COLUMBIA CLIMATE SCHOOL
CENTER FOR INTERNATIONAL EARTH SCIENCE INFORMATION NETWORK

Data Release Notes

Name of the dataset	GRID3 COD - Health Facilities v3.0
Name of the file	*GRID3_COD_health_facilities_v3_0.gpkg*
Date of data release	October 21, 2024
File format	OGC Geopackage
Dataset version	3.0
Abstract	This document outlines the methodology and data sources used during the production of the GRID3 COD - Health Facilities v3.0 dataset. The dataset consists of health facility points with name, location, health zone, and health area, among other attributes, for thirteen provinces in the Democratic Republic of the Congo (COD). Limitations and use constraints are provided. This operational dataset has not been fully validated by government officials or ministries. The current version supersedes the GRID3 COD - Health Facilities v2.0. The following changes were made: ● Tshopo and Mongala Provinces have been added.
Dataset citation	Center for International Earth Science Information Network (CIESIN), Columbia University and Ministère de la Santé Publique, Hygiène et Prévention, Democratic Republic of the Congo 2024. GRID3 COD - Health Facilities v3.0. New York: GRID3. https://doi.org/10.7916/1c3h-tc02 . Accessed <DAY MONTH YEAR>.
Terms of use	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 2024. The Trustees of Columbia University in the City of New York.
Data license	The data and accompanying document are licensed under a Creative Commons Attribution 4.0 International License, CC BY 4.0 (http://creativecommons.org/licenses/by/4.0) and specified in legal code (http://creativecommons.org/licenses/by/4.0/legalcode).
Contacts and data queries	The authors of this dataset appreciate feedback regarding the data, including suggestions, discovery of errors, difficulties in using the data, and format preferences. For dataset-related questions, please send an email to: info@ciesin.columbia.edu

I. Data inputs / methodology

To create this dataset, CIESIN developed a consistent data schema and methodology to harmonize the data from eleven different provinces collected as five different province groups. The province groups are legacy groupings corresponding to the order in which the original data was collected.

- Province group 1: Haut-Katanga, Kasaï, Kasaï-Oriental, Kinshasa, and Lomami
- Province group 2: Haut-Lomami and Tanganyika
- Province group 3: Ituri and Kwilu
- Province group 4: Maniema
- Province group 5: Kasaï-Central
- Province group 6: Tshopo and Mongala

Province Group 1: Haut-Katanga, Kasaï, Kasaï-Oriental, Kinshasa, and Lomami

From January to July 2021 with the support of provincial and national health authorities, local healthcare workers (“head nurses”, “health zone management staff”, and “head doctors of the health zones”) and GRID3 GIS specialists (“mappers” and “provincial coordinators”) engaged in a participatory mapping process in Haut-Katanga, Kasaï, Kasaï-Oriental, Kinshasa, and Lomami. This mapping process occurred at the level of the health zone (an operational unit made up of approximately 15-20 health areas).

Mappers were deployed to health zones in teams of two for approximately nine days where they trained the health area head nurses on data collection using the Geospatial Tracking System (GTS), an Open Data Kit (ODK)-based application. The head nurses routinely work in their respective areas and have a good understanding of the health facilities, settlements, and points of interest (POIs, such as schools and religious centers) within their health areas.

While the head nurses collected data, the GRID3 mappers worked with the health zone management team to validate and modify data from the field. After all data were collected, cleaned, and integrated into final geospatial layers, the health zone head doctor validated the preliminary data.

From July 2021 through October 2022, the mappers and provincial coordinators worked with CIESIN staff to consolidate the data and address inconsistencies found in the data (e.g. spelling errors, accuracy issues, duplicates, etc). This data was used to produce basemaps at the health area level that were shared back with each health zone and province authorities for a second round of validation.

From November through January 2022, the in-country GIS team worked with CIESIN staff to integrate these corrections into a final geodatabase.

This work was also done with the participation and supervision of the Direction du Système National d'Information Sanitaire (DSNIS). The Agence Nationale d'Ingénierie Clinique, de l'Information et de l'Informatique de Santé (ANICiiS) also played an important role in the area of data governance. This work was part of the GRID3 Mapping for Health project.

Province Group 2: Haut-Lomami and Tanganyika

Haut-Lomami and Tanganyika settlement data was originally created through an extensive fieldwork exercise spanning from July to September 2019.

Phase 1: 2019 Field data collection

With the support of provincial and national health authorities, local healthcare workers and GRID3 GIS specialists engaged in a participatory mapping process in Haut-Lomami and Tanganyika from July to September 2019. This mapping process occurred at the health zone level.

Mappers deployed to the health zones organized participatory mapping meetings with local healthcare workers, and trained head nurses to collect settlement, health facility, and other points of interest in their respective health areas using an ODK-based smartphone application. Mappers then used this information to delineate health area boundaries. This data was then sent back to CIESIN for additional quality checks.

Phase 2: Data integration

The data collected in the field proved to be highly accurate and nearly complete when compared to the master list of health facilities managed by the DSNIS. To ensure 100% completion, CIESIN integrated a few points provided by data providers such as Acasus, PROSANI, and WHO.

Province Group 3: Ituri and Kwilu

Ituri and Kwilu health facility point data were created using a variety of data inputs and validated through two field missions. DHIS2 integration was also conducted.

Table 1- Data sources

Acronym	Source Name	Kwilu	Ituri
PEV Acasus	Programme Elargi de Vaccination (PEV) / Acasus	X	X
BLSQ	Bluesquare	X	
DSNIS	Division du Système National d'Information Sanitaire (DSNIS)	X	X
GRID3	Ministère de la Santé Publique (MoH) / Center for International Earth Science Information Network (CIESIN)	X	X
ESPK ECV	Ecole de Santé Publique de Kinshasa (ESPK) / Enquête de Couverture Vaccinale (ECV)		X

Acronym	Source Name	Kwilu	Ituri
OMS DSNIS	Organisation Mondiale de la Santé (OMS) / Division du Système National d'Information Sanitaire (DSNIS)		X
PNLP	Programme National de Lutte contre le Paludisme (PNLP)	X	
PNLTHA UCLA	Programme National de Lutte contre la Trypanosomiase Humaine Africaine (PNLTHA) / Programme de Recherche et de Formation en Santé, University of California, Los Angeles (UCLA)	X	
OMS ISS	Organisation Mondiale de la Santé (OMS) / Supervision Formative Intégrée (ISS)	X	

In Kwilu, the data collection was completed by a wide range of actors

- PNPL: data collected by IMA World Health on the behalf of the PNLP as part of bednet distribution campaign, between June and August 2022. Local health zone teams collected GPS points of the health facilities in their distribution areas.

- PEV Acasus: time series data from the multiple rounds of EPI surveillance of fixed vaccination sessions since 2020. The data was collected by EPI staff.

- DSNIS: primary data collected by head nurses (Infirmiers Superviseurs) trained by the Bureau de Cartographie, using Garmin GPS devices. The data was collected by Infirmier Superviseurs.

- OMS ISS: data collected by WHO consultants since 2019 as part of the Integrated Supportive Supervision (ISS) to health facilities.

- PNLTHA UCLA: data collected by a team of enumerators coordinated by UCLA on the behalf of the National Sleeping Sickness Control Programme in 2018.

- Bluesquare: compilation of data collected in the field by various partners.

- ESPK_GIRD3: mix of data directly collected in the field or digitalized under the close supervision of the health team (in order to complete some data gaps when funds were not available to collect the data in the field).

A first GRID3 verification mission was carried out in October-November 2021. Two GRID3 mappers traveled to Kikwit and Bandudu, respectively, to present the consolidated data in digital and printed format. The majority of the verifications were carried out in collaboration with antenna-level and provincial-level staff. Additional targeted data collection was conducted to fill in pre-identified data gaps.

A second GRID3 mission was organized one year later, in October-November 2022 after PLNP data was incorporated. Therefore, a second round of data validation took place in the Kikwit and Bandudu antenna to update and consolidate the health facilities database.

In Ituri, the data collection was completed by the following actors:

- GRID3: mix of data directly collected in the field or digitalized under the close supervision of the health team (in order to complete some data gaps when funds were not available to collect the data in the field).

- OMS DSNIS: data set compiled by the WHO GIS team in support of the DSNIS, in collaboration with the OpenStreetMap (OSM) community and other technical partners after the Ebola outbreak in August 2018.

- PEV Acasus: time series data from the multiple rounds of EPI surveillance of fixed vaccination sessions since 2020. The data was collected by EPI staff.

- DSNIS: data collected by head nurses (Infirmiers Superviseurs) trained by the Bureau de Cartographie, using Garmin GPS devices.

- ESPK ECV: data collected by Kinshasa School of Public Health (ESPK) in 2020 and 2021 as part of the Vaccination Coverage Survey (ECV).

The data was verified through a mission organized in October 2022. Four GRID3 mappers traveled to Bunia and Aru, respectively, to validate the health facility data in all health zones of the two antennas. During this mission, the GRID3 mappers worked with local health teams to validate and/or make corrections when needed. Primarily the focus was to ensure completeness and accuracy of the data.

All data edits from both missions were sent back to CIESIN for final verification before integration and publication. CIESIN ensured that the final layers were free of spelling errors and that each health area had at least one health facility.

Province Group 4: Maniema

The Maniema data was created through an extensive fieldwork data collection conducted by the Kinshasa School of Public Health (Ecole de Santé Publique de Kinshasa, ESPK) and supplemented with additional data from the Elongated Programme for Immunization (EPI/PEV) collected as part of their routine surveillance/ monitoring; as well as fieldwork supported by the Centers for Disease Control (CDC) and CIESIN in 2023.

Table 2- Data sources

Source Name	Data type/ format	Input data year	Source Acronym
Fieldwork data collected by the Kinshasa School of Public Health (ESPK) in collaboration with GRID3 and CIESIN.	Spatial points	2024	ESPK GRID3
Fieldwork data collected by the Programme Elargi de Vaccination (PEV) in collaboration with the Centers for Disease Control (CDC) and CIESIN	Spatial points	2023	PEV CDC CIESIN
Ecole de Santé Publique de Kinshasa (ESPK) / Enquête de Couverture Vaccinale (ECV)	Spatial points	2020 - 2022	ESPK ECV
Health facility data from the DHIS2	Tabular	2023	DHIS2
Programme Elargi de Vaccination (PEV) / Acasus	Spatial points	2019-2023	PEV Acasus

Phase 1: Data integration

In 2022-2024, CIESIN received health facilities data from the PEV/ Acasus, PEV/ CDC, and ESPK ECV. These datasets were assembled and explored, cleaned, and matched against existing microplans from Maniema, from prior years and DHIS2's list. The combined data sets created the list of health facilities to collect during fieldwork.

Phase 2: Field data collection and data processing

ESPK with the support of provincial and national health authorities, local healthcare workers and GRID3 GIS specialists engaged in data collection from October 2023 to January 2024. Teams of enumerators were deployed to each health zone, and liaised with local authorities and local health workers to collect health facility's names and geographic points using an ODK-based application. This data was then sent back to CIESIN for processing, integration, and quality checks.

Province Group 5: Kasaï-Central

A comprehensive geospatial survey was conducted by ESPK between March and May 2024 in collaboration with GRID3 and partners. The survey team collected names, geospatial locations, and relevant attribute information of health facilities that conform the network of health service points in the Kasaï-Central province.

Table 3 - Data sources

Source Name	Data type/ format	Input data year	Source Acronym
Fieldwork data collected by the Kinshasa School of Public Health (ESPK) in collaboration with GRID3 and CIESIN.	Spatial points	2024	ESPK GRID3

CIESIN proceeded with data processing, integration, and harmonization steps to create a clean health facilities dataset. Data processing steps include standardization of health facility names and types, as well as clustering to detect anomalies in spatial accuracy and/or attribute information. Visual inspection using Maxar's satellite imagery (2022-2024) was conducted when relevant.

A consistent data schema prioritizing information preservation was also developed to include data fields as per common use cases. A standardized data schema will enable seamless integration across subsequent versions of GRID3 health facility datasets in the DRC. Common generalized values were created for all fields and the original values in each dataset were mapped to these.

Province group 6: Tshopo and Mongala

A comprehensive geospatial survey was conducted by ESPK between March and May 2024 in collaboration with GRID3 and partners. Similar as in Maniema, the CIESIN team assembled a baseline, and the survey team collected names, geospatial locations, and relevant attribute information of health facilities in collaboration with health authorities in Tshopo and Mongala.

Table 4 - Data inputs

Source Name/ description	**Data type/ format**	**Input data year**	**Source Acronym**
Programme Elargi de Vaccination (PEV) / Acasus	Spatial points	2024	PEV ACASUS
Health facility data from the DHIS2	Tabular	2024	DHIS2
Fieldwork data collected by the Kinshasa School of Public Health (ESPK) in collaboration with GRID3 and CIESIN.	Spatial points	2024	ESPK GRID3

Phase 1: Data integration

CIESIN extracted, cleaned, and harmonized health facility locations and names contained in all data sources, as specified in Table 4. We also matched the data against DHIS2 records in order to identify potential data gaps to be completed during fieldwork. We utilized this preliminary dataset as the baseline.

Phase 2: Field data collection and data processing

A comprehensive geospatial survey was conducted by ESPK between March and May 2024 in collaboration with GRID3 and partners. The survey team collected names, geospatial locations, and relevant attribute information of health facilities that conform to the network of health service points in the Tshopo and Mongala provinces.

CIESIN conducted data processing, integration, and harmonization to create an enhanced health facilities dataset. Data processing steps included standardization of names and types, as well as clustering to detect anomalies in spatial accuracy and/or attribute information. Visual inspection using satellite imagery was conducted as needed.

A consistent data schema prioritizing information preservation was also developed to include data fields as per common use cases. A standardized data schema will enable seamless integration across subsequent versions of GRID3 health facility datasets in the DRC.

II. Dataset Description

The *GRID3 COD - Health Facilities v3.0* dataset consists of spatial point data with attribute information (see table 5). The data are available for download in OGC Geopackage format contained in a zip file; metadata files are also included.

Table 5- Codebook

Variable Names	**Type**	**Definition**
OBJECTID	numeric	Software- generated unique code
pays	text	Country name french official UN member state spelling
iso3	text	ISO3 code
province	text	DHIS2 province name
prov_uid	text	The identifier of the province in the DHIS2
antenne	text	Antenna name
zonestante	text	DHIS2 health zone name
zs_uid	text	The identifier of the health zone in the DHIS2
airesante	text	Health area name

Variable Names	Type	Definition
as_uid	text	The identifier of the health area in the DHIS2
localite	text	Settlement name
essnom1	text	Abbreviated version of the health facility type (esstype) and health facility name (essnom2)
esstype	text	Health facility type generalized by GRID3 from original raw data
essnom2	text	Health facility name without type
typeorig	text	Health facility type from original raw data
categorie	text	Category of health facility
frigo	text	Refrigeration available
frigofct	text	Refrigeration functioning at time of field work
vaccfixe	text	Is this site a fixed vaccination site
dhis2	text	DHIS2 health facility code
date	text	Year of data collection or last edit/modification
source	text	Institution or project providing point data for this dataset
lat	numeric	Latitude in decimal degrees
lon	numeric	Longitude in decimal degrees
precision	numeric	Precision of GPS coordinates indicating the tolerance in meters
origine	text	Method of coordinate derivation
grid3id	text	Internal GRID3 ID

III. Known Data Limitations and Disclaimer

The spatial accuracy of the health facility data is dependent on both the accuracy of the point data collected in the field as well as on the correctness of the edits made to the collected data throughout the validation process. In general, it was assumed that data collected in the field was more accurate than the one compiled through secondary data sources. Temporal mismatches exist among the point datasets and the satellite imagery used to perform quality checks. This may lead to health facilities not being identified or the inclusion of abandoned health facilities. Likewise, spelling mistakes and/or mismatches may have occurred due to colloquial variations on how data points are referred to in the field.

This operational dataset has not been fully validated by government officials or ministries.

CIESIN, Columbia University, and its co-authors follow procedures designed to ensure that data disseminated by the project are of reasonable quality. If, despite these procedures, users encounter apparent errors or misstatements in the data, they should contact CIESIN, info@ciesin.columbia.edu.

CIESIN, Columbia University, its co-authors, and their sponsors do not guarantee the accuracy, reliability, or completeness of any data provided. We provide these 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.

IV. Acknowledgments

CIESIN and its co-authors thank the following institutions that provided input data and/or assistance with data production:

Acasus, Switzerland
Agence Nationale d'Ingénierie Clinique et du Numérique de la Santé (ANICNS) , DRC
Bill & Melinda Gates Foundation, USA
Bluesquare, Belgium
Bureau Central du Recensement (BCR), DRC
Caritas, USA
Centers for Disease Control and Prevention (CDC), USA
Direction d'Etudes et Planification (DEP), DRC
Direction des Soins de Santé Primaires (DSSP), DRC
Division du Système National d'Informations Sanitaires (DSNIS), DRC
Division Provinciale de la Santé (DPS) de Kinshasa, Kwilu, Kasaï, Kasaï-Central, Kasaï-Oriental, Lomami, Haut-Lomami, Tanganyika, Haut-Katanga, Ituri, Maniema, Tshopo et Mongala.
Ecole de Santé Publique de Kinshasa (ESPK), DRC
Gavi, the Vaccine Alliance, Switzerland
Geospatial Evaluation and Observation Lab (geoLab), College of William & Mary, USA
Global Affairs Canada (GAC), Canada
Global Good, USA
IMA World Health, DRC
Initiative Régionale de Documentation et d'Accompagnement Communautaire au Développement (IDRAC Sarl), DRC
International Federation of Red Cross and Red Crescent Societies (IFRC), Switzerland
International Medical Corps (IMC), USA
Médecins Sans Frontières (MSF), Switzerland
Ministère de l'Environnement et Développement Durable (MEDD), DRC
Ministère de la Santé publique, Hygiène et Prévention, DRC
Novel-T, Switzerland
Open Street Map (OSM), DRC
PATH, USA
Programme Elargi de Vaccination (PEV), DRC

Programme National de Lutte contre le Paludisme (PNLP), DRC
Référentiel Géographique Commun (RGC), DRC
Soins de Santé Primaires en Milieu Rural (SANRU), DRC
The International Organization for Migration (IOM), DRC
United Nations Children Fund (UNICEF), USA
United Nations Development Programme (UNDP), USA
United Nations Office for Project Services (UNOPS), Denmark and DRC CO
United Nations Office for the Coordination of Humanitarian Affairs (OCHA), USA
United Nations Organization Stabilization Mission in the Democratic Republic of the Congo (MONUCSO), DRC
University of California, Los Angeles (UCLA) DRC Health Research and Training Program, USA
VillageReach, USA
World Health Organization (WHO), Switzerland (HQ), Brazzaville (AFRO), Kinshasa (CO)
World Resources Institute (WRI), USA

Funding for this work was provided by GRID3 under grants INV-044979 GRID3 - Phase 2 Scaling, and FAE/GRID3/001/2024 Soutien à la mise en oeuvre des activités du Projet de Vaccination des Enfants Zéro Dose du Fonds d'Accélération de l'Équité (FAE) en RDC.