Energy Strategy Reviews 65 (2026) 102213

Contents lists available at ScienceDirect

Energy Strategy Reviews

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

Which customers belong together? An enhanced off-grid clustering algorithm for cost-effective rural electrification

Pedro Ciller [a,b,*], Sara Lumbreras [a], Andrés González-García [a]

[a] Institute for Research in Technology (IIT), ICAI, Comillas Pontifical University, Madrid, 28015, Spain
[b] Department of Quantitative Methods, ICADE, Comillas Pontifical University, Madrid, 28015, Spain

ARTICLE INFO

Handling Editor: Jesse Van Griensven Thé

Keywords:
Rural electrification
Large-scale planning
Clustering
Mini-grids
Off-grid systems

ABSTRACT

With 750 million people lacking access to electricity, cost-effective rural electrification is essential. A critical challenge for rural electrification projects is determining whether to cluster consumers to best serve them with standalone systems, mini-grids, or grid extensions. While state-of-the-art models offer advanced solutions, current clustering algorithms often rely on simplified cost estimators and rigid, bottom-up approaches, limiting their accuracy and adaptability.

This paper introduces a clustering algorithm that advances the state of the art by thoroughly evaluating the space of potential off-grid clustering solutions (i.e., the algorithm excludes extensions of the power grid as alternatives) and enhancing the accuracy of cost estimations. Applied to the Cajamarca region in Peru, it reduced electrification costs by 6.16% compared to a traditional state-of-the-art clustering method. Qualitatively, the method produced smaller, better-sized mini-grids and more appropriate allocations of standalone systems, demonstrating planning accuracy for sustainable energy access. An additional sensitivity analysis was performed, demonstrating the algorithm's ability to consistently deliver more cost-efficient and flexible electrification solutions, thereby contributing to sustainable energy access.

1. The rural electrification challenge

Approximately 750 million people lack access to electricity [1]. Electricity access is a key driver for developing underserved regions, with vital services such as education and healthcare relying on electricity to function [2]. Target 7.1 of the seventh Sustainable Development Goal (SDG 7) seeks to achieve universal energy access by 2030 [3], and this should be done in an environmentally responsible way [4–6]. Achieving this goal will require significant financial and institutional efforts. Estimates indicate an annual investment of 35–40 billion USD is needed to provide universal electricity access by 2030 [7].

Establishing an electrification agenda for underserved regions is a complex and challenging process. Factors such as regulatory challenges, financial constraints, and planner preferences must be carefully considered [8]. Electrification plans typically combine grid extensions with off-grid alternatives such as mini-grids and isolated systems. Traditional grid extensions generally suit areas close to the power grid and with a high density of consumers, whereas off-grid systems tend to be a good alternative for places where consumers are scattered or far from the main grid. Developing a robust techno-economic plan—determining which consumers should be electrified using standalone systems, mini-grids, or grid extensions while minimizing overall costs—is essential for advancing the electrification agenda. It is essential to utilize economic resources efficiently to promote development and accelerate progress toward universal energy access, and this should be done by setting the right institutional drivers [9–11] and green initiatives [12].

Computer-based software models have proven effective in minimizing costs in techno-economic planning. The goal of these models is to assist planners in determining the optimal electrification strategy for a region ranging in size from a single village to an entire country, recommending where traditional grid extensions should be implemented and where mini-grids or isolated systems offer the best alternatives. Models need to calculate the costs of technical designs for grid extensions, mini-grids, and isolated systems for the electrification solution they propose. This is achieved by solving multiple interrelated subproblems, which include the grouping or clustering of consumers into mini-grids and grid extensions, the optimization of generation designs for off-grid systems, and the optimization of distribution networks for mini-grids and grid extensions [13].

* Corresponding author. Institute for Research in Technology (IIT), ICAI, Comillas Pontifical University, 28015, Madrid, Spain.
E-mail addresses: pciller@comillas.edu (P. Ciller), slumbreras@comillas.edu (S. Lumbreras), agonzalez@comillas.edu (A. González-García).

https://doi.org/10.1016/j.esr.2026.102213
Received 4 December 2025; Received in revised form 5 March 2026; Accepted 15 March 2026
Available online 14 May 2026

The grouping or clustering of consumers into systems is a significant challenge in the development of electrification plans [14]. The cost related to a combination of consumers into isolated systems, mini-grids, and grid extensions should accurately reflect the electrification costs involved in these systems, but the calculation of such detailed costs for potentially thousands of different alternatives is computationally challenging and entails many difficulties. For example, numerous algorithms [15] and software tools [16] exist that optimize the generation cost of a single mini-grid or a hybrid renewable system [17,18]. There are also algorithms that optimize the temporal operation of an individual mini-grid [19], and deal with demand response and uncertainty [20]. Similarly, some algorithms in the literature can be adapted to optimize the network cost of mini-grids [21]. However, in large-scale electrification planning, the clustering of consumers may consider thousands of potential mini-grids, making it computationally infeasible to optimize the generation and network design of each one from scratch.

One way to avoid the hurdles related to clustering consumers is not to operate with individual consumers, but with villages or settlements of terrain cells. This approach, which many geospatial models employ, reduces computational complexity at the expense of compromising solution accuracy. Examples of such models include GEOSIM, which clusters villages around development poles (villages with high scores based on specific indicators) using a Huff-model-based algorithm [22], which assigns probabilities of villages being electrified together with a development pole based on distances and other criteria. Another example is OnSSET, whose light version employs a clustering algorithm that merges adjacent cells into clusters [23]. Similarly, Ref. [24] outline a methodology that creates buffers around populated areas, polling units, and schools, clustering overlapping zones together. Network Planner is a tool that operates at the village level and uses a variant of the Kruskal algorithm to determine the network layout of villages electrified through power grid extensions [25,26].

A few models operate at the consumer level instead, offering more detailed solutions and handling algorithmic challenges. For example, the latest version of OnSSET [27,28] uses DBSCAN, a density-based clustering method [31], to group individual consumers. The most recent version of IntiGIS [32] applies a clustering method that incorporates consumer density and distance-based criteria. Ref. [33] proposes clustering based on consumer density and geometric criteria. However, these methods do not incorporate cost criteria, which may result in suboptimal cost minimization.

To our knowledge, the Reference Electrification Model (REM) is the only rural electrification planning tool that clusters individual consumers while performing cost comparisons among potential solutions, selecting the lowest-cost option [34]. REM works with individual buildings and is the only one optimizing mini-grid and grid extension designs with electrical constraints, such as power flows and voltage drops. It also optimizes generation designs for mini-grids and standalone systems by simulating hourly dispatch of potential designs [35]. REM has been utilized to develop rural electrification master plans for institutions such as the World Bank (WB), the Inter-American Development Bank (IADB), and the Asian Development Bank (ADB) [36]. REM-based plans include projects in Rwanda [37], Mozambique [38], Indonesia [39], Ecuador [40], Bolivia [40], Haiti [41], and Paraguay [42].

REM's clustering also has drawbacks, which also appear in the rest of the state-of-the-art models that operate with individual consumers. First, it estimates the network costs of candidate mini-grids and grid extensions without fully accounting for the nonlinear complexities of network designs, which are governed by electrical constraints such as Kirchhoff's laws and voltage drop limits. This can result in a significant mismatch between the estimated network cost and the actual network cost, leading to suboptimal decisions regarding cluster merging.

Additionally, REM's clustering does not perform an in-depth evaluation of the space of potential electrification solutions. The algorithm starts with each individual consumer being its own isolated cluster. Then, the model performs cost comparisons among pairs of nearby clusters to determine whether it is worth electrifying these clusters together (and then it joins the clusters) or if these clusters are better electrified separately (and it does not join the clusters). However, this procedure overlooks the potential benefits of connecting more than two clusters in a single iteration.

To sum up, several geospatial electrification models cluster with algorithms that have important methodological limitations. Many approaches operate at coarse spatial resolution or rely on density- and distance-based heuristics that do not explicitly incorporate cost comparisons during clustering. Even advanced tools such as REM, which cluster individual consumers and compare electrification alternatives, depend on oversimplified cost estimators and greedy bottom-up procedures that evaluate only local pairwise mergers. As a result, clustering decisions may be based on inaccurate cost approximations, and the algorithm may fail to fully explore the solution space. This implies a research gap for a clustering methodology that accurately and systematically estimates costs and explores the space of potential clustering solutions.

This paper addresses these limitations by proposing an enhanced off-grid clustering algorithm (i.e., extensions of the power grid are not considered) that combines systematic exploration of the clustering solution space with an accurate, computationally efficient network cost estimator. The algorithm introduces an exploratory clustering phase that constructs a hierarchical structure of candidate solutions, avoiding premature termination. It also applies a regression-based network cost estimator to incorporate electrical constraints into clustering decisions without incurring prohibitive computational costs.

By integrating these two methodological advances, the proposed clustering algorithm enables reliable cost comparisons across cluster configurations and yields measurable cost reductions in rural electrification planning. The algorithm is implemented in the REM model and applied to a case study in the Cajamarca region (Peru) to demonstrate its effectiveness against the current REM clustering.

Table 1 summarizes the clustering algorithms used by the main large-scale rural electrification planning tools and the algorithm presented in this paper, noting the spatial resolution of each tool (i.e., buildings, settlements, or GIS cells) and the cost estimates used during clustering (if any). Some tools have evolved over time, and they are listed several times to reflect their evolution.

The clustering algorithm presented in this paper could have a significant impact on real electrification projects. It is estimated that an annual investment of 35 billion USD is needed to achieve universal access to electricity by 2030 [49], and rural electrification planning models play a crucial role in ensuring the efficient use of these necessary investments. Any improvement in the algorithms of planning tools, such as clustering, could lead to significant savings in real-world electrification plans based on the outcomes of these tools [50,51].

The main objective of this paper is to present a clustering algorithm that groups consumers who are better electrified with the same off-grid system. The algorithm performs a thorough evaluation of the space of potential clustering solutions and uses accurate estimations of the costs involved in the final electrification solution.

The level of achievement of this objective is measured in terms of cost reduction in the final electrification solution. Adequately exploring the space of potential clustering solutions and using accurate cost estimations are desirable attributes of the clustering algorithm, whose effectiveness is evaluated by the cost reduction in the final electrification solution.

The remainder of this paper is structured as follows: Section 2 presents an overview of the REM model, and Section 3 presents its current clustering algorithm. Section 4 describes the proposed clustering algorithm, and Section 5 applies it to a case study. Section 6 presents conclusions and future research directions.

Table 1
Rural electrification planning tools and their clustering algorithms.

Reference	Tool	Spatial resolution	Clustering algorithm	Cost estimations in clustering
This paper	REM (enhanced off-grid clustering)	Buildings	Geometric-based (Delaunay triangulation) and cost-based heuristic	Accurate estimations for off-grid systems, quick estimations for grid extensions
[34]	REM (traditional off-grid clustering)	Buildings	Geometric-based (Delaunay triangulation) and cost-based heuristic	Quick estimations
[27]	Detailed OnSSET	Buildings	Density-based (DBSCAN)	–
[23]	Light OnSSET	Settlements	The tool is compatible with the clustering algorithm described in Ref. [43], which aggregates GIS cells to form population settlements.	–
[44]	Initial version of OnSSET	GIS cells	None (it assumes each cell is its own cluster)	–
[32]	IntiGIS (most recent version)	Buildings	Density-based (DBSCAN) and geometric-based (MST)	–
[45]	IntiGIS 2	GIS cells	None (it assumes each cell is its own cluster)	–
[46]	IntiGIS 1	GIS cells	None (it assumes each cell is its own cluster)	–
[29,30]	GISEle	GIS cells	Density-based (DBSCAN)	–
[25,26]	Network Planner	Settlements	Geometric-based (MST) for grid extensions. It evaluates whether to connect settlements within the same grid extension.	Quick estimations
[47,48]	GEOSIM	Settlements	Probabilistic. Gravity-based (Huff Model). It clusters villages around a few, labeled as “Development Poles”. Later, it can create electric clusters by joining nearby settlements to Development Poles until the levelized cost of electricity does not decrease any further.	Quick estimations (for electric clusters)

2. The Reference Electrification Model

REM is a large-scale geospatial electrification planning model designed to develop the least-cost techno-economic plan for electrifying underserved regions. The model identifies areas suitable for grid extensions and zones where off-grid systems offer the most viable solutions, such as mini-grids and standalone systems. REM calculates the optimal generation designs for each mini-grid and standalone system included in the electrification solution. Additionally, it optimizes the distribution network layout for both mini-grids and grid extensions featured in the solution.

Before REM determines the final electrification mode of each consumer, it groups the consumers into clusters that represent potential mini-grids, grid extensions, or combinations of isolated systems. The clustering process is inherently complex, as the optimal electrification alternative for each consumer—whether a mini-grid, grid extension or standalone system—is not predetermined.

Using an iterative process, the clustering algorithm begins by grouping consumers into potential mini-grids and grid extensions. Initially, each consumer is treated as an individual cluster. The algorithm then connects nearby clusters if doing so results in a lower-cost electrification solution compared to electrifying them separately. The algorithm operates in two stages. In the first stage, known as off-grid clustering (hereafter, *traditional off-grid clustering*), it assumes that only off-grid solutions are viable. This stage generates off-grid clusters, representing the least-cost combination of consumers for mini-grids and standalone systems. In the second stage, known as grid-extension clustering, the algorithm incorporates grid extensions as viable electrification options. This stage starts with the off-grid clusters and computes the grid-extension clusters. Ultimately, the clustering algorithm forms potential mini-grids and grid extensions, represented by off-grid and grid-extension clusters. However, it does not assign a definitive electrification mode to each consumer. Instead, it constructs a hierarchical structure of clusters, as illustrated in Fig. 1.

The model determines the final electrification solution by comparing costs across the three levels of the hierarchical structure illustrated in Fig. 1. A grid-extension cluster (belonging to the on-grid level) may ultimately be electrified using a combination of mini-grids and isolated systems if the cost of off-grid solutions for the clusters within the grid-extension cluster is lower than the cost of extending the grid to the entire cluster. Conversely, an off-grid cluster may be electrified with a grid extension if this is the least-cost solution for that cluster.

The clustering process aims to minimize the electrification cost for each cluster. This is challenging because the number of potential clusters is vast in large-scale planning, and performing detailed designs for every potential cluster would require unmanageable computation time.

Although knowing the exact generation and distribution network cost would be ideal for clustering decisions, calculating them exactly for all potential mini-grids and grid extensions is computationally infeasible for large-scale regions. To address this, REM relies on cost estimations for generation and network costs when evaluating whether to connect two clusters. In a preprocessing stage, REM optimizes the generation designs for a representative set of mini-grids deemed relevant for the electrification plan. These representative mini-grid generation designs and associated costs are stored in a lookup table. Suppose the clustering algorithm requires the generation cost of a mini-grid not included in this table. In that case, it estimates the cost by interpolating from the nearest generation designs in the lookup table. When evaluating the connection of two clusters during traditional off-grid clustering, REM uses the lookup table to estimate generation costs and calculates incremental network costs by considering the line connecting the centroids of the clusters. After clustering consumers into potential mini-grids and grid extensions, the model computes accurate costs to finalize the electrification solution by determining precise generation costs for off-grid systems and producing detailed network designs for mini-grids and grid extensions. Fig. 2 depicts the sequential flow of REM. The second, third, and fourth blocks represent REM's algorithms, while the first and last blocks pertain to data processing stages that format the input and output data appropriately.

The first block of REM is data preparation. The model requires a substantial amount of data to operate, which includes the geospatial coordinates of each individual consumer and their hourly demand profiles, the location of the existing network, the generation and

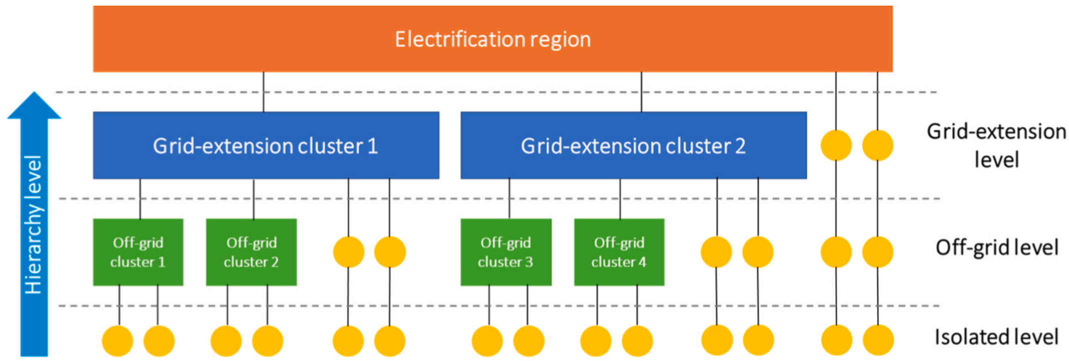


Fig. 1. Hierarchical structure of clusters. Source: adapted from Ref. [52].

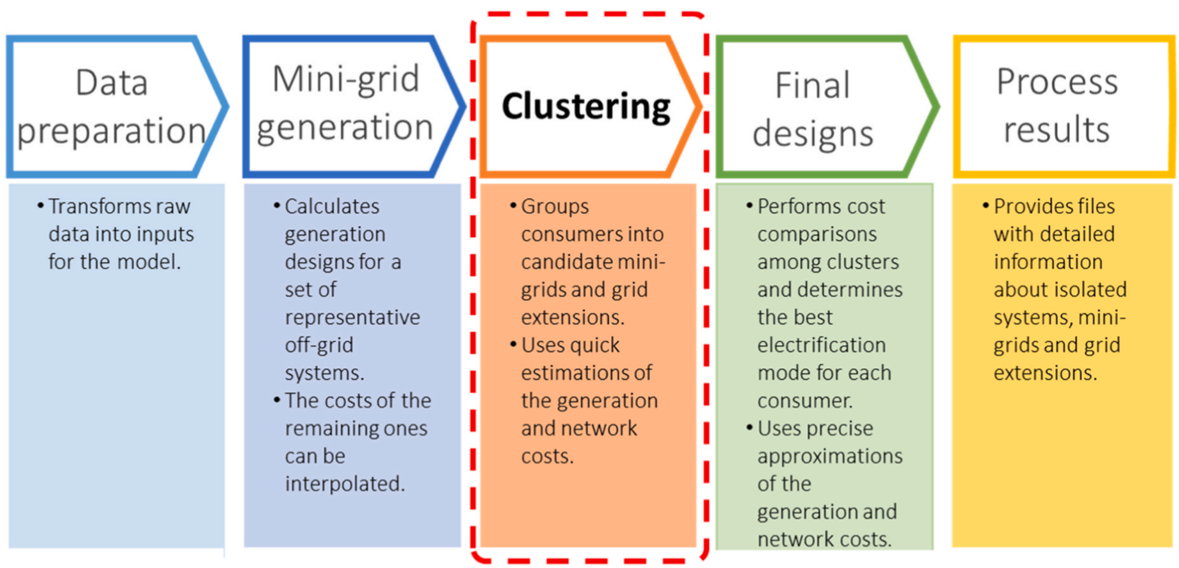


Fig. 2. High-level structure of REM.

network catalogs of available components, solar irradiance, and several techno-economic parameters, such as component lifetimes and discount rates.

The second block is the mini-grid generation. In this block, REM optimizes the generation designs of several mini-grids and isolated systems from scratch, which are representative of the case study. The optimized generation designs and their corresponding costs are stored in a lookup table, which is used later to interpolate the generation costs of the remaining off-grid systems if needed.

The third block corresponds to the clustering algorithm, which groups the consumers into potential mini-grids and grid extensions by creating a hierarchical structure of clusters. The algorithm utilizes the look-up table calculated in the second block to interpolate the generation cost of the potential mini-grids it considers, and it estimates the incremental network cost of potential mini-grids and grid extensions using fast approximations that avoid calculating the entire distribution network of the systems.

In the fourth block, REM evaluates the hierarchical structure of clusters. The model calculates accurate network designs for potential mini-grids and grid extensions. REM still interpolates the generation costs from the look-up table in this step, because this approximation is sufficiently accurate for large-scale electrification as shown in Ref. [53].

The calculation of network designs for candidate mini-grids and grid extensions is performed using the Reference Network Model (RNM). RNM is a tool that optimizes the distribution network design for a given set of consumers and their corresponding demands [54]. RNM incorporates power flow calculations and standard electrical constraints, such as allowable voltage drops. It also accounts for topographical features, such as altitude raster maps and restricted or penalized zones [55]. While REM utilizes RNM to calculate the network layout for mini-grids included in the final electrification solution, applying RNM to optimize the network design for each potential mini-grid during clustering is computationally prohibitive in large-scale areas.

Finally, the model generates geospatial files and tables that provide a detailed description of the electrification solution. The files include a cost breakdown and the amount of non-served energy of each system, the network layout of mini-grids and grid extensions, and the generation design of each isolated system and mini-grid.

The next sections provide a detailed description of the traditional off-grid clustering of REM, offering a critical overview of the algorithm and highlighting its limitations.

3. Clustering in REM

As mentioned in section 2, REM creates a hierarchical structure of clusters following a sequential process. Initially, each consumer is its own isolated cluster. Then, REM determines whether it is worth merging nearby clusters by performing cost comparisons in a three-step process. The first step determines the candidate connections among consumers. The second step is called traditional off-grid clustering, and it assumes temporarily that only individual isolated systems and mini-grids are viable electrification solutions. The third step is named grid-extension clustering, and it incorporates extensions of the power grid as potential electrification solutions. The remainder of this section explains the first and the second steps of REM's clustering, which are the relevant parts for this paper.

3.1. Candidate connections

The REM clustering starts with all the consumers being their own isolated clusters. Then, the clustering computes the Delaunay triangulation of the consumers, which divides the set of consumers into non-overlapping triangles that cover the convex hull of consumers. The Delaunay triangulation represents the potential connection among clusters. Fig. 3 shows the consumers from the case study, which will be described in Section 5, along with their corresponding Delaunay triangulation.

The Delaunay triangulation includes several properties that make it a good fit for clustering nearby consumers. Some of its properties include that the minimum spanning tree (MST) of a set of points is always contained in their Delaunay triangulation, ensuring that no arc of the MST is missed. Arcs of the Delaunay triangulation never intersect, and consumers are always connected to their closest consumers [56].

This triangulation has been used to cluster consumers when designing distribution networks [54,57], and there are cases where there is a close match among the arcs of a Delaunay triangulation and the real lines of the electric network [58].

3.2. Traditional off-grid clustering

Having calculated the Delaunay triangulation, the traditional off-grid clustering iterates through its edges, starting with the shortest edges and progressing to the longest. For each edge, the algorithm performs a cost comparison between two configurations (see Fig. 4) to decide whether the connected clusters should be electrified together.

The costs associated with configuration 1 include all the costs of clusters 1 and 2 electrified separately as different off-grid systems, and the costs related to configuration 2 include all the costs of clusters 1 and 2 electrified together as a single off-grid system. One of these costs is the cost of the distribution network, which is difficult to calculate accurately for each potential cluster considered during the traditional off-grid clustering process.

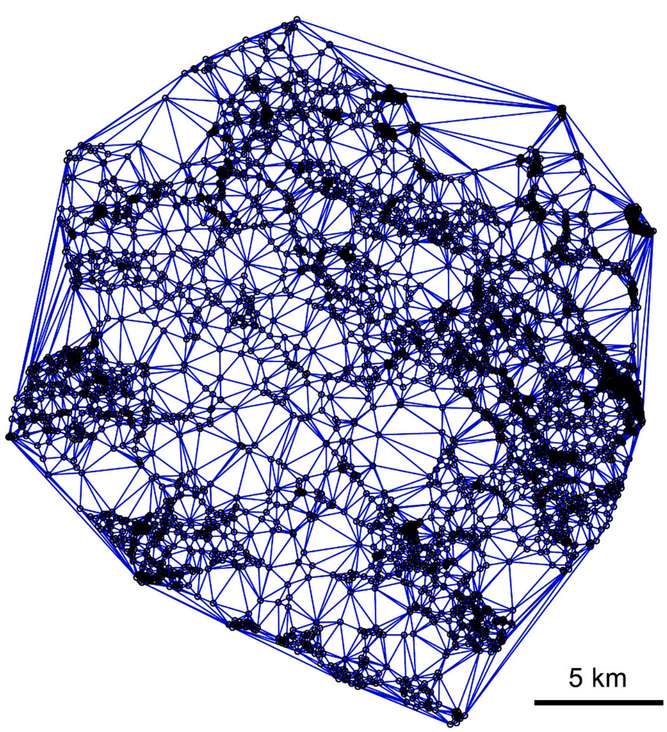


Fig. 3. Delaunay triangulation of the consumers considered for the case study. Consumers are represented with black circumferences and the arcs of the Delaunay triangulation are plotted with blue lines.

Instead of calculating the detailed network cost for each cluster, REM estimates the additional network cost between configurations 1 and 2 as the cost of a line connecting the centers of both clusters. However, the cost of an interpolated line is not always an accurate estimator of the incremental network cost. Network designs exhibit nonlinear behavior due to electrical constraints, such as Kirchhoff's laws and the maximum allowable voltage drop. Calculating power flows is necessary to optimize a network design, and this task becomes more complex with a wide range of network components. An incorrect estimation of the incremental network cost may lead to merging clusters that are better electrified separately or not merging clusters that would benefit from being electrified with the same system.

Even if network costs were accurately calculated, a significant portion of the clustering solution space may remain unexplored because the traditional off-grid clustering in REM halts when local cost comparisons no longer appear to benefit the merging of clusters. This unexplored portion could be essential to consider.

To summarize, the traditional off-grid clustering in REM, which can be considered the most advanced in the state of the art, has two fundamental drawbacks. The first is its reliance on a network cost estimator that does not account for the complexity of network designs, leading to suboptimal decisions about merging or separating clusters. The second is the bottom-up, local logic, which relies on greedy decisions and is sensitive to local minima, making it challenging to explore the space of potential clustering solutions thoroughly.

The next section introduces a novel clustering method, named *enhanced off-grid clustering*, which addresses these two main drawbacks. This algorithm thoroughly evaluates the space of potential clustering solutions to produce a proposed clustering solution.

4. Enhanced off-grid clustering: an alternative for smart customer grouping

This section presents the enhanced off-grid clustering, which aims to overcome the two main flaws of traditional off-grid clustering: its reliance on a faulty network cost estimator and its inability to thoroughly evaluate the space of clustering solutions. The enhanced off-grid clustering process follows a sequential approach, which is explained in the remainder of this section. Fig. 5 shows the current and proposed REM clusterings, illustrating how enhanced off-grid clustering integrates, as well as its sequential steps.

The sequential steps of the enhanced off-grid clustering are explained in the remainder of this section.

4.1. Exploratory clustering

The first obstacle to improving the traditional off-grid clustering is to ensure that the space of potential clustering solutions is properly explored, thereby avoiding cases where a local minimum prevents the algorithm from reaching the optimal solution.

This obstacle is overcome in the initial stage of enhanced off-grid clustering, known as exploratory clustering. The exploratory clustering will store many potential clustering solutions that form a hierarchical structure and are representative of the space of potential clustering solutions. Later, the enhanced off-grid clustering will determine the clustering solution by performing an accurate cost evaluation of the stored clustering solutions.

The exploratory clustering also begins with an initial clustering where all consumers are their own isolated clusters, and it examines whether it is worthwhile connecting nearby clusters using the same cost

Fig. 4. The two configurations considered. Configuration 1 includes two clusters that are electrified separately, whereas configuration 2 considers that both clusters are electrified together. The green dots represent consumers, and the green segment represents a distribution line connecting the two clusters.

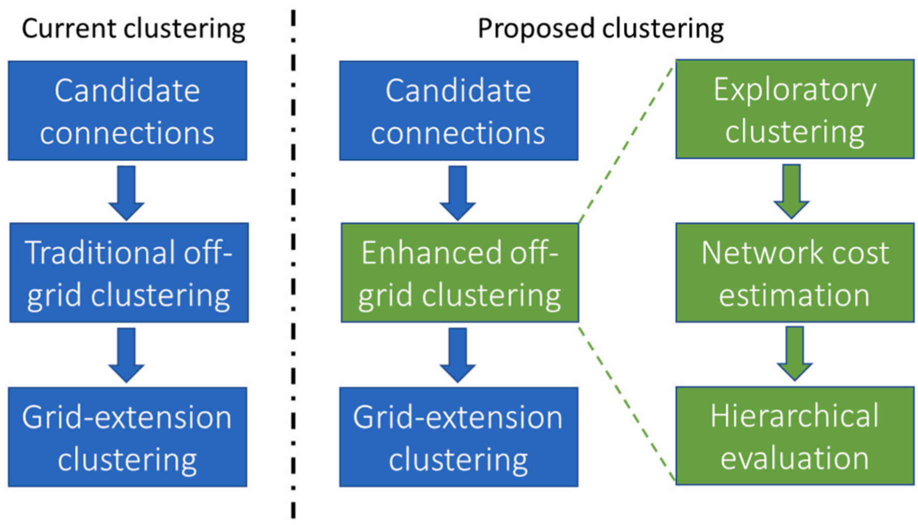


Fig. 5. REM current and proposed clustering flow diagrams.

comparisons as in the traditional off-grid clustering. The algorithm stores a clustering solution each time a certain number of clusters are connected.

Once the exploratory clustering has merged all connections among nearby clusters based on the cost comparisons, it will add a cost margin to the comparisons so that connections that are not worth connecting by a difference lower than the cost margin are also connected.

The cost margin is progressively increased, so all potential connections among nearby clusters are activated at the end of the exploratory clustering unless a cluster becomes large enough that an LV mini-grid would no longer be a viable electrification solution for that cluster. The reason to increase the cost margin progressively, rather than directly merging each pair of nearby clusters, is to maintain an order of connections where pairs of clusters that would benefit more from being electrified together are prioritized over the remaining ones.

The exploratory clustering calculates the values of the cost margin considering the worst-case scenario where a connection was not activated (i.e., the algorithm stores the maximum value that the differences between the costs of configuration 2 and configuration 1 take). The additional values are provided by a logarithmically spaced vector that starts at one and ends at one hundred times the worst-case cost, where a connection was not activated. Finally, the cost margin is set to a very large value to ensure that all candidate connections are activated.

The reason for using a logarithmically spaced vector instead of a linearly spaced one is to ensure that the cost margin increases slowly in the initial steps, where a small increase in the cost margin may lead to several connections being activated. However, the authors acknowledge that the method used to increase the cost margin may not be the most effective one, and further research regarding this point should be performed. Additional details regarding the implementation of the cost margin are provided in the pseudocode of the exploratory clustering in appendix B, and additional sensitivity analyses regarding parameters used in the exploratory clustering are provided in appendix D.

It should be noted that the cost margin mechanism is a procedure that ensures no relevant part of the space of clustering solutions is left unevaluated, but it is not intended to be a fine-tuned optimization procedure that guarantees an optimal ordering of the activation of the inactive arcs of the Delaunay triangulation (it would be impossible to determine such optimal ordering without accurate network costs). In other words, the critical point of the cost margin is to ensure that no large parts of the space of potential solutions are left unevaluated; the specific numerical form of the relaxation schedule is secondary, as long as it achieves broad and systematic coverage. Appendix D shows that an alternative strategy, where all the inactive connections of the Delaunay triangulation are activated in an additional pass, yields results similar to those of the procedure described in this section, indicating that the proposed method does not rely on specific parameter values to work.

The criteria used to ensure that the resulting mini-grid is low-voltage are inherited from RNM, which is the tool that ultimately designs all networks for mini-grids and their extensions to the power grid. REM does not perform a low-voltage network design for a mini-grid if the mini-grid has more than 4000 consumers or the consumers are outside of a square of side length 8 km.

Fig. 6 illustrates an example of exploratory clustering, showing how the electrification cost varies with the number of connected clusters. The exploratory clustering would traverse the space of potential clustering solutions, exploring solutions beyond the local minima reached when 100 clusters are connected, and storing them in the hierarchical structure of clusters. In this example, the traditional off-grid clustering would not analyze solutions where more than 100 clusters are connected.

The sizes of the clusters grow as nearby clusters are merged into larger clusters. Therefore, the exploratory clustering stores solutions

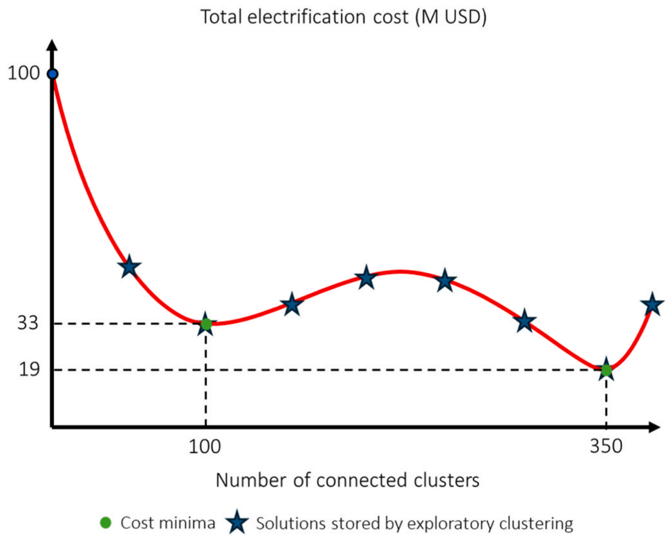


Fig. 6. Example of intermediate solutions stored by the exploratory clustering. The traditional off-grid clustering would fail to evaluate solutions beyond the first cost minimum.

where cluster sizes range from very small to considerably large. Fig. 7 illustrates this point with an example of several intermediate clustering solutions that the enhanced off-grid clustering could consider, showcasing the additional solution space explored by this modification.

It should be noted that exploratory clustering utilizes the same cost estimations as traditional off-grid clustering. Therefore, exploratory clustering aims to adequately explore the space of potential clustering solutions, but it does not properly evaluate the cost of distribution networks related to mini-grids. The second step of the enhanced off-grid clustering addresses this issue.

4.2. Network cost estimation

This section presents an overview of the network cost estimator that will be used later to obtain accurate network costs for the clusters. The network cost estimator algorithm was introduced in Ref. [59]; further details regarding this algorithm can be found in that reference.

It is essential to incorporate a network cost estimator that strikes a balance between cost accuracy and computational efficiency, as large-scale planning often involves evaluating thousands of potential clusters.

There are methods that quickly estimate the network cost of a potential mini-grid. They apply rules of thumb or geometric-based algorithms, which generally calculate the MST of a set of consumers to obtain the distribution network [44]. However, these methods share the same issues as the current network cost estimator that traditional off-grid clustering applies: they neglect electrical constraints, such as power flows and voltage drops, thereby misestimating the real cost of a distribution network.

Some algorithms optimize a distribution network through detailed power flow calculations, ensuring that electrical requirements are satisfied, albeit at the expense of higher computation time compared to geometric-based algorithms [54,60]. The detailed network optimization of each potential mini-grid, stored in the hierarchical structure of clusters developed by exploratory clustering, is computationally infeasible for large-scale planning.

We have developed a method that strikes a balance between accuracy in cost estimation and computational speed. This method provides quick estimates of network costs for mini-grids, taking into account the electrical implications that are critical for the feasibility of the designs.

This method follows a sequential procedure. The first step is to obtain a set of mini-grids that is representative of the case study. The network cost estimator leverages the hierarchical structure that exploratory clustering has already created, utilizing it as input. The network cost estimator performs a k-medoids clustering [61] of all the potential mini-grids stored in the hierarchical structure to group mini-grids that are similar in terms of MST length and electric demand, which are features that measure the geometric and electric nature of the mini-grid networks, respectively.

The second step involves obtaining the network cost of each representative mini-grid by designing its network in detail, calculating the layout, and determining the cost of all components. The network design is calculated using RNM, a model that incorporates electrical constraints into the design and can handle a catalog with multiple network lines for each voltage level. The RNM model was introduced in Section 2, where we explained that RNM is the model REM uses after the clustering process to design the networks of mini-grids and grid extensions for the final electrification solution.

The third step determines the explanatory variables used in the model to estimate the network costs, which is a piecewise linear regression model. The explanatory variables are obtained by considering a set of candidate variables, such as the length of the MST, the number of mini-grid consumers, the total network demand, or the electric consumption of mini-grid consumers (among others). Then, a statistical procedure called hierarchical regression is applied, which selects the most significant variables [62]. Further details regarding the application of the network cost estimator to the case study analyzed in section 5 and the linear models involved are provided in appendix A.

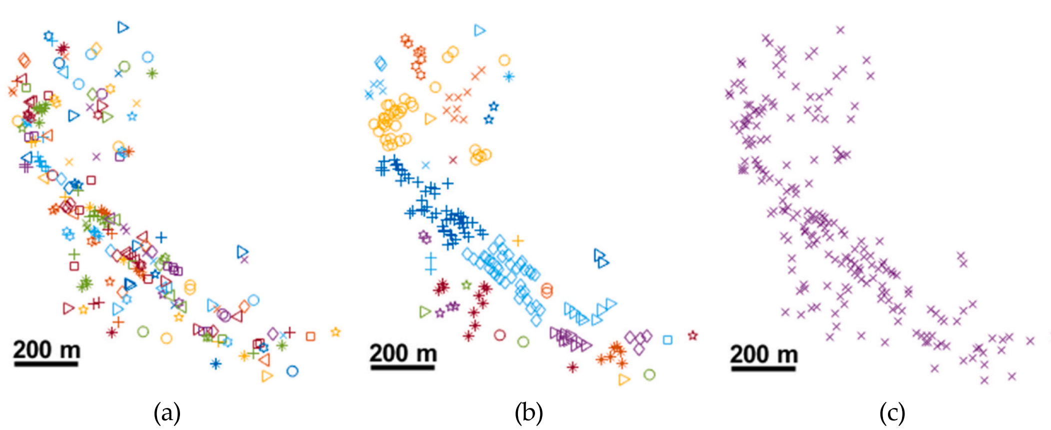


Fig. 7. Example of intermediate solutions with clusters of different sizes. Nearby consumers who belong to the same clusters are represented with the same symbol and color.

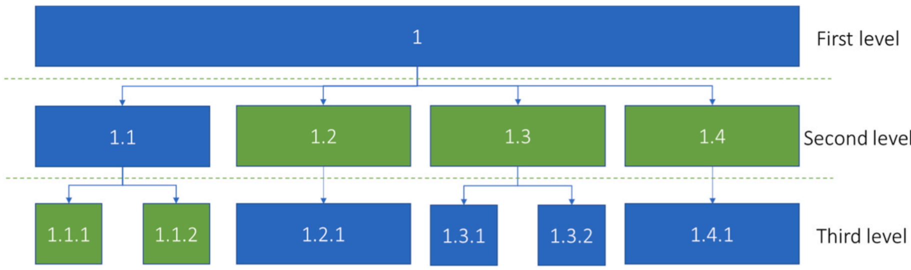


Fig. 8. Example of the hierarchical structure of clusters. Cost comparisons are performed among different levels to determine the final clustering solution (colored in green).

4.3. Hierarchical evaluation

The final step of the enhanced off-grid clustering performs a comprehensive cost comparison among the different levels of the hierarchical structure of clusters previously obtained in the exploratory clustering. The network cost estimator adjusted in section 4.2 is applied to determine the network cost of the potential mini-grids.

Fig. 8 shows an example of the hierarchical structure of clusters, where the combination of clusters that forms the final clustering solution is colored in green. To reach the final clustering solution, the electrification cost of cluster 1.1 (representing the least-cost option between a mini-grid or a combination of isolated systems for all its consumers) would be compared with the combined electrification costs of clusters 1.1.1 and 1.1.2. Similarly, the electrification costs of other clusters at the second level of the hierarchy are compared with the sum of the electrification costs of their corresponding clusters from the third level. These comparisons yield an off-grid electrification solution that could include a mix of clusters from the second and third levels. Finally, the total cost of this off-grid electrification solution is compared with the electrification cost of cluster 1, which represents the entire first-level cluster. The least-cost option among these alternatives becomes the final clustering solution.

The hierarchical evaluation selects a final clustering solution that may combine clusters from several levels stored in the hierarchical solution. Therefore, the final clustering solution may have an estimated electrification cost lower than the cost of every intermediate clustering solution (i.e., level) stored in the hierarchical structure.

It should be highlighted that enhanced off-grid clustering does not guarantee finding the global optimal solution. There are several parts of the algorithm where global optimality cannot be ensured. For instance, looping through the arcs of the Delaunay triangulation from the shortest to the longest may not be the optimal order to test whether nearby clusters should be merged. There may be intermediate clustering solutions that are not included in the hierarchical structure of clusters (further research should be devoted to determining which clustering solutions should be stored in the hierarchical structure of clusters), and the network cost estimator used may incur small inaccuracies when determining the network cost of a potential mini-grid.

Table 2 presents the primary differences between traditional off-grid clustering and enhanced off-grid clustering, highlighting the key strengths of the latter.

Table 2
Comparison between traditional off-grid and enhanced off-grid clusterings.

	Traditional off-grid clustering	Enhanced off-grid clustering
Space of solutions exploration	The algorithm terminates when it reaches the first local minimum	Performs a thorough exploration of the space of potential solutions
Cost evaluations	It uses inaccurate estimations of the network costs	It uses accurate estimations of all costs involved
Final clustering solution	The algorithm terminates when it reaches the first local minimum	It is the best combination of several intermediate solutions

5. Case study: Assessing the potential of smart clustering

This section demonstrates the effectiveness of enhanced off-grid clustering through several case studies. First, we apply the algorithm to a reference case located in Cajamarca. We also conduct several sensitivity analyses, varying input parameters to demonstrate that enhanced off-grid clustering systematically outperforms traditional off-grid clustering.

5.1. Reference case

To assess the potential benefits of enhanced off-grid clustering in a real-world setting, we apply the enhanced off-grid clustering proposed in this paper to the Cajamarca region in northern Peru. Cajamarca encompasses approximately 6700 consumers distributed across an area of about 400 km^2. Fig. 9 illustrates the location of the Cajamarca region within Peru. This section compares the results of the enhanced off-grid clustering to the outcomes of the traditional off-grid clustering. For this case study, REM does not consider grid extension solutions, as the comparison focuses exclusively on two algorithms that cluster consumers into potential off-grid systems, not grid extensions.

The consumer locations were manually extracted from satellite imagery using Google Earth. Fig. 10 shows the hourly demand profile employed in this case study, which aligns with the profile presented in Ref. [35]. This demand profile was derived by dividing the combined demand profile provided in Ref. [63] by the total number of consumers considered.

Table 3 shows the off-grid generation components used in the case study. The solar irradiance was obtained from Ref. [68], and the average diesel price is 0.5 USD/l.[1]

Table 4 shows the network components (lines) used in the case study. The catalogue of network lines is based on [35,69], which analyzed similar case studies. All the lines have a failure rate of 0.133 failures/yr, a preventive maintenance cost of 2.8 USD/(yr*km), and a corrective maintenance cost of 427 USD/failure.

Fig. 11 shows the electrification solution with the traditional off-grid and enhanced off-grid clustering. Both solutions appear very similar, but mini-grids tend to be more prominent in traditional off-grid clustering than in enhanced off-grid clustering. For example, we have highlighted a section of the solution in red. REM electrifies the highlighted consumers with one mini-grid when applying traditional off-grid clustering. Still, the model uses eight mini-grids and some standalone systems to electrify the same consumers when it applies enhanced off-grid clustering.

Fig. 12 shows the cumulative number of consumers per cluster for clusters with fewer than 250 consumers (we limit the number of consumers for the sake of clarity). The distributions are very similar.

[1] The average diesel price is similar to the ones we have considered in studies developed in African countries. Although this value may seem low, it is realistic. We should note that the diesel fuel could be subsidized, which could lead to a very inexpensive average diesel price.

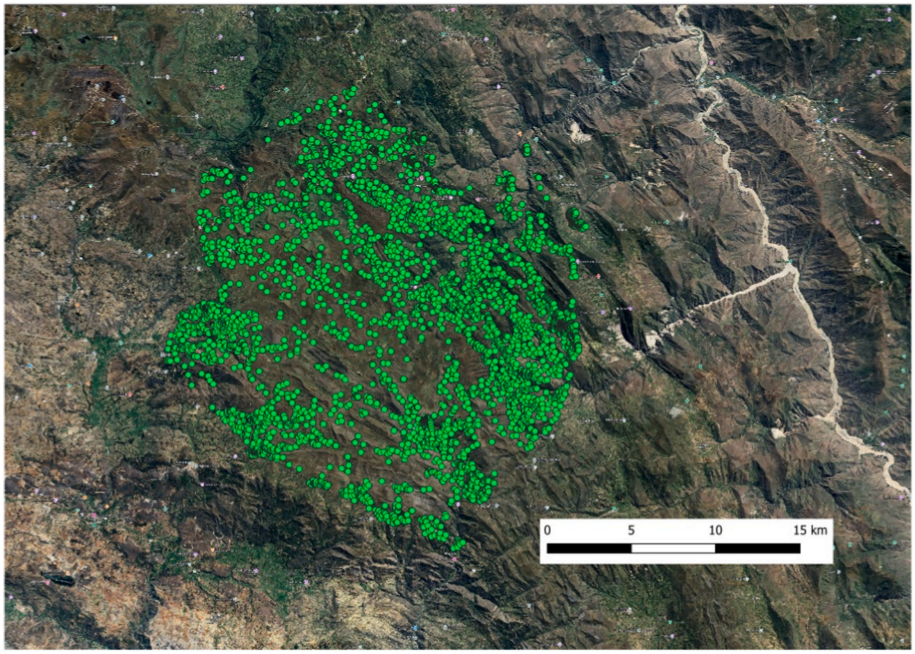


Fig. 9. Consumers of the case study, which are represented with green dots.

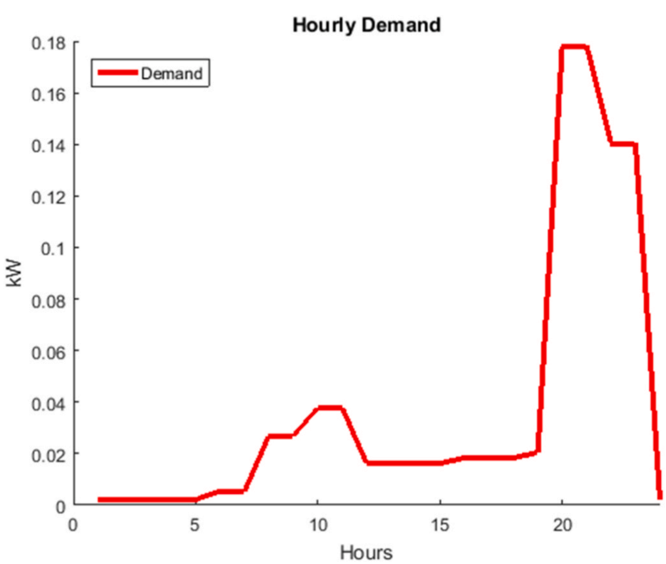


Fig. 10. Hourly demand profile. Source [35].

Table 3
Generation components for the off-grid systems.

Component	Model	Rated capacity (ies)	Rated voltage	Reference
Diesel generator	-	10 kW, 100 kW, 200 kW, 600 kW, and 1500 kW	-	The data are based on the expertise of the UEA Lab, field trips, and interviews
Solar panel	First Solar FS-497	97.5 W	-	[64]
Battery	Lead-acid Trojan L16RE	1021 AH	2 V	[65]
Inverter	Sunny boy 5000 TL	4.6 kW	-	[66]
Charge controller	Sunny island 6.0H	4.6 kW	-	[67]

Table 4
Network components (lines) for the mini-grids.

Name	R (Ω/km)	X (Ω/km)	Ampacity (A)	Investment Cost (USD/km)
Mole_s	20.37	1.58	22	1173
Gopher_s	8.41	1.41	38.33	2019
Weasel_s	6.99	1.38	43	2250
Ferret_s	5.21	1.32	51.67	2687
Weasel	1.16	0.23	129	6407
Ferret	0.87	0.22	155	7595
Rabbit	0.7	0.21	178	8544
Horse	0.5	0.19	225	10,500
Dog	0.35	0.19	271	12,285
Dingo	0.23	0.18	346	14,763
Lynex	0.2	0.17	384	15,872
Jaguar	0.18	0.17	411	16,577
Panther	0.17	0.17	420	16,800
Zebra	0.09	0.15	636	24,380

Approximately 8.25% of the clusters have 250 consumers or more with traditional off-grid clustering, whereas that figure drops to 1.49% with enhanced off-grid clustering (those clusters are not shown in Fig. 12). These numbers suggest that the traditional off-grid clustering is generating a few oversized clusters, which enhanced off-grid clustering replaces with smaller, more efficient clusters.

Table 5 shows the electrification costs obtained using traditional and enhanced off-grid clusterings. The final cost is 6.16% more expensive with traditional off-grid clustering than with enhanced off-grid clustering, representing a significant improvement that translates to 53,228 USD per year in savings in this case study of Cajamarca.

Table 6 presents the computation times required by REM to process the cases outlined in Table 5. These times were measured using a Lenovo computer equipped with 16 GB of RAM, a 12th Gen Intel(R) Core(TM) i5-1235U (1.30 GHz) processor, and running on Windows 11 Enterprise (64-bit).

The results show that enhanced off-grid clustering requires more computation time than traditional off-grid clustering. This is primarily due to the additional time needed to calculate accurate network designs for the mini-grids identified as representative by the network cost estimator, which accounts for 18 min and 12 s of the total computation time.

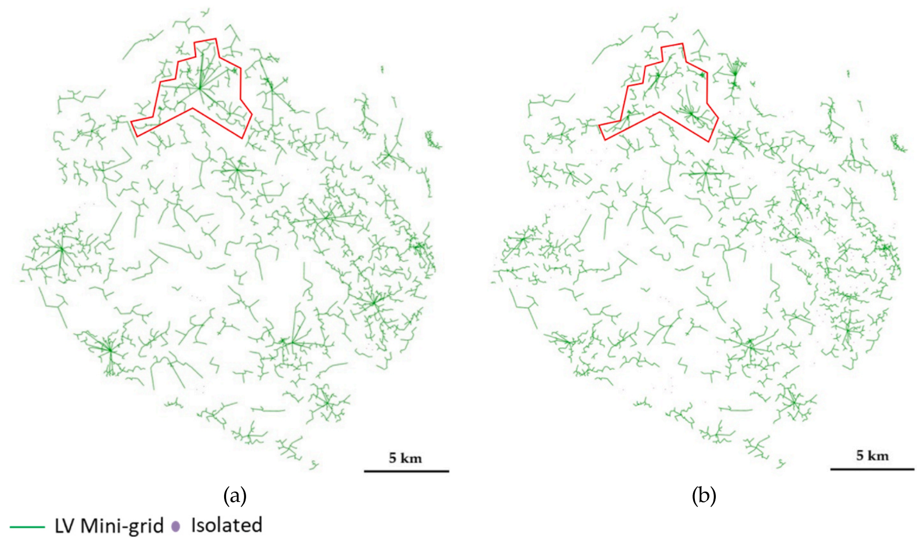


Fig. 11. Electrification solutions where candidate mini-grids are calculated with (a) the traditional off-grid clustering and (b) the enhanced off-grid clustering.

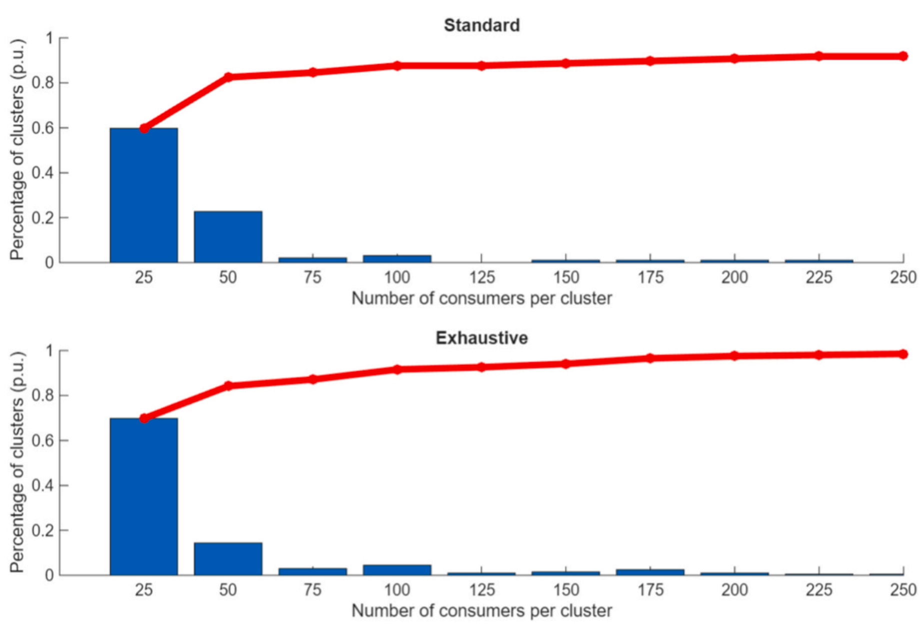


Fig. 12. Number of consumers per cluster for clusters of less than 250 consumers.

Table 5
Electrification solution summary for the two clustering algorithms. The last column contains the percentage increment between the "All" columns of the table.

	Consumers grouped with traditional off-grid clustering			Consumers grouped with enhanced off-grid clustering			Δ All (%)
	Mini-grids	Isolated	All	Mini-grids	Isolated	All	
Number of consumers	6666	22	6688	6582	106	6688	0.00
Fraction of consumers	1	0	1	0.98	0.02	1	0.00
CAPEX per consumer (USD/yr)	77.71	108.24	77.82	62.14	108.24	62.87	−19.21
OPEX per consumer (USD/yr)	50.6	167.46	50.99	56.07	167.46	57.83	13.41
CNSE per consumer (USD/yr)	0.31	0.85	0.32	0.46	0.85	0.46	43.75
Final cost per consumer (USD/yr)	128.63	276.54	129.12	118.66	276.54	121.16	−6.16
Total CAPEX (USD/yr)	518,046	2381	520,427	408,974	11,473	420,447	−19.21
Total OPEX (USD/yr)	337,330	3684	341,014	369,024	17,751	386,775	13.42
Total CNSE (USD/yr)	2093	19	2112	3013	90	3103	46.92
Final cost (USD/yr)	857,469	6084	863,553	781,011	29,314	810,325	−6.16
Fraction of demand served (p.u.)	0.999	0.998	0.999	0.999	0.998	0.999	0.00
Cost per kWh of demand served (USD/kWh)	0.378	0.804	0.379	0.348	0.804	0.356	−6.07

Table 6
Computation times of the cases in the format hours:minutes:seconds. The last column contains the percentage increment between the first and second columns of the table.

	Consumers grouped with traditional off-grid clustering	Consumers grouped with enhanced off-grid clustering	Δ (%)
Look-up table	1:16:45	1:16:45	0
Clustering	0:00:58	0:26:45	2646.42
Final designs	00:22:05	0:26:49	21.45

This extra step ensures greater accuracy in the clustering process, albeit at the cost of increased processing time. However, in all cases, the computation time is more than affordable for a project of these dimensions. Further analysis regarding the computational complexity of traditional and enhanced off-grid clusterings is provided in appendix C.

The network cost estimator described in Ref. [59] enables the enhanced off-grid clustering algorithm to track the estimated electrification costs of all intermediate solutions evaluated during the exploratory clustering, regardless of whether these solutions are stored in the hierarchical structure of clusters generated by the algorithm. This capability provides valuable insights into how electrification costs evolve as clusters merge.

Fig. 13 illustrates how the estimated electrification cost changes during the latter stages of the exploratory off-grid clustering. The blue line represents the connections activated by the traditional off-grid clustering, while the red line corresponds to additional connections introduced by the exploratory off-grid clustering to ensure a thorough exploration of the solution space.

Notably, when two large off-grid clusters are merged, the network cost of the resulting cluster can increase sharply due to constraints such as the maximum allowable voltage drop (i.e., this is what is likely happening in the last connections of Fig. 13). This phenomenon results in a significant increase in the estimated electrification cost, as indicated by the steep upward segments of the red curve in Fig. 13. These insights highlight the trade-offs involved in merging larger clusters and the importance of accurately estimating network costs in the decision-making process.

The estimated electrification cost at the end of the traditional off-grid clustering is 868,591 USD/yr, which is very close to the final cost provided in Table 5 (863,553 USD/yr, only 0.58% higher). The latter cost is calculated by designing detailed network layouts for all mini-grids using the RNM. This similarity suggests that the network cost estimator produces highly accurate predictions of the costs that RNM calculates in its final step.

Several configurations with estimated costs lower than 868,591 USD/yr are evaluated during the traditional off-grid clustering. For instance, after 6510 connections, the estimated cost is 855,902 USD/yr. However, the electrification cost achieved with enhanced off-grid clustering (810,325 USD/yr, as shown in Table 5) is significantly lower than the cost of any intermediate solutions evaluated during the exploratory off-grid clustering process. This indicates that the final solution provided by enhanced off-grid clustering is a combination of multiple levels within the hierarchical structure.

5.2. *Sensitivity analysis*

We perform a sensitivity analysis by modifying two input parameters of the case study. This sensitivity analysis provides insights into how these parameters influence the overall electrification costs and configurations.

The first parameter is the minimum number of consumers that a cluster must have in order to be electrified as a mini-grid (i.e., there can be no mini-grids with fewer consumers than this parameter). Planners sometimes want to avoid small mini-grids in their electrification strategy, and this constraint incorporates this point in the REM logic. This parameter ranges from 5 to 50 in the sensitivity analysis, analyzing the lower end of the mini-grid size restriction commonly used in electrification planning for underserved regions [34,50,70]. In South Saharan Africa, the bulk of mini-grid communities have between 50 and 1000 customers [70], while smaller communities are served by standalone systems. In the last mile of Latin American countries, especially in the Amazonian rainforest and Andean regions, smaller villages centralize productive and community services for nearby consumers, and smaller microgrids (usually diesel-powered in the past) are still common. This sensitivity analysis shows the impact of this restriction on the overall performance of the electrification costs.

The second input parameter is the use of a DC solar kit as an electrification solution for individual consumers. DC solar kits generally provide less energy than AC individual systems, but they tend to be less

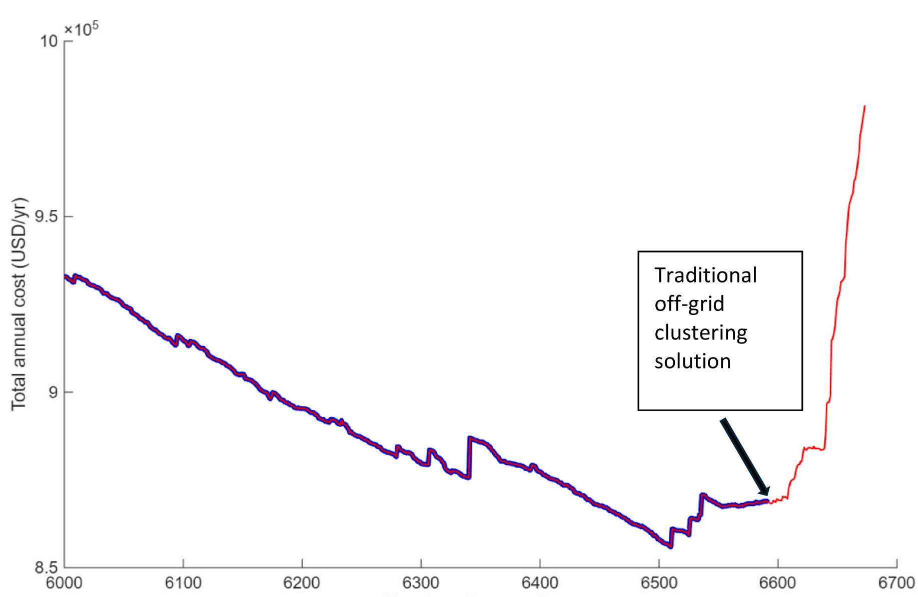


Fig. 13. Variation of the estimated electrification cost in the exploratory clustering.

Table 7
Parameters of the solar kit.

Parameter	Value
Investment and operation cost (USD/yr)	172.8
Management cost (USD/yr)	0
Per-consumer cost (USD/yr)	0
Solar capacity (kWp)	0.18
Battery capacity (kWh)	0.978
Lifetime (yr)	5
CNSE (USD/kWh)	0.98

Table 8
Final costs (USD/yr) obtained with traditional off-grid clustering.

		Minimum number of consumers for mini-grids				
		5	10	15	25	50
Solar kits	**No**	863,553	871,003	898,350	936,508	1,039,820
	Yes	859,966	860,045	863,700	869,489	893,749

Table 9
Final costs (USD/yr) obtained with enhanced off-grid clustering.

		Minimum number of consumers for mini-grids				
		5	10	15	25	50
Solar kits	**No**	810,325	822,289	855,375	892,018	961,569
	Yes	797,991	799,654	806,096	815,685	853,021

expensive. The characteristics of the solar kits used in this analysis –which come from the expertise of the UEA Lab in Latin America and are similar to other commercial solar kits [71]– are detailed in Table 7.

Table 8 presents the final electrification costs for several scenarios that combine the possibility of using DC solar kits with several minimum-number-of-consumers constraints for mini-grids (the case study shown in Table 5 does not allow solar kits and has a minimum number of consumers for mini-grids of 5). The final electrification cost includes investment and operational costs, as well as penalties for non-served energy. The clustering of consumers is calculated using the traditional off-grid clustering algorithm for all scenarios shown in Table 8.

As anticipated, the final costs increase when the minimum number of consumers required for a cluster to qualify as a mini-grid is raised. Additionally, the cost reductions resulting from introducing solar kits are more pronounced when the minimum number of consumers for a mini-grid is set higher. This reflects the greater flexibility of solar kits in scenarios where fewer consumers can be grouped into mini-grids.

Table 9 presents the final electrification costs achieved using the enhanced off-grid clustering algorithm for the same scenarios provided in Table 8. The results demonstrate that enhanced off-grid clustering consistently outperforms the traditional off-grid clustering across all scenarios. Specifically, the enhanced off-grid clustering reduces the electrification cost by approximately 4.5% to 7.5% compared to the traditional off-grid clustering, highlighting its effectiveness in identifying more cost-efficient solutions.

We can track the changes in estimated electrification costs as the exploratory clustering connects consumers across different scenarios (as we did for the base case in Fig. 13). Fig. 14 illustrates these changes under two scenarios: (a) solar kits are included, and the minimum number of consumers for mini-grids is 10, and (b) solar kits are excluded, and the minimum number of consumers for mini-grids is 50.

In both cases, the blue line represents the connections the traditional off-grid clustering algorithm performs. The red line corresponds to the additional connections introduced by the exploratory clustering to ensure a thorough exploration of the space of potential clustering solutions. This expanded exploration allows the algorithm to identify better solutions that the traditional off-grid clustering may not have considered.

In case (a), the traditional off-grid clustering activates an excessive number of connections, resulting in a clustering solution that is less optimal than some of the intermediate solutions explored during the process. This phenomenon mirrors what occurs in the base case but with a more pronounced effect. In case (b), the traditional off-grid clustering fails to activate certain connections that would lead to a better electrification solution, thereby missing opportunities to improve the overall outcome.

At this point, it is reasonable to wonder about how cost differences between the enhanced and traditional off-grid clusterings would behave if other sensitivity analyses were performed. Indeed, different variations in cost reductions could be obtained if different sensitivity analyses were performed, such as modifying demand profiles or introducing uncertainty in consumer locations. However, several methodological reasons support the general claim that enhanced off-grid clustering outperforms off-grid clustering.

Firstly, enhanced off-grid clustering obtains an accurate cost estimation for each cluster stored in the hierarchical structure, whereas traditional off-grid clustering lacks precise cost estimations for the clusters it evaluates. Secondly, enhanced off-grid clustering thoroughly explores the space of potential clustering solutions, whereas traditional off-grid clustering may reach a local minimum and fail to explore a significant portion of the solution space. Finally, enhanced off-grid clustering has the capability to provide a solution that combines clusters from multiple layers of the hierarchical structure of clusters, whereas traditional off-grid clustering is unable to provide a solution with such a combination of clusters.

To summarize, a clustering algorithm that thoroughly covers the

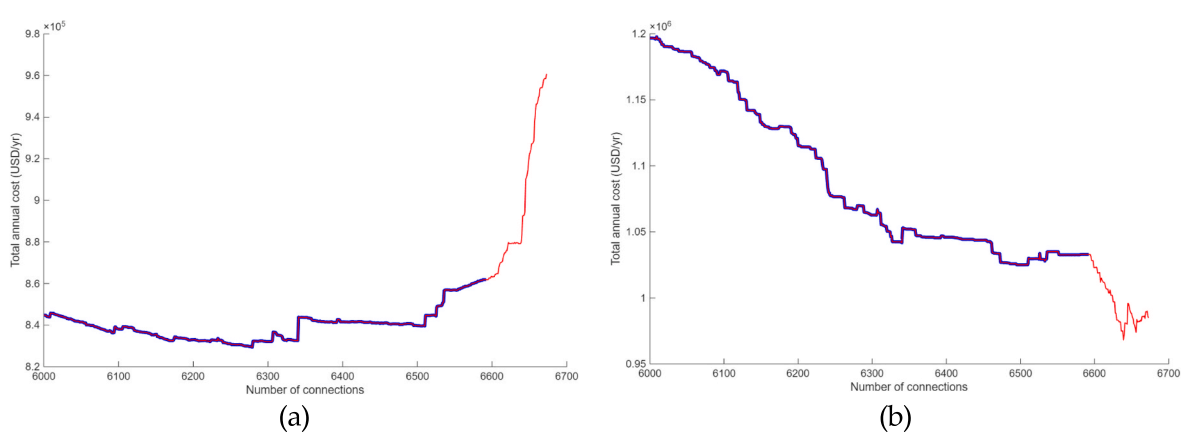


Fig. 14. Variation of the estimated electrification cost in the exploratory clustering (sensitivity analysis).

space of potential solutions, evaluates and combines many of them with accurate cost estimations to provide a final clustering solution, should yield a better performance than a clustering algorithm that may fail to go through the entire space of potential solutions, does not calculate accurate estimations of network costs, and cannot combine clusters from solutions that are separated in the space of clustering solutions.

5.3. Discussion

The results from Cajamarca suggest that improvements in clustering logic have significant economic impacts in real electrification programs. In the scenarios analyzed, the enhanced off-grid clustering reduces total electrification costs by 4.5–7.5%. Although this study focuses exclusively on off-grid systems and does not consider grid extensions—which play a crucial role in national electrification strategies—the magnitude of the savings is illustrative.

Large-scale electrification programs often involve hundreds of thousands or millions of consumers. For example, it was estimated that Rwanda needed 3 billion USD to electrify its entire population [72], so even a modest cost reduction in planning efficiency can have meaningful macroeconomic implications.

A fundamental part of the cost improvements achieved stems from incorporating accurate network cost estimates into the clustering algorithm. Distribution networks exhibit nonlinear behavior, governed by power flows and voltage drop constraints. The sharp increase in estimated costs observed in Fig. 13 during large-cluster mergers highlights how electrical constraints can create "cost cliffs" that purely geometric models fail to anticipate.

These results indicate that other rural electrification planning tools that rely primarily on geometric methods and do not explicitly incorporate detailed electrical constraints in their final network designs should consider upgrading their algorithms to address this. Otherwise, they may generate mini-grids and grid extensions that appear cost-effective with simplified modelling but require substantial redesign once electrical feasibility is enforced.

The methodological improvements presented in this paper come at the expense of additional computation time, and it is questionable whether the proposed algorithm could be applicable in a national-scale case with potentially hundreds of thousands of consumers.

In the Cajamarca case, the clustering time increased from less than 1 min under the traditional algorithm to approximately 27 min with the enhanced one. Most of this additional time is devoted to calibrating and applying the network cost estimator, whose computational complexity does not depend on the total number of consumers. Moreover, both traditional and enhanced clustering procedures exhibit the same computational complexity with respect to the number of consumers, implying similar scalability (further details regarding the computational complexity are provided in appendix C).

In practice, large electrification regions can be decomposed into subregions (i.e., districts or provinces) and solved in parallel, a strategy already applied in national-scale REM studies. This parallelization would mitigate computational burdens, facilitating the application of enhanced clustering to large territories. Therefore, we are optimistic about the application of the enhanced clustering in cases with more consumers, but additional tests are needed to reach firm conclusions on this point.

Overall, the enhanced clustering demonstrates that a thorough exploration of the solution space and improved representation of electrical constraints can yield tangible cost reductions without prohibitive computational penalties. These findings reinforce the importance of integrating accurate modelling into large-scale planning tools and imply that improvements to clustering can meaningfully enhance rural electrification plans.

6. Conclusions, limitations and future research lines

This section presents the conclusions, limitations, and further research lines. First, we describe the conclusions extracted from the application of the enhanced off-grid clustering to the reference case located in Cajamarca and the scenarios studied in the sensitivity analysis. Then, we describe the main limitations of the enhanced off-grid clustering and several research lines related to this algorithm.

6.1. Conclusions

This paper addresses a key research gap in rural electrification planning: the lack of cost-driven clustering algorithms capable of operating at the individual-consumer level while accounting for accurate cost calculations that incorporate economies of scale and discrete infrastructure components. The existing clustering algorithms typically rely on coarse spatial aggregation, simplified cost representations, or greedy heuristics prone to local optima, limiting their ability to produce cost-effective consumer aggregation into mini-grids.

To overcome these limitations, we developed an enhanced off-grid clustering algorithm that thoroughly explores the solution space of potential off-grid configurations. The method incorporates a refined cost estimator that captures nonlinearities in network designs and the effect of technical constraints such as voltage drops. By constructing a hierarchical structure of candidate clusters, the algorithm evaluates multiple configurations and selects the most cost-effective solution.

The algorithm is integrated into the Reference Electrification Model (REM) and validated in a real case study in Cajamarca (Peru), covering approximately 6700 consumers over 400 km^2 with heterogeneous geographic conditions derived from satellite imagery. Quantitatively, the enhanced algorithm reduced total electrification costs by 6.16% compared to the traditional REM off-grid clustering method. Sensitivity analyses, including scenarios with the availability of standalone solar kits and varying minimum mini-grid sizes, confirmed the robustness of the approach, with cost reductions consistently ranging from 4.5% to 7.5%.

Qualitatively, the algorithm produced smaller, more appropriately sized mini-grids and incorporated standalone systems where necessary, reflecting its ability to apply accurate cost estimates to optimize mini-grid sizes. Overall, these findings highlight the crucial role of accurate and flexible clustering algorithms in rural electrification planning. By addressing the key shortcomings of existing approaches, the proposed method represents a significant step forward in leveraging computational tools for sustainable energy access.

6.2. Limitations and future research lines

This subsection presents the main limitations of enhanced off-grid clustering, which are closely related to future research lines that should be developed for further improving the algorithm.

The first limitation is related to the cost margins that the exploratory clustering applies to ensure that the full space of potential clustering solutions is properly explored. Further analysis is required to determine whether the current cost margins applied are efficient or if other methods could provide better results.

Another limitation of enhanced off-grid clustering is related to determining which intermediate clustering solutions are stored in the hierarchical structure of layers. The current version of enhanced off-grid clustering stores a clustering solution each time that a certain number of clusters are merged, but this procedure may be inefficient. Further research should evaluate whether cost-based rules, such as storing a solution in the hierarchical structure if its cost is lower than the cost of all the current solutions in the hierarchical structure, could improve the results of enhanced off-grid clustering.

The main limitation of enhanced off-grid clustering is that it only groups consumers into potential mini-grids, without considering the possibility of grid extensions in the process. Additional research should be devoted to this point, and it may be necessary to extend the network cost estimator used by the enhanced off-grid clustering to handle extensions of the power grid.

Besides, the enhanced off-grid clustering is highly dependent on the accuracy of the network cost estimator. If the network cost estimator failed to provide accurate cost estimates, the enhanced off-grid clustering would misestimate the cost of clustering solutions, potentially leading to a significant gap between the estimated and actual network cost in the final electrification solution. This gap would have two direct consequences in the selection of clusters. On the one hand, the algorithm may select clusters whose real network cost is underestimated. On the other hand, the algorithm may avoid selecting clusters that should appear in the electrification solution because their real network costs are overestimated.

In addition, the computation time of enhanced off-grid clustering should be tested in cases with a larger number of consumers. The enhanced off-grid clustering required significantly more computation time than off-grid clustering for the case study analyzed in this paper. Further analysis is needed to investigate how computation time scales with the number of consumers.

Another future research point concerns the consideration of additional factors beyond costs (such as emissions). At the policy level, there was a swift move to significantly accelerate renewable deployment and improve long-term sustainability indicators [73] as structural transitions toward cleaner energy systems can enable economic growth while reducing emissions [74]. Studies that include several factors beyond techno-economics – such as emissions [75,76] and health effects [77]– have provided useful analyses and results. These studies underscore that achieving scalable, cost-effective rural electrification requires frameworks that integrate techno-economic optimization with broader sustainability objectives.

Ultimately, enhanced off-grid clustering relies on detailed, high-quality data to operate effectively. The algorithm requires the location of each individual consumer as well as their hourly demand profile to operate. A detailed catalog of components available for electrification is also required. The acquisition of accurate input data in underdeveloped countries is a challenging task, and the quality of the algorithm's outputs strongly depends on having reliable and accurate input datasets.

Author contributions

Conceptualization, P.C. and S.L.; Investigation, P.C., S.L. and A.G.-G.; Data curation: A.G.-G.; Methodology, P.C. and S.L.; Writing—original draft, P.C.; Writing—review & editing, S.L. and A.G.-G.; Software, P.C.; Resources: A.G.-G.; Visualization, P.C.; Validation, P.C.; Supervision, S. L.; All authors have read and agreed to the published version of the manuscript.

Declaration of generative AI and AI-assisted technologies in the manuscript preparation process

The authors have utilized ChatGPT to refine certain sections based on a specific outline provided by them, to enhance the flow of specific sections, and to gain insights from relevant papers. A GPT tool with the “Elements of style” was used to improve style with minimal changes based on the rules of this book. Grammarly was used to improve grammar and to simplify some convoluted expressions. The authors revised all these changes carefully and are responsible for the final version of the text.

Funding

This research received no external funding

Declaration of competing interest

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

Appendix A. Network cost estimator models for the case study

This appendix contains information regarding the network cost estimator used for the reference case and sensitivity analysis presented in section 5. The network cost estimator is applied as part of the enhanced off-grid clustering algorithm to ensure that accurate network costs are used in the hierarchical evaluation of clusters.

The network cost estimator applies a piecewise linear regression model to estimate the network cost of potential mini-grids. Depending on the mini-grid's MST length, a different linear regression model is used to estimate its network cost. Table A.1 shows which linear regression model would be used depending on the MST length of the mini-grid.

Table A.1
Mini-grid assignment to a linear regression model.

Assignation condition (meters)	Model
MST length <997.1	1
997.1 ≤ MST length ≤7410.4	2
7410.4 < MST length	3

The coefficients and explanatory variables of the first linear regression model are presented in Table A.2, which applies to mini-grids with an MST length of less than 997.1 m. The intercept variable refers to a constant term.

Table A.2
explanatory variables of the first linear regression model. The intercept variable refers to a constant term.

	Estimate	Standard Error	t Statistics	p value
(Intercept)	−0.889	0.938	−0.9478	0.348
MST length	204.7549	1.7197	119.0643	5.46E-61

The coefficients and explanatory variables of the second linear regression model are presented in Table A.3, which applies to mini-grids with an MST length between 997.1 and 7410.4 m.

Table A.3
explanatory variables of the second linear regression model. The intercept variable refers to a constant term. The variables $\mu_{2,0}$ and $\mu_{0,2}$ are the central electric moments of order 2 with respect to the coordinates x and y of the consumers, respectively.

	Estimate	Standard Error	t Statistics	p value
(Intercept)	66.8688	30.1841	2.2154	0.0306
MST length	161.9765	13.9779	11.588	7.73E-17
$\mu_{2,0}$	1.33E+05	1.70E+04	7.8431	1.01E-10
$\mu_{0,2}$	1.32E+05	1.39E+04	9.5102	1.63E-13

The coefficients and explanatory variables of the third linear regression model are presented in Table A.4, which applies to mini-grids with an MST length greater than 7410.4 m.

Table A.4
explanatory variables of the third linear regression model. The intercept variable refers to a constant term. The variables $\mu_{2,0}$ and $\mu_{0,2}$ are the central electric moments of order 2 with respect to the coordinates x and y of the consumers, respectively.

	Estimate	Standard Error	t Statistics	p value
(Intercept)	−2.41E+03	518.0993	−4.6577	2.47E-05
MST length	488.0854	25.7436	18.9595	3.29E-24
$\mu_{2,0}$	4.28E+04	3.14E+03	13.6366	2.65E-18
$\mu_{0,2}$	3.68E+04	3.81E+03	9.6595	6.28E-13

Appendix B. Pseudocodes

This appendix includes pseudocodes for three different algorithms. The first algorithm is the traditional off-grid clustering of REM, which is presented in section 3.2. The second and third algorithms are exploratory clustering and hierarchical evaluation of clusters, which are parts of the enhanced off-grid clustering and are introduced in sections 4.1 and 4.3, respectively.

Algorithm: traditional off-grid clustering

Inputs: Customer Coordinates, Customer Demand Profiles, Delaunay Triangulation, Generation Lookup Table, Network Catalog
Outputs: Activated Arcs of Delaunay Triangulation, Clusters

```
Delaunay.Sorted = sort(Delaunay)//Sort Delaunay arcs according to their lengths (from the shortest to the longest)
Delaunay.Unconnected = true//Initially, all Delaunay arcs are labeled as not connected
Clusters = Initialize(Customer_Coordinates, Customer_Demand_Profiles)//Initialize the initial clustering configuration (each consumer is its own isolated cluster)
Do while (New_Connections = = false)
        New_Connections = false
        for i in Delaunay.Sorted AND Delaunay.Unconnected//Loops through the unconnected arcs of the Delaunay triangulation (from the shortest to the longest)
                [Cluster_1, Cluster_2] = get clusters(i)//Get clusters connected by the arc of the Delaunay triangulation
                if(Cluster_1 = = Cluster_2)//Check if cluster 1 is the same as cluster 2
                        Delaunay.Unconnected[i] = false//Label arc i of Delaunay triangulation as connected
                        Continue//Go through the next arc of Delaunay triangulation
                end if
                //Calculate costs of configuration 1 (clusters electrified separately) and configuration 2 (clusters electrified together)
                [Cost_Config_1, Cost_Config_2] = calculate config cost(Cluster _1, Cluster_2, Network_Catalog, Generation_Lookup_Table)
                Low_Voltage_Minigrid = check low voltage(Cluster_1, Cluster_2)//Check if the combined clusters 1 and 2 could be electrified in LV
                if(Cost_Config_2<Cost_Config_1 AND Low_Voltage_Minigrid = = true)
                        Delaunay.Unconnected[i] = false//Label arc i of Delaunay triangulation as connected
                        Cluster_1 = merge(Cluster_1, Cluster_2)//Consumers from clusters 1 and 2 are now assigned to the same cluster
                        New_Connections = true
                        break//Restart evaluating potential connections between clusters, starting from the shortest inactive arcs of the Delaunay triangulation
                end if
        end for
end do while
```

Algorithm: exploratory clustering

Inputs: Customer Coordinates, Customer Demand Profiles, Delaunay Triangulation, Generation Lookup Table, Network Catalog, Cost Margin Nelements[1], Cost Margin Multiplier[2], Connections To Store[3]
Outputs: Hierarchical Structure of Clusters (HSoC), Activated Arcs of Delaunay Triangulation

```
Delaunay.Sorted = sort(Delaunay)//Sort Delaunay arcs according to their lengths (from the shortest to the longest)
Delaunay.Unconnected = true//Initially, all Delaunay arcs are labeled as not connected
Activated_Connections = 0//Initialize the number of activated connections of the Delaunay triangulation to zero
Cost_Margin = array[Cost Margin Nelements]//Create an array with the extra cost margin
Cost_Margin = 0//Initialize all elements of the extra cost margin array to zero
cm = 1//Initialize iterator for the Cost_Margin array
```

(continued on next page)

(*continued*)

Algorithm: exploratory clustering

```
Cost_Config_Diff = 0//Initialize variable that will store the maximum cost difference between configuration 2 (clusters electrified together) and configuration 1 (clusters electrified separately)
HSoC=Initialize(Customer Coordinates, Customer Demand Profiles)//Initialize the hierarchical structure with the initial clustering configuration (each consumer is its own isolated cluster)
Do while (cm < Cost Margin Nelements+1)//Algorithm terminates when the iterator of Cost_Margin array surpasses its maximum number of elements (Cost Margin Nelements)
    New_Connections = false
    for i in Delaunay.Sorted AND Delaunay.Unconnected//Loops through the unconnected arcs of the Delaunay triangulation (from the shortest to the longest)
        [Cluster_1, Cluster_2] = get clusters(i)//Get clusters connected by the arc/edge of the Delaunay triangulation
        if(Cluster_1 = = Cluster_2)//Check if cluster 1 is the same as cluster 2
            Delaunay.Unconnected[i] = false//Label arc i of Delaunay triangulation as connected
            Continue//Go through the next arc of Delaunay triangulation
        end if
        //Calculate costs of configuration 1 (clusters electrified separately) and configuration 2 (clusters electrified together)
        [Cost_Config_1, Cost_Config_2] = calculate config cost(Cluster _1, Cluster_2, Network_Catalog, Generation_Lookup_Table)
        Low_Voltage_Minigrid = check low voltage(Cluster_1, Cluster_2)//Check if the combined clusters 1 and 2 could be electrified in LV
        //Stores maximum value of Cost_Config_2 - Cost_Config_1 of all evaluated arcs of Delaunay triangulation
        Cost_Config_Diff = max(Cost_Config_Diff, Cost_Config_2-Cost_Config_1)
        if(Cost_Config_2<Cost_Config_1+Cost_Margin[cm] AND Low_Voltage_Minigrid = = true)
            Delaunay.Unconnected[i] = false//Label arc i of Delaunay triangulation as connected
            Cluster_1 = merge(Cluster_1, Cluster_2)//Consumers from clusters 1 and 2 are now assigned to the same cluster
            New_Connections = true
            Activated_Connections = Activated_Connections+1//Add one to the number of activated connections
            if(Activated_Connections IS MULTIPLE OF Connections To Store)
                HSoC = add(Clustering_Solution, HSoC)//Include current clustering solution in the
                hierarchical structure
            end if
            break//Restart evaluating potential connections between clusters, starting from the shortest inactive arcs of the Delaunay triangulation
        end if
    end for

    if(New_Connections = = false)//The extra cost margin is raised if no new connection is activated in a whole pass through unconnected arcs of the Delaunay triangulation
        if(cm = = 1)//Define second-to-last elements of the Cost_Margin array
            //Second-to-penultimate elements of Cost_Margin are logarithmically spaced between zero and Cost Margin Multiplier times the maximum cost
            Cost_Margin[2:Cost Margin Nelements-1] = logspace(0,Cost Margin Multiplier *Cost_Config_Diff)
            //The last element of Cost_Margin is set to infinite (or a very large value) to ensure that all connections are activated
            Cost_Margin[Cost Margin Nelements] = infinite
        elseif(cm = = Cost
        Margin Nelements)
            HSoC = add(Clustering_Solution, HSoC)//Include the current clustering solution in the hierarchical structure
        end if
        cm = cm+1//Increase the iterator of the Cost_Margin array
    end if
end do
while
```

[1] Cost Margin Nelements: parameter that determines the number of elements of the Cost_Margin array. In the cases shown in section 5, its value was set to 10.
[2] Cost Margin Multiplier: parameter that determines the intermediate values of the Cost_Margin array. In the cases shown in section 5, its value was set to 100.
[3] Connections To Store: parameter that determines the number of connections that need to be activated before a clustering solution is stored in the hierarchical structure. In the cases shown in section 5, its value was set to 100.

Algorithm: hierarchical evaluation

Inputs: Hierarchical structure of clusters (HSoC), Generation Lookup Table, Network Cost Estimator, Solar Kits
Outputs: Clustering Solution

```
For iLayer = 1 to iLayer = number of layers of HSoC
    For iCluster = 1 to iCluster = number of clusters in iLayer
        //Calculate cost of each cluster of the HSoC
        HSoC[iLayer, iCluster].Cost = get cluster cost(HSoC[iLayer, iCluster], Generation Lookup Table, Network Cost Estimator, Solar Kits)
    end for
end for
Clustering_Solution = HSoC[1]//Initialize clustering solution to the first layer of the HSoC
For iLayer = 2 to iLayer = number of layers of HSoC
    Above_Layer = HSoC[iLayer]
    Bottom_Layer = HSoC[iLayer-1]
    For iCluster = 1 to iCluster = number of clusters of Above_Layer
        Above_Cost = HSoC[iLayer, iCluster].Cost
        Bottom_Clusters = get clusters(HSoC, iLayer,iCluster)//Get clusters from iLayer-1 contained in iCluster from iLayer
        Bottom_Cost = sum(HSoC [iLayer-1,Bottom clusters])//Sum costs of all clusters from iLayer-1 contained in iCluster from iLayer
        if(Above_Cost ≤ Bottom_Cost)
            //Replace clusters from iLayer-1 with the cluster that contains them from iLayer
            Clustering_Solution = replace(Clustering_Solution, Bottom_Clusters, HSoC[iLayer, iCluster])
        else
```

(*continued on next page*)

(*continued*)

Algorithm: hierarchical evaluation

```
                //If the bottom-level clusters have a lower cost, then these clusters are the ones used for further comparisons in the HSoC
                HSoC[iLayer, iCluster] = HSoC[iLayer-1,Bottom_Clusters]
            end if
    end for
end for
```

Appendix C. Computational complexity

This appendix analyzes the computational complexity of traditional and enhanced off-grid clusterings, providing insights into how computation times scale for cases of different sizes.

C.1. Traditional off-grid clustering

The traditional off-grid clustering loops through the unconnected arcs of the Delaunay triangulation, from shortest to longest, and compares costs to determine whether it is worth electrifying the two clusters at the ends of the Delaunay triangulation arcs together.

Each time the algorithm connects an arc of the Delaunay triangulation (i.e., it determines that the clusters at its ends should be electrified together), it restarts the loop from the shortest unconnected arc of the Delaunay triangulation. The algorithm terminates when no arc of the Delaunay triangulation is connected in a complete loop.

To determine the computational complexity of traditional off-grid clustering, we consider the worst-case scenario (i.e., the scenario where the algorithm would make the maximum number of iterations through the arcs of the Delaunay triangulation). This scenario happens when each loop through the arcs of the Delaunay triangulation connects no arc but the last unconnected one.

In this scenario, the first loop would only connect the last arc of the Delaunay triangulation. Then, the traditional off-grid clustering would perform a second loop starting from the shortest unconnected arc of the Delaunay triangulation. In this second loop, the algorithm would only connect the second-to-last arc of the Delaunay triangulation. Then, the traditional off-grid clustering would perform a third loop that would only connect the third-to-last arc of the Delaunay triangulation and so on.

Therefore, the first loop would evaluate all the arcs of the Delaunay triangulation. The second loop would evaluate all the arcs of the Delaunay triangulation but the last one, which was connected in the first loop. The third loop would evaluate all the arcs of the Delaunay triangulation, except for the last two, which were connected in the first and second loops, and so on.

If n is the total number of arcs of the Delaunay triangulation, then the first loop evaluates n arcs, the second loop evaluates $n-1$ arcs, the third loop evaluates $n-2$ arcs, and so on until the last loop, which would only evaluate one arc. The total number of evaluated arcs would be $n+(n-1)+(n-2)+\ldots+1=\frac{n(n+1)}{2}$. This implies that the computational complexity of the traditional off-grid clustering is $O\left(\frac{n(n+1)}{2}\right)=O(n^2)$.

It is convenient to express the computational complexity in terms of the number of consumers rather than the number of arcs in the Delaunay triangulation, as we are comparing several algorithms. The number of arcs of a bidimensional Delaunay triangulation depends linearly on the number of its vertices [78]. Therefore, if m is the number of consumers in a case, then the computational complexity of the traditional off-grid clustering is $O(m^2)$.

C.2. Enhanced off-grid clustering

Enhanced off-grid clustering consists of three steps that operate sequentially; therefore, it is necessary to analyze them separately to draw conclusions regarding the computational complexity of the algorithm.

C.2.1. Exploratory clustering

The first step is exploratory clustering, which is similar to traditional off-grid clustering. The main difference is that the algorithm does not terminate if no arc is connected through a whole loop through the unconnected arcs of the Delaunay triangulation. Instead, the cost margin is increased, and the algorithm performs a new loop through the unconnected arcs of the Delaunay triangulation.

The worst-case scenario for enhanced off-grid clustering would be similar to the one analyzed for traditional off-grid clustering, but it would also include several loops through the unconnected arcs of the Delaunay triangulation where no loop is activated, thereby increasing the cost margin for additional loops.

Since the number of increases allowed for the cost margin is constant, then the computational complexity of the enhanced off-grid clustering is $O\left(\frac{n(n+1)}{2}+\text{Cost Margin Nelements}-1\right)=O(n^2)$, being n the number of arcs of the Delaunay triangulation and "Cost Margin Nelements" the number of elements of the cost margin array (this parameter is described in appendix B). This implies that the computational complexity of the exploratory clustering is $O(m^2)$, being m the number of consumers in a case.

C.2.2. Network cost estimation

The algorithm used for network cost estimation offers good performance in terms of computational time required for large cases. This procedure was introduced in Ref. [59], where it was applied to a case study with 1,598,842 consumers. The computation time needed for that case study was approximately 22 min –the specifications of the computer used are described in Ref. [59]–, which is similar to the time needed for the case presented in this paper (18 min and 12 s, the characteristics of the computer used are mentioned in section 5).

We can conclude that the network cost estimation requires a computation time that is approximately constant regardless of the number of consumers, so the computational complexity of this method is $O(1)$.

C.2.3. Hierarchical evaluation

The hierarchical evaluation step traverses the layers of the hierarchical structure of clusters, performing cost comparisons among them to determine the least-cost solution. The algorithm performs an initial nested loop to accurately calculate the cost of each cluster, and then it performs another nested loop to compare the costs among layers.

We will determine the computational complexity of this algorithm by evaluating the worst-case scenario (i.e., the one that requires the maximum number of iterations). The maximum number of clusters that a layer of the hierarchical cluster structure is equal to the number of consumers, and the maximum number of layers that the hierarchical structure of clusters could have depends linearly on the number of arcs of the Delaunay triangulation (i.e., one layer is included in the hierarchical structure of clusters each time that a number of connections that is a multiple of the parameter "Connections To Store" is activated; the parameter "Connections To Store" is described in appendix B).

Since the number of Delaunay arcs also depends linearly on the number of consumers, we can conclude that the computational complexity of the hierarchical evaluation is $O(m^2)$, where m is the number of consumers of the case.

C.3. Clustering comparison

Table C.1 summarizes the computational complexities calculated in this appendix in terms of the number of consumers m. The computational complexity of the traditional off-grid clustering is $O(m^2)$, and the computational complexity of the enhanced off-grid clustering is $O(m^2) + O(1) + O(m^2) = O(m^2)$.

Table C.1
Computational complexities of the traditional and enhanced off-grid clustering algorithms.

	Traditional Off-grid clustering	Enhanced off-grid clustering	
	$O(m^2)$	$O(m^2)$	(Exploratory clustering)
		$O(1)$	(Network cost estimation)
		$O(m^2)$	(Hierarchical evaluation)
Total complexity	$\boldsymbol{O(m^2)}$	$\boldsymbol{O(m^2)}$	

Both algorithms have an $O(m^2)$ complexity, which implies that the computing time would grow proportionally to the square of the number of consumers, and both algorithms should exhibit similar performances for a large case. In practical terms, the traditional off-grid clustering should always require fewer computation time than the enhanced off-grid clustering, because the exploratory clustering on its own involves more cost comparisons than the traditional off-grid clustering.

As the computational complexity of both clustering algorithms is $O(m^2)$, it is worth studying parallelization strategies that could ease the computational burden for large-scale cases that could include potentially millions of consumers. Although there is no straightforward way to parallelize the traditional and enhanced off-grid clustering algorithms, a large region can be split into several subregions (e.g., districts), and each subregion can be solved as an independent electrification problem. This parallelization strategy (i.e., splitting a large region into subregions) has been applied in REM to deal with national-scale electrification planning cases [50].

Appendix D. Additional sensitivity analyses

This appendix provides additional sensitivity analyses for several parameters used by the exploratory clustering. Specifically, we will study how the number of connections that need to be activated before a clustering solution is stored in the hierarchical structure (parameter “*Connections To Store*” of the pseudocode of the exploratory clustering in appendix B) affects the final electrification solution.

We will also study the impact of the cost margin parameters, performing a sensitivity analysis on the number of elements (parameter “*Cost Margin Nelements*” of the pseudocode of the exploratory clustering in appendix B) and its intermediate values (parameter “*Cost Margin Multiplier*” of the exploratory clustering in appendix B). Finally, we will analyze additional ways to create the cost margin (such as distributing its points in a linear rather than a logarithmic space) and conclude that the enhanced off-grid clustering systematically outperforms the traditional off-grid clustering.

D.1. Connections that need to be activated so that a clustering solution is stored in the hierarchical structure

The cases shown in section 5 include one solution in the hierarchical structure each time that 100 connections of the Delaunay triangulation are activated. In principle, decreasing this parameter should yield an improvement in the electrification solution, as more intermediate clustering solutions would be available to adjust the network cost estimator and to select when determining the final clustering solution. Conversely, increasing this parameter would reduce the number of intermediate clustering solutions, leading to worse final electrification solutions.

Table D.1 compares the traditional and enhanced clustering electrification costs for several values of the number of connections needed to store a solution in the hierarchical structure of clusters.

Table D.1
Comparison between traditional and enhanced off-grid clustering for several values of the connections needed to store a solution in the hierarchical structure of clusters.

Connections needed to store a solution in the hierarchical structure	Traditional off-grid clustering electrification cost (USD/yr)	Enhanced off-grid clustering electrification cost (USD/yr)	Δ Cost (%)
10	863,553	807,555	−6.48
25	863,553	808,810	−6.34
50	863,553	810,769	−6.11
100 (section 5 cases)	863,553	810,325	−6.16

(*continued on next page*)

Table D.1 (*continued*)

Connections needed to store a solution in the hierarchical structure	Traditional off-grid clustering electrification cost (USD/yr)	Enhanced off-grid clustering electrification cost (USD/yr)	Δ Cost (%)
150	863,553	818,393	−5.23
200	863,553	820,375	−5.00

The results in Table D.1 behave as expected, with the electrification cost decreasing as the number of connections to the store parameter decreases. The electrification cost for 50 connections to store is slightly higher than for 100 connections to store, which is probably due to differences between actual and expected network costs.

D.2. Cost margin parameters

The cost-margin mechanism is not intended to serve as a hyperparameter that “optimizes” the final solution, as in simulated annealing temperatures or learning rates. Its role is to provide a bounded relaxation mechanism that prevents premature termination of exploratory clustering and ensures that the algorithm generates a sufficiently diverse set of candidate cluster mergers to populate the hierarchical structure. More simply, the goal of the cost margin is to guarantee that no substantial portion of the space of potential clustering solutions is left unexplored; the particular numerical structure of the relaxation schedule is secondary, as long as it provides thorough coverage.

Considering the abovementioned role of the cost margin, conservative defaults that balance coverage and computational tractability were selected regarding the cases shown in section 5. Specifically, the cases shown in section 5 use a cost margin array with 10 points, with logarithmic spacing, and the maximum cost difference is multiplied by 100 to obtain the intermediate points of the array.

A 10-point cost margin provides a practical discretization that spans the relaxation range without incurring unnecessary computational overhead, and a 100-multiplier for the maximum cost difference provides a conservative upper bound that ensures that even relatively expensive candidate connections are included in the hierarchical structure before reaching the last element of the cost margin array.

Table D.2 shows a comparison among the traditional and enhanced off-grid clustering algorithms where the cost margin array has a different number of values. The results showcase that there are no significant variations among the electrification costs that the enhanced off-grid clustering provides.

Table D.2
Comparison between traditional and enhanced off-grid clustering for several values of the number of points in the cost margin array.

Number of points in the cost margin array	Traditional off-grid clustering electrification cost (USD/yr)	Enhanced off-grid clustering electrification cost (USD/yr)	Δ Cost (%)
5	863,553	812,378	−5.93
10 (section 5 cases)	863,553	810,325	−6.16
20	863,553	811,474	−6.03

The Cajamarca cases provided in section 5 multiply by 100 the maximum cost difference among cluster comparisons to determine the intermediate values of the cost margin.

Table D.3 compares traditional and enhanced off-grid clustering algorithms, where the cost margin array has a different cost difference. The enhanced off-grid clustering yields similar results across all values tested in the sensitivity analysis.

Table D.3
Comparison between traditional and enhanced off-grid clustering for several values of cost difference used for cost margin calculation.

Cost difference used for cost margin calculation	Traditional off-grid clustering electrification cost (USD/yr)	Enhanced off-grid clustering electrification cost (USD/yr)	Δ Cost (%)
50	863,553	810,417	−6.15
100 (section 5 cases)	863,553	810,325	−6.16
150	863,553	811,694	−6.01

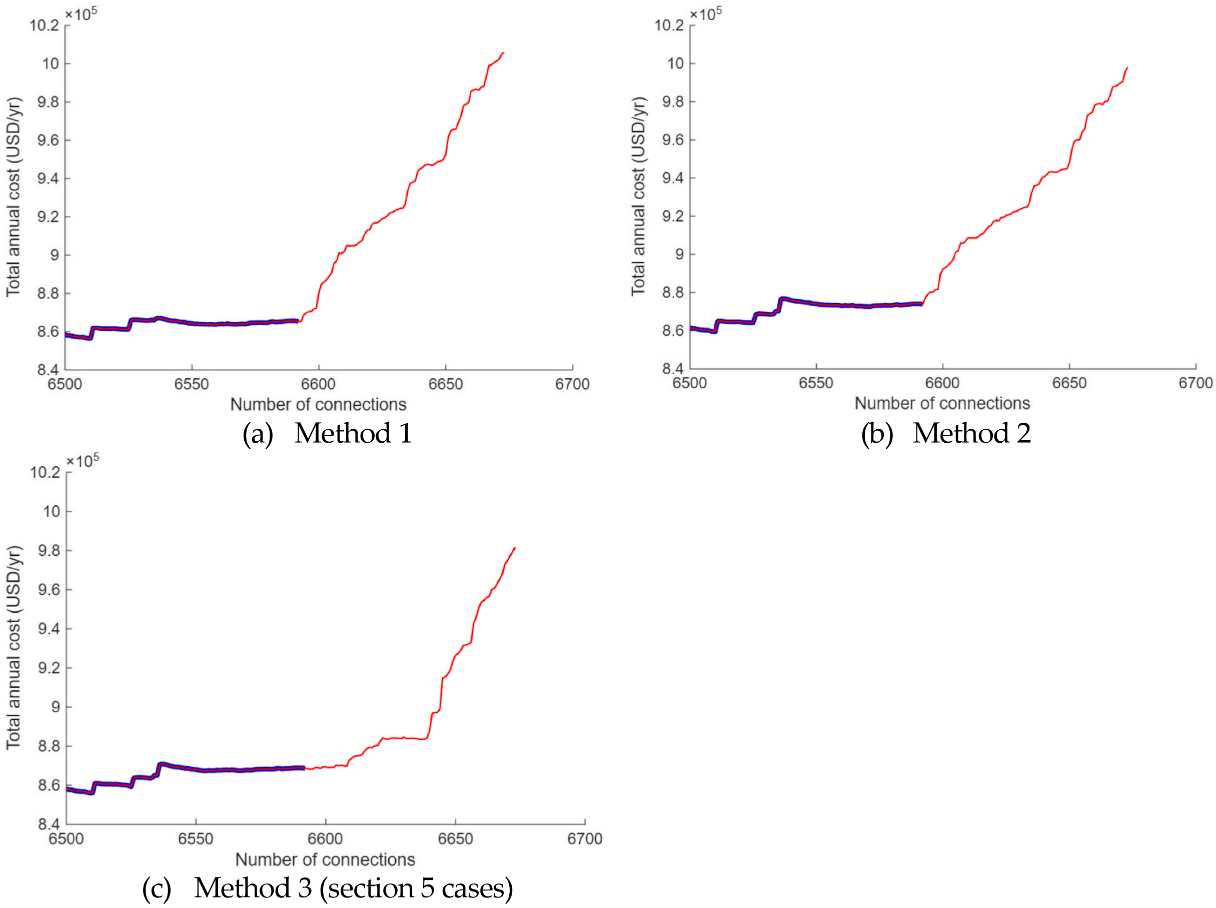
Finally, we will use two different methods, different from the one applied in the cases of section 5, to compute the cost margin. The first method we use computes the cost margin using a linearly spaced array rather than a logarithmically spaced one.

The second method we apply will activate all connections in the Delaunay triangulation in an additional pass, once all viable connections have been activated. This method is equivalent to setting the cost margin array to a two-element vector with the first element equal to 0 and the second element equal to infinity (or a very large value). The results of these methods are shown in Table D.4.

Table D.4
Comparison between traditional and enhanced off-grid clustering for several methods used to calculate the cost margin. Method 1 uses a linearly spaced array, method 2 forces connections in an additional pass, and method 3 is the one used in the cases presented in section 5.

Method used for cost margin calculation	Traditional off-grid clustering electrification cost (USD/yr)	Enhanced off-grid clustering electrification cost (USD/yr)	Δ Cost (%)
1	863,553	812,286	−5.94
2	863,553	812,403	−5.92
3 (section 5 cases)	863,553	810,325	−6.16

We can track changes in the estimated electrification costs as the exploratory clustering connects consumers (as we did in section 5 in Figs. 13 and 14). Figure D.1 illustrates these changes for the three methods provided in Table D.4.



(a) Method 1

(b) Method 2

(c) Method 3 (section 5 cases)

Fig. D.1. Variations in the estimated electrification costs in the enhanced off-grid clustering. Method 1 uses a linearly spaced array. Method 2 activates the remaining connections in an additional pass. Method 3 uses a logarithmically spaced array.

The most relevant part of the images shown in Figure D.1 corresponds to the part of the curves plotted in red, which corresponds to the connections that are activated when the cost margin is progressively increased. The red curves show that method 3, which corresponds to the logarithmically spaced cost margin, reaches an estimated electrification cost of 981,725 USD/yr at the end of the red curve. Methods 1 and 2 reach an estimated electrification cost of 1,005,830 USD/yr and 998,055 USD/yr, respectively. Therefore, the logarithmically spaced cost margin seems to provide slightly better results than the other methods tested in section 5.

Data availability

Data will be made available on request.

References

[1] International Energy Agency, World Energy Outlook 2024, 2024.

[2] I.A. Olanrele, A.I. Lawal, S.O. Dahunsi, A.A. Babajide, J.O. Iseolorunkanmi, The impact of access to electricity on education and health sectors in Nigeria's rural communities, Entrep. Sustain. Issues 7 (30) (2020) 3016–3035, https://doi.org/10.9770/jesi.2020.7.4.

[3] Our World in Data, Ensure Access to Affordable, Reliable, Sustainable and Modern Energy for all, Our World Data, 2023 [WWW Document], https://www.undp.org/energy/our-work-areas/energy-access/access-electricity. (Accessed 22 October 2024).

[4] N. Duan, The quest for balance: the journey of energy, environment and sustainable development, Energy Environ. Sustain 1 (2025) 100001, https://doi.org/10.1016/j.eesus.2024.12.001.

[5] G. Li, J. Luo, S. Liu, Performance evaluation of economic relocation effect for environmental non-governmental organizations: evidence from China, Economics 18 (2024) 20220080, https://doi.org/10.1515/econ-2022-0080.

[6] Z. Xu, Harmony between nature and humanity: a systematic environmental governance approach to address climate change, Energy Environ. Sustain. 1 (2025) 100002, https://doi.org/10.1016/j.eesus.2025.01.001.

[7] Sustainable Energy Hub, United nations development programme [WWW Document]. U. N. Dev. Programme, https://www.undp.org/energy/our-work-areas/energy-access/access-electricity, 2024. (Accessed 22 October 2024).

[8] Femi Oluwatoyin Omole, Oladiran Kayode Olajiga, Tosin Michael Olatunde, Challenges and successes in rural electrification: a review of global policies and case studies, Eng. Sci. Technol. J. 5 (2024) 1031–1046, https://doi.org/10.51594/estj.v5i3.956.

[9] Y. Dai, X. Tong, X. Jia, Executives' legal expertise and corporate innovation, Corp. Govern. Int. Rev. 32 (2024) 954–983, https://doi.org/10.1111/corg.12578.

[10] T. Ji, S. Wang, J. Wang, J. Zhang, Beyond technology: role of technology-organization-environment-human factors in autonomous driving adoption for new energy vehicles, J. Retailing Consum. Serv. 88 (2026) 104552, https://doi.org/10.1016/j.jretconser.2025.104552.

[11] J. Zhang, J. Li, Y. Yan, Z. Xie, Concentration in cross-border research collaborations and MNCs' knowledge creation in a host country, Strateg. Manag. J. 47 (2026) 555–582, https://doi.org/10.1002/smj.70025.

[12] S. Jiang, X. Yuan, Does green credit promote real economic development? Dual empirical evidence of scale and efficiency, PLoS One 20 (2025) 1–26, https://doi.org/10.1371/journal.pone.0326961.

[13] P. Ciller, S. Lumbreras, Electricity for all: the contribution of large-scale planning tools to the energy-access problem, Renew. Sustain. Energy Rev. 120 (2020) 109624, https://doi.org/10.1016/j.rser.2019.109624.

[14] J. Morrissey, Achieving universal electricity access at the lowest cost: a comparison of published model results, Energy Sustain. Dev. 53 (2019) 81–96, https://doi.org/10.1016/j.esd.2019.09.005.

[15] R. Luna-Rubio, M. Trejo-Perea, D. Vargas-Vázquez, G.J. Ríos-Moreno, Optimal sizing of renewable hybrids energy systems: A review of methodologies, Solar Energy 86 (4) (2012) 1077–1088, https://doi.org/10.1016/j.solener.2011.10.016. https://linkinghub.elsevier.com/retrieve/pii/S0038092X11003835.

[16] Sunanda Sinha, S.S. Chandel, Review of software tools for hybrid renewable energy systems, Renewable and Sustainable Energy Reviews 32 (2014) 192–205, https://doi.org/10.1016/j.rser.2014.01.035. https://linkinghub.elsevier.com/retrieve/pii/S136403211400046X.

[17] M. Abdoos, A. Shahee, H. Salaripoor, R. Zahedi, H. Yousefi, A.A. Sohrabi, Assessing technical, economic, and environmental impacts of solar-powered wastewater treatment with battery storage, Energy Convers. Manag. X 28 (2025) 101355, https://doi.org/10.1016/j.ecmx.2025.101355.
[18] M.H. Gholizadeh, H. Yousefi, A. Hajinezhad, M. Abdoos, Optimization of the economic-technical model for hydrogen production with an approach to utilizing solar power plants and waste-to-energy conversion, Fuel Commun 24 (2025) 100144, https://doi.org/10.1016/j.jfueco.2025.100144.
[19] Q. Meng, S. Hussain, F. Luo, Z. Wang, X. Jin, An online reinforcement learning-based energy management strategy for microgrids with centralized control, IEEE Trans. Ind. Appl. 61 (2025) 1501–1510, https://doi.org/10.1109/TIA.2024.3430264.
[20] Q. Meng, X. Jin, F. Luo, Z. Wang, S. Hussain, Distributionally robust scheduling for benefit allocation in regional integrated energy system with multiple stakeholders, J. Mod. Power Syst. Clean Energy 12 (2024) 1631–1642, https://doi.org/10.35833/MPCE.2023.000661.
[21] P.S. Georgilakis, N.D. Hatziargyriou, A review of power distribution planning in the modern power systems era: models, methods and future research, Elec. Power Syst. Res. 121 (2015) 89–100, https://doi.org/10.1016/j.epsr.2014.12.010.
[22] D.L. Huff, A probabilistic analysis of shopping center trade areas, Land Econ. 39 (1963) 81–90, https://doi.org/10.2307/3144521.
[23] A. Korkovelos, B. Khavari, A. Sahlberg, M. Howells, C. Arderne, The role of open access data in geospatial electrification planning and the achievement of SDG7. An OnSSET-Based case study for Malawi, Energies 12 (2019) 1395, https://doi.org/10.3390/en12071395.
[24] P. Blechinger, C. Cader, P. Bertheau, Least-cost electrification modeling and Planning–A case study for five Nigerian federal states, Proc. IEEE 107 (2019) 1923–1940, https://doi.org/10.1109/JPROC.2019.2924644.
[25] U. Akpan, Technology options for increasing electricity access in areas with low electricity access rate in Nigeria, Socioecon. Plann. Sci. 51 (2015) 1–12, https://doi.org/10.1016/j.seps.2015.05.001.
[26] F. Kemausuor, E. Adkins, I. Adu-Poku, A. Brew-Hammond, V. Modi, Electrification planning using network planner tool: the case of Ghana, Energy Sustain. Dev. 19 (2014) 92–101, https://doi.org/10.1016/j.esd.2013.12.009.
[27] A. Sahlberg, A. Korkovelos, C. Kabongo, C. Trujillo, B. Khavari, F. Fuso Nerini, Attention to detail: exploring effects of model resolution and complexity in geospatial electrification modelling, Discov. Energy 6 (2026) 8, https://doi.org/10.1007/s43937-025-00117-0.
[28] A. Sahlberg, A. Korkovelos, C. Kabongo, C. Trujillo, B. Khavari, F.F. Nerini, Attention to detail – exploring effects on technology selection in geospatial electrification modelling (preprint), Review (2023), https://doi.org/10.21203/rs.3.rs-3043251/v1.
[29] S. Corigliano, T. Carnovali, D. Edeme, M. Merlo, Holistic geospatial data-based procedure for electric network design and least-cost energy strategy, Energy Sustain. Dev. 58 (2020) 1–15, https://doi.org/10.1016/j.esd.2020.06.008.
[30] G.T.F. Vinicius, C. Silvia, D. Aleksandar, B. Massimo, M. Marco, Rural electrification planning based on graph theory and geospatial data: a realistic topology oriented approach, Sustain. Energy Grids Netw. 28 (2021) 100525, https://doi.org/10.1016/j.segan.2021.100525.
[31] M. Ester, H.-P. Kriegel, J. Sander, X. Xu, A density-based algorithm for discovering clusters in large spatial databases with noise, Proc. Second Int. Conf. Knowl. Discov. Data Min. (1996) 226–231.
[32] M. Torres-Pérez, J. Domínguez, L. Arribas, J. Amador, P. Ciller, A. González-García, A geospatial clustering algorithm and its integration into a techno-economic rural electrification planning model, Eng. Appl. Artif. Intell. 137 (2024) 109249, https://doi.org/10.1016/j.engappai.2024.109249.
[33] A. Leonard, S. Wheeler, M. McCulloch, Geospatial clustering and network design for rural electrification in Africa, in: 2020 IEEE PES/IAS Powerafrica. Presented at the 2020 IEEE PES/IAS Powerafrica, IEEE, Nairobi, Kenya, 2020, pp. 1–5, https://doi.org/10.1109/PowerAfrica49420.2020.9219908.
[34] P. Ciller, D. Ellman, C. Vergara, A. Gonzalez-Garcia, S.J. Lee, C. Drouin, M. Brusnahan, Y. Borofsky, C. Mateo, R. Amatya, R. Palacios, R. Stoner, F. de Cuadra, I. Perez-Arriaga, Optimal electrification planning incorporating On- and off-grid technologies: the reference electrification model (REM), Proc. IEEE 107 (2019) 1872–1905, https://doi.org/10.1109/JPROC.2019.2922543.
[35] P. Ciller, F. de Cuadra, S. Lumbreras, Optimizing off-grid generation in large-scale electrification-planning problems: a direct-search approach, Energies 12 (2019) 4634, https://doi.org/10.3390/en12244634.
[36] Waya Energy, Waya energy projects [WWW Document], https://waya-energy.com/projects, 2024. (Accessed 29 October 2024).
[37] IIT-Comillas, Assistance to the government of Rwanda to develop a national electrification strategy (NES) and related investment plans [WWW Document], https://www.iit.comillas.edu/proyectos/mostrar_proyecto.php.es?nombre_abreviado=Rwanda%20NES%2017, 2017. (Accessed 4 April 2018).
[38] IIT-Comillas, Mozambique geospatial options analysis towards universal electrification [WWW Document], https://www.iit.comillas.edu/proyectos/mostrar_proyecto.php.en?nombre_abreviado=MIT-NEP-MOZAMBIQUE-2018, 2019. (Accessed 27 September 2020).
[39] IIT-Comillas, Least-cost electrification plan of Maluku-Papua region of Indonesia [WWW Document], https://www.iit.comillas.edu/proyectos/mostrar_proyecto.php.es?nombre_abreviado=AsianDB_Indonesia, 2019. (Accessed 19 December 2020).
[40] IIT-Comillas, Development of the universal access to electricity strategy in Bolivia, based on a geo-referenced electricity access plan [WWW Document], https://www.iit.comillas.edu/proyectos/mostrar_proyecto.php.es?nombre_abreviado=BOLIVIA-PAUNE, 2022. (Accessed 16 April 2022).
[41] IIT-Comillas, Haiti national electrification analysis - consultancy services for the assessment of electrification options [WWW Document], https://www.iit.comillas.edu/publicacion/proyecto/es/HAITI-PAUNE/Haiti_National_electrification_analysis_-_Consultancy_services_for_the_assessment_of_electrification_options, 2023. (Accessed 29 October 2024).
[42] IIT-Comillas, Development of the national electrification plan of Paraguay [WWW Document], https://www.iit.comillas.edu/publicacion/proyecto/es/PARAGUAY-PAUNE/Elaboraci%c3%b3n_del_plan_nacional_de_electrificaci%c3%b3n_de_Paraguay, 2024. (Accessed 29 October 2024).
[43] B. Khavari, A. Korkovelos, A. Sahlberg, M. Howells, F. Fuso Nerini, Population cluster data to assess the urban-rural split and electrification in Sub-Saharan Africa, Sci. Data 8 (2021) 117, https://doi.org/10.1038/s41597-021-00897-9.
[44] D. Mentis, M. Howells, H. Rogner, A. Korkovelos, C. Arderne, E. Zepeda, S. Siyal, C. Taliotis, M. Bazilian, A. de Roo, Y. Tanvez, A. Oudalov, E. Scholtz, Lighting the world: the first application of an open source, spatial electrification tool (OnSSET) on Sub-Saharan Africa, Environ. Res. Lett. 12 (2017) 085003, https://doi.org/10.1088/1748-9326/aa7b29.
[45] L. Romero Otero, Sistemas De Información Geográfica Y Electrificación Rural. Análisis, Desarrollo Y Estudio De Caso Con Intigis. Master Thesis, Universidad Complutense. Faculty of Geography and History, Madrid, 2016 (in Spanish).
[46] I. Pinedo Pascua, Intigis: Propuesta Metodológica Para La Evaluación De Alternativas De Electrificación Rural Basada En SIG, Universidad Politécnica de Madrid. School of Agricultural Engineering, Madrid, 2012. PhD thesis (in Spanish).
[47] Innovation Energie Developpement, CAP-REDEO Results Report, 2010.
[48] Innovation Energie Developpement, Geosim [WWW Document], https://www.ied-sa.com/es/planificacion/geosim.html, 2021. (Accessed 22 February 2026).
[49] United Nations Development Programme, Proposed global roadmap shows how universal access to sustainable energy can be achieved by 2030, transforming the lives of millions [WWW Document]. U. N. Dev. Programme, https://www.undp.org/press-releases/proposed-global-roadmap-shows-how-universal-access-sustainable-energy-can-be-achieved-2030-transforming-lives-millions, 2021. (Accessed 1 October 2026).
[50] Rwanda Energy Group (REG), The National Electrification Plan: Report on Definition of Technologies (On-grid and off-grid) at Village Level, 2019.
[51] Andrés González García, Kristin Dietrich, Pedro Ciller Cutillas, Clara Pérez-Andújar, Rafael Palacios, Plan nacional de electrificación de última milla en Paraguay, Banco Interamericano de Desarrollo División de Energía (2025).
[52] P. Ciller Cutillas, Clustering-Related Improvements in the Reference Electrification Model, Master thesis, Universidad Pontificia Comillas, Madrid, 2016. School of Engineering.
[53] P. Ciller, The Rural Electrification Planning Problem: Strategies and Solutions, Universidad Pontificia Comillas, Madrid, 2021. Ph.D. thesis.
[54] C. Mateo Domingo, T. Gómez San Román, Á. Sanchez-Miralles, J.P. Peco Gonzalez, A. Candela Martinez, A reference network model for large-scale distribution planning with automatic street map generation, IEEE Trans. Power Syst. 26 (2011) 190–197, https://doi.org/10.1109/TPWRS.2010.2052077.
[55] C. Drouin, Geospatial Cost Drivers in computer-aided Electrification Planning: the Case of Rwanda, Massachusetts Institute of Technology. Department of Mechanical Engineering, 2018. Master thesis.
[56] J. Peco, Modelo De Cobertura Geográfica De Una Red De Distribución De Energía Eléctrica, PhD thesis (in Spanish), Universidad Pontificia Comillas, Madrid, 2001. School of Engineering.
[57] A. Navarro, H. Rudnick, Large-Scale distribution planning — part II : macro-optimization with voronoi's diagram and tabu search, IEEE Trans. Power Syst. 24 (2009) 752–758, https://doi.org/10.1109/TPWRS.2009.2016594.
[58] A.B. Birchfield, T. Xu, K.M. Gegner, K.S. Shetye, T.J. Overbye, Grid structural characteristics as validation criteria for synthetic networks, IEEE Trans. Power Syst. 32 (2017) 3258–3265, https://doi.org/10.1109/TPWRS.2016.2616385.
[59] P. Ciller, S. Lumbreras, A. González-García, Network cost estimation for mini-grids in large-scale rural electrification planning, Energies 14 (2021) 7382, https://doi.org/10.3390/en14217382.
[60] C. Mateo, F. Postigo, F. de Cuadra, T.G.S. Roman, T. Elgindy, P. Dueñas, B.-M. Hodge, V. Krishnan, B. Palmintier, Building large-scale U.S. synthetic electric distribution System models, IEEE Trans. Smart Grid 11 (2020) 5301–5313, https://doi.org/10.1109/TSG.2020.3001495.
[61] L. Kaufman, P.J. Rousseeuw, Finding Groups in Data: an Introduction to Cluster Analysis, Wiley Series in Probability and Mathematical Statistics, Wiley, Hoboken, N.J, 1990.
[62] E.J. Pedhazur, Multiple Regression in Behavioral Research: Explanation and Prediction, third ed., Harcourt Brace College Publishers, Forth Worth, 1997.
[63] K. Villanueva Saberbein, L. Aye, Technical and financial feasibility of a stand-alone photovoltaic system for rural electrification in the andean south region of Peru, J. Sustain. Dev. 5 (2012), https://doi.org/10.5539/jsd.v5n11p32.
[64] First Solar, First solar FS-497 datasheet [WWW Document], http://www.firstsolar.com/-/media/First-Solar/Documents/PD-5-401-04_Series_4_Datasheet.ashx, 2017. (Accessed 3 February 2017).
[65] Trojan Battery Company, Trojan L16RE2V Datasheet, 2017 [WWW Document], http://www.trojanbattery.com/pdf/datasheets/L16RE2V_TrojanRE_Data_Sheets.pdf. (Accessed 3 February 2017).
[66] SMA Solar Technology AG, Sunny Boy 5000TL Datasheet, 2017 [WWW Document], http://files.sma.de/dl/15330/SB5000TL-21-DEN1551-V20web.pdf. (Accessed 3 February 2017).
[67] SMA Solar Technology AG, Sunny Island 6.0H datasheet [WWW Document], http://files.sma.de/dl/2485/SI_6H8H-AEN131411W.pdf, 2017. (Accessed 3 February 2017).

[68] NREL, Pvwatts Calculator, 2017 [WWW Document], http://pvwatts.nrel.gov/. (Accessed 3 February 2017).
[69] A. Gonzalez-Garcia, R. Amatya, R. Stoner, J.I. Perez-Arriaga, Evaluation of Universal Access to Modern Energy Services in Peru. Case Study of Scenarios for Electricity Access in Cajamarca, Enel Foundation Working Papers, 2016.
[70] A. González-García, P. Ciller, S. Lee, R. Palacios, F. de Cuadra García, J.I. Pérez-Arriaga, A rising role for decentralized solar minigrids in integrated rural electrification planning? Large-Scale, Least-cost, and customer-wise design of grid and off-grid supply systems in Uganda, Energies 15 (2022), https://doi.org/10.3390/en15134517.
[71] Eco Worthy, 200W 12V complete off grid solar kit [WWW Document]. My Store, https://ecoworthysolar.com/products/200w-off-grid-solar-kits-with-100w-mono-solar-panel, 2025. (Accessed 10 August 2025).
[72] C. de Abajo, S. Díaz-Pastor, A. González, I. Pérez-Arriaga, A Business Plan to Achieve Full Electrification in Rwanda Under the Integrated Distribution Framework (IDF), 2021.
[73] S. Mohammaddini, A. Saifoddin, A. Zahedi, M. Abdoos, A. Hajinezhad, Impact of energy taxes and eco-innovation on sustainable energy markets, Clean. Eng. Technol. 29 (2025) 101097, https://doi.org/10.1016/j.clet.2025.101097.
[74] A.N. Gazik, H. Yousefi, M. Abdoos, Sustainability under the threat: the nexus of economic growth, energy consumption, and CO2 emission during COVID-19 pandemic (case study: united Kingdom), Environ. Chall. 20 (2025) 101285, https://doi.org/10.1016/j.envc.2025.101285.
[75] M.K. Aghababaei, A. Saifoddin, A. Zahedi, M. Abdoos, Macroeconomic impact of energy transition: a comparative study of developed and developing countries, Energy Strategy Rev. 62 (2025) 101910, https://doi.org/10.1016/j.esr.2025.101910.
[76] M.H. Ghodusinejad, M. Abdoos, H. Yousefi, Optimal design of hybrid cooling/heating system using TLBO: Techno-economic evaluation for office buildings, Energy Rep. 15 (2026) 109075, https://doi.org/10.1016/j.egyr.2026.109075.
[77] T.A. Adagunodo, A.A. Aremu, O.G. Bayowa, A.I. Ojoawo, A.O. Adewoye, T. E. Olonade, Assessment and health effects of radon and its relation with some parameters in groundwater sources from shallow aquifers in granitic terrains, southeastern axis of Ibadan, Nigeria, Groundw. Sustain. Dev. 21 (2023) 100930, https://doi.org/10.1016/j.gsd.2023.100930.
[78] Elisha Sacks, Delaunay triangulation [WWW Document], https://www.cs.purdue.edu/homes/cs53100/slides/delaunay.pdf, 2023. (Accessed 1 December 2026).