COLUMBIA CLIMATE SCHOOL
CENTER FOR INTERNATIONAL EARTH SCIENCE
INFORMATION NETWORK

Data Release Notes

Name of the dataset	GRID3 COD - Health Zones v3.0
Name of the file	GRID3_COD_health_zones_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 for constructing the GRID3 COD - Health Zones v3.0 dataset. The dataset consists of health zone boundaries with name, location, and other related 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 Zones 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 Zones v3.0. New York: GRID3. https://doi.org/10.7916/hb9r-8p34. 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 location and names of health facilities, settlements, and points of interest (POIs, such as schools and religious centers) within their health (or catchment) areas.

While the head nurses collected data in their health areas, the 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 (i.e. spelling errors, gaps and overlaps, other inconsistencies). This data was used to produce basemaps at the health area-level and shared back with every health zone and province for a second round of validation.

From November 2022 through January 2023, the in-country GIS team worked with CIESIN staff to integrate these corrections into a final geodatabase. Final health zones polygons were generated by dissolving health areas based on the corresponding attribute.

This work was 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

The Haut-Lomami and Tanganyika health area data was originally created through an extensive fieldwork exercise from July to September 2019 with additional data added from the National Malaria Elimination Programme in the DRC (PNLP).

Table 1- Data sources

Source	Description
PNLP	Programme National de Lutte contre le Paludisme (National Malaria Elimination Programme) of the DRC
ESP-UCLA	Public health school (École de Santé Publique) of the University of California, Los Angeles.
MSP GRID3 CIESIN	DRC's Ministry of Public Health (Ministère de Santé Publique) and the Center for International Earth Science Information Systems)
DSNIS	Division du Système National d'Informations Sanitaires

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 were deployed to the health zones, where they organized participatory mapping meetings with local healthcare workers. They also trained head nurses to collect data on settlements, health facilities, and other points of interest in their respective health areas using an ODK-based application. Mappers then used this information to delineate health area boundaries, using previously existing data to guide this effort. This data was then sent back to CIESIN for additional quality checks.

Phase 2: Integration of data from the PNLP

In October 2021, GRID3 received access to a large geospatial dataset collected during a bednet distribution campaign by IMA World Health. This dataset was combined and consolidated with previous GRID3 settlement data for Haut-Lomami and Tanganyika. The health zone and health area boundaries were then refined based on the health area and health zone attributes of PNLP-derived point clusters and the GRID3 friction surface layer (non-published). Final health zones polygons were generated by dissolving health areas based on the corresponding attribute.

Province Group 3: Ituri and Kwilu

Ituri and Kwilu health area boundaries were derived from clustering of PLNP data points. Health area and health zone boundaries were verified with local health teams through in-person missions in both provinces.

Between 2021 and 2022, GRID3 received PLNP data covering the provinces of Ituri and Kwilu. This data consisted of household-level GPS points with village, health area, and health zone attributes; and collected by IMA World Health, an implementing partner of the PNLP, during province-wide bed net distribution campaigns. In Ituri, over 1,165,000 household points were received from a bed net distribution campaign conducted in June 2021. In Kwilu, over 1,191,000 household points were received from a bed net distribution campaign conducted in July 2022. These household points were processed and mapped against the GRID3 settlement extents datasets as a way of validation. Further, PLNP-derived settlement points were combined with other data sources in order to compile a geodatabase as complete as possible.

CIESIN used the following process to delineate preliminary health area boundaries (see steps below). Subsequently, health areas were dissolved to create health zones, based on the corresponding attributes.

- Assessment of already existing health area boundaries. Available point data available (settlements, health facilities, other points of interest available with relevant attribute information) with health area and health zone attributes were overlaid with the boundaries to determine the percent match and select the best layer available, in concert with the local health authorities. In Kwilu, the 2018 data produced by UCLA on behalf of the PNLTHA - and already integrated into the DSNIS national database - was considered the best available data. In Ituri, a combination of recent WHO and OSM boundary work was used as reference only.

- Creation of a friction surface layer. Several input datasets --e.g. elevation, slope, land cover, water bodies (rivers, streams, permanent bodies of water, lakes, etc.), and road infrastructure (primary, secondary, tertiary road networks) were combined to construct a surface (raster or grid, at ~100 m x 100m resolution) with a travel costs at the pixel level --which quantifies the amount of time that it takes to travel from one point to another. Rivers, permanent bodies of water, or cliffs were used as physical barriers and considered obstacles to be circumvented.

- Inclusion of the best available health areas to the friction layer. In addition to natural and man-made barriers, the best available health area boundaries were added to the friction surface to avoid modifying or redrawing existing boundaries. If extensive work had already been conducted and validated in an area, and only a few points fell outside of the original boundaries, no changes were made to the boundaries, as the amount of data available did not justify modifications. Conversely, if a sufficient quantity of recent and reliable data points fell outside current health area limits, these were adjusted accordingly.

- Creation of village catchment areas. The granularity of the data collected during the ITN distribution campaign of the PNLP (door-to-door distribution of bednets) at the provincial scale allowed us to create buffers around each survey location in order to generate village catchment areas (settlement contours).

- Creation of health areas. Each village catchment area was dissolved into health area boundaries, based on their health area attribute. An additional visual inspection identified and corrected minor issues and ensured the boundaries were topologically correct.

- Preliminary health zone boundaries (pre-alpha) were created by dissolving health area polygons based on health zone attributes. This data was presented to local health officials (Médecins Chefs de Zone, Infirmiers Superviseurs) to make sure all the health zones were complete, and health areas fell within their respective health zone boundaries - and/or to make the necessary adjustments when necessary.

In Kwilu, the data was verified through a mission organized in October 2022. Two GRID3 mappers traveled to Kikwit and Bandudu, respectively, to present the cartographic improvements driven by the use of the PNLP data and to verify that the boundaries were correctly delineated. The majority of this work was carried out in collaboration with both antenna and provincial level staff.

In Ituri, the data was verified through a mission also in October 2022. Four GRID3 mappers traveled to Bunia and Aru, respectively, to validate the health area boundary data in all health zones. During this mission, the GRID3 mappers worked with local health teams, where boundary layers were validated and edited.

All modifications from both missions were sent back to CIESIN for final verification before publication. CIESIN ensured that the final layers were free of spelling and topological errors.

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 PLNP.

Table 2- Data sources

Source Name/ description	Data type/ format	Input data year
Fieldwork data collected by the Kinshasa School of Public Health (ESPK) in collaboration with GRID3 and CIESIN.	Spatial points and qualitative feedback	2024
Pre-Distribution Registration Survey (PDRS) from the National Malaria Control Programme (PNLP) collected as part of the anti-malaria campaigns in the Democratic Republic of the Congo	Polygons derived from spatial points	2023

Phase 1: Integration of data from the PNLP

In October 2021, CIESIN received access to a large settlement point dataset collected during a bednet distribution campaign by IMA World Health. This dataset was explored, cleaned, and matched against health area and health zone lists within DRC's DHIS2. This attribute information was used to produce a preliminary, draft boundary to be validated with the help of local and provincial health authorities.

Phase 2: Field data collection

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. GRID3 GIS specialists were deployed to each health zone, and liaised with local authorities and health workers to validate the list of health areas within each health zone and validate (or make corrections) to the preliminary draft boundary produced in Phase 1. Corrections were sent back to CIESIN for processing and quality checks, and validated against point-data attributes collected as part of the same fieldwork.

Province group 5: Kasaï-Central

A comprehensive geospatial survey was conducted by ESPK between March and May 2024 in collaboration with GRID3 and partners. Similar as with Maniema, the survey team collected names, geospatial locations, and relevant attribute information of spatial points to derive health zones and health areas.

Table 3- Data sources

Source Name/ description	Data type/ format	Input data year
Data collected in the field between March - May 2024 by the Kinshasa School of Public Health (ESPK) in collaboration with GRID3 and CIESIN.	Polygons derived from spatial points	2024

CIESIN delineated health area boundaries by clustering attribute information contained in the data collected in the field. Natural features such as rivers and ridges, as well as roads and railroad lines were considered as barriers to delimit boundaries, when applicable. GIS tools/ environments were used to clean topology and isolated attribute errors, harmonize and standardize data, and resolve other geometry discrepancies. Spatial polygons were matched against the national DHIS2 database to ensure full interoperability across databases. Once clean health areas were obtained, they were dissolved based on attribute information to create clean health zones.

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 zone datasets in the DRC.

Province group 6: Tshopo and Mongala

The Tshopo and Mongala data were added through an extensive fieldwork data collection conducted by ESPK and supplemented with additional data from the PLNP.

Table 4- Data sources

Source Name/ description	Data type/ format	Input data year
Fieldwork data collected by the Kinshasa School of Public Health (ESPK) in collaboration with GRID3 and CIESIN.	Spatial points and qualitative feedback	2024
Pre-Distribution Registration Survey (PDRS) from the National Malaria Control Programme (PNLP) collected as part of the anti-malaria campaigns in the Democratic Republic of the Congo	Polygons derived from spatial points	2021 - Tshopo 2023 - Mongala

Phase 1: Integration of data from the PNLP

We delineated health boundaries by clustering attribute information about health areas and health zones contained within the Pre-Distribution Registration Survey (PDRS) data from the National Malaria Control Programme (PNLP), collected in Tshopo and Mongala in 2021 and 2023 respectively. We utilized this preliminary dataset as the baseline during fieldwork.

Phase 2: Field data collection and data processing

ESPK with the support of provincial and national health authorities, local healthcare workers (i.e. head nurses, health zone management staff, and head doctors of the health zones), and GRID3 GIS specialists engaged in data collection from March to May 2024. GRID3 GIS specialists were deployed to each health zone, and liaised with local authorities and local health workers to validate the list and geometry of health areas within each health zone. Corrections were sent back to CIESIN for processing.

CIESIN consolidated both PLNP and fieldwork data, giving priority to the latter. We considered natural features such as rivers and ridges, as well as roads and railroad lines as barriers to delimit borders, when applicable. We used GIS tools/ environment to clean topology and isolate attribute errors, harmonized and standardized data, and resolved other geometry discrepancies. We also matched the spatial polygons against the national DHIS2 database to ensure interoperability across databases. Once we obtained clean health areas, we dissolved them based on attribute information and created clean health zones. The current dataset has not been adjusted against neighboring, provincial boundaries previously released by GRID3, yet.

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 zone datasets in the DRC.

II. Dataset Description

The *GRID3 COD - Health Zones v3.0* dataset consists of spatial polygon 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 pyramid, used by the DSNIS
antenne	text	Antenne name
zonesante	text	DHIS2 health zone name
zs_uid	text	The identifier of the health zone in the DHIS2
airesante	text	DHIS2 health area name
as_uid	text	The identifier of the health area in the DHIS2
date	text	Year of data collection or last edit/modification
edit_par	text	Editor of the original data
source	text	Institution or project providing input data for this dataset
grid3id	text	Internal GRID3 ID

III. Known Data Limitations and Disclaimer

The spatial accuracy of the health zones data is dependent on both the accuracy of the input data collected in the field as well as on the correctness of the edits made to the data throughout the validation process. In general, it was assumed that the field collected data were the most accurate data available. Temporal mismatches exist among the point datasets and the satellite imagery used to perform quality checks. Spelling mistakes and/or mismatches may have occurred due to colloquial variations on how data points are referred to in the field.

This dataset should be considered operational; it 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 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.