Zambia land use and land cover field data set

Michelle C. A. Picoli[1], Kenny Helsen[1], Haggai Mulenga[2]

[1]WeForest asbl/vzw – Cantersteen 47, 1000 Brussels, Belgium

[2]WeForest Zambia – PO Box 70591, Ndola, Zambia

{michelle.picoli,kenny.helsen,haggai.mulenga}@weforest.org

Abstract. *This paper presents a data set of land use and land cover collected in Muchinga and Copperbelt provinces, Zambia, in 2023. The land use and land cover field data are essential for the training and validation of classification algorithms. However, open-field data is scarce. The data set provides information on five land use and land cover classes (Forest land, Cropland, Grassland, Wetland, and Other land) for 697 points. The data were collected in a fieldwork campaign that took place between May and June 2023. The data collected in situ were geographically corrected using PlanetScope images with a spatial resolution of 3 m. This data set contributes to the understanding of land use dynamics and provides essential information for environmental studies, land use planning, public policy, and decision-making.*

Resumo. *Este artigo apresenta um conjunto de dados de uso e cobertura da terra coletados em campo nas províncias de Muchinga e Copperbelt, Zambia, entre Maio e Junho de 2023. Dados de campo de uso e cobertura do solo são essenciais para treinamento e validação de algorítimos de classificação. Porém, dados de campo abertos são escassos. O conjunto de dados fornece informações de 5 classes de uso e cobertura da terra (Floresta, Agricultura, Gramíneas, Área Úmida, e Outros usos) para 697 pontos. Os dados de referência foram coletados em um trabalho de campo que ocorreu entre Maio e Junho de 2023. Os dados coletados in situ foram corrigidos geograficamente usando imagens PlanetScope com resolução espacial de 3 m. O conjunto de dados de uso e cobertura do solo contribui para a compreensão da dinâmica do uso do solo e fornece informações essenciais para estudos ambientais, planejamento do uso do solo, políticas públicas e tomada de decisões.*

1. Data Description

Land use and land cover (LULC) data is fundamental for understanding environmental changes. Although satellite images provide information about land use and land cover, data sets are still needed to train and validate the models. In addition to producing accurate maps, field data can reduce the large disparity that is currently observed between different available LULC maps [Fritz et al., 2017].

In particular, the datasets described in this paper provide information on land use and land cover for two provinces in Zambia, based on the classes proposed by the Ministry of Lands, Natural Resources, and Environmental Protection of Zambia

[MLNREP, 2016]. The field campaign collected geo-referenced LULC field data to serve as reference data to validate the LULC maps generated by WeForest.

This data set includes 697 points collected between May and June 2023 in Muchinga and Copperbelt provinces of Zambia, and is available in ESRI shapefile format, published at https://doi.org/10.5281/zenodo.8318287 [Picoli et al., 2023].

The ESRI shapefile format consists of four files (.shp, .shx, .dbf, and .prj). The .shp (shape format) contains the feature geometry; .shx (shape index format) has a positional index of the feature geometry; .dbf (attribute format) includes columnar attributes for each shape in dBase format; and .prj (projection description) has the information of coordinate reference systems [ESRI, 2023]. This data set can be opened in software like QGIS, R, and Python, among others.

Each of these points has a geographic location and a labeled land use and land cover class. In addition to latitude, longitude, and the label, there are four other columns in the attribute table. The "id" column associates each sample point collected with an indicator; "year" refers to the year the data were collected; the EPSG column refers to the coordinate system associated with the shapefile file; and the "code" column refers to the code associated with each LULC class.

The land use and land cover classes collected in the field and their definitions defined by the Ministry of Lands, Natural Resources and Environmental Protection of Zambia [MLNREP, 2016] are described below:

- Forest land: "This is land covered both by natural and planted forest meeting the threshold of 10% canopy cover growing over a minimum area of 0.5 ha with trees growing above 5m height";
- Cropland: "Land actively used to grow agriculture (annual and perennial) crops that may be irrigated or rain feed for commercial, peasant, and small-scale farms around urban and rural settlements";
- Grassland: "Land that includes wooded rangeland that may be covered mainly by grasslands, plains, dambos, and pans found along major river basins and water channels";
- Wetland: "Land that is water logged, may be wooded, such as marshland, perennially flooded plains, and swampy areas that may be recognized and classified as such by RAMSAR";
- Other lands: "Barren land covered by natural bare earth/soil such as sandy dunes, beach sand, rocky outcrops, and may include old open quarry sites".

The code associated with each of the classes is: "Forest land" 1; "Cropland" 2; "Grassland" 3; "Wetlands" 4; "Other land" 5. Figure 1 shows a map location and the classes of the 697 samples. Figure 2 presents the pictures taken in situ of each LULC class collected.

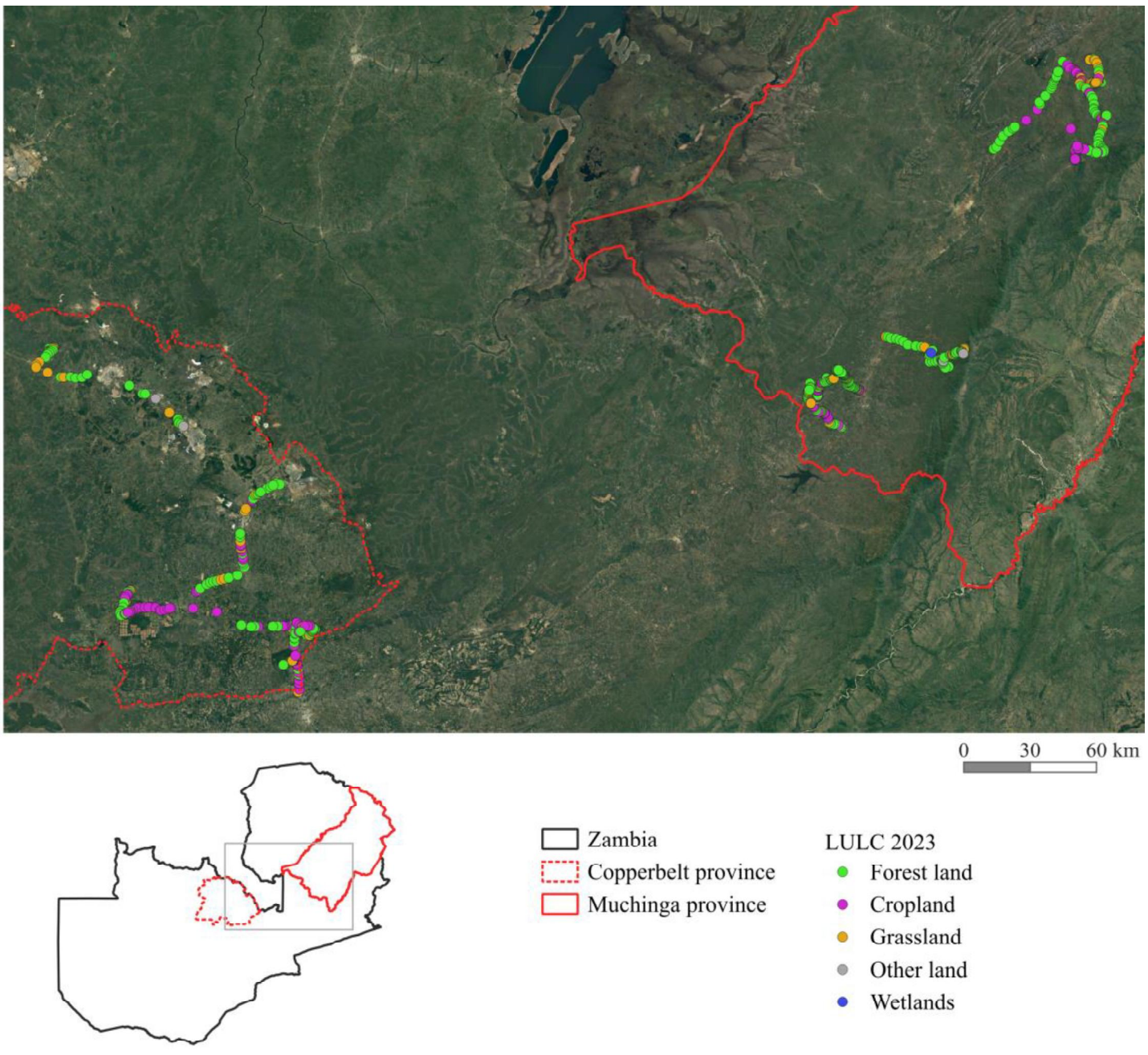

Figure 1. Location of field data collected in Muchinga and Copperbelt provinces, Zambia.

Figure 2. Examples of land use and land cover classes of the data set: (a) Forest land", (b) "Cropland"; (c) "Grassland"; (d) "Wetlands"; (e) "Other land". The “Other land” (e) in the picture is a granite inselberg.

2. Method

2.1. Area of data set collection

Data were collected in the provinces of Copperbelt and Muchinga. According to the most recent official data from the Forestry Department of the Zambia Ministry of Lands and Natural Resources in 2014, in the Muchinga province, the land cover area distribution of ~8.7 Mha was: Forest land 73.3%, Grassland 11.2%, Cropland 6.8%, Wetland 8.1%, Settlements 0.5%, and Other land 0.1% [FAO and MLNR, 2016]. In the Copperbelt province of the ~3.1 Mha, the land cover area distribution was: Forest land 60.5%, Grassland 18.3%, Cropland 17.1%, Wetland 2.5%, Settlements 1.4%, and Other land 0.2% [FAO and MLNR, 2016]. The Zambian climate is characterized as predominantly sub-tropical, with three seasons: a hot and dry season from mid-August to mid-November, a rainy wet season from mid-November to April, and a cool dry season from May to mid-August [World Bank Group, 2023].

2.2. Field data collection

A field campaign was carried out between May and June 2023. The points were collected within the vicinity of areas where the international NGO 'WeForest', in collaboration with the Zambian NGO 'WeForest Zambia' implement forest landscape restoration projects (https://www.weforest.org/programme/miombo-belt/). Altogether, 418 data points were collected, containing information on land use and land cover in the province of Muchinga and 279 data points in the province of Copperbelt. Due to site accessibility limitations, the 'convenience sampling' design was used. This means that the samples were collected alongside main, secondary, and tertiary roads. Table 1 presents the number of samples collected per class.

Information was collected through the KoboCollect application on a mobile phone. PlanetScope images from June 2023 of Color InfraRed composition (CIR: red (R), green (G), near-infrared (NIR)) were used for geometric correction of some data collected in the field.

Table 1. LULC samples collected in the field campaign.

Label	Samples
Forest land	412
Cropland	161
Grassland	109
Wetland	5
Other land	10

2.3. Land use and land cover classes

The data collected in the field were labeled with the land use and land cover classes defined by the Ministry of Lands, Natural Resources, and Environmental Protection of Zambia [MLNREP, 2016]. As the samples were collected alongside main, secondary, and tertiary roads, some adjustments regarding location needed to be made as sometimes the collected points were close to the roads. Therefore, visual interpretation of the PlanetScope false-color composite (R, G, NIR) images with ~3m spatial resolution was also used to adjust the samples location. These images were downloaded through the Planet QGIS Plugin. The Planet Plugin allows QGIS users to explore, stream, and download Planet imagery and Planet Basemaps.

3. Usage Notes

The land use and land cover dataset contributes to the understanding of land use dynamics and provides essential information for environmental studies, land use planning, public policy, and decision-making.

The data collected in the field can be used for land use and land cover classification, which needs data to train and validate classification algorithms. In turn, land use and land cover classification maps can assess land use and land cover change, collaborating on projects that involve agricultural expansion and forest regeneration

assessment, carbon calculation, and even Nationally Determined Contribution (NDC) compliance, signed by the Zambian government in 2015 in the Paris Agreement.

References

ESRI (2023) “GIS Dictionary”. https://support.esri.com/en-us/gis-dictionary, August.

Food and Agriculture Organization of the United Nations (FAO) and Forestry Department of Zambia Ministry of Lands and Natural Resources of Zambia (MLNR) (2016) “Integrated Land Use Assessment Phase II – Report for Zambia”. https://prais.unccd.int/sites/default/files/2018-08/ILUA%20II_Final%20Report_Zambia_19062016.pdf

Fritz, S., See, L., Perger, C., McCallum, I., Schill, C., Schepaschenko, D., Duerauer, M., Karner, M., Dresel, C., Laso-Bayas, J. C. and Lesiv, M. (2017). “A global dataset of crowdsourced land cover and land use reference data”. Sci Data 4, 170075 https://doi.org/10.1038/sdata.2017.75

Ministry of Lands, Natural Resources and Environmental Protection of Zambia (MLNREP) (2016) “Zambia’s Forest Reference Emissions Level Submission to the UNFCCC”. https://redd.unfccc.int/files/2016_submission_frel_zambia.pdf

Picoli, M., Mulenga, H., Van de Loock, D., Watts, N., Zulu, C., Ndabala, R., Mwale, A., Mazimba, G., Zimba, F., Kantumoya, L., Kasanda, I. and Lumbwe, F. (2023). “Zambia land use and land cover field samples” (Version 1) [Data set]. Zenodo. https://doi.org/10.5281/zenodo.8318287

World Bank Group (2023) “Climate Change Knowledge Portal”. https://climateknowledgeportal.worldbank.org/country/zambia/climate-data-historical, August.