

Data Release Statement
GRID3 DRC Kongo Central Settlements Points, Version 1
March 2020

Abstract

This document outlines the basic methods and data sets used to construct a spatial database with settlement points and place names in Kongo Central, Democratic Republic of the Congo (DRC). Limitations and use constraints are provided.

Data Set Citation

Bureau Central du Recensement (BCR), Democratic Republic of the Congo, and Center for International Earth Science Information Network (CIESIN), Columbia University. 2020. GRID3 DRC Kongo Central Settlement Point, Version 1. Palisades NY: Geo-Referenced Infrastructure and Demographic Data for Development (GRID3). DOI: <https://doi.org/10.7916/d8-42f6-7v44>. Accessed DAY MONTH YEAR

Data Use Constraints

Users are free to use, copy, distribute, transmit, and adapt the work for commercial and non-commercial purposes, without restriction, as long as clear attribution of the source is provided.

Copyright 2020. The Trustees of Columbia University in the City of New York. The data and accompanying document is licensed under a Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0>) and specified in legal code (<http://creativecommons.org/licenses/by/4.0/legalcode>).

Contacts and Data Queries

GRID3 appreciates feedback regarding this data set, such as suggestions, discovery of errors, difficulties in using the data, and format preferences.

Please contact: Geo-Referenced Infrastructure and Demographic Data for Development (GRID3) Email: data.queries@grid3.org

Table of Contents

1. Introduction	3
2. Input Data	3
2.1 Settlement polygon layer	3
2.2 Source point layers for naming.....	4
3. Method	4
3.1 Settlement point layers.....	4
3.2 Generating a master list: Spatially matching data points (names) with settlement locations.....	4
4. Validation Method	5
5. Dataset Descriptions	5
6. Known Data Limitations	5
7. Disclaimer	5
8. Acknowledgments.....	6
9. References	6

1. Introduction

This document explains the methodology and data pertaining to the Kongo Central Settlements Version 01 database, which provides a comprehensive set of settlement points and their names to spatially locate and identify settlement features in Kongo Central. This work has been undertaken as part of the Geo-referenced Infrastructure and Demographic Data for Development (GRID3) initiative in Democratic Republic of the Congo (DRC). GRID3 is a multi-country and multi-donor partnership that aims to support the collection, storage, and application of geospatial data for informed decision-making. The initiative is funded by the Bill & Melinda Gates Foundation and the United Kingdom's Department for International Development. It is implemented by WorldPop/Flowminder and the United Nations Population Fund, and coordinated by the Center for International Earth Science Information Network (CIESIN).

2. Input Data

Two types of input data were used during the development of the GRID3 Kongo Central settlements database: settlement polygons and settlement point data.

2.1 Settlement polygon layer

The settlement polygon layer was produced by Oak Ridge National Laboratory (ORNL) and covers the provinces of Kongo Central, Kwilu, Kwango, and Mai-Ndombe (Oak Ridge National Laboratory, 2018) in western DRC. The settlement polygons are proprietary data with a restricted license. Due to this limitation, the present database uses the centroid of the features extracted from satellite imagery that exclusively fall within Kongo Central.

ORNL characterised building density into three (3) classes: built-up areas, small settlement areas, and hamlets. These three classes of settlement agglomerations are presented below:

Built-up areas (BUAs)

A built-up area (BUA) is an area of urbanisation with moderately-to-densely-spaced buildings and a visible grid of streets and blocks.

Small Settlements (SSAs)

A small settlement (SSA) is a settled area of permanently inhabited structures and compounds of roughly a few hundred to a few thousand inhabitants. The housing pattern in SSAs is an assemblage of family compounds adjoining other similar habitations.

Hamlets

A hamlet is a single-family compound or several compounds or sleeping houses in isolation from small settlements or urban areas.

2.2 Source point layers for naming

Settlement point data were obtained from different sources, such as government entities, non-governmental organisations, private companies, citizen communities, and independent contractors (Table 1). Besides geographic location, the main attribute of these layers is place names.

Table 1. Settlement point data sources

ACRONYM	NAME OF THE ORGANISATION	YEAR
BCR	Bureau Central du Recensement	2018
CAID	Cellule d'Analyses des Indicateurs de Développement	1994 - 2016
OSM	OpenStreetMap Foundation	2018
SNIS	National Health Information System	2018
UCLA	University of California, Los Angeles	2019
UNICEF	United Nations Children's Fund	2018

3. Method

This section outlines the basic method used to construct a spatial database with settlement points with place names in Kongo Central. The method presented below describes the various steps followed during the development of the GRID3 DRC settlement polygon database, in which a compilation of independent point data was used to name polygon centroids based on location—that is, any valid point within or in close proximity to a settlement polygon is assumed to be the name of the settlement. The aim was to name as many BUAs and SSAs as possible. The naming of hamlets was left as optional, based on valid data available.

3.1 Settlement point layers

Point data were gathered from multiple sources, with various degrees of accuracy. A series of data-cleaning steps were performed to remove identical duplicates (i.e. data points with the exact same name and exact same geographic coordinate), white spaces, and special characters.

3.2 Generating a master list: Spatially matching data points (names) with settlement locations

The near analysis tool in ArcGIS (v10.6) was used to match names with settlement locations, based on a specified spatial threshold. In this case, the near analysis evaluated whether a point fell inside or outside a 100 metre buffer of any given settlement polygon. When valid points were found (i.e. points within the 100 metre buffer), name attributes from the corresponding source were assigned to the respective polygon centroid. This generates a master list of points near settlements.

4. Validation Method

A variety of validation steps were used to improve the representativeness of the feature extraction centroids and attribute data.

First, GRID3 compared the polygons against the points and identified what was missing or incorrect in the points, checking for misalignments, potential errors, and data gaps. GRID3 took this information to the DRC government to see if it could fill in the gaps, as well as checked non-governmental sources like OpenStreetMap. If none of these sources had the needed information, GRID3 worked with a partner that is already in the field to collect the missing data.

After field collection, GRID3 repeated the exercise of linking place names with settlement centroids to get a best-possible settlement layer that incorporated all inputs. Finally, GRID3 created the comprehensive settlement point layer.

5. Dataset Descriptions

The data are in shapefile format and consist of a point shapefile created from the polygon centroids with the settlement attributes.

Extent: Democratic Republic of the Congo: Admin Level 1 Boundaries for Kongo Central (Référentiel Géographique Commun, 2017)

Coordinate system: GCS WGS 1984

6. Known Data Limitations

The ORNL settlement centroids have not been visually inspected for false positives. A point may exist for a location where there is no settlement. Temporal mismatches exist among the point datasets, the ORNL settlement polygons, and the satellite imagery used to perform quality checks. This may lead to settlements not being identified or the inclusion of abandoned settlements.

Another limitation is that GRID3 may not have included all possible data point layers to name all extracted settlement polygons. The research team is confident that all known available data sources were gathered, though it is possible one was missed.

7. Disclaimer

CIESIN, Columbia University and the GRID3 initiative follows procedures designed to ensure that data disseminated by project are of reasonable quality. If, despite these procedures, users encounter apparent errors or misstatements in the data, they should contact GRID3 at data.queries@grid3.org.

CIESIN, Columbia University, and their sponsors do not guarantee the accuracy, reliability, or completeness of any data provided. We provide this data without warranty of any kind

whatsoever, either expressed or implied and shall not be liable for incidental, consequential, or special damages arising out of the use of any data provided.

8. Acknowledgments

GRID3 thanks the UCLA-DRC Health Research and Training Program, the Kinshasa School of Public Health, and the Bureau Central du Recensement. GRID3 also thanks Oak Ridge National Laboratory for providing the settlement data. GRID3 thanks OpenStreetMap and its contributors for the work carried out in the DRC.

Funding for the development and dissemination of this data set was provided by the Bill & Melinda Gates Foundation and the United Kingdom Department for International Development.

9. References

Bureau Central du Recensement, Democratic Republic of the Congo. (2018, 10). Localites [shapefile].

Cellule d'Analyses des Indicateurs de Développement. (1994 to 2016). Settlement Names Across DRC [shapefile]. Kinshasa/Gombe, Democratic Republic of the Congo.

Barau, I., et al. (2014). Improving Polio Vaccination Coverage in Nigeria Through the Use of Geographic Information System Technology. *The Journal of Infectious Diseases* 210(suppl_1): S102-S110, <https://doi.org/10.1093/infdis/jiu010>.

Oak Ridge National Laboratory. (2018, Feb. 7). LandScan HD DRC version 1.1[GDB].
OpenStreetMap. (2018). Cities, Towns, Suburbs, and Villages [shapefile]. Geofabrik.
SNIS, B. S. (2018, Jul). Villages SNIS [shapefile]. UCLA.

Référentiel Géographique Commun. (2017, May 30). Democratic Republic of the Congo: Admin Level 1 Boundaries. Référentiel Géographique Commun.

UNICEF. (2018, May 10). UNICEF WASH Villages [shapefile].

University of California Los Angeles. (2019, Feb). Healthposts geocoordinates for Bas Congo [Excel].