Clustering-related improvements in the Reference Electrification Model

Pedro Ciller Cutillas
Institute for Research in Technology (IIT)
ICAI School of Engineering, Comillas Pontifical University
Madrid, Spain
pciller@comillas.edu

***Abstract*— The Reference Electrification Model is a large-scale rural electrification planning tool that decides the best electrification mode – stand-alone system, microgrid or grid extension – for each customer. It has already been applied to real case studies such as Bihar (India) and Cajamarca (Peru). The aim of this paper is to describe several improvements that are related to the clustering process used in the Reference Electrification Model. These enhancements can be divided into three different groups: electrical model improvements, improvements to the logic of the algorithm and improvements to the interpolation process. A real case study is performed using data from Cajamarca (Peru). Results show a significant improvement compared to the original clustering process, being the final solution approximately 40% less expensive than the one obtained before implementing the improvements.**

***Keywords*—Rural electrification, clustering, Reference Electrification Model, Cajamarca.**

1. Introduction

According to reference [1], there are roughly 1.6 billion people worldwide who lack access to electricity, mainly in rural areas, where very often entire villages are yet to be electrified. The International Energy Agency (IEA) forecasts that this figure will be higher than 1.2 billion people by 2030 if adequate policies are not implemented [2]. In fact, achieving basic universal access to energy by 2030 will require an investment of 48 million dollars each year [3].

Establishing the planification agenda of a large-scale rural area is a complex issue that requires a multidisciplinary approach since several factors need to be considered. The policymaker has to be aware of different aspects such as total cost, long-term sustainability, regulatory issues and local constraints and preferences.

In order to successfully overcome the challenge of electrifying nationwide areas the cooperation of large energy companies will be necessary. Hence, an adequate business model is required. This business model has to consider:

- The proper low-cost technologies that will be used to effectively implement the planning decisions. The technologies can be divided into three different groups:
  - Stand-alone systems [4], which provide electricity to only one customer. These systems can use AC or DC technologies as well as several energy technologies, such as solar, diesel, wind, biomass, hydraulic or hybrids.
  - Microgrids [5], which provide electricity to several customers through a distribution network with a grid-independent generation system. Economies of scale allow a per-customer cost that is lower than in stand-alone systems. In addition, a wider array of generation options can be considered.
  - Network connection [6], which is the most worldwide extended way of providing access to electricity.
- Social biases and preferences. The preferences of the customers have to be considered in the decision-making process [7]. Customers may prefer a certain electrification mode or a specific generation technology.
- Regulatory framework. The existence of policies that promote certain renewable energies such as solar or wind have to be considered. Moreover, a reduction in the consumption of fossil fuels such as kerosene may be a target to achieve.
- Different agents involved. Although the policymaker and large energy companies play a key role, other players such as local entrepreneurs and NGOs have to be considered as well. The financial structure of the business model may vary depending on several factors such as the contracts with the consumers and the ownership of the grid and microgrids [8].

Hence, the development of a successful rural electrification strategy is a complex issue that benefits from rigorous computational tools that help the policymaker in the decision-making process.

2. The reference electrification model

One of the most advanced state-of-the-art tools for rural electrification is the Reference Electrification Model (REM), which was introduced in reference [9]. REM considers specific input data for each region, determining where the power grid should be extended and where off-grid systems should be placed. REM makes use of a heuristic algorithm that minimizes total costs when planning the power system in a rural electrification problem. The costs considered include both financial and social costs (which are incurred if demand is not fully satisfied). REM also calculates technical designs for grid extension and off-grid systems using a greenfield Reference Network Model (RNM) to perform the network designs [10].

A. Inputs and outputs

REM's inputs can be divided into several groups:

- Demand. Location of the customers, electrical appliances of each customer and their usage habits.
- Existing network. Layout of the medium-voltage feeders, reliability of each feeder and cost of energy grid at each feeder.
- Network components catalogue. Electrical and economic parameters of lines and transformers.
- Local generation components catalogue. Electrical, economic and reliability parameters related to solar panels, batteries, diesel generators, inverters and charge controllers.
- Weather. Solar irradiation and temperature profiles.
- Financial and cost parameters. These parameters include economic figures that are used in calculations performed by RNM. Fuel cost (diesel) and non-served energy cost are also considered by REM.
- Territory boundaries: REM allows the possibility of dividing the customers into several regions and "solving" these regions in parallel to reduce the computation time.

Regarding the outputs, REM provides a nearly-optimal combination of stand-alone systems, microgrids and grid extension designs. Generation designs and network designs are provided for microgrids. Network designs are provided for grid-extension designs.

B. REM's architecture

REM is divided into five blocks that perform sequential tasks:

1. Pre-processing. This block extracts data related to some inputs described in subsection 2.A. Specifically, REM extracts information associated with customers and the existing electrical network from satellite imagery. The demand profile of each customer is calculated considering the corresponding electrical appliances and usage habits.
2. Microgrid generation designs. REM uses a heuristic algorithm in order to calculate quasi-optimal generation designs for representative combinations of customers, considering their demand profiles. The generation technologies considered by REM are solar panels, batteries and diesel generators. REM builds a look-up table that stores the generation cost of these designs. The reason for using a look-up table is that calculating local generation designs for each candidate microgrid would be unmanageably long. Only a subset of representative designs are accurately calculated and costs related to the remaining ones are obtained using interpolation techniques.

   The generation costs that are stored in the look-up table account for cost related to elements of the microgrid (such as solar panels, batteries and diesel generators) plus a penalty added for the demand that is not met by the generation devices.
3. Clustering. This block groups customers into a hierarchical structure of clusters. The clustering process is based on a bottom-up greedy logic that joins clusters if the estimated cost of being electrified together (connected) is lower than the estimated cost of being electrified separately. Robust cost estimations are critical in order to obtain adequate clusters.
4. Final designs. REM explores the structure of clusters and calls RNM in order to calculate the costs associated to network designs. REM proposes a combination of stand-alone systems, microgrids and grid-extension designs in this block. In REM, microgrids are not connected to the network and they always operate in islanded mode. Furthermore, each microgrid has a single centralized generation system that provides AC electricity in REM.
5. Post-processing and reports. REM generates figures, tables and files that describe the best electrification mode of each customer. REM also provides relevant statistics related to the final designs.

3. State of the Art

There are several tools in the literature that can be used for rural electrification purposes. These tools can be classified in several groups according to their scope.

A. Hybrid system generation design tools

These tools perform a similar task to the second block of REM, which is described in subsection 2.B. They do a more detailed analysis than REM and they usually offer more options regarding the generation technologies and the objective functions at the expense of a higher computational time.

iHOGA (improved Hybrid Optimization by Genetic Algorithms) is a piece of software that optimizes generation selection and operation for a hybrid power system using a genetic algorithm [11]. iHOGA is an improved version of HOGA (Hybrid Optimization by Genetic Algorithms). iHOGA is able to consider several objective functions performing multicriteria optimization.

Hybrid Optimization of Multiple Energy Resources (HOMER) is a tool that selects and sizes the generation and storage devices of a microgrid or a grid-connected system [12]. In order to do that, HOMER simulates the operation of the system over one year for each possible generation configuration of the search space. HOMER is frequently used and referenced in the literature ([13]–[15]). However, HOMER is not able to consider several objective functions and it only focuses on minimizing total cost.

Distributed Energy Resources Customer Adoption Model (DER-CAM) is a tool that chooses which generation technologies should be adopted in a microgrid or a grid-connected system and how that technologies should be operated for minimizing costs [16]. DER-CAM considers a wider range of generation technologies and the possibility of considering additional objective functions such as a CO_2 emissions but the optimal placement of generation is not considered.

B. Hybrid system simulation tools

Other pieces of software simulate the behavior of a given generation design for a microgrid with the purpose of providing information to the user. Hybrid2 ([17], [18]) is a tool that forecasts the behavior of a hybrid power system using time series and probabilistic methods in order to perform an economic analysis. Hybrid2 offers a large number of dispatching strategies and models more generation technologies than REM, but it is currently unsupported and will not work on operating systems newer than Windows XP [19, p. 2].

C. Large-scale rural electrification planning tools

There are also large-scale rural electrification tools that are aimed at villages, communities and even countries. However, there is a scarcity of these integrated planning tools due to the complexities and difficulties of addressing the problem.

Network Planner is a tool that selects the best electrification mode of a rural community [20]. It is very similar to REM, but REM considers a broader range of catalogue options when planning the network designs. For example, REM allows considering several medium-voltage lines for network designs whereas Network Planner is only able to consider one type of medium-voltage line. Moreover, Network Planner groups customers in aggregated communities giving less detailed information about final designs. A case study where Network Planner is used is described in reference [21].

Logiciel d'Aide à la Planification d'Électrification Rurale (LAPER) is a tool aimed at sustainable rural electrification of large regions. It is also very similar to REM, but it is able to consider more generation technologies and some non-economic criteria [22]. However, LAPER aggregates consumers into villages and does not perform network designs inside villages. LAPER also considers less catalogue options when planning the network designs.

D. Viability project analysis tools

RETScreen is a tool that performs viability studies of projects related to renewable energy, energy efficiency and cogeneration. RETScreen can perform analysis such as risk analysis or emissions reduction analysis and covers a wider spectrum of projects [23]. RETScreen first version was targeted at on-grid projects but the current version is capable of handling hybrid and stand-alone systems.

Model for the Analysis of Sustainable Energy Roadmaps for all (MASTER4all) is a tool that considers several generation, transportation and distribution options, selecting a final solution that depends on several objective functions. MASTER4all is able to consider emission restrictions, maximum budget constraints and incentives for providing energy access service [24].

E. Network design tools

Village Power Optimization model for Renewables (VIPOR) is a tool whose goal is to design a distribution network for isolated power systems. In order to do that, VIPOR considers an initial way of connecting consumer nodes and generation and iteratively adds and removes connections using a simulated annealing algorithm until a nearly-optimal solution is reached. VIPOR's algorithm has a similar structure to the clustering process performed in REM. VIPOR considers additional costs related to geography, which is something currently not supported by REM. VIPOR's algorithm is thoroughly described in reference [25]. The main drawbacks of VIPOR are inflexible assumptions about generation sites and poor performance with growing problem sizes. Moreover, VIPOR does not allow the inclusion of electrical constraints such as voltage-drop requirements. An example of rural electrification design where VIPOR is used can be seen in reference [26].

On the other hand, Reference Network Models are computational tools that have been created with the purpose of helping the regulator estimate the costs related to a distribution network in an already-electrified country, where the possibility of substituting grid-connected systems by islanded microgrid systems is not necessarily allowed. However, they can provide assistance when calculating distribution networks associated with microgrids and grid extensions as in the case of REM.

Moreover, Reference Network Models have heuristic clustering algorithms that are worth considering when calculating clusters in REM.

The clustering process performed in the Network Performance Assessment Model (NPAM) [27] bears some resemblance to the one carried out in REM. In NPAM, a cluster starts with one representative node and more nodes are added if electrical and distance conditions are satisfied. Once these conditions are not satisfied, a new cluster is created and the process is repeated. This clustering process is performed first with low voltage customers and later including transformers. The clustering process in the ANETO model [28] divides the whole territory into a squared-cell grid, assigning each customer to its corresponding cell. Each cell is classified as urban, rural or disseminated for reliability reasons. Adjacent cells may be grouped, joining their customers into the same cluster. In the Reference Network Model (RNM) [10], customers in urban areas are grouped into settlements considering criteria related to power and distance. A squared grid is obtained in order to calculate the boundary of each settlement.

Some classic clustering techniques have also been applied to heuristic procedures in the traditional design of distribution networks, although this is not very frequent. As a more sophisticated approach, iterative clustering algorithms based on k-means are used in references [29], [30] to determine the optimal location of medium-voltage/low-voltage transformers in Reference Network Models.

4. Original clustering process in REM

This section describes the original clustering process in REM. Part of this section is a summary of chapter 4 of reference [9], but original analyses have been added in subsection D.

A. Minimum spanning tree

The original clustering process starts by calculating the minimum spanning tree (MST) that connects every customer of an analysis region. Each arc of the MST can be activated in order to join the customers on either side of the connection into one cluster. Initially, all connections are inactive and each customer is in its own cluster of one customer. Then, the off-grid clustering and the on-grid clustering processes are carried out.

B. Off-grid clustering process

Arcs of the MST are ordered according to length, from the shortest to the longest. Candidate connections are then evaluated following that order. REM activates a connection if the cost of a line that joins the two clusters that are located at the ends of that connection added to the generation cost of a microgrid whose customers are the sum of the customers of both clusters is strictly lower than the generation cost of two microgrids whose customers are the customers of each of the clusters. The generation cost of these microgrids is taken from the look-up table. Once this procedure has been performed, customers are distributed into off-grid clusters.

C. On-grid clustering process

Connections that are still inactive are evaluated again from the shortest to the longest. This time, the costs associated with five different configurations are calculated in order to determine if the candidate line is going to be activated or not.

In configuration 1, cluster 1 is connected to the already existing network and a line between clusters 1 and 2 is set. Costs related to configuration 1 are:

- The cost of a medium-voltage or low-voltage line that goes from the main network to the center of cluster 1. This line is designed so that it satisfies the sum of the demands of both clusters.
- The cost of a medium-voltage/low-voltage transformer designed to support the sum of the demands of both clusters.
- The cost of a low-voltage line that goes from cluster 1 to cluster 2. This line is designed with enough capacity to satisfy the demand of cluster 2.
- Network energy and non-served energy costs associated with clusters 1 and 2.

Configuration 2 is built by exchanging clusters 1 and 2 in configuration 1. Configurations 1 and 2 are shown in Figure 1:

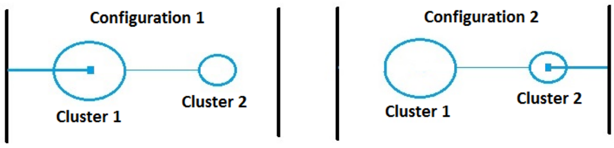

Figure 1: Graphical representation of configurations 1 and 2.

Configuration 3 connects cluster 1 to the network and assigns cluster 2 to a microgrid. The costs associated with configuration 3 are:

- The cost of a medium-voltage or low-voltage line that goes from the network to the center of cluster 1. This line is designed with enough capacity to satisfy the demand of cluster 1.
- The cost of a medium-voltage/low-voltage transformer that satisfies the demand of cluster 1.
- Network energy and non-served energy cost related to cluster 1.
- The cost of the microgrid associated with cluster 2. This cost is taken from the look-up table.

Configuration 4 is built by exchanging clusters 1 and 2 in configuration 3. In configuration 5, both clusters are connected to the already-existing network separately. The costs associated with this configuration are:

- The cost of a medium-voltage or low-voltage line that goes from the network to the center of cluster 1. This line is designed so that it satisfies the demand of cluster 1.
- The cost of a medium-voltage or low-voltage line that goes from the network to the center of cluster 2. This line supports the demand of cluster 2.
- The cost of a medium-voltage/low-voltage transformer that supports the demand of cluster 1.
- The cost of a medium-voltage/low-voltage transformer that supports demand at cluster 2.
- Network energy and non-served energy costs associated with clusters 1 and 2.

Configurations 3, 4 and 5 are shown in Figure 2:

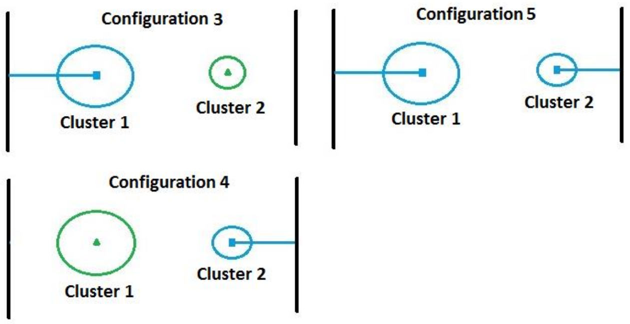

Figure 2: Graphical representation of configurations 3, 4 and 5.

The original on-grid clustering process in REM allows the user to determine which configurations should be considered using two different options. If option 1 is selected by the user REM evaluates the cost of the 5 configurations described. If the least-cost configuration was the one where one of the clusters is linked to the main grid and the other is linked to the first one (that is, configurations 1 or 2) the candidate line is activated. This is reasonable, as configurations 1 and 2 are the only ones where both clusters are connected with a line. However, if the least-cost configuration was any of the others (3, 4, or 5), then the candidate line is not activated.

If option 2 is selected by the user, REM evaluates the costs of configurations 1, 2 and 5. If the least-cost configuration was configuration 1 or 2, then the candidate line is activated. If the least-cost configuration is configuration 5 then the candidate line is not activated. This option was introduced as a rapid way of avoiding the issues related to clustering that are described in subsection 4.D. In fact, option 2 performed better than option 1 [9].

Once this procedure has been performed, off-grid clusters are distributed into on-grid clusters. Each on-grid cluster may contain several off-grid clusters. The clustering process provides a hierarchical structure of on-grid clusters, off-grid clusters and isolated customers. This is shown in Figure 3:

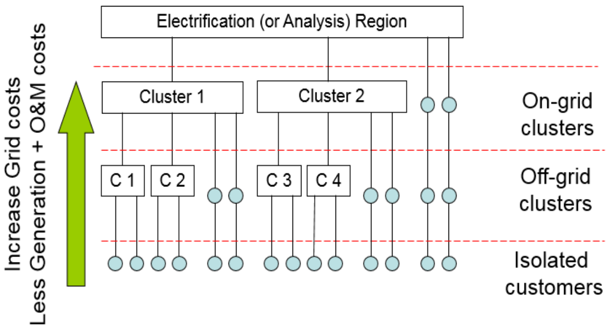

Figure 3: Hierarchical structure of clusters.

It is worth noticing that at this stage it is still undetermined whether an on-grid or an off-grid cluster will be electrified by means of a microgrid or a grid extension. In order to decide that, REM compares the total cost of a grid extension design for an on-grid clusters with sum of the total costs of a microgrid for each off-grid cluster that is contained in that on-grid cluster with the additional costs of stand-alone systems for each isolated customer that belong to that on-grid cluster. If the grid extension design is less expensive, then the whole on-grid cluster is electrified as a grid extension. However, if the grid extension design is more expensive then each off-grid cluster is electrified with its own microgrid and isolated customers are electrified with stand-alone systems.

D. Critical analysis

The solutions provided by REM rely on the comparison between RNM's design of a grid extension for an on-grid cluster and RNM's designs of microgrids for each off-grid cluster that belongs to the corresponding on-grid cluster. Hence, imbalances in either the off-grid clustering or the on-grid clustering are translated into non-optimal results.

Both clustering processes are based on a single pass among the arcs of the MST. The local decisions made are sensitive to the sizes of clusters involved, the discrete catalogue of components and inter-node distances. Moreover, the final results provided by REM depend heavily on the initial local decision carried out in the off-grid clustering. In addition, it should be noted that big off-grid clusters are never created. The main reason for this is related to the behavior of the per-customer generation costs of microgrids that is stored in the look-up table. These costs are calculated using a heuristic algorithm, and they are not always a strictly decreasing function of the number of customers. Sometimes this happens because the heuristic algorithm ends in a local minimum, and considering a discrete catalogue of generation components causes this issue too. Indeed, off-grid clusters stop joining where the generation per-customer costs of microgrids starts increasing.

If option 1 is selected by the user, big on-grid clusters are not created neither. The catalogue that was used for designing networks when calling RNM was the same that was used in the clustering processes. Since decisions are made bottom-up, the less expensive line of this catalogue may be too expensive when deciding if it is worth connecting two small clusters, and therefore on-grid clusters that are provided are usually small.

If option 2 is selected by the user, either on-grid clusters are not created either a big on-grid cluster that contains a large amount of the customers appears. This behavior happens even if the best electrification mode of these customers was not a grid extension. In practice, user option 2 worked better than option 1 because both clusters were always connected to the network in the configurations that were compared using option 2. This mitigates the effect of using RNM's network catalogue in the clustering process.

Therefore, off-grid clusters were usually small and on-grid clusters were either small or extremely big. When determining the best electrification mode of each customer, this fact leads REM to choose a large number of isolated customers or a large number of customers connected to the grid even if that was not optimal.

5. Improvements to the Clustering Process

Several clustering improvements have been implemented in REM in order to avoid the undesirable consequences described in subsection 4.D. These changes have been divided into three different groups: electrical modeling improvements, improvements to logic of the algorithm and improvements to the interpolation process.

A. Electrical modeling improvements

1. <u>Addition of power losses</u>.

   Power losses may be non-negligible in rural electrification [31]. In the original REM, costs associated with lines losses were not considered. In order to obtain more accuracy, the cost related to power losses has been included in the calculations performed in the clustering process.

2. <u>Addition of operation and management (O&M) costs.</u>

   The last step to be considered when implementing a microgrid system is O&M [32]. In order to capture economies of scale that were not present in the original REM model, O&M costs are introduced. Now, each microgrid has an associated O&M cost that depends on its number of customers. These costs account for technicians as well as equipment pieces that fail and need to be replaced. In the literature, the O&M per-customer cost of an off-grid system is usually a monotonically decreasing function of the number of customers [33]. Moreover, the decreasing rate becomes smaller when the number of customers increases. O&M costs are modeled in that way in REM using the function:

$$f(m) = \frac{A\left(1-e^{-\frac{m}{k}}\right)}{m} + B \qquad (1)$$

   where m represents the number of customers of the microgrid. Coefficients A, B and k are calculated based on user-defined parameters. With these data, a system of equations is solved in order to obtain the coefficients A, B and k.

   Furthermore, this new operation and management costs are added not only in the clustering process but in the costs that are used in order to decide whether clusters will be connected to the grid or form microgrids after calling RNM.

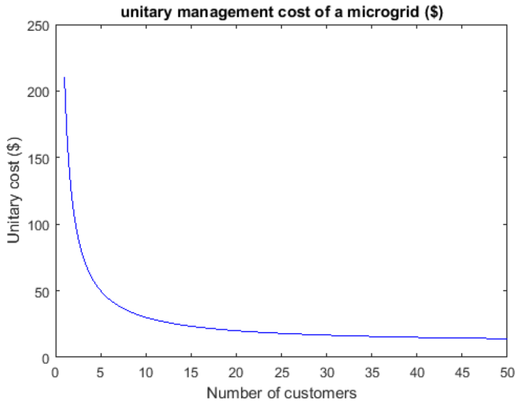

Figure 4: Example of per-customer operation and management costs of a microgrid.

3. <u>Using peak demand instead of average demand for calculations.</u>

In the original REM, average demands were considered when calculating the sizes of transformers and lines. The main reason for doing this is that using the peak demand would have implied even smaller off-grid and on-grid clusters, worsening the issues described in subsection 4.D. However, it makes more sense to use peak demand instead of average demand once the main problems of the original REM clustering have been properly addressed.

B. Improvements to logic of the algorithm

1. <u>Using demand-weighted centers of clusters when calculating distances instead of the minimum distance between the two clusters.</u>

In the original REM, distance between two clusters was always calculated as the minimum of all the distances between a customer that belonged to the first cluster and a customer that belonged to the second cluster. It is clear that using that shortest-customer-to-customer distance is the most favorable case, since it corresponds to the less expensive line. Experience with the current version of REM (which has implemented all the improvements described in this section) has shown that this distance may be excessively optimistic in some cases where a very large off-grid cluster was created. In order to avoid that, distances among clusters are now calculated as the distances among their demand-weighted centers in the current version of REM (a demand-weighted center is a weighted center in which each vertex has a weight that is equal to the peak demand of the corresponding customer).

2. <u>Addition of configurations 1' and 2' to the five configurations of the on-grid clustering process described in section 4</u>.

Configuration 1’ is very similar to configuration 1, but this time the line that joins both clusters is a medium-voltage line instead of a low-voltage line. This implies that two medium-voltage/low-voltage transformers are required instead of 1. Configuration 2’ is built by exchanging clusters 1 and 2 in configuration 1’.

Configuration 1 is now compared to configuration 1'. If the total cost associated with configuration 1' is lower than the total cost related to configuration 1, then configuration 1' is considered in the remaining calculations instead of 1. Configuration 1' costs include:

- Procurement of grid energy and non-served energy cost associated with clusters 1 and 2.
- Cost of a medium-voltage line from the center of cluster 1 to the nearest medium-voltage line, carrying power of the total peak demand of clusters 1 and 2 together.
- Cost of a medium-voltage/low-voltage transformer with enough capacity to satisfy the total peak demand for cluster 1.
- Cost of a medium-voltage/low-voltage transformer with enough capacity to satisfy the peak demand for cluster 2.
- Cost of a medium-voltage line from the center of cluster 1 to the center of cluster 2, carrying the power of the peak demand for cluster 2.
- Operation and management cost of a grid extension whose customers are the sum of customers of clusters 1 and 2.

Configurations 1’ and 2’ are shown in Figure 5:

Figure 5: Graphical representation of configurations 1’ and 2’.

Configuration 1’ is less expensive than configuration 1 when the demand of cluster 2 is large enough so that satisfying it through a medium-voltage line is more efficient than using a low-voltage line for that purpose. Analogous reasoning is applied to configurations 2 and 2’.

3. <u>Addition of several loops through arcs of the MST in both off-grid and on-grid clustering processes.</u>

Going through arcs of the MST several times avoids missing potential connections that should be activated. This happens because the second time that an arc is evaluated both clusters that are connected by that arc may be significantly bigger than the first time that the same arc was evaluated. Hence, both off-grid and on-grid clustering now loop through all unconnected arcs of the MST until no arc is connected in the last loop. It should be noted that this greedy heuristic is not necessarily optimal, but it approximates to the optimal solution better than the former procedure. Specifically, the total cost savings between applying this improvement and not applying it to the current REM version in the case study described in section 7 is around 10%.

C. Improvements to the interpolation process

REM stores the generation designs in a look-up table, as well as their respective generation costs for a certain number of customers and demand patterns. Each axis of the look-up table corresponds to a customer type. Customer types are characterized by their demand patterns. The generation technologies considered in REM are solar panels, batteries and diesel generators.

1. Look-up table interpolation with one customer type.

If there is only one customer type, linear interpolation was performed in the original REM in order to estimate the generation cost of designs that have not been calculated when creating the look-up table. In the current version of REM, a different approach is performed.

Firstly, a partial smoothed look-up table is obtained ensuring that that the per-customer generation costs strictly decrease when the number of customers increase. In order to do that, the generation costs of some designs are modified if necessary. Secondly, a continuous two-part piecewise function is obtained. Both parts of the function have the expression of equation 1.

The reason for using a two-part function instead of a one-part function is that the shape of the curve is extremely important for low values of customers. If the shape of this part of the function is not adequate big off-grid clusters will never be created, and having this part right with a one-part function turns out to be extremely complicated.

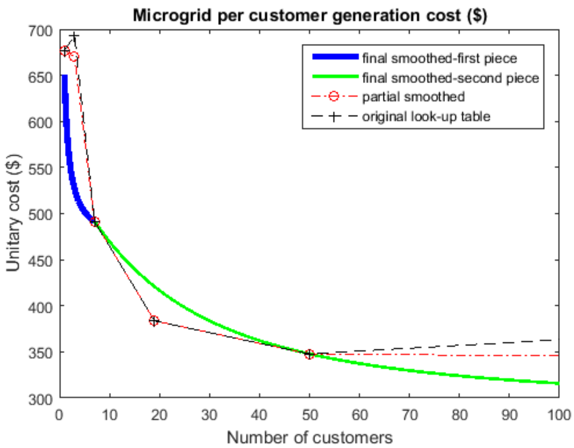

Figure 6: Smoothed look-up table. The original look-up table generation per-customer costs are higher for three customers than for one customer. That implies that no off-grid clusters with more than one customer were produced in the original REM.

2. Look-up table interpolation with several customer types.

If there are several customer types, multi-dimensional interpolation is performed to estimate cost of designs that are not stored in the look-up table. In the original REM, a hypersurface that has the same dimensions as the number of customer types was adjusted using data of all designs stored in the look-up table. Then, the corresponding point of the hypersurface was selected in order to estimate data of a design. That hypersurface went through all points of the look-up table. The process of calculating a hypersurface, which was carried out each time that an interpolation was performed, required a lot of computational time.

In order to avoid this, multilinear interpolation is used. Multilinear interpolation has been successfully applied to both theoretical [34] and practical [35] problems. Its main advantage is the low amount of computational time that is required each time that a multilinear interpolation is performed, since only the values of boundary points of the hyperinterval that the point belongs to are used to perform the calculation. More interpolation techniques such as multidimensional spline interpolation [36] or Chebyshev polynomials [37] have been considered, but they are harder to implement and they require more computational time for each interpolation.

3. Usage of a continuous catalogue in the clustering process.

In order to avoid ending up with small off-grid and on-grid clusters, the catalogue used in the clustering process is now a continuous version of the discrete catalogue used in designs performed by RNM. Specifically, it assumes that a line or a transformer which capacity is exactly the capacity required in the clustering process always exists. The remaining electrical and economic parameters associated with that line or transformer are obtained through a linear interpolation between the two closest-capacity lines or transformers of the discrete catalogue.

Moreover, a line and a transformer of capacity zero are added to the continuous catalogue in order to allow interpolation through capacity values lower than the capacity of the lowest-capacity line and transformer of the discrete catalogue. The remaining values of the zero-capacity line and transformer parameters are calculated considering the corresponding parameters of the discrete lowest-capacity line and transformer. If a line or transformer with capacity greater than the capacity of the maximum-capacity discrete line or transformer is required, values of electrical and economic parameters are obtained through extrapolation considering the corresponding values of the maximum-capacity discrete line and transformer.

6. CASE STUDY

In Peru, rural electrification is developed by the Ministry of Energy and Mines through the National Plan for Rural Electrification that was approved in 2013. This plan considers the time period 2014-2022 with the goal of achieving universal access to electricity by the end of this period.

The case study developed in this paper is included in the National Plan for Rural Electrification. Specifically, it corresponds to the Michiquillay District Encañada, which belongs to the region of Cajamarca (Peru). This district is located in the Andes Mountains. The altitudes of this zone is between 2200 and 4100 meters, and its area is approximately 400 km^2. It has approximately 6700 households and several connection points to the 11 kV projected network.

In the case study, the network infrastructure mentioned in the Electrification Plan for Cajamarca (2008-2017) is supposed to be installed already. The network is assumed to have enough capacity to satisfy all the demand. All the customers of the case study are assumed to be non-electrified.

This case study has been analyzed with the original version of REM in reference [38]. In order to establish a fair comparison between both REM versions the input parameters that were used with the original REM are maintained.

Due to the inaccessibility of the district, the average local price of diesel is relatively high (2$/l). The network catalogue that is used is the same that was considered in [38]. Grid supply is assumed to be 100% reliable and network energy cost has been set to 0.045 $/kWh.

Network lifetime is assumed to be 40 years and the discount rate is set to 10%. The demand growth rate is set at 1%.

This case of study considers that customers have the following appliances:

1. Two lights and a phone charger (critical demand).
2. One additional light for 50% of the customers (non-critical demand).
3. One fan for 20% of the customers (non-critical demand).
4. One television for 30% of the customers (non-critical demand).

Failing to satisfy critical demand is highly penalized. The remaining appliances are regarded as non-critical and failing to satisfy their demand is not highly penalized. Costs of critical and non-critical non-served energy are 10 and 1.5 $/kWh respectively.

Since the O&M costs introduced in subsection 5.A were not available in the original REM version they are set to zero for the case study.

7. RESULTS

The original version of REM obtained a solution for the case study where the best electrification mode of each customer was either to be isolated or to be connected as a microgrid.

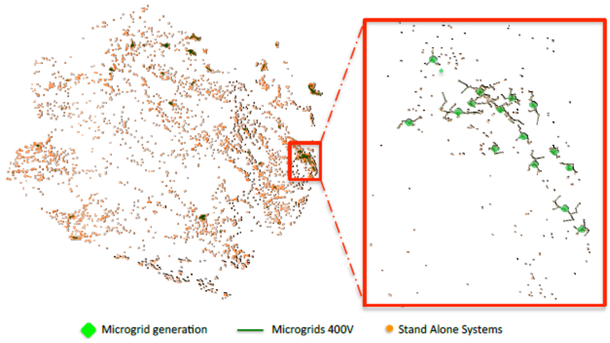

Figure 7: Solution provided by the original REM, taken from reference [38].

A total number of 4779 customers obtained stand-alone systems with solar panels of 180 pW and a battery bank of 960 Wh that supplies 93% of their demand. The average yearly cost of a stand-alone system is 203.88 $/customer. In addition to that 1917 customers are connected forming microgrids and 83 of these microgrids have 10 customers or more. The average yearly cost of electrifying a customer connected to a microgrid is 149.28 $/customer. Microgrids supply 94.8% of their demand.

The total yearly cost of this solution adds up to $1,258,592.3. However, problems described in subsection 4.D are present here: neither big off-grid clusters nor big on-grid clusters are created. Network designs performed by RNM indicate that connecting a small number of customers to the network is more expensive than having them isolated or connected as microgrids. Hence, REM ends up connecting no customers to the network.

The new version of REM provides a different solution to the case of study. Customers that are close to the network connection points are electrified with grid extension designs whereas customers that are further from the network are electrified either as stand-alone systems or connected forming microgrids.

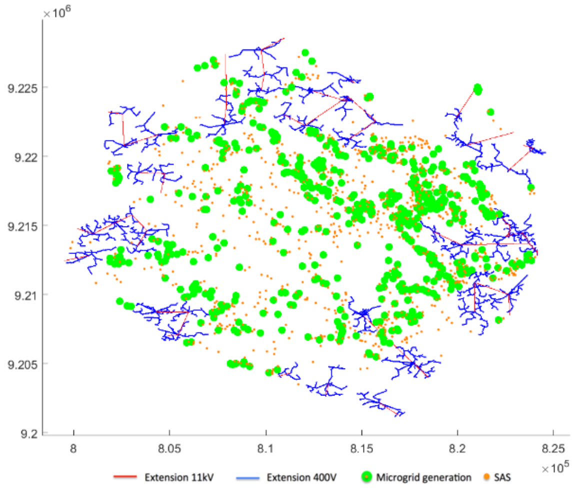

Figure 8: Solution provided by the current version of REM.

In the results, 816 customers were assigned to stand-alone systems with solar panels of 180 pW and a battery bank of 960 Wh that supplies 93% of their demand. The average yearly cost of a stand-alone system is 203.88 $/customer. Another 1565 customers are connected forming microgrids, and 6 microgrids have 10 customers or more. The average yearly cost of electrifying a customer connected to a microgrid is 159.92 $/customer. Microgrids supply more than 95% of their demand.

The remaining 4307 customers are connected to the grid. The average yearly cost of electrifying a customer connected to the grid is 72.56 $/customer. A total of 20 grid extension designs have been suggested in the current REM best solution. Each of these designs has a total number of customers equal or greater than 31.

The total yearly cost of the solution provided by the current REM is $729,156. It is clear that the solution provided by the current REM outperforms the solution provided by the original REM. Moreover, it does not make any sense to have no customer connected to the network when the network reliability is 100%, the grid energy cost is 0.045 $/kWh and the diesel price is relatively high (2$/l), and this is what happened with the original REM.

8. CONCLUSIONS AND FUTURE RESEARCH

The clustering process implemented REM has been analyzed, showing that it had several defects that had a huge impact in the quality of the final solutions provided by REM. In order to avoid that undesired behavior, several clustering improvements have been presented and implemented in REM. These enhancements have been classified into three types according to their nature: electrical model improvements, improvements to the logic of the algorithm and improvements to the interpolation process.

Results of a small but representative case of study have been obtained with both the original version of REM and the current version of REM. It is clear that the current REM version behaves in a more consistent way, obtaining a solution that is approximately 40% less expensive than the one obtained with the original REM for the same case study.

However, there is still work to be done. The addition of topography to REM is a task that would allow much more realistic modeling. RNM already supports this, so this improvement has to be implemented in a way that is consistent with RNM.

Moreover, the additional upstream reinforcements that would be required when connecting customers to the already-existing network are not considered in the current version of REM. This task is far from trivial and no state-of-the-art rural electrification tool has succeeded in it so far.

9. REFERENCES

[1] Hisham Zerriffi, *Rural Electrification. Strategies for Distributed Generation*. Springer, 2011.

[2] Organisation for Economic Co-operation and Development, *World energy outlook 2013*. Paris: International Energy Agency, 2013.

[3] Fatih Birol, ‘Energy for all: financing access for the poor’, Oslo, Oct. 2011.

[4] S. Khennas, ‘Best practices for sustainable development of micro hydro power in developing countries’, The World Bank, 21640, Aug. 2000.

[5] Jose Ignacio Pérez Arriaga, *Tecnologías para el desarrollo humano de las comunidades rurales aisladas*. Real Academia de Ingeniería, 2012.

[6] S. Pachauri, B. J. van Ruijven, Y. Nagai, K. Riahi, D. P. van Vuuren, A. Brew-Hammond, and N. Nakicenovic, ‘Pathways to achieve universal household access to modern energy by 2030’, *Environ. Res. Lett.*, vol. 8, no. 2, p. 24015, Jun. 2013.

[7] Francisco Javier Santos Pérez, ‘Metodología de ayuda a la decisión para la electrificación rural apropiada en países en vías de desarrollo’, Universidad Pontificia Comillas, Madrid, 2015.

[8] ‘Sustainability - Business Models for Rural Electrification’, European Commission, Energy Facility, 2012.

[9] Douglas Ellman, ‘The Reference Electrification Model: A Computer Model for Planning Rural Electricity Access.’, Massachusetts Institute of Technology, 2015.

[10] C. Mateo Domingo, T. Gómez San Román, Á. Sanchez-Miralles, J. P. Peco Gonzalez, and A. Candela Martinez, ‘A Reference Network Model for Large-Scale Distribution Planning With Automatic Street Map Generation’, *IEEE Trans. Power Syst.*, vol. 26, no. 1, pp. 190–197, Feb. 2011.

[11] ‘iHOGA software’. [Online]. Available: http://personal.unizar.es/rdufo/index.php?option=com_content&view=article&id=2&Itemid=105&lang=en. [Accessed: 09-Apr-2016].

[12] Tom Lambert, Paul Gilman, and Peter Lilienthal, *Micropower System Modeling with HOMER*. John Wiley & Sons, 2006.

[13] R. Sen and S. C. Bhattacharyya, ‘Off-grid electricity generation with renewable energy technologies in India: An application of HOMER’, *Renew. Energy*, vol. 62, pp. 388–398, 2014.

[14] A. Mohamed and T. Khatib, ‘Optimal sizing of a PV/wind/diesel hybrid energy system for Malaysia’, in *Industrial Technology (ICIT), 2013 IEEE International Conference on*, 2013, pp. 752–757.

[15] L. Zhu and X. Yang, ‘Design and Simulation for Microgrid System Based on Homer Software’, *Adv. Mater. Res.*, vol. 361–363, pp. 1874–1877, Oct. 2011.

[16] O. Bailey, C. Creighton, R. Firestone, C. Marnay, and M. Stadler, ‘Distributed Energy Resources in Practice: A Case Study Analysis and Validation of LBNL’s Customer Adoption Model’, *Lawrence Berkeley Natl. Lab.*, 2003.

[17] J. F. Manwell, A. Rogers, G. Hayman, C. T. Avelar, J. G. McGowan, U. Abdulwahid, and K. Wu, *Hybrid2: a hybrid system simulation model: theory manual*. Citeseer, 1998.

[18] E. lan Baring-Gould, ‘The Hybrid System Simulation Model’, 1996.

[19] ‘Hybrid2, Wind Energy Center.’ [Online]. Available: http://www.umass.edu/windenergy/research/topics/tools/software/hybrid2. [Accessed: 12-Apr-2016].

[20] R. H. Han, ‘Network Planner Landing Page’, *Network Planner*. [Online]. Available: http://networkplanner.modilabs.org/. [Accessed: 10-Apr-2016].

[21] Francis Kemausuor, Edwin Adkins, Isaac Adu-Poku, Abeeku Brew-Hammond, and Vijay Modi, ‘Electrification planning using Network Planner tool: The case of Ghana’, *Energy Sustain. Dev.*, vol. 19, pp. 92–101, 2014.

[22] Rainer Fronius and Marc Gratton, ‘Rural Electrification Planning Software (LAPER)’, Electricité de France - Research and Development, 2001.

[23] ‘RETScreen Clean Energy Project Analysis Software | Open Energy Information’. [Online]. Available: http://en.openei.org/wiki/RETScreen_Clean_Energy_Project_Analysis_Software. [Accessed: 12-Apr-2016].

[24] A. González–García, I. J. Pérez–Arriaga, and A. Moreno-Romero, ‘MASTER4all: A decision support model for roadmaps towards sustainable energy universal access for policymakers and utilities’.

[25] T. W. Lambert and D. C. Hittle, ‘Optimization of autonomous village electrification systems by simulated annealing’, *Sol. Energy*, vol. 68, no. 1, pp. 121–132, 2000.

[26] ‘Design and Analysis of SPV-Diesel Hybrid System for Rural Electrification in Odisha.’, *Int. J. Sci. Eng. Res.*, vol. 4, no. 12, pp. 129–131, 2013.

[27] M. B. Larsson, ‘The Network Performance Assessment Model: A new framework for regulating the Electricity network Companies’, 2005.

[28] J. a. Garcia Conejo, G. Tevar Bartolome, A. Gomez Exposito, and M. Rodriguez Montanes, ‘ANETO: A System for the Automatic Generation of Theoretical Network Models’, *9th Int. Conf. Electr. Power Qual. Util.*, Oct. 2007.

[29] A. Navarro and H. Rudnick, ‘Large-Scale Distribution Planning — Part I : Simultaneous Network and Transformer Optimization’, *IEEE Trans. Power Syst.*, vol. 24, no. 2, pp. 744–751, 2009.

[30] L. González-sotres, C. Mateo, Á. Sánchez-miralles, and M. A. Miró, ‘Large-Scale MV / LV Transformer Substation Planning

Considering Network Costs and Flexible Area Decomposition', *Power Deliv. IEEE Trans. On*, vol. 28, no. 4, pp. 2245–2253, 2013.

[31] K. Gautam and V. Bhuria, 'Rural Electrification with Loss Minimization Through different Strategies for Sustainable Infrastructure Development', *Int. J. Eng. Res. Appl. DERA ISSN*, pp. 2248–9622, 2012.

[32] Venu Aggarwal, Alison Fahey, Hana Scheetz Freymiller, CC Huang, Simonne Li, Stephen Moilanen, Chikara Onda, Nathan Ratledge, Stephanie Speirs, and Jacqueline Wong, 'Rural Energy Alternatives in India: Opportunities in Financing and Community Engagement for Renewable Energy Microgrids Projects.', Princeton University, 2014.

[33] L. M. Carrasco, L. Narvarte, and E. Lorenzo, 'Operational costs of A 13,000 solar home systems rural electrification programme', *Renew. Sustain. Energy Rev.*, vol. 20, pp. 1–7, Apr. 2013.

[34] Loukas Grafakos and Terence Tao, 'Multilinear interpolation between adjoint operators.', *J. Funct. Anal.*, vol. 199, pp. 379–385, 2003.

[35] Fernando de Cuadra García, 'El programa general de la optimización de diseño por ordenador: Aplicación de técnicas de ingeniería del conocimiento.', Universidad Pontificia Comillas, 1990.

[36] C. Habermann and F. Kindermann, 'Multidimensional Spline Interpolation: Theory and Applications', *Comput Econ*, vol. 30, no. 2, pp. 153–169, Sep. 2007.

[37] B. Heer and A. Maussner, *Dynamic General Equilibrium Modeling: Computational Methods and Applications*. Springer Science & Business Media, 2011.

[38] Andres Gonzalez-Garcia, Reja Amatya, Robert Stoner, and Ignacio Perez-Arriaga, 'Evaluation of universal access to modern energy services in Peru. Case study of scenarios for Electricity Access in Cajamarca.' Enel Foundation, 2015.