage.

Points d'intérêt (POI)

Points géo-référencés qui représentent des infrastructures, des bâtiments et des points de repère. Ces points localisent les structures et les services essentiels à la santé et au bien-être de la société, tels que, par exemple, les établissements de santé, les écoles, les marchés, les banques, les entrepôts et les puits.

Polygone

Une représentation numérique des limites d'une localité. Les polygones décrivent également l'étendue d'une localité et peuvent inclure d'autres informations d'attribut, telles que le nombre total de bâtiments dans la zone donnée.

Modèle de probabilité

Une représentation mathématique d'un phénomène aléatoire. ¹

Raster

Dans sa forme la plus simple, une matrice de carreaux (ou pixels) organisée en lignes et colonnes (ou un quadrillage / carroyage) où chaque carreau contient une valeur représentant des informations, telles que la température ou le nombre de bâtiments. Les rasters comprennent des modèles d'élévation numériques, des photographies aériennes numériques, des images provenant de satellites, des images numériques ou même des cartes numérisées. ²

Télédétection

Processus de détection et de surveillance des caractéristiques physiques d'une zone en mesurant son rayonnement réfléchi et émis à distance de la zone ciblée. Des capteurs spéciaux sur des satellites ou des avions collectent des images de la Terre par télédétection, qui aident les chercheurs à « détecter » les informations sur la Terre. ³

Localité

Une zone peuplée de structures et de complexes habités en permanence qui peuvent aller de petits villages ruraux à de grandes zones urbaines.

Données de localité

Informations sur l'emplacement et les caractéristiques des structures potentiellement habitées qui sont utilisées comme données d'entrée pour produire des estimations de population à haute résolution.

Couche de localité

Un jeu de données qui fournit des points ou des polygones des localités, avec leurs noms, afin de les localiser, identifier et visualiser spatialement.

Cartographie des localités

La collecte, la création et l'harmonisation de plusieurs sources de données représentant des lieux habités par des personnes.

Modèles supervisés

Un sous-ensemble de l'apprentissage automatique (machine learning). Les modèles supervisés sont entraînés avec de petites quantités de données pré-classifiées, par opposition aux méthodes de classification non supervisées qui n'utilisent pas des données d'apprentissage. Un modèle aide à prédire le comportement de quelque chose dans le monde réel (appelé « système »). Le modèle explique comment le système fonctionne et démontre l'effet des différentes composantes du système.

Point de localité

Représentation de l'emplacement d'une localité et de ses données d'attribut de base.

Modèle de régression Poisson avec inflation à zéro (ZIP)

Utilisé pour modéliser les données d'effectif qui ont un excès d'effectifs à zéro. La théorie suggère que les zéros en excès peuvent être modélisés indépendamment des effectifs positifs. Ainsi, le modèle ZIP comporte deux parties, un modèle de Poisson pour les effectifs positifs et un modèle LOGIT pour prédire les excès d'effectifs à zéro. ⁴

1. Institut national des normes et de la technologie. "Modèle de probabilité". <https://csrc.nist.gov/glossary/term/Probability-model>.

2. Définition basée sur l'Institut de recherche sur les systèmes environnementaux. "What is raster data?".

<https://desktop.arcgis.com/en/arcmap/10.3/manage-data/raster-and-images/what-is-raster-data.htm>.

3. Définition basée sur le Service géologique des États-Unis. "What is remote sensing and what is it used for?". https://www.usgs.gov/faqs/what-remote-sensing-and-what-it-used?qt-news_science_products=7#qt-news_science_products

4. Définition basée sur l'UCLA Institute for Digital Research & Education Statistical Consulting. "Zero-Inflated Poisson Regression | R Data Analysis Examples". <https://stats.idre.ucla.edu/r/dae/zip/>.

Résumé

Afin d'identifier les priorités de développement d'un pays, il est crucial de comprendre autant que possible les lieux où les gens vivent et travaillent. Les données que nous recueillons sur ces localités peuvent être utilisées pour guider un large éventail d'activités, qu'il s'agisse de planifier le budget d'un gouvernement, de distribuer des moustiquaires dans une communauté ou de mener une intervention rapide en cas de catastrophe. Plus les données de localités sont précises et détaillées, plus ces interventions peuvent être efficaces.

Création de la couche de localité

Traditionnellement, la cartographie de recensement et les cartes opérationnelles sont la principale source de données sur les localités. Dans certains cas, ces données sont des cartes papier et des listes non géo-référencées incomplètes ou obsolètes. En revanche, le travail de GRID3 (Geo-Referenced Infrastructure and Demographic Data for Development) se concentre sur la création d'une couche des localités améliorée et exhaustive qui peut permettre d'obtenir une image réelle des communautés, facilitant ainsi la production d'analyses plus précises et plus efficaces.

La première étape de la création d'une couche de localité exhaustive consiste à obtenir les données existantes (points ou polygones) auprès des autorités gouvernementales, ainsi que l'autorisation de publier ces données. Une fois cela fait, GRID3 effectue une première analyse exploratoire pour se familiariser avec le contenu de ces ensembles de données. Ensuite, les données sont évaluées pour déterminer si elles couvrent toute la zone d'intérêt. Bien que les données gouvernementales sont souvent celles qui servent de référence, elles sont souvent incomplètes. Ainsi, la prochaine étape consiste à compléter les données dont dispose GRID3 en utilisant des données provenant de diverses sources supplémentaires, notamment des ONG, des entreprises privées, des efforts de terrain menés par la population, des entités des Nations Unies et des prestataires indépendants. Une fois que les données ont été recueillies et évaluées, elles sont prêtes à être nettoyées. Cela implique de parcourir de nombreux fichiers, beaucoup d'entre eux dans différents formats et avec des doublons dans et entre ces fichiers. Ayant compris ce qui est contenu dans chaque fichier, GRID3 peut commencer à y intégrer les données gouvernementales. GRID3 examine comment les données dans les fichiers sont représentées, puis normalise les mots utilisés pour représenter les différents types de données. Cela permet entre autres l'identification et la suppression des doublons. Enfin, la structure tabulaire des lignes et des colonnes utilisées au sein d'un fichier de données (ou « schéma de base de données ») est standardisée afin que les fichiers contenant les différentes données soient placés dans un format uniforme.

Les étendues des localités étaient précédemment fournies par le laboratoire national d'Oak Ridge. Actuellement, les données d'Ecopia Landbase Africa, alimentées par les empreintes des bâtiments de Maxar, sont utilisées pour la production des limites des localités qui aident à définir les étendues.. L'étape suivante consiste à valider les données. Tout d'abord, GRID3 compare les polygones aux points et identifie ce qui est manquant ou incorrect dans les points, en vérifiant les défauts d'alignement, les erreurs potentielles et les lacunes dans les données. Ensuite, la comparaison est

inversée: les points sont comparés aux polygones pour voir ce qui manque dans les polygones. Enfin, GRID3 crée une couche de localité exhaustive . Cette couche combine les étendues des localités avec les attributs de la couche de points. La couche de localité exhaustive peut se présenter sous la forme d'une couche de points ou d'une couche de polygones. Cette couche est ensuite envoyée à une série de réviseurs. Après avoir effectué les ajustements nécessaires, GRID3 finalise la couche.

Approches de classification

Bien que de tels ensembles de données peuvent fournir des détails sur la présence et l'emplacement des localités, ils manquent souvent d'informations qui peuvent différencier un lieu dans un noyau urbain d'une périphérie croissante d'une ville, ou une périphérie croissante d'une ville d'un village éloigné. Cependant, la taille et la forme des structures, ainsi que la disposition et la position des structures (par rapport à d'autres infrastructures ou caractéristiques), peuvent transmettre des informations sur différentes utilisations des terres ou activités économiques. GRID3 explore de nouvelles méthodes pour identifier ces modèles et extraire des informations supplémentaires qui peuvent améliorer notre compréhension des localités à l'échelle des structures individuelles et des quartiers. Plus précisément, GRID3 développe deux approches qui utilisent les empreintes de bâtiments, des couches de données géospatiales et des algorithmes d'apprentissage automatique pour classer les structures et les zones locales.

Le niveau le plus élémentaire de classification des localités est par type de bâtiment (résidentiel ou non résidentiel). Les caractéristiques de chaque structure peuvent donner des indices sur l'utilisation d'un bâtiment. Pour prédire avec précision le type de bâtiment, GRID3 utilise plusieurs méthodes d'apprentissage automatique. Chaque méthode fournit une prédiction et ces prédictions sont combinées entre les modèles dans un ensemble qui contribue à améliorer la prédiction globale. Les méthodes statistiques utilisées sont des modèles supervisés ; c'est-à-dire que les modèles sont entraînés avec de petites quantités de données pré-classifiées. GRID3 utilise les données d'OpenStreetMap et d'autres ensembles de données de bâtiments étiquetés.

La distinction binaire entre résidentiel et non résidentiel est importante pour soutenir des modèles de population plus précis et pour guider les services qui doivent trouver des zones résidentielles. Cependant, ces deux classes en disent peu sur le type de quartier dans lequel se trouve un bâtiment et, par extension, sur la manière dont ce quartier s'intègre dans les modèles d'établissement de la zone géographique plus large. Pour cette information, GRID3 développe un deuxième modèle de classification au niveau de la zone. Cette approche commence une fois de plus avec les ensembles de données d'empreinte de bâtiment à haute résolution ; mais dans ce cas, GRID3 se concentre sur les schémas qui peuvent être observés dans les bâtiments (plutôt que sur les structures individuelles). Lorsqu'elles sont considérées ensemble, la taille, la densité, les formes et les orientations des structures créent une « texture » visuelle qui varie à travers le paysage, indiquant des différences dans l'utilisation des terres ou le type de quartier. Pour créer ces classifications de quartier, le motif et la texture sont quantifiés à plusieurs échelles spatiales sur toute une région d'intérêt en calculant la variation des tailles et des formes, ainsi que de nombreuses autres mesures sur une grille régulière d'emplacements. Des modèles de mélange gaussien et d'autres méthodes d'apprentissage automatique sont utilisés pour trouver des groupes de schémas similaires et pour classer les types sur une surface carroyée de résolution de 100 x 100 mètres.

GRID3 continue de développer cette ligne de recherche - les efforts du programme incluront à terme l'élargissement de la gamme d'attributs générés par (et donc les données d'entrée utilisées par) le modèle de classification, ce qui améliorera le pouvoir prédictif. Bien que la recherche sur la classification en soit encore à ses débuts, elle offre déjà une perspective radicalement différente - et améliorée - sur les empreintes des bâtiments et aide à situer ces empreintes dans des schémas de localités plus larges.

Applications

Les produits GRID3 sur les localités ont un large éventail d'utilisations. Ils peuvent générer des données open source, fournir un soutien aux applications de données pour assurer un impact efficace, ou permettre une formation pour renforcer les bases géospatiales nationales pour le développement futur basé sur des preuves et la prise de décision humanitaire. Ces produits s'intègrent également dans les autres travaux du programme sur le recensement, les modèles de population, la délimitation des limites administratives et opérationnelles, et l'identification des emplacements des infrastructures. Les données sur les établissements de GRID3 ont déjà commencé à avoir des impacts dans plusieurs domaines différents du développement durable; des exemples de ces impacts sont décrits ci-dessous.

- Alors que la plupart des localités dans l'est de la République démocratique du Congo (RDC) sont visibles sur les images satellite, beaucoup n'ont pas de noms qui leur sont attachés. GRID3 a travaillé avec divers partenaires pour changer cela et a jusqu'à présent identifié avec succès les noms de toutes les localités dans un ensemble sélectionné de zones de santé prioritaires dans les provinces du Haut-Lomami et du Tanganyika.
- GRID3 a également amélioré la cartographie des localités dans l'est de la RDC en organisant des sessions de cartographie participatives. En créant une carte des localités plus complète dans les deux provinces, GRID3 a permis aux agents de santé d'améliorer l'exhaustivité des localités incluses dans leurs micro-plans, permettant d'inclure des populations cibles supplémentaires dans les futures campagnes de vaccination.
- Au Nigeria, GRID3 a aidé les équipes de vaccination à surveiller la couverture des localités, à réduire le nombre de localités manquées et à améliorer les résultats des équipes. Ce travail a permis de générer des cartes de micro-plans basés sur des systèmes d'information géographique (SIG) pour les équipes de vaccination et de visualiser les localités.
- GRID3 a travaillé en étroite collaboration avec le gouvernement zambien pour créer une couche de localité exhaustive qui fournira le référentiel le plus complet à ce jour pour les noms et les emplacements des localités du pays. Un large éventail d'avantages résulte de cet ensemble de données encore en développement, notamment: une aide à la planification d'un recensement à l'échelle nationale, des campagnes de vaccination plus efficaces et une capacité accrue de réponse aux catastrophes.

Conclusion

Les activités de GRID3 ajoutent de la valeur aux efforts visant à comprendre les localités, à la fois en améliorant les techniques de collecte/analyse des données et en améliorant les données elles-mêmes. De nombreux défis doivent encore être relevés dans la campagne de collecte de données complètes et exploitables dans ce domaine. Ceux-ci comprennent: la capture de localités temporaires, les déplacements temporaires de population, les déplacements de population à long terme, les campements nomades, diurnes et saisonniers; distinguer les populations résidentielles des populations non résidentielles; concilier les différences entre les conceptions officielles et locales des lieux, des noms de lieux et des limites des localités; et négocier les sensibilités politiques et les questions de sécurité à la fois à l'intérieur et entre les pays autour des lieux, des noms de lieux et des limites des localités.

Dans les années à venir, GRID3 continuera à travailler avec des partenaires pour rendre ses données actuelles plus conviviales et pour chercher de nouvelles façons de collecter et de valider les informations sur les localités.

Introduction

Afin d'identifier les priorités de développement d'un pays, il est crucial de comprendre autant que possible les lieux où les gens vivent et travaillent. Les données que nous collectons sur ces localités peuvent être utilisées pour éclairer un large éventail d'activités, qu'il s'agisse de planifier le budget d'un gouvernement, de distribuer des moustiquaires dans une communauté ou de mener une intervention rapide en cas de catastrophe. Plus les données de localité sont précises et détaillées, plus ces interventions peuvent être efficaces.

Les localités sont des lieux dynamiques qui existent souvent dans de multiples contextes : les localités peuvent être définies par des limites administratives/politiques, des espaces fonctionnels, leurs relations avec d'autres lieux ou la façon dont les gens utilisent et expérimentent un espace.

Et pourtant, s'assurer que les données de localités sont exactes et détaillées n'est pas simple. Les localités sont des lieux dynamiques qui existent souvent dans de multiples contextes : elles peuvent être définies par des limites administratives ou politiques, des espaces fonctionnels, leurs relations avec d'autres lieux, ou la façon dont les gens utilisent et expérimentent un espace. Pour ces raisons et bien d'autres, les méthodes utilisées pour mesurer les localités doivent être aussi dynamiques que les localités elles-mêmes.

Les travaux GRID3 sur les localités

GRID3 (Geo-Referenced and Demographic Data for Development) est un programme financé par une subvention de la Fondation Bill & Melinda Gates et du Bureau du Commonwealth et du Développement des affaires étrangères du Royaume-Uni (FCDO). Il est mis en œuvre par CIESIN de l'Université de Columbia (Center for International Earth Science Information Network) le Fonds des Nations Unies pour la population (UNFPA), WorldPop à l'Université de Southampton et Flowminder. La mission principale de GRID3 est de créer des solutions de données spatiales qui rendent les objectifs de développement réalisables. Pour ce faire, GRID3 s'appuie sur l'expertise de ses partenaires au sein du gouvernement, des Nations Unies, des universités et du secteur privé pour concevoir des solutions géospatiales adaptables et pertinentes basées sur les capacités et les besoins de développement des pays en développement.

GRID3 définit une localité comme une zone habitée de structures et de complexes habités en permanence qui peuvent aller des petits villages ruraux aux grandes zones urbaines. Les données de localité, à leur tour, fournissent des informations sur l'emplacement et les caractéristiques des structures potentiellement habitées et sont utilisées comme données d'entrée pour produire des estimations de population à haute résolution.

Les données peuvent être dérivées d'images aériennes/satellites à haute résolution ou d'autres cartes de bâtiments, et peuvent être représentées de plusieurs manières: points/polygones de localité, modélisation probabiliste ou catégorisation des quartiers au sein d'une localité.

- **Points/polygones:** Les points indiquent l'emplacement d'une localité et des données attributaires de base (par exemple le nom de l'agglomération). Les polygones représentent les limites d'une localité ; ils décrivent également son étendue (c'est-à-dire la couverture géographique) et peuvent inclure d'autres informations d'attribut, telles que le nombre total de bâtiments dans la zone donnée.
- **Modèle de probabilité:** A Une représentation mathématique d'un phénomène aléatoire. Pour prédire l'occurrence des localités, un modèle de Poisson avec inflation du zéro (ZIP) est utilisé.⁵
- **Catégorisation des quartiers au sein d'une localité:** Alors que les deux représentations de données de localité précédentes se concentrent sur la localisation des localités dans l'espace et l'identification des attributs clés, des informations supplémentaires peuvent être associées à ces emplacements. La catégorisation des quartiers au sein d'une localité est l'application de méthodes pour identifier et classer les zones locales ou quartiers dans une localité en différents types en fonction de la taille, de la forme et de la disposition des structures, ainsi que de leur relation avec les infrastructures et d'autres caractéristiques du paysage.

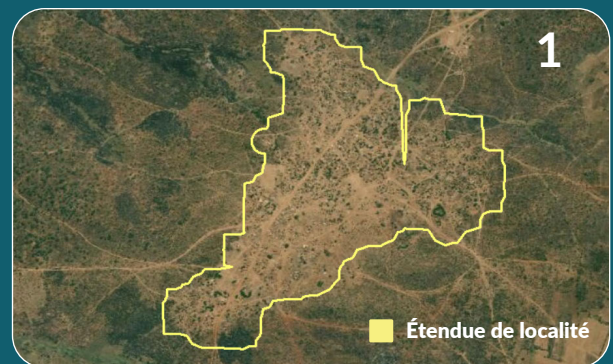
Représentations des données de localité

1. Points/polygones

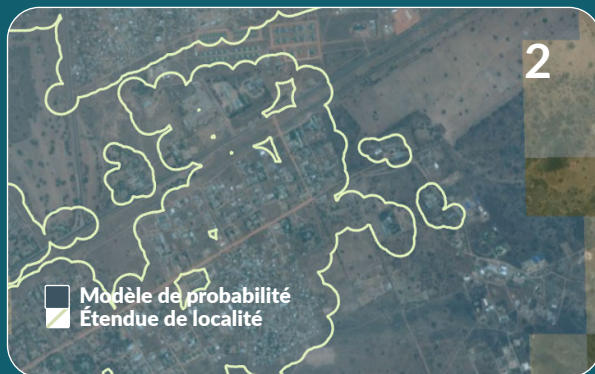
2. Modèle de probabilité

3. Catégorisation des quartiers au sein des localités

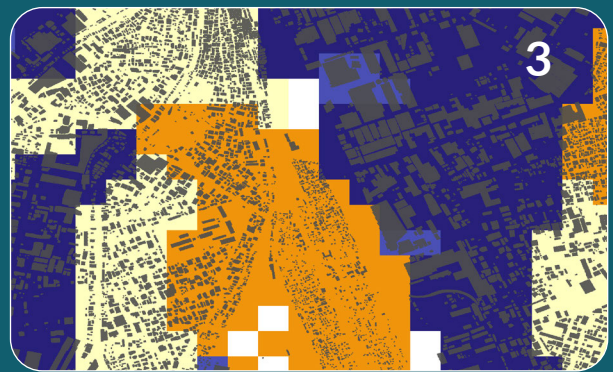
Types de zone affichés dans différentes couleurs et superposés avec des empreintes de bâtiments. En utilisant uniquement des informations extraites des emplacements des bâtiments et de leurs schémas, cela peut aider à distinguer différentes zones d'implantation au sein d'une ville.



Étendue de localité: CIESIN, Columbia University dans le cadre du programme GRID3 utilisant des cartes vectorielles Ecopia Powered by © 2020 Maxar; image de base: Esri



Étendue de localité: Oak Ridge National Laboratory; Modèle de probabilité: Markus Walsh; image de base: Esri



Empreintes de bâtiments © 2020 Maxar Technologies, Ecopia.IA

5. Utilisé pour modéliser les données de comptage qui ont un excès de comptage nul. En outre, la théorie suggère que les zéros en excès sont générés par un processus distinct des valeurs de comptage et que les zéros en excès peuvent être modélisés indépendamment. Ainsi, le modèle ZIP comporte deux parties, un modèle de comptage de Poisson et le modèle LOGIT pour prédire les excès de zéros.

Traditionnellement, la cartographie de recensement et les cartes opérationnelles sont la principale source de données des localités. Dans certains cas, ces données sont des cartes papier et des listes non géo-référencées incomplètes ou obsolètes. Récemment, dans le cadre de la modernisation du recensement, la cartographie numérique est encouragée; mais cette approche n'est pas encore largement utilisée. En revanche, le travail de GRID3 se concentre sur la création d'une couche de localité améliorée et exhaustive⁶ qui peut permettre une image réelle des communautés, ce qui facilite à son tour des analyses plus précises et plus efficaces. Plutôt que de générer de nouvelles données par un travail de terrain à grande échelle coûteux et chronophage, GRID3 harmonise les données qui ont déjà été collectées auprès d'organisations gouvernementales et non gouvernementales et les intègre aux données de terrain ciblées nouvellement collectées.

Une fois qu'un ensemble de données de localité nettoyé et standardisé est en main, GRID3 effectue l'harmonisation et l'intégration décrites ci-dessus via la cartographie des localités. Il s'agit de la collecte, de la création et de l'harmonisation de plusieurs sources de données représentant des lieux. GRID3 travaille en collaboration avec les parties prenantes nationales pour intégrer de nouvelles données dans les cartes de localité, un processus qui implique souvent l'utilisation de la télédétection⁷ et des données de terrain GPS pour générer des données de localité précises, à jour et complètes. La création, la liaison et/ou l'harmonisation des cartes des localités peuvent améliorer les initiatives de développement de multiples façons. Il facilite le recensement en aidant les gestionnaires de recensement à savoir où vivent les gens. Il est essentiel pour délimiter les limites et identifier les emplacements des infrastructures (par exemple, en aidant à identifier tous les points connus pour faire partie d'une zone administrative ou sanitaire donnée). Et, plus généralement, cette cartographie est essentielle à la planification durable, car les localités cartographiées et nommées avec précision permettent de suivre les progrès des activités de développement, des interventions et des programmes dans des secteurs critiques tels que l'agriculture, la santé et l'éducation.

Structure of this paper

Cet article décrit l'approche de GRID3 pour la cartographie et la collecte de données sur les localités. La première section décrit le processus de création d'une couche de localité exhaustive. La deuxième section décrit les motivations et les différentes approches de la classification des localités. La troisième section traite des diverses applications réelles de la cartographie des localités de GRID3 pour les initiatives de développement. La quatrième section conclut le document par une enquête sur les autres avantages potentiels du travail des localités, les lacunes actuelles des méthodes de GRID3 et les plans pour les futurs travaux du programme dans ce domaine.

6. Une couche de localité est un jeu de données qui fournit des points ou des polygones avec leurs noms pour localiser, identifier et visualiser dans l'espace ces localités. Une couche de localité exhaustive combine les étendues des localités avec les attributs (par exemple, les noms de lieux ou les unités administratives) de la couche de points. La couche de localité exhaustive peut se présenter sous la forme d'une couche de points ou d'une couche de polygones.

7. « La télédétection est le processus de détection et de surveillance des caractéristiques physiques d'une zone en mesurant son rayonnement réfléchi et émis à distance de la zone ciblée. Des caméras spéciales sur les satellites collectent des images de la Terre par télédétection, qui aident les chercheurs à « détecter » les informations sur la Terre. Voir la référence du Service géologique des États-Unis ci-dessus.

Création de la couche de localité

De nombreux pays en développement s'appuient sur des cartes obsolètes, incomplètes ou inexactes pour prendre leurs décisions politiques, ce qui entraîne une planification inefficace et inéquitable, ainsi qu'une répartition inégale des ressources. Cependant, de nouvelles méthodes sont en cours d'élaboration pour créer de nouvelles cartes de localités améliorées qui peuvent aider à dénombrer la population au sol et à améliorer le développement. GRID3 aide à diriger l'effort de collecte, d'organisation et d'utilisation des données qui peuvent faire des cartes de localités des outils plus efficaces pour le développement. Un élément clé du travail de GRID3 dans ce domaine consiste à créer une couche de localités exhaustive composée d'étendues de polygones et de données de points d'attributs.

Évaluation des données existantes

La première étape de la création d'une couche de localité exhaustive consiste à obtenir des données (points ou polygones) des autorités gouvernementales (par exemple, des cartes administratives ou des cartes opérationnelles) ainsi que l'autorisation de publier ces données. Après cela, GRID3 effectue une première analyse exploratoire pour se familiariser avec le contenu des ensembles de données (établir, par exemple, le format des données). Ensuite, les données sont évaluées pour déterminer si elles couvrent l'ensemble de la zone d'intérêt. GRID3 cherche à voir si une zone a peu de points de données parce qu'elle n'est pas peuplée, ou parce que GRID3 manque de données. GRID3 cherche également à voir si les données sont plus complètes dans certaines zones que dans d'autres.

Alors que les données gouvernementales sont souvent le type d'information qui fait le plus autorité, les données sont souvent incomplètes - par exemple, les données ne couvriront que la moitié du pays, ou elles couvriront l'ensemble du pays mais avec des niveaux de granularité inégaux (par exemple, une région ou province aura 100 points de localité, tandis qu'une autre en comptera 100 000). Ainsi, la prochaine étape consiste à compléter les données dont dispose GRID3 en utilisant des données provenant de diverses sources supplémentaires, notamment des ONG, des entreprises privées, des efforts de terrain menés par la population, des entités des Nations Unies et des prestataires indépendants⁸. GRID3 réitère ce processus autant de fois que nécessaire, en comblant autant que possible les lacunes dans les données. Au cours de cette étape, GRID3 peut retourner auprès du gouvernement pour obtenir plus de données, continuer à rechercher d'autres données auprès de sources non gouvernementales, effectuer un travail de terrain ciblé pour obtenir des données, ou s'associer avec une organisation ou une entreprise qui est déjà dans le pays et obtenir des données de leur part.

8. If there are several providers of raw settlement data (points and polygons) and conflicting information is provided, field work is conducted to resolve the issues, or the data are reviewed by a local stakeholder. If this is not possible, data provided by the country are defined as the authoritative source and will take precedence; but the schema provides space for alternate names.

Création de la base de données des emplacements des localités (points)

Une fois que les points de données ont été recueillis et évalués, ils sont prêts à être nettoyés. Cela implique de parcourir de nombreux fichiers, beaucoup d'entre eux dans différents formats et avec des doublons dans et entre les fichiers. Ayant compris ce qui est contenu dans chaque fichier, GRID3 peut commencer à intégrer les données des points de localité existants. GRID3 examine la manière dont les données dans les fichiers sont représentées (par exemple, le mot « village » peut être orthographié différemment dans différents fichiers), puis normalise les mots utilisés pour représenter différents types de données. Cela permet entre autres l'identification et la suppression des doublons. Enfin, la structure tabulaire des lignes et des colonnes utilisées au sein d'un fichier de données (ou « schéma de base de données ») est standardisée afin que les fichiers contenant les différentes données soient placés dans un format uniforme. Cela implique de fusionner des centaines de fichiers et de normaliser les noms de colonnes et les types de données afin que les tableaux de chaque fichier contiennent des données pouvant être consolidées de manière uniforme, chacune contenant un identifiant unique, des noms alternatifs, l'emplacement géographique du point de données et un identifiant de localité unique.

Création de la base de données des emplacements des localités (points)

Les étendues des localités étaient précédemment fournies par Oak Ridge National Laboratory⁹. Actuellement Ecopia Landbase Africa, alimenté par les empreintes de bâtiments de Maxar¹⁰ (c'est-à-dire les contours numérisés des bâtiments) permettent la production d'une limite qui aide à définir les étendues.¹¹ Cela se fait soit en agrégeant des entités situées à moins de 25m à 100m les unes des autres, soit en utilisant la densité du bâtiment pour dessiner des contours autour de groupes d'empreintes de bâtiments. Les étendues de localité sont divisées en quatre catégories : 1. les zones bâties, 2. les petites zones de peuplement, 3. les hameaux et 4. les zones de hameaux.

9. Voir Oak Ridge National Laboratory. <https://www.ornl.gov/>.

10. Voir Maxar Technologies, Inc. and Ecopia Tech Corporation. 2020. Ecopia Landbase Africa powered by Maxar. DigitizeAfrica.ai.

11. Certaines zones peuvent avoir une grande population mais une petite étendue, ce qui indiquerait que de nombreuses personnes sont regroupées dans une petite zone - la densité de population est donc importante. Certaines zones peuvent avoir la même population mais une plus grande étendue, ce qui signifie que la population est répartie sur une plus grande zone.

Four Categories of Settlement Extents

1. Polygones des zones bâties

Une zone d'urbanisation avec des bâtiments modérément dense, espacés et une maille visible des rues et des îlots.

2. Polygones des petites localités + points

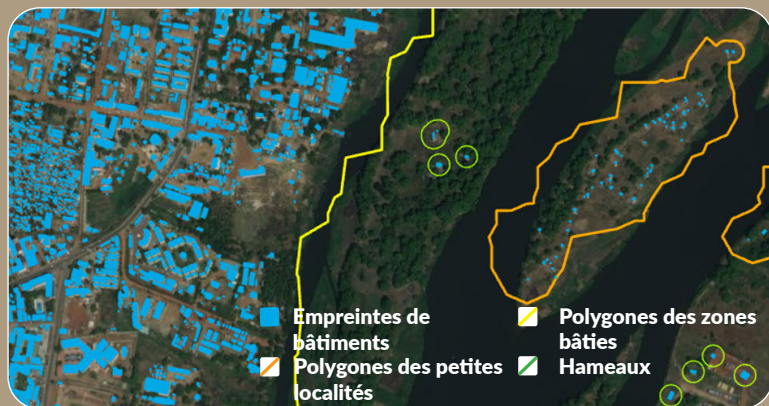
Une zone peuplée de structures et de complexes habités en permanence d'environ quelques centaines à quelques milliers d'habitants. Le modèle de logement dans les SSA est un assemblage de complexes familiaux jouxtant d'autres habitations similaires.

3. Polygones des hameaux + points

Un complexe unifamilial ou plusieurs complexes ou maisons de couchage isolés des petites agglomérations ou des zones urbaines.

4. Polygones des zones de hameau

Groupe de deux hameaux ou plus, chaque hameau se trouvant à moins de 200 mètres d'un autre hameau de la HA.



Étendue de localité: CIESIN, Columbia University dans le cadre du programme GRID3; Empreinte du bâtiment: Ecopia Landbase Africa propulsé par Maxar; Image: Esri

*Définitions basées sur Barau, I., M. Zubairu, M.N. Mwanza et V.Y. Marin. 2014. Améliorer la couverture vaccinale contre la poliomyélite au Nigéria grâce à l'utilisation de la technologie du système d'information géographique. Le Journal des maladies infectieuses (novembre) : S102-S110.

Validation et travail du terrain

À ce stade, GRID3 a collecté un ensemble de points de données qui ont été nettoyés, normalisés et placés dans un schéma de base de données standard, ainsi qu'un ensemble de polygones qui ont été créés en collaboration avec l'un des partenaires de GRID3. Cependant, aucun ensemble de données n'est parfait et l'étape suivante consiste à valider les données.

Tout d'abord, GRID3 compare les polygones aux points et identifie ce qui est manquant ou incorrect dans les points, en vérifiant les défauts d'alignement, les erreurs potentielles et les lacunes dans les données. GRID3 transmet ces informations au gouvernement partenaire pour voir s'il peut combler les lacunes, ainsi que pour vérifier les sources non gouvernementales comme OpenStreetMap¹². Si aucune de ces sources ne dispose d'informations nécessaires, GRID3 peut travailler avec un partenaire déjà présent sur le terrain pour collecter les données manquantes, ou effectuer lui-même le travail de terrain.

Ensuite, la comparaison est inversée : les points sont comparés aux polygones pour voir ce qui manque dans les polygones. GRID3 inspecte visuellement ces points par rapport à l'imagerie satellite, ce qui permet 1) d'identifier les localités qui ont été manquées par l'extraction, puis 2) de créer des polygones pour eux.

GRID3 répète l'exercice de liaison des noms de lieux avec l'étendue des localités pour obtenir la meilleure couche de localité possible qui intègre toutes les entrées. Les contrôles visuels et logiques sont automatisés afin de signaler les incohérences entre les deux couches (par exemple, les endroits qui sont affichés dans la grille de raster¹³ mais qui ne font pas partie de la localité de base, et vice versa). Les conflits sont résolus en coordination avec les acteurs locaux. Le cas échéant, une consultation générale avec les parties prenantes locales est effectuée pour valider les résultats.

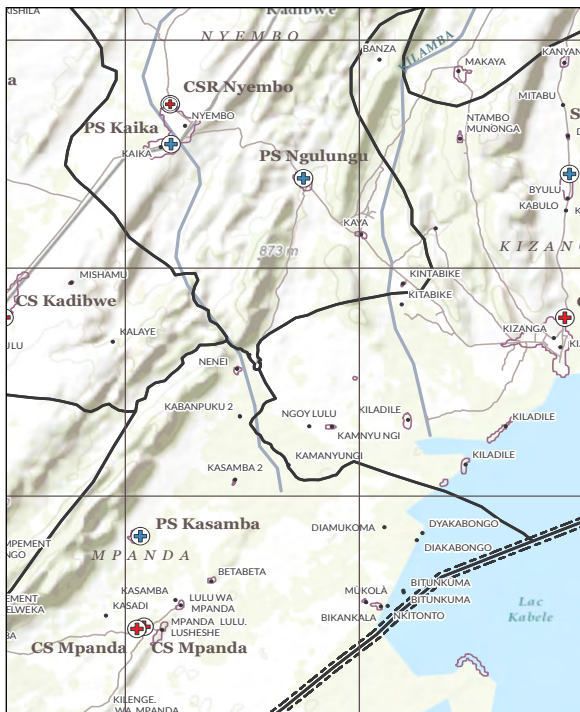
12. Voir OpenStreetMap. <https://www.openstreetmap.org/about>.

13. "Dans sa forme la plus simple, les données raster se composent d'une matrice de carreaux (ou pixels) organisée en lignes et colonnes (ou une quadrillage/carroyage) où chaque carreau contient une valeur représentant des informations, telles que la température. Les données raster sont des photographies aériennes numériques, des images provenant de satellites, des images numériques ou même des cartes numérisées. Voir la référence de l'Environmental Systems Research Institute ci-dessus.

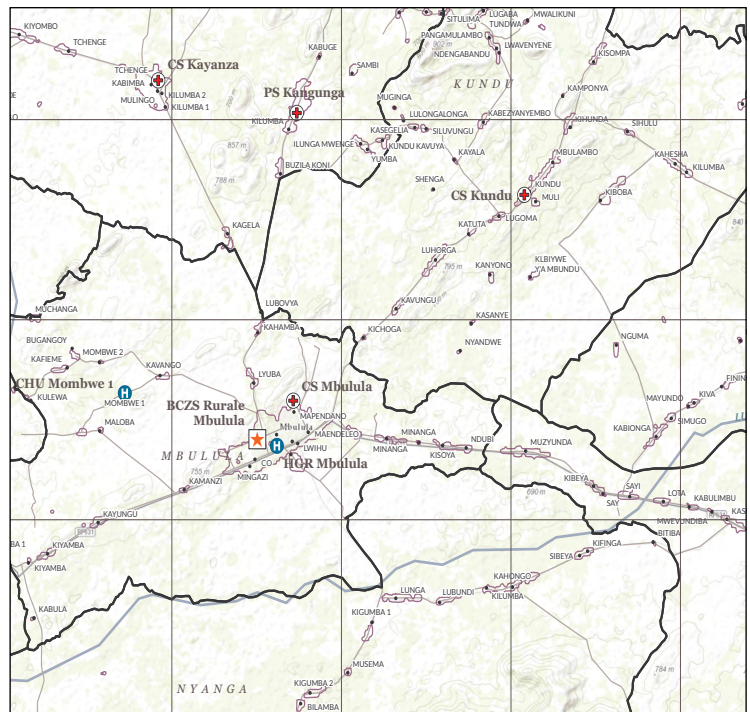
Couche de localité exhaustive

Enfin, GRID3 crée une couche de localité exhaustive. Cette couche combine les étendues de localité avec les attributs (par exemple, les noms de lieux ou les unités administratives) de la couche de points. La couche de localité exhaustive peut se présenter sous la forme d'une couche de points ou d'une couche de polygones. Cette couche est envoyée à une série de réviseurs. Après avoir effectué les ajustements nécessaires, GRID3 finalise la couche de données.

ZS de Bukuma, province du Haut-Lomami, RDC



ZS de Mbulula, province de Tanganyika, RDC



Sources: Esri, Airbus DS, USGS, NGA, NASA, CGIAR, N Robinson, NCEAS, NLS, OS, NMA, Geodatastyrelsen, Rijkswaterstaat, GSA, GEoland, FEMA, Intermap et la communauté des utilisateurs SIG

13. "In its simplest form, a raster consists of a matrix of cells (or pixels) organised into rows and columns (or a grid) where each cell contains a value representing information, such as temperature. Rasters are digital aerial photographs, imagery from satellites, digital pictures, or even scanned maps." See Environmental Systems Research Institute reference above.

GRID3 Settlement Mapping Process

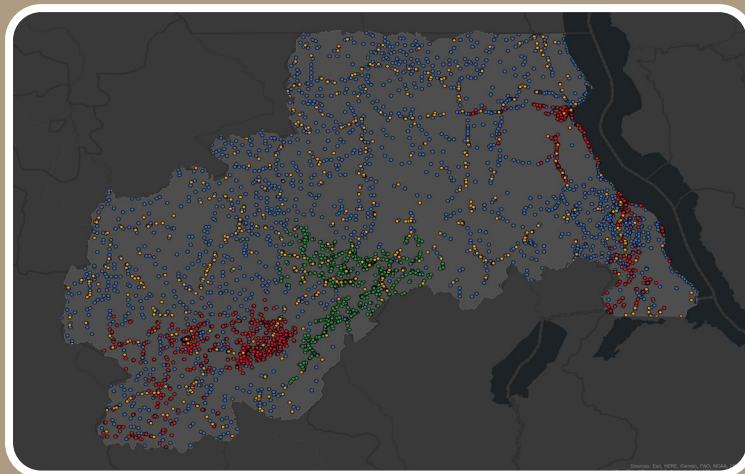
1

Évaluer les données existantes

Obtenir des données de points et de polygones des autorités gouvernementales

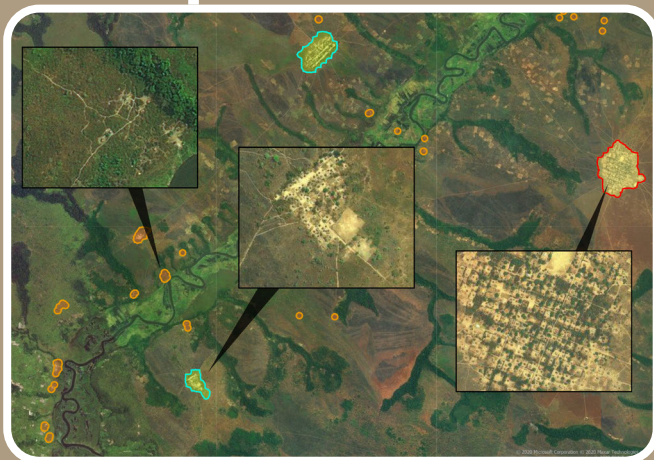
Si les données sont incomplètes, remplir les lacunes dans les données avec une variété de sources supplémentaires, y compris celles d'autorités gouvernementales, d'ONG, d'entreprises privées, par des groupes locaux, des entités des Nations Unies et des prestataires indépendants.

- Hameaux
- Petites zones de peuplement
- Zones construites

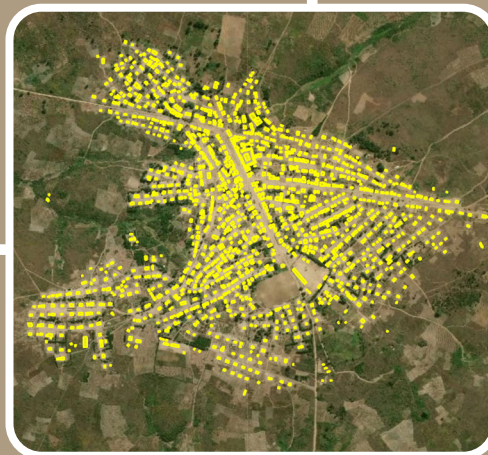


- Source 1 ■
- Source 2 ■
- Source 3 ■
- Source 4 ■

Sources: Esri, HERE, Garmin, FAO, NOAA, USGS, contributeurs OpenStreetMap et la communauté des utilisateurs SIG



Sources: © 2020 Microsoft Corporation © 2020 Maxar © CNES (2020) Distribution Airbus DS



Sources: Esri, Maxar, GeoEye, Earthstar Geographics, CNES/Airbus DS, USDA, USGS, AeroGRID, IGN et la communauté des utilisateurs SIG

2

Créer des bases de données (Emplacements et régions)

Standardiser quels mots sont utilisés pour représenter différents types de données.

Regrouper les éléments qui tombent dans un rayon de 25 à 100 m l'un de l'autre, ou en utilisant la densité des bâtiments pour dessiner des contours autour de groupes d'empreintes de bâtiments.

3

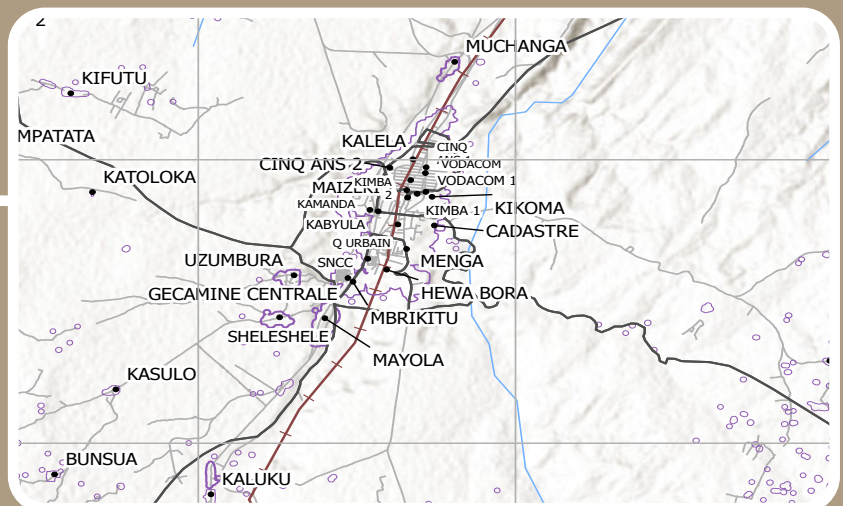
Validation et travail sur le terrain

Comparer les polygones avec les points et identifier ce qui est manquant ou incorrect dans les points, en vérifiant les désalignements, les potentielles erreurs et lacunes dans les données; et vice versa.

4

Couche de localité

Cette couche combine les étendues de localité avec les attributs de la couche de points. La couche de localité exhaustive peut être sous la forme d'une couche de points ou d'une couche de polygones.



Sources : Esri, Airbus DS, USGS, NGA, NASA, CGIAR, N Robinson, NCEAS, NLS, OS, NMA, Geodastyrelsen, Rijkswaterstaat, GSA, GEoland, FEMA, Intermap et la communauté des utilisateurs SIG

Approches de classification

Le monde continue de s'urbaniser: 55 % de la population mondiale vit aujourd'hui dans des zones urbaines, un chiffre qui devrait atteindre 68 % d'ici 2050.¹⁴ Mais les zones urbaines ne sont pas homogènes (que ce soit entre localités ou au sein des zones construites) et donc leurs besoins et les défis de développement ne sont pas les mêmes. Les progrès informatiques nous permettent d'acquérir une compréhension plus nuancée des zones urbaines. Comme mentionné dans la section précédente, les algorithmes d'apprentissage automatique sont capables de détecter et de cartographier automatiquement toutes les caractéristiques d'un bâtiment à partir d'images satellite et aériennes à haute résolution, et ces empreintes fournissent à leur tour une carte haute résolution des zones potentiellement habitées (identifiables par la présence de structures détectables).

GRID3 explore de nouvelles méthodes pour identifier ces modèles et extraire des informations supplémentaires qui peuvent améliorer notre compréhension des localités à l'échelle des structures individuelles et des quartiers.

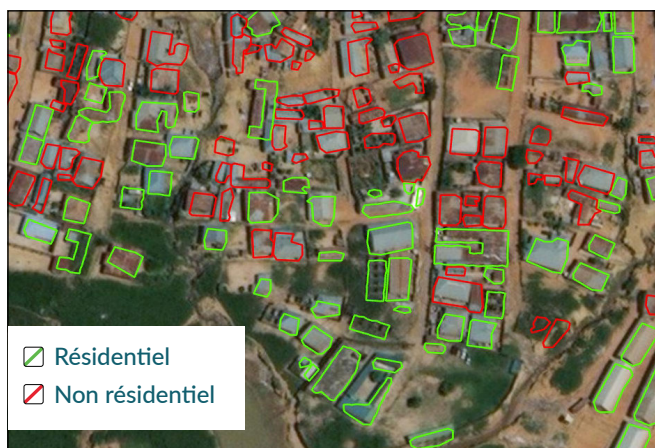
Alors que de tels ensembles de données peuvent fournir des détails sur la présence et l'emplacement des localités, ils manquent souvent d'informations qui peuvent différencier un lieu dans un noyau urbain d'une périphérie croissante d'une ville, ou une périphérie croissante d'une ville d'un village éloigné. Cependant, la taille et la forme des structures, ainsi que la disposition et la position des structures (par rapport à d'autres infrastructures ou caractéristiques), peuvent transmettre des informations sur différentes utilisations des terres ou activités économiques. GRID3 explore de nouvelles méthodes pour identifier ces schémas et extraire des informations supplémentaires qui peuvent améliorer notre compréhension des localités à l'échelle des structures individuelles et des quartiers. Plus précisément, GRID3 développe deux approches (décrites ci-dessous) qui utilisent les empreintes de bâtiments, des couches de données géospatiales et des algorithmes d'apprentissage automatique pour classer les structures et les zones locales. De telles images améliorées des localités peuvent à leur tour améliorer notre étude des modèles d'urbanisation, améliorer notre compréhension des similitudes et des différences trouvées dans les contextes locaux, aider à distinguer les zones urbaines et rurales, et rendre la modélisation de la population et les enquêtes de terrain plus robustes (les données sur les localités sont une contribution clé aux modèles de population ascendants de GRID3).

Classification des localités en fonction des types de bâtiment (résidentiel ou non résidentiel)

Le niveau le plus élémentaire de classification des localités est par type de bâtiment (résidentiel ou non résidentiel). Les caractéristiques de chaque structure (telles que leur taille et leur forme ainsi que la présence d'infrastructures et d'autres caractéristiques d'intérêt) peuvent donner des indices sur l'utilisation d'un bâtiment.

14. Voir United Nations Department of Economic and Social Affairs. 2018. 2018 Revision of World Urbanization Prospects.

Pour prédire avec précision le type de bâtiment, GRID3 utilise plusieurs méthodes d'apprentissage automatique. Chaque méthode fournit une prédiction et ces prédictions sont combinées entre les modèles dans un ensemble¹⁵ qui permet d'améliorer la prédiction globale. Les méthodes statistiques utilisées sont des modèles supervisés; c'est-à-dire que les modèles sont entraînés avec de petites quantités de données pré-classifiées. GRID3 utilise les données d'OpenStreetMap et d'autres ensembles de données de bâtiments étiquetés.



© 2020 Maxar Technologies, Ecopia.AI



© 2020 Maxar Technologies, Ecopia.AI

Une approche d'apprentissage automatique basée sur une classification binaire et basée sur des objets est utilisée pour entraîner un modèle afin de prédire les structures de bâtiments en tant que types résidentiels ou non résidentiels.

La distinction binaire entre résidentiel et non résidentiel est importante pour soutenir des modèles de population plus précis et pour guider les services qui doivent trouver des zones résidentielles. Cependant, ces deux classes en disent peu sur le type de quartier dans lequel se trouve un bâtiment et, par extension, sur la manière dont ce quartier s'intègre dans les schémas de localités de la zone géographique plus large. Dans ce cas, GRID3 développe un deuxième modèle de classification au niveau de la zone.

Trouver des zones similaires à partir des schémas de construction (classification des quartiers)

GRID3 développe des méthodes qui peuvent aller au-delà de la classification résidentielle/non résidentielle. A nouveau, GRID3 commence avec les ensembles de données d'empreinte de bâtiment à haute résolution; mais dans ce cas, GRID3 se concentre sur les modèles qui peuvent être observés dans les bâtiments (plutôt que sur les structures individuelles). Lorsqu'elles sont vues ensemble, la taille, la densité, les formes et les orientations des structures créent une « texture » visuelle qui varie à travers le paysage, indiquant des différences dans l'utilisation des terres ou le type de quartier.

15. Une combinaison de prédictions entre les modèles utilisée pour améliorer la prédiction globale et pour éviter « d'enchevêtrer » tout modèle individuel dans l'ensemble.

Pour créer ces classifications de quartier, le motif et la texture sont quantifiés à plusieurs échelles spatiales sur toute une région d'intérêt en calculant la variation des tailles et des formes, ainsi que de nombreuses autres mesures sur une grille régulière d'emplacements. Des modèles de mélange gaussien et d'autres méthodes d'apprentissage automatique sont utilisés pour trouver des groupes de schémas similaires et pour classer les types sur une surface carroyée de 100 x 100 mètres de résolution.¹⁶

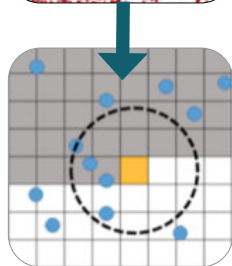
GRID3 continue de développer cette ligne de recherche - les efforts du programme incluront à terme l'élargissement de la gamme d'attributs générés par (et donc les données d'entrée utilisées) le modèle de classification, ce qui améliorera le pouvoir prédictif. Bien que la recherche sur la classification en soit encore à ses débuts, elle offre déjà une perspective radicalement différente - et améliorée - sur les empreintes des bâtiments et aide à situer ces empreintes dans des schémas de localités plus larges.

Quantification et classification des modèles



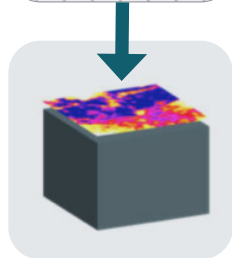
Étape 1 : Couche de localité

La couche de localité est composée d'empreintes de bâtiments; le but est de classer les zones de schémas similaires.



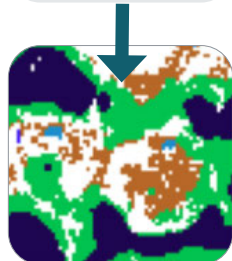
Étape 2 : Calculs de fenêtre mobile

À chaque emplacement de 100 m x 100 m sur la carte de localité, une série de calculs quantifie la taille, la forme, le regroupement des bâtiments et d'autres mesures des motifs. Les calculs sont effectués dans des «fenêtres» de tailles différentes autour des emplacements.



Étape 3 : Statistiques de fragmentation

Collectivement, les calculs sont appelés « pile de statistiques de fragmentation ». A chaque endroit sur la carte, il y a une plage de valeurs décrivant les modes de localité locaux.



Étape 4 : Classement non supervisé

Des méthodes de classification non supervisées sont utilisées pour explorer les statistiques de fragmentation et pour trouver des regroupements de données similaires. Ces regroupements sont cartographiés vers les emplacements du monde réel pour produire un classement des localités.

Images créées par Warren C. Jochem, en utilisant les données de construction d'Ecopia Landbase Africa

16. GRID3 fonctionne sur un carroyage 3 secondes d'arc, ce qui se rapproche en réalité de 90 m x 90 m. GRID3 arrondit nominalement les chiffres jusqu'à 100m x 100m.

Applications

Les produits de localité de GRID3 ont un large éventail d'utilisations. Ils peuvent générer des données open source, fournir un soutien aux applications de données pour assurer un impact efficace, ou permettre de renforcer les fondations géospatiales nationales pour un futur développement fondé sur des décisions humanitaires. Ces productions s'intègrent également dans les autres travaux du programme sur le recensement, comme les modèles de population, la délimitation des limites et frontières et l'identification des emplacements des infrastructures.

Les données sur les localités de GRID3 ont déjà commencé à avoir des impacts dans plusieurs domaines différents du développement durable. Certains de ces impacts sont décrits ci-dessous.

Prestation de services et opérations

Les campagnes de vaccination, l'acheminement de la chaîne d'approvisionnement et la réponse aux catastrophes/réduction des risques, peuvent tous être renforcés à l'aide des données GRID3 sur les localités.

Par exemple, dans l'est de la République démocratique du Congo (RDC), l'instabilité politique et les insuffisances d'infrastructures associées ont rendu difficile la création de cartes de localités précises pour la région. Alors que la plupart des localités de l'est de la RDC sont visibles sur les images satellite, beaucoup n'ont pas de noms qui leur sont attachés. Pour changer cela, la Fondation Bill & Melinda Gates, GRID3, PATH¹⁷ et Novel-T¹⁸ se sont associés à la campagne de vaccination contre la polio du ministère de la Santé de la RDC pour identifier les noms de toutes les localités dans un ensemble sélectionné de zones de santé prioritaires dans les provinces du Haut-Lomami et Tanganyika. Des smartphones ont été fournis aux vaccinateurs pour enregistrer les données nécessaires sur la vaccination des enfants ainsi que saisir le nom des localités où ils se sont rendus, la zone de santé et l'unité administrative dans laquelle se trouvait cette localité. Avec environ 300 vaccinateurs voyageant pendant trois à quatre jours dans chacune des zones de santé sélectionnées, GRID3 s'est appuyé sur les ressources humaines considérables déployées sur le terrain pour augmenter substantiellement, et avec succès, le nombre de noms des localités collectés, ainsi que pour améliorer les connaissances sur les emplacements des localités. Les données dérivées de ce projet seront utilisées non seulement par le ministère de la Santé pour les futures campagnes de vaccination contre la poliomyélite, mais seront également susceptibles d'améliorer la vaccination de routine, les campagnes de masse contre la rougeole et la lutte contre le paludisme, et la prestation des services de santé en général. Les données serviront de carte de base à toute ONG ou organisation humanitaire travaillant dans ces zones de santé, ainsi qu'à l'Institut National de la Statistique (INS) qui s'apprête à réaliser le premier recensement national de la RDC depuis 1984.

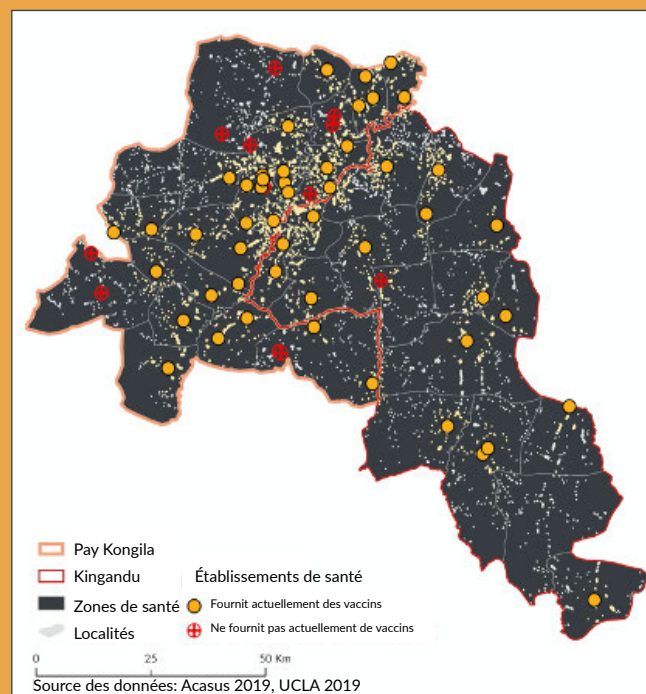
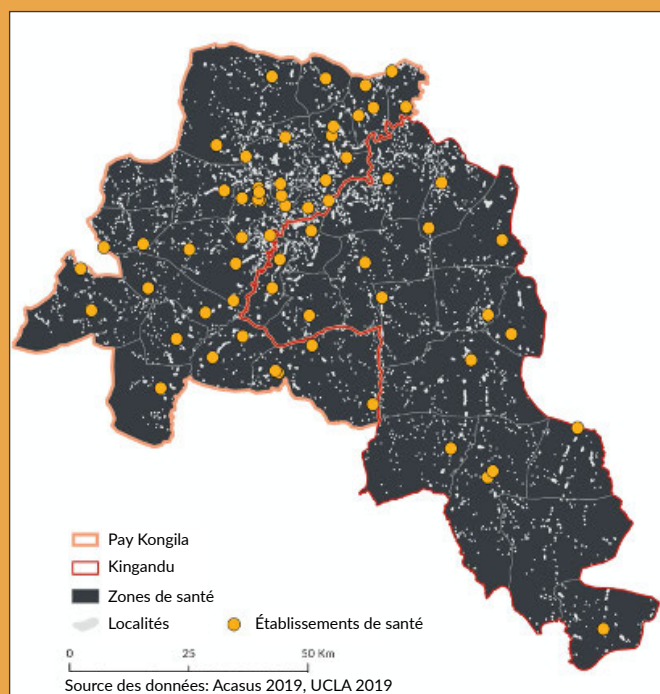
17. PATH: The Bill & Melinda Gates Foundation's logistical/coordinating partner in Haut-Lomami and Tanganyika provinces. Voir <https://www.path.org/where-we-work/democratic-rep-of-the-congo/>.

18. Novel-T: partenaire technique configurant les téléphones Android et fournissant l'application de collecte de données. Voir <http://www.novel-t.ch/fr#Accueil>.

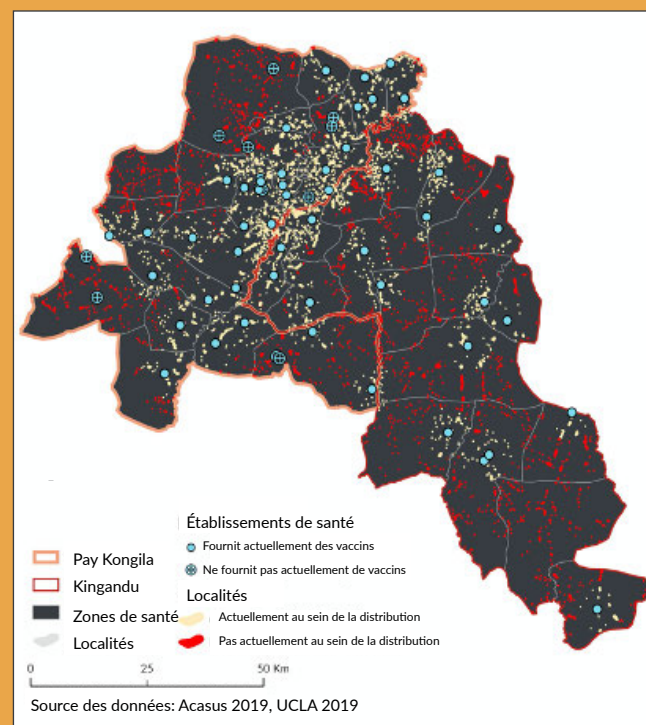
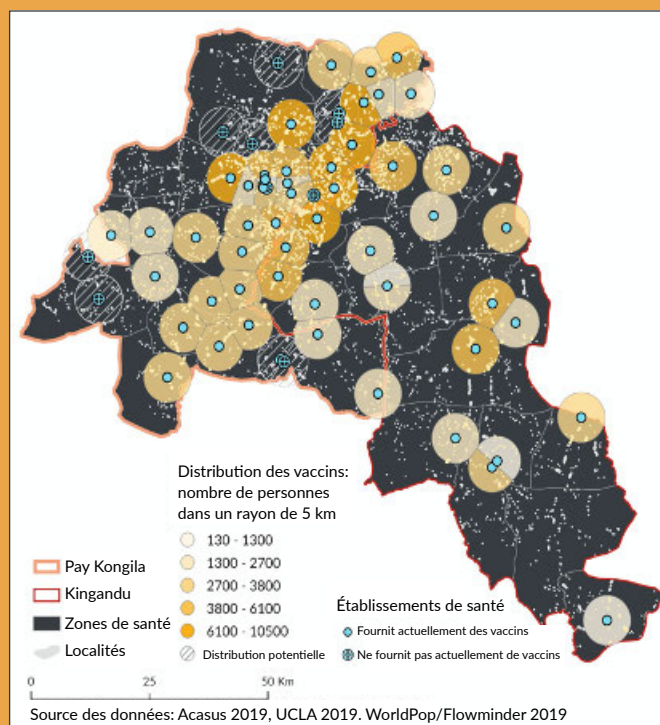
GRID3 a permis aux agents de santé d'améliorer l'exhaustivité des localités incluses dans leurs micro-plans, ce qui a permis d'inclure des populations cibles supplémentaires dans les futures campagnes de vaccination.

GRID3 a également amélioré la cartographie des localités dans l'est de la RDC en organisant des sessions de cartographie participatives. Les agents de santé d'une zone de santé spécifique, idéalement lors de leur réunion mensuelle ou de leurs activités de routine, ont assisté à la session initiale de deux jours pour identifier et cartographier les villages sans nom. Pour guider ce processus, deux cartographes GRID3 avec une formation SIG ont reçu toutes les données qui avaient été collectées sur la région jusqu'à présent (via des images satellite téléchargées sur les ordinateurs portables des stagiaires, ainsi que les extractions de caractéristiques des localités dont la qualité avait été améliorée par les chercheurs GRID3 basés aux États-Unis). Les cartographes GRID3 ont utilisé les connaissances locales des agents de santé ainsi que leurs propres compétences techniques pour combiner ces couches et localiser les localités. Pour identifier les villages inconnus ou les villages dont l'emplacement et les noms restaient incertains à la fin de la session de cartographie participative, GRID3 a équipé les agents de santé de smartphones et les a formés sur une application de collecte de données facile à utiliser fournie par Novel-T. Les agents de santé sont retournés dans les zones de santé qui leur avaient été assignées et ont collecté des informations sur les localités restées anonymes. Les deux cartographes GRID3 ont collecté ensuite ces données supplémentaires et créé une couche de localité exhaustive. Une fois ces travaux terminés, une réunion de validation avec le Médecin Chef de Zone a été organisée afin d'obtenir l'approbation des nouveaux noms des localités. En créant une carte des localités plus complète dans les deux provinces, GRID3 a permis aux agents de santé d'améliorer l'exhaustivité des localités incluses dans leurs micro-plans, permettant d'inclure des populations cibles supplémentaires dans les futures campagnes de vaccination.

Améliorer la prestation des services de santé au Kwilu, en RDC



Une planification efficace de la livraison des vaccins utilise des couches de données combinées et mises à jour comprenant : les noms de lieux, l'étendue des localités, les estimations de population, l'emplacement et l'infrastructure disponible dans les centres de santé.



Lorsque toutes les informations sont consolidées, il est possible d'observer et de quantifier les localités et les populations hors de portée des zones d'intervention des centres de santé. Grâce à ces informations, les postes de santé mobiles peuvent être correctement positionnés et planifiés pour une couverture efficace.

GRID3 a effectué un travail similaire au Nigeria.¹⁹ Depuis 2012, la Fondation Bill & Melinda Gates a soutenu les efforts d'éradication de la poliomyélite via l'application de la technologie SIG - cette méthode a permis aux équipes de vaccination de surveiller correctement la couverture des localités, de réduire le nombre de localités manqués et d'améliorer les performances des équipes. Des données détaillées (y compris les limites des quartiers, les localités et les points d'intérêt²⁰ ont été recueillies dans dix États du nord du Nigéria où d'importantes épidémies de poliomyélite ont eu lieu. Ce travail de terrain, en plus de l'extraction d'information à partir d'images satellites, a permis de générer des cartes de micro-plan SIG pour les équipes de vaccination et de visualiser les localités. Le tableau de bord du système de suivi des vaccinations (VTS) a été développé pour suivre les équipes de vaccination et produire des rapports identifiant les localités manqués, ce qui a permis aux superviseurs d'évaluer les performances de l'équipe et d'identifier les localités qui devaient être revisités pour les vaccinations de suivi. Une enquête de couverture post-campagne a révélé que les états du Nigéria où les micro-plans SIG et le VTS étaient utilisés avaient de meilleurs résultats de vaccination et une immunité de la population plus élevée. En conséquence, les efforts de collecte de données ont été étendus aux autres états en 2016 et ces technologies ont également été appliquées dans les campagnes ultérieures de vaccination contre la poliomyélite. De fin 2019 au début 2020, le VTS a été amélioré et utilisé pour des campagnes supplémentaires, notamment contre la fièvre jaune, la rougeole et d'autres maladies.

Informations et recherches

Les données sur les localités de GRID3 et les efforts associés fournissent des informations cruciales pour la recherche sur le développement, notamment: l'identification des domaines prioritaires pour les interventions de développement, la réalisation d'études de marché et les emplacements pour des investissements potentiels, et la génération de données sur la trajectoire et l'ampleur de la croissance des localités.

Avec le soutien de GRID3, la Zambie progresse vers la création d'une couche géospatiale représentant la répartition de ses localités. L'amélioration de la cartographie des localités est une question d'une importance particulière dans le pays; La Zambie est profondément rurale et se compose majoritairement de terres appartenant à des traditions. Un petit pourcentage de ces terres a été officiellement cartographié. Pour compliquer encore les choses, il peut être difficile de savoir ce qu'il faut cartographier en premier lieu (un village zambien peut parfois être aussi petit qu'un seul ménage). Pour relever ces défis, GRID3 a travaillé en étroite collaboration avec le gouvernement zambien pour créer une couche de localité exhaustive qui servira de référentiel le plus complet à ce jour pour les noms et les emplacements des localités du pays.

Pour créer un niveau appartenant au gouvernement, GRID3 a collecté des points de données (à partir de diverses sources gouvernementales. La majorité des noms de localité ont été extraits d'un exercice de cartographie du gouvernement en 2010, au cours duquel des points d'intérêt (POI) à l'échelle nationale ont été collectés. GRID3 a catégorisé les POI de l'exercice de cartographie par type afin d'extraire les points qui représentaient les villages. GRID3 a standardisé des données provenant de diverses sources afin de les compiler en un seul fichier.

19. Pour plus d'informations sur les points suivants, voir Touray, K., P. Mkanda, S.G. Tegegn, P. Nsubuga, T.B. Erbetto, R. Banda, A. Etsano, F. Shuaib et R.G. Vaz. 2016. Tracking Vaccination Teams During Polio Campaigns in Northern Nigeria by Use of Geographic Information System Technology: 2013-2015. *Journal of Infectious Diseases* (May): S67-72).

20. Les points d'intérêt (POI) font référence à des points géo-référencés qui représentent des infrastructures, des bâtiments et des points de repère. Ces points localisent les structures et les services essentiels à la santé et au bien-être de la société, tels que, entre autres, les établissements de santé, les écoles, les marchés, les banques, les entrepôts et les puits.

Une collaboration étroite avec le gouvernement zambien garantit non seulement que GRID3 a un meilleur accès aux données et aux ressources, mais plus important encore, cela favorise l'appropriation zambienne du projet. Un large éventail d'avantages résultera de cet ensemble de données encore en développement, dont: une aide à la planification d'un recensement à l'échelle nationale, des campagnes de vaccination plus efficaces et une capacité accrue de réponse aux catastrophes.

Autres impacts des travaux de GRID3 sur les localités :



Conception de programmes et d'interventions

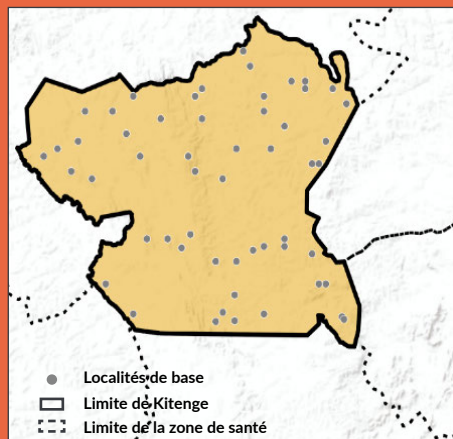
Les données de GRID3 peuvent aider à identifier les emplacements optimaux pour de nouvelles installations publiques ou services publics; établir une base pour des plans de travail rationalisés et plus efficaces; et aider à favoriser une conception centrée sur la population.



Avantages à long terme

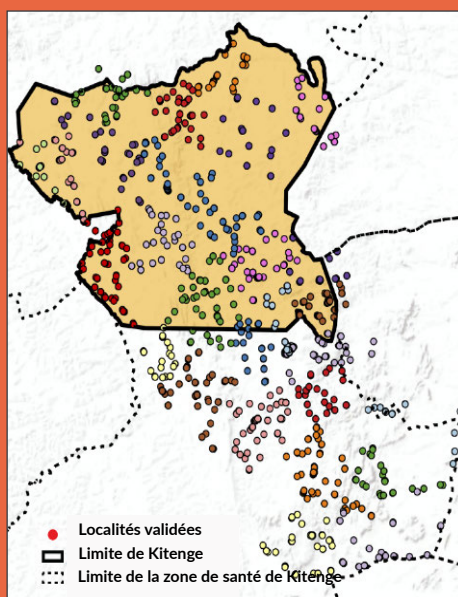
Au fil du temps, les données de GRID3 et les efforts qui y sont associés peuvent conduire à de meilleurs résultats de développement à long terme, y compris la réduction de divers types de taux de mortalité et d'infection, une augmentation des investissements et de la croissance de l'emploi, un accès accru aux réseaux et aux services, et d'autres éléments relatifs aux Objectifs de Développement durable des Nations Unies.

Étude de cas : Renforcement des campagnes de vaccination avec de meilleures cartes d'établissement en RDC



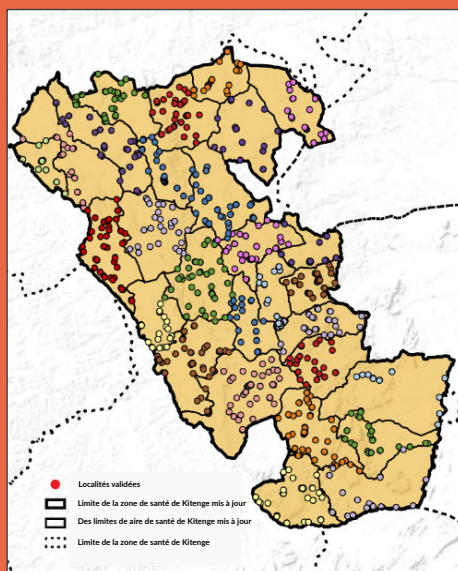
Souvent, les praticiens sont obligés de se contenter de cartes de localités incomplètes ou inexactes. Cette carte de référence de l'est de la RDC a été considérée comme inadéquate pour soutenir une récente campagne de vaccination contre la poliomyélite.

Sources: Programme Elargi de Vaccination (PEV), Ministère de la Santé Publique, RDC, 2019; Walsh, Chen, Wu, Simbila et al, 2018; Novel-T (VTS / POI Tracker), 2019; Bureau Central Recensement, 2018; Référentiel Géographique Commun, 2018; Système National d'Information Sanitaire, 2018; UCLA, 2018; UNICEF, 2018; OMS, 2018; Fondation Bill & Melinda Gates, 2019



GRID3 et des partenaires locaux ont rassemblé toutes les sources de données de localité GIS ; évalué leur qualité à l'aide d'images satellite et d'extractions de caractéristiques de bâtiments ; et les ont consolidées dans un nouveau fond de carte harmonisé. Ensuite, les données satellitaires et l'apprentissage automatique ont été utilisées pour localiser des localités auparavant non cartographiées. Le résultat a été une carte des localités plus fiable que celle que les vaccinateurs contre la poliomyélite pouvaient utiliser dans leurs efforts pour atteindre tout le monde dans la région.

Sources: Programme Elargi de Vaccination (PEV), Ministère de la Santé Publique, RDC, 2019; Walsh, Chen, Wu, Simbila et al, 2018; Novel-T (VTS / POI Tracker), 2019; Bureau Central Recensement, 2018; Référentiel Géographique Commun, 2018; Système National d'Information Sanitaire, 2018; UCLA, 2018; UNICEF, 2018; OMS, 2018; Fondation Bill & Melinda Gates, 2019



Les campagnes de vaccination sont organisées en fonction des limites de santé. En identifiant quelles zones de santé sont associées à chaque localité, il devient possible de corriger les erreurs de limites et de délimiter les limites manquantes. Dans ce cas, une erreur majeure dans la limite de la zone de santé de Kitege a été corrigée et des limites de zone de santé validées ont été générées. Ces limites claires et validées permettent aux gestionnaires locaux de planifier des campagnes de vaccination qui atteignent efficacement tout le monde.

Sources: Programme Elargi de Vaccination (PEV), Ministère de la Santé Publique, RDC, 2019; Walsh, Chen, Wu, Simbila et al, 2018; Novel-T (VTS / POI Tracker), 2019; Bureau Central Recensement, 2018; Référentiel Géographique Commun, 2018; Système National d'Information Sanitaire, 2018; UCLA, 2018; UNICEF, 2018; OMS, 2018; Fondation Bill & Melinda Gates, 2019

Conclusion

Les activités de GRID3 ajoutent de la valeur aux efforts visant à comprendre les localités, à la fois en améliorant les techniques de collecte/d'analyse des données et en améliorant les données elles-mêmes. En combinant diverses données, en renforçant les capacités et en comprenant les besoins, GRID3 aide à établir de nouveaux flux de données importants pour les emplacements et les utilisations des bâtiments non seulement pour ses pays partenaires, mais potentiellement pour d'autres pays à des stades de développement similaires dans le monde.

De nombreux défis doivent encore être relevés dans la campagne de collecte de données complètes et exploitables sur les localités. Ceux-ci incluent :

- Capturer les localités temporaires (tels que les camps de personnes déplacées à l'intérieur du pays, de réfugiés et de travailleurs migrants), déplacements temporaires de population (par ex. déplacement après une catastrophe), déplacements de population à long terme (par ex. croissance urbaine), campements nomades, diurnes et saisonniers .
- Distinguer les populations résidentielles des populations non résidentielles.
- Réconcilier les différences entre les conceptions officielles et locales des lieux, des noms de lieux et des limites des localités. Actuellement, l'approche de GRID3 est essentiellement une conception spatiale des localités et néglige les expériences de terrain des lieux.
- Negotiating political sensitivities and security issues both within and between countries
Négocier les sensibilités politiques et les problèmes de sécurité à l'intérieur et entre les pays autour des lieux, des noms de lieux et des limites des localités . GRID3 envisage des questions telles que : Quels endroits ont été inclus dans les anciennes couches de données, et devraient-ils/ être inclus dans les nouvelles? Quels noms (et dans quelle langue) les lieux sont-ils donnés Dans quelle mesure les communautés autochtones et ethniques minoritaires sont-elles représentées ?

Dans les années à venir, GRID3 s'efforcera de relever ces défis en continuant à travailler avec des partenaires pour rendre ses données actuelles plus conviviales et en recherchant de nouvelles façons de collecter et de valider ces informations.

Remerciements et attribution

Auteurs contributeurs

Ce document a été préparé par Corey Sobel, Jolynn Schmidt, Warren C. Jochem, Kevin Tschirhart, Emilie Schnarr et Olena Borkovska, avec des contributions supplémentaires fournies par Susana Adamo, Sandra Baptista, Sophie Delaporte, Justine Dowden, Paola Kim-Blanco, Matthew Heaton, Attila Lazar, Etienne Leue, Marc Levy, Chris Lloyd, Maxwell Madzikanga, Polly Marshall, Chisimdi Onwuteaka, Silvia Renn, Cathy Riley, Markus Walsh, Greg Yetman et Apphia Yuma.

Recommended Citation

Center for International Earth Science Information Network (CIESIN), Columbia University; Center for International Earth Science Information Network (CIESIN), Columbia University; Flowminder Foundation; United Nations Population Fund (UNFPA); WorldPop, University of Southampton. 2020. Cartographie et classification des localités. Palisades, NY: Georeferenced Infrastructure and Demographic Data for Development (GRID3). [DOI TBA]. Consulté JOUR MOIS ANNÉE.

License and Copyright

Copyright ©2020. The Trustees of Columbia University in the City of New York. This document is licensed under the Creative Commons Attribution 4.0 International License.

Works Cited

Barau, I., M. Zubairu, M.N. Mwanza, and V.Y. Seaman. 2014. Improving polio vaccination coverage in Nigeria through the use of geographic information system technology. *The Journal of Infectious Diseases* (November): S102–S110.

Environmental Systems Research Institute. “What is raster data?”. <https://desktop.arcgis.com/en/arcmap/10.3/manage-data/raster-and-images/what-is-raster-data.htm>.

Maxar Technologies, Inc. and Ecopia Tech Corporation. 2020. Ecopia Landbase Africa powered by Maxar. [DigitizeAfrica.ai](https://digitizeafrica.ai).

National Institute of Standards and Technology. “Probability model”. <https://csrc.nist.gov/glossary/term/Probability-model>.

Novel-T. <http://www.novel-t.ch/en#Home>.

Oak Ridge National Laboratory. <https://www.ornl.gov/>.

OpenStreetMap. <https://www.openstreetmap.org/about>.

PATH. <https://www.path.org/where-we-work/democratic-rep-of-the-congo/>.

Programme Elargi de Vaccination (PEV), Ministry of Public Health, DRC, 2019; Walsh, Chen, Wu, Simbila, et al, 2018; Novel-T (VTS / POI Tracker), 2019; Bureau Central Recensement, 2018; Référentiel Géographique Commun, 2018; Système National d'Information Sanitaire, 2018; UCLA, 2018; UNICEF, 2018; WHO, 2018; Bill and Melinda Gates Foundation, 2019.

Touray, K., P. Mkanda, S.G. Tegegn, P. Nsubuga, T.B. Erbetto, R. Banda, A. Etsano, F. Shuaib, and R.G. Vaz. 2016. Tracking Vaccination Teams During Polio Campaigns in Northern Nigeria by Use of Geographic Information System Technology: 2013-2015. *Journal of Infectious Diseases* (May: S67-72).

UCLA Institute for Digital Research & Education Statistical Consulting. "Zero-Inflated Poisson Regression | R Data Analysis Examples". <https://stats.idre.ucla.edu/r/dae/zip/>.

United Nations Department of Economic and Social Affairs. 2018. 2018 Revision of World Urbanization Prospects.

United States Geological Service. "What is remote sensing and what is it used for?" https://www.usgs.gov/faqs/what-remote-sensing-and-what-it-used?qt-news_science_products=7#qt-news_science_products.

Lectures complémentaires

Dougherty, L., M. Abdulkarim, F. Mikailu, T. Usman, K. Owolabi, K. Gilroy, A. Naiya et al. 2019. From paper maps to digital maps: enhancing routine immunisation microplanning in Northern Nigeria. *BMJ Global Health* 4.

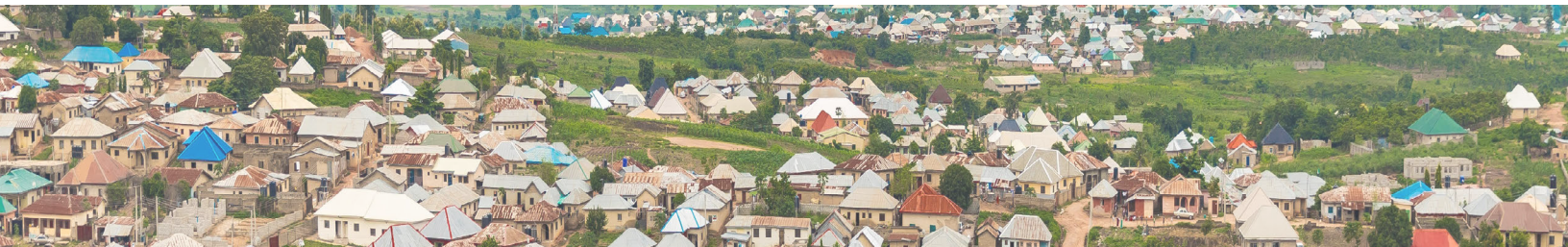
Jochem, W.C., T.J. Bird, and A.J. Tatem. 2018. Identifying residential neighbourhood types from settlement points in a machine learning approach. *Computers, Environment and Urban Systems* 69: 104-113.

United Nations Children's Fund. 2018. Guidance on the use of geospatial data and technologies in immunization programs: overview and managerial considerations for in-country strengthening (October).

Weber, E.M., V.Y. Seaman, R.N. Stewart, T.J. Bird, A.J.Tatem, J.J. McKee, B.L. Bhaduri, J.J. Moehl, and A.E. Reith. 2018. Census-independent population mapping in northern Nigeria 204: 786-798.

Historique des versions papier

- Table des matières mise à jour à la page 2
- Suppression de la carte sous Microplan de vaccination contre la polio pour le quartier de Dundubus dans l'État de Jigawa, au Nigéria à la page 19. Il s'agit d'une carte des clusters à poste fixe contre la rougeole créée par Novel-T.
- Logos mis à jour et section À propos des partenaires à la page 25



GRID3 (Geo-Referenced Infrastructure and Demographic Data for Development) travaille avec les pays pour générer, valider et utiliser les données géospatiales sur la population, les lieux de peuplement, les infrastructures, et les limites infranationales. GRID3 combine l'expertise des partenaires du gouvernement, des Nations Unies, d'universitaires et le secteur privé pour concevoir des solutions géospatiales adaptables et pertinentes en fonction des capacités et des besoins de développement de chaque pays.

 grid3.org

 info@grid3.org

 [@GRID3Global](https://twitter.com/GRID3Global)

About the partners: Le programme GRID3 est financé par une subvention de la fondation Bill & Melinda Gates et le Bureau du Commonwealth et du Développement des affaires étrangères du Royaume-Uni (FCDO). Il est mis en œuvre par le Center for International Earth Science Information Network (CIESIN), le Fonds des Nations Unies pour la population (UNFPA), WorldPop de l'Université de Southampton et Flowminder.

BILL & MELINDA
GATES foundation



Center for International Earth
Science Information Network
EARTH INSTITUTE | COLUMBIA UNIVERSITY



WorldPop FLOWMINDER.ORG