POLITECNICO
MILANO 1863

SCUOLA DI INGEGNERIA INDUSTRIALE
E DELL'INFORMAZIONE

Exploring Synergies and Complementarities between Geospatial Electrification Modeling and an Energy System Optimization Model for Off-Grid Energy Planning in Kenya

Tesi di Laurea Magistrale in
Energy Engineering - Ingegneria Energetica

Authors: **Edoardo Curia, Valentina Tomasso**

Student ID: 252638, 246848
Advisor: Prof. Nicolo' Stevanato
Co-advisors: Prof. Alex Muumbo
Academic Year: 2025-26

Abstract

Universal access to electricity remains a critical challenge in rural Sub-Saharan Africa, where mini-grids present a viable alternative to national grid expansion. Effective off-grid energy planning typically relies on either Geospatial Electrification Models (GEMs), such as OnSSET, which offer broad spatial analyses, or Energy System Optimization Models (ESOMs), like MicroGridsPy, which provide high-resolution techno-economic optimization. Despite their inherent complementarities, integrating these models poses significant challenges, particularly regarding the consistent characterization of high-resolution electrical demand.

This thesis proposes a synergistic framework to soft-link OnSSET and MicroGridsPy, using Kenya as a comprehensive case study. To bridge the gap in temporal granularity, high-resolution stochastic demand archetypes that account for wealth, latitude and specific cooling needs are integrated into OnSSET to enhance its demand definition and hybrid system modeling. Subsequently, 39 optimal mini-grid sites initially identified by the geospatial tool are individually optimized using MicroGridsPy. A calibrated Gaussian distribution of households is applied within the ESOM to ensure load coherence between the two software environments without sacrificing its settlement-level resolution.

The comparative analysis reveals that OnSSET systematically overestimates the Levelized Cost of Energy (LCOE) for mini-grids compared to the dedicated optimization tool. MicroGridsPy generally prioritizes capital-intensive renewable technologies, such as photovoltaics and battery storage, to achieve a lower long-term LCOE. However, specific operational circumstances, such as intermittent seasonal cooling demands, can shift the optimal configuration back toward diesel generation to prevent battery underutilization. Crucially, the implementation of demand archetypes into OnSSET successfully reduces the divergence in system sizing and cost estimations between the two models. Ultimately, the sequential application of these tools leverages the spatial scale of GEMs alongside the technological rigor of ESOMs, offering a more robust and coherent methodology for rural electrification planning.

Keywords: mini-grids; gems; esoms; onsset; microgridspy; kenya

Abstract in lingua italiana

L'accesso universale all'elettricità è una sfida cruciale nell'Africa Subsahariana rurale, dove i sistemi decentralizzati offrono una valida alternativa all'espansione della rete nazionale. La pianificazione energetica si affida tipicamente a modelli geospaziali (GEMs) come OnSSET, per l'ampia analisi spaziale, o a modelli di ottimizzazione (ESOMs) come MicroGridsPy, per l'ottimizzazione tecno-economica ad alta risoluzione. Nonostante le complementarità, l'uso integrato di questi strumenti pone sfide significative nella definizione coerente della domanda elettrica.

Questo lavoro propone un framework sinergico applicato al caso studio del Kenya. Per colmare il divario nella risoluzione temporale, in OnSSET sono stati integrati archetipi di carico ad alta risoluzione basati su fasce di reddito, latitudine ed esigenze di raffrescamento. Successivamente, 39 siti ottimali per mini-grid individuati dal modello geospaziale sono stati ottimizzati singolarmente con MicroGridsPy. Nell'ESOM è stata applicata una distribuzione Gaussiana delle famiglie, calibrata per garantire la coerenza del carico tra i due strumenti senza perdere la risoluzione a livello di singolo insediamento.

L'analisi comparativa rivela che OnSSET sovrastima sistematicamente l'LCOE delle mini-grid rispetto allo strumento dedicato. MicroGridsPy predilige generalmente tecnologie rinnovabili ad alto costo iniziale e sistemi di accumulo per ottenere un LCOE inferiore nel lungo termine. Tuttavia, carichi stagionali intermittenti possono spostare la configurazione ottimale verso la generazione diesel per prevenire il sottoutilizzo delle batterie. L'introduzione degli archetipi in OnSSET riduce con successo la divergenza di dimensionamento e di costo tra i modelli. L'approccio sequenziale sfrutta così la scala spaziale dei GEMs e il rigore tecnologico degli ESOMs, offrendo una metodologia robusta per la pianificazione elettrica rurale.

Parole chiave: mini-grids; gems; esoms; onsset; microgridspy; kenya

Contents

1 | Introduction

1.1 Universal electricity access

Universal access to electricity has been a challenge for the past decades, and continues to be so nowadays despite the great progress made throughout the years: according to the World Bank [1], as of 2023 8.4% of the world population lacked electricity access; Sub-Saharan Africa (SSA) bears a disproportionate share of this burden, as only 53.3% of the population can rely on some form of connection. Looking at the future, the lack of access is a critical bottleneck, as Sub-Saharan countries are expected to experience a great economic growth, and having an inadequate and unsustainable electrical infrastructure severely hinders this development. Furthermore, relying on unsustainable electrification paths would result heavily damaging to the environment. For these reasons, achieving universal access to reliable, clean and affordable electricity is part of the Sustainable Development Goals, as SDG 7, as it promotes both economic growth and its sustainability, from both an environmental and an economical standpoint.

The contrast with international levels of electricity access is even more stark when looking at the electrification status of rural areas of SSA, reaching a low 33.1%, indicating how rural zones are of major interest in this field. The lack of connection is mostly due to the scattered locations of rural settlements, each quite distant from the national grid, even if in the past years grid expansion has progressed incredibly fast in some countries: an example is Kenya, where the population with electricity access has risen from 15.2% in 2000 to 76.2% in 2023 [1]. Large distances and challenging morphology mean that in rural areas electricity is still a luxury only few people can afford. In the past few decades, though, an alternative to the expansion of the national grid has emerged: mini-grids are decentralized systems that exploit local resources in order to supply electricity to even the most isolated communities; originally, mini-grids were producing electricity by solely relying on traditional fuel generators, but renewable technologies such as solar photovoltaics, small hydro and wind have emerged, effectively reducing the need to stock large amounts of fuels and enhancing the decentralizing aspect. A back-up system is needed though, either relying on fuel based generators or battery storage systems, as renewable power is non-dispatchable and is not always available: the solar resource, for example, is not present during night hours, when most of the households use electricity; hence batteries can help shift the power produced by the RES to the hours where it is most needed.

1.2 Energy modeling for electricity access in rural areas

Energy modeling is a process that consists of representing and sizing energy systems in order to meet given requirements, such as electrical or thermal demand, while optimizing either costs or emissions of the system and exploiting given resources. Different approaches of energy modeling exist, mainly grouped in either bottom-up or top-down: the first focuses on technological detail, while the latter prioritizes an economy-wide approach.
In the context previously described, the energy modeling process is used for various purposes, ranging from the sizing of photovoltaic projects in rural areas to national strategy planning of energy mix optimization.
Focusing on bottom-up models, for off-grid sizing two main groups of software are available:

- **Energy System Optimization Modeling (ESOMs):** aim to define the optimal configuration of a hybrid energy system that supplies energy to a localized set of users, normally a single community, as well as the optimal operation and dispatch strategy to minimize costs
- **Geospatial Electrification Modeling (GEMs):** based on GIS (Geospatial Information Systems) data, focus instead on identifying the least cost solution, among a set of predefined options, to meet the energy demand of clusters of non-electrified populations across vast areas, unit by unit.

Both types of tools have to satisfy a given demand load by choosing the least cost solution. Software like Homer [2] and MicroGridsPy [3] fall into the first category, requiring as input demand the hourly load of a limited number of users; in some models the distribution system of the mini-grid is also accounted for, as it is in the case for the latest MicroGridsPy version [4] or for OnSSET for MiniGrids (OMG) [5], in order to offer a more comprehensive depiction of the costs of the entire system, but they lack the geospatial information electrification platforms offer. Still, even with similar approaches and inputs, the different model behind the software makes them appropriate for different purposes: for example, OMG can be more appropriate for regional planning and identifying sites for

plausible mini-grid installations, while MicroGridsPy excels at detailed feasibility studies and cost analysis [6].
Geospatial tools like OnSSET (Open-source Spatial Electrification Tool) [7] and GISEle (GIS for Electrification) [8] instead are preferable when analyzing the possible least cost electrification strategies over a large area. OnSSET is the software behind the Global Electrification Platform (GEP) [9], which offers a quick first analysis of different scenarios for possible future national paths. GISEle differs as it can have a more detailed demand input, while still being a GIS based tool, making it more oriented towards regional electrification planning.

1.3 Models Integration in Literature

Although ESOMs and GEMs are designed for different analytical scopes, their contrasting strengths can be effectively utilized in a collaborative framework. A joint application of the models may allow researchers to take advantage of the broad spatial outlook offered by geospatial models and the rigorous technological detail of optimization tools. By soft-linking geospatial and energy system optimization models, trade-offs between spatial scale and temporal granularity can be overcome, reaching a significantly higher level of techno-economic accuracy and analytical definition.
In literature, various studies concerning the soft-linking between GEMs and ESOMs are reported, most of which propose the joint use of software working on a large, national scale. For example, the research developed by *Moksnes et al.*(2017) [10] proposed the combination of OnSSET and OSeMOSYS [11], a national ESOM, to investigate possible pathways to reach electrification targets with an optimal energy mix of the centralized grid. A similar study, soft-linking OnSSET and OSeMOSYS, is detailed in *Pappis et al.*(2021) [12]. Instead, only a few examples can be found regarding the integration of high-temporal resolution demand in geospatial electrification models: specifically, the research conducted by *Silinto et al.*(2025) [13] focuses on the integration of the demand generation software RAMP [14] in the GISEle geospatial model, while the study by *Balderrama et al.* (2020) [15] investigates the population of a micro-grid's look-up table with numerous MicroGridsPy optimizations to be implemented in OnSSET as an alternative technology in the final choice. The latter is one of the few publicly available studies dedicated to propose a methodology to directly soft-link a geospatial model and an energy system optimization model focused on small scale systems.

The combined use of software like OnSSET and MicroGridsPy, with such a different approach to energy modeling despite both being bottom-up defined, can be useful for a more accurate cost analysis and scaling of off-grid energy systems for remote locations. The geospatial analysis offered by OnSSET can be complemented by MicroGridsPy's great temporal resolution, detailed demand definition and system specific energy balance optimization.

1.3.1 Integrating detailed load demand in a GIS model

As previously mentioned, the study presented in *Silinto et al.*(2025) [13] introduced in the geospatial electrification model GISEle the RAMP model for demand construction: the stochasticity of the input creation allows for a more representative description of the users, starting from a basket of appliances and generating a load profile that varies during the temporal horizon of the project. The GISEle model is generally used to analyze a limited area, so such representation of load demand is not excessively computationally intensive. A similar logic cannot be applied to the OnSSET tool, which instead is used for national case studies and would hence require a huge computational cost.

1.3.2 Incorporation of system optimization in OnSSET

The research conducted in *Balderrama et al.*(2020) [15] proposes an alternative way to introduce high-temporal resolution demand in a geospatial model; moreover, the techno-economic optimization for energy system models is also introduced. RAMP is used in order to generate a pool of 330 load scenarios; then, energy system optimization is conducted via the MicroGridsPy tool, changing relevant parameters such as diesel cost, technology investment costs and allowed lost load, generating a comprehensive solution pool dataset. Finally, the dataset is implemented in OnSSET as representative of micro-grid solutions: each settlement is assigned the LCOE of the closest solution (based on given parameters), in order to then be able to compare it with other technologies.
Such a complex framework links MicroGridsPy and OnSSET, but it is hard to replicate and does not improve the temporal detail and description of the load demand of the OnSSET tool itself. Moreover, the computational cost of populating the demand and optimization datasets is exceptionally high.

1.3.3 The issue of demand characterization: GEMs and ESOMs pose different challenges

One of the major issues with energy modeling is the accurate depiction of the demand load, which obviously influences the energy balance and hence the sizing and costs of the techno-economic solution. In developing contexts such as rural SSA it is especially difficult

to reliably foresee the load once electrification takes place, both due to the remoteness of the areas of interest and the unpredictability of the demand evolution during the project. For these reasons, *Stevanato et al.*(2023) [16] proposed a set of archetypes of electrical demand profiles for rural Sub-Saharan communities load estimation, generated with a stochastic approach with the RAMP tool. In the same study, these archetypes were integrated in MicroGridsPy in order to offer a fast load profile definition, as it will be explained thoroughly in Paragraph 2.1.1.
In the study, it is suggested that the integration of archetypes in geospatial electrification platforms may be of use for a more accurate demand depiction for system modeling.

1.4 Thesis purpose and structure

As previously mentioned, a geospatial electrification model like OnSSET and a minigrid-specific energy system optimization tool such as MicroGridsPy show great potential synergy because of the apparent complementarity of their approaches; unfortunately the research developed to soft-link the MicroGridsPy and OnSSET tools does not offer a direct improvement of demand definition in the geospatial software, instead opting to create a look-up table of micro-grid solutions and adding it to the pool of possible technologies to compare on OnSSET.
The present study aims at filling the gap between the temporal resolution offered by the two software by first integrating archetypes in OnSSET, as it is still an unexplored approach, and then comparing the results of a set of settlements solutions with MicroGridsPy; the final objective is to propose a framework to sequentially use the two tools to fully exploit their strengths, while guaranteeing consistency between the analyses.
The thesis is structured as follows: the mathematical formulation of the models of interest is explored; then the methodology followed to propose a consistent analysis with the two tools is explained in depth; finally, after an introduction of the case study, the corresponding results are analyzed and then summarized in final conclusions.

2 | Mathematical formulation of the models

To better understand the synergies and complementarities between MicroGridsPy and OnSSET, a more in depth look at their mathematical formulation is due. In the following paragraphs, both software's core logic are described, highlighting the points of interest for our analysis.

2.1 MicroGridsPy formulation

MicroGridsPy is a bottom-up, open-source optimization tool developed for the planning of energy systems in remote and under served areas. The model is implemented in a python-based optimization framework, Linopy. It is a useful instrument in rural contexts, as it considers the challenges of limited data availability, high uncertainty and strong economic constraints. In the latest version of the software, the workflow includes the design of the distribution grid topology and its cost assessment, but the version used during this analysis [17] lacks this feature.

MicroGridsPy provides an optimized dispatch strategy, assuming perfect foresight, utilizing an hourly time resolution for both demand and resource availability definition: indeed, it requires as inputs the hourly time series of both the available renewable resources (e.g. solar potential, wind capacity factor) and the electric load the system will be expected to meet. Regarding the latter, a more in depth discussion can be found in the following subsection.

As previously mentioned, MicroGridsPy is an optimization tool and it has as objective function the minimization of the Net Present Costs (NPC) (2.1) while satisfying the electric demand.

$$NPC = \sum_{y=0}^{N} \frac{C_{inv,y} + C_{OM,y} + C_{fuel,y} - SV_y}{(1+d)^{y-y_0}} \tag{2.1}$$

- NPC: Net Present Cost of the project
- $C_{inv,y}$: total investment costs at year -y
- $C_{OM,y}$: total operation and maintenance costs at year -y
- $C_{fuel,y}$: total cost of fuel at year -y
- SV_y: total salvage value at year -y
- d: discount rate

MicroGridsPy mainly considers the integration of renewable resources technologies, energy storage systems and back-up generation systems, in order to be able to meet the load demand even in temporary absence of resource availability.

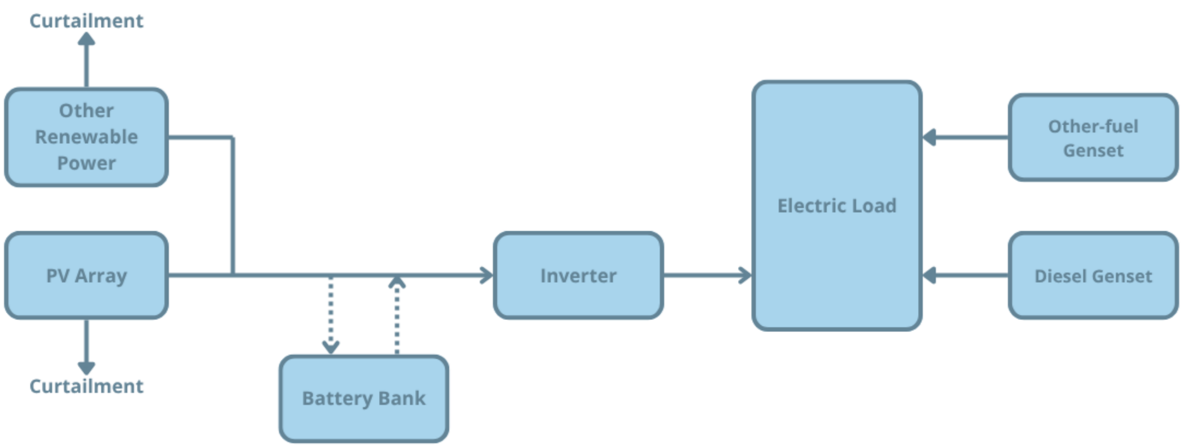


Figure 2.1: Example of reference energy system as implemented by MicroGridsPy

All technologies are characterized by a basic cost structure and technical parameters, which also account for possible inverters and rectifiers according to the type of load the system is required to meet.
In order to ensure that the supply meets the demand of the system, MicroGridsPy controls the system in each time-step via an energy balance constraint, as defined in Equation (2.2).

$$\sum_r E^{\text{res}}_{t,r} + \sum_g E^{\text{gen}}_{t,g} + E^{\text{imp}}_t - E^{\text{exp}}_t + P^{\text{dis}}_t - P^{\text{ch}}_t + \ell_t = D_t \quad \forall t \tag{2.2}$$

- $\sum_r E^{\text{res}}_{t,r}$: total energy production by RES
- $\sum_g E^{\text{gen}}_{t,g}$: total energy production by generators
- E^{imp}_t and E^{exp}_t: energy imported from/exported to the grid (if grid connection is enabled)

- P_t^{dis} and P_t^{ch}: battery discharge/charge
- ℓ_t: lost load (unserved energy)
- D_t: electrical demand

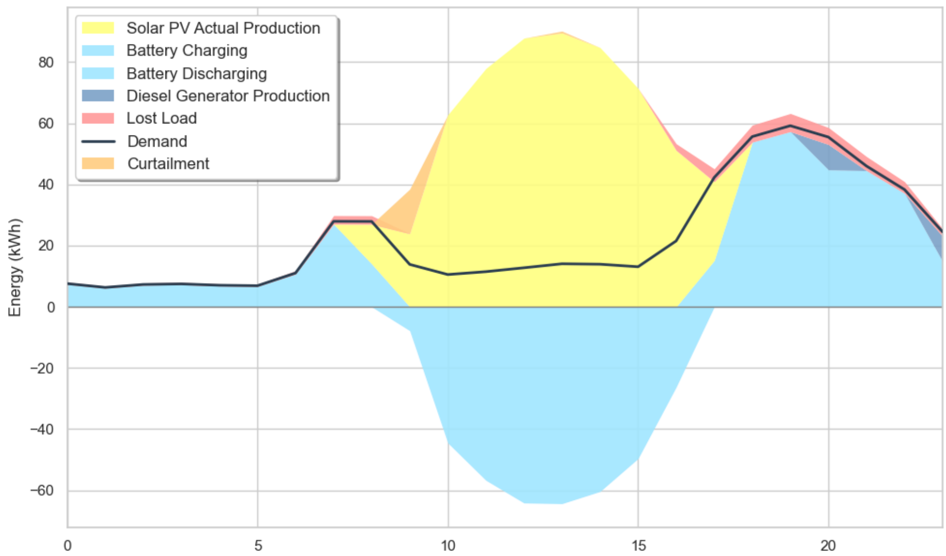


Figure 2.2: Example of the optimized dispatch strategy for a given project

2.1.1 Demand Definition

The load demand is defined for every year of the project on an hourly basis: the software requires as input a timeseries of 8760 values of hourly demand per year, in *Wh*. While a personalized demand file can be uploaded by the user, archetypes of load profiles of Sub-Saharan Africa users are also available, as introduced by the study by *Stevanato et al.*(2023) [16]. These archetypes are useful in case no information about user specific electric appliances is available, since they were built via the RAMP stochastic model and have as a target the depiction of rural areas of SSA countries.
The construction of the load curves is based on a previous study on the M-LED platform by *Falchetta et al.*(2021) [18], which proposed a methodology to delineate potential electricity demand of un-electrified settlements in absence of reliable data; within their work, the authors contextualized and validated their findings through an extensive review and comparison with existing energy demand literature (Figure 2.3). Among the considered studies, it is important to highlight the tier based classification proposed by ESMAP (Energy Sector Management Assistance Program) [19], upon which the OnSSET tool demand definition is based: while it is in line with studies concerning load variation in urban contexts, rural settlement characterization results only partially aligned with similar studies.

Regarding the M-LED study, the survey-based demand definition results consistent with existing literature regarding both urban and rural non-electrified contexts. As the latter are the focus of the present study, part of it is devoted to the introduction of a coherent demand definition in the GEM tool (see Section 3.2). More considerations about the differences between the MTF and M-LED estimates of potential electricity demand are explored in Section 5.1.

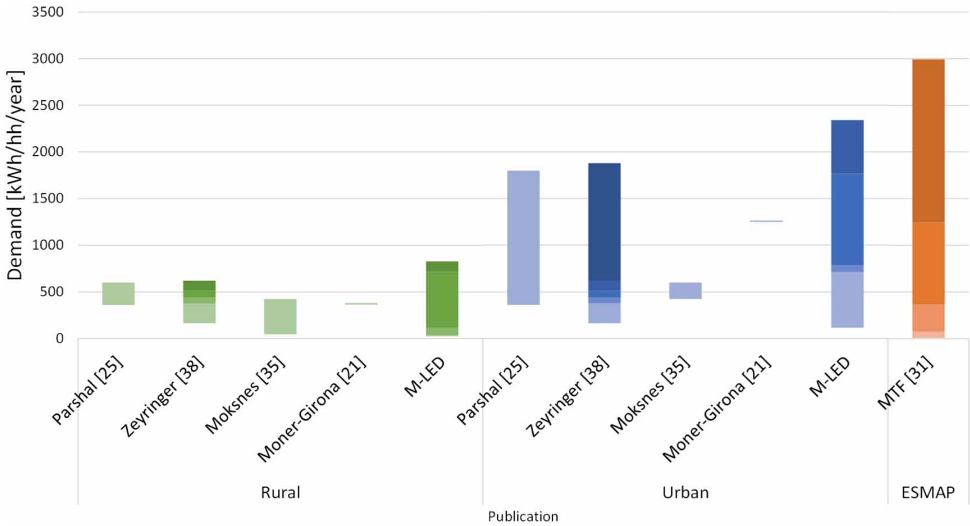


Figure 2.3: Comparison of the estimated potential electricity demand of non-electrified households [18]

The archetypes introduced in *Stevanato et al.*(2023) [16] allow to represent combinations of a variety of users, including load curves for different types of residential demand, health-care centers and educational facilities. Delving deeper regarding residential consumption, 100 possible archetypes are available: other than the wealth of the specific household, the geographical position of the settlement and the climate conditions are also accounted for.

- **Wealth Tiers:** five wealth categories of household are proposed, each considered to own a different basket of appliances, created based on literature review of electricity use in rural SSA
- **Latitude Zones:** latitude is considered as sunrise and sunset hours influence appliances utilization time within a household
- **Climate Zones:** as suggested by *Falchetta, Mistry*(2021) [20], in the nearby future cooling demand in rural SSA is likely to grow considerably due to growing heat stress caused by climate change and it is fundamental to include it in demand estimation practices for energy modeling. The reference study for archetypes hence proposes four different cooling periods, based on the computation of the *Cooling Degree Days*

(CDD) reported by *Falchetta, Mistry*(2021) [20]: all year (AY), no cooling (NC), October to March (OM) and April to September (AS). For further information in this regard, please refer to Paragraph 3.3.2.

In MicroGridsPy, it is possible to choose the number of schools, and households and healthcare centers represented by each archetype; then, the software computes the aggregated demand by summing the hourly-defined loads of each user, using it as the final load curve to satisfy. It is worth noting that, as MicroGridsPy is a planning tool, the yearly growth of the served community is also required in order to compute the hourly load demand for all the years of the project.

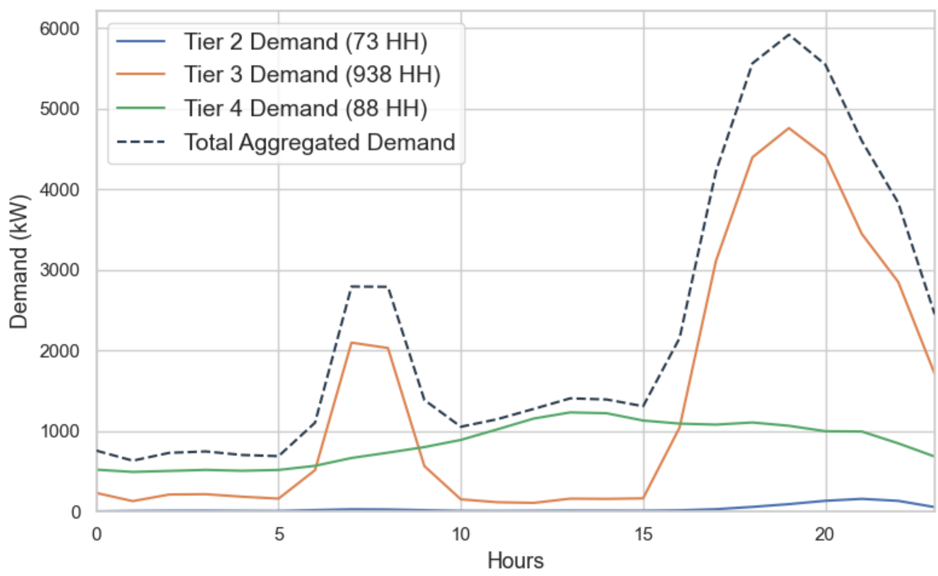


Figure 2.4: Example of aggregated demand for a settlement (AY - lat [-10, 10])

2.2 OnSSET formulation

OnSSET is a GIS-based optimization tool developed to support electrification planning and decision making to achieve energy access goals in currently non-electrified locations; like MicroGridsPy, it adopts a bottom-up approach, hence providing high technological detail without accounting for complex, economy-wide interactions.
The primary distinguishing feature of OnSSET is its high spatial definition, for which detailed geospatial inputs are required, such as distance from the existing grid infrastructure, population clusters (settlements) and renewable resource potential maps (including solar irradiation and wind speed).
The software initially processes the geospatial input data through a calibration that in-

cludes the verification of the data using national statistics, subsequently modeling the starting year electrification status. It then calculates the optimal LCOE associated with different electrification paths: national grid extension, mini-grids and stand-alone solar systems; it later compares the options, finally choosing for each settlement the one with the minimum LCOE. The extension of the grid is handled with a progressive logic, explained in detail in Paragraph 2.2.3.
It is important to highlight how OnSSET operates on a single target year: when optimizing the electrification path, it accounts for all the capacity to be installed in the target year; when used on multiple time-steps the tool operates on a sequential logic, finding the optimal solution in the first time-step and then using these results to plan the electrification in the final year. Unlike MicroGridsPy, OnSSET does not operate with perfect foresight, meaning the optimization of the electrification plan in a timestep does not account for the entire time-horizon.

2.2.1 Demand Definition

The way the software handles the demand definition has to be addressed with particular attention, since a key part of Chapter 3 revolves around the modification performed on the demand definition to ensure a consistent analysis.
In its standard configuration, OnSSET defines the demand by associating a tier to each settlement: as previously mentioned, the geospatial tool bases the yearly electricity consumption tiers on the classification proposed by ESMAP [19] of the *Multi-Tier Framework*, as reported in the table below.

	Tier 1	**Tier 2**	**Tier 3**	**Tier 4**	**Tier 5**
$[kWh/hh/y]$	≥ 4.5	≥ 73	≥ 365	$\geq 1'250$	$\geq 3'000$

Table 2.1: Multi-Tier Framework classification of household consumption [19]

When a custom demand is not provided, the user defines one tier for each of the following possible categories:

- urban
- big rural
- small rural

The software does not require the urban/rural categorization as an input, just the urban ratio in the starting and final years; this ratio is used to calibrate the population by defining a yearly urban growth rate as follows:

$$r_u = \left(\frac{u_f \cdot pop_f}{u_i \cdot pop_i}\right)^{1/project\_life} \tag{2.3}$$

- r_u: yearly urban growth rate
- u_i and u_f: urban ratio in the starting year and as expected in the final year of the analysis
- pop_i and pop_f: population in the starting year and as projected in the final year of the analysis
- $project\_life$: time-frame of the analysis

The tool initially categorizes all settlements as "rural" and rearranges them from largest to smallest, population-wise; then, the "urban" classification is assigned to settlements in descending order, while keeping track of the resulting cumulative sum of urban population. When the sum reaches the total urban population in the current time-step, derived from the urban ratio, all remaining settlements are then categorized as "rural". The *big rural* and *small rural* differentiation is imposed through the *rural cut-off*, defined as the minimum population above which a rural settlement is considered "big".
OnSSET also offers an alternative way to define the demand that allows the user to provide the values of per household demand for each settlement as input; with this information the software can safely assign a tier to each settlement. This second way is the *Custom Demand* definition; Chapter 3 details how GIS data was used to define the demand, using the *custom demand* method, ensuring input consistency with MicroGridsPy and increasing the detail of the analysis.
It is important to note that, regardless of what type of demand input is considered, OnSSET uses the values of demand assigned to the tiers by the user as the lower limits of each consumption bin. Indeed, once a settlement has an assigned per household demand, the software compares that value with the input bins and attributes it the corresponding tier: for example, if a settlement presents a per household consumption of 500 kWh/y and the input bins for Tier 3 and Tier 4 are respectively 100 kWh/y and 600 kWh/y, OnSSET assigns Tier 3 to the settlement. The tier assignment determines not only the normalized curve that is linked to the settlement (see Paragraph 2.2.2), but the *Average to Peak Ratio* too: it is defined as the ratio between the average and peak loads as presented by the curves, and it is used to size the capacity required by technologies such as the national grid and hydro mini-grids, as well as the transmission and distribution systems.

2.2.2 Hybrid Systems Modeling

Particular attention needs to be given to how the software models the hybrid mini-grids. While MicroGridsPy is a tool dedicated to optimizing a single decentralized system, OnSSET analyzes thousands of potential mini-grids sites, hence it cannot afford to run a full-fledged optimization for each of them; for this reason, the system sizing does not happen for every individual settlement, but a look-up table is compiled for each type of hybrid system with sample mini-grid solutions. Each of these solutions is computed via the combination of three key parameters that vary in defined ranges:

- **Global Horizontal Irradiance (GHI) / Wind Speed:** the software identifies the minimum and maximum values across the country and creates a sequence of fixed-step values within the limits
- **Diesel Price:** a sequence of values is delineated similarly to the GHI / wind speed levels
- **Tiers:** values from 1 to 5, each representing a settlement electricity demand tier

The resulting look-up tables, one each for the solar and wind resources, are 3D matrices of all possible combinations and OnSSET computes the optimal LCOE and composition of PV/wind turbines, diesel generator and battery system capacities for them all. Each settlement is assigned the LCOE and sizing corresponding to the closest combination of RES, diesel price and tier; the system is then scaled to the settlement electricity demand, as every sample mini-grid is sized to cover a total yearly load of 10'000 kWh.
With the aim of modeling hybrid systems, hourly timeseries of renewables resources availability are required, both for solar irradiation and wind speed. The software requires a singular timeseries representing the country for each renewable resource, then scaling it with the values in the RES ranges previously described as follows:

$$GHI_{scaled}(t) = \frac{GHI(t)}{\sum_{h}^{8760} GHI(h)} \cdot g \cdot 1000 \quad \forall t \in \{1, ..., 8760\} \tag{2.4}$$

**Multiplication by 1000 is done to convert from W/m^2 to kW/m^2*
A similar scaling method is employed for the wind speed timeseries as well.
Each aforementioned consumption tier is linked with a specific normalized daily load curve, consisting of a series of hourly shares of average daily demand; the series is then repeated 365 times to represent an entire year. Lower tiers present lower average to peak load ratios, as most of the demand is concentrated in evening hours, while higher tiers are characterized by a more even profile due to the use of electricity intensive appliances

during the day, as it is depicted in Figure (2.5).

In order to switch from hourly shares to actual hourly demand load, each share in the entire series is multiplied by the average daily demand, computed as:

$$daily\_demand = \frac{yearly\_demand}{365} \tag{2.5}$$

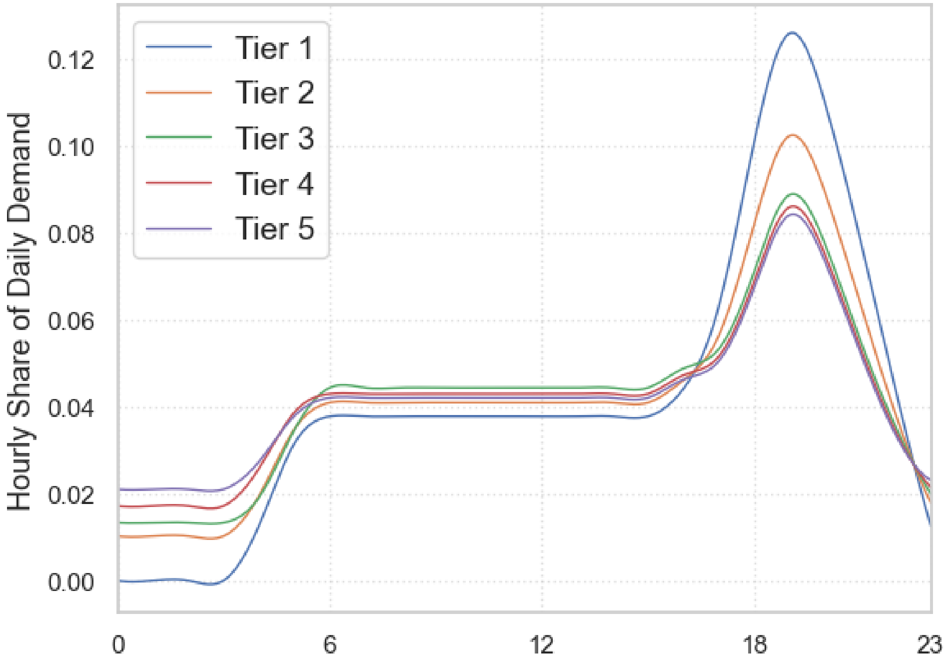


Figure 2.5: Normalized daily load curves for hybrid modeling, OnSSET

The dispatch strategy of the mini-grids is not optimized as it is the case for MicroGridsPy, instead following through a defined logic.

The main objective of the dispatch strategy is to satisfy the demand for every hour of the year while exploiting the potential of renewables and reducing the use of the diesel generator, that is limited to a user-defined maximum production share. Every hour, the RES exploitation is prioritized, meaning the renewable power generation is subtracted from the load demand; the remaining load is the *Net Load* (P_{net}), which can either be positive (if the demand was not fully satisfied) or negative (if the demand was fully satisfied and there is leftover renewable power).
To ensure a full exploitation of the renewable resource, the rest of the strategy is split in three operating bands, based on the period of the day.

Band 1: Daytime (04:00 - 17:00)

During the day the system prioritizes the dispatch of RES and batteries; if $P_{net} < 0$, the leftover RES power is used to charge the batteries, while if $P_{net} > 0$ the batteries are discharged to supplement the deficit unfilled by the RES; the diesel generator is turned on only in case the batteries fully discharge before satisfying the load.

Band 2: Evening (17:00 - 23:00)

This band prioritizes the use of the diesel generator; since solar power generation is negligible during evening hours, the diesel generator is dispatched at a minimum of 40% capacity (to avoid low efficiency) to fulfill the load and to fully charge the batteries for the night; in the eventuality that the diesel generator might be insufficient, the batteries are dispatched.

Band 3: Night (23:00 - 04:00)

This band prioritizes the dispatch of the batteries; ideally the BESS were charged during the previous two time bands but, in case the state of charge might be insufficient, the diesel generator activates once more (at minimum 40% capacity), not only to help meet the load, but also to fully charge the batteries so that it can shut down as soon as possible.

The last concept required to fully grasp the hybrid systems modeling is how OnSSET optimizes the capacities of RES, BESS and diesel generator to reach the lowest LCOE. To solve this multi-variable optimization problem, OnSSET utilizes a stochastic Python-based metaheuristic method from the SciPy library known as *Differential Evolution* (DE). *Differential Evolution* is a population-based evolutionary algorithm designed to search a multi-dimensional space by iteratively evaluating and evolving combinations of the three capacity sizes until the global minimum LCOE for the sample mini-grid is reached.
Utilizing this evolutionary approach inherently introduces a trade-off: it sacrifices computational speed if compared to traditional differential solvers but it compensates with mathematical robustness. Since DE generates and maintains a large population of candidate solutions across the search space, it is less likely to become trapped in sub-optimal local minima.
By employing this method, OnSSET is able to generate the look-up tables with exceptionally high empirical probability of having the capacity configurations associated with the true global minimum of LCOE.

2.2.3 Grid Extension Logic

What can be considered as the "core" of the OnSSET software is its grid extension logic, specifically how it simulates the progressive expansion of the grid during the years by deciding where to extend the grid and where it's more optimal to adopt off-grid solutions. Before detailing the extension algorithm, it is necessary to briefly address how OnSSET decides, for a given settlement in a given timestep, which technology to adopt: the decisive parameter is the LCOE, as the option resulting with the lowest LCOE is picked, but there are some intricacies concerning how the software decides to compare the different technologies with the grid extension option specifically. The grid LCOE of unelectrified settlements in a given timestep is calculated separately from other technologies' and it is deeply intertwined with the extension logic.
As a first step, the software compiles a list of the unelectrified settlements, ordering them by proximity to the existing grid. It must be highlighted that OnSSET prioritizes densification of the grid, meaning that all new households present in previously grid-connected settlements will be automatically connected to the national network.
There are user-defined limits regarding both the new connections to the grid and its capacity expansion: as mentioned, the software prioritizes grid densification over fresh new connections, so it is possible for these limits to be saturated before the expansion starts if the densification is too demanding; if that were the case, the software cannot proceed with the expansion algorithm and the grid will not be extended in that timestep.

Each settlement is assigned a minimum off-grid LCOE chosen by comparing all the decentralized options, as well as a "base" grid LCOE computed without accounting for the external lines, hence considering only generation and internal distribution costs: by subtracting the base grid LCOE from the minimum off-grid one, it is possible to compute the "cost margin" of choosing grid over off-grid in $\$/kWh$ (eq. (2.6)); this value is then multiplied by the sum of the discounted generation for all remaining years as shown in Equation (2.7a). The resulting parameter is called $max\_investment$ and it represent the total cost margin in \$ between grid and off-grid that can be used to extend the transmission lines, which were left out of the calculation for this precise reason.

$$cost\_margin = LCOE_{off\_grid}^{min} - LCOE_{grid}^{base} \tag{2.6}$$

This parameter is converted into a distance by dividing it by the cost of the transmission lines per unit length; the resulting value is named max_dist, and it represent the maximum distance from the grid that a settlement can be at before it becomes more convenient to pick the best off-grid option over extending the grid.

$$\begin{cases} max\_investment = cost\_margin \cdot \sum_{y}^{N} \frac{el\_gen}{(1+d)^{y-y_0}}, & \text{(2.7a)} \\ max\_dist = \frac{max\_investment}{lines\_cost}. & \text{(2.7b)} \end{cases}$$

It is now possible to explain how the iterative process of extending the grid works: from the previously mentioned list of unelectrified settlements, the software begins an iteration by scanning the list in descending order, checking every settlement by quickly computing max_dist until it finds one whose minimum distance from the grid is lower than max_dist; that settlement is promptly connected to the grid and its required capacity and number of households are subtracted from the aforementioned limits, being how many new settlements can be connected to the grid and how much new grid capacity can be added. The newly electrified settlement is then added as a grid node, eventually creating also intermediate nodes if the distance from the grid is greater that 0.75 km; since the grid has changed, the iteration is concluded and the software starts over from top of the list that is now one settlement short. The iterative process continues until one of the following three conditions occurs:

- There are no more villages closer than max_dist
- All settlements from the list are electrified
- One of the limits reaches saturation

For a simplified visual representation of the grid-extension algorithm, please refer to the flowchart reported in Figure 2.6.

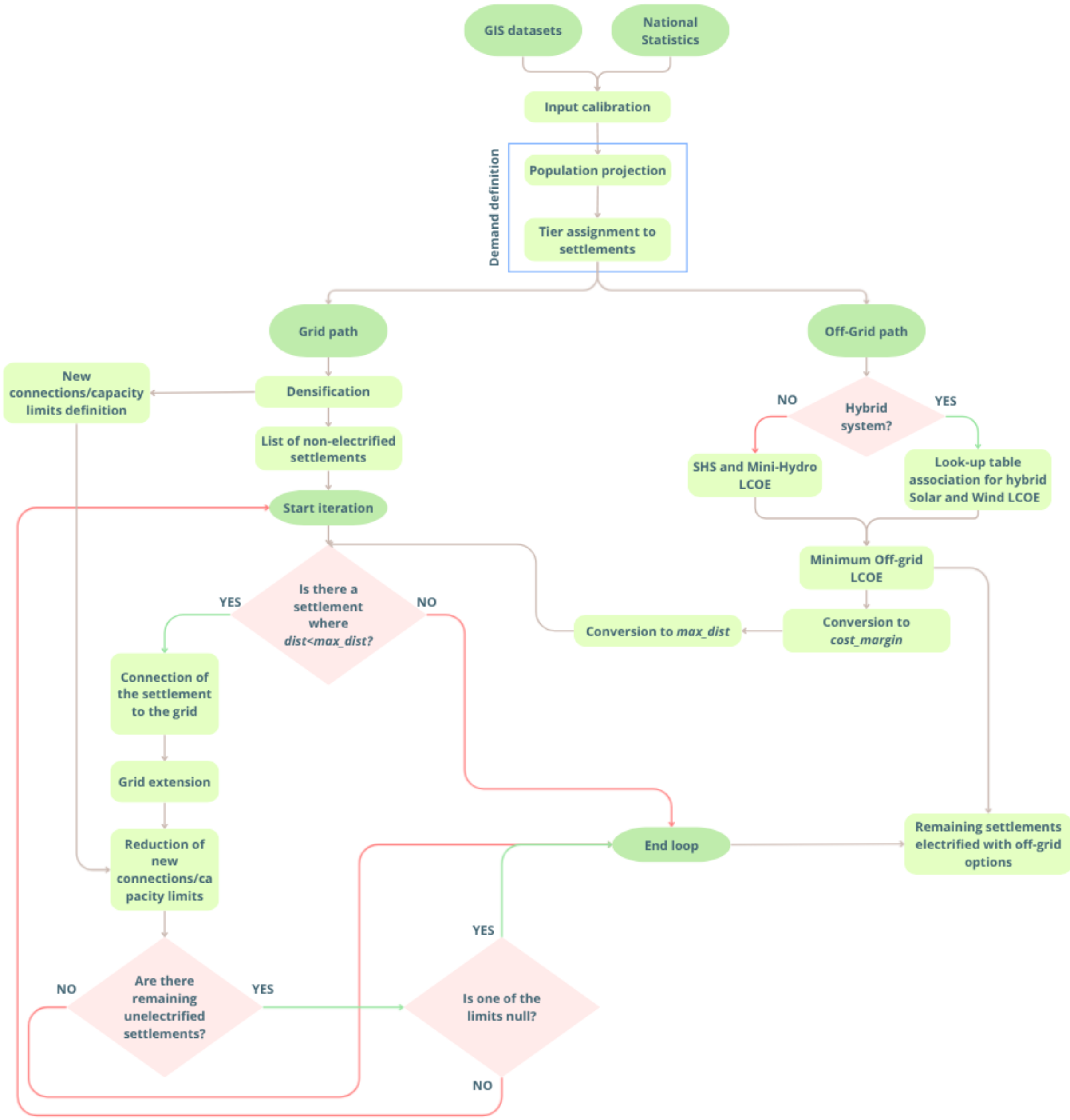


Figure 2.6: OnSSET grid extension algorithm flowchart

2.2.4 OnSSET versions

OnSSET is a dynamic open-source project, meaning that over the past decade it was continuously updated and improved, ever since it was first published. The first version [21], still used for didactic purposes to understand the core of the OnSSET algorithm, includes the following technology choices:

- Grid
- Solar/Diesel Stand-Alone Systems
- Solar/Diesel/Wind/Hydro mini-grids (no hybrids)

The modeling of the solar and wind systems was introduced at a basic level, only based on geospatial data regarding solar global horizontal irradiance (GHI) and wind capacity factor. In later versions, a more detailed modeling of wind and solar based systems was gradually implemented, leading to the current framework described in Section 2.2: it integrates the diesel technology as a back-up system for hybrid mini-grids and adding the use of battery storage systems; the variable availability of the renewable resource is accounted for, as well as a standard dispatch strategy and normalized load curves to size the system components with more accuracy.

Despite these improvements in hybrids modeling, the load demand curves implemented in OnSSET are only five, each linked to one of the defined tiers of consumption: each settlement is assigned one of these tiers, lacking any location-specific detail that may influence the load curve, such as the cooling needs. In MicroGridsPy, as it has been previously described, there is instead the possibility to build a demand curve that reflects a more variable behavior, accounting for different types of users and locations.

As mentioned in Chapter 1, a mini-grid dedicated tool has been released by the developers behind OnSSET: OnSSET for MiniGrids (OMG) is a software based on geospatial information that has as purpose the identification of the least cost energy system for a limited area. It introduces also the computation of the distribution grid mapping, including in its costs analysis also the ones related to non-generation factors. OMG is designed for fast optimization, offering a rapid estimation of capacities to install, the distribution network extension and their related costs. While being an alternative to software like MicroGridsPy, it is considered to be more apt for feasibility studies rather than energy system modeling [6].

3 | Methodology

This chapter will first focus on the gathering and elaboration of geospatial data required as inputs by OnSSET; then, the method used to introduce Sub-Saharan demand archetypes in the geospatial software will be explained; finally, all the steps followed in order to ensure a consistent analysis of OnSSET and MicroGridsPy results will be described.

3.1 Geospatial data preparation

OnSSET is a tool based on geospatial analysis, therefore it requires specific GIS data input, detailed in Table 3.1. Note that not all listed datasets are mandatory, such as the existing grid infrastructure, but for a more thorough analysis it is recommended to include as many as retrievable.

It should be noted that layers such as elevation and land cover may be required in future versions of OnSSET, since the source code includes the option to apply a grid penalty factor that depends on morphology that may hinder the extension of the grid; as of the currently available versions, this functionality is not implemented, so the aforementioned datasets are not necessary for in this analysis.

Most of the listed layers can be found in online repositories, with little modifications required: it needs to be ensured that the coordinate reference system (CRS) is EPSG:4326, and minimal adjustments to reduce the dataset size to the analyzed country extent can be carried out.

Before being used as input for OnSSET, all the layers listed in Table 3.1 are processed via an additional software, the *Extraction Notebook* [22]: for raster datasets, the extraction notebook assigns to each settlement either the mean or most repeated value in the area of the cluster; shapefiles are instead processed in order to compute meaningful values for OnSSET input (for example: the road network, for each cluster, is used to compute the distance from the closest road).

Existing electrical infrastructure

As mentioned in Chapter 1, in some cases the fast development of the electrical infrastructure in Sub-Saharan Countries means even relatively recently updated available datasets are outdated: to ensure an up-to-date electrification plan, it may be necessary to retrace revised maps on QGIS to obtain a usable format.

Custom Demand

For this analysis, the layer was generated based on the relative wealth index (RWI). The RWI is an indicator of the relative standard of living within countries: it is available on online repositories and it spans a range from -3 to +3, with 0 representing the average wealth level of the country. The range specific to the country in analysis is then split in five equal bins and each is linked with a specific yearly household consumption value in $[kWh/hh/y]$; the values used during this analysis will be detailed in Paragraph 3.3, as they were chosen to offer a consistent input with MicroGridsPy. It is worth highlighting that the demand associated with each settlement is computed by the Extraction Notebook as the weighted average of the custom demand raster file, where the weights are defined by the proportional area of the settlement covered by pixels of each demand value.

Existing Mini-grids

Concerning the existing mini-grids, it should be noted that in OnSSET they are considered only to calibrate the currently electrified population but no information regarding the systems' size or characterization is required or accounted for.

Dataset	Type	Description
Administrative boundaries	Polygon	Delineates the boundaries of the analysis
Population clusters	Polygon	Identifies the settlement area, as well as its population count
Solar GHI	Raster	Global Horizontal Irradiation [kWh/m2/y], used to identify the suitability of PV systems
Wind velocity	Raster	Wind velocity [m/s], used to identify the suitability of aerogenerators
Nighttime lights	Raster	Used to identify and spatially calibrate the currently electrified population
Custom demand	Raster	User defined electricity demand in each settlement
Substation	Point	Current substation infrastructure used to identify and spatially calibrate the currently electrified population
Existing HV lines	Line	Used to identify and spatially calibrate the currently electrified population
Existing MV lines	Line	Used to identify and spatially calibrate the currently electrified population
Road network	Line	Current road infrastructure used to identify and spatially calibrate the currently electrified population
Travel time	Raster	Visualizes spatially the travel time required to reach from any individual cell to the closest town with more than 50'000 people
Distribution transformers	Point	Current transformer infrastructure used to identify and spatially calibrate the currently electrified population
Hydro-power potential	Point	Mini/small hydro power potential, dataset developed at KTH dESA including environmental, social and topological restrictions
Existing mini-grids	Point	Location of existing mini-grids, used to identify and spatially calibrate the currently electrified population

Table 3.1: Geospatial data required by OnSSET

3.2 Integration of Archetypes in OnSSET

In order to achieve greater time-resolution for hybrid mini-grid sizing in OnSSET and to significantly increase the input consistency between the two software, the archetypes load profiles used for MicroGridsPy's demand definition are implemented in OnSSET; they serve as a replacement for the default load curves linked to the tiers, as explained in Paragraph 2.2.2.
MicroGridsPy's demand is defined with hourly resolution, while the load curves originally implemented in OnSSET are of hourly shares of average daily demand then repeated for all year: to maintain the same format, the archetypes load curves are normalized by dividing the hourly values by the total yearly demand; the conversion of the timeseries in absolute values is then slightly modified, as instead of multiplying each hourly share by the average daily demand, it is multiplied by the yearly consumption.
This integration also increases the diversification of the settlements, since what originally depended only on the tier (5 possible values) now depends on 3 different variables (cooling period, latitude and RWI), as described in Paragraph 2.1; it is important to note that the core structure of the OnSSET algorithm was not modified despite this introduction, regarding the size of the look-up table this change only impacted 1 of the 3 variables that constitute the look-up tables, namely the tiers.
The archetypes implementation was made possible by combining the 3 parameters in a single one, the *Archetype ID*, that can assume values ranging from 0 to 99, each of which corresponding to a unique archetype. To be able to perform this connection it was necessary to exploit the way the load curves files for each archetype are saved in MicroGridsPy's repository, and rather than use the name of the file it was convenient to focus on their ordination; the files follow the format *Cooling-Lat-Tier* and are ordered sequentially. The sequence progresses by incrementing the variables in a nested fashion: the *Tier* index increments first, followed by the *Latitude bin*, and finally the *Cooling period*, which increments only after all combinations of the previous two have been exhausted.
The formula used to assign the *Archetype ID* to the corresponding curve is:

$$Archetype\ ID = (x \cdot 25) + (y \cdot 5) + z \tag{3.1}$$

- **x, y, z:** positional indexes of respectively *Cooling Period*, *Latitude* and *Tier*
- there are 25 possible combinations of latitude and consumption tiers
- 5 are the possible values the consumption tier can assume

With such implementation, it is possible to easily access each curve via the *Archetype ID* and to assign the corresponding ID to each settlement according to its specific combination of the three categories.

3.2.1 Average to Peak Ratios

In Paragraph 2.2.1, concerning the demand definition in OnSSET, it is highlighted how the tier assignment influences not only the load curves for hybrid mini-grids modeling but the sizing of other technologies as well. Archetypes' load curves present very different shapes compared to the ones used by the original OnSSET (Figure 3.1), hence a modification to remain consistent was due: as illustrated in Figure 3.2, the *Average to Peak Ratios* of the various combinations of cooling periods and latitude bands do not vary significantly from an average value for each consumption tier; for this reason, no structural changes to the source code were necessary, as the only modification consisted in replacing the default values associated with each tier with their respective archetype averages.

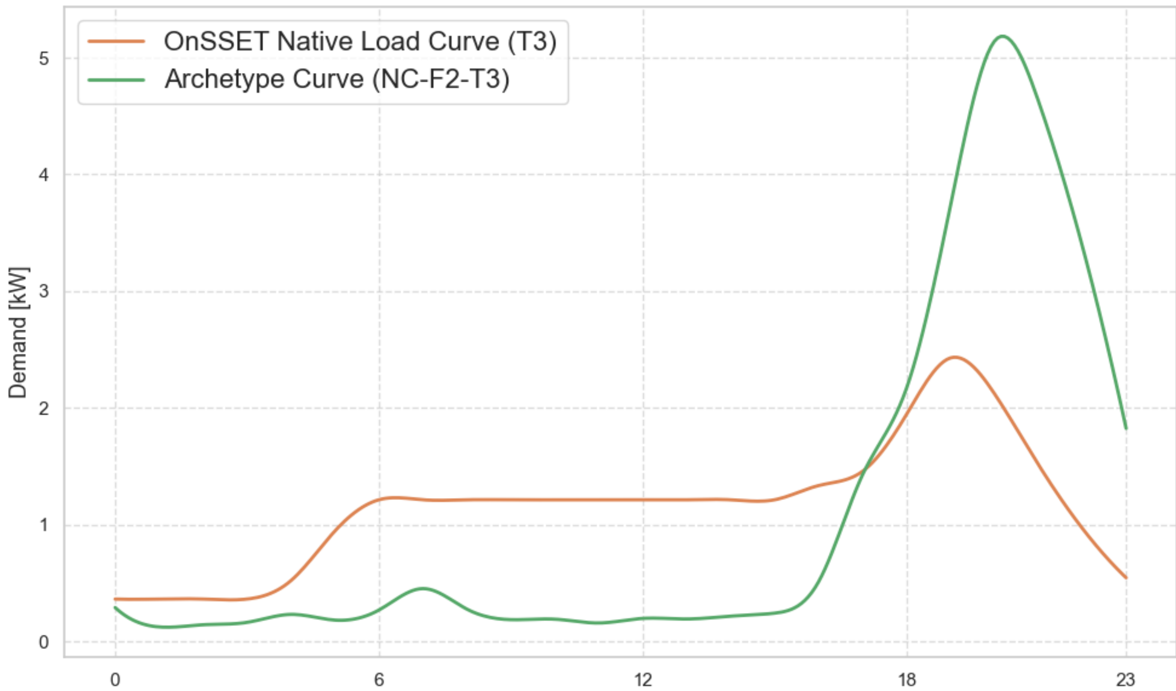


Figure 3.1: Comparison between load curves shapes, example for Tier 3 (reference settlement of 10'000 kWh/y)

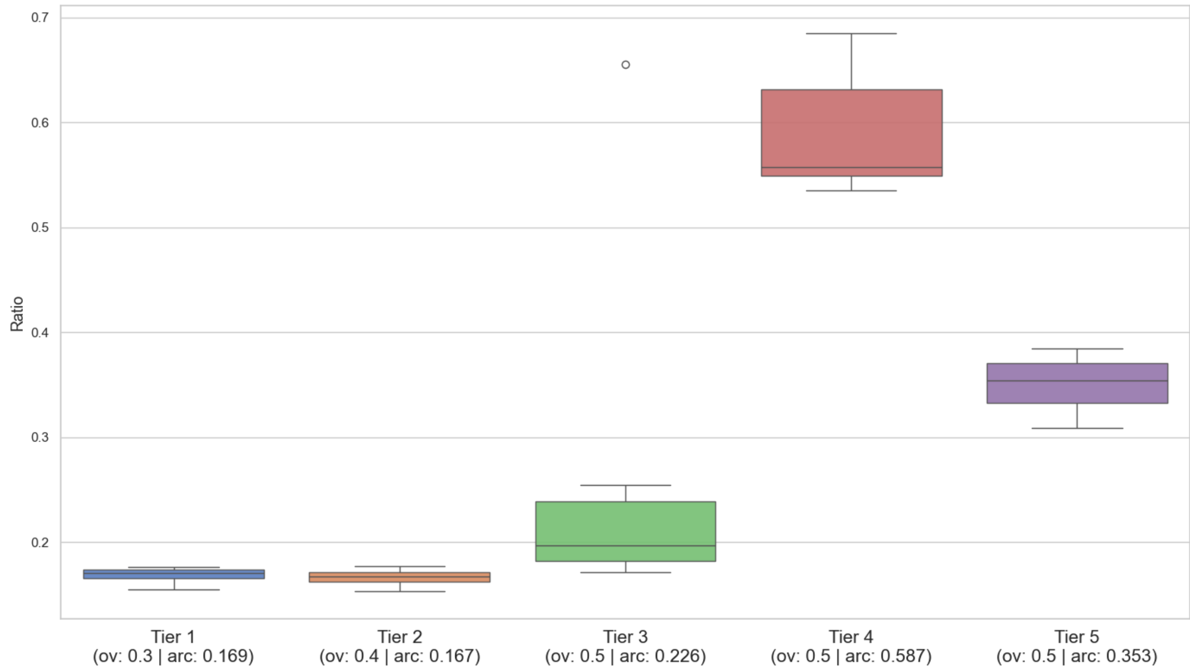


Figure 3.2: Average to Peak Ratios of the five wealth tiers
OV: original OnSSET curves *average to peak ratio* values
ARC: average archetype curves *average to peak ratio* values

3.3 OnSSET and MicroGridsPy: a consistent analysis

This section provides an overview of the measures employed in this study to ensure a consistent analysis between two software; these are necessary due to the many differences embedded in the tools' approaches to modeling.

3.3.1 Techno-economic characterization of Technologies

All common input parameters regarding the techno-economic characterization of the technologies (solar photovoltaics, diesel generator, battery storage systems) were set to be the same, with the due conversions. The considered input data are reported in Tables A.2, A.3.

3.3.2 Cooling Period category assignment

One of the defining characteristics of the archetype curves is the classification based on cooling needs, which may vary from settlement to settlement depending on local temperatures. According to the expected use of appliances such as fans and air conditioning

relative to different climates, each settlement can be classified in any of the following categories: No Cooling (NC), All Year (AY), October to March (OM) or April to September (AS).
In the study introducing archetypes [16], the classification is based on the *Cooling Degree Days* (CDD) computation as described by *Falchetta and Mistry*(2021) [20]; the used correlations, upon which the present research is also based, are reported in Table 3.2. The Cooling Degree Days are defined as *"the number of degrees that the daily temperature is above an arbitrarily defined comfort temperature T_{base}"* [20].

Condition	**CDDs**
$T_{\text{max}} \leq T_{\text{base}}$	$CDD = 0$
$T_{\text{avg}} \leq T_{\text{base}} < T_{\text{max}}$	$CDD = \frac{(T_{\text{max}} - T_{\text{base}})}{4}$
$T_{\text{min}} \leq T_{\text{base}} < T_{\text{avg}}$	$CDD = \left[\frac{(T_{\text{max}} - T_{\text{base}})}{2} - \frac{(T_{\text{base}} - T_{\text{min}})}{4}\right]$
$T_{\text{min}} \geq T_{\text{base}}$	$CDD = T_{\text{avg}} - T_{\text{base}}$

Table 3.2: Cooling Degree Days calculation methodology [20]
T_{base} arbitrarily defined at 26°C

Calculating the CDD for every settlement would be extremely computationally intensive, as individual time-series of daily T_{max}, T_{avg} and T_{min} would be required; it was possible to mitigate this issue by aggregating the daily-resolution data into a monthly indicator named *Cooling Degree Month* (CDM). The CDM is computed on QGIS by applying the relations reported in Table 3.2 for monthly values of T_{max}, T_{avg} and T_{min}, instead of daily. A further aggregation is then required as the archetypes cooling periods present a temporal resolution in a biannual basis; the approach simply consists of adding the CDMs for each of the two seasons, resulting in new parameters named AS_{CDM} and OM_{CDM}.
As a final step, the actual classification of the cooling period for each settlement is done by comparing these fields with a user-defined threshold following the boolean logic reported below.

$$\text{Cooling Category} = \begin{cases} AY & if\ AS_{CDM} > Threshold \text{ and } OM_{CDM} > Threshold \\ AS & if\ AS_{CDM} > Threshold \text{ and } OM_{CDM} < Threshold \\ OM & if\ AS_{CDM} < Threshold \text{ and } OM_{CDM} > Threshold \\ NC & if\ AS_{CDM} < Threshold \text{ and } OM_{CDM} < Threshold \end{cases} \tag{3.2}$$

3.3.3 Yearly electricity consumption per household

The *Custom Demand* input file is created using values that are coherent with the demand files of MicroGridsPy. As mentioned in Section 3.1, the elaboration of this layer is based on the subdivision of the RWI in five bins: each of those bins is assigned a specific yearly household consumption, depending on latitude and cooling category.
The values used as yearly household consumption for those categories are extracted from the corresponding archetype file: each of these is essentially a series of 8760 values of hourly demand for an ensemble of 100 households, so the yearly household consumption is computed for every file as:

$$D_{hh}^{CC-F-T}\left[\frac{kWh}{hh\cdot y}\right] = \frac{\sum_{h}^{8760} D_{hourly}^{CC-F-T}(h)}{total\_households} \tag{3.3}$$

- D_{hh}^{CC-F-T}: household yearly demand in CC cooling category, F latitude band and T wealth tier
- D_{hourly}^{CC-F-T}: hourly demand for the reference settlement of 100 households of CC cooling category, F latitude band and T wealth tier
- $total\_households$: number of households in the reference settlement (100)

It is worth to note that, as highlighted in the reference paper for archetypes integration [16], the reference settlement is composed of 100 households to account for stochasticity of use of appliances.

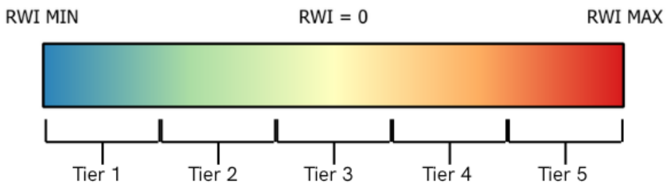


Figure 3.3: RWI bins and corresponding tier of consumption

It is important to highlight that a strong underlying assumption is made with this procedure: it is supposed that the country's average household is classified as Tier 3.

In Chapter 4 the detailed values of household consumption assigned to each tier are reported, according to both latitude and cooling category, specifically to the case study

analyzed.
As final step, the logic OnSSET uses to assign tiers to the settlements once the demand is given (described in Paragraph 2.2.1) was modified: the tier linked to the settlement, instead of being the one with the immediately lower corresponding value of demand, is the closest; hence, in the example previously used the settlement becomes of Tier 4.

3.3.4 Households distribution

An essential difference between OnSSET and MicroGridsPy is that while the first can only refer to a single tier of consumption to represent a settlement, the latter can better describe it by considering the different residential consumption levels within it.
Limiting the demand definition in MicroGridsPy to overcome this intrinsic difference by assigning a unique tier to all households would be detrimental: first, it would mean to ignore one of the tool's strengths (demand definition on high detail) to comply to a software lacking this feature. More importantly, a critical discrepancy renders this option unsuitable for comparison: even with an identical single tier representation, the two tools would compute different total demand. This occurs because OnSSET assigns a tier based on the demand bin but it applies the actual *Custom Demand* value for the per-household yearly load, while MicroGridsPy directly uses the central value of the bin; the difference between the two values, however small it may be, is significant as it adds up for every household in the settlement.
To address both issues it was decided to adopt an empirical solution; as proposed by *Stevanato et al.*(2023) [16] in the case study section of the paper, the Gaussian distribution based on RWI is chosen to describe the settlements in this research: the mean value of RWI was calculated within each cluster's boundaries, while its standard deviation was extended to county limits. A further consideration is required to ensure a reasonably similar average demand, since the presence of households of different tiers is bound to perturb it significantly: what was previously described as the mean RWI value, from which the mean demand derives, is now the initial value of a virtual mean RWI that is shifted by an algorithm so that the resulting weighted average household consumption is as close as possible to the one considered by OnSSET.

Mathematical formulation of the calibrated distribution

- N_{hh}: total number of households in the settlement
- μ_{RWI}: real mean value of RWI in the settlement
- σ_{RWI}: standard deviation of RWI in the county

- b_i: RWI bins, with $i \in [0, 5]$ ($b_0 = -\infty$, $b_5 = +\infty$)
- D_i: household consumption in the -*ith* tier
- D_{target}: actual average household consumption in the settlement
- $\Phi\left(\frac{b_i - \mu_{RWI}}{\sigma_{RWI}}\right)$: cumulative distribution function of the $-ith$ bin

As the RWI is assumed to have a normal distribution in the settlements, the probability that a household belongs to the -*ith* tier is computed as:

$$P_i(\mu_{RWI}) = \Phi\left(\frac{b_i - \mu_{RWI}}{\sigma_{RWI}}\right) - \Phi\left(\frac{b_{i-1} - \mu_{RWI}}{\sigma_{RWI}}\right) \tag{3.4}$$

The expected average yearly household consumption $E[D(\mu_{RWI})]$ computed with the distribution is given by the weighted sum of the probabilities multiplied by the consumption of each tier:

$$E[D(\mu_{RWI})] = \sum_{i=1}^{5} P_i(\mu_{RWI}) \cdot D_i \tag{3.5}$$

The objective function in order to achieve the calibration is hence the following:

$$\sum_{i=1}^{5} P_i(\mu_{RWI}) \cdot D_i - D_{target} = 0 \tag{3.6}$$

The number of households belonging to each tier is finally computed as the total number of households multiplied by its calibrated probability:

$$N_i = P_i(\mu^*_{RWI}) \cdot N_{hh} \tag{3.7}$$

Where μ_{RWI} is the calibrated mean value of RWI found with the objective function.

A visual representation of an example of the calibration can be found in Figure 3.4, as well as the difference between the number of households assigned in each tier.

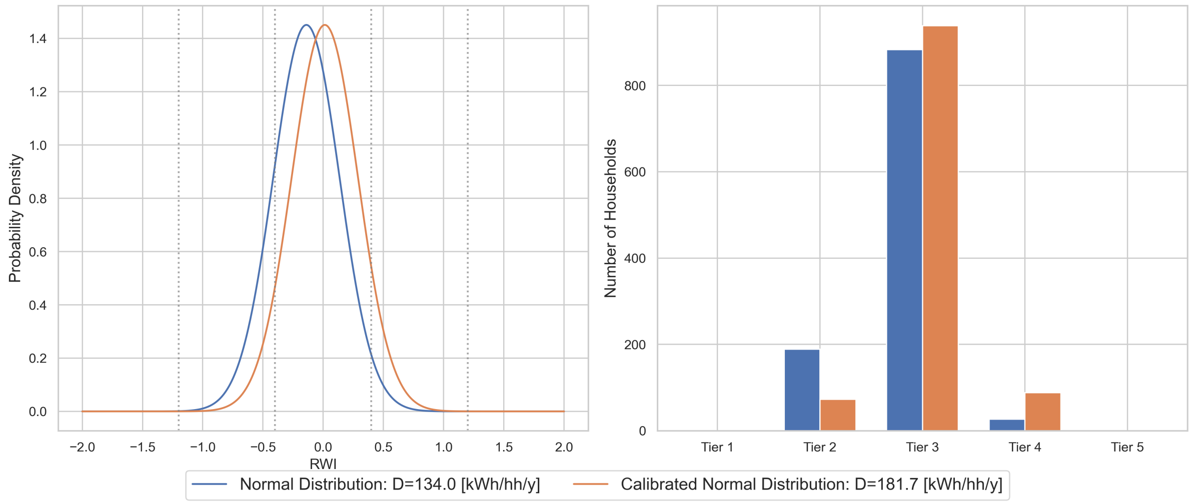


Figure 3.4: Comparison between households distributions for a given settlement

3.3.5 Comparison of OnSSET and MicroGridsPy results

MicroGridsPy offers as results the specific electricity production mix of the settlement of interest, listing the installed capacity of each technology in every time-step and a detailed cost analysis as well; the latter highlights the LCOE in the final year of the project, the NPC (the minimization of which is the objective function of MicroGridsPy) and the actualized investment and operation and maintenance costs, both variable and fixed. In the OnSSET source code, instead, there is a less detailed depiction for each settlement: the installed capacity of hybrid technologies is the total generation capacity, not including the storage system and not specifying the generation mix; moreover, the LCOE and investment costs, which are the results used for economic comparison, are inclusive of the distribution network within the settlement, therefore accounting for costs that are not considered in the used version of MicroGridsPy. The OnSSET code was therefore slightly modified so that the results file also includes the economic parameters related solely to generation and storage, as well as the specific installed capacities of solar, wind, diesel and battery technologies.

On a final note, the diesel cost in the settlement was also extrapolated from OnSSET in order to feed it as MicroGridsPy input: it is computed with the Szabo formula (3.8), accounting for fuel transportation costs to reach remote areas.

$$C_{fuel} = \left(p_{fuel} + \frac{2 \cdot p_{fuel} \cdot truck\_consumption \cdot travel\_time}{truck\_volume} \right) \frac{1}{LHV_{fuel} \cdot \eta} \tag{3.8}$$

- C_{fuel} : actual fuel cost in the settlement

- p_{fuel} : average national fuel price
- *truck consumption:* fuel consumed by the truck to transport the fuel to the settlement location
- *travel time:* time required to reach the settlement from the closest urban center
- *truck volume:* truck fuel tank
- LHV_{fuel}: lower heating value of the fuel
- η: efficiency of the fuel based generator

4 | Case Study

4.1 Electricity Access in Kenya

This analysis uses Kenya as a case study to evaluate the implementation of archetypes in OnSSET and the subsequent comparison with MicroGridsPy's results. Kenya is one of the largest economies in SSA, but while it has reached important achievements in electrification over the past two decades it still presents a critical situation in remote, rural areas. Indeed, despite the great electrification level if compared to other SSA countries, the northern regions are barely connected to the national grid and great part of electrification is due to off-grid systems, either mini-grids or stand-alone systems. This has been brought to attention by the IEA (International Energy Agency) in the *2024 Energy Policy Review* [23], where it has been reported that almost one in five Kenyans access electricity via off-grid technologies, and Kenya has the highest share of households, globally, that accesses electricity via Solar Home Systems (SHS).

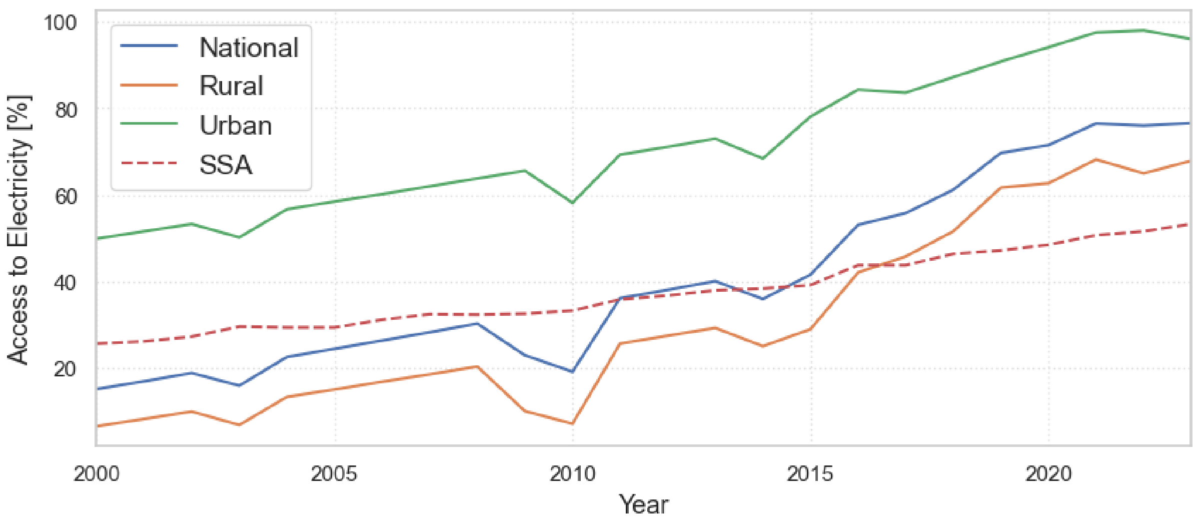


Figure 4.1: Electricity Access [%] in Kenya, from WorldBankData [1]

The electricity mix in Kenya relies majorly on renewable resources, with RES making up nearly 90% of power generation according to the IEA [24]. This is due to the great renewable potential in geothermal, solar, wind and hydro resources across the entire Country; this potential, combined with the harsher morphology of northern territories, makes off-grid systems a reasonable alternative to the expansion of the national grid; in addition, it must be considered that the centralized network is neither particularly affordable, with the end-user price at 0.256 USD/kWh being one of the highest of the continent [23], nor reliable, with transmission losses amounting to over 20% [25] and reporting frequent system outages.

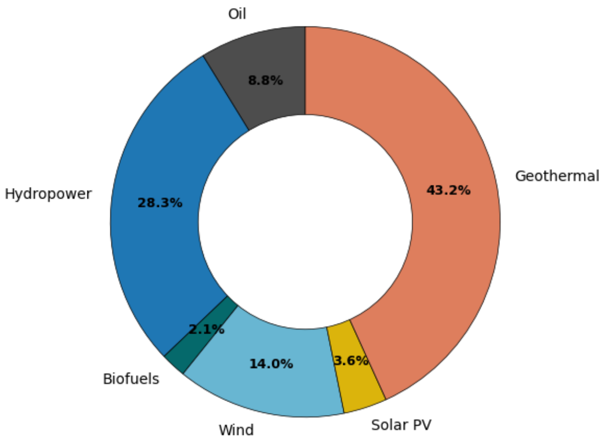


Figure 4.2: Kenya Electricity Production Mix - IEA 2024 [24]

The *Kenya Vision 2030* [26] development program includes as one of its goals to reach 100% electrification status nationwide by 2030: to achieve so, mini-grids solutions have already been implemented, with the main regulating authority REREC (Rural Electrification and Renewable Energy Corporation) being active since the enactment of the Energy Act of 2019 [27]. Figure 4.3 illustrates the current status of national HV transmission lines, as of the latest report by KETRACO [28], as well as the mini-grids projects currently operating in otherwise unreached zones (whose sizing and location data were kindly provided by REREC): it is noticeable how most of the electrical infrastructure is developed in the southern and coastal regions, where most of the population resides, while the northern region is experiencing electricity access mostly due to off-grid systems.

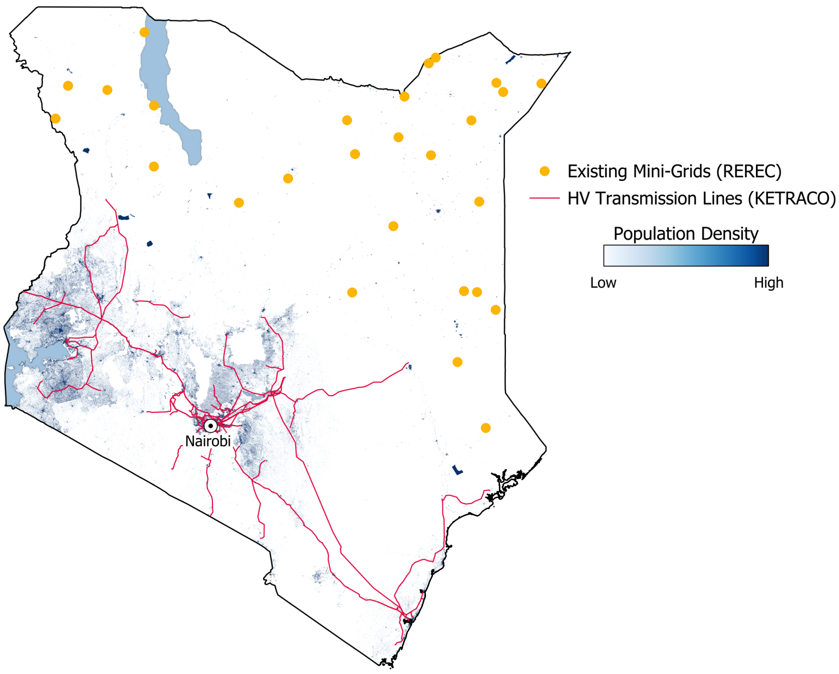


Figure 4.3: Current Electricity Infrastructure in Kenya

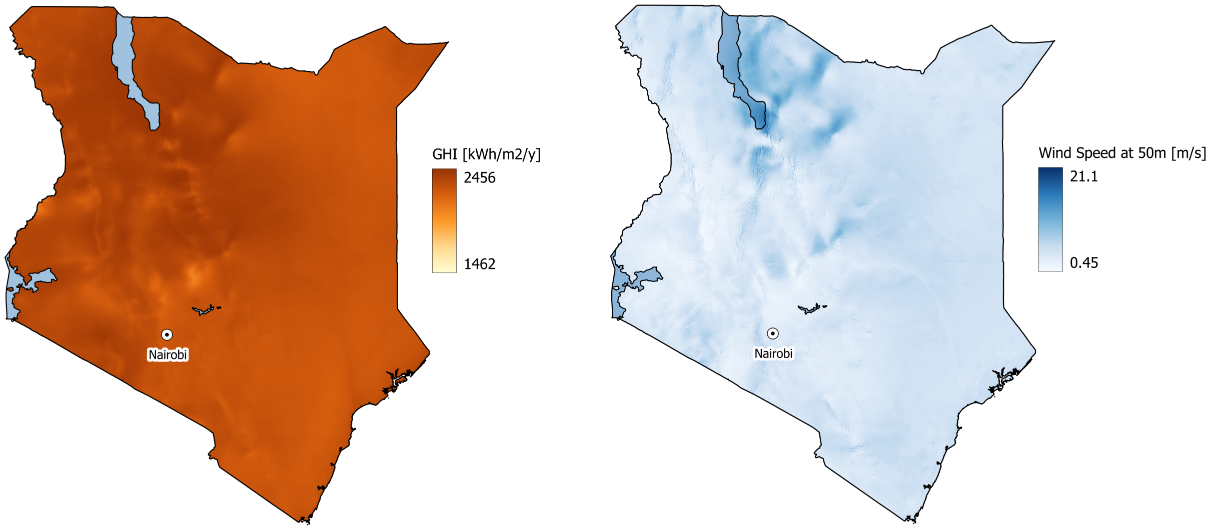
4.2 Renewable Potential

In the previous section, it has been mentioned how Kenya presents great renewable potential and a high degree of its utilization. While in remote areas SHS are a valid solution due to the all year availability of solar resource, the grid itself is dependent on large power plants based on renewable resources: exceptionally windy areas such as the Ngong Hills or the southern coasts of Lake Turkana are exploited with wind farms; important geothermal power plants are scattered all across the Rift Valley; solar fields make up a large portion of electricity generation as well, and hydro-electric power plants present a major production despite the frequent and harsh droughts that hit the Country.

The implemented mini-grids are shifting as of late from diesel-based to hybrid systems to both enhance the decentralization aspect and exploit RES. Indeed, in the economic analysis conducted using OnSSET, hybrid mini-grids are a prominent choice, as it will be evident from the results showcased in Chapter 5.

As mentioned in the mathematical formulation section, OnSSET requires time-series of renewables availability to size mini-grids for the look-up table, and these would be ideally required to represent the whole Country. The solar resource is more or less evenly distributed across Kenya, and therefore the country average GHI and temperature time-series available on *Renewables.ninja* [29] [30] are suitable to actually represent the Country. It is important to remember though that the construction of mini-grid solutions is still based

on a scaled version of the time-series over a range of country-specific GHI values, as explained in Paragraph 2.2.2.
A similar dataset is not available for the wind speed time-series: the chosen location is in the windiest area of Kenya, Lake Turkana, which is then scaled with set values from the wind-speed raster layer for the look-up table population.



(a) Global Horizontal Irradiation (b) Wind Speed at 50m

Figure 4.4: Renewable resources availability in Kenya

4.3 Geospatial Data Gathering

As previously mentioned, online repositories of geospatial data are often not up to date with recent developments of electrical infrastructure due to its fast evolving pace in developing economies such as Kenya. In order to avoid an outdated representation of the Country, companies such as REREC, KETRACO and KPLC were consulted to obtain recent updates for decentralized systems, transmission network and distribution grid respectively. Data regarding the distribution systems were kindly made available by Kenya Power in shapefile format: only the northern region was provided, with recent additions of power substations and medium voltage lines. KETRACO supplied an advance copy of the *Kenya Transmission Master Plan 2025-2044* [28], that reported the updated transmission network lines in image format; this required retracing the maps to include it in the QGIS datasets. Finally, REREC provided the list of state owned mini-grids and their details (generation sizing, location), which had to be manually pinpointed in QGIS.

4.4 Household Consumption Tiers

As anticipated in Paragraph 3.3.3, the yearly household consumption tiers classification is proposed to be country-dependent in order to reflect the modular demand constructed as input for MicroGridsPy. The values assigned to each tier for each category are reported in Table 4.1. Each reported value is computed as explained in Equation (3.3).

	AS	**AY**	**NC**	**OM**	**Average**
Tier 1	39,55	39,40	38,47	39,70	38,92
Tier 2	51,98	50,01	52,21	51,17	52,04
Tier 3	108,36	130,03	89,58	107,94	92,53
Tier 4	795,86	842,37	754,76	795,69	761,27
Tier 5	1355,73	1474,72	1230,75	1351,28	1249,56
Sett. Count*	0	10752	351568	41214	

Table 4.1: Electricity consumption by tier and cooling category [kWh/hh/y]
**Assuming a CDM threshold of 6°C*

Note that Kenya lies entirely within the F2 latitude band; if that were not the case, it would be advisable to also account for different latitude categorization in assigning consumption levels to tiers.

In this study, the chosen threshold for the activation of the cooling categories is CDM $\geq$ 6°C.

The *Custom Demand* layer is now fully defined, as these values are linked to the RWI bins, but OnSSET still requires an additional information in order to link each settlement to its specific archetype: since the core logic of the tool remains untouched, the wealth tier assignment once the average household consumption is still as described in Paragraph 3.3.3. The defining values of the bins are only five and not differentiated by cooling category or latitude, and the assignment logic has been modified to choose the closest tier of consumption: considering this, the values chosen to represent the tiers are the averages for each tier, weighted on the number of settlements assigned to the cooling categories. These values are reported in the final column of Table 4.1.

5 | Results

This chapter explores the outcomes of applying the proposed methodological framework to the Kenyan context. First, a direct comparison between the original OnSSET version and the proposed upgrade is reported; then, an in-depth analysis is conducted on a sample group of settlements, comparing mini-grid sizing and implementation costs resulting from the OnSSET tool and MicroGridsPy optimization; finally, this chapter discusses whether the archetypes integration improves the results' coherence with the ones obtained with the energy system optimization model.
The time-frame of the analysis is 21 years, starting from 2025 up to 2045. Please note that wind technologies were not included in the comparison, as they did not prove economically viable in any case within the OnSSET analysis. In future studies, it would be of interest to consider them in the technology mix in MicroGridsPy regardless of OnSSET results.

5.1 OnSSET results: a comparison between configurations

To fully grasp the differences between the original OnSSET version and the one proposed by this study, a first comparison is due. A direct comparison of the summary results from three separate configurations is reported:

- **MTF:** in this scenario, the tool is not modified and the *Custom Demand* is built with consumption tiers as proposed by the Multi-Tier Framework (MTF);
- **Custom Demand:** archetype curves are not introduced in the algorithm yet but the demand is constructed to be coherent with MicroGridsPy inputs, as described in Paragraph 4.4;
- **Archetypes:** archetypes are implemented in the tool to offer more representation of settlements; the input demand is the same as in the previous case.

As shown in Figure 5.1, an analysis done assuming a demand based on the MTF categorization results in greater installed capacity for relatively small differences in amount of new connections: this is due to the much greater energy intensity the MTF assumes, resulting in higher loads for a similar number of connections. It must also be considered that, as highlighted in Paragraph 2.1.1, the subdivision and associated consumption values proposed by ESMAP are not aligned with the M-LED based description of currently non-electrified, rural settlements, which are the primary focus of the tools employed in this study. The first logical step towards a consistent analysis with MicroGridsPy for

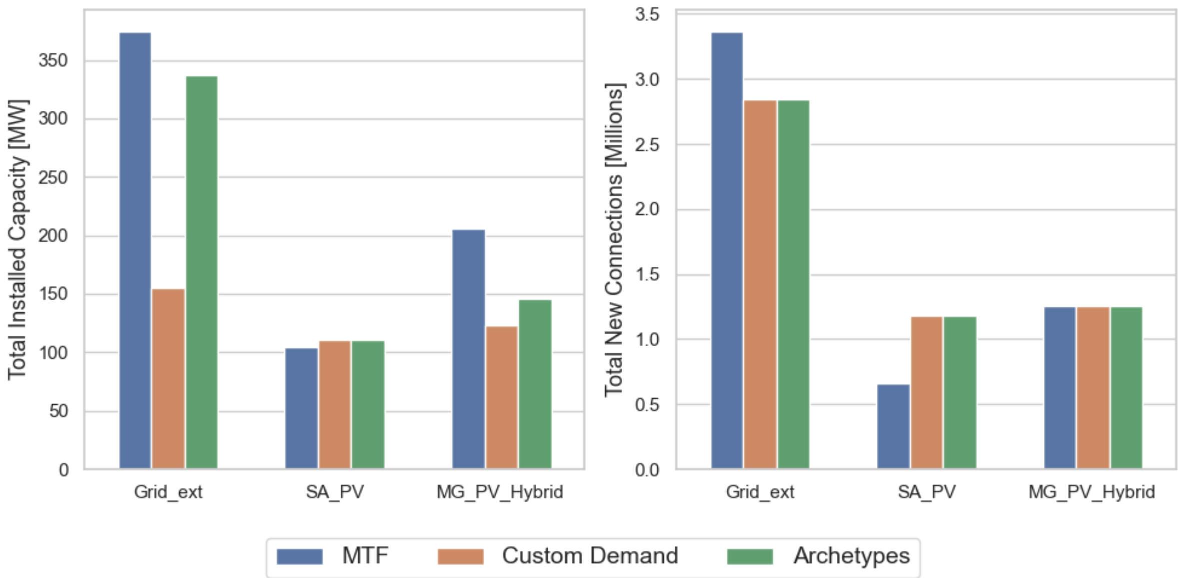


Figure 5.1: Comparison of Installed Capacity and New Connections in the three configurations

electrification of rural areas is therefore to refer to a comparable input demand. Notably, when comparing the other two configurations, the magnitude of their differences depends heavily on the electrification technology. Looking strictly at decentralized systems, the variations are slight but telling: the *Archetypes* case requires a greater installed capacity of hybrid systems to cover a comparable number of connections. The reason behind this result is explored in Paragraph 5.3.

Conversely, focusing on grid extension, it is noticeable how the capacity required by the *Archetypes* setup is almost comparable with the original *MTF* configuration, even if the number of new connections is actually lower: this is due to the updated values of *average to peak ratios* (see Paragraph 3.2.1), which are generally lower than the ones originally implemented in OnSSET therefore requiring greater new grid capacity to satisfy the connected settlements.

5.2 Energy System Optimization in Geospatially Significant Sites

As output for the case study, OnSSET found over 1'600 optimal mini-grid locations. The workflow of this analysis envisions using MicroGridsPy for energy system optimization in locations suggested by the geospatial tool. For the present study, as it focuses on proposing a complementary use of the two software, a limited amount of proposed locations is considered: for every county, two possible sites for hybrid mini-grids are analyzed, being the ones requiring the largest and smallest capacities of power generation. In some counties it emerged that it was either never optimal to install decentralized systems or just in one specific case, hence resulting in a pool of significant sites smaller than expected. As highlighted in Figure 5.2, the sites selected by OnSSET are concentrated in the northern counties, as well as in remote areas along the Tanzanian border; in central and coastal zones instead the existing grid is intensified and extended, so it often proves to be the most convenient choice.

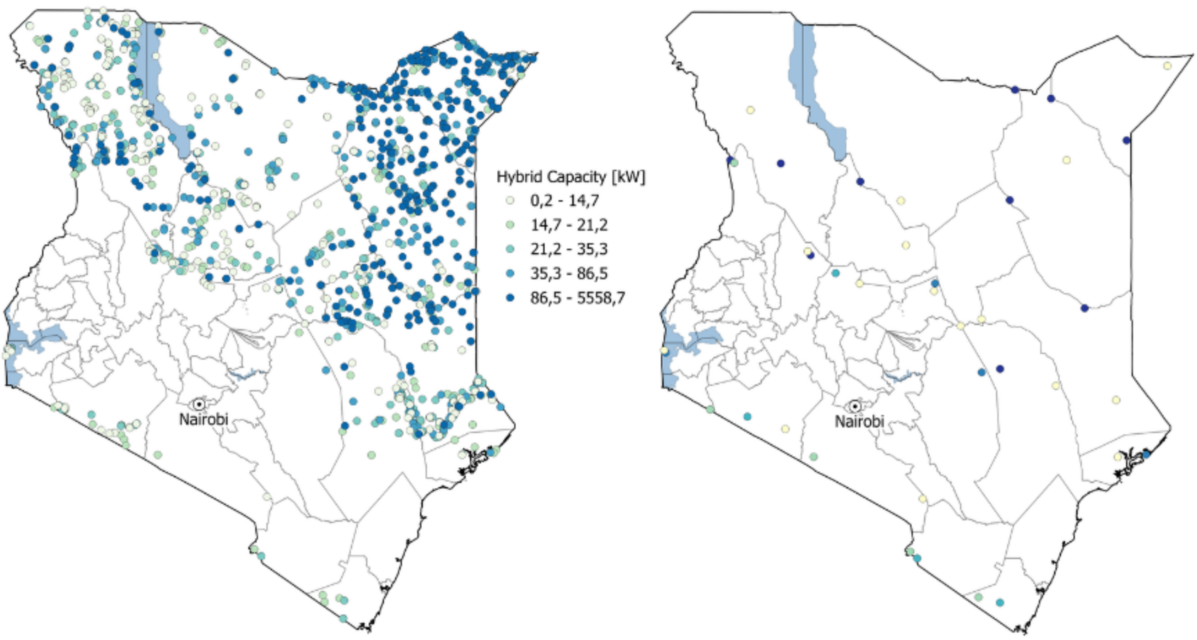


Figure 5.2: Optimal mini-grid sites as pinpointed by OnSSET and selected locations

It is worth noting that a few constraints were considered when choosing the sites of mini-grids:

- As mentioned, the main selection logic was based on the installed hybrid capacity in order to represent different sizes of settlements: the minimum and maximum sizes, by means of installed generation capacity, of mini-grid per each county was considered

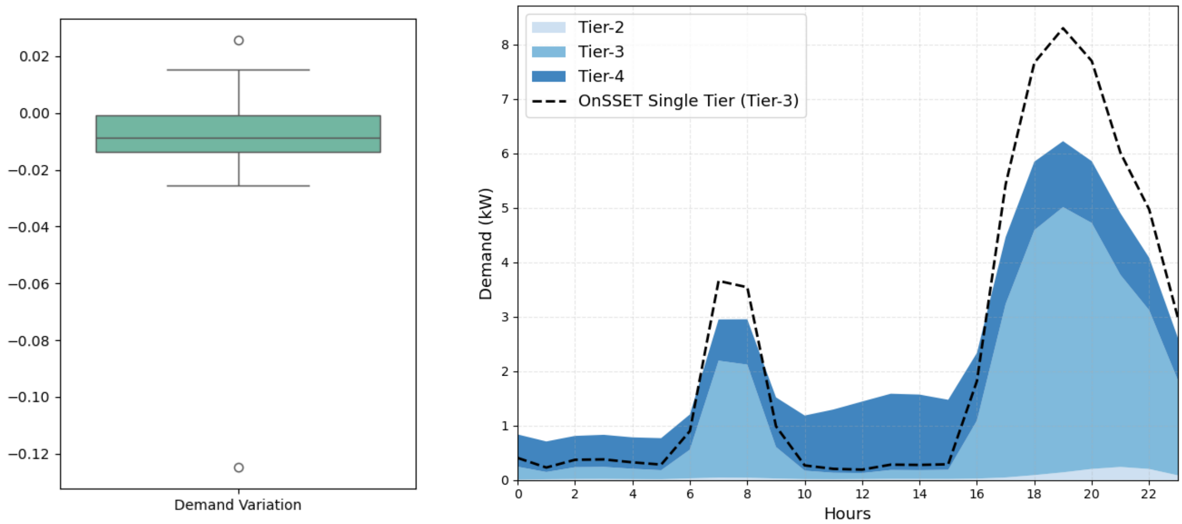
- Only settlements that are not connected to existing mini-grids were considered: OnSSET, while accounting for the connection of communities to existing mini-grids, does not offer any information regarding the electrification status or capacity required to cover the electrified settlement; due to the lack of this information and the difficulty of acquiring it, it was not possible to utilize the *Brownfield Mode* offered by MicroGridsPy, which allows to account for existing infrastructure in the given site

- The minimum capacity resolution considered by MicroGridsPy is 1 kW, so smaller settlements pinpointed by OnSSET had to be excluded in order to have meaningful results from the system optimization; the arbitrary threshold is a minimum final year demand of $1'000$ kWh/y, set after a first analysis of optimization results

5.2.1 Coherence of load curves

In Chapter 3 one of the main topics discussed was the coherence of residential demand representation between the two tools. The proposed method was a Gaussian distribution of households in MicroGridsPy, calibrated so that the average household consumption corresponded to the one accounted for by OnSSET. This method proved to be successful when comparing both the total settlement demand in the final year of the analysis and the load curves resulting from the two software. In Figure 5.3a the reduced variation in final yearly demand is highlighted, calculated as the divergence of MicroGridsPy's demand from the one computed by OnSSET: only a few outliers are present, one of which of relatively small significance; the other, presenting a -12% variation in demand assessment, has been excluded from the subsequent result comparison.

In most cases, the shape of the load curve as considered by MicroGridsPy via the household distribution balances the differences during peak hours with OnSSET by distributing the load during the day, eventually resulting in a net final yearly demand that is coherent with the geospatial tool, as illustrated for a reference settlement in Figure 5.3b.
In the eventuality that a settlement is classified in OnSSET as Tier 4, due to the general shape associated to this consumption level the evening peaks are severely underestimated compared to MicroGridsPy: the latter considers demand curves also linked with other tiers, resulting in a greater load in hours of absence of solar availability while still maintaining the yearly demand at OnSSET levels (Figure 5.4). Such cases generally result in a much greater installed storage capacity to cover for night-time demand, therefore



(a) Demand deviation in MicroGridsPy from OnSSET.

(b) Comparison of OnSSET and MicroGridsPy load curves representing the same settlement.

Figure 5.3: Analysis of demand deviations and load curve comparisons between OnSSET and MicroGridsPy.

requiring greater investments as well compared to OnSSET. Indeed, in the case reported in Figure 5.4 it is evident how the evening load as considered by the geospatial tool is lower than MicroGridsPy's demand by 200 kW on average, hence causing a drastic surge in the storage and back-up systems capacities installed when optimizing the settlement with the ESOM.

Since in this analysis only two settlements are represented by this specific consumption tier, it is not possible to establish a major trend in this regard, hence these settlements are treated as outliers.

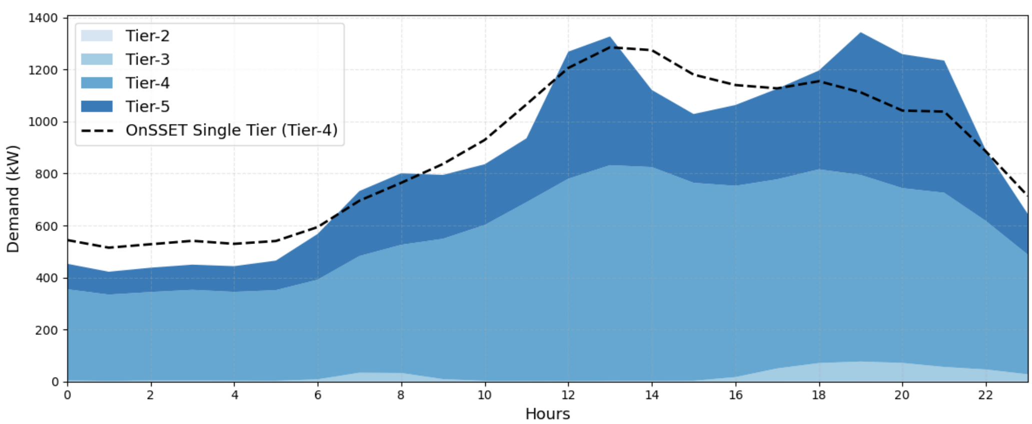


Figure 5.4: Tier 4 settlement load curve comparison

A similar reasoning applies to settlements assigned to Tier 2, which also represent only two cases in the analysis pool. However, a significant difference emerges: MicroGridsPy demand curves for these settlements do not deviate from OnSSET's curves as sharply as in the Tier 4 case (Figure 5.5). This is because the weight of Tier 3 households in a Tier 2 settlement is less disruptive than the impact of Tier 5 households in a Tier 4 setting. Regardless, every technology shows at least a 50% increase in installed capacity, which does not align with the expected proportion given the minimal variation in the demand curve; this makes OnSSET's optimization appear more efficient. This deviation is partially due to how OnSSET defines and rescales the diesel generator in low-demand settlements, often resulting in fractional values lower than unity that MicroGridsPy cannot replicate, given its minimum unit capacity of 1 kW; more importantly though, it is driven by the dispatch strategy the GEM employs. OnSSET considers a parameter called *Loss of Power Supply Probability* (LPSP), representing the allowed annual unmet demand as a percentage share of total load (10%). For small settlements characterized by low base loads and high night-time peaks, this parameter can be exploited to effectively "cut" the peak during critical days with low solar irradiation. This allows the diesel generator to be sized significantly more leniently than in the stochastic ESOM. In the following sections, all considerations will be made mainly regarding only Tier 3 to highlight possible trends in a more easily accessible manner; exhaustive results including outliers are reported in *Annex B*.

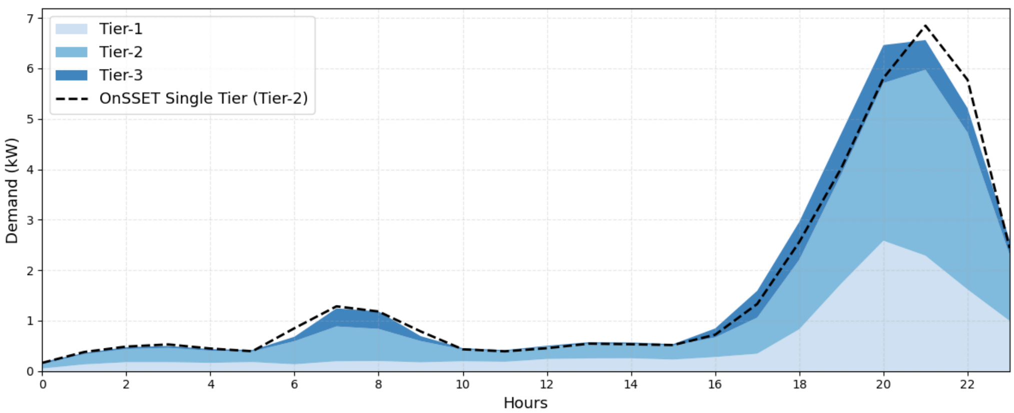


Figure 5.5: Tier 2 settlement load curve comparison

As a final consideration regarding the load curves, it is worth highlighting the differences in yearly household consumptions between OnSSET and MicroGridsPy when using a mono-tier representation in the latter: this issue has been introduced in Paragraph 3.3.4, in which it was proposed the use of a gaussian distribution to overcome it. Figure 5.6 reports the variation observed in the case study, representative of the above described discrepancy. The variation is computed as follows:

$$\Delta consumption = \frac{D_{hh,monotier} - D_{hh,custdem}}{D_{hh,custdem}} \tag{5.1}$$

- $D_{hh,monotier}$: yearly household demand associated to the single archetype
- $D_{hh,custdem}$: yearly average household demand as computed for OnSSET input
- $\Delta consumption$: difference in household consumption

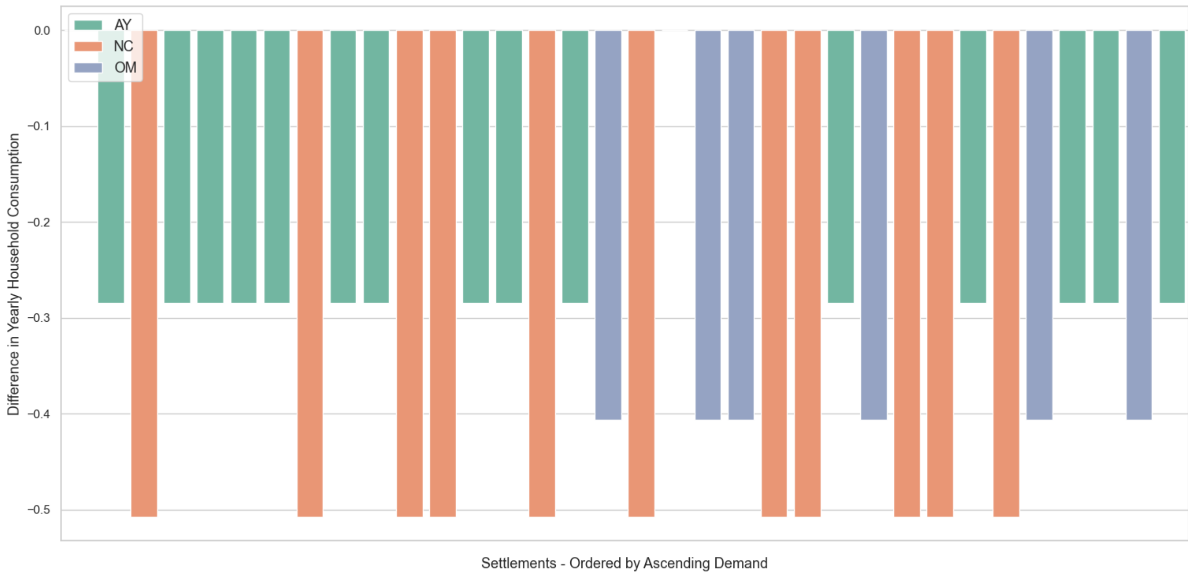


Figure 5.6: Household consumption difference when using the mono-tier representation in MicroGridsPy

It is noticeable how the mono-tier representation consistently underestimates the demand considered by the geospatial electrification tool, for which the input demand considers instead different wealth levels within the settlement's boundaries. Moreover, it is evident how this issue is more evident when a settlement is assigned either OM or NC as cooling category, as the differences between the settlement consumption considered by OnSSET and the one linked to the specific tier become more evident.

5.2.2 Results comparison

In the present study a pool of 39 sites pinpointed by OnSSET is considered: for each, an individual system optimization with location specific parameters and timeseries was obtained with MicroGridsPy. The comparison between the system optimization and the solution of the GEM tool is based on economic parameters, being the LCOE and the

investment cost associated with generation only, and sizing results.
Please note that all deviations reported in the following analysis are computed to highlight how MicroGridsPy's optimization diverges from OnSSET's initial estimations. The variations are calculated as follows:

$$\Delta X = \frac{X_{mgpy} - X_{onsset}}{X_{onsset}} \tag{5.2}$$

Where X_i is the value of the considered parameter.
From a preliminary analysis, an apparent contradiction becomes evident: as shown in Figure 5.7, which reports the deviation of MicroGridsPy's results from OnSSET's, the system optimization tool shows a systematic decrease in LCOE (-13.79% on average) while simultaneously having considerably higher investment costs across almost all settlements ($+29.51\%$ on average). This result may seem counter-intuitive as the two parameters are expected to be closely correlated, but it must be noted that the LCOE computation considers both investment and operation and maintenance costs. This divergence actually highlights the strength of a dedicated optimization model that operates on a single settlement: because MicroGridsPy foresees the future benefit of heavily investing in renewable energies rather than relying on a diesel generator, the higher investment cost is more than compensated by the cumulative OPEX reduction over the 20-year time frame.

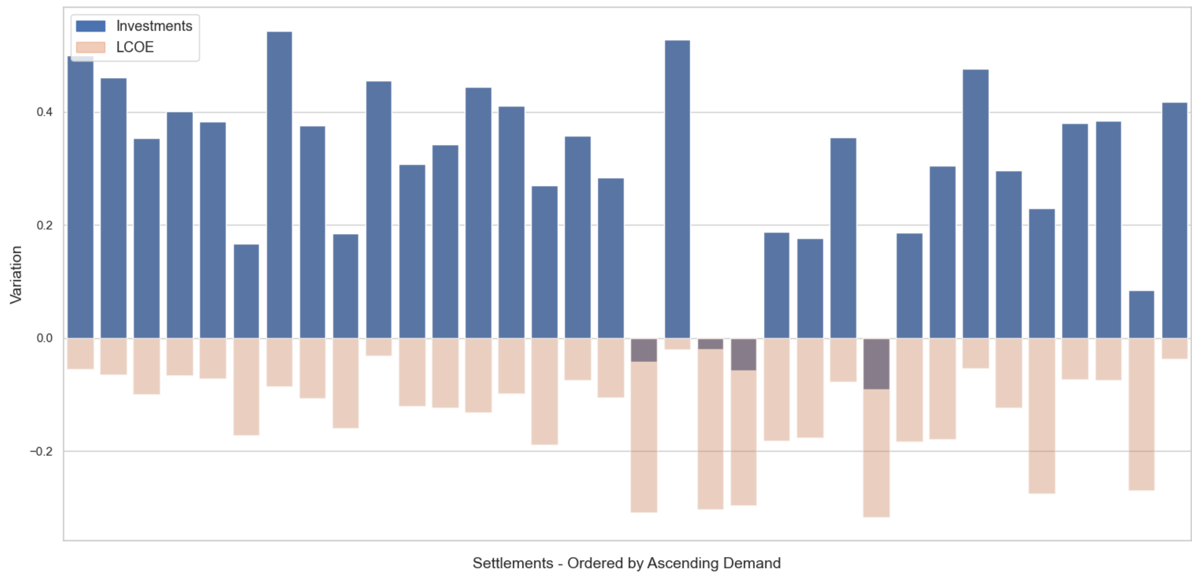


Figure 5.7: Investment Costs and LCOE variations between MicroGridsPy and OnSSET

Moreover, concerning the LCOE estimation, in Figure 5.8 it is emphasized that not only MicroGridsPy obtains a steadily lower LCOE for nearly all analyzed settlements, but the range of their values is considerably more restricted; this highlights the higher consistency provided by a dedicated optimization model if compared to OnSSET's pre-calculated sizing allocation based on differential evolution and employment of the look-up tables.

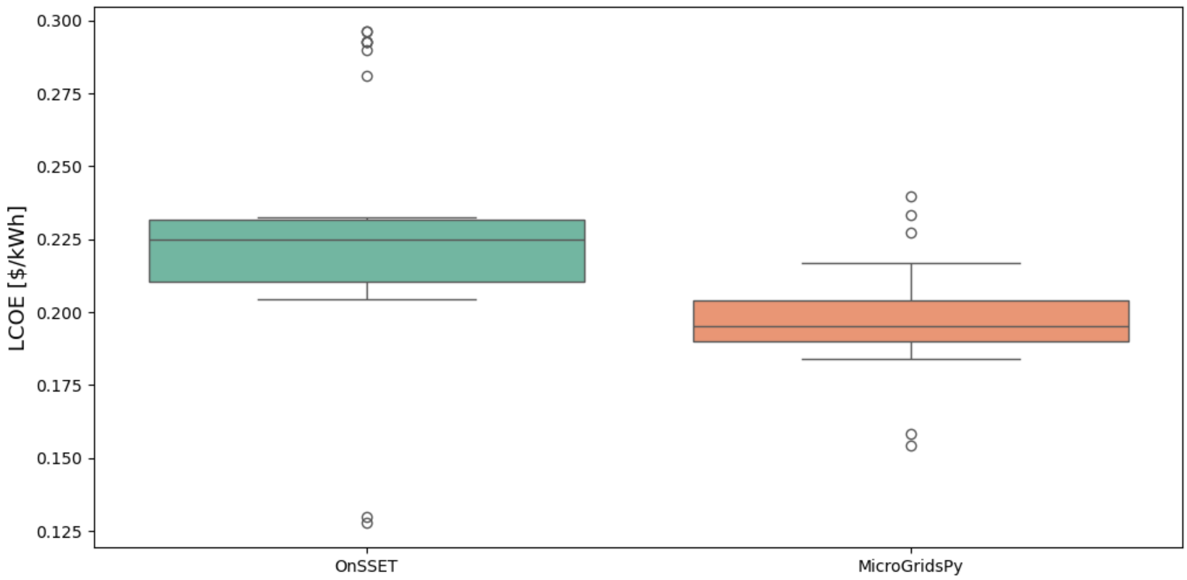


Figure 5.8: LCOE computed by OnSSET and MicroGridsPy

These results pinpoint a criticality of the OnSSET model, being that it consistently overestimates the LCOE of settlements when employing mini-grids, hence resulting in them being less economically appealing; for a tool such as OnSSET, which bases the electrification planning upon the most convenient choice using the LCOE as decisive parameter, this overestimation is of critical interest. A potential synergic integration between the two software could be using MicroGridsPy to offset OnSSET's LCOE to make it more accurate; this could be accomplished by performing an analysis on a number of sample settlements in a similar fashion to the one proposed in this thesis, then it could be theoretically possible to calculate a correction factor and use it to adjust OnSSET's results. Further considerations might be required, such as how the size of the settlement influences the LCOE gap between the models, but that goes beyond the scope of this work.

This economic behavior is directly explained by the differences in system sizing between the two models. When analyzing the overall installed capacities, a distinct trend emerges: MicroGridsPy's optimization is heavily skewed toward oversizing the solar generation. On average, it installs 43% more PV capacity compared to OnSSET. Interestingly, the overall average capacities for backup and storage remain unexpectedly stable, with the diesel generator showing a mere 5% increase and the BESS capacity showing a 0.35% difference. This massive expansion of the PV array alone justifies the spike in initial investment costs, while the abundance of excess solar generation drastically reduces fuel

consumption over the project's lifetime, thereby lowering the LCOE.
However, looking only at these overall averages for backup and storage hides a much more complex, context-dependent reality. Based on the results presented thus far, it might seem that a key difference between the two models is that the ESOM seems to favor renewable resources exploitation while OnSSET relies more heavily on the diesel generator. This statement, while accurate for most settlements, proves to be fairly reductive: prioritizing RES is the optimal solution in most cases, but some specific circumstances may allow for completely different approaches that flip the standard. The prime example of this behavior is shown in the impact of the cooling category on the installation preferences between the diesel generator and BESS; from the Figures 5.9 - 5.10 it is evident that the *OM* category has a counter-tendency compared to the other categories. The latter show a 4% average decrease in diesel generator capacity and a 6% average increase in BESS capacity, that perfectly matches the aforementioned trend of RES reliance of MicroGridsPy; on the other hand the settlements with cooling period being *October to March* heavily prioritize the diesel generator and greatly reduce BESS installation. This divergence perfectly explains the unexpectedly stable averages of the backup and storage systems mentioned earlier. The strong counter-tendency of the *OM* settlements effectively offsets the capacity shifts seen in the remaining categories, making the overall average variations appear much closer to zero than the individual trends would suggest.

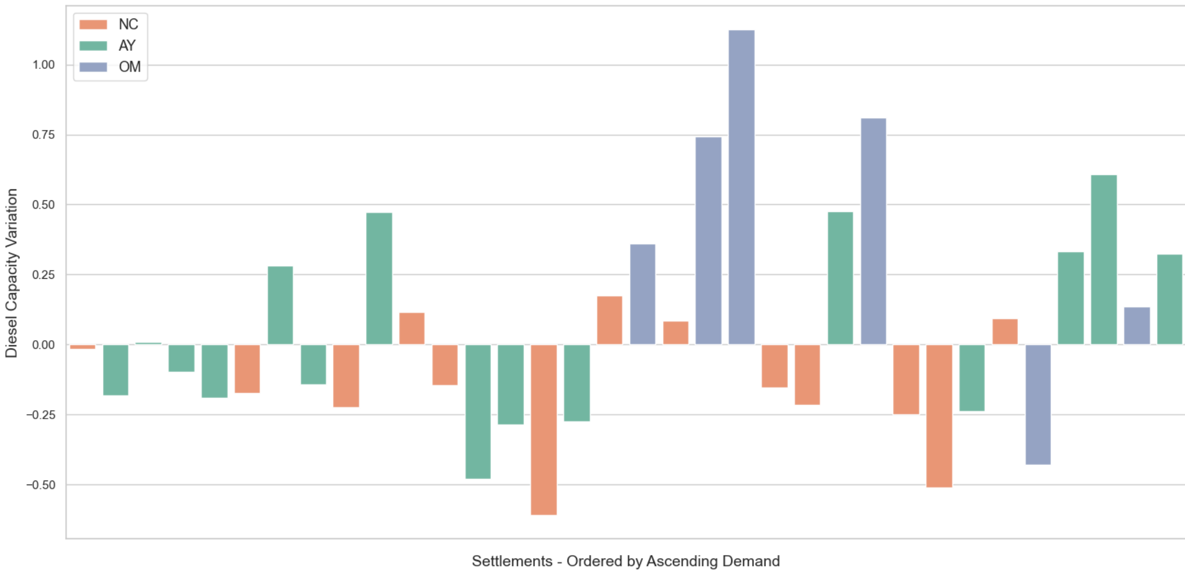


Figure 5.9: Diesel capacity variation between MicroGridsPy and OnSSET

The logic behind this choice comes from the intermittence of *OM* cooling (it is plausible to be the same for the *AS* cooling period, but it doesn't show in this analysis): having active cooling only for half of the year practically means that the model needs to satisfy

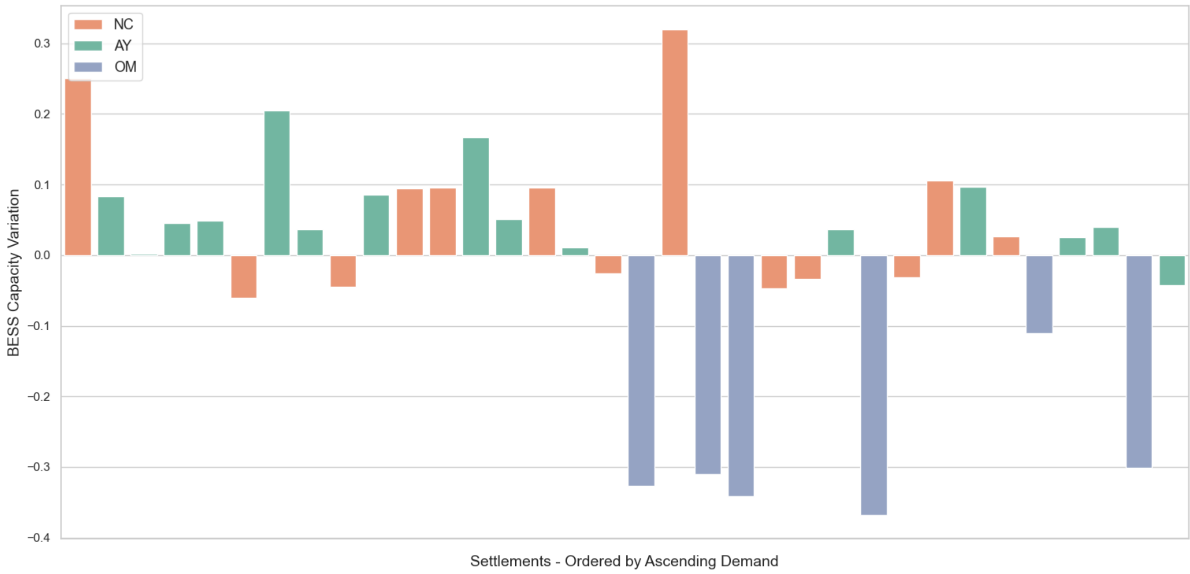


Figure 5.10: BESS capacity variation between MicroGridsPy and OnSSET

two different demand curves, since the impact of cooling is far from negligible; so while OnSSET is able to cover the demand increase in the October to March range, in doing so it ends up being over-dimensioned for the remaining months. MicroGridsPy, on the other hand, adapts to the situation at hand by recognizing the wasted BESS capacity and opting for a more diesel-heavy alternative. On a more technical note, this concept can be visualized by plotting the LCOE behavior versus equivalent operating hours as shown in Figure 5.11; from this graph it is possible to infer that while batteries generally show a lower LCOE, that is not the case for an especially short segment of low equivalent hours, where the diesel generator becomes the optimal choice. It must be noted that this graph completely ignores the PV modules and their influence on the resulting LCOE. This distinction is highly relevant, as MicroGridsPy evaluates the dynamic interactions among all three technologies and incorporates them directly into its optimization logic. Therefore, while this graph is not a perfect representation of the model's inner workings, it still provides a useful depiction of the general LCOE trend. From these results, it can be inferred that accounting for all system parameters and their cascading effects ultimately points toward a balanced integration of the three technologies as the optimal configuration.

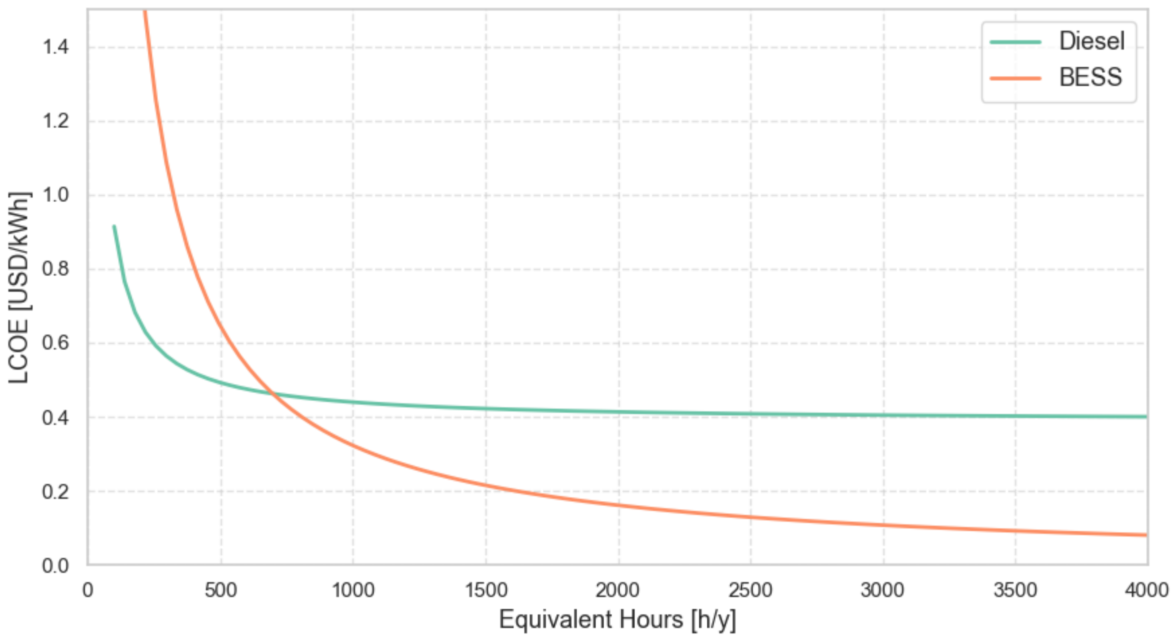


Figure 5.11: Diesel and battery storage system comparison based on equivalent hours

5.3 Influence of Archetypes introduction on soft-linking

The introduction of archetype load curves in OnSSET produced different outcomes compared to the version using the native load profiles; this implementation still showed substantial differences from individual system optimizations with a dedicated tool, although presenting the trends as analyzed in the previous section. As a final part of this study, a direct comparison was conducted between the results offered by the original OnSSET configuration with the native curves (*Custom Demand*) and by the modified version (*Archetypes*) with the optimizations obtained from the use of MicroGridsPy. The same sample of settlements is analyzed, excluding the mentioned outliers for coherence; the representation of the results with no exclusions is reported in *Annex B*.
Please note all data reported in the analyzed plots are computed as follows:

$$\Delta X = \frac{X_{onsset} - X_{mgpy}}{X_{mgpy}} \tag{5.3}$$

Where X_i is the value of the considered parameter.
When comparing the divergence from MicroGridsPy's results, it is useful to keep in mind the different load profiles the geospatial tool considers: typically archetype profiles present a lower *average to peak ratio*, showing a reduced load during daytime and accounting for greater peaks in evening hours when compared to native curves (Figure 5.12).

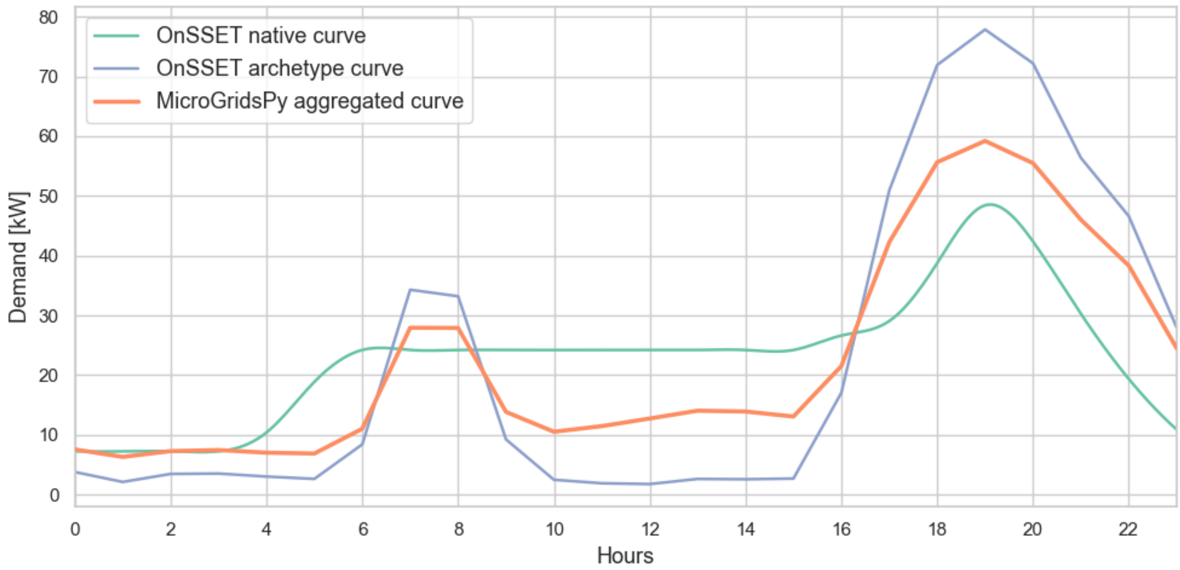


Figure 5.12: Comparison of load curves for a given Tier 3 settlement

A first striking contrast between the two OnSSET configurations is the resulting LCOE: as discussed in the previous paragraph, the *Archetypes* version showed a steady overestimation of the LCOE compared to the MicroGridsPy optimum. As illustrated in Figure 5.13, the use of native curves instead presented the opposite trend, that is a constant underestimation of the parameter: it results in the possible localization of non-optimal mini-grids sites, where instead it could be more convenient to either extend the grid or implement another technology. The increase of the LCOE with archetypes integration is an interesting but expected outcome, as it is due to the greater disparity in load distribution in archetype curves (shown in Figure 5.12) which results in a greater installed capacity for a similar daily demand. It must be highlighted that, while the inversion of the relation is evident, the actual result is that the magnitude of the divergence from MicroGridsPy's calculations is lower when implementing the new curves.

The increased coherence in LCOE estimation is a consequence of the installed capacity of both generation and storage technologies: as pictured in Figures 5.14 - 5.15, the use of native curves resulted in a severe underestimation of required capacity compared to the energy system optimization model. Regarding photovoltaics, a visible improvement is observed in *Archetypes* even if in most cases the capacity is still underestimated due to lower diurnal loads; considering the storage system estimation, the reduction of divergence is extremely evident, as the evening hours are similarly represented both in peaks and load distribution when compared to MicroGridsPy.
A further consideration is worth to be made considering the trends presented by the

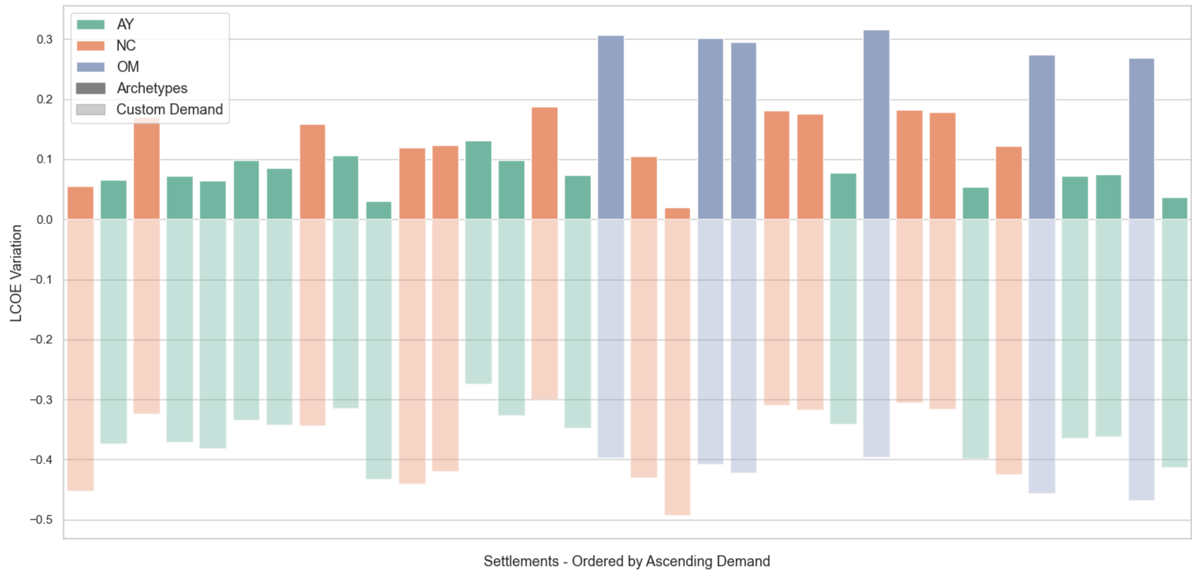


Figure 5.13: LCOE difference between OnSSET versions and MicroGridsPy

settlements assigned the *OM* cooling category. When OnSSET bases its optimization on the native load curves, no seasonality is accounted for as the same typical day profile is repeated throughout all year; with the introduction of archetypes it is instead possible to represent it, affecting both generation and storage capacities as analyzed in Paragraph 5.2.2.

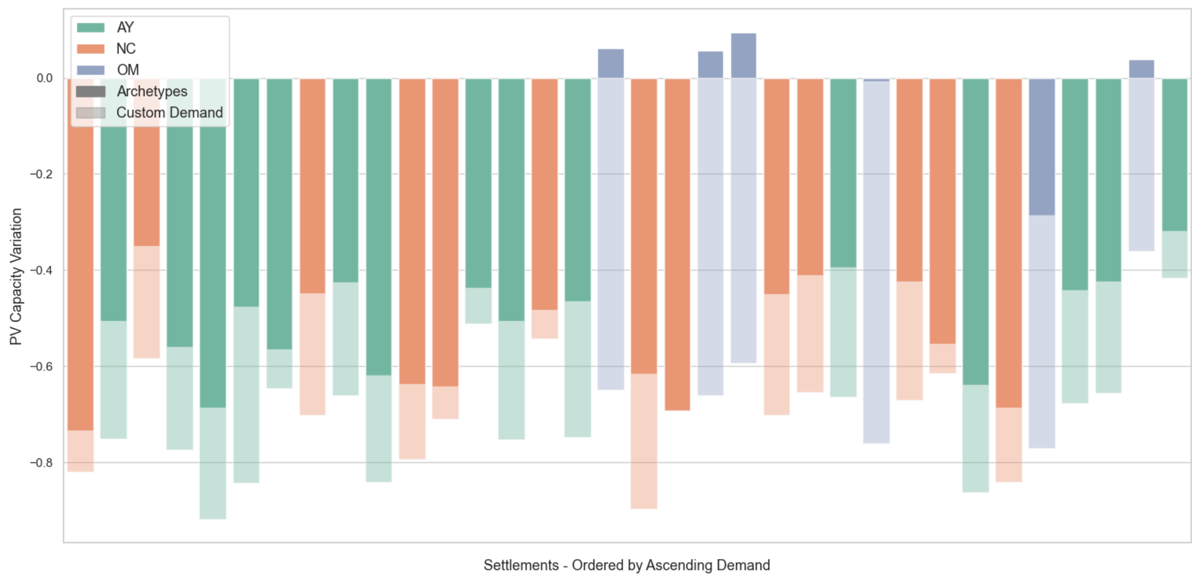


Figure 5.14: PV capacity difference between OnSSET versions and MicroGridsPy

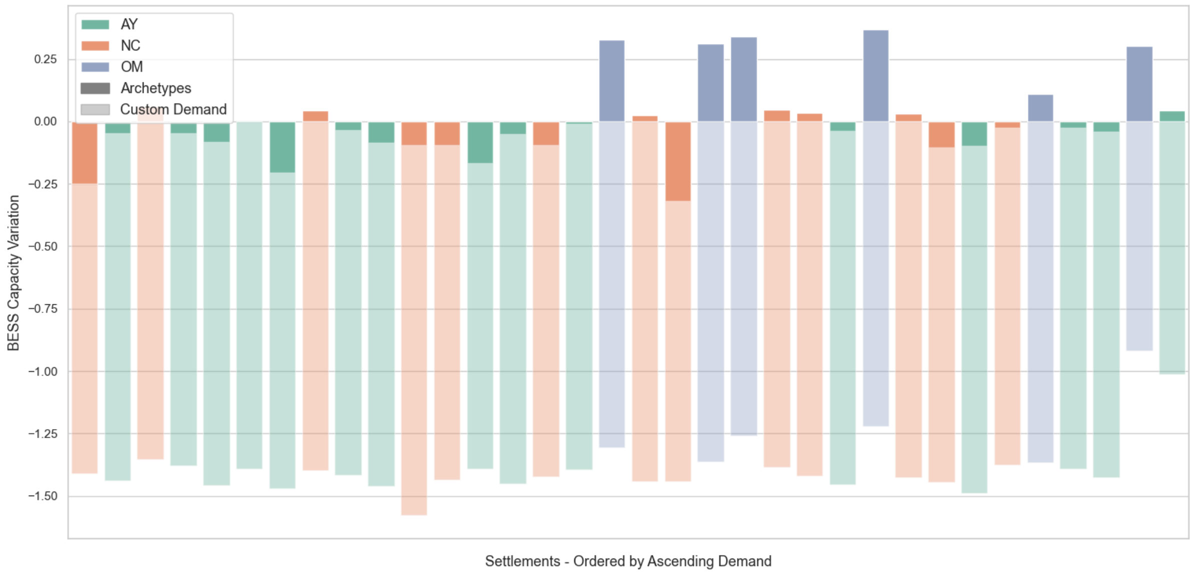


Figure 5.15: BESS capacity difference between OnSSET versions and MicroGridsPy

6 | Conclusions and future developments

The present study showed that a synergistic and complementary use of MicroGridsPy and OnSSET is possible, and the proposed methodological framework has been proven to enhance the coherence of results: the introduction in OnSSET of a greater resolution in demand definition allowed for an improvement in the hybrid system modeling, resulting in outcomes more consistent with the dedicated energy system optimization model.
Although the introduction of demand archetypes enabled a consistent analysis with MicroGridsPy, the predefined dispatch strategy and, more importantly, the static optimization which characterize OnSSET still hinder mini-grid modeling. Methodologically, the introduction of a calibrated Gaussian distribution to represent diverse household wealth levels within MicroGridsPy while maintaining input consistency with the GEM proved highly successful. This approach highlighted the limitations of OnSSET's single tier settlement classification compared to the complex reality of rural communities. However, it must be noted that energy system optimization is not the main goal of the tool, whose primary objective is instead to find the optimal electrification strategy on a national level. Therefore, as illustrated during this analysis, a sequential use of the software is highly recommended: utilizing OnSSET for the identification of suitable sites for mini-grid implementation, followed by MicroGridsPy for the optimization of the system in those locations.

As discussed, OnSSET steadily overestimates the LCOE of mini-grids compared to MicroGridsPy, hence possibly assuming other technologies to be more cost-effective in given locations. The development and introduction of an LCOE correction factor in the OnSSET algorithm could overcome this issue, delivering results that are more coherent with actual optimal energy systems.
Regarding system sizing, the optimization model consistently favors capital intensive renewable technologies, oversizing PV arrays to drastically reduce lifetime fuel consumption and OPEX. However, this study highlighted that contextual nuances, specifically intermittent seasonal cooling demands like the *October to March* category, can completely invert this trend, forcing a reliance on diesel generation to avoid battery underutilization.

Furthermore, the impact of the archetypes extended beyond decentralized systems: their lower average to peak ratios significantly increased the generation capacity required for national grid extension.

It is important to note that this research is only a step in increasing demand resolution in the geospatial software, as important loads such as health centers and educational facilities, and other possible productive uses of electricity, are not accounted for in this work. Future developments may also include a more in-depth analysis of possible trends across different settlement consumption tiers. In this case study, only communities to which OnSSET assigned Tier 3 were analyzed; however, since this represents the average settlement wealth level, it is arguably the tier of greatest interest. Moreover, it should be noted that in this case study, no settlement was classified in the *April to September* cooling category: while it is plausible that it presents the same characteristics as the *October to March* classification, it could be of interest to analyze how it influences the results.

It must be emphasized that multiple functions offered by MicroGridsPy could not be replicated in the OnSSET analysis, most notably the use of *Brownfield Mode*. This is of particular interest as it allows the model to account for already existing mini-grid capacity in a given settlement, and while it is possible to include implemented systems in the geospatial analysis there is no accounting for actual capacities. Since Kenya and other Sub-Saharan countries already have numerous active mini-grids, for the joint use of the two tools it would be of interest to introduce such considerations in OnSSET.

Finally, for simplicity, this study excluded the possible connection of the decentralized systems to the grid, but it could be a topic of interest for future studies.

Bibliography

[1] World Bank Group. World Bank Data, 2023.

[2] Homer Energy Pro - Microgrid Optimization Software.

[3] N. Stevanato, G. Pellecchia, I. Sangiorgio, D. Shendrikova, C. A. Soares, R. Mereu, and E. Colombo. Planning third generation minigrids: Multi-objective optimization and brownfield investment approaches in modelling village-scale on-grid and off-grid energy systems. *Renewable and Sustainable Energy Transition*, 2023.

[4] A. Onori. MicroGridsPy Planning - GitHub Repository, 2026.

[5] A. Sahlberg, A. Korkovelos, and J. Cantor. OnSSET for MiniGrids - GitHub Repository.

[6] A. Onori. Fit For Purpose: comparing open-source mini-grid modeling tools. Knowledge Brief, 2026.

[7] Dimitrios Mentis, Mark Howells, Holger Rogner, Alexandros Korkovelos, Christopher Arderne, Eduardo Zepeda, Shahid Siyal, Costantinos Taliotis, Morgan Bazilian, Ad de Roo, Yann Tanvez, Alexandre Oudalov, and Ernst Scholtz. Lighting the world: the first application of an open source, spatial electrification tool (onsset) on sub-saharan africa. *Environmental Research Letters*, 12(8):085003, jul 2017.

[8] Silvia Corigliano, Tommaso Carnovali, Darlain Edeme, and Marco Merlo. Holistic geospatial data-based procedure for electric network design and least-cost energy strategy. *Energy for Sustainable Development*, 58:1–15, 2020.

[9] World Bank Group and ESMAP. Global Electrification Platform, 2025.

[10] Nandi Moksnes, Alexandros Korkovelos, Dimitrios Mentis, and Mark Howells. Electrification pathways for Kenya–linking spatial electrification analysis and medium to long term energy planning. *Environmental Research Letters*, 12(9):095008, sep 2017.

[11] Mark Howells, Holger Rogner, Neil Strachan, Charles Heaps, Hillard Huntington, Socrates Kypreos, Alison Hughes, Semida Silveira, Joe DeCarolis, Morgan Bazillian,

and Alexander Roehrl. OSeMOSYS: The Open Source Energy Modeling System: An introduction to its ethos, structure and development. *Energy Policy*, 39(10):5850–5870, 2011. Sustainability of biofuels.

[12] Ioannis Pappis, Andreas Sahlberg, Tewodros Walle, Oliver Broad, Elusiyan Eludoyin, Mark Howells, and Will Usher. 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(4), 2021.

[13] B. F. Silinto, D. Edeme, S. Corigliano, A. Dimovski, M. Merlo, C. Zuidema, and A. Faaij. An integrated geospatial modelling framework of microgrid sizing for rural electrification planning. *MethodsX*, 2025.

[14] Francesco Lombardi, Sergio Balderrama, Sylvain Quoilin, and Emanuela Colombo. Generating high-resolution multi-energy load profiles for remote areas with an open-source stochastic model. *Energy*, 177:433–444, 2019.

[15] J.G. Peña Balderrama, S. Balderrama Subieta, F. Lombardi, N. Stevanato, A. Sahlberg, M. Howells, E. Colombo, and S. Quoilin. Incorporating high-resolution demand and techno-economic optimization to evaluate micro-grids into the open source spatial electrification tool (OnSSET). *Energy for Sustainable Development*, pages 98–118, 2020.

[16] N. Stevanato, I. Sangiorgio, R. Mereu, and E. Colombo. Archetypes of rural users in sub-saharan africa for load demand estimation. *IEEE PES/IAS Power Africa*, 2023.

[17] A. Onori. MicroGridsPy Development Linopy Version, 2025.

[18] G. Falchetta, N. Stevanato, M. Moner-Girona, D. Mazzoni, E. Colombo, and M. Hafner. The M-LED platform: advancing electricity demand assessment for communities living in energy poverty. *Environmental Research Letters*, 2021.

[19] ESMAP. Beyond Connections - Energy Access Redefined, 2015.

[20] Giacomo Falchetta and Malcolm N. Mistry. The role of residential air circulation and cooling demand for electrification planning: Implications of climate change in sub-saharan africa. *Energy Economics*, 99:105307, 2021.

[21] A. Sahlberg, B. Khavari, A. Korkovelos, W. Usher, D. Mentis, and C. Arderne. OnSSET - Online Course Version.

[22] A. Sahlberg. GIS extraction notebook, 2025.

[23] IEA. Kenya 2024 - Energy Policy Review. Technical report, International Energy Agency, 2024.

[24] IEA. International Energy Agency - Kenya, 2025.

[25] EPRA. Least Cost Power Development Plan. Technical report, Ministry of Energy and Petroleum, 2024.

[26] Government of Kenya. Kenya Vision 2030, 2008.

[27] Government of Kenya. The Energy Act, 2019.

[28] KETRACO. Transmission Master Plan 2025-2044, 2025.

[29] S. Pfenninger and I. Staffel. Long-term patterns of European PV output using 30 years of validated hourly reanalysis and satellite data. *Energy*, (114):1251–1265, 2016.

[30] S. Pfenninger and I. Staffel. Using bias-corrected reanalysis to simulate current and future wind power output. *Energy*, (114):1224–1239, 2016.

[31] World Bank Group. Kenya - Global Electrification Platform, 2025.

[32] Google. Administrative Boundaries, 2025.

[33] Kenya Power and Lightning Company. Kenya Electricity Network, 2025.

[34] Kenya Power and Lightning Company. Kenya Primary Substations, 2025.

[35] Kenya Roads Board. Kenya - Roads, 2023.

[36] ICPAC. Kenya Night Time light, 2025.

[37] Solargis s.r.o. Global Solar Atlas, 2025. Global Solar Atlas 2.0, a free, web-based application developed and operated by the company Solargis s.r.o. on behalf of the World Bank Group, utilizing Solargis data, with funding provided by the Energy Sector Management Assistance Program (ESMAP).

[38] Technical University of Denmark. Global Wind Atlas, 2025. Global Wind Atlas version 4.0, a free, web-based application developed, owned and operated by the Technical University of Denmark (DTU). The Global Wind Atlas version 4.0 is released in partnership with the World Bank Group, utilizing data provided by Vortex, using funding provided by the Energy Sector Management Assistance Program (ESMAP).

[39] N. N Davis, J. Badger, A. N. Hahmann, B. O. Hansen, N. G. Mortensen, M. Kelly, X. G. Larsén, B. T. Olsen, R. Floors, G. Lizcano, P. Casso, O. Lacave, A. Bosch, I. Bauwens, O. J. Knight, A. Potter van Loon, R. Fox, T. Parvanyan, S. B. K. Hansen,

D. Heathfield, M. Onninen, and R. Drummond. The Global Wind Atlas: A high-resolution dataset of climatologies and associated web-based application. *Bulletin of the American Meteorological Society*, 2023.

[40] A. Korkovelos. Sub Saharian Africa - Small and mini hydropower potential, 2017.

[41] HarvestChoice and International Food Policy Research Institute (IFPRI). Travel Time to Markets in Africa South of the Sahara, 2018.

[42] Kenya Power and Lightning Company. Kenya - Distribution Transformers, 2025.

[43] Meta. Relative Wealth Index, 2021.

[44] Government of Kenya. Kenya National Bureau of Statistics, 2026.

[45] Ministry of Energy and Petroleum (EPRA). Kenya Energy Compact, 2025.

[46] Ministry of Energy and Petroleum (EPRA). Biannual Statistics Report, 2023.

[47] Global Petrol Prices, 2023.

[48] World Bank Group. Mini Grids for Half a Billion People, 2022.

[49] International Renewable Energy Agency (IRENA). Renewable Power Generation Costs, 2024.

A | Appendix A

Layer	Type	Source	Year
Population Clusters	Polygon	EnergyData [31]	2025
Administrative Boundaries	Polygon	GADM [32]	2020
Existing HV Network	Line	EnergyData [33] and KETRACO	2025
Existing MV Network	Line	EnergyData [33] and KETRACO	2025
Power Substation	Point	EnergyData [34] and KPLC	2025
Road Network	Line	EnergyData [35]	2023
Night Time Lights	Raster	IGAD Climate Prediction and Applications Center (ICPAC) [36]	2025
GHI	Raster	Global Solar Atlas [37]	2025
Wind Speed at 50m	Raster	Global Wind Atlas [38] [39]	2025
Hydro Power Potential	Point	EnergyData [40]	2017
Travel Time	Raster	Harvard Dataverse [41]	2015
Service Transformers	Point	EnergyData [42] and KPLC	2025
RWI	csv	Meta [43]	2021
Mini-Grids	xlsx	REREC	2025

Table A.1: GIS data collection

Parameter	Value	Source	Year
Calibration			
Population at y0	53'330'978	Kenya National Bureau of Statistics [44]	2019
Grid electrification ratio at y0	76.2%	World Bank Group [1]	2023
Urban ratio at y0	32%	World Bank Group [1]	2023

Table A.2

Parameter	Value	Source	Year
Grid urban ratio at y0	96%	World Bank Group [1]	2023
Grid rural ratio at y0	67.9%	World Bank Group [1]	2023
Min nightlights for grid connection	0		
Min population for grid connection	50		
Max MV line distance for grid connection	2 km		
Max transformer distance for grid connection	1 km		
Min nightlights for mini-grid connection	0		
Min population for mini-grid connection	0		
Max MV line distance for mini-grid connection	0.5 km		
Notebook			
Intermediate year electrification target	100%	Kenya Vision 2030 [26]	2008
Final year electrification target	100%	Kenya Vision 2030 [26]	2008
Population projected in final year	70'179'943	Kenya National Bureau of Statistics [44]	2019
Urban ratio in final year	40%		
People per household (urban)	3.2	Kenya National Bureau of Statistics [44]	2019
People per household (rural)	4.1	Kenya National Bureau of Statistics [44]	2019
Rural cut-off size	100 hh		
Grid Parameters			
Grid generation cost in final year	0.1 $\$/kWh$		

Table A.2

Parameter	**Value**	**Source**	**Year**
Grid power plant capital cost	3'853 $\$/kWh$	Kenya Energy Compact [45]	2025
Grid losses	24%	IEA Energy Policy Review [23]	2024
Grid emission factor	500 $gCO2/kWh$	EPRA Biannual Statistics Report [46]	2023
Annual new connections limit	690'000	Kenya Energy Compact [45]	2025
Annual new capacity limit	343 MW	Global Electrification Platform (GEP) [31]	2025
HV line capacity	132 kV		
HV line cost	53'000 $\$/km$	GEP [31]	2025
Grid MV line capacity	33 kV		
Grid MV line cost	7'000 $\$/km$	GEP [31]	2025
Grid MV line max length	50 km		
Grid MV line A limit	274 A		
Grid LV line capacity	0.24 kV		
Grid LV line cost	4'250 $\$/km$	GEP [31]	2025
Grid LV line max length	0.8 km		
Grid service transformer type MV/LV	50 kVA		
Grid service transformer cost	4'250 $\$/unit$	GEP [31]	2025
Grid max nodes per transformer	95		
HV/MV transformer type	23'000 kVA		
HV/MV transformer cost	575'000 $\$/unit$	GEP [31]	2025
Grid discount rate	8%		
Off-Grid Technologies Parameters			
Min households for mini-grid	100		
Min distance from MV lines for mini-grid consideration	0 km		
Diesel price	1.27 $\$/lt$	Global Petrol Prices [47]	2025

Table A.2

Parameter	Value	Source	Year
PV cost	834 $\$/kW$	Mini-Grids for Half a Billion People [48]	2022
Battery cost	142 $\$/kWh$	Mini-Grids for Half a Billion People [48]	2022
Inverter cost	928 $\$/kW$	Mini-Grids for Half a Billion People [48]	2022
Diesel gen cost	261 $\$/kW$	GEP [31]	2025
Wind cost	2'800 $\$/kW$	GEP [31]	2025
Inverter life	15 y		
Diesel life	20 y		
PV life	25 y		
LPSP max	0.1		
Max diesel share	0.5		
MG hydro capital cost	4'000 $\$/kW$	Renewable Power Generation Costs [49]	2024
MG MV line cap	33 kV		
MG MV line cost	7'000 $\$/km$	GEP [31]	2025
MG MV line A limit	274 A		
MG LV line cap	0.24 kV		
MG LV line cost	4'250 $\$/km$	GEP [31]	2025
MG LV line max length	0.8 km		
MG serv tra type MV/LV	50 kVA		
MG serv tra cost	4'250 $\$/unit$	GEP [31]	2025
MG max nodes per tra	95		
Mini-grid discount rate	8%		
SA PV cost ($<20W$)	9620 $\$/kWp$	GEP [31]	2025
SA PV cost ($20<W<50$)	8780 $\$/kWp$	GEP [31]	2025
SA PV cost ($50<W<100$)	6380 $\$/kWp$	GEP [31]	2025

Table A.2

Parameter	Value	Source	Year
SA PV cost ($100<W<1000$)	4470 $\$/kWp$	GEP [31]	2025
SA PV cost ($>1000W$)	6950 $\$/kWp$	GEP [31]	2025
SHS life	5 y		
Stand Alone discount rate	8%		

Table A.2: OnSSET Input Data

Parameter	Value	Comment
Project settings		
Total project duration	20 y	
Discount rate	8%	
Currency	USD	
Time resolution	hourly	
Optimization goal	NPC	minimization
Select system configuration	RES+BESS+DG	
Electricity distribution	AC	
Min RES penetration	50%	
Battery independence	0 $days$	
Max lost load fraction	10%	
Advanced settings		
Model formulation	MILP	
Unit commitment approach	no	
Project type	greenfield	
Investment strategy	capacity expansion over time	
Inv step duration	10 y	
Grid connection	no	
Resource assessment		
Nominal capacity solar	1’000 W	
Panel tilt angle	35°	
Panel azimuth angle	180°	
Ground reflectance	0.2	
Temperature coefficient	-0.5 $\%°C$	
Nominal model operating temperature	45$°C$	

Table A.3

Parameter	Value	Comment
Ambient temperature at NMOT	20°C	
Solar irradiance at NMOT	800 W/m^2	
Demand assessment		
Yearly demand growth (Urban)	2.52%	as computed by OnSSET
Yearly demand growth (Rural)	0.75%	as computed by OnSSET
Solar PV parameters		
Investment cost step 1	0.834 $/$W$	
Investment cost step 2	0.834 $/$W$	
Specific O&M cost	1.5 %inv	
Lifetime	25 y	
Unit CO2 emission	0 $kgCO_2/W$	
Connection to microgrid	separate inverter from BESS	
Inverter η	93%	
Inverter nominal cap	1'000 W	
Inverter lifetime	15 y	
Inverter cost	0 $/$W$	included in panel inv cost
Battery parameters		
Investment cost step 1	0.142 $/$Wh$	
Investment cost step 2	0.142 $/$Wh$	
Specific electronic inv cost	20 %inv	
Specific O&M cost	2 %inv	
Discharge η	92%	
Charge η	92%	
Initial state of charge	50%	
Depth of discharge	80%	
Max discharge time	5 h	
Max charge time	4 h	
Battery cycles	2000 $cycles$	
Lifetime	10 y	
Unit CO2 emission	0 $kgCO_2/kWh$	
Inverter η DC to AC	93%	
Inverter η AC to DC	93%	

Table A.3

Parameter	**Value**	**Comment**
Inverter nominal capacity	1'000 W	
Inverter lifetime	15 y	
Inverter cost	0.928 $\$/W$	
Diesel parameters		
Nominal η	33%	
Specific inv cost	0.261 $\$/W$	
Specific O&M cost	10%	
Lifetime	20 y	
Fuel LHV	9944.5485 Wh/l	
Fuel CO2 emission	2.68 $kgCO_2/l$	
Fuel cost type	fixed price	computed with Szabo formula
Enable partial load	no	

Table A.3: MicroGridsPy Input Data

B | Appendix B

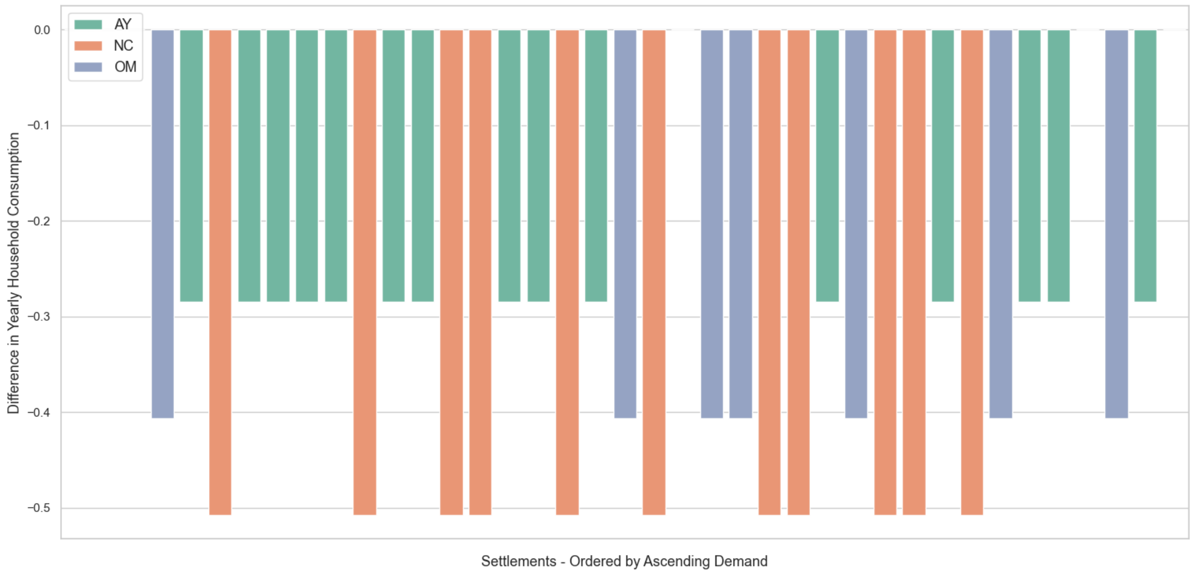


Figure B.1: Household consumption difference when using the mono-tier representation in MicroGridsPy (including outliers)

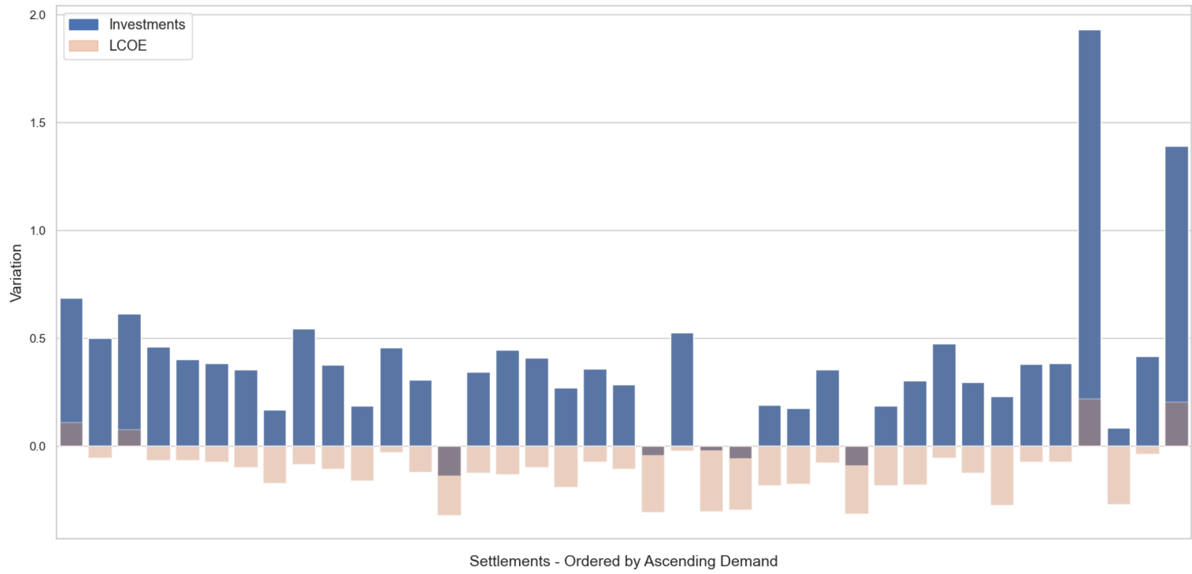


Figure B.2: Investment Costs and LCOE variations between MicroGridsPy and OnSSET (including outliers)

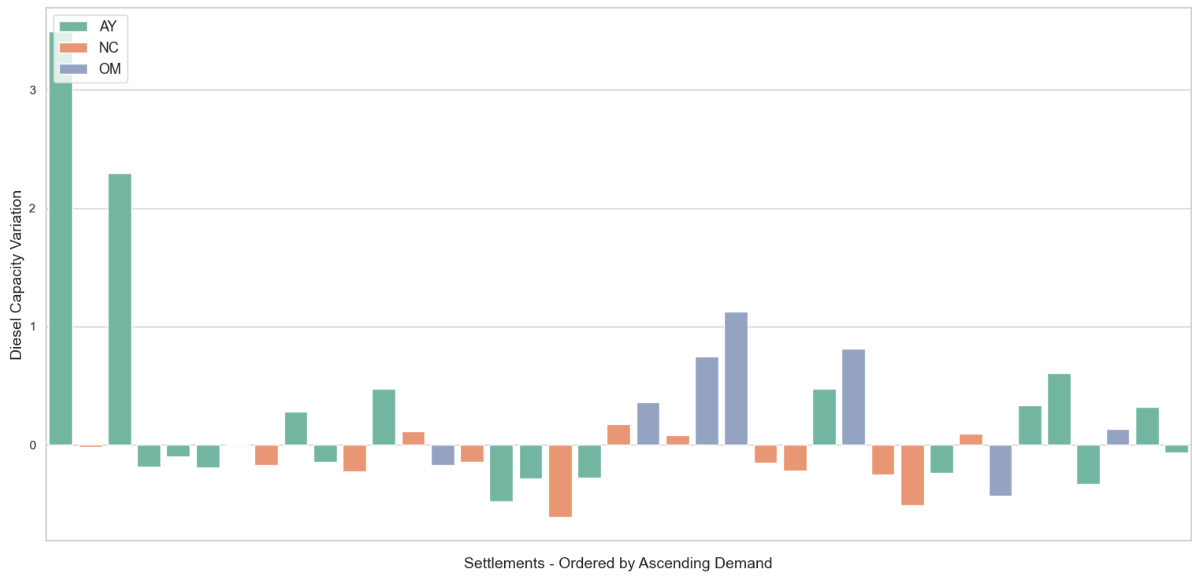


Figure B.3: Diesel capacity variation between MicroGridsPy and OnSSET (including outliers)

Figure B.4: BESS capacity variation between MicroGridsPy and OnSSET (including outliers)

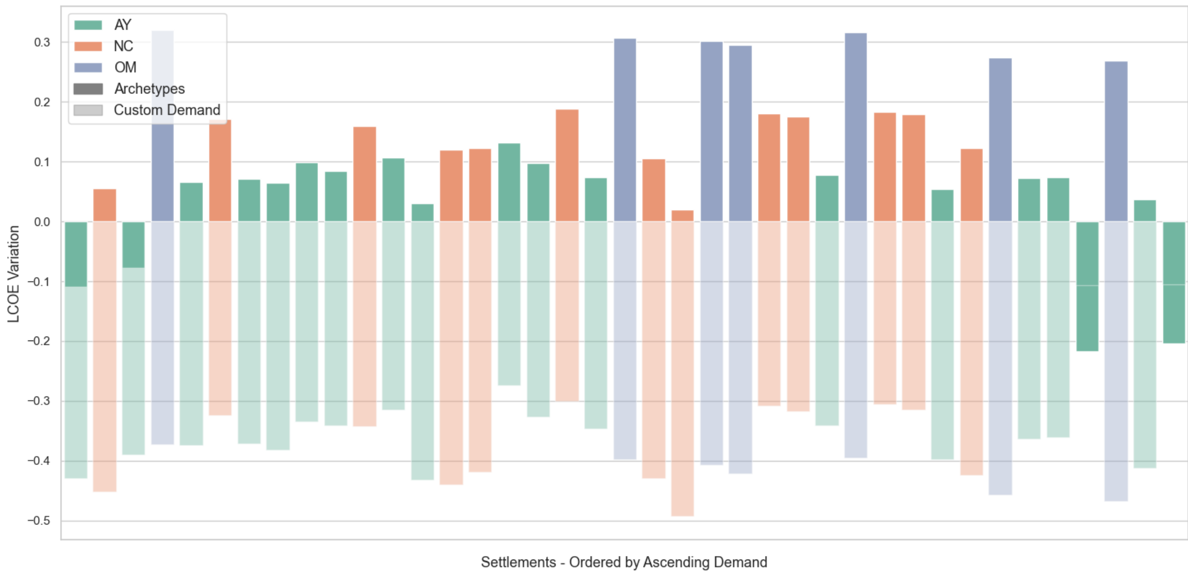


Figure B.5: LCOE difference between OnSSET versions and MicroGridsPy (including outliers)

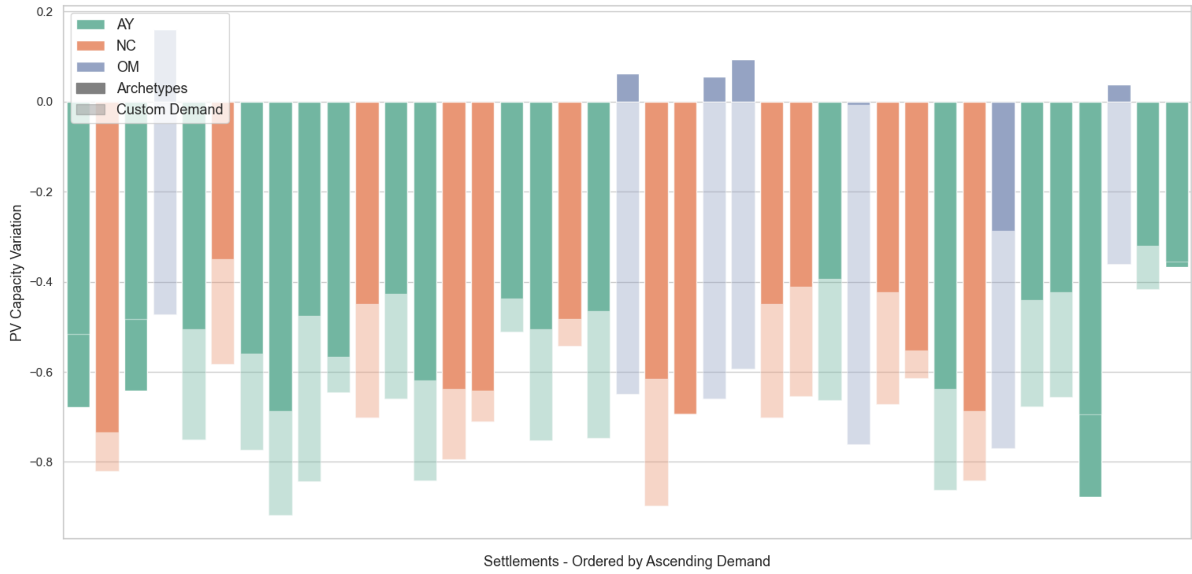


Figure B.6: PV capacity difference between OnSSET versions and MicroGridsPy (including outliers)

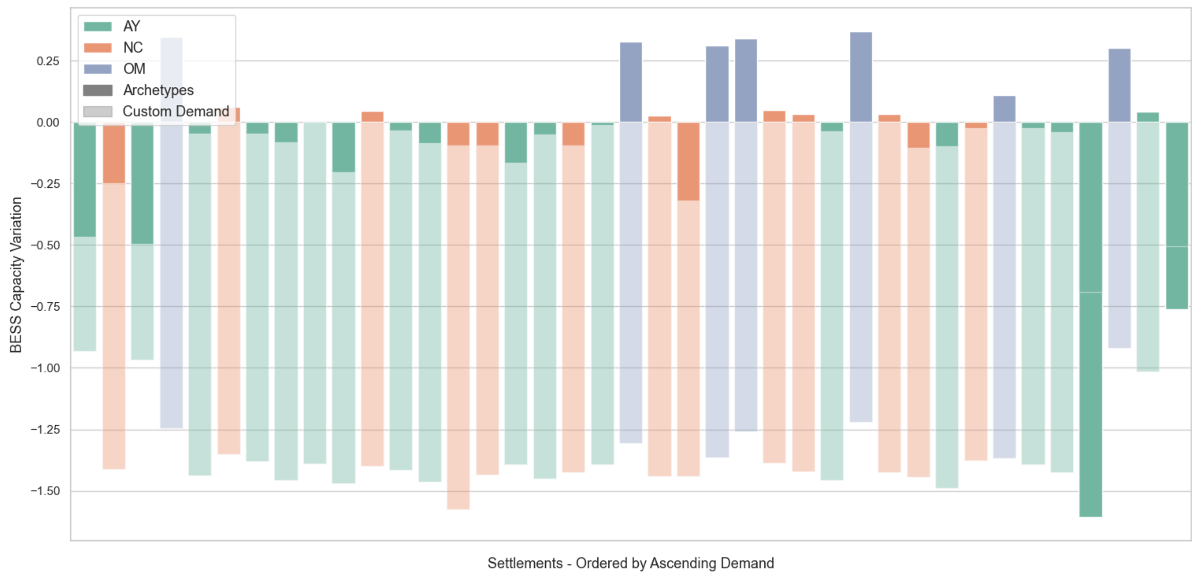


Figure B.7: BESS capacity difference between OnSSET versions and MicroGridsPy (including outliers)

List of Figures

List of Tables