Applied Energy 338 (2023) 120906

Contents lists available at ScienceDirect

Applied Energy

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

Optimizing planning and operation of renewable energy communities with genetic algorithms

Florencia Lazzari [a,b,*], Gerard Mor [b], Jordi Cipriano [b], Francesc Solsona [a], Daniel Chemisana [c], Daniela Guericke [d]

[a] Department of Computer Science and Industrial Engineering, University of Lleida, Jaume II 69, 25001 Lleida, Spain
[b] International Center for Numerical Methods in Engineering. Building Energy and Environment Group. CIMNE-Lleida. Pere de Cabrera 16. Office 2G, 25001 Lleida. Spain
[c] Applied Physics Section of the Environmental Science Department, University of Lleida, Jaume II 69, 25001 Lleida, Spain
[d] High-Tech Business and Entrepreneurship Department, University of Twente, P.O. Box 217, 7500 AE, Enschede, The Netherlands

HIGHLIGHTS

- Optimization of renewable energy communities considering environment and economy.
- Combinatorial optimization for participant selection.
- Multi-objective optimization of solar energy allocation.
- Results show high avoided CO2 emissions and low paybacks for all participants.

ARTICLE INFO

Keywords:
Renewable Energy Communities
Solar energy
Optimization
Genetic Algorithm

ABSTRACT

Renewable Energy Communities (REC) have the potential to become a key agent for the energy transition. Since consumers have different consumption patterns depending on their habits, their grouping allows for a better use of the resource. REC provide both economic and environmental benefits. However, its potential drastically diminishes when grouping of prosumers and energy al- location is performed improperly, as the energy generated ends up not being consumed. Given the importance of extracting the maximum potential of REC, this study presents a tool to assist in both the planning and the operation phases. We present a combinatorial optimization method for participant selection and a multi-objective (MO) optimization of solar energy allocation. Specific Genetic Algorithms (GA) were developed including problem-specific approaches for reducing the search space, encoding, techniques for space ordering, fitness functions, special operators to replace duplicate individuals and decoding for equality constraints. The performance of the novel solution approach was exper- imentally proved with an electrical solar installation and electricity consumers from Northern east Spain. The results show that the developed tool achieves energy sharing in REC with low solar energy excess, high self-consumption and high avoided CO2 emissions while assuring low payback periods for all partic- ipants. This tool will be essential to increase revenues of REC schemes and boost their beneficial environmental impact.

1. Introduction

To tackle the global environmental crisis, the emergence of a new power system, close to citizens and sustainable, is imminent. Putting citizens at the center will bring the energy issue into public debate, generating concern about the electricity consumption. It will enable a more equitable and democratic model, in which energy generation facilities will shift from being owned by a few companies to being owned by citizens. In addition, to ensure that this new power system is decarbonized, it should be powered by renewable energy sources (RES).

Renewable Energy Communities (REC) represent the participation of a col- lective in the power system through renewable energy generation facilities placed near consumers [1]. REC offer citizens a means of co-ownership of energy sources that provides environmental and economic benefits [2]. The environmental ben- efit is obtained through the

* Corresponding author.
E-mail address: florencia.lazzari@udl.cat (F. Lazzari).

https://doi.org/10.1016/j.apenergy.2023.120906
Received 16 November 2022; Received in revised form 13 February 2023; Accepted 22 February 2023
Available online 10 March 2023

Nomenclature

Abbreviations

DSO	distribution system operator
ESS	energy storage systems
EU	european union
GA	genetic algorithm
MO	multi-objective
NSGA-II	non-dominated sorting genetic algorithm-II
P2P	peer-to-peer
PV	solar photovoltaics
REC	renewable energy communities
RES	renewable energy sources Indices

Indices

i	participant
j	position inside an individual (possible solution)
t	time [hours]
w	objective function

Parameters

K	number of participants forming the REC
M	number of time steps in one analysis period
N	number of available participants in the neighbourhood
pp	electricity purchase price [$/kWh]
ps	solar excess electricity sale price [$/kWh]
Q	cardinality of the search space
W	number of objective functions
Y	number of time steps in one year

Variables

$\beta^i(t)$	normalized hourly allocation coefficient
$\widetilde{E}_s(t)$	solar electricity consumed [kWh]
ϕsc	self-consumption
ϕss	self-sufficiency
$\overrightarrow{x}$	combination of participants from a set of candidates
B	matrix of the energy allocation coefficients
$C_n^i(t)$	cost of energy consumed from new sources, once the REC is set up [$]
$C_o^i(t)$	cost of energy consumed from old sources, before the REC is set up [$]
$E_s^i(t)$	hourly solar electricity corresponding to each participant [kWh]
$E_c^i(t)$	energy consumed by each participant [kWh]
$E_e^i(t)$	solar energy excess dispatched to the grid [kWh]
$E_g^i(t)$	energy imported from the grid [kWh]
$E_s(t)$	hourly energy generated by the PV system [kWh]
$E_s^i(t)$	allocated solar energy generated by the PV system [kWh]
f_w	objective functions
I^i	investment of each participant
P^i	profit for a whole year of each participant
R^i	payback of each participant [years]

decarbonization of the power system and reduction of the transmission losses. The economical benefit comes from direct financial returns to the involved members. Since REC are formed by a combination of.

consumers with different daily electric consumption patterns, their grouping in sharing schemes optimizes the energy generation fitting [3]. Furthermore, the co-ownership of energy sources fosters consumer-empowerment, increases the energy sovereignty, helps to reduce energy poverty and boosts the local econ- omy [4].

In this study, the focus is on REC formed by residential prosumers, since it is one of the most promising markets. Amongst RES, solar photovoltaics (PV) is particularly attractive due to its low cost, high reliability and consolidated technical service providers [1]. Therefore, the focus will be on REC addressed to residential consumers and fed by solar PV systems, hereafter referred to as REC.

The actors involved in a REC have complex dynamics; consumer behaviour is highly stochastic, PV generation is intermittent and the price of electricity varies over time. Therefore, proper policy frames are necessary to provide the expected benefits of energy sharing [5]. Energy communities are now defined in the Clean Energy Package, and the Renewable Energy Directive 2018/2001 sets the framework for REC [2]. These recent favourable regulatory frameworks are boosting the creation of REC in the European Union (EU) [3,6].

Besides these advances in the EU regulation, REC in the EU by 2021 only contributed 7 % of nationally installed capacities of renewables, estimated at. GW [7]. The fact that REC still have a marginal presence is due to inaccu- racies in the planning and operation stages. The planning phase is the period during which appropriate dimensions of community generation and flexible re- sources must be estimated [8]. Requirements in the planning phase are: how many peers are necessary to form efficient groups or who are the ideal part- ners [9]. Later, during the operation phase, it is critical to optimize the energy sharing among the involved customers (considering their investment, how their daily profile fits the generation or any other criteria as long as there is an agreement). These design requirements and energy allocation needs are generally addressed by simplifying the problem. However, specific optimization tools can improve these inaccurate procedures, improving the design and management of REC in a way that helps to achieve community goals and, therefore, promoting a successful rolling out of REC. All in all, the transition to a renewable energy model is leading to new opportunities for citizens but if they are not provided with the right tools, the high potential of REC will not be tapped [6].

When analysing the existing literature on REC optimization the first clear distinction is made between studies focused on the planning phase and studies addressing the operation phase [8]. In general, when focusing on the planning stage, the analysis is conducted from the perspective of Distribution System Operators (DSOs), that are mainly interested in voltage stability and power flow. Ghiani et al. [10] present the design of a REC in which the sizing of the generation is addressed using power flow simulation. Vahidinasab et al. [11] studied both the power flow optimization problem and the economic profitability of the community as a whole. Weckesser et al. [12] carried out an extensive study of REC and their potential impact on different topologies of the electricity network (city, village). In summary, most researchers propose detailed power. flow studies simulating the design of REC to guarantee the grid's stability.

Considering the impact of the REC on the electric grid is essential. But in the planning stage, the physical viability of the REC should be complemented with optimization studies from the community's point of view. Zarei et al. [13] identify the best combination of participants to form a REC assessing energy- efficient behaviors in a social network. They concluded that different combina- tions of the participants could considerably increase the energy savings rate. In this sense, the human energy-related behaviors have a significant impact.

On the other hand, most literature studies addressing the operation stage provide tools to work on online markets using the Peer-to-Peer (P2P) energy trading modality. They generally use human-on-the-loop strategies that give feedback through apps, using systems similar to the betting ones. In a P2P model, participants buy or sell energy directly with each other. Ye et al. [14] designed an online algorithm to tackle cost-aware energy sharing among resi- dents in a REC. This included the cost of purchasing electricity from the main grid, and the cost of charging and discharging Energy Storage Systems (ESS). Liu et al. [15] developed P2P energy trading management approaches of RES integrated with energy storage of hydrogen and battery vehicles for power sup- ply to a REC. Finally, Rodrigues et al. [16] investigated the management of a P2P energy sharing network considering different ESS ownership structures.

However, when looking at real life implementations, these P2P transaction models are unfeasible under most current regulations. In the particular case of Spain, where our research is centered, the legal framework doesn't allow for an online update of energy allocation or P2P transactions.

We present a method to optimize the energy sharing in a REC. This research makes several contributions to the state of the art of REC planning and opera- tion. It contributes to fill the gap between theoretical methodologies presented in the literature and realistic tools. To do so, we present a combinatorial op- timization for participant selection and a multi-objective (MO) optimization of solar energy allocation. The study shows the implementation of Genetic Algo- rithms (GA) along with the novel heuristic designed to address the complexities of the two problems. The novel heuristic consists of: a procedure to reduce search space, encoding to represent possible solutions, technique for space ordering, fitness functions to achieve the stated objectives, special operators to replace duplicate individuals and decoder as repair mechanism to modify indi- viduals which don't comply equality constraints. The main distinctions from state-of-the-art studies are the following.

- This method is applicable in real scenarios, more specifically in countries where the legal framework establishes that the distribution of RES genera- tion needs to be done based on allocation coefficients (such as the Spanish and French legal framework).
- The methodology is developed with a prosumer-driven perspective, instead of the typical DSO's point of view. Providing vital information for the decision making of consumers eager forming a REC.
- The objective of the tool is not limited only to the exploitation of RES potential or the economic profitability. Unlike many studies, our aim is focused on guaranteeing both low PV surplus and fair distribution of benefits.
- The economic objective is achieved by considering each investor individ- ually, which is rarely done in the literature and often a requirement in practice.
- The methodology is validated with real data. This allows us to understand that the developed method can work successfully in real scenarios.

2. Optimization models

Since there is no standard regulation and each country defines their specific technical implementation, we will focus on the electricity market of Spain as an example of implementation. However, the presented model can be applied to other countries with minor modifications. In Spain, the RD 244/2019 decree presents new possibilities for prosumers by enabling collective self-consumption [6]. It simplifies administrative procedures and improves the economic viability of the PV installations by recognizing the right of prosumers to sell the surplus electricity to the grid [3]. The decree presents different regimes according to the installed capacity [17]. Among these, we will focus on the scenario for in- stallations from 15 to 100 kW, which is the most attractive for the residential sector.

In this case, the regulation allows a simplified monthly net billing remuneration mechanism. The difference between PV generation and the energy consumption is obtained hourly. A positive result is considered a surplus (or excess) that benefits from a revenue price. Negative values are subtracted from the hourly consumption. The computation of these differences is made monthly. The customer will obtain an energy saving benefit when the energy generated is greater than the consumed and an income from the surpluses sold to the grid. These revenues are limited to the same quantity as the monthly energy cost.

According to the regulation, the hourly solar electricity for billing purposes corresponding to each participant $E_s^i(t)$, is:

$$E_s^i(t) = \beta^i(t)E_s(t) \tag{1}$$

where $E_s^i(t)$ is the total hourly energy generated by the PV system; and $\beta^i(t)$ is the normalized hourly allocation coefficient in time t (in hours) and participant i. The allocation coefficients are signed in an agreement by all the participants and notified to the DSO. $\beta^i(t)$ can be different for each hour, provided that the sum of them over all K participants forming the REC is one:

$$\sum_i^K \beta^i(t) = 1\forall t, \tag{2}$$

This enables a customized distribution of the generated energy. These coef- ficients may be determined according to the billing power of each participant, their economic contribution to the PV installation, or any other agreed crite- rion. The coefficients must be established *a priori*, before energy is generated and consumption is produced. They can be modified every 4 months. All these rules and regulation constrains have been taken into consideration hereafter to develop the optimization methods.

This study presents optimization algorithms to enhance energy sustainability and economic profitability of the participants of a REC. Energy sustainability is considered by minimizing the solar energy generated and not consumed by the REC (which is dispatched to the electricity grid). This is supported by the fact that energy from the grid has higher $CO2$ emission rates than the local solar. energy production. On the other hand, economic profitability is considered by. looking for individual payback periods which are acceptable and at similar levels for all participants.

Fig. 1 shows the process flow for the optimization of the planning and oper- ational phases of the REC, and the algorithms designed to be applied in each phase. Both in the planning and the operation phases, the required inputs are the investment of each participant, historical data of the energy generated by the PV collective installation, the electricity consumption of each participant, and the price signal of the addressed electricity market. The method comprises two optimization algorithms for the planning phase: i) Selection and ii) Al- location. The Selection algorithm aims to select the optimal combination of participants that minimizes the surplus generation to be delivered to the grid. The Allocation algorithm aims to determine the optimum $\beta^i(t)$ assigned to each participant i in an hourly granularity. This last algorithm minimizes the solar energy excess and ensures a fair investment return for all the participants. In the operational phase, only the Allocation algorithm is executed.

The energy consumed by each participant i at time t is defined as $E_c^i(t)$ (eq. (3)).

$$E_c^i(t) = E_g^i(t) + E_s^i(t) - E_e^i(t) \tag{3}$$

were $E_g^i(t)$ is the energy imported from the grid, $E_s^i(t)$ is the allocated solar energy generated by the PV system, and $E_e^i(t)$ is the solar energy excess dispatched to the grid. In this study $E_c^i(t)$ and $E_s^i(t)$ are known values. $E_e^i(t,\beta)$ and $E_g^i(t,\beta)$ are calculated according to Equations (4) and (5), respectively.

$$E_e^i(t;\beta) = \left(\beta^i(t)E_s(t) - E_c^i(t)\right)^+ \tag{4}$$

$$E_g^i(t;\beta) = \left(\beta^i(t)E_s(t) - E_c^i(t)\right)^- \tag{5}$$

where $()^+$ and $()^-$ represent the positive and negative part respectively of the calculation inside the (Macaulay) brackets. Notice that $E_s^i(t) =$

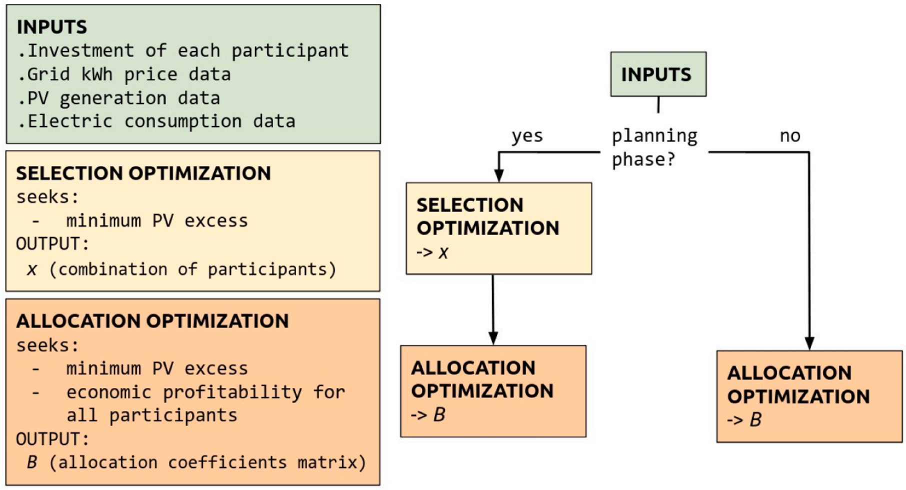


Fig. 1. Scheme showing the optimizations and the couplings.

$\beta^i(t)E_s(t)$.

In the planning phase, historical data sets of energy consumed by each par- ticipant are available. In some cases, the PV system is already installed, and historical data sets of the generated power are also available. The energy gen- erated should be simulated when the PV system is not yet installed. In this phase, $E_s(t)$ is defined as the hourly mean of the solar energy generated or simu- lated, calculated over each hour slot within each month.) $E_c^i(t)$ is also defined as the hourly mean of the energy consumption for each participant calculated over hourly slots within each month. Still, here, two types of days are distinguished to calculate the hourly means: weekday and weekend. This categorization is done because consumption behaviour significantly varies these days.

In the operational phase, two different scenarios can be distinguished ac- cording to the energy allocation procedure defined by each national regulation:

i) the *a priori* scenario, where the allocation coefficients should be determined before the billing period (Spanish electricity market); and ii) the *a posteriori* scenario, where the energy allocation coefficients are determined after the billing period (French electricity market). In the first case, the *a priori* scenario, naive forecasting of the previous period is performed, and the hourly mean values are determined following the same procedures as in the planning phase. In the second case, the *a posteriori* scenario, the energy consumed and generated is already metered. Therefore, no forecasting of $E_c^i(t)$ and $E_s(t)$ is needed.

2.1. Selection optimization

The Selection optimization has to be executed in the planning phase of the REC. It is typified as a combinatorial optimization aiming to find the best combination of participants from a set of candidates. As an example, a possible. solution could be $\vec{x} = (1, 21, 80)$, representing the combination of participants 1,

21 and 80. To choose among the possible combinations, the objective function to be minimized is the solar energy excess $E_e^i(t)$ of each participant (eq. (6)), throughout the whole analysis period, comprising M time steps (in hours) and over the total amount of participants K.

$$min_{\vec{x}} \sum_t^M \sum_i^K E_e^i(t) where \vec{x} \in \{\vec{x_1}, \vec{x_2}, \cdots, \vec{x_Q}\} \tag{6}$$

where Q is the cardinality of the search space and is calculated as the combination $\binom{N}{K}$; N is the number of all available participants in the neighbourhood; and K is the number of participants to form the REC. This means that when choosing, for example $K = 6$ participants among $N = 100$, the order of magnitude of the cardinality is 109. Search spaces with this cardinality can't be explored using brute force. Therefore, specific combinatorial optimization techniques, combined with REC domain expertise, should be smartly taken into account when imple- menting the algorithm to solve this problem.

2.2. Allocation optimization

The Allocation optimization is designed to be executed in each optimization period (4 months in Spain). The objective is to find the β s that minimize the difference among individual payback periods and the solar energy excess. Therefore, two target functions should be implemented.

Several variables related to the economic costs are considered to minimize the difference between individual payback periods (herein- after referred to as *paybacks*). $C_0^i(t)$ is the cost per time unit (i.e. hour), for each participant *i*, of the energy consumed from *old* sources (the grid), before the REC is set up and there has been no generation of solar energy yet (eq. (7)). It is calculated considering the energy consumed and the time varying electricity purchase price *pp*.

$$C_o^i(t) = p_p(t)E_c^i(t) \tag{7}$$

The energy cost of each participant, once the REC is established and the solar energy generation is allocated, is defined as the cost of *new* sources $C_n^i(t)$ per time unit (eq. (8)). This energy cost is affected by the revenues of the *solar excess* dispatched to the grid, which is sold at the *solar excess* sale price *ps*.

$$C_n^i(t) = p_p(t)E_g^i(t) - p_s(t)E_e^i(t) \tag{8}$$

In general, $pp \sim 3\ ps$. This implies that self-consuming solar energy is always more convenient than selling it.

The aggregated profit for a whole year of each participant is expressed as *Pi*

$$P^i = \sum_t^Y \left(C_o^i(t) - C_n^i(t)\right) \tag{9}$$

where Y is the overall number of time steps (hours) in a year. On the other hand, the payback of each participant is expressed in years as R^i (eq. (10)).

$$R^i = \frac{I^i}{P^i} \tag{10}$$

where I^i is the investment of each participant.

Our objective is to minimize the difference among the paybacks of the different participants and the exported solar energy excess. This MO optimization is stated in eq. (11).

$$
\begin{aligned}
&\min_B\{f_1(B),f_2(B)\}\\
&s.t. B = \beta^i(t) \in \mathbb{R}^{MxK}_{[0,1]}\\
&\sum\nolimits_i^K \beta^i(t) = 1 \forall t = t_1 \cdots t_M\\
&f_1(B) = \sum\nolimits_i^K E_e^i\\
&f_2(B) = \sum\nolimits_i^K exp\left(R^i\right)
\end{aligned}
\tag{11}
$$

where $f_1(B)$ is the *solar excess* function, $f_2(B)$ is the *payback* function; and B is the matrix of the energy allocation coefficients $\beta i(t)$. B contains all the K participants in the REC for every time step in the analysed period (M time steps, in hours). To complete the problem statement, the constrain in eq. (2) must be satisfied independently for each time step (represented by each row of the matrix).

$$
B = \begin{pmatrix} \beta^1(t_1)\cdots\beta^i(t_1)\cdots\beta^K(t_1) \\ \vdots \\ \beta^1(t_M)\cdots\beta^i(t_M)\cdots\beta^K(t_M) \end{pmatrix} s.t. \sum_i^K \beta^i(t) = 1, \forall t = t_1 \cdots t_M
$$

Our goal is to minimize both objective functions simultaneously. However, it is impossible to find a solution that optimizes two conflicting objectives. Instead, the solution to a MO problem is a number of points belonging to the objective function space, these are called Pareto optimal solutions. For a Pareto optimal solution, no objective can be improved without degrading the other objective. Without additional information on subjective preferences, there are an infinite number of Pareto optimal solutions. All of them can be considered equally good, and none of them is preferred over the others.

With the above in mind, the Pareto set of best solutions will be searched first. Afterwards, an ideal or utopian point will be drawn in the objective function space. Finally, the best solution will be selected from the Pareto set of candidate solutions as the one that minimizes the euclidean distance to the ideal point. Therefore, the selected solution will be as close as possible to the ideal solution. The main outcomes of this optimization will be: (i) the selected solution; and.

(ii) the complete set of Pareto solutions found. These outcomes can be delivered to the decision makers so they can choose among all the possible scenarios.

3. Solution method

The Selection and Allocation optimization problems are high-dimensional, non derivable and NP-Hard problems. There is no polynomial-time algorithm to find the solution when the problem is of these characteristics. In the worst case, one would need to evaluate all possible solutions in an exact optimization approach to prove optimality. However, the cardinality of the search spaces being considered grows rapidly (due to the curse of dimensionality), making it computationally intractable to use an exact optimization approach.

Therefore we trade optimality for speed, using an approximation algorithm. Ad-hoc heuristic methods are often ineffective because they are handicapped by their biased set of rules. Stochastic optimization approaches are the alternative approach for solving these types of problems. Stochastic optimizations cannot guarantee optimal solutions. However, they outperform traditional deterministic search methods when applied to complex problems.

There is a large variety of stochastic optimization approaches. None of them can be said to be generally superior to all the others. When selecting a solu- tion method, the specific features of the problem to be solved must be taken into consideration. We chose Genetic Algorithms (GA) to solve the two opti- mization problems presented. There is a scientific community working on GAs which shows consensus on the fact that GAs are empirically good at provid- ing near-optimal solutions in cases where: 1) the function to be optimized is non-differentiable, 2) the function evaluation has low computational cost, 3) complex constraints are involved (including equality constraints), 4) the search space is not smooth (for example in combinatorial optimization, there is a need to optimize over a discrete domain), and 5) for MO optimization (when the relation among the functions is not known in advance). If it is possible to define an heuristic where the representation in which the genetic operators work at their best (which requires tuning and domain knowledge), then the GA shows.

fast convergence and requires low computational time to produce high-quality solutions [18].

GA are inspired by natural selection, relying on mutation, crossover and selection of individuals (which represent the candidate solutions) [19]. Algo- rithm 1 shows the pseudocode describing a general GA. The heuristic defined by the GA designer includes several aspects. Firstly, the encoding scheme that transforms possible solutions into strings. Secondly, the objective function that maps problem solutions to fitness values. Thirdly, the population size (*Np*). Then, the method to select parents (*roulette-wheel selection, tournament selec- tion, rank selection*, among others). Furthermore, the crossover type should also be specified. This means, choose the way in which parents will share some of its genetic information with their offspring. In addition, the mutation rate should be selected accordingly. A high rate will turn the GA into random search. In- stead, a low rate could leave parts of the search domain unexplored. Finally, the stopping criterion may be set to a predetermined number of generations or to run until the fitness of the best individual is better than some user-defined threshold or when evolutionary process is not changing significantly.

Algorithm 1: GA pseudocode

```
initialize Np individuals {xi }
encode individuals with specific representation
evaluate objective function for each individual
while not (stopping criterion)
  for k = 1 to Np
    select parents from {xi }
    use crossover to create a new child ck
    r ← random number between 0 and 1
    if r < mutation rate
       mutate ck
     end if
     evaluate objective f u n c t i o n using ck
   next child
   replace duplicate individuals in {xi } U {ci }
   {xi } ←best Np individuals from {xi } U {ci }
 next generation
```

According to the Law of Conservation of Information [20], it is pointless to attempt to design a GA that is better than random search, unless you can incorporate problem-specific information in the algorithm. This translates in the fact that, if the GA is not correctly implemented, it will end taking large computational time or premature convergence to a local minimum (this ap- plies especially in problems with large cardinality). Performance improvements hinges on using prior information to match procedures to problems [21]. The performance of a GA highly depends on the definition of: the function assigning fitness values to possible solutions, an encoding (mapping) mechanism between.

the problem and algorithm domains, the representation, and special operators to handle constraints. There are no hard-and-fast rules to define them. The understanding of the problem is critical, so that a problem-specific heuristic is constructed in a way that leads to successful results. There are infinite heuristics to represent any optimization problem. However, it must be carefully designed to match the structure of the target problem, because if the heuristic fails, the GA will fail lamentably [22].

To motivate the particular heuristic proposed, studies considering different combinatorial and MO problems were analyzed. Combinatorial problems ana- lyzed include the traveling salesman problem, the graph coloring problem, the minimum spanning tree problem, the job shop scheduling problem, the knapsack problem, and the bin packing problem [18,21,23]. On the other hand, MO prob- lems incorporating the concept of Pareto dominance were considered [24,25].

3.1. GA for selection optimization

The objective of the Selection optimization is to find the best combination of participants to form the REC. This type of optimization is called combinatorial optimization, where independent variables are restricted to a set of discrete values.

Searching for all possible solutions is often the only way to solve combinato- rial problems. But when solving real-life problems, search spaces tend to grow rapidly, making it impossible to carry out an exhaustive search. Therefore, the aim here is not to find the exact solution but a reasonable, near-optimal, feasi- ble solution in an acceptable timescale. In this sense, studies in the literature revealed that GAs have a great potential to solve a wide range of combinato- rial problems [18]. This is mainly because GAs use an intelligent way to seek through the domain, avoiding the infeasible brute-force search.

However, for this type of problems with large cardinality, the implementa- tion should avoid large computational time or premature convergence to a local.

minimum. We propose the introduction of specific knowledge by limiting the number of participants (K) by a maximum defined as $Kmax(K \leq Kmax)$. First, the periods in the data set where the solar generation peaks occur, *tsmax*, are found by using derivatives. We then obtain $Es(tsmax)$, the maximum solar electricity generated per day, and $Ec(tsmax)$, the electricity consumption for that time ($tsmax$). Finally, the ratio of self-consumption φsc (eq. (12)) for each user is obtained. φsc measures the usage of the generated solar electricity.

$$\varphi_{sc} = \frac{\widetilde{E}_s(\mathrm{t})}{E_s(\mathrm{t})} \tag{12}$$

where $\widetilde{E}_s(t)$ is the solar electricity consumed $\left(\widetilde{E}_s = Es - Ee\right)$. To estimate the φsc here, $E_s^i = E_c^i$ was used. This is, we assumed that all the energy consumed will be solar. Then $\overline{\varphi sc}$ is calculated as the temporal mean over all periods *tsmax* and over all participants. Finally, the maximum number of participants *Kmax* is estimated as the inverse of the self-consumption percentage of the generated electricity (eq. (13)).

$$K_{max} = \left\lceil (\overline{\varphi_{sc}})^{-1} \right\rceil \tag{13}$$

This approach to estimating the maximum number of participants only con– siders the sustainable side (solar surplus). It is supported by the fact that participants' investments are usually not a barrier because installation costs are currently low [3]. Therefore, we can affirm that the maximum number of participants in the REC depends only on their consumption capacity and the generation capacity of the PV installation. Once the maximum number of participants is determined, we search for the most suitable consumers to take advantage of the available solar energy.

Different ways to represent combinatorial candidate solutions exist. The chosen encoding determines the size of the space. Encoding was defined as an integer vector $\overrightarrow{x}$ of size K_{max}, representing the participants forming the REC. It was designed to avoid redundant possible solutions (in GA terminology; individuals). An encoding which allows duplicate individuals (two individuals representing the same solution) can generate an entire population of clones, leading to premature convergence.

For example, $\overrightarrow{x}_1 = (4,8,7)$ and $\overrightarrow{x}_2 = (7,4,8)$ are duplicated individuals. Both represent the same feasible solution. The coordinates are ordered in ascending order. In the same example, $\overrightarrow{x}_1$ and $\overrightarrow{x}_2$ are replaced by $\overrightarrow{x}_0 = (4,7,8)$. This process is performed in the replace duplicate individuals step, in Algorithm 1, line 15.

In addition, infeasible solutions must be removed. For example $\overrightarrow{x}_3 = (5,5,40)$ is a valid individual, but in real life, only one participant 5 exists. Therefore, all duplicated participants in one solution vector are replaced by 0. Such a vector should be transformed into $\overrightarrow{x}_3 = (0,5,40)$. This approach avoids infeasible solutions.

When analysing GA implementations of similar combinatorial problems, an important aspect is the ordering of the search domain. Therefore, the participants are ordered according to their % self-consumption (eq. (12)) before the GA is executed. This led neighbors in the search space to come closer in the fitness space.

According to eq. (1), it is clear that the solar allocation depends on the βs. Since we don't know in advance the ideal hourly $\beta i(t)$ for each participant, they are established to be proportional to their hourly consumption $E_c^i(t)$. In this way, the only independent variables participating in the optimization are the number and chosen participants. This strategy is based on a "divide and conquer" technique because if the $\beta i(t)$'s were not fixed, the problem-optimization complexity could increase excessively.

The parameters of the GA for the *Selection* optimization are summarized in Table 1. The population size, mutation rate and number of generations for the stopping criterion were obtained doing a sensitivity analysis, carrying out multiple runs of the algorithm with different values and comparing the outcome [26]. The algorithm performance was defined as the best fitness at termination. The tuning algorithm performed 20 runs with different values for.

each parameter. The whole implementation was done in R language, using the package GA (version 3.2) [27,28].

3.2. GA for allocation optimization

The objective of the Allocation optimization is to find the best B matrix, which contains the $\beta i(t)$ coefficients for each participant for all the hours in the optimization period. The search is performed considering economic and environmental aspects using the *payback* and the *solar excess* functions, defined in eq. (11). As stated previously, this is a MO problem. Two methods can be used to solve MO problems: *a priori* and *a posteriori* methods.

The *a priori* methods require preference information from the decision maker before the solution process, in order to define the objectives' relative importance. This information is used to produce a single scaled objective function (such as a weighted sum). Following this approach, the problem is transformed into a single-objective optimization and thus, allowing it to be solved. The *a priori* methods reduce the search space and may not be able to find all the available solutions. Therefore, these methods are not desirable for exploring all the do- main. Moreover, in real-world problems, these techniques generally lead to poor quality solutions because obtaining the objective function is a non-trivial task. In addition, in our case, we don't know in advance the trade-off between the *payback* and *solar excess* functions because it is different for each combination of participants. Besides, the units of both functions are different.

In *a posteriori* methods, a representative set of Pareto optimal

Table 1
Implementation of Selection optimization.

Encoding	Integer (through Binary -Gray code)
Size of individuals	*Kmax*
Population size	80
Selection of parents	Linear Rank selection
Crossover type	Uniform binary
Mutation rate	0.08
Fitness function	f_1, with $\beta i(t)$'s fixed proportionally to their $E_c^i(t)$
Stopping criterion	50 generations without changes
Output	Combination of participants with minimum *solar excess*

solutions is first found, and then the decision maker chooses one of them. These methods do not require preference information. Instead, they produce a set of elements of the Pareto optimal set. GAs are popular within the *a posteriori* methods because they simultaneously deal with a set of possible solutions (the so-called population) and can find several members of the Pareto optimal set in a sin- gle "run" by incorporating the concept of Pareto dominance into the selection mechanism [24,25]. Therefore, *a posteriori* is the most appropriate method to be used in our scenario.

Non-dominated Sorting Genetic Algorithm-II (NSGA-II) [29] has become a standard approach and it is one of the most commonly used *a posteriori* method.

NSGA-II emerged as an early approach, and several enhancements were made over the years. NSGA-II assigns fitness of each individual by considering non– domination level and crowding distance.

In the non-domination, every individual is ranked in Pareto fronts based on its performance regarding the objective functions. To construct the Pareto fronts, the domination between individuals is considered. A feasible solution $\overrightarrow{x}_1 \in \Omega$ (domain) is said to *dominate* another solution $\overrightarrow{x}_2 \in \Omega$, if.

$\forall w \in \{1, \ldots, W\}, f_w\ (\overrightarrow{x}_1) \leq f_w\ (\overrightarrow{x}_2)$, and (14).

$\exists w \in \{1, \ldots, W\}, f_w(\overrightarrow{x}_1) < f_w\ (\overrightarrow{x}_2)$, (15).where W is the number of objective functions which in our case is $W = 2$. Ω is the domain composed of all the feasible parameters which in our case is equal to $\Omega = B \in R^{MxK}_{[0,1]} RMxK$. If one individual dominates another, then: i) for the two objective functions, the value of the dominating individual is lower or equal to the other one, and ii) for one of the functions, the value of the dominating individual is the lowest one. The set of Pareto optimal solutions are those that are not dominated by any other feasible solutions.

Each individual is compared with the rest of the individuals in the popula- tion. A list of dominant individuals and the number of dominants is obtained. Based on this information, a sorting process finds the different Pareto fronts in which all individuals are ranked.

In one Pareto front, all individuals have the same performance. The crowding distance is used to distinguish among individuals in the same front. It is defined as the average distance to the nearest neighbours along each objective function dimension. Individuals in more crowded regions of the objective function space are assigned the worst crowding distance, encouraging population diversity.

The main procedure of the NSGA-II algorithm is then as follows: a) all individuals from the best fronts with the best crowding distance are chosen to be the next parent population, b) tournament selection for parents, c) offspring through crossover, and d) mutation. This whole process is repeated until the stopping criterion is reached. The parameters of the NSGA-II algorithm for.

the Allocation optimization are summarized in Table 2. The parameter tuning was obtained doing a sensitivity analysis [26]. The algorithm performance was defined as the best fitness at termination. The tuning algorithm performed 20 runs with different values for each parameter. The whole implementation was done in R language, using the package nsga2R (version 1.1).

Regarding the encoding, each individual is composed of real numbers be- tween 0 and 1, representing the hourly $\beta i(t)$ of each participant. The problem with this encoding arises when looking for individuals that satisfy the equality constraint for $\beta i(t)$ s in eq. (2). If we randomly generate an initial population in a continuous search domain, we have an essentially zero probability of obtaining individuals satisfying equality constraints. Therefore, given that rejecting infeasible individuals would eliminate all the individuals, we apply a *decoder* to convert them into feasible solutions [30]. Then, each individual in the population is transformed by an instruction set for building a candidate solution that always satisfies the problem constraint.

This decoder consists of a normalization of the individual, obtaining $x_j^{decoded}$ (eq. (16)).

$$x_j^{decoded} = \frac{x_j}{\sum_j^K x_j} \tag{16}$$

where j represents the position inside the individual. This decoder distributes the population non-uniformly in the search space, with a peak density around $1/K$. This introduces bias in the search, that can be problematic. However, in this case it is advantageous, because this accumulation point is in the subspace we are looking to benefit (equidistribution of energy), which makes it a good way of focusing the space of search.

Once all the optimal Pareto set is found, the *ideal point* is defined by drawing in the objective function space the (usually utopian) scenario comprising the minimum for both objective functions, as shown with a cross in Fig. 2. To help the decision maker, the *best point* solution is selected from the Pareto set of candidate solutions as the one that minimizes the euclidean distance to the *ideal point* (highlighted with a rhombus).

Table 2
Implementation of Allocation optimization.

Encoding	R[0,1]
Size of individuals	$M \times K$
Population size	200
Selection of parents	Tournament selection
Crossover type	Simulated binary crossover
Mutation rate	0.2
Fitness functions	f_1 and f_2
Stopping criterion	500 generations
Outputs	1)Pareto set of possible scenarios 2) Matrix $B \in R^{MxK}_{[0,1]}$

4. Results and discussion

The case study comprises electricity consumption data from smart meter readings of 128 households and solar generation data from 1 PV installation. The data sets contain information generated for the year 2021. All the house- holds and the solar generation installation are from the city of Barcelona, Spain. In this case study, the number of participants to be selected (obtained with eq. (13)) was 7. That is, the selection optimization had to select 7 participants out of 128 possible participants.

4.1. GA domain-ordering performance

Since GAs are nondeterministic, different results may be obtained every time they are executed. Therefore, a convergence study should be conducted.

to determine if our algorithm is reliable and avoid biased conclusions. The same test was performed 40 times using the Monte Carlo method, seeded with different random numbers. Fig. 3 and Fig. 4 shows how the self-consumption ordered domain implementation of the GA outperforms the non-ordered one. Both figures are the convergence study of the *solar excess* (f_1) carried out for the two variants.

Fig. 3 shows in a log scale, the *solar excess* corresponding to the best candi- date solutions, with and without ordering the domain. It can be observed that over the generations, on average, the solar excess decreases faster in the case of the ordered domain.

Fig. 4 shows the histograms of f_1 for the final selected combinations on the Monte Carlo analysis for both implementations (with and without ordering the.

domain). When comparing both histograms, we can conclude that, firstly, both algorithms reached the same *best solution* (combination of participants with minimum solar excess). This can be determined since both the ordered and the non-ordered achieved an equal minimum solar excess of 3.75 kWh. However, the ordered one performs better because, for this implementation, the amount of outputs with the mentioned

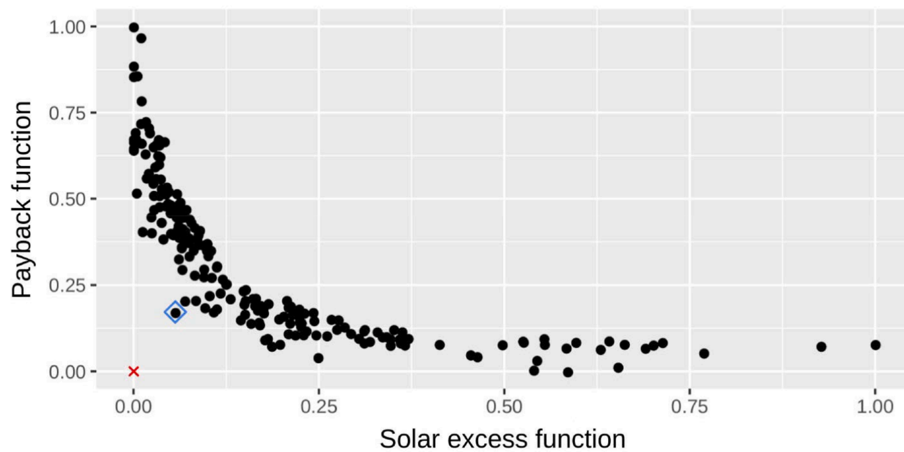


Fig. 2. . Normalized solutions plot on objective functions space. The cross represents the ideal point and the rhombus remarks the best scenario found by minimizing the euclidean distance to the ideal point.

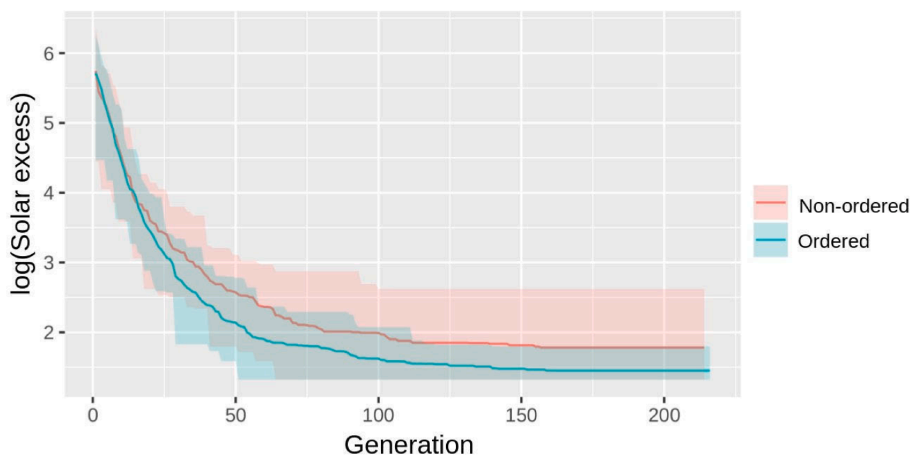


Fig. 3. Monte Carlo convergence analysis. In red, GA with non-ordered domain. In blue, GA with self-consumption ordered domain. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

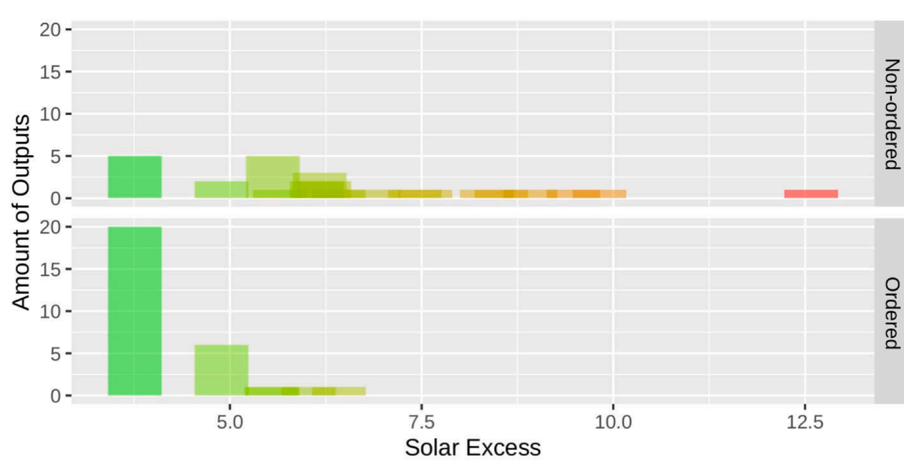


Fig. 4. Histogram of $f1$ for the final selected combination on the Monte Carlo analysis. The top graph shows the GA with free domain. The bottom graph shows the GA with self-consumption ordered domain.

minimum solar excess is higher. This is shown by a significant peak of outcomes for the lowest value of f_1 reached. Moreover, the ordered implementation shows minor variance, meaning it is more robust. To sum up, the novel proposed optimization algorithm can find combinations more concentrated in low ranges of f_1.

Finally, an aspect generally criticised of GAs is that they are time-consuming. However, in our case, the average processing time of convergence was $\sim$ 6 min-.

utes when running in a 16 threads processor - AMD2700x.

These results show that the proposed approach successfully detects combi- nations of participants with low *solar excess* and a high degree of convergence in an acceptable time to be used in real applications.

4.2. Comparison to traditional REC

A general evaluation of the results of the whole optimization is carried out. Regarding environmental objectives, the aim is to guarantee that the overall *solar excess* is low. This is, to minimize f_1. In the eco- nomic aspect, all par- ticipants have invested differently and each has a different capacity to take advantage of solar generation due to distinct consumption patterns (for exam- ple, a REC formed by family households and student apartments). However, all of them want a reasonable and similar investment payback period. This is, to minimize f_2.

The performance of the optimization methods is analyzed by comparing the obtained REC with two other REC configurations. The first one is named profitable REC, and it is designed to look at equitable profitability for all par- ticipants. The usual strategy used to define this profitable REC is choosing participants out of a random selection and then using solar allocation based on investment ($\beta^i = \frac{I^i}{\sum_k I^k}$, this is the common way to calculate allocation coefficients in Spain). The second one is named sustainable REC, and its main objective is to extract the maximum potential of renewable sources. The usual strategy to define the sustainable REC is choosing participants out of a random selection but including a higher number of participants. In this case, we decided to double the number of participants and to allocate the solar energy considering only the solar excess ($\beta^i = \frac{E_c^i}{\sum_i E_c^i}$, this is the common way to cal-.

culate allocation coefficients in France).

To perform the comparison, different parameters are calculated. The eco- nomic parameters are related to the payback period. The mean, the maximum and the difference between the maximum and minimum (Δ) are calculated. The environmental parameters are the *solar excess* delivered to the grid, the avoided CO2 emissions (using the CO2 emission factor of the Spanish electricity mix, 0.357 kg $CO2/kWh$ [31]) and the self-consumption (eq. (12) calculated over the sunny hours). In addition, the self-sufficiency is calculated (eq. (17)), even though it is not a strictly environmental parameter, to provide more information in the assessment of the *sustainable* and *novel* REC.

$$\varphi_{ss} = \frac{\widetilde{E}_s(t)}{Ec(t)} \tag{17}$$

where $Ec(t)$ is the energy consumed and $\widetilde{E}_s(t)$ is the solar consumption, defined as the solar electricity assigned and consumed ($\widetilde{E}_s = Es(t) - Ee$).

The results are shown in Fig. 5.

When looking at the environmental parameters, the first thing that stands out is that both the *sustainable* and the *novel* REC obtain similar results in three of the four parameters: low *solar excess*, high self-consumption and high reduction of CO2 emissions. This shows a good environmental performance of the *novel* and *sustainable* REC. In contrast, if we look at the *profitable* REC, it can be observed that it achieves a poor environmental performance (high *solar excess*, low self-consumption and low reduction of CO2 emissions). Finally, when comparing the *novel* and *sustainable* REC, it is seen that the *sustainable* REC will find it more difficult to fulfil the amount of electricity needed with solar energy (because the number of participants has doubled) and this is translated into a lower self-sufficiency.

When comparing the economic parameters, the first thing we observe is that for the three REC configurations, the mean payback is acceptable (considering 7 years as an acceptable payback). However, the average payback by itself is not a good measure on which to base de- cisions, since despite having a reasonable average payback, it can hide a large variance, generating disparity between the participants. When observing in more detail, the max and delta payback of the *sustainable* REC are significantly high, showing an important profitability disparity of the participants. The *profitable* REC shows, in this respect, better outcomes. However, the *novel* REC has the lowest max and delta payback. This is because, although the *profitable* REC allocation focused on the economic aspect, the participants were randomly selected. Therefore, their capacity to minimise the *solar excess* delivered to the grid (at a significantly lower selling price than the purchased price) is much lower than the capacity of the *novel* REC. Overall, the *novel* REC performed the best in the economic tests.

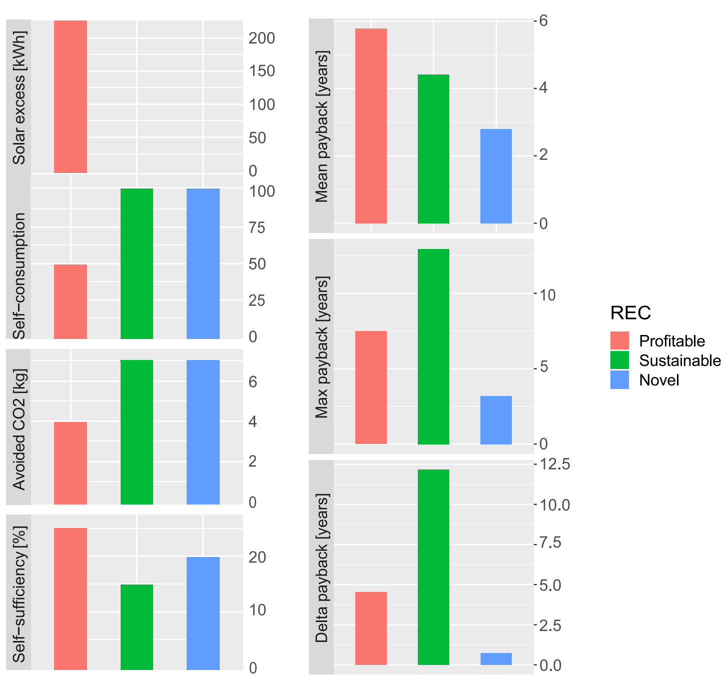


Fig. 5. Environmental and economic performance of the different REC:*profitable*, *sustain- able* and *novel*.

4.3. Profitable, sustainable and novel REC configurations performance

Figs. 6, 7 and 8 show the results obtained in the 3 REC scenarios (*prof- itable*, *sustainable* and *novel*) in a more detailed study. Figures a) represent the hourly mean aggregated energy for the whole REC during a day. The colours distinguish between *solar excess*, *solar* and *grid consumption*. *SC* and *SS* stands for self-consumption and self-sufficiency respectively. The graphs on the left side of figures b) show the investment of three selected participants. In order to better understand the behaviour for a wide variety of participants, the ones with minimum, median and maximum paybacks where chosen. The graph on the right side of figures b) show the solar energy consumed and excess (delivered to the grid) for the three participants. The darker area represents the solar energy consumed by each participant and the lighter area represents the *energy excess*, sold to the grid at price *ps*. The dotted line represents the electricity-grid purchase price *pp*.

Fig. 6 shows the results obtained for the *profitable* REC. Looking at Fig. 6b,

a possible problem with this type of allocation arises. By distributing the energy proportionally to the investment, the participant who invested the most, receives more energy than anyone else. This seems like a fair choice and, in fact, leads to acceptable paybacks. The problem appears when this participant is not the one with the highest consumption during peak solar generation hours. In these cases, the energy received by this participant will contribute to the solar energy excess. This leads to the aggregated consumption pattern shown on Fig. 6a, where it is reflected that the *solar excess* is very significant with respect to the solar energy consumed by the REC.

Fig. 7 shows the analysis for the *sustainable* REC. In Fig. 7a it is observed that adding a high amount of participants generates a consumption peak that matches the peak of solar generation. Moreover, complementing with the indi- vidual consumption on the right, it is evidenced that the allocated generation is adjusted to the consumption of each participant. So that, there is no individual.

surplus, leading to zero total surplus and the self-consumption is maximized. But when analyzing Fig. 7b in detail, we can see that there were participants who consumed little with high investment (right-top graph), and others that with small investment consumed a lot (right-top graph). What happened is that the first participant will obtain an extremely high payback and the second a very low payback. In this way, it is understood that although this model fulfils sustainability objectives, at the same time it encourages low investment and high energy consumption.

Finally, Fig. 8 shows the results for the *novel* REC. From 8a, it can be seen that using the Selection optimizer, a group of participants whose peak con– sumption coincides with the peak of solar generation has been found, obtaining a high self-consumption without the need of recruiting many participants. Fur- thermore, when compared with the two previous REC, we must highlight that the difference between the participant with the minimum and maximum pay- back is very small. From 8b it is clear what strategy the Allocation optimizer uses to achieve this: the participant who has invested the most receives more solar energy when the price of grid electricity is most expensive (in the after- noon). In this way, these graphs show how the two novel optimizers manage to form REC groups of participants with their allocation coefficients, meeting the objectives of sustainability and equidistributed and reasonable profit.

From the results, it can be appreciated that the *novel* REC model presented generated using the Selection and Allocation optimizers proposed in this study, self-consumes all of the solar energy generated, thus reducing the *solar excess* and maximising the reduction of CO2 emissions. At the same time, all partici-.

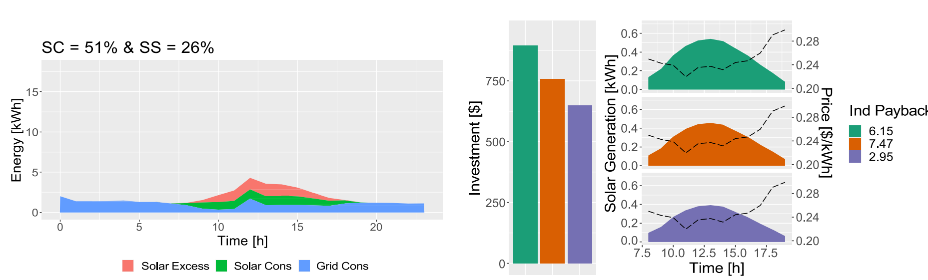
pants obtained a very acceptable and equitable payback. This means that the.

proposed method can optimize the energy generated in a neighborhood, manag- ing this energy so that all participants obtain the same benefits without budget discrimination. Hence, the presented optimization method configures a REC that meets both economic and sustainable objectives.

Studies focused on similar combinatorial problems were analysed to moti-vate this particular implementation. Reference [21] describes a complete survey of the recent advances of adapting metaheuristic algorithms when applied to combinatorial problems. And [23] presents a GA along with a novel heuristic to solve the Bin Packing Problem in which individuals are generated through sorting descending order according to the value of volume, length, width and height of each products respectively. The reason to do this is that when the search space has some regularity, recombination tends to perform well. Based on these previous works we proposed a new model, using GAs, that outperforms, in an acceptable computational time our minimization problem. Despite GAs solutions being not deterministic, in our case, the presented results provide a solution (group of participants with their allocation coefficients) with reasonable performance.

This novel development has been validated with data from a collective solar PV installation and hundreds of electricity consumption profiles from house- holds placed in Barcelona, in Northern east Spain. The results show that, regarding the environmental aspects, the novel algorithm successfully obtained a REC with low *solar excess* (1 kWh in a year), high self-consumption (100 %) and high avoided CO2 emissions (7 kg/day). And, regarding the economic as- pects, savings were evenly spread among all members and low paybacks for all participants (mean payback of 2.8 years, max payback of 3.1 years and delta payback of 0.7 years). Overall, the proposed approach can be applied without much computational cost.

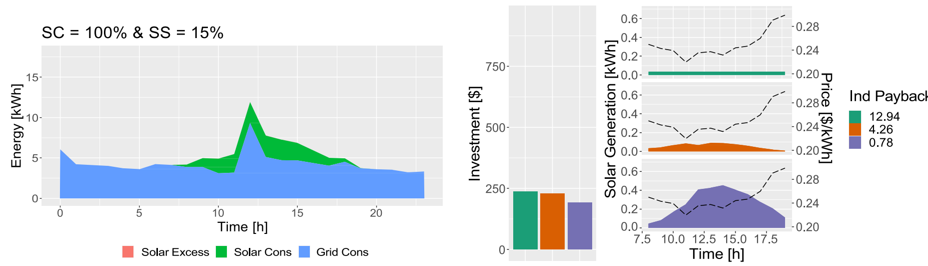
Nevertheless, some limitations should be highlighted. One of the limitations comes from the current Spanish legislation, which doesn't allow the change of the allocation coefficients within four months. Since updating these coefficients is subject to this regulation, our results will



(a) Hourly mean aggregated energy for the whole REC.

(b) Left: investment of three participants. Right: PV allocated for each one (dark fill colour: PV energy which is used; light color: PV energy excess; dotted line: price signal).

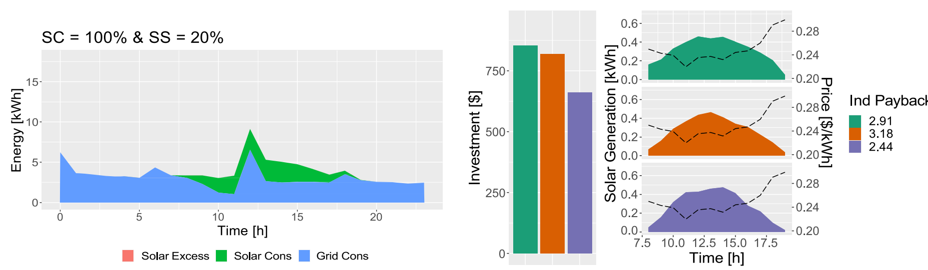
Fig. 6. . Analysis for *profitable* REC: aggregated for the whole REC and individualized for the participants with the lowest, mean and minimum payback.



(a) Hourly mean aggregated energy for the whole REC.

(b) Left: investment of three participants. Right: PV allocated for each one (dark fill colour: PV energy which is used; light color: PV energy excess; dotted line: price signal).

Fig. 7. Analysis for sustainable REC: aggregated for the whole REC and individualized for the participants with the lowest, mean and minimum payback.



(a) Hourly mean aggregated energy for the whole REC.

(b) Left: investment of three participants. Right: PV allocated for each one (dark fill colour: PV energy which is used; light color: PV energy excess; dotted line: price signal).

Fig. 8. Analysis for *novel* REC: aggregated for the whole REC and individualized for the participants with the lowest, mean and minimum payback.

have an error margin that will depend on the constancy of the participants in terms of their consumption habits. For example, suppose some participants were to change their consumption patterns within that period. In that case, the restructuring of energy allocation could only be done at the beginning of the next period. Such a limitation would be reduced if the legislative framework reduced the period for allocation updates. Therefore, we believe that the impact of this study will be even more relevant when this change in the regulation is applied. On the other hand, more sophis- ticated consumption forecasting models could be used to avoid such unexpected consumption changes [32].

5. Conclusions and future work

A transition towards a more sustainable energy model where citizenship will play a central role is taking place in all European countries. REC have the potential to become a key instrument for this citizen-driven energy transition. However, if the grouping of prosumers and their respective allocation of energy is not implemented smartly, the high potential of REC will remain untapped. Furthermore, the lack of optimized design and operation of REC may lead to in- creased energy generation but decreased self-consumption, with many surpluses dispatched to the grid and lowering economic profits. This study presents and validates a set of optimization procedures to assist the REC' participants in both the planning and the operation phases and, therefore, support them to get the maximum environmental and economic profit from their collective initiative. Considering the unpredictable energy usage patterns, radically intermittent characteristics of RES, and dynamic electricity price, it would be difficult for participants of a REC to intelligently share their energy with others and thus minimize the overall cost of the whole community. This research contributes to filling the gap regarding tools for the better development of REC. Specific GAs were developed to find the best combination of users and allocation of the generated solar energy. The presented framework is a practical and extensible approach which will assist in decision-making for REC.

More specifically, this study presents a GA including problem-specific com- ponents designed to enable the algorithm to address the complexities of the problems. Regarding the Selection optimization, different novel aspects were presented. Firstly, a procedure to calculate the maximum number of partici- pants to reduce the search space. Secondly, a novel encoding to represent the possible solutions. Thirdly, a novel technique for space ordering to generate regularity in search space according to self-sufficiency. Moreover, the mathe- matical definition of the fitness function that allowed us to achieve the stated objective. Finally, the definition of a special operator to replace duplicate and infeasible individuals that avoids having duplicate individuals. For the Alloca- tion optimization, the following problem-specific concepts were added. First of all, a novel encoding to represent the possible solutions was presented. Then, the definition of two fitness functions was developed. And finally, a decoder as repair mechanism to modify infeasible individuals which do not comply the equality constraint was designed.

To conclude, this methodology sets the basis for the design of tools to help REC participants increase their economic revenues and their positive impact on the environment. If people are not advised with the right tools, the great potential of REC will not be realized. Therefore, this study is essential to provide important guidance for stakeholders.

Finally, further work could assess more in deep REC. For example, consider the inclusion of energy storage systems (such as batteries or electric vehicles), or different consumer typologies (such as those with commercial or industrial loads), or the use of other energy sources to hybridize (as could be the wind). All these possible add-ons would make the tool more versatile, allowing it to adapt to a wider range of REC scenarios, becoming a highly valuable asset for tomorrow's power system.

CRediT authorship contribution statement

Florencia Lazzari: Conceptualization, Methodology, Software, Writing – original draft. **Gerard Mor:** Conceptualization. **Jordi Cipriano:** Supervision. **Francesc Solsona:** Writing – review & editing. **Daniel Chemisana:** Writing – review & editing. **Daniela Guericke:** Conceptualization, Writing – review & editing.

Declaration of Competing Interest

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

Data availability

The data that has been used is confidential.

Acknowledgments

This work was developed during the PhD thesis of F. Lazzari. D. Chemisana thanks ICREA for the ICREA Acad'emia. This work emanated from research conducted with the financial support of the European Commission through the POCTEFA project EKATE, grant agreement EFA312/19 and the H2020 project ePLANET, grant agreement 101032450.

References

[1] Capellán-Pérez I, Campos-Celador Á, Terés-Zubiaga J. Renew-able energy cooperatives as an instrument towards the energy transition in spain. Energy Policy 2018;123:215–29.
[2] Caramizaru A, Uihlein A. Energy communities: an overview of energy and social innovation. Publications Office of the European Union; 2020.
[3] Prol JL, Steininger KW. Photovoltaic self-consumption is now profitable in spain: Effects of the new regulation on prosumers' internal rate of return. Energy Policy 2020;146:111793.
[4] Hanke F, Guyet R, Feenstra M. Do renewable energy communities deliver energy justice? exploring insights from 71 european cases. Energy Res Soc Sci 2021;80:102244.
[5] Walker A, Kwon S. Design of structured control policy for shared energy storage in residential community: a stochastic optimization ap- proach. Appl Energy 2021;298:117182.
[6] Inês C, Guilherme PL, Esther M-G, Swantje G, Stephen H, Lars H. Regulatory challenges and oppor- tunities for collective renewable energy prosumers in the eu [Online]. Available: Energy Policy 2020;138:111212. https://www.sciencedirect.com/science/article/pii/S0301421519307943.
[7] "State of the Energy Union report," https://ec.europa.eu/commission /presscorner/detail/en/IP 21 5554, accessed: 2022-12-01.
[8] Kazmi H, Munné-Collado I, Mehmood F, Syed TA, Driesen J. To- wards data-driven energy communities: a review of open-source datasets, models and tools. Renew Sustain Energy Rev 2021;148:111290.
[9] Duvignau R, Heinisch V, Göransson L, Gulisano V, Pap- atriantafilou M. Benefits of small-size communities for continuous cost- optimization in peer-to-peer energy sharing. Appl Energy 2021;301:117402.
[10] Ghiani E, Giordano A, Nieddu A, Rosetti L, Pilo F. Planning of a smart local energy community: the case of Berchidda Municipality (Italy). Energies 2019;12(24):4629.
[11] Vahidinasab V. Optimal distributed energy resources planning in a com- petitive electricity market: Multiobjective optimization and probabilistic design. Renew Energy 2014;66:354–63.
[12] Weckesser T, Dominković DF, Blomgren EM, Schledorn A, Madsen H. Renewable energy communities: Optimal sizing and distribu- tion grid impact of photo-voltaics and battery storage. Appl Energy 2021;301:117408.
[13] Zarei M, Maghrebi M. Targeted selection of participants for energy efficiency programs using genetic agent-based (gab) framework. Energ Effi 2020;13(5):823–33.
[14] G. Ye, G. Li, D. Wu, X. Chen, and Y. Zhou, "Towards cost minimiza- tion with renewable energy sharing in cooperative residential communities," IEEE Access, vol. 5, pp. 11 688–11 699, 2017.
[15] Liu J, Yang H, Zhou Y. Peer-to-peer trading optimizations on net- zero energy communities with energy storage of hydrogen and battery ve- hicles. Appl Energy 2021;302:117578.
[16] Rodrigues DL, Ye X, Xia X, Zhu B. Battery energy storage sizing optimisation for different ownership structures in a peer-to-peer energy sharing community. Appl Energy 2020;262:114498.
[17] Manso-Burgos Á, Ribó-Pérez D, Alcázar-Ortega M, Gómez- Navarro T. Local energy communities in Spain: Economic implications of the new tariff and variable coefficients. Sustainability 2021;13(19):10555.
[18] M. Mutingi and C. Mbohwa, "Grouping genetic algorithms," Advances and Applications. Switzerland: Springer International Publishing, vol. 243, 2017.
[19] Z. Michalewicz and M. Schoenauer, "Evolutionary algorithms for constrained parameter optimization problems," Evol. Comput., vol. 4, no. 1, p. 1–32, mar 1996. [Online]. Available: https://doi.org/10.1162/evco.1996.4.1.1.
[20] Dembski WA, Marks II RJ. Conservation of information in search: measuring the cost of success. IEEE Trans Syst Man, and Cybernetics-Part A: Syst Humans 2009;39(5):1051–61.
[21] C. Munien, S. Mahabeer, E. Dzitiro, S. Singh, S. Zungu, and A. E.-S. Ezugwu, "Metaheuristic approaches for one-dimensional bin packing prob- lem: A comparative performance study," IEEE Access, vol. 8, pp. 227 438– 227 465, 2020.
[22] Simon D. Evolutionary optimization algorithms. John Wiley & Sons; 2013.
[23] X. Li, Z. Zhao, and K. Zhang, "A genetic algorithm for the three- dimensional bin packing problem with heterogeneous bins," in IIE Annual Conference. Proceedings. Institute of Industrial and Systems Engineers (IISE), 2014, p. 2039.
[24] Coello CAC, Lamont GB, Van Veldhuizen DA, et al. Evolutionary algorithms for solving multi-objective problems, vol. 5. Springer; 2007.
[25] Miettinen K. Nonlinear multiobjective optimization, vol. 12. Springer Science & Business Media; 2012.
[26] Eiben AE, Smit SK. Parameter tuning for configuring and an- alyzing evolutionary algorithms. Swarm Evol Comput 2011;1(1):19–31.
[27] Scrucca L. Ga: A package for genetic algorithms in r. J Stat Softw 2013;53:1–37.
[28] ——, "On some extensions to ga package: hybrid optimisation, parallelisa- tion and islands evolution," arXiv preprint arXiv:1605.01931, 2016.
[29] Deb K, Pratap A, Agarwal S, Meyarivan T. A fast and elitist mul- tiobjective genetic algorithm: Nsga-ii. IEEE Trans Evol Comput 2002;6(2):182–97.
[30] Koziel S, Michalewicz Z. Evolutionary algorithms, homomor- phous mappings, and constrained parameter optimization. Evol Comput 1999;7(1):19–44.
[31] "Ministerio para la transición ecológica," 2021. [Online]. Available: https://energia.gob.es.
[32] Lazzari F, Mor G, Cipriano J, Gabaldon E, Grillone B, Chemisana D, et al. User behaviour models to forecast electricity consumption of residential customers based on smart metering data. Energy Rep 2022;8:3680–91.