Analyzing grid extension suitability: A case study of Ethiopia using OnSSET

Citation for the original published paper (version of record):
Temesgen, A., Wassie, Y., Ahlgren, E. (2024). Analyzing grid extension suitability: A case study of Ethiopia using OnSSET. Energy Strategy Reviews, 52. http://dx.doi.org/10.1016/j.esr.2023.101292

(article starts on next page)

Energy Strategy Reviews 52 (2024) 101292

Contents lists available at ScienceDirect

Energy Strategy Reviews

journal homepage: www.elsevier.com/locate/esr

Analyzing grid extension suitability: A case study of Ethiopia using OnSSET

Adugnaw Lake Temesgen [a,b], Yibeltal T. Wassie [b], Erik O. Ahlgren [b,*]

[a] *Addis Ababa University, Addis Ababa Institute of Technology, School of Electrical and Computer Engineering, Addis Ababa, Ethiopia*
[b] *Chalmers University of Technology Department of Space, Earth and Environment Division of Energy, Technology, SE- 412 96, Gothenburg, Sweden*

ARTICLE INFO

Handling Editor: Mark Howells

Keywords:
Electrification planning
Grid extension
Geospatial factors
OnSSET
GIS tools
Unelectrified areas

ABSTRACT

Several geospatial factors influence the suitability of national power grid expansion, especially in remote areas. Since previous studies have neither explicitly examined the level of influence of these factors nor provided a clear spatial representation of their impact, this paper examines how geospatial factors (distance from substation and road, terrain slope, elevation, and land cover) influence grid extension suitability to unelectrified settlements in the context of Ethiopia. Open-access and remote-sensing datasets are used together with OnSSET geospatial modeling analysis methodology. A spatial grid extension suitability map is developed to display areas that are most suitable, semi-suitable, and less suitable for grid extension. Results show that terrain slope is the most significant contributor to grid extension suitability, accounting for 40.8 % of the total score. The findings reveal that the geospatial factors studied, aggregately, might increase the total investment cost of grid extension by 2.3 %–29 % across Ethiopia. The results also show that 45 % and 85 % of Ethiopia's population live within 10 km distance from high-voltage and projected medium-voltage lines, respectively. The study underscores that rather than focusing exclusively on distances from existing grid infrastructures, it is important to take into account the aforementioned geospatial factors affecting investment costs for grid extension planning.

1. Introduction

Access to electricity is a fundamental pillar for socioeconomic development, underscored by the United Nation's (UN's) Sustainable Development Goal 7 (SDG7) to ensure affordable and reliable energy for all by 2030 [1]. Electricity access significantly influences living standards, education, healthcare, and agricultural productivity [2], but 759 million still lacked basic electricity access in 2019, predominantly in Sub-Saharan Africa (SSA), where over half the population lacks electricity access [3] in contrast to the UN's goal of universal energy access and projected economic growth [4]. In light of this, the imperative to extend electricity access to underserved population figures prominently in the policy agendas of many developing countries. However, the pathways to achieve this goal diverge, primarily revolving around grid extension and off-grid systems.

The de facto electrification strategy is to expand existing national power grids and transmission networks [5], supported by initiatives such as Power Africa [6]. [1]Grid expansion refers to the construction of additional transmission lines to meet the increasing electricity demand. The feasibility of grid-based electrification relies on the correlation between existing grid coverage and population distribution [7]. The distance from households to the nearest grid connection point significantly impacts electricity transmission costs for two reasons: longer transmission lines entail higher capital costs, and also energy losses, for a given connection [8]. Thus, proximity to the grid typically translates to lower connection costs [9].

On the demand side, grid extension-based electrification pathways often face challenges from low electricity consumption by settlements [10]. This can be attributed to several factors, including low population density, sluggish economic growth, and low-income levels of residents: leading to limited demand for electricity. Consequently, the levelized cost of electricity ([2]LCOE) for grid-based electrification tends to be high, making it economically less viable.

Grid extension suitability is also affected by various [3]geospatial

* Corresponding author.
E-mail addresses: adugenetlake@gmail.com, adugnaw@chalmers.se (A.L. Temesgen), tebikew@chalmers.se (Y.T. Wassie), erik.ahlgren@chalmers.se (E.O. Ahlgren).

[1] In this paper, the terms "grid extension" and "grid expansion" are used interchangeably.
[2] The **levelized cost of electricity** (**LCOE**), represents the average cost per unit of electricity generated that would be required to recover the costs of building and operating a generating plant over its lifetime. LCOE is used for investment planning and to compare different methods of electricity generation on a consistent basis.
[3] The term "geospatial" in this paper refers to information that is directly related to specific geographic locations.

https://doi.org/10.1016/j.esr.2023.101292
Received 11 May 2023; Received in revised form 25 November 2023; Accepted 27 December 2023
Available online 2 February 2024

factors, including but not limited to, access to roads, land cover, and topography [8]. For instance, road distance guides electrification choices in Directorate General of Energy (DGE) – Energy Division (Togo) [11]. Additionally, higher elevations entail greater expenses for construction and transportation [12]. Notably, these geospatial elements, including distance from power grid, can account for a significant portion of the initial investment cost, with estimates reaching up to 30 % [12].

Grid extension suitability in this study refers to the comparative suitability of extending the grid to different settlements, assessed by evaluating the grid extension penalty (GEP) and LCOE. GEP represents extra costs per unit grid extension length due to supply-side factors (distance from substation and road, terrain, elevation and land cover). LCOE covers total electricity costs, including capital and operating expenses. High GEP or LCOE values indicate higher extension expenses, while low values indicate better suitability for grid extension.

Moreover, the dependency of geospatial analysis outcomes on the resolution of utilized datasets bears crucial significance [13]. Notably, employment of lower-resolution datasets can potentially undermine the accuracy of the GEP multiplier and, thus, possibly result in a skewed preference for grid extension as a more favorable electrification approach. This underscores the necessity of embracing finer-scale data to capture the intricate dynamics of geospatial factors accurately.

In this context, the study by Mentis et al. [14] holds significance. The authors conducted geospatial planning studies to determine the least-cost electrification mix in Ethiopia for universal electricity access. Employing a population grid with a spatial resolution of 2.5 × 2.5 km, their findings indicate that grid extension was the cost-effective choice for nearly 89 % of the newly electrified population by 2030 for the base case scenario (for the electricity access targets 50 (rural)-300 (urban) kWh/capita/year). Similarly, Korkovelos et al. [8] undertook a comparable analysis, utilizing high-resolution population data (30 × 30 m grid size) in Malawi. Their emphasis on detailed data capture was geared towards accurately accounting for the characteristics of settlements, thereby yielding more precise estimations of grid extension suitability. Their work showed, under the baseline scenario, that by 2030 off-grid PV emerged as the least-cost option for 67.4 % of the population and grid extension only for the remaining 32.6 %.

Furthermore, although previous studies [8,12,14–16] have made strides towards factoring in geospatial parameters, none of them has illuminated the precise spatial influence of these factors on individual unelectrified settlements. Additionally, they have not probed into the cumulative weight of these factors, nor have they delineated the specific contributions of each factor to the overall suitability of grid extension. This gap in comprehensive spatial representation hampers effective settlement prioritization for grid expansion, potentially leading to inefficiencies in electrification initiatives and undermining larger-scale national strategies and investments.

Addressing the problems, therefore, requires grid-based electrification planning that is grounded on comprehensive and thorough understanding of the various geospatial factors and their influence on the suitability of grid-extension to unelectrified settlements. This will enable the government and other stakeholders to make informed electrification planning decisions. The main objective of this paper is, thus, to examine how various geospatial factors influence the suitability of grid extension, specifically in the context of areas where the population lacks access to electricity. With this objective, the paper seeks to answer the following research questions.

- How is the unelectrified population distributed in relation to the national power grid infrastructures?
- How and to what extent do geospatial factors influence the suitability of grid extension to unelectrified population settlements?
- How does the suitability of grid expansion, in terms of LCOE, vary across unelectrified settlements as electricity demand target changes?

This study is novel in two ways. Firstly, it methodologically complements previous studies on electrification planning [5,8,12,14–17] by incorporating the projected medium-voltage lines, high-resolution populations, and nighttime light data into the modeling, particularly for grid-based electrification planning. Secondly, unlike earlier studies, it explicitly investigates how geospatial factors affect grid-extension suitability across various demand levels (electrification tiers), thus contributing to the understanding of this relationship amidst other influences.

This research employs Ethiopia as a case due to its size (area of 1.1 million km^2), sizable (population over 114 million people) and considerable gaps in electricity access (only 51.1 % had access to electricity in 2020 [18]). The country also has a pronounced rural-urban disparity with only 39.4 % of the rural population having access to electricity compared to 93.2 % in urban areas. These statistics underscore the challenges the country faces in achieving universal electricity access. The Government of Ethiopia (GoE) launched the second National Electrification Plan (NEP 2.0) in 2019, targeting universal access by 2025, with 65 % through the grid and 35 % via off-grid solutions [19]. As of 2019, the government reported providing electricity to 33 % of its population through grid expansion and 11 % through off-grid solutions [19,20]. Notably, electricity tariffs in Ethiopia were subsidized at around 0.0187 USD/kWh from 2005 to 2017, rising to 0.0765 USD/kWh in 2021 [21]. Despite these relatively affordable tariffs, a major barrier to electricity access lies in the high connection fee, which amounts to USD 150 per household connection, relative to the low household income [20]. This last-mile connection cost is a key economic factor that inhibits wider access, particularly in SSA [22]. A related barrier is the high cost of electrical appliances such as CFL/LED lights, that are subject to 25 % import duty and 15 % VAT taxes [20].

The remainder of this paper is organized as follows. Section two details the methodology employed, including the choice of modeling tool and the techniques used for obtaining and processing input data. This methodology is then applied to the case country under investigation. The results are then presented and analyzed in the Results and Analysis section. Section four provides a discussion of the results together with the novelty of the developed methodology and the strengths and weaknesses of the paper. The final section presents the major conclusions drawn from the research with the answers to the research questions posed at the outset of the research.

2. Methodology

2.1. Research approach

It is evident from the introductory section that well-informed grid extension planning requires a comprehensive understanding of the influence of several factors including geophysical factors, population density, and level of electricity demand [8,14–16]. Therefore, in this study, a causal comparative research approach is applied to analyze the relationships and level of influences of the various geospatial factors on grid extension suitability in unelectrified areas of Ethiopia. Causal research, also known as causal-comparative research, identifies the extent and nature of cause-and-effect relationships between two or more variables [23]. It is typically used to determine the influence of changes in explanatory variables on the outcome variable, in this case, the influence of geospatial factors and the level of electricity demand on grid expansion suitability. In this study the approach is divided into four phases: modeling tool selection, data acquisition, data processing, and data analysis and visualization. The following subsections provide description of each phase. Fig. 1 shows a simplified flowchart of the methodology and steps followed.

2.2. Modeling tool selection

The selection of an appropriate energy-modeling tool depends on the

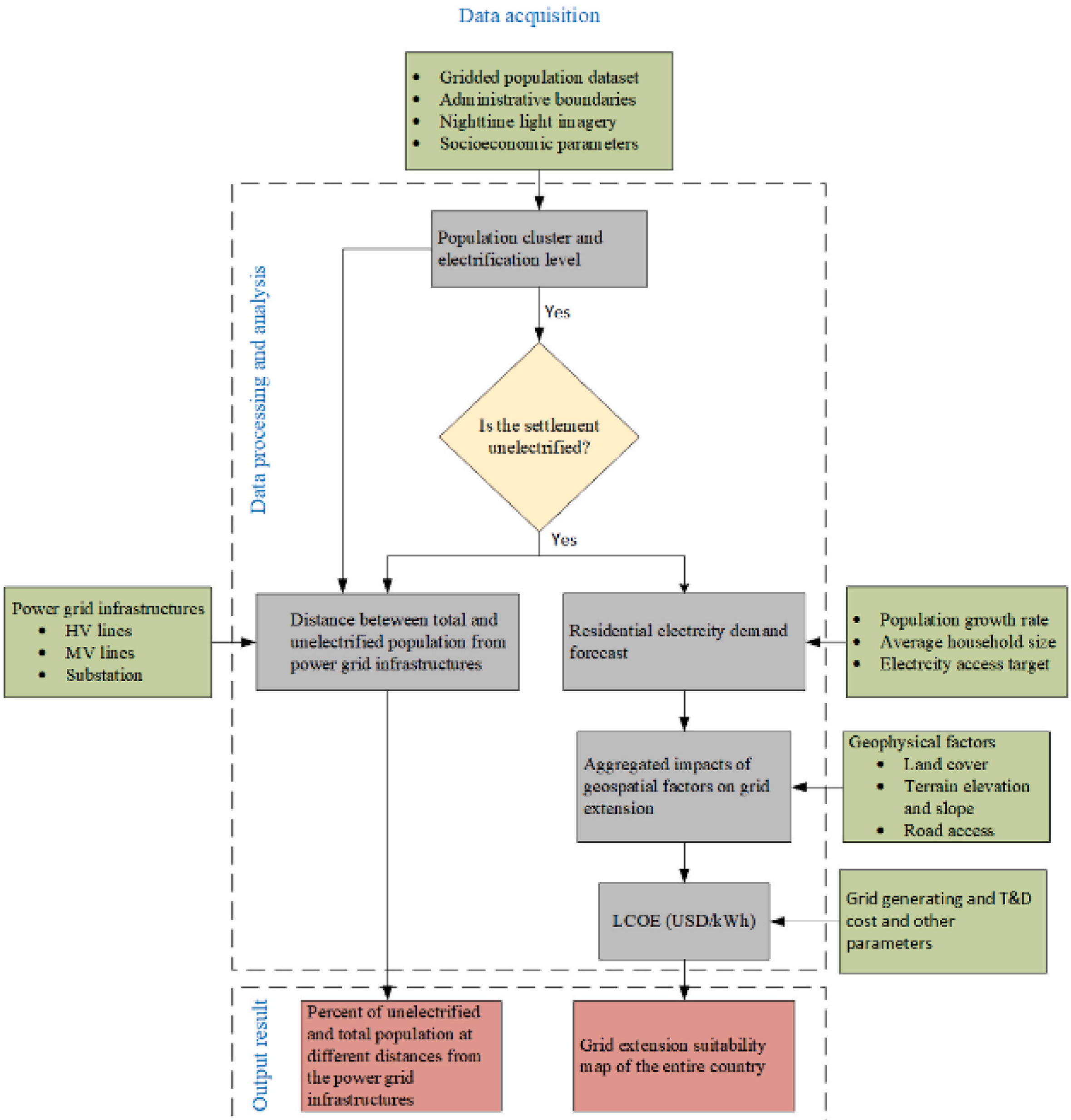

Fig. 1. Flowchart of methodological approach. Diagram color shows: olive green "input data", gold and grey "data process" and red "results of the analysis".

specific questions that need to be answered and the available data. For this particular study, the OnSSET (Open-Source Spatial Electrification Toolkit) was chosen for several reasons. Firstly, it is an open-source platform [24], making it easily accessible from the GitHub repository and customizable to tailor specific research needs. Secondly, the tool supports nationwide spatial analysis, which meets the study's objective of examining the factors that affect grid extension suitability across unelectrified settlements in the entire country. The modeling tool is programmed to determine the extent of influence of geospatial factors and identify areas that are suitable, semi-suitable, and unsuitable for grid extension. This modeling tool is applied to Ethiopia.

Thirdly, in contrast to traditional energy modeling tools, such as TIMES, MESSAGE, and OSeMOSYS, OnSSET has the unique capability of accounting for topography-related costs in grid extension-based electrification pathways [15]. Traditional energy supply models typically lack this capability of accounting for location-specific peculiarities [25]. Fourthly, OnSSET is thoroughly documented in the literature [8] and on the developers' website [26]. All these qualities make OnSSET a preferred choice for nationwide spatial analysis of electrification efforts, particularly in developing countries like Ethiopia with diverse terrain and varying degrees of access to electricity.

2.3. Data acquisition

The second step in conducting this study is to acquire the necessary input GIS datasets and other parameters presented in Appendix Tables A1, A2, A3, A4, and A5. The datasets used in this study are obtained from open-source platforms. These datasets are further discussed in detail in the following subsections.

2.3.1. Gridded population dataset

The gridded population dataset serves as the basis for generating population clusters/settlements [17]. This data is used to determine where populations are situated within the study area's territory. There are various free, open-source gridded population datasets available, like the Global Human Settlement Layer (GHSL), WorldPop, and the

High-Resolution Settlement Layer (HRSL). This study used HRSL as it has demonstrated greater prowess in recognizing the footprint of built-up areas/buildings in both urban and rural settings than GHSL and WorldPop population dataset [27], owing to its resolution of approx. 30 m (1 arc-second) grid size and the use of subnational census data together with high-resolution (0.5 m) satellite imagery [27,28]. However, since the most recent population data from the 2020 census were not available, population data from 2018 were utilized and adjusted to reflect the total population in 2020. This raster dataset is produced by a joint venture between Meta (formerly Facebook) Connectivity Lab and Columbia University's center for International Earth Science Information Network (CIESIN). Fig. 2 discloses Ethiopia's 2018 HRSL population. It is divided into five categories that differentiate between densely and sparsely populated areas.

2.3.2. Nighttime light imagery

We utilized nighttime light (NTL) imagery in conjunction with population data, following the approach proposed by Falchetta et al. [29], to estimate electrification status of settlements and, consequently, the national electrification rate. This methodology has demonstrated effectiveness in previous research [30–34]. The satellite imagery dataset was sourced from the National Oceanic and Atmospheric Administration (NOAA) and comprises cloud-free composite VIIRS nighttime lights (NOAA/VIIRS/DNB/MONTHLY_V1/VCMSLCFG). As the satellite detects and gathers nocturnal light from a variety of sources, including lights from fires, flares, the sun and moon, boats, and blooming effects surrounding major cities, the best result of electricity rate estimation is obtained by using a dataset that has been cleaned of such noise [14].

To reduce short-term fluctuations and enhance data quality, we generated the 2020 annual composites of NTL from the VIIRS-DNB straylight-corrected monthly composite images utilizing the Google Earth Engine (GEE) platform, which is a cloud-based computing platform that allows users to analyze geospatial data based on Earth science data [34]. Following the approach by Falchetta et al. [29], we processed annual median composite scenarios using five different lower-bound noise floors: 0.25, 0.27, 0.28, 0.30, and 0.35 $\mu W\ cm^{-2} \cdot sr^{-1}$. These values represent the minimum intensity below which any light is considered noise or ephemeral. This ensures that only electricity-generated light is included in our analysis. We tested various noise floors in an effort to identify the one that offered the most precise estimation of Ethiopia's national electrification rate in comparison to the World Bank's reported electrification rate. Fig. 3 displays the NTL composite for the case study country. The data in this figure represents the annual median radiance, and it has been filtered using a threshold of 0.27 $\mu W \cdot cm^{-2} \cdot sr^{-1}$. Additionally, Appendix Figure A1 provides further insights into nighttime light variations using both the minimum threshold (0.25 $\mu W\ cm^{-2} \cdot sr^{-1}$) and the maximum threshold (0.35 $\mu W\ cm^{-2} \cdot sr^{-1}$). These thresholds provide insights into variations in nighttime light levels within the case study country.

The population density map (Fig. 2) aligns closely with the processed nighttime light (NTL) map (Fig. 3), highlighting a strong correlation between population density and NTL intensity. Higher population density corresponds to brighter NTL, signifying greater electrification, while lower population density areas exhibit dimmer or no NTL, indicating lower electrification levels. This connection emphasizes NTL's effectiveness as a proxy for assessing electricity access where detailed data may be limited.

2.3.3. Administrative boundaries

The administrative boundaries were used to extract the nighttime light data from the NOAA VIIRS satellite data and delimit the analysis to the study area. By doing this, only the nighttime light data inside the administrative boundaries is accounted for in the analysis. This also ensures that the population cluster is contained within these boundaries, thus limiting the maximum area of the cluster [35]. In this study, the entire country territory (level 0) was applied for clipping the annual nighttime light raster data, and the first-level administrative boundaries (level 1) were employed for generating the population cluster. Both datasets were obtained from the Global Administrative Areas (GADM) version 4.0. The GADM gives high-resolution administrative boundaries

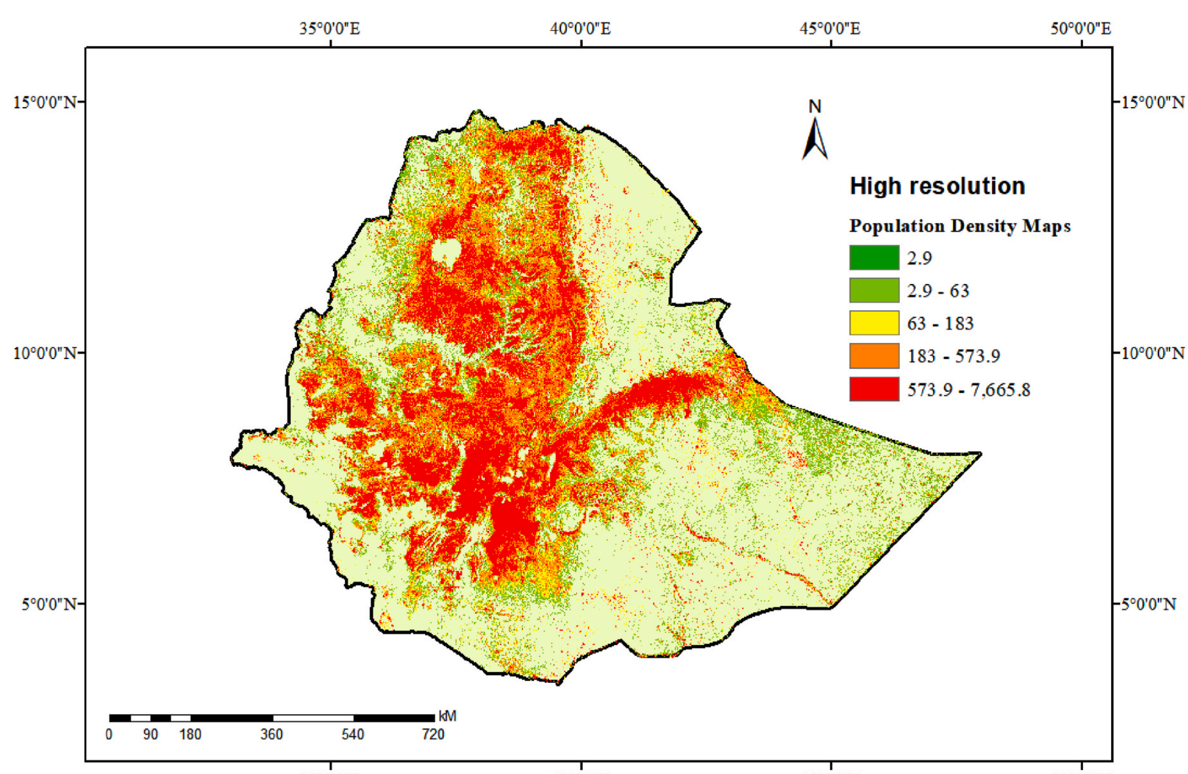

Fig. 2. Population distribution in Ethiopia (at a resolution of 30 m) based on HRSL 2018 raster data [28].

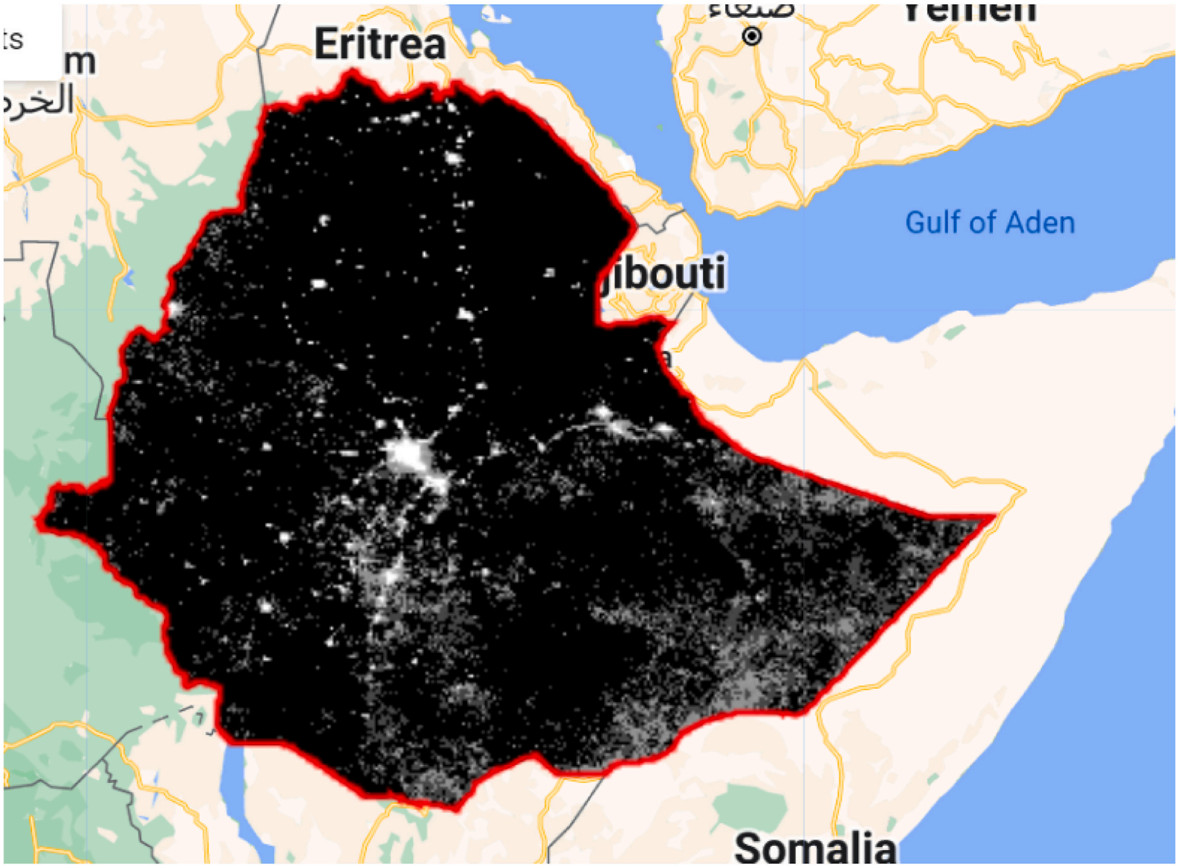

Fig. 3. Annual NTL composite in Ethiopia, filtered using 0.27 μW $cm^{-2} \cdot sr^{-1}$.

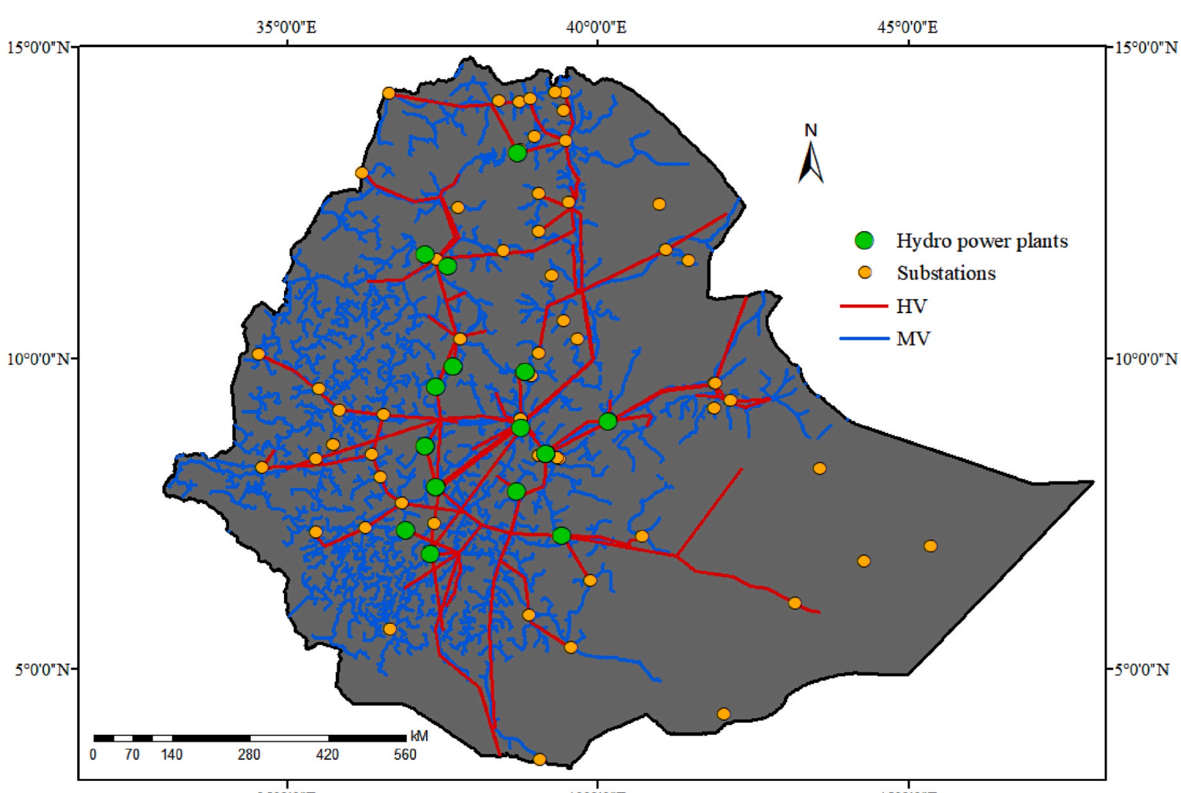

Fig. 4. Existing HV and projected MV lines, and substations in Ethiopia [37,40].

at different levels using the WGS84 datum coordinate reference system and longitude/latitude [36].

2.3.4. Power grid infrastructures

The spatial distribution of the existing power grid infrastructures was used in order to examine where the unelectrified population lives in relation to it. To that end, publicly accessible georeferenced datasets of High Voltage (HV) transmission lines, projected Medium Voltage (MV) lines, and substations, were used as input data. These datasets can be obtained from Energydata.info, an open data platform created by the World Bank and its partners with the aim of filling a data gap in the energy sector [37]. The Energydata.info portal sources its data from multiple reliable and verified sources, such as Africa Infrastructure Country Diagnostic (AICD), OpenStreetMap (OSM) contributors, and World Bank projects archive and IBRD maps, and international organizations [38]. The data used for this case study was updated in 2017.

It is worth noting that the availability of these datasets varies by country and the type of infrastructure. In specific instances, such as the case study area, the MV lines dataset is unavailable on the energydata.info platform. Consequently, we sourced this information from the Gridfinder web application [39]. This application uses machine learning and publicly available datasets, such as nighttime data and MODIS land cover data to drive MV network distribution lines. The Gridfinder Tool estimates the MV lines by utilizing a many-to-many variant of Dijkstra's algorithm and integrating road networks as a cost function. The algorithm aims to predict distribution networks based on the assumption that lines tend to follow roads [39]. Fig. 4 displays the HV, projected MV lines, and substations in the study country.

The red lines signify HV lines linking remote hydropower plants (depicted in green) to substations (highlighted in orange). The blue lines, on the other hand, depict the distribution of the projected Medium Voltage (MV) lines within cities and towns. These substations encompass both HV to MV converters and diesel-based units.

2.3.5. Land cover

To evaluate the suitability of land cover for grid extension, we employed the MODIS/Terra + Aqua Type MCD12Q1 Version 6 land cover dataset, obtained from Earth Data [41]. This dataset provides global land cover information at a resolution of 500 m [42]. Maps were created using data from the Moderate Resolution Imaging Spectroradiometer (MODIS), and land cover classification values were made based on a range of 1–17 [26]. Fig. 5 shows the land cover in Ethiopia in 2020, classified into different categories based on the MODIS/Terra + Aqua data. When calculating the GEP, the OnSSET GIS extraction model reclassifies these values, with corresponding weights listed in Table A2.

This study utilized a reclassification approach [12], detailed in Appendix Table A2, to assess the suitability of different land cover types for grid extension. This method integrates the Analytic Hierarchy Process to determine the significance (weight) of each land cover type in terms of grid extension suitability. Notably, land cover types like open shrublands, cropland/natural vegetation mosaics, savannas, grasslands, and barren areas are deemed highly suitable for grid extension, while water bodies and permanent wetlands are considered the least suitable for such expansion.

2.3.6. Elevation and terrain slope

To assess the influence of elevation and terrain slope on the suitability of grid extension [11], the Advanced Spaceborne Thermal Emission and Reflection Radiometer Global Digital Elevation Model (ASTER GDEM) elevation dataset is used. This dataset, produced by the Japanese Ministry of Economy, Trade, and Industry and NASA, uses stereo-pair images from the Terra satellite to generate a digital elevation model with a horizontal resolution of 15 m [8,43]. The OnSSET GIS extraction model then utilizes this information to create a map of terrain slope, a sub-product of DEM, which is used to identify restriction zones and determine the suitability for grid extension.

2.3.7. Road access

The impact of road infrastructure access (major roads, highways,

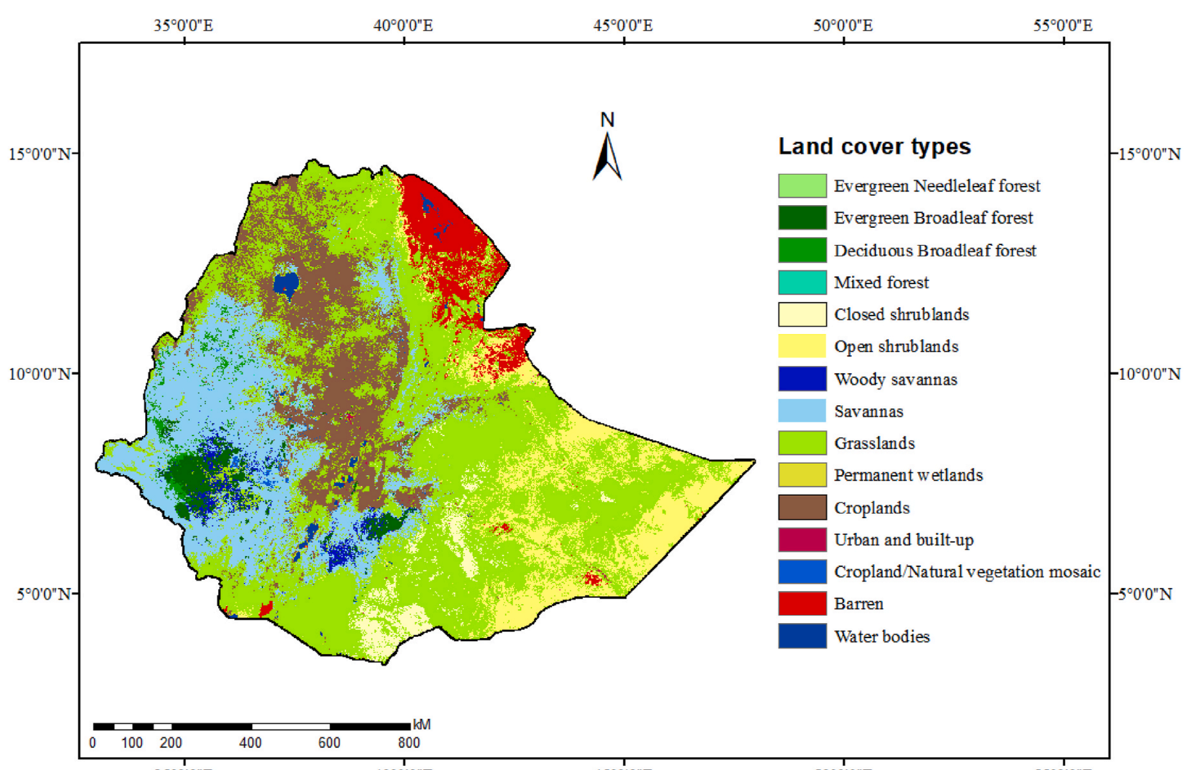

Fig. 5. MODIS/Terra + Aqua land cover types in Ethiopia in 2020 [42].

primary, and secondary roads) on the grid extension suitability is assessed by evaluating its proximity to roads [9,11]. The road scale should at least allow for the use of pickup/trucks [8]. The road data used in this study is obtained from the OpenStreetMap project via Geofabrik's server, which is updated daily [44].

2.4. Data processing

2.4.1. Population cluster and electrification rate

At this stage, the population clusters are generated, and distribution of electrified and unelectrified settlements are identified. A population cluster also known as a settlement is a geospatial unit representing human settlements formed by contiguous groups of people, which serves as a basis for the analysis. Population clustering in this context is a technique for converting high-resolution raster population data into vector-based population clusters with unique characteristics indicating population, electrification status, and urban-rural categorization [45]. The following three GIS datasets are utilized to generate the population cluster and ascertain the current electrification status [26]; more information about the datasets is given in Appendix Table A1.

- Gridded population dataset (raster layer)
- Nighttime light imagery (annual median nighttime light (raster)), cleaned from background, biomass burning, cloud cover, stray and aurora light
- Administrative boundaries (vector layer)

To establish population clusters for analysis, researchers have several options available. Two common methods include using the PopCluster QGIS plugin developed by KTH-dES [35] or the Clustering method written in Python developed by Khavari et al. [27] and available on the GitHub repository [45]. Both methods allow the user to set a threshold for the population layer below which is counted as zero. Alternatively, the second method allows to set thresholds for both the population layer and the nighttime light. It also allows users to calibrate population data using "start year" (base year) population figures. At the end of this process, the population clusters are polygonized within the inner administrative border to ensure that the clusters do not spill over into different administrative territories, enabling leaders and policymakers to focus on a specific area/region [27].

In this study, the Clustering method [27] was used. The simplified

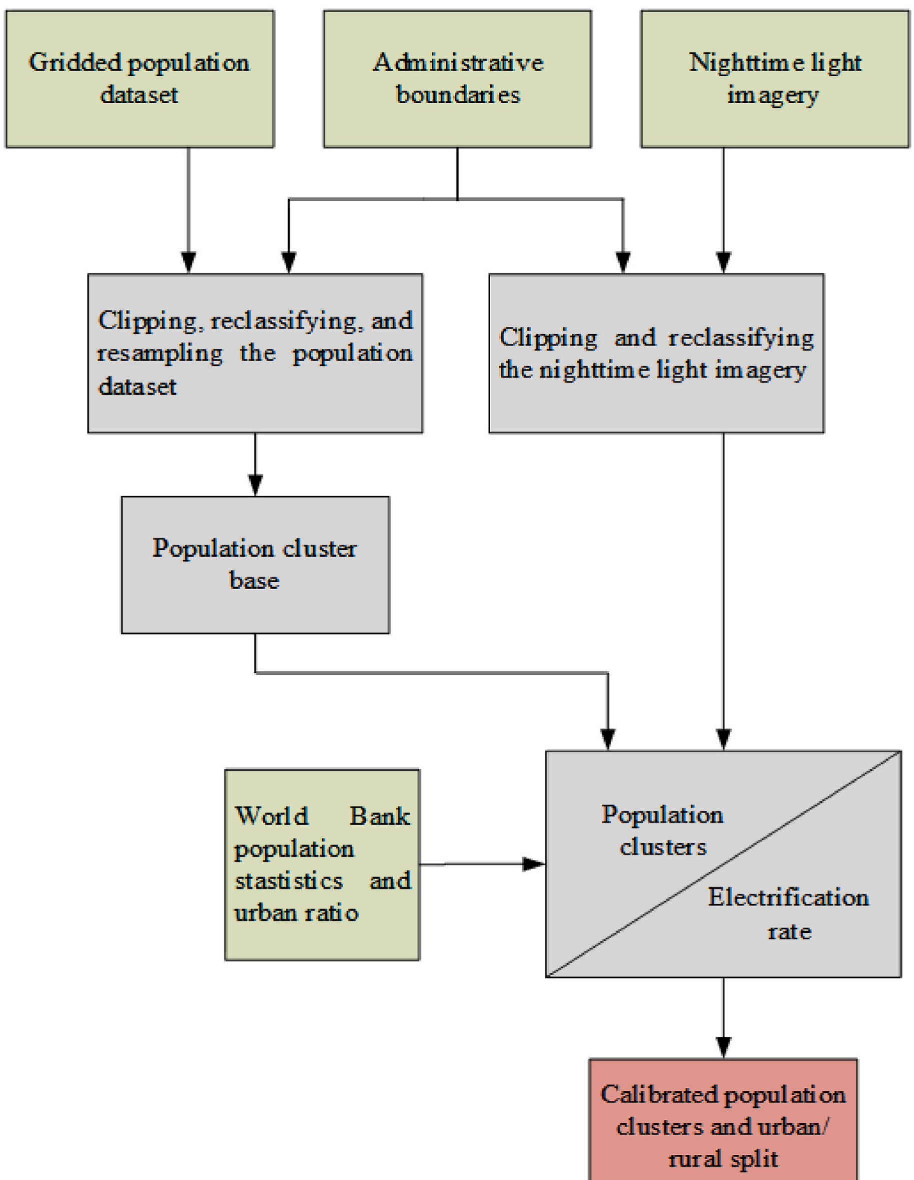

Fig. 6. The simplified population clustering process diagram, adopted from Ref. [27].

steps of the process are depicted in Fig. 6. A threshold of zero was used for both the population and nighttime light data in the analysis. The population raster dataset (a raster with cell size of 30 m) is resampled to produce an output raster with cell size of 90 m [45]. Clusters are created by multiplying a resampled population with rasterized disaggregated administrative boundaries and then converting the resulting raster layer into polygons, each polygon representing a cluster. The size of the clusters is determined by the population's spatial distribution and administrative boundaries. Each cluster denotes a specific region/location. The Clustering method allowed us to analyze the demographic composition of each cluster, including the number of people, their electrification status, and the urban/rural split.

In order to detect already electrified populations, OnSSET uses a heuristic approach utilizing a combination of spatial nighttime light brightness and population datasets to identify lit clusters and classify them as electrified (1) or unelectrified (0) [8], [14], [26]. Specifically, a settlement is considered as electrified if it fulfills the criteria of having a population density of at least 1 person per square kilometer and a nighttime light brightness value exceeding the minimum threshold described in Section 2.3.2. Then the number of electrified people is determined by summing the population with the count of individuals within each cluster who inhabit areas where light sources are detected, signifying those with access to electricity [45]. This enables us to assess access to electricity at a cluster level, information that is often difficult to obtain. The rate of access to electricity at the national level for the year 2020 is then computed with equation (1). The Clustering method relies on thresholds used by Eurostat to classify settlements into urban and rural areas [27].

$$\text{Access to electricity at national level} = \frac{\sum_{i=1}^{n} \textit{Electrified population in each cluster}_i}{\sum_{i=1}^{n} \textit{Population in each clusters}_i} x100\% \tag{1}$$

where i and n represent the population cluster.

Accordingly, the pixelated population of Ethiopia was divided into 809,087 clusters by the OnSSET Clustering method. This was accomplished by merging eight adjacent grid cells of the HRSL population dataset.

2.4.2. Distance from the power grid

The OnSSET grid extension algorithm begins by identifying all population settlements within the study area. Each settlement and power grid infrastructure (HV line, projected MV line, and substation) is denoted by its x and y position coordinates. The algorithm then iteratively traverses a subset of the tree nodes using Locality-Sensitive Hashing (LSH) to identify the nearest HV line, projected MV line, and substation to each settlement [8]. LSH is a method for approximating the nearest neighbor search problem in high-dimensional spaces. It works by creating a hash table of the data points, and then using the hash values to quickly identify points that are likely to be close to the target point.

Once the nearest power grid infrastructure (HV line, projected MV line, or substation) is identified for a population settlement, the algorithm uses the coordinates of both the settlement and the power grid infrastructure to calculate the distance between them. This allows the algorithm to determine the proportion of unelectrified and total population in relation to the grid infrastructure by computing the ratio of the sum of unelectrified population to the total population of the country and the ratio of the total population located at various distances from the grid infrastructure to the total population of the country, respectively.

2.4.3. Residential electricity demand

Population projections for 2030, along with different electricity access targets, referred to as "tiers," are used to determine future residential electricity demand scenarios. We used the current average electricity consumption per household in Ethiopia, approx. 335 kWh/year [46,47], as a benchmark to adapt two consumption tiers. From the Global Tracking Framework [48], we adapted tier 1 (38.7 kWh/household/year) and tier 3 (803 kWh/household/year), which are lower and higher than Ethiopia's average consumption, respectively. These tiers are assumed to be applied uniformly across the country. Initially, two separate access tiers were considered for urban and rural areas. However, the author notes that the urban areas of the country have nearly achieved electricity access, allowing for the use of a single, uniform tier without sacrificing any important information. The electricity demand for unelectrified settlements is calculated by multiplying the target electricity access tier by the projected population for the year 2030, taking into account the average household size in rural and urban areas.

2.5. Data analysis and visualization

At this phase, the suitability of grid extension to unelectrified settlements was evaluated by analyzing the GEP and LCOE. Accordingly, a spatial grid extension suitability map showing the distribution of areas most suitable, semi-suitable and least suitable for grid-based electrification is developed.

2.5.1. Aggregate effects of geospatial factors

The combined and weighted contribution of geospatial factors to grid extension suitability is estimated through Equation (2). Using this equation, we also evaluated the weighted contribution of each geospatial factor to the total grid extension suitability score. The input parameters were first classified between 1 and 5, with higher values indicating more suitable conditions, as shown in Table A2. The weight assigned to each geospatial factor represents the importance of that feature in determining the grid expansion penalty (GEP). The default weights provided by OnSSET were adopted in this study based on their extensive usage by the Global Electrification Platform in analyzing electrification strategies for 58 countries, many of which are SSA countries including Ethiopia [49]. However, the Analytic Hierarchy Process can be used to determine the weight assigned to each factor [12]. The GEP multiplier for each settlement is calculated using Equation (3) [26]. This equation calculates the grid penalty to increase the grid cost in areas with high slope angles, unsuitable land cover, higher substation distance, higher road distance, or high elevation, [8,14,15].

$$\begin{aligned} \textit{Combined classification} = (&0.30 * \textit{slope angle} + 0.20 \\ &* \textit{land cover suitability} + 0.20 \\ &* \textit{substation distance} + 0.15 * \textit{road distance} \\ &+ 0.15 * \textit{elevation height}\ ) \end{aligned} \tag{2}$$

$$\textit{Grid expansion penalty (GEP)} = 1 + \frac{e^{(0.85*|1-\textit{Combined classification}|)} - 1}{100} \tag{3}$$

The GEP is a crucial factor in the optimization of the grid extension process, as it estimates the additional cost (penalty) of extending the grid to unelectrified settlements. Settlements with high GEP values are considered less suitable for grid extension due to their higher penalty, while those with low GEP values are more suitable. To visually represent this information, the GEP was then mapped and settlements were categorized into three classes of suitability (most suitable, semi-suitable, and less suitable) based on their GEP values, using quantiles. This map may provide decision-makers and energy companies with a better understanding of areas where grid expansion would be relatively feasible from the supply side perspective.

2.5.2. LCOE

The LCOE for each population settlement was calculated using a cost model following Nerini et al. [50] and Mentis et al. [12]. The LCOE calculation takes into account both supply and demand side factors including capital expenses, operating costs, and expected project lifespan. In the LCOE calculations, we considered three key parameters interlinked with costs: I) the targeted level of electricity demand expressed in kWh/household/year; II) population density, quantified in households/km2; III) local grid connection specifics, encompassing the distance from the nearest grid (km), translatable into wire costs and the average national cost of grid electricity ($/kWh). Details of the cost model are given in the Appendix Table A4 and Table A5. Based on the LCOE values, settlements are classified into three suitability categories as most suitable, semi-suitable, and less suitable. Settlements with a low LCOE value were deemed most suitable for grid extension while those with an intermediate LCOE value were considered semi-suitable, and settlements with a high LCOE were considered less suitable. Finally, the data analysis results are presented in maps, graphs, and figures for better understanding and easier visualization.

It is evident from the detailed description of the methodology section that this study is based on high-quality and multiple datasets to the greatest extent possible, including high-resolution population datasets, NTL datasets, varying demand levels; and as many robust and relevant data analysis methods as possible, including GEP, LCOE, and visual-spatial representations.

Table 1
Geospatially estimated electricity access at different noise threshold floors.

NTL data noise floor ($\mu W \cdot cm^{-2} \cdot sr^{-1}$)	Estimated electricity national access	Urban and rural access respectively	Compared to the national statistics from World Bank
0.25	69.1 %	98.7 % and 61.4 %	Overestimated (+18 %)
0.27	51.3 %	97 % and 39.3 %	Overestimated (+0.2 %)
0.28	44.1 %	96 % and 30.4 %	Underestimated (−7%)
0.30	33.8 %	93.4 % and 18.2 %	Underestimated (−17.3 %)
0.35	24.6 %	88.4 % and 8.1 %	Underestimated (−26.5 %)

The results indicate that a noise floor of 0.27 μW cm^{-2}·sr^{-1} provides the best estimate, with a slight overestimation of the national electrification rate by 0.2 % compared to the World Bank's reported rate of 51.1 % for Ethiopia. In contrast, a noise floor of 0.28 μW cm^{-2}·sr^{-1} underestimates electrification rates by 7 %. The study highlights the importance of selecting an appropriate noise floor when using NTL data for estimating access to electricity. Additionally, the study suggests that using a noise floor of 0.30 μW cm^{-2}·sr^{-1} or higher can provide better accuracy in determining access to electricity in urban areas, while a noise floor of 0.27 μW cm^{-2}·sr^{-1} or higher yields better results for estimating access in rural areas.

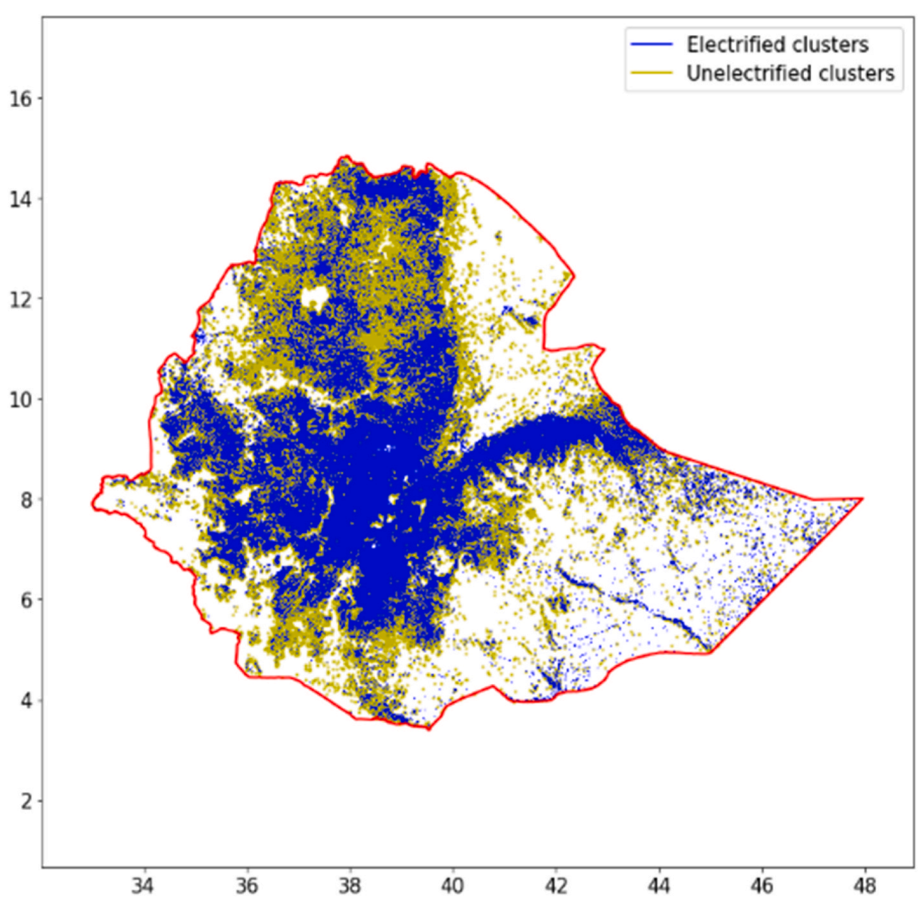

Fig. 7. OnSSET based electricity access, applying 0.27 μW cm^{-2}·sr^{-1} lower bound noise floor, resulting in an estimated electricity access of 51.3 %.

3. Results and Analysis

3.1. Geospatially estimated electrification rate

Fig. 7 provides a visual representation of the distribution of electrified and unelectrified populations at the cluster level, as determined by the methodology described in section 2.4.1. The blue areas indicate the electrified population, while the yellow areas correspond to the unelectrified population. To accurately estimate the electrification rate, the study applied five different noise filter floors to find the noise floor that minimizes the discrepancy between the estimated access to electricity and the World Bank report, as explained in section 2.3.2. Table 1 presents the estimated access to electricity obtained from these filters, which varies from 69.1 % at the lowest noise filter value (0.25 μW cm^{-2}·sr^{-1}) to 24.6 % at the highest noise filter value (0.35 μW cm^{-2}·sr^{-1}).

Fig. 8 shows the subnational electrification estimates of the 11 regional states/administrative regions. These results are compared with the regional electricity access rates available in the GoE NEP 2.0 report, specifically for the five regional states. The subnational electrification estimation results generally align with the GoE NEP 2 report (see Figure A3.2 in [20]), with variations observed in specific regions. For instance, Addis Ababa's perfect electrification estimate of 100 % closely corresponds to the GoE NEP 2's 99.9 % rate. While Amhara's 41.8 % estimate falls below GoE NEP 2's 51.3 %, a correlation is still evident. Conversely, our model's estimates for Oromia (69.20 %) and SNNP (71.49 %) surpass the GoE NEP 2 rates of 63.3 % and 37.9 %, respectively. The Tigray variation is notable, with our estimate at 61.08 % compared to the GoE NEP 2's 87.7 %. These comparisons highlight the model's insights and the significance of using diverse sources for comprehensive assessments.

3.2. Power grid infrastructures proximity to population settlements

The results of the study presented in Section 2.4.2 indicate that a significant portion of Ethiopia's population resides in close proximity to power grid infrastructure. Specifically, Fig. 9 shows that 85 % of the population lives within 10 km of the projected medium-voltage (MV) transmission lines, while 45 % of the population lives within the same distance of the high-voltage (HV) lines. Combining the two figures reveals that 87.7 % of the population lives within 10 km of either HV or the projected MV lines. Additionally, the study shows that almost all electrified population in Ethiopia is located within 5 km of existing grid lines, while nearly all the unelectrified populations are located within a 25 km from existing grid lines. These findings have important implications for the planning and prioritization of electrification projects in Ethiopia, particularly with regards to extension of the power grid.

Unfortunately, this study did not include off-grid electrification solutions such as mini-grids due to lack of reliable nationwide data. As a result, the geographic overlap between the grid and off-grid systems, as well as the distribution of Ethiopia's population relative to these off-grid power generation and distribution systems are not analyzed in this study.

Limited access to data on MV lines presents a challenge in assessing their contribution to electrification planning. However, the inclusion of projected MV lines in our analysis improves the understanding of the

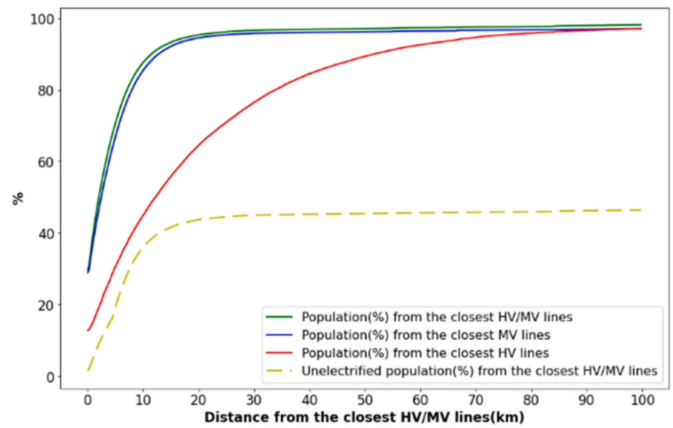

Fig. 9. Share of the population as a function of distance from the closest HV/ projected MV line. The total share of the unelectrified population is 48.7 %.

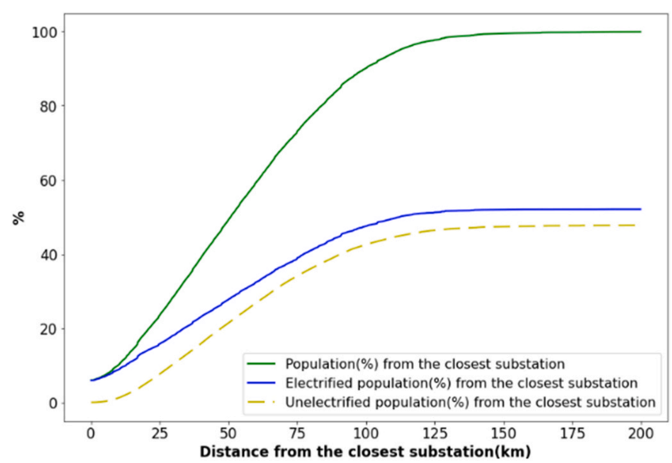

Fig. 10. Share of the population as a function of distance from the closest substation. The total share of the unelectrified population is 48.7 %.

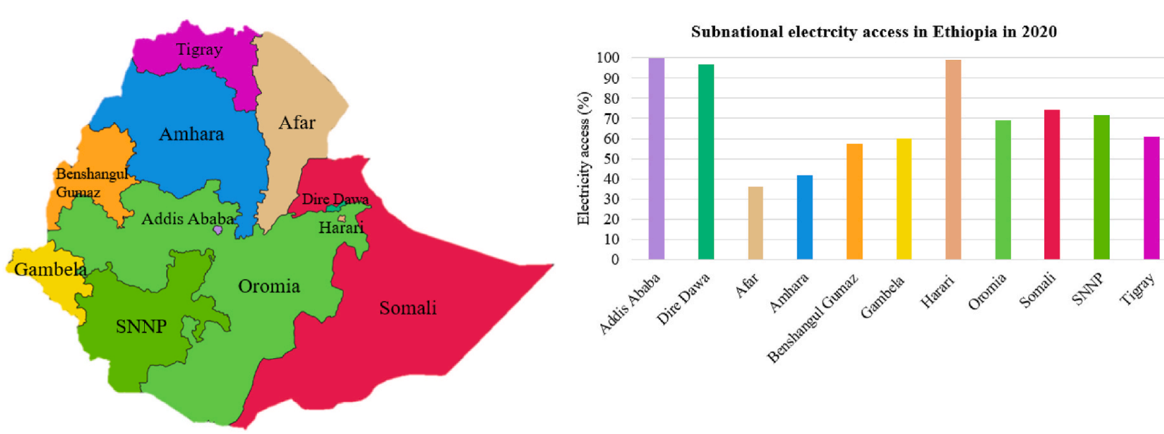

Fig. 8. Estimated access to electricity at subnational level.

distribution of the total and unelectrified population in relation to the existing grid infrastructure. This strengthens the robustness of our findings and provides a clearer picture of the proportion of the population residing in close proximity to the grid. In addition, the use of high-resolution population data (30 m) enhances the accuracy of our analysis, enabling the identification of areas without access to electricity at a finer resolution.

Fig. 10 presents the percentage of the total and unelectrified population within different distances from the closest substation. The results reveal that more than 90 % of the total population and nearly all the electrified population are located within 100 km of the nearest substation. Furthermore, almost all the unelectrified population is located within 125 km of the nearest substation.

3.3. Grid extension suitability and penalty

Fig. 11 illustrates the relative contribution of the five geospatial factors to the grid extension suitability, as determined from equation (2) in section 2.5.1. The results indicate that terrain slope contributes the most to the grid extension suitability, accounting for 40.8 % of the total score. This is because the majority of settlements in Ethiopia, 94 %, have a terrain slope of less than 10°, which is considered the most suitable class in this study, as presented in Table A2 in the Appendix. In contrast, the distance from the substation contributes the least (5.7 %) to the suitability score. This is because only a small fraction of the settlements falls under the most suitable class, which indicates that the distance from the substation has a high contribution to the GEP multiplier, resulting in a significant impact on the capital cost of grid extension.

The impact of the five geospatial factors on grid extension costs was evaluated by calculating the GEP multiplier using Equation (3). According to the analysis results, the combined impact of these factors could raise the cost of grid extension across unelectrified settlements in Ethiopia by an average of 8.6 %, ranging from 2.3 % to 29 % depending on the location. The variability in GEP values across settlements is likely due to location-specific factors, implying that the cost of grid extension may vary based on the specific conditions of the settlement. The GEP multiplier is also used as an indicator of the suitability of grid extension, as a penalty is the inverse of suitability. Settlements with a lower GEP value are more suitable for grid extension, while those with a higher GEP value are less suitable. Fig. 12 presents a spatial representation of the GEP multiplier. The figure shows the relative suitability of extending the grid network across the country. Settlements with a GEP value below 1/3 quantile (green) are deemed most suitable for grid extension, those with a GEP value in the range between 1/3 and 2/3 quantiles (yellow) are semi-suitable, and those with a GEP value in the upper 2/3 quantile (red) are unsuitable for grid extension due to the high penalty cost involved.

The most suitable class of settlements requires lower investment costs than other settlements, while the less suitable class of settlements would incur high investment costs. There is a significant variation in the distribution of grid extension suitability across the country.

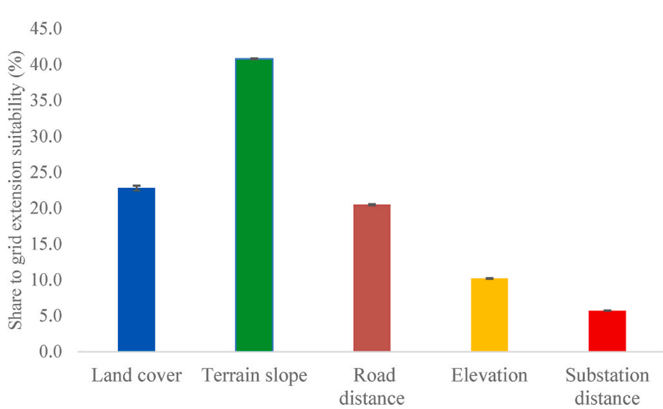

Fig. 11. The contribution of each geospatial factor to the grid extension suitability.

The semi-suitable locations are distributed throughout the country. It is worth noting that proximity to electrified settlements is not the only determining factor for grid extension suitability. There are areas near electrified settlements but not suitable for grid extension due to other geospatial factors than just distance. On the other hand, there are also areas that are semi-suitable for grid extension despite being far from electrified settlements. This is due to other factors, such as favorable local topography, suitable land cover, and convenient road access.

3.4. Grid extension suitability in terms of LCOE

Fig. 13 presents the findings of the geospatial analysis using the LCOE as a primary metric for grid extension suitability. The LCOE was calculated for two distinct electricity access targets: tier-1 (38.7 kWh/household/year) and tier-3 (803 kWh/household/year) as adapted from the World Bank's Global Tracking Framework [48]. The results shown by the two maps in Fig. 13 provide a comprehensive visual representation of the regions that are most suitable (green), semi-suitable (yellow), and less suitable (red) for grid extension for both access targets.

Semi-suitable locations are scattered across the country. Fig. 13 also shows that as the demand target increases from tier-1 to tier-3, the proportion of people falling under the most suitable category increases by 13 %. Conversely, the proportion of people falling under the less suitable category decreases by 62 % as the demand shifts from tier-1 to tier-3.

4. Discussion

4.1. Distribution of the total and unelectrified population in relation to the grid

Approximately 87.7 % of Ethiopia's population resides within 10 km of the HV or projected MV grid lines, with 85 % within the 10 km distance of the projected MV lines. Our findings also show that almost the entire unelectrified population (48.7 % of the total) resides within 25 km of power grid lines. This finding is in agreement with the Ethiopian government's NEP 2.0 report, which states that 90 % of the population lives within 10 km distance from the existing MV lines [20]. A study by Arderne et al. [39] using nighttime light data to map power system networks calculated that 97 % of the world's population is located within 10 km from an MV line. The authors, however, noted that this estimate varies greatly among regions and income levels, with SSA having the highest proportion of individuals beyond 10 km from an MV line. Our results suggest that the distribution of unelectrified population within a reasonably short (10 km) distance from HV/the projected MV lines could make grid extension viable, considering other factors and available capital.

4.2. National electrification rate estimation using geospatial analysis

The electrification rate estimation using the NTL imagery and population data shows that for a lower bound noise filter of 0.27 μW $cm^{-2}{\cdot}sr^{-1}$, the estimated electricity access rate is relatively accurate with +0.2 % margin of error compared to the World Bank report. However, a 0.35 μW $cm^{-2}{\cdot}sr^{-1}$ noise floor leads to larger discrepancies with a −26.5 % margin of error. This result is in line with earlier studies [27,29], which similarly underestimated electricity access rate in Ethiopia by 26.8 % and 20 %, respectively. It is worth noting that the use of a noise floor less than 0.35 μW$\cdot cm^{-2}{\cdot}sr^{-1}$ in estimating electrification rate resulted in estimates higher than those reported in previous studies by Falchetta et al. [29] and Khavari et al. [27]. The underestimation of the electricity access rates, when lower bound noise floors of above 0.27 μW $cm^{-2}{\cdot}sr^{-1}$ are used, could be due to several factors, including under

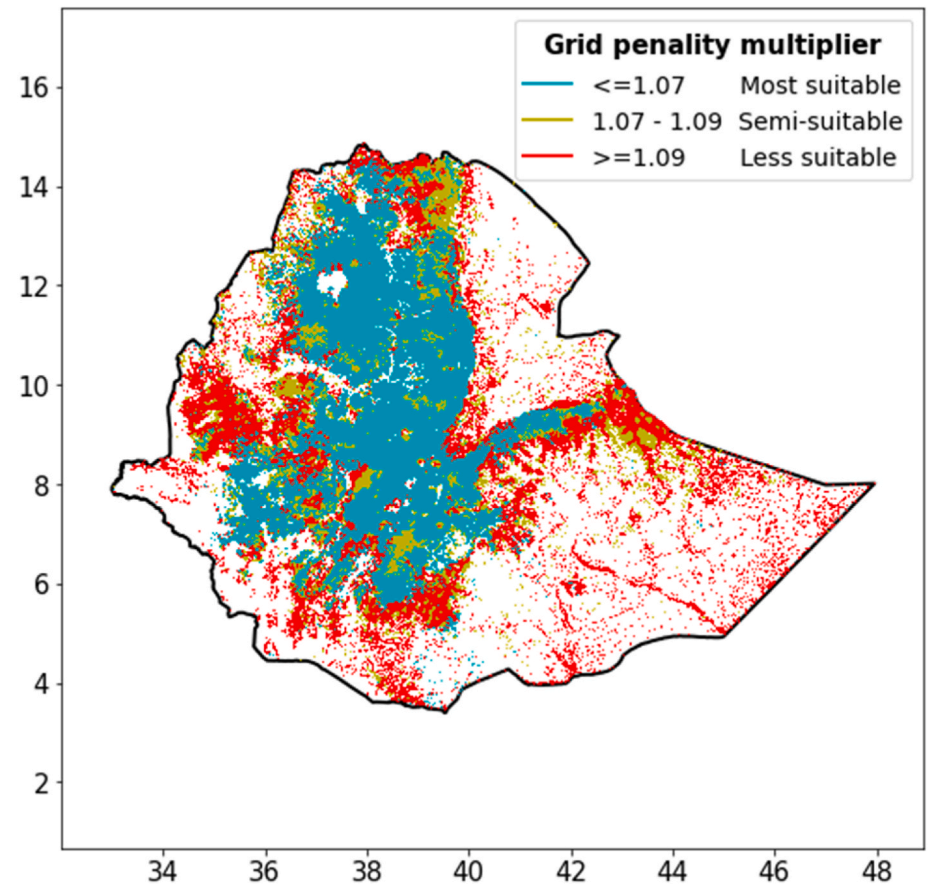

Fig. 12. Grid extension penalty (GEP) map.

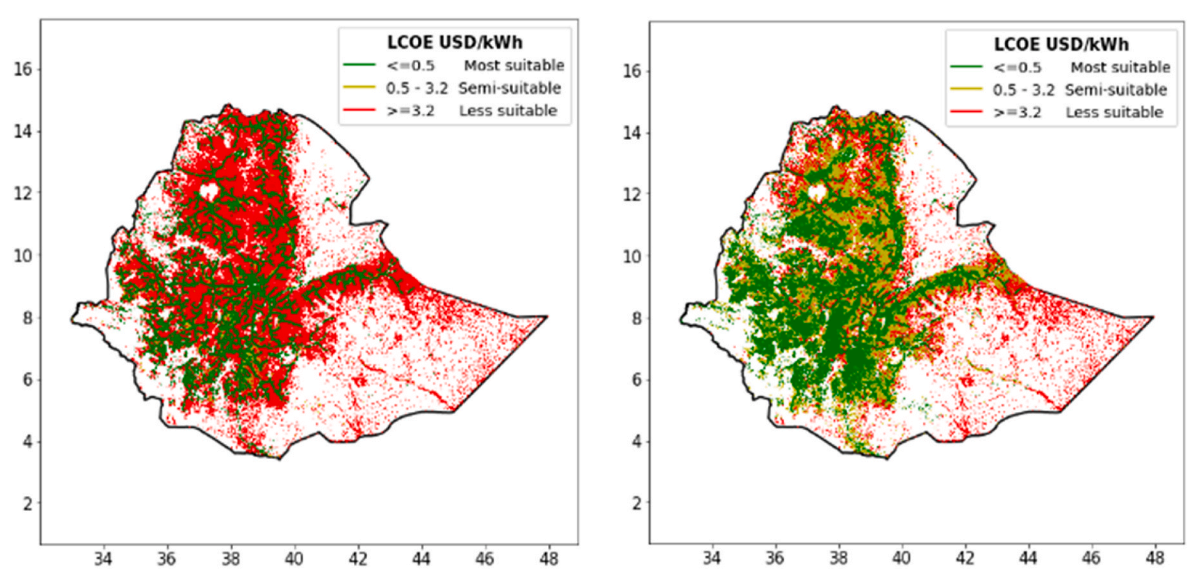

Fig. 13. Relative grid extension suitability in Ethiopia for two different electricity demand targets (tiers): 38.7 kWh/household/year (left) and 803 kWh/household/year (right).

detection of standalone solutions such as solar-home-systems (SHS), or low levels of lighting in highly distributed communities [29]. The method's reliability is indicated by its alignment with the World Bank's data, affirming its practical application potential.

4.3. The influence of geospatial factors on grid extension suitability

Our findings show that among the considered geospatial factors, terrain slope and land cover contribute the most to the suitability of grid extension to unelectrified settlements (see Fig. 11). Terrain slope has the most significant contribution to the grid extension suitability or contributes the least to the GEP multiplier, compared to the other four geospatial factors. Land cover is the second significant factor that contributes to the suitability of grid extension to unelectrified settlements. This is because the majority of settlements in Ethiopia are located in cropland, grassland, and savanna areas and according to the suitability classification outlined in section 2.3.5, grasslands, and savannas were classified as the most suitable and croplands were classified as semi-suitable for grid extension.

The distance from existing substations, in contrast, contributes the least to the grid extension suitability, accounting for 5.7 % of the total suitability score. This is due to that only a minuscule fraction of settlements falls under the most suitable class, while the vast majority of settlements (97 %) are found in the least suitable class of distance from substations. Thus, the distance of settlements from existing substations plays a crucial role in increasing or decreasing the GEP multiplier and is perhaps the primary limiting factor for grid extension in Ethiopia.

According to the results presented in section 3.3, the combined impact of the five geospatial factors can result in a substantial rise in the investment cost by up to 29 %, depending on the location. This unequivocally illustrate the critical role of geospatial factors in grid extension planning; neglecting them can lead to nearly 30 % cost increases. A comparable study [12] showed that the GEP multiplier could escalate the grid extension investment cost by 30 %.

4.4. Novelty and significance of the study

This study introduces several novel aspects to the field. Firstly, it methodologically complements prior studies [8,14–17] which have implicitly considered geospatial factors in grid extension planning, especially in the context of grid vs. off-grid electrification. However, these studies, except [16], have used lower-resolution population datasets with a spatial resolution of 1 sq. km. This reliance on lower-resolution data can lead to an underestimation or misrepresentation of geospatial factors due to the complexities present at finer scales. Consequently, critical nuances in grid extension suitability might be overlooked, potentially resulting in suboptimal investment choices and inefficiencies in electrification planning. With finer spatial granularity (30 m for terrain, 500 m for land cover, and 30 m for population distribution), we are able to capture intricate local variations that impact grid-extension suitability. By leveraging such high-resolution datasets, our study minimizes the risk of overlooking critical geospatial factors, ultimately contributing to more effective and accurate electrification planning.

Secondly, this study stands out by incorporating not only HV lines but also the projected MV lines and substations into the modeling process, unlike the studies referenced as [14,15], and [17]. These studies primarily focused solely on HV lines when evaluating electrification planning in Nigeria, Ethiopia, and SSA, respectively. This comprehensive inclusion of multiple voltage levels and infrastructure elements enhances the accuracy of our electrification planning analysis. It allows us to capture a more nuanced representation of the grid network and its potential extensions, leading to more robust and precise planning outcomes. Thirdly, this study breaks new ground by explicitly analyzing the spatial impact of geospatial factors on grid-extension suitability, factoring in diverse energy demand targets or electrification tiers.

Although the research employed Ethiopia as a case study, the robustness of the methodology applied and the datasets and analyses employed make the findings generalizable and applicable to other developing countries that have sizable populations without access to electricity. In addition to its scientific contribution, this study offers valuable insights to decision-makers, government bodies, and utility companies, guiding investment decisions towards areas with the highest potential for benefiting from grid extension. This targeted approach enhances the overall resource efficiency of electrification efforts.

4.5. Limitations of the study

One limitation pertains to the reliance on open-access datasets, which might not consistently capture the most up-to-date information. This reliance on open-access datasets could potentially introduce limitations in terms of data accuracy, affecting the precision of the study's findings and conclusions. A particular noteworthy limitation within this context pertains to our use of the Gridfinder Tool for assessing population proximity to MV (Medium Voltage) lines. It is important to acknowledge that predictive tools like Gridfinder, relying on remote sensing techniques and publicly available datasets, such as nighttime data and MODIS land cover data, to establish MV network distribution lines, may not achieve the same level of reliability as data directly sourced from utility companies. This discrepancy is evident when considering the contrast between the relatively low nightlight imagery in Ethiopia and Gridfinder's estimates of densely distributed MV lines in the west of Addis Ababa, as illustrated in Fig. 4. This could potentially affect the accuracy of the grid extension suitability analysis conducted using the LCOE for each settlement. However, this does not affect the grid extension suitability analysis solely based on geospatial factors.

Besides, the study does not take into account off-grid solutions such as SHSs and mini-grids in accelerating the electrification in remote areas. Grid extension feasibility might not fully represent the complete electrification landscape. By excluding consideration of off-grid solutions, the study might not capture the potential contributions of these alternative approaches to rapidly expand electrification in remote and underserved areas. In addition, the study's scope is confined to the economic viability of grid extension and does not address other challenges (political, regulatory, socio-cultural factors, etc.) that often significantly influence large-scale electrification projects.

Moreover, a noteworthy limitation of the study pertains to its reliance on nighttime satellite imagery, collected at 1:30 a.m. local time when many rural households may not have their interior lighting switched on, thereby remaining invisible from an external perspective. This limitation can result in the underestimation of actual electricity use in rural areas, especially for activities such as indoor lighting, refrigeration, or electricity for irrigation, which might not be detectable by the satellite sensors. This can lead to the omission of actual electricity use in the estimation. A previous study [30] also highlighted this challenge, indicating that the NOAA VIIRS sensor's capability to detect electricity availability might be compromised in areas with minimal or absent outdoor lighting. Even in cases where electrified villages emit subdued light, the sensor's sensitivity may not be sufficient for accurate detection. This potential challenge influences the accuracy of electrification rates estimation. The study underscores the need for a nuanced understanding of the limitations inherent in the use of nighttime imagery for electrification assessment and acknowledges the potential for discrepancies in the results due to these constraints. Additionally, the accuracy of electrification estimation is also contingent upon the noise filter applied to the nighttime light imagery, which can inadvertently affect the identification of electrified areas.

5. Conclusions and future work

This paper investigates how and to what extent geospatial factors (distance from substation and road, terrain slope, elevation, and land cover) influence grid expansion suitability in Ethiopia. Using high-resolution nighttime light (NTL) imagery and population data with the OnSSET methodology, the research examines the proportion of the country's population in relation to the power grid infrastructures and the suitability of extending the grid to unelectrified areas.

The findings show that a significant portion of the population resides near the existing power grid lines, with 87.7 % living within a 10 km distance from a high-voltage lines and/or projected medium-voltage line, and 85 % residing within a 10 km distance from the projected medium-voltage lines. Nearly all unelectrified population (about 49 % of the total) lives within 25 km of power grid lines. This proximity suggests the potential cost-effectiveness of grid extension-based electrification.

The study finds that the combined impact of the five geospatial factors (GEP multiplier) increases grid extension costs by 2.3 %–29 %, depending on location. Terrain slope and land cover are the factors contributing the most to the suitability of extending the grid to unelectrified settlements, while distance from substations contributes the least to the grid extension suitability. This emphasizes the necessity of considering geospatial factors for accurate and efficient grid extension planning. Neglecting these factors could lead to cost underestimations and ineffective electrification strategies. Consequently, the Ethiopian Electric Utility (EEU) and Government of Ethiopia (GoE) should integrate these factors in planning, moving beyond the conventional focus on proximity to existing infrastructure. Incorporating spatial analyses of grid extension suitability enhances planning efficacy.

Furthermore, the study demonstrates that grid extension suitability, gauged by LCOE, highlights the need to consider both electricity demand and population when assessing grid extension feasibility. This underscores the significance of geospatial factors in grid planning and informs priority areas for extension or off-grid solutions. This approach optimizes costs and expedites electricity access in unelectrified regions. Future research could explore demand estimation, cost-effective alternatives, and investment strategies for universal electrification.

CRediT authorship contribution statement

Adugnaw Lake Temesgen: Conceptualization, Methodology, Software, Validation, Formal analysis, Conceptualization, Methodology, Software, Validation, Formal analysis. **Yibeltal T. Wassie:** Supervision, Writing – review & editing, Supervision, Writing – review & editing. **Erik O. Ahlgren:** Supervision, Writing – review & editing, Supervision, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Appendix. All the input GIS datasets, parameters and variables used in this paper are listed below

Table A1
Input GIS datasets.

Dataset	Data source	Data type	Spatial resolution	Temporal coverage	The year used in analysis
Population	HRSL (High Resolution Settlement Layer) [28]	Raster	30 m	2003–2020	2018
Administrative boundaries	GADM administrative areas V.4.0 [36]	Vector/polygon	–	2018–2022	2018
Night-time light	VIIRS DNB night-time lights [51]	Raster	~500 m (at equator)	2012–2021	2020
HV and substation	EnergyData.info [37]	Vector/polygon	–	2012–2017	2017
MV lines	gridfinder.org [40]	Vector/polygon	–	–2020	2020
Land cover	MODIS Land Cover Product (MCD12Q1) V6 [42]	Raster	–500 m	2001-	2020
Terrain elevation and slope	GDEM (NASA and Japan Space Systems) [43]	Raster	30 m	2009–2019	2019
Roads	Geofabric [44]	Vector/lines	–	–2018	2018

Table A2
Geospatial factors classification and weight [26,49].

Geospatial factor	Weight	Suitability index				
		5 (suitable)	4	3	2	1 (unsuitable)
Slope (degree)	30 %	0–10	10.1–20	20.1–30	30.1–40	>40
Land cover [a]	20 %	7,9,10,14,16	2,4	1,3,5,12,13,15	6,8	11,17
Distance to substation (km)	20 %	0	0.5	1	5	>10
Distance to road (km)	15 %	0	5	10	25	>50
Digital elevation (m)	15 %	<500	500–1000	1000–2000	2000–3000	>3000

[a] Further classification can be obtained at https://lpdaac.usgs.gov/products/mcd12q1v006/.

Table A3
Major socioeconomic parameters of the study area [18,46,47,52].

Parameters	Unit	Year	Study base year value
Population	Million persons	2020	114.96
Medium population growth	Percent	2030	2.2
High population growth	Percent	2030	2.5
Average household size, urban	People	2030	4.4
Average household size, rural	People	2030	5.2
Urban ratio start year	Percent	2020	21.7
Urban ratio end year	Percent	2030	26.9
Total electricity access	Percent	2020	51.1
Urban electricity access	Percent	2020	93.2
Rural electricity access	Percent	2020	39.4

Table A4
Grid generating and T&D cost. Sources [12,50,53–56].

Parameter	Value	Unit
HV line (69–132 kV)	53000	$USDkm^{-1}$
MV line (11–33 kV)	7000	$USDkm^{-1}$
LV line (0.2–0.4 kV)	4250	$USDkm^{-1}$
HV to MV substation (1000 kVA)	25000	USD/unit
MV to LV substation (400 kVA)	10000	USD/unit
Service transformer (50 kVA)	4250	USD/unit
Generating cost	0.09	$USDkWh^{-1}$
Additional connection cost per household connected to grid	150	USD/HH
T&D losses	7–29	% of capital cost/year
O&M costs of distribution	2	% of capital cost/year

Table A5
Other model parameter. Sources [12,18,52].

Parameter	Value	Unit
Annual new grid connection limit	534,000	Households/year
Annual grid generation limit	389	MW/year
Discount rate	8	Percent
Lifetime	30	years

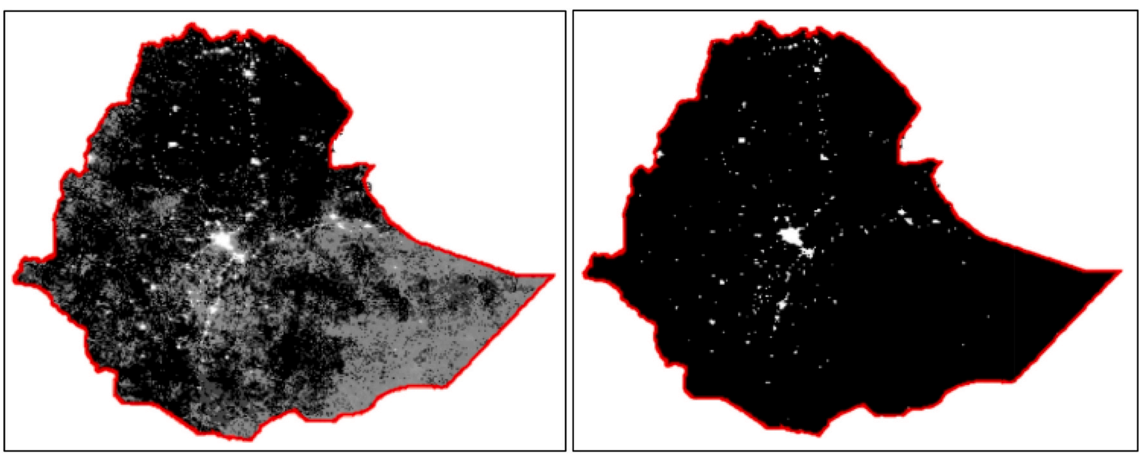

Fig. A1. Filtered nighttime light of 2020 using 0.25 μWcm-2sr-1 (left) and 0.35 μWcm-2sr-1 (right) lower bound noise floor.

References

[1] United Nations Development Programme, SDG 7: Affordable and clean energy. https://www.undp.org/sustainable-development-goals/affordable-and-clean-energy. (Accessed 12 October 2022).

[2] F. Kemausuor, E. Adkins, I. Adu-Poku, A. Brew-Hammond, V. Modi, Electrification planning using Network Planner tool : the case of Ghana, Energy Sustain Dev 19 (2014) 92–101, https://doi.org/10.1016/j.esd.2013.12.009.

[3] IEA, IRENA, UNSD, World Bank, WHO, Tracking SDG 7: The Energy Progress Report. World Bank, Washington DC. © World Bank. License: Creative Commons Attribution—NonCommercial 3.0 IGO (CC BY-NC 3.0 IGO). Executive Summary and Chapter 1, 2021, pp. 1–29. https://www.irena.org/publications/2021/Jun/Tracking-SDG-7-2021. March 23.

[4] M.V. Rocco, E. Fumagalli, C. Vigone, A. Miserocchi, E. Colombo, Enhancing energy models with geo-spatial data for the analysis of future electrification pathways: the case of Tanzania, Energy Strategy Rev. 34 (2021) 100614, https://doi.org/10.1016/j.esr.2020.100614.

[5] P. Ciller, D. Ellman, C. Vergara, A. Gonzalez-Garcia, S.J. Lee, C. Drouin, et al., Optimal electrification planning incorporating on- and off-grid technologies- and reference electrification model (REM), Proc. IEEE 107 (2019) 1872–1905, https://doi.org/10.1109/JPROC.2019.2922543.

[6] African Development Bank Group, Annual Report 2014, ADB-ADF/BG/AR/2014, Regional and Non-Regional Member Countries, 2014, pp. 13–184. Chapter 1-7, https://www.afdb.org/fileadmin/uploads/afdb/Documents/Generic-Documents/Annual_Report_2014_-Full.pdf. (Accessed 18 November 2022).

[7] P. Bertheau, A.S. Oyewo, C. Cader, C. Breyer, P. Blechinger, Visualizing national electrification scenarios for sub-saharan African countries, Energies 10 (2017) 1–20, https://doi.org/10.3390/en10111899.

[8] A. Korkovelos, B. Khavari, A. Sahlberg, M. Howells, C. Arderne, The role of open access data in geospatial electrification planning and the achievement of SDG7. An onsset-based case study for Malawi, Energies 12 (2019), https://doi.org/10.3390/en12071395.

[9] The division of energy systems analysis at KTH in collaboration with SNV, Electrification Pathways for Benin A Spatial Electrification Analysis Based on the Open Source Spatial Electrification Tool, OnSSET), 2018, pp. 2–13. http://www.onsset.org/uploads/1/8/5/0/18504136/electrification_pathways_for_benin.pdf. (Accessed 20 December 2022).

[10] C. Bastholm, Micro-grids supplied by renewable energy : Improving technical and social feasibility [Licentiate dissertation]. [Uppsala]: Acta Universitatis Upsaliensis, Available from: https://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-375499, 2019. (Accessed 2 October 2022).

[11] S. Watchueng, R. Jacob, A. Frandji, Planning Tools and Methodologies for Rural Electrification 56 (2010) 31–48. https://www.club-er.org/files/fichiers/Planning_Tools_and_Methodologies_for_Rur041218015603.pdf. (Accessed 15 May 2022).

[12] D. Mentis, M. Howells, H. Rogner, A. Korkovelos, C. Arderne, E. Zepeda, et al., Lighting the World: the first application of an open source, spatial electrification tool (OnSSET) on Sub-Saharan Africa, Environ. Res. Lett. 12 (2017), https://doi.org/10.1088/1748-9326/aa7b29.

[13] B. Khavari, A. Sahlberg, W. Usher, A. Korkovelos, F. Fuso Nerini, The effects of population aggregation in geospatial electrification planning, Energy Strategy Rev. 38 (2021) 100752, https://doi.org/10.1016/j.esr.2021.100752.

[14] D. Mentis, M. Andersson, M. Howells, H. Rogner, S. Siyal, O. Broad, et al., The benefits of geospatial planning in energy access - a case study on Ethiopia, Appl. Geogr. 72 (2016) 1–13, https://doi.org/10.1016/j.apgeog.2016.04.009.

[15] D. Mentis, M. Welsch, F. Fuso Nerini, O. Broad, M. Howells, M. Bazilian, et al., A GIS-based approach for electrification planning-A case study on Nigeria, Energy Sustain Dev 29 (2015) 142–150, https://doi.org/10.1016/j.esd.2015.09.007.

[16] S.R. Isihak, Achieving universal electricity access in line with SDG7 using GIS-based Model: an application of OnSSET for rural electrification planning in Nigeria, Energy Strategy Rev. 45 (2023) 101021, https://doi.org/10.1016/j.esr.2022.101021.

[17] S. Szabó, K. Bódis, T. Huld, M. Moner-Girona, Energy solutions in rural Africa: mapping electrification costs of distributed solar and diesel generation versus grid extension, Environ. Res. Lett. 6 (2011), https://doi.org/10.1088/1748-9326/6/3/034002.

[18] World Bank, Access to electricity (% of population) - Ethiopia. https://data.worldbank.org/indicator/EG.ELC.ACCS.ZS?locations=ET. (Accessed 23 March 2022).

[19] Power Sector Reform - Ethiopia Proposed Roadmap for Implementation, Ministry of Water Irrigation and Energy Power Sector Reform Roadmap - Ethiopia, 2020.

[20] FDRE, Federal Democratic Republic of Ethiopia, MoWIE, Ministry of Water, Irrigation and Electricity, "Light to all, National Electrification Program 2.0 Light to All Integrated Planning for Universal Access. https://www.powermag.com/wpcontent/uploads/2020/08/ethiopia-national-electrification-program.pdf, 2019. (Accessed 25 January 2022). Chapter 1-4.

[21] I. Pappis, A. Sahlberg, T. Walle, O. Broad, E. Eludoyin, M. Howells, et al., Influence of electrification pathways in the electricity sector of Ethiopia—policy implications linking spatial electrification analysis and medium to long-term energy planning, Energies 14 (2021), https://doi.org/10.3390/en14041209.

[22] M. Moner-Girona, K. Bódis, J. Morrissey, I. Kougias, M. Hankins, T. Huld, et al., Decentralized rural electrification in Kenya: speeding up universal energy access, Energy Sustain Dev 52 (2019) 128–146, https://doi.org/10.1016/j.esd.2019.07.009.

[23] B. Johnson, Toward a new classification of nonexperimental quantitative research, Educational Researcher, vol. 30, no. 2, , pages 3–13. JSTOR (2001). http://www.jstor.org/stable/3594335.

[24] J.G. Peña Balderrama, S. Balderrama Subieta, F. Lombardi, N. Stevanato, A. Sahlberg, M. Howells, et al., Incorporating high-resolution demand and techno-economic optimization to evaluate micro-grids into the Open Source Spatial Electrification Tool (OnSSET), Energy Sustain Dev 56 (2020) 98–118, https://doi.org/10.1016/j.esd.2020.02.009.

[25] M. Howells, H. Rogner, N. Strachan, C. Heaps, H. Huntington, S. Kypreos, et al., OSeMOSYS: the open source energy modeling system. An introduction to its ethos, structure and development, Energy Pol. 39 (2011) 5850–5870, https://doi.org/10.1016/j.enpol.2011.06.033.

[26] OnSSET. OpeN, Source Spatial Electrification Toolkit (2022). http://www.onsset.org/. accessed March 25.

[27] B. Khavari, A. Korkovelos, A. Sahlberg, M. Howells, F. Fuso Nerini, Population cluster data to assess the urban-rural split and electrification in Sub-Saharan Africa, Sci. Data 8 (2021) 1–11, https://doi.org/10.1038/s41597-021-00897-9.

[28] Ciesin, High Resolution Settlement Layer N, 2022. https://ciesin.columbia.edu/data/hrsl/. accessed March 20.

[29] G. Falchetta, S. Pachauri, S. Parkinson, E. Byers, A high-resolution gridded dataset to assess electrification in sub-Saharan Africa, Sci. Data 6 (2019), https://doi.org/10.1038/s41597-019-0122-6.

[30] Min B, Gaba KM. Tracking Electrification in Vietnam Using Nighttime Lights 2014: 9511–29. https://doi.org/10.3390/rs6109511..

[31] T.R. Kiran Chand, K.V.S. Badarinath, C.D. Elvidge, B.T. Tuttle, Spatial characterization of electrical power consumption patterns over India using temporal DMSP-OLS night-time satellite data, Int. J. Rem. Sens. 30 (2009) 647–661, https://doi.org/10.1080/01431160802345685.

[32] N. Levin, C.C.M. Kyba, Q. Zhang, A. Sánchez de Miguel, M.O. Román, X. Li, et al., Remote sensing of night lights: a review and an outlook for the future, Remote Sens. Environ. 237 (2020) 111443, https://doi.org/10.1016/J.RSE.2019.111443.

[33] B. Min, K.M. Gaba, O.F. Sarr, A. Agalassou, Detection of rural electrification in africa using DMSP-OLS night lights imagery, Int. J. Rem. Sens. 34 (2013) 8118–8141, https://doi.org/10.1080/01431161.2013.833358.

[34] W.M. Dlamini, L.C. Dlamini, Spatial assessment and monitoring of household electricity access and use using nighttime lights and ancillary spatial data: a case of Eswatini, African Geogr Rev 41 (2022) 299–317, https://doi.org/10.1080/19376812.2021.1897023.

[35] B Khavari, A Korkovelos, KTH-dESA/PopCluster. https://github.com/OnSSET/PopCluster. (Accessed 20 January 2022).

[36] GADM, GADM Database of Global Administrative Areas V.4.0. https://gadm.org/download_country.html. (Accessed 28 March 2022).

[37] Energydata, Energy data, ENERGYDATAINFO. https://energydata.info/, 2022.

[38] S.-S.A.-E.T. Network, Sub-Saharan africa - electricity transmission network, ENERGYDATAINFO. https://energydata.info/dataset/sub-saharan-africa-electricity-transmission-network-2007, 2022.

[39] C. Arderne, C. Zorn, C. Nicolas, E.E. Koks, Predictive mapping of the global power system using open data, Sci. Data 7 (2020) 19, https://doi.org/10.1038/s41597-019-0347-4.

[40] C. Arderne, Predictive mapping of the global power system using open data, Predict Mapp Glob Power Syst Using Open Data [Data Set] Nat Sci Data 7 (2022). Number Artic 19) Zenodo 2020, https://gridfinder.org/. accessed April 21.

[41] Earth data, NASA Earth Observation, Nasa, 2020, 2012, http://earthobservatory.nasa.gov/. (Accessed 30 April 2022).

[42] D. Sulla-menashe, M.A. Friedl, User Guide to Collection 6 MODIS Land Cover (MCD12Q1 and MCD12C1) Product (2018), https://doi.org/10.5067/MODIS/MCD12Q1.006. NASA EOSDIS L Process DAAC.

[43] T Tachikawa, M Kaku, M Material, T Corp, A Iwasaki, ASTER GDEM Version 3 Validation Report 2015, 2022, pp. 1–12. https://asterweb.jpl.nasa.gov/gdem.asp. (Accessed 2 April 2022).

[44] GEOFABRIK, OpenStreetMap. https://www.geofabrik.de/, 2018. (Accessed 19 April 2022).

[45] B Khavari, F Fuso Nerini, A Korkovelos, A Sahlberg, M Howells, ASTER GDEM Version 3 Validation Report 2015. https://github.com/OnSSET/Clustering. (Accessed 14 February 2022).

[46] IEA, IRENA, UNSD, WB, WHO, Tracking SDG 7: The Energy Progress Report 2019, Washington DC. https://www.irena.org/media/Files/IRENA/Agency/Publication/2019/May/2019-Tracking-SDG7-Report.pdf?rev=30fdda95ad074b6eb27c8aa4347605e8, 2019. (Accessed 29 March 2022).

[47] IEA, Energy Statistics Data Browser, IEA, Paris. https://www.iea.org/data-and-statistics/data-tools/energy-statistics-data-browser, 2023. (Accessed 17 March 2022).

[48] S.G. Banerjee, M. Bhatia, E. Portale, J. Schers, D. Dorner, G.E. Azuela, et al., Global tracking Framework, Sustainable Energy All, World Bank 3 (289) (2013). https://documents1.worldbank.org/curated/en/603241469672143906/pdf/778890GTF0full0report.pdf. (Accessed 4 December 2022).

[49] B Stewart, A Korkovelos, A Sahlberg, B Khavari, The Global Electrification Platform. https://github.com/global-electrification-platform. (Accessed 14 April 2022).

[50] F.F. Nerini, O. Broad, D. Mentis, M. Welsch, M. Bazilian, M. Howells, A cost comparison of technology approaches for improving access to electricity services, Energy 95 (2016) 255–265, https://doi.org/10.1016/j.energy.2015.11.068.

[51] Earth Observation Group (EOG), Annual VNL V2. https://eogdata.mines.edu/products/vnl/. (Accessed 3 April 2022).

[52] United Nations | DESA Population Division, World Population Prospects - Population Division - United Nations, 2022. https://population.un.org/wpp/. May 13.

[53] META ESMAP. Model for Electricity Technology Assessment - Integrating Externalities into Electricity Supply Decisions, 2014, pp. 1–6. https://www.esmap.org/sites/esmap.org/files/ESMAP_META_March2014.pdf.

[54] R Karhammer, A Sanghvi, E Fernstrom, M Aissa, J Arthur, J Tullock, et al., Sub-Saharan Africa: Introducing Low Cost Methods in Electricity Distribution Networks, Energy Sector Management Assistance Program (ESMAP), Technical Paper 104 (427) (2006). https://documents1.worldbank.org/curated/en/333511468204868864/pdf/415990AFR0Elec1n0ESMAP010401PUBLIC1.pdf.

[55] World Bank, Reducing the Cost of Grid Extension for Rural Electrification ESM227, 2000.

[56] BJ Van Ruijven, J Schers, DP van Vuuren, Model-based scenarios for rural electrification in developing countries, Energy 38 (2012) 386–397. https://doi.org/10.1016/j.energy.2011.11.037.