Energy 295 (2024) 131033

Contents lists available at ScienceDirect

Energy

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

Fair investment strategies in large energy communities: A scalable Shapley value approach

Raquel Alonso Pedrero [a,*], Paolo Pisciella [a], Pedro Crespo del Granado [a,b]

[a] *Department of Industrial Economics and Technology Management, Norwegian University of Science and Technology, Trondheim, Norway*
[b] *Department of Smart Grids, SINTEF, Trondheim, Norway*

ARTICLE INFO

Keywords:
Energy communities
Cooperative game theory
Investment models
Investment strategies
Payoff allocation

ABSTRACT

Energy communities based on joint investments of energy assets might enable the democratisation of the power system by enabling all individuals to become active participants. Adopting this investment strategy requires ensuring fair economic value distribution among investors. Some studies propose using the Shapley value, given its uniqueness in fairly allocating value. However, this method suffers from scalability issues due to its computational demands. This paper introduces the Nested Shapley value as a new sharing mechanism to allocate profits to members of large coalitions fairly, thereby addressing the fairness-scalability dilemma. In conjunction with a cooperative investment model for communities with multi-dwelling buildings, the method is applied to assess individual preferences toward investment strategies in a real-world case study. The Nested Shapley value is demonstrated to encourage residents to opt for joint investments over individual strategies. Also, if combined with adequate governance structures, this payoff allocation could lead to the selection of resource-efficient investment strategies within residential communities. Furthermore, the Nested Shapley value is proven to satisfy two of the four fairness axioms defined for the Shapley value. A third one is also satisfied, albeit under specific conditions.

1. Background and motivation

Energy communities (ECs) play a crucial role in promoting the adoption of distributed energy resources and fostering active involvement of residential end users in the power system. These initiatives are built on the voluntary participation of consumers in energy-related activities, offering a range of social, economic, and environmental benefits [1]. Examples of these activities include providing flexibility services to stakeholders (e.g., distribution system operators) and facilitating peer-to-peer electricity exchanges.

Adopting ECs demands attention to design considerations, notably [2]: (i) financing mechanisms for energy assets, such as PV panels, storage technologies, and smart meters; and (ii) strategies for sharing local resources and economic outcomes among community members.

ECs can fund their energy assets from diverse sources, ranging from community members to local actors, retailers, and system operators. When community members cannot bear the initial capital expenditure, third-party financing becomes an option [3]. Alternatively, end users might opt for self-investing, either individually or collectively [4].

Furthermore, ECs must establish a resource and value allocation framework that guarantees the return on investment. In partnerships with third-party investors, contracts may be required to set specific terms, such as rates per unit of electricity delivered to the community. Alternatively, in scenarios where members individually own small-scale assets, adopting a local energy market could be a viable option, as it allows community members to trade energy products and define prices using market mechanisms.

However, defining how to allocate economic benefits or local resources in ECs with joint investments may not be straightforward. Collective investing consists of community members pooling funds for the assets, leading to a division of asset ownership. This leads to questions such as: *How should local energy resources be distributed among members with joint asset ownership? How does this allocation translate into the electricity bill of each member?* Essentially, members contributing to the community should receive economic rewards, whereas those who utilise local resources should be billed accordingly [5].

Differences in treatment among asset owners might lead to the division of the community into smaller communities. Members may be less willing to cooperate if they perceive unfair division of benefits and costs. Therefore, ensuring that participants are charged or rewarded fairly is vital for the overall efficiency and sustainability of the EC. In addition to fostering cooperation, choosing an effective cost/benefit allocation method brings several advantages [6]: it prevents free-rider behaviour, provides economic signals for more cost-efficient behaviours,

* Corresponding author.
E-mail address: raquel.a.pedrero@ntnu.no (R. Alonso Pedrero).

https://doi.org/10.1016/j.energy.2024.131033
Received 5 October 2023; Received in revised form 14 February 2024; Accepted 17 March 2024
Available online 19 March 2024

can guarantee economic efficiency, and strengthens the social acceptance of ECs.

Cooperative game theory delves into the dynamics among rational agents collaborating to optimise their utilities. A primary aim of this mathematical framework is to develop methods to distribute the coalition's returns among members in a manner that satisfies certain criteria. Typically, these methods ensure *efficiency* and *individual rationality*. The former indicates that the value is fully allocated among members, while the latter means that members benefit more from joining the coalition than from operating alone. Nevertheless, these two features do not guarantee the full satisfaction of members. One reason can be when members could achieve higher payoffs by forming smaller subcoalitions. Payoffs that avoid this situation are said to be in the *core* of the game and to stabilise the coalition. The other reason is that agents are not treated fairly; for instance, if two agents contribute the same to the coalition but do not get the same payoff.

The Shapley value is a cost allocation technique that stands out for its unique ability to guarantee fair treatment in coalitions [7]. Therefore, adopting Shapley values in ECs could mitigate governance issues derived from the unfair distribution of resources. Nevertheless, a notable drawback of the method is its computational intensity, which grows exponentially as the number of agents in the coalition increases, restricting its use in large-scale communities.

The aim of this paper is to tackle the issue of fairness in large-scale coalitions where applying the traditional Shapley value is impractical. The paper introduces the *Nested Shapley value*, a method designed based on the Shapley value that reduces the computational overhead and fairly distributes value in large-scale coalitions. To demonstrate its practical application, the method is applied to a real case involving a multi-dwelling building neighbourhood consisting of 250 apartments. In this case study, the Nested Shapley value is utilised together with an investment optimisation model to compare the final payoffs of residents across three potential investment strategies in energy assets: (i) individual investments, (ii) collective investments on a building-by-building basis, and (iii) a unified investment by the entire neighbourhood. By computing the division of payoffs, it is possible to compare the willingness to adopt the three strategies. The paper also dives into the importance of governance structures in establishing investment strategies in ECs.

The remaining sections are structured as follows. First, the paper presents existing literature on value allocation methods in ECs, with a particular focus on the application of cooperative game theory and the Shapley value. Thereafter, the foundational theories behind the Shapley value are elaborated in Section 3. Then, the definition of the Nested Shapley value is detailed in Section 4. The specifics and results of applying the method to the real-world case study are presented in Sections 6 and 7, respectively. The paper concludes with final remarks in Section 8.

2. Literature review

Self-investment in ECs typically occurs in two forms: *individual self-investments* where community members individually invest in resources [8], and *joint self-investments* where instead members collectively invest in shared assets [9]. Individual self-investment can lead to adopting *local energy markets*, where participants trade energy products to optimise their utilities individually [10]. In such a setup, local resources are priced following a market structure, which can be classified into decentralised markets, where participants directly interact with each other, and centralised markets, which require a mediator that enables the interactions [11].

Conversely, ECs that opt for joint self-investments cannot rely on marketplaces to price the resources since members own the shared assets collectively. Here, the challenge lies in determining a fair allocation of costs and revenues among members. Chan et al. [2] highlight the importance of rate design for adopting energy communities and explain heuristic methods to divide the value generated, for instance, based on ownership shares. A comprehensive review of traditional tariff designs is presented in [6] to identify their applicability at the local level. Introducing heuristic tariff structures has the advantage of providing a transparent and stable procedure to understand their involvement in the community. Nevertheless, they cannot guarantee desired properties. For instance, they may fail to ensure that participants' payoffs cover the cooperative value (*efficiency* property) since these are often established before operations. Another example is that they might not adequately reflect each participant's contribution, potentially leading some participants to perceive certain tariffs as unfair. This differs from allocation methods proposed in cooperative game theory that are systematically designed to optimise specific objectives.

2.1. Cooperative game theory for value allocation in energy communities

Cooperative game theory is a mathematical framework that studies the strategies and willingness to collaborate of agents who engage in cooperation. In the energy sector, cooperative game theory has been applied to activities such as cooperation between actors in the gas sector [12], allocation of firm energy rights in hydro plants [13], and allocation of transmission grid costs to grid users [14].

Coalitional game theory is expected to be a key analytical framework for designing future smart grids, where there is a need to coordinate many autonomous and smart-control assets [15]. For example, Saad et al. [16] proposes an energy trading algorithm that autonomously forms coalitions of microgrids to minimise the power losses in the distribution grid.

One main application of cooperative game theory is to determine solution methods for allocating the payoff of coalitions among agents. Among all the methodologies, the Shapley value [7] stands out for capturing the unique number reflecting the marginal contribution of each agent in a coalition considering all possible sub-coalitions. Essentially, the value is the only value in the payoff space that ensures axiomatic fairness. As such, it is a popular method to apply, also among energy community studies. For instance, Tveita et al. [17] performs a sensitivity analysis on the energy storage capacity in a community to evaluate if different storage capacities may affect participants' preferences on Shapley values over another allocation method. Also, the Shapley value was deployed by Han et al. [18] to divide the benefits obtained by a group of prosumers owning storage technologies. The cooperative game proposed by the authors reflects one of the characteristics of this method: the Shapley value is not necessarily in the core. In other words, it does not guarantee the formation of stable coalitions.

The study by Abada et al. [19] explores the stability of EC under different technical conditions. Interestingly, they show the Shapley value lies within the core when the community has a concave investment cost function of solar panels. This occurs because the game becomes convex under this condition. However, Fleischhacker et al. [20] demonstrates that the Shapley value may not be at the game's core when the community has a limited area suitable for PV panels. The authors show how two managerial strategies can help to guarantee the long-term durability of ECs in this case: expanding solar resources and adopting renting payments. Alternatively, Cheng et al. [21] presents a method that modifies the Shapley Value to ensure a stable coalition even under non-convex games.

Table 1 compares several studies according to their focus on stability and fairness concerns, the techniques employed and the number of agents assessed. Note that the table only includes papers that utilise the Shapley value or approximation methods, excluding studies that examine other methodologies associated with stability concerns, such as the nucleolus. While these studies make significant contributions to our understanding of the Shapley value's application in energy communities, most research (i) predominantly concentrates on issues related to stability, and (ii) operates under the assumption of relatively small coalitions, often comprising a limited number of participants.

Table 1
Review of some papers applying the Shapley value to energy communities.

Paper	Stability	Fairness	Techniques used	No. agents
Fioriti et al.	x	x	Shapley-Core, Shapley-Nucleolus, MinVar-Nucleolus	10
Gjorgievski et al.	x	x	Virtual net-metering, MinVar, Shapley and heuristic rules	15 agents (600 ECs)
Fleischhacker et al.	x		Nash bargaining and Shapley	4
Chis & Koivunen			Shapley	2 groups
Han et al.	x		Shapley and Nucleolus	Up to 14
He et al.			Shapley and two-factor method* (proposed)	3
Cremers et al.		x	Stratified expected value (proposed), adaptative sampling and last marginal contribution	200
Lee et al.		x	Asymptotic Shapley value (proposed)	2
Norbu et al.		x	Last marginal contribution	200
Tveita et al.	x		Shapley and Nucleolus	4
Abada et al.	x		Shapley, MinVar and heuristic rules	6
Cheng et al.	x		Shapley and Fine-tuned value* (proposed)	3

Regarding their focus, none of the reviewed studies offer a comprehensive formal mathematical analysis of the fairness axioms of the proposed methods. Gjorgievski et al. [22] delve into various social interpretations of 'fairness' and categorise methods according to their alignment with the concept of marginal contribution. However, this study does not examine the relationship between the allocation methods explored and the fairness axioms satisfied by the Shapley value.

Regarding the size of the coalition, computing the Shapley values for large coalitions implies a high computational burden that makes the problem intractable. Consequently, the applicability of traditional Shapley values methodologies in ECs is limited in real-life communities [23]. To overcome the computational challenge, different authors have considered approximated versions of the Shapley Value. A possible method is to derive the *last marginal contribution*, where it only requires computing players' marginal contribution to the entire coalition. This is calculated as the difference between the total value of the EC and the value of the community without that player. This method was, for instance, applied by Norbu et al. [24] to distribute value in a community of two hundred players. This allocation method significantly reduces the computational time and burden compared to the traditional Shapley value, but simplifies the value by considering only the contribution to the grand coalition. Thus, not guaranteeing fairness properties.

Alternatively, Lee et al. [25] derives the *asymptotic Shapley Value* for a cooperative game formed by prosumers, for which its computation only requires knowledge of the number of participants and statistical information of agents and resources. Nevertheless, the method is not applicable to communities with flexible assets like storage technologies. Furthermore, Cremers et al. [26] provide a comprehensive review of approximation methods for the Shapley Value and proposes the *stratified expected value*. The stratified expected value is based on the concept of stratum, which is a group of subcoalitions with the same cardinality. This value can be defined as the expected average contribution of an agent i within all the possible strata, assuming that the demands of the rest of the agents are equal. The study demonstrates the effectiveness of solving allocation problems in communities with shared assets. Nonetheless, the method's formulation, examples and empirical validation are all situated within the specific setting of energy communities without extending the discussion to its applicability in other games or energy community settings.

2.2. Contributions of the paper

There is a general agreement that adopting appropriate value distribution methods within ECs is a fundamental design feature for their real-life implementation. As such, there has been a progressive development in examining value allocation methods that can be applied in this context, particularly focusing on stabilising communities. However, fairness in the context of communities might also need attention, as end users perceiving unfair treatment may also be willing to separate from the coalition.

Focusing on the property of fairness is particularly relevant in scenarios where participants cannot divide into subcoalitions or leave the grand coalition. This is the case in contexts of common property regimes, where the resources are indivisible, and all participants have no other option but to collaborate. In such setups, agreements among subcoalitions are all ineffective as resource management is linked to the decisions of the grand coalition [27]. Since there is forced stability and agents are better off in the coalition than without engaging in any activity, the main potential disagreement between members may arise from fairness issues. This scenario applies to energy communities with multi-dwelling buildings, where communal areas – essential for installing energy assets – are under a communal property regime. The unique nature of these shared spaces, which are neither divisible nor exclusively owned, necessitates that all decisions regarding their use for asset installation are made jointly, emphasising the need for fair allocation of payoffs.

Consequently, this paper focuses on determining a fair allocation mechanism that tackles the *scalability-fairness dilemma* of the Shapley value: while fairness is crucial for the long-term viability of ECs, the computational burden of the Shapley value has limited its applicability in large coalitions.

Some studies have already developed promising tools to approximate the Shapley value. Nonetheless, most of these techniques were developed assuming specific designs of ECs (e.g., without storage technologies), limiting its applicability. This is the case, for example, of the asymptotic Shapley Value. Additionally, out of the examined literature on approximation methods for the Shapley value, no contribution includes a discussion on whether their methods uphold any of the fairness axioms associated with the Shapley value.

Regarding the specific literature about ECs, most studies assume the existence of single-family houses and neglect the possibility of adopting ECs in neighbourhoods with multi-dwelling buildings [3]. Interestingly, this particular type of building is conducive to adopting joint investments, given that the available area for installing assets is shared among residents. Research on investments and distribution of benefits among residents from apartment buildings is a topic overlooked in the literature.

This study contributes to the past literature by:

- Proposing the novel method of Nested Shapley value for approximating Shapley values. This method is proven to satisfy three of the four fairness axioms. Also, the method is versatile enough to be applied to games of any nature, not only ECs.
- Formulating an investment optimisation model adapted to ECs with multi-dwelling buildings, including operations and peer-to-peer interactions. This formulation is designed to guarantee proper value distribution among residents of buildings with shared ownership of the installing areas.
- Using the Nested Shapley value to analyse four investment strategies based on their value proposition to individual participants in a large-scale neighbourhood. A numerical case study uses a real-life case from a multi-dwelling neighbourhood in Austria.

- The study addresses democratic governance structures to guarantee the acceptance and legitimacy of investment strategies in multi-dwelling neighbourhoods.

3. Theoretical background

The upcoming section provides an introduction to two methodologies that are closely related to the development of the Nested Shapley value. Initially, the mathematical description and fairness axioms of the Shapley value will be presented. Secondly, the allocation method of the Owen value will be introduced. Presenting the Owen value in this section is because the proposed Nested Shapley value is based on the same principle: clustering agents.

3.1. The shapley value

In cooperative game theory, a coalitional game formally consists of a set of players $\mathcal{N}$, named the grand coalition, along with a characteristic function v, which assigns a value to each subset of agents $i \in \mathcal{N}$. According to the definition of the Shapley value, the payoff that a player i receives in a coalitional game is computed as

$$\Phi_i = \sum_{S \subseteq \mathcal{N} \setminus \{i\}} \frac{|S|!(|\mathcal{N}|-|S|-1)!}{|\mathcal{N}|!} (v(S \cup \{i\}) - v(S)) \tag{1}$$

where $v(S \cup \{i\}) - v(S)$ describes the marginal contribution of i in subcoalition S. Therefore, the Shapley value is usually defined as the weighted marginal contribution of a player for possible permutations of sub-coalitions. The nomenclature used throughout the paper is included in Appendix A.

The Shapley value's uniqueness is that it is the only payoff allocation that satisfies the following fairness axioms:

1. (Efficiency) $\sum_{i \in \mathcal{N}} \Phi_i = v(\mathcal{N})$. It guarantees cost recovery as the sum of all players' Shapley value equals the total value of the grand coalition.
2. (Symmetry) For every $S \subset \mathcal{N}$, if $v(S \cup \{i\}) = v(S \cup \{j\})$ then $\Phi_i = \Phi_j$. This ensures that two players contributing equally to any coalition receive the same payoff.
3. (Linearity) $\Phi_i(v+w) = \Phi_i(v) + \Phi_i(w)$. If the game played by the coalition is the combination of multiple games, the payoff of any agent is the sum of the payoffs for each game.
4. (Null Player) If a player $i \in \mathcal{N}$ is such that $v(S \cup \{i\}) = v(S)$ for each $S \subseteq \mathcal{N} \setminus \{i\}$ then $\Phi_i = 0$. Therefore, agents that do not add value to any of the possible coalitions receive a zero payoff.

Despite satisfying these properties, computing the Shapley values for all agents in a coalitional game necessitates evaluating the cost value function of $2^{|\mathcal{N}|}$ instances. Therefore, the computational burden grows exponentially with the number of agents. This gives rise to the fairness-scalability dilemma of the Shapley value: while it guarantees fair allocations, it is not scalable for large coalitions. To illustrate, a coalition of 100 agents would require computing $1.26 \cdot 10^{30}$ instances of the characteristic function.

3.2. The Owen value

Another cost allocation method in cooperative game theory is the Owen value [28]. This solution concept extends the original Shapley value by assuming cooperative games with pre-existing coalitional structures of agents. This methodology is sometimes conceived as the *coalitional value of the Shapley value* [29].

A coalitional structure, denoted as $P = \{P_1, P_2, \dots, P_p\}$, represents a partition of the set of players $\mathcal{N}$. Each element of the coalitional structure is referred to as a union. A coalition structure satisfies two conditions: first, the union of all unions forms the grand coalition $\cup_{a=1}^{p} P_a = \mathcal{N}$; and second, each agent is associated with only one union of the coalitional structure, i.e., $P_a \cap P_b = \emptyset$ for $a \neq b$.

Computing the Owen value involves a two-step process. Initially, the Shapley values of each union $P_a \in P$ are computed based on the game played amongst them. Then, the payoff each coalition receives is split among its members, once again using the Shapley value. In this latter stage, the payoff of an agent is computed considering its contribution to all possible subsets within its union and with external unions. The Owen value of player i which belongs to the union $P_a \in P$, i.e., $i \in P_a \in P$ is

$$\phi_i^O = \sum_{Q \subset P \setminus \{P_a\}} \sum_{S \subset P^a \setminus \{i\}} \frac{|Q|!(|P|-|Q|-1)!\,|S|!(|P^a|-|S|-1)!}{|P|!\,|P^a|!} \times M_i(Q,S) \tag{2}$$

where $M_i(Q,S) = c(Q \cup S \cup \{i\}) - c(Q \cup S)$.

This method has been axiomatically characterised and satisfies the properties of efficiency, additivity, the dummy player, and symmetry -both within individual unions and between different unions. Additionally, it is less computationally expensive than the Shapley value, as it does not consider games played between subsets of agents from different unions. Nonetheless, the Owen value still has a computational challenge related to the number of unions considered. If this number is large enough, computing the Shapley value of each union might still be intractable. The following section presents the Nested Shapley value. Although inspired by the concept in the Owen value of forming clusters of agents, the proposed methodology tackles the limitation of the first step.

4. The Nested Shapley value

The Nested Shapley value is introduced here as an allocation method that approximates the Shapley Value by applying a clustering algorithm to group agents into predefined clusters. Specifically, the agents, $i \in \mathcal{N}$, are organised into nested coalitions that form an ordered nodal tree with q layers and M nodes. All the nodes of the tree are contained in set $F = \{1, \dots, M\}$. The structure of the ordered tree is defined as follows:

1. Each node n is associated with a coalition of agents C_n,
2. The coalition of agents for the root node corresponds to the grand coalition, $C_1 = \mathcal{N}$.
3. The coalitions of agents associated with the nodes in the last layer (i.e., the leaf nodes), $n \in F^q$, contain a single agent i, such that $|C_n| = 1$ for all $n \in F^q$.
4. Every node n containing only one element belongs to the set of nodes F^s.
5. Every node in F^s belongs to the set of nodes in the last layer F^q.
6. Every node n is linked to a parent node $m \in F_n^P$, except the root node; and a set of children nodes $s \in F_n^C$, except for the leaf nodes.
7. The coalition of agents associated with node n is the union of the coalitions of agents associated with its child nodes $s \in F_n^C$, such that $\bigcup_{s \in F_n^C} C_s = C_n$ and $C_m \cap C_s = \emptyset$ where m and s share the same parent node n,

Fig. 1 illustrates an example of ordered tree with $q = 3$ and $M = 6$ for a coalition of four agents $\mathcal{N} = \{1, 2, 3, 4\}$.

In the methodology to calculate the Nested Shapley value, three interconnected elements play essential roles:

1. **Nodal Shapley value** (ϕ_n^*): Calculated for each node n in the tree structure, this value is computed as a standard Shapley value considering the coalition of nodes F_m^C, where m is the parent node of n. Therefore, it considers only the characteristic functions of coalitions of nodes belonging to the parent set.
2. **Corrected Shapley value** (ψ_n^*): This value adjusts the Nodal Shapley values to ensure that the sum of the Corrected Shapley values for all leaf nodes associated with each agent equals the total value of the grand coalition $v(\mathcal{N}) = v(C_1)$. Essentially, it guarantees meeting the efficiency property.

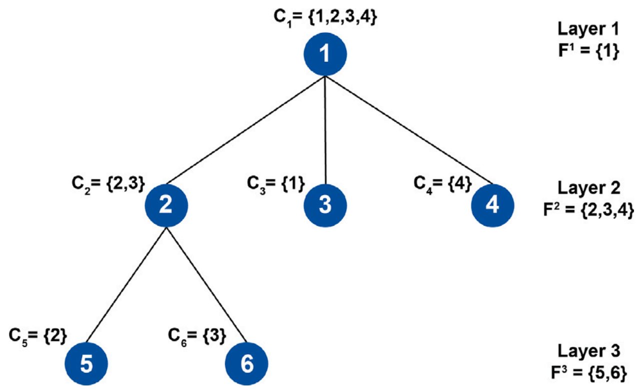


Fig. 1. Example of an ordered tree.

3. **Nested Shapley value** (ψ_i): This is the final payoff assigned to each agent i. It equals the Corrected Shapley value assigned to the leaf node to which the agent i belongs.

The following equations capture the formal relationship between these three elements:

$$\phi_n^* = \sum_{\substack{S \subseteq F_m^C \setminus \{n\} \\ m \in F_n^P}} \frac{|S|!(|F_m^C| - |S| - 1)!}{|F_m^C|!}\left(v\left(\bigcup_{s\in S} C_s \cup \{C_n\}\right) - v\left(\bigcup_{s\in S} C_s\right)\right) \quad (3)$$

$$\psi_n^* = \frac{\phi_n^*}{v(C_m)}\psi_m^*, \qquad \forall\, n \setminus \{1\}, \forall\, m \in F_n^P \quad (4)$$

$$\psi_i = \psi_n^*, \qquad n \in F^q \cap F_i \quad (5)$$

where F_i corresponds to the set of all nodes associated with agent i.

In Eq. (4), the adjustment applied to the Nodal Shapley value ensures that the sum of the Corrected Shapley values of each child node in F_m^C equals that of their parent node m. This recursive adjustment throughout the tree confirms that the Nested Shapley value satisfies the efficiency property, as detailed in Appendix D.1. Furthermore, given that the Corrected Shapley value of a node ψ_n^* is tied to that of its parent node ψ_m^*, obtaining the Corrected and Nodal Shapley values necessitates a top-down computational approach. This process begins with the root node and extends downward to the leaf nodes, calculating the Corrected Shapley values at each level. Algorithm 1 outlines the procedure for calculating the Nested Shapley values for agents organised in an ordered tree.

Algorithm 1 Nested Shapley Value Algorithm

compute $v(\mathcal{N})$
for $n = 1, \dots, M$ **do** for each $S \subset F_n^C$ compute $v(\cup_{m\in S} C_m)$
end for
for $n = 1, \dots, M$ **do** compute ϕ_n^*
end for
$\psi_1^* = v(C_1) = v(\mathcal{N})$
for $n = 2, \dots, M$ **do**
  for $m \in F_n^P$ **do** compute $\psi_n^* = \frac{\phi_n^*}{v(C_m)}\psi_m^*$
  end for
end for
for $i = 1, \dots, N$ **do**
  for $n \in F^q \cap F_i$ **do** $\psi_i = \psi_n^*$
  end for
end for

Unlike the Owen value, the Nodal Shapley value of a node is computed for nodes sharing the same parent, such that coalitions between members of different clusters are not included. This reduces the number of coalitions considered, thereby enhancing the computational burden. Also, to the authors' knowledge, no existing literature presents the Owen value in a way that involves nesting clusters of agents within an ordered tree.

Furthermore, the Nested Shapley value method reduces the number of instances of the characteristic function compared to applying the Shapley value method from $2^{\mathcal{N}} - 1$ to $\sum_{n\in F\setminus F^s}(2^{|F_n^C|} - 1)$. In particular, the number of computations, z, can be expressed in a closed-form formulation as Eq. (6) (refer to Appendix B for the derivation of this equation). This formulation is derived considering symmetrical ordered trees, where each node has the same number of child nodes c.

$$z = \frac{c^{q-1} - 1}{c - 1}(2^c - 1) \quad (6)$$

The computational burden of the Nested Shapley value with symmetric trees depends on the number of layers, q, and players, N, considered. Fig. 2 compares the computational burden of the Nested Shapley value with that for the Shapley value and the stratified expected value from Cremers et al. [26].

The Nested Shapley value presents computational advantages compared to the other two methods. For example, in a coalition of 125 agents, the Nested Shapley value requires 916 instances, assuming an ordered tree structure with $q = 4$ and $n = 5$. In contrast, calculating the traditional Shapley value for the same coalition demands $4.25 \cdot 10^{37}$ instances, and the stratified expected value needs 31,125 (calculated as $2N^2 - N$).

Additionally to reducing the number of computations and the efficiency property, the Nested Shapley value satisfies the null player and symmetry properties. The proofs are included in Appendices D.2 and D.3, respectively. In particular, the symmetry property holds if agents that equally contribute to every possible sub-coalition share the same parent node in the last layer. Note the extreme case where the number of leaf nodes of the agents equally contributing to the coalition (associated with the same parent node) is large enough to compromise the computation of the Nodal Shapley value. Two solutions can be adopted in this case: (i) continue generating children nodes symmetrically such that the additional nodes in each layer have the same Corrected Shapley value, this only holds for an even number of agents, or (ii) divide the Corrected Shapley value of the parent node among them equally.

Furthermore, ensuring holding the symmetry property leads to a paradox: determining which agents contribute equally would necessitate calculating the Shapley value. In essence, to ensure that the Nested Shapley value guarantees the symmetry property, one would need to compute the Shapley values that the Nested Shapley method aims to approximate. Nevertheless, the symmetry property indicates how the agents should be clustered together. Namely, agents with similar properties are likely to contribute similarly to coalitions and should thus be grouped into the same nodes along the tree. The following section explores using the k-means algorithm to construct the ordered trees, obtaining clusters that are internally homogeneous and mutually different, thereby approximating the symmetry property.

4.1. Ordered tree using the k-means algorithm

The k-means algorithm [30] divides the agents into coalitions based on shared attributes. If these attributes influence the coalition's value, constructing the ordered tree via the k-means algorithm can aid in satisfying the symmetry property. Thus, this clustering method can provide a systematic process to determine the coalitions of agents C_n associated with each node in the ordered tree.

Formally, the k-means algorithm divides agents into a set of p partitions, $P = \{P_1, P_2, \dots, P_p\}$, based on a set of predefined characteristics $X = \{x_1, x_2, \dots, x_N\}$.

Typically, the criterion in the k-means algorithm is to allocate the data points x of all agents into p different clusters in such a way

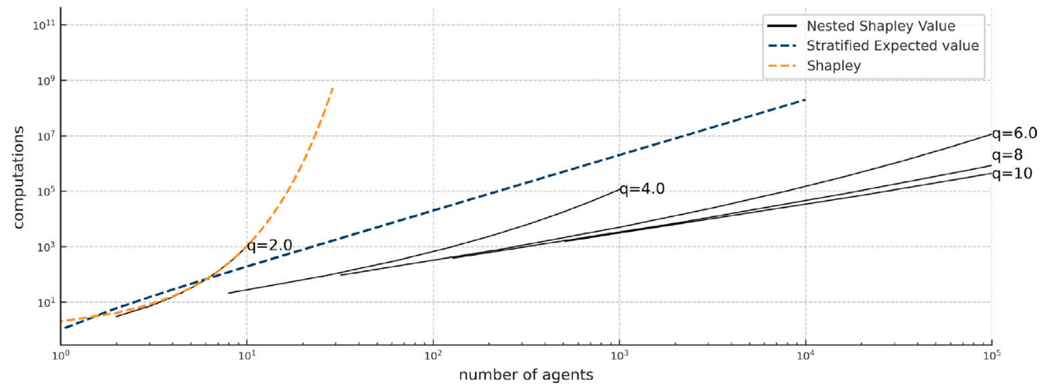


Fig. 2. Computational comparison between the Nested Shapley value, the stratified expected value, and the true Shapley value as a function of the number of players N. The computation for the Nested Shapley value is calculated for symmetric trees with different numbers of layers q. Note that the x-axis is logarithmic.

that the total sum of the squared Euclidean distances between each feature vector x in X and the centroid μ_a (i.e., the cluster centre) of its respective cluster P_a is minimised. Here, μ_a is the average of all points x in a partition a. Mathematically, the algorithm aims to find the partition P that minimises this total:

$$\arg\min_P \sum_{a=1}^{p} \sum_{x \in P_a} || x - \mu_a ||^2 \tag{7}$$

In the ordered tree, each partition corresponds to a node. The children nodes of a particular node are generated by applying the k-means to the set of agents associated with that node. Following the example from Fig. 1, node one is divided into three partitions/nodes using the k-means algorithm.

Moreover, the k-means algorithm requires defining the number of partitions p beforehand. In the context of the Nested Shapley value method, this number p should not exceed a maximum $\overline{p}$. This constraint arises because increasing the number of partitions exponentially increases the computational burden of calculating the Nodal Shapley values ϕ_n^*. Therefore, p should be chosen to balance the trade-off between computational efficiency and the best granularity of the partitioning. This can be done by setting a maximum number of clusters $\overline{p}$.

To determine which granularity of partitions p^* could perform better, the silhouette method is implemented. This is a well-known technique that evaluates the quality of clusters by analysing how similar the properties of agents are to the points of their clusters. A high value of the silhouette coefficient indicates that the agent is well-matched to the assigned cluster. Specifically, it is set such that $p^* = \min\{p \in \{1,\dots,\overline{p}\} : \rho_{ip} \geq 0 \ \forall i \in P_p \text{ and } \sigma_p \geq 0.5\}$. Here, ρ_{ip} refers to the silhouette coefficient of an agent, and σ_p is the average silhouette coefficient for all the agents.

The k-means method is recursively applied to nodes whose size of the coalition $|C_n|$ exceeds $\overline{p}$. This process continues until generating child nodes no longer meet this constraint. Then, for the child nodes with $|C_n| < \overline{p}$, leaf nodes are created for each agent in C_n. Using the example in Fig. 1 and assuming a $\overline{p} = 3$, node two meets this constraint such that the leaf nodes five and six are generated for each agent in C_2. The detailed algorithm outlining the construction of the ordered tree based on the k-means can be found in Appendix E.

Notably, if one aims to approximate the symmetry property, the feature vectors in X must be selected, considering which agents' characteristics influence their contribution to coalitions.

4.2. Numerical example of the Nested Shapley value

To demonstrate the applicability of the algorithm, consider a cooperative game with $\mathcal{N} = \{\text{A, B, C, D, E}\}$, where agents can be clustered in different ordered trees. In this example, four different ordered trees are considered, as shown in Fig. 3.

First, the values of the characteristic function for the subcoalitions $S \subset F_n^C$ for each node $n = 1,\dots,M$ are calculated together with the value of the grand coalition to compute the Nodal Shapley value. The resulting characteristic function values computed in each tree are presented in Table 2.

Any coalition of five agents necessitates $2^5 - 1 = 32$ instances of the characteristic function to compute the true Shapley values. However, applying the Nested Shapley Value, this number is reduced to $z = 14$ for trees 1, 2 and 3 and $z = 18$ for tree 4.

Then, the algorithm proceeds by calculating first the Nodal (Eq. (3)) and Corrected (Eq. (4)) Shapley values. The results obtained for the numerical example are included in Tables 3 and 4, respectively.

The Corrected Shapley value of the leaf nodes equals the Nested Shapley value of its associated agent as stated in Eq. (5). The resulting Nested Shapley values of the agents in this game are shown in Table 5.

The results of the numerical example demonstrate the importance of selecting the ordered tree appropriately. In this case, tree 3 presents the largest relative difference to the true Shapley value for agent E (45%), while the lowest tree with the lowest maximum relative difference is number 4 with only 14% for agent B. A larger example for evaluating the computational accuracy of the Nested Shapley value is presented in

Table 2
Characteristic function values. Each column indicate an ordered tree.

S	1	2	3	4	S	1	2	3	4
(A, B)	23	23			(A, E)			17	
(A, B, C)	26	26		26	(A,)	7	7	7	7
(A, B, C, D)	43	43		43	(B, C)	18	18	18	
(A, B, C, D, E)	45	45	45	45	(B, D)				28
(A, B, C, E)	26	26		26	(B, D, E)				33
(A, B, D, E)			40		(B, E)				12
(A, C)	16	16		16	(B,)	8	8	8	8
(A, C, D)				36	(C,)	6	6	6	6
(A, C, D, E)			42	42	(D, E)	26	26	26	26
(A, C, E)				19	(D,)	19	19	19	19
(A, D)			30		(E,)	2	2	2	2
(A, D, E)			36						

Table 3
Nodal Shapley values.

Node, n	1	2	3	4
1	45.00	45.00	45.00	45.00
2	23.00	23.00	32.33	15.08
3	19.50	2.50	6.33	6.92
4	2.50	19.50	6.33	19.92
5	9.17	10.67	10.00	3.08
6	10.67	6.17	20.50	8.50
7	6.17	9.17	5.50	7.50

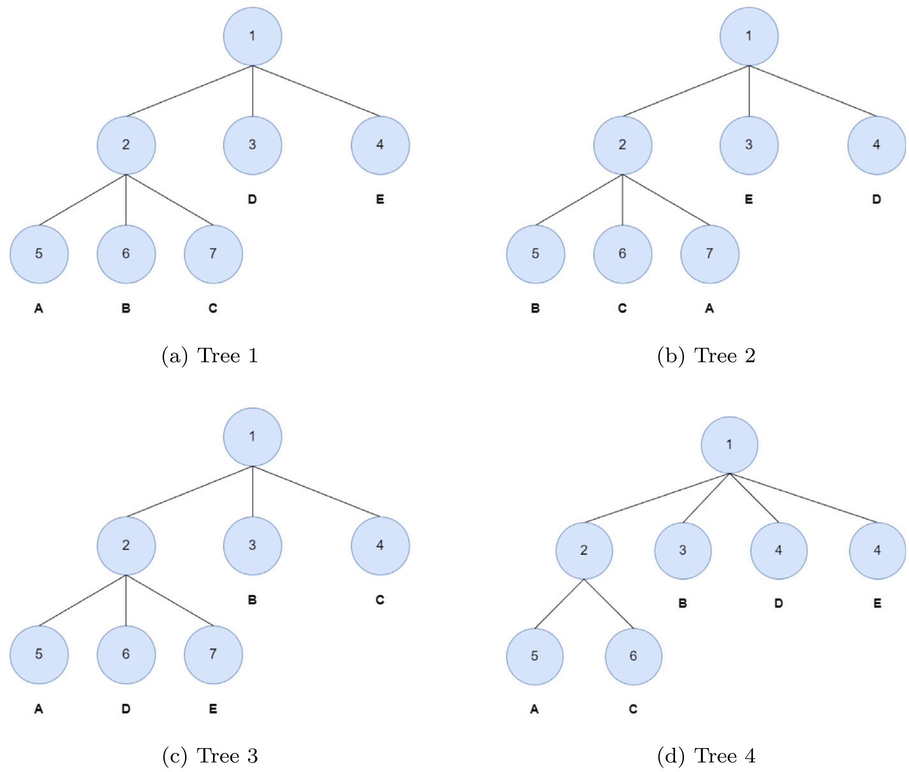


Fig. 3. Trees used for the numerical example.

Table 4
Corrected Shapley values.

Node, n	1	2	3	4
1	45	45	45	45
2	23	23	32.33	15.08
3	19.5	2.5	6.33	6.92
4	2.5	19.5	6.33	19.92
5	8.11	9.44	8.99	3.08
6	9.44	5.46	18.41	8.01
7	5.46	8.11	4.94	7.07

Table 5
True Shapley value for the game and Nested Shapley values obtained applying different ordered trees.

	Shapley	Nested Shapley			
		1	2	3	4
A	7.83	8.11	8.11	8.98	8.01
B	8.08	9.44	9.44	6.33	6.92
C	7.92	5.46	5.46	6.33	7.07
D	19.75	19.50	19.50	18.41	19.92
E	3.42	2.50	2.50	4.94	3.08

Appendix C. The method is compared not only with the true Shapley value but with the stratified expected value presented in [26].

The code for computing the Nested Shapley value and these examples can be found in the GitHub repository.[1]

5. Joint-investment model for communities with multi-dwelling buildings

This section presents the joint-investment model for communities with multi-dwelling buildings. The Nested Shapley value would be employed to divide the costs from this model among community members.

[1] https://github.com/raquelal94/NestedShapley.

Assuming a coalition of dwellings, $h \in \mathcal{H}$, with a cardinality H, interested in reducing the total annual electricity costs by installing energy assets in the shared areas, such as rooftops and garages. The objective of the neighbourhood is to minimise its investment and operational costs as follows:

$$\min_{x,y,b} c(\mathcal{H}) = \text{AIC} + \text{AOC} \tag{8}$$

where

$$\text{AIC} = \overbrace{\sum_{g\in G} a_g (c_g^P + c_g^I)\, x_g^{PG}}^{\text{Generation investment costs}} + \overbrace{\sum_{s\in S} a_{s\in S}(c_s^P x_s^{PS} + c_s^C x_s^C)}^{\text{ESS investment costs}} \tag{9}$$

$$\text{AOC} = \overbrace{c_s^{MP} x_s^{PS} + c_s^{MC} x_s^C}^{\text{ESS Maintenance}} + \overbrace{\sum_{i\in I} \alpha_i \sum_{t\in T}\sum_{h\in \mathcal{H}} \left[(c^G + c_{it}^{Spot} + c^{Tax})\, y_{ith}^G\right]}^{\text{Individual electricity bill}} + \overbrace{\sum_{i\in I} \alpha_i \sum_{t\in T} (c^G + c_{it}^{Spot} + c^{Tax})\, y_{it}^{GA}}^{\text{Shared-areas electricity bill}} \tag{10}$$

The annualised investment costs (AIC) for the generation technologies, $g \in G$ and energy storage systems (ESSs), $s \in \mathcal{S}$, are captured in Eq. (9). To consider the operational level in the investment decisions, the model includes the annualised operational cost (AOC). These cost includes the maintenance costs for ESSs, c^{MP} and c^{MC}, based on their power rates and capacity, x_s^{PS} and x_s^C, the sum of individual electricity bills and the electricity costs incurred in shared areas (Eq. (10)). The individual electricity bills of each households is proportional to their consumption from external sources, y^G, and the sum of grid costs, c^G, spot price, c^{Spot}, and tax costs, c^{Tax}. Additional retailing costs are not included, as they might differ depending on the contracted retailer. Imported electricity for meeting the demand of the shared areas, y^{GA} are assumed to be subject to the same cost structures as the residents of the coalition. The demand of these areas will occur only for charging the batteries, thus no additional loads are considered (e.g., lightning of corridors, elevators).

To reduce the model complexity, the model assumes specific representative times of the year, I. As such, the operations are scaled with the seasonal scaling factor α_i, based on the representative time periods selected (e.g., week per month, week per season, day per month). For simplicity, the model does not include the option of selling excess power to the main grid and additional operational costs related to smart controlling the assets (e.g., energy management systems). The detailed nomenclature of the model can be found in Appendix A.2.

Furthermore, the capacity for installing generation technologies, x^{PG} is restricted by Eq. (11) to consider the physical limits of the buildings, $\overline{G}$, in the coalition, such as the rooftop area. Typically, investment models for ECs allocate specific areas for installing assets to individual agents. In such models, if an agent is not part of the coalition, its available areas cannot be used by others. However, this approach does not apply to multi-dwelling buildings, where common areas such as rooftops are collectively owned. To address this, the model assumes a fixed, shared space for installing the generation assets available for the entire coalition considered.

$$x_g^{PG} \leq \overline{G}_g, \quad \forall g \tag{11}$$

$$x_s^{C} \leq H\overline{S}_s, \quad \forall s \tag{12}$$

$$x_s^{PS} = \lambda x_s^{C}, \quad \forall s \tag{13}$$

This feature is crucial when evaluating the contributions of individual members. In apartment buildings, the value derived from the capacity to install assets in common areas must be collectively attributed to all residents. In contrast, for single-family houses, the value of such capacity is considered an individual contribution. This distinction is crucial for assessing the contributions of individual community members, as the value delivered by available installation areas should be accounted for differently depending on the housing type.

Furthermore, Eq. (12) ensures the maximum installed capacity of storage systems, $\overline{S}$, is proportional to the number of dwellings in the coalition. Here, differently from generation assets, the size of the coalition will determine the available area. This is based on the assumption that batteries are installed in the parking places owned by coalition members. Also, the installed capacity of a battery, x_s^C, is proportional to how much power it can inject, x_s^{PS}. Eq. (13) shows this relationship, using λ as the constant of proportionality.

At the operational level in shared areas, the grid imports from the main grid, y^{GA}, discharges from storage technologies, y^D, and the electricity production must cover the needs for charging the ESSs, y^C, and exports to the residents in the community, y^{EA}. This balance is captured in Eq. (14). In this equation, the electricity production is calculated as the product of a normalised generation profile, G, with the total capacity installed, x_g^{PS}. Moreover, Eq. (15) states that the electricity imported by all dwellings equals the power exported from the shared areas. The ψ factor is included to account for any losses in exporting this power. Then, for each resident, its demand can be met with electricity from common areas, y^A, and the main grid, y^G, as shown in Eq. (16).

$$y_{th}^{GA} + \sum_{s\in S} y_{sit}^{D} + \sum_{g\in G} G_{git}\, x_g^{PG} \geq \sum_{s\in S} y_{sit}^{C} + y_{it}^{EA}, \quad \forall i, \forall t \tag{14}$$

$$\sum_{h\in \mathcal{H}} y_{ith}^{A} = \psi\, y_{it}^{EA}, \quad \forall i, \forall t \tag{15}$$

$$y_{ith}^{G} + y_{ith}^{A} = L_{ith}, \quad \forall i, \forall t, \forall h \tag{16}$$

In regards to the ESS, their state-of-charge, y^{SOC} must always be lower than the capacity invested, x_s^C (Eq. (17)), and there must be intertemporal balance. In the balance Eqs. (18) and (19), the state-of-charge of a battery depends on the state in the previous time, and the charge and discharge at that time. Losses from charging and discharging are captured by the efficiency factors η^C and η^D, respectively. In Eq. (20), the conversion factor β transforms the charging and discharging rate to energy.

$$y_{sit}^{SOC} \leq x_s^{C}, \quad \forall s, \forall i, \forall t \tag{17}$$

$$y_{si1}^{SOC} = SOC_{si} + \eta^C y_{si1}^{C} - 1/\eta^D y_{si1}^{D}, \quad \forall s, \forall i \tag{18}$$

$$y_{sit}^{SOC} = y_{si(t-1)}^{SOC} + \eta^C y_{sit}^{C} - 1/\eta^D y_{sit}^{D}, \quad \forall s, \forall i, \forall t \setminus 1 \tag{19}$$

$$y_{sit}^{C},\ y_{sit}^{D} \leq x_s^{PS}\ \beta, \quad \forall s, \forall i, \forall t \tag{20}$$

Finally, the model includes Eq. (21) where the binary variable b avoids charging and discharging simultaneously in the same time period. This equation makes the problem nonlinear.

$$0 = b_{sit}\ y_{sit}^{C} + (1 - b_{sit}) y_{sit}^{D}, \quad \forall s, \forall i, \forall t \tag{21}$$

To linearise the model, Eq. (21) was transformed using the big M method as in Eqs. (22) and (23), where M is sufficiently large to ensure it does not constraint the original constraints on charging and discharging capacities (Eq. (20)).

$$y_{sit}^{C} \leq M \cdot b_{sit}, \quad \forall s, \forall i, \forall t \tag{22}$$

$$y_{sit}^{D} \leq M \cdot (1 - b_{sit}), \quad \forall s, \forall i, \forall t \tag{23}$$

6. Case study and investment strategies

The case study for applying the Nested Shapley value and the joint-investment model is a demo site from the EU Horizon 2020 project "Sustainable Plus Energy Neighbourhoods" Sy.nikia[2] located in Salzburg, Austria. This neighbourhood comprises 11 closely situated buildings, which add up to 250 dwellings. For a detailed overview of the neighbourhood's technical infrastructure, such as building envelopes and heating/cooling systems, refer to Andresen et al. [32].

Joint investments in solar PV and batteries were assumed to be installed in shared spaces such as rooftops and garages. According to the project's technical assessment by Andresen et al. [32], the neighbourhood's maximum photovoltaic capacity is 667 KWp, utilising south-facing PV panels. Additionally, for each dwelling within a building, a small-scale battery of 4 kWh can be integrated.

For the PV panels, hourly generation profiles were simulated on the renewables.ninja platform[3] [33,34] assuming 10% system loss and a tilt of 35 degrees. The simulation used historical irradiation and temperature data from 2019. Concurrently, day-ahead prices for 2019 were retrieved from the ENTSO-E Transparency platform [35].

Since the demo project is still in the development phase, actual electricity consumption data is unavailable. To address this, the LoadProfileGenerator software[4] [36] was used to generate load profiles for 251 households. These profiles were based on demographic details (e.g., number of family members, occupation) and geographical factors (e.g., irradiation, temperature). The total electricity demand of the neighbourhood resulted in 1238 MWh, averaging around 5000 kWh/y per household. For a detailed breakdown of other investment and operational costs, please refer to Table 6. Finally, the annuity factors for each technology were calculated using an interest rate of 6%.

The load, prices and solar profiles used for this case study are in the GitHub repository.[5]

2 https://www.synikia.eu/.
3 https://www.renewables.ninja/.
4 https://www.loadprofilegenerator.de/references/.
5 https://github.com/raquelal94/NestedShapley.

Table 6
Investments and operational parameters assumed in the case study.

Investment costs					Operational costs		
	Capital cost	*Maintenance costs*	*Lifetime*	*Source*		*Price* (€/kWh)	*Source*
PV	1 038 €/kWp 103 €/kWp[a]		25 years	Fleischhacker et al. [9]	Grid tariff	0.072	Eurostat [31]
ESS	10 €/kW 1 200 €/kWh	0.5 €/kW/a 0.5€/kWh/a	15 years	Fleischhacker et al. [9]	Taxes	0.0645	Eurostat [31]

[a] Installation costs.

6.1. Case definition based on possible business models

Four investment strategies are proposed as available to communities formed by multi-dwelling buildings. These strategies are different from each other based on two primary criteria: the nature of asset ownership, either individual or joint, and the size of the coalitions ranging from individual buildings to groups of buildings. For simplicity and to facilitate a clearer interpretation of the results, feed-in tariffs are not considered in any of the cases. The four models of investment strategies are:

- **Business-as-usual (BAU):** In this strategy, there are no investments in renewables or energy storage. Each apartment covers its electricity demand by purchasing directly from the grid.
- **Individual Dwellings (ID)**: Here, each household determines its optimal investment to meet its electricity demand. Any excess energy produced is not shared with others and is instead curtailed.
- **Energy Communities in Buildings (EC-B)**: In this approach, energy communities are established on a building-by-building basis. All residents within a building share a common smart meter and a unified electricity bill. These ECs decide on the optimal generation and storage capacities to install in shared areas. Any surplus electricity is not traded with the neighbouring buildings.
- **Energy Community in Neighbourhood (EC+)**: Under this strategy, all buildings come together to form a single EC. Shared areas across all buildings are utilised collectively, and there is a unified electricity bill for the entire neighbourhood. Importantly, peer-to-peer energy exchanges between buildings are permitted.

7. Results

The results for the joint investment and the application of the Nested Shapley value are presented in the following section. First, the overall costs and the operational decisions at the community level for the different cases are presented and discussed. The section continues by examining the individual electricity bills of agents which initiates the stability issues addressed through the adoption of democratic governance structures. The section concludes by specifying the details on the Nested Shapley method used in the case study.

7.1. Costs and investments in the neighbourhood

Table 7 presents the capacities of the energy assets and the annual costs borne by the entire neighbourhood for each investment strategy. In the EC-B and EC+ strategies, no batteries were installed. This differs from the ID, as individual households needed to exploit the flexibility from batteries to save their excess production, as peer-to-peer was not an option. Moreover, there was an evident reduction in solar photovoltaic capacity in the EC-B and EC+ cases compared to the ID case, implying that peer-to-peer transactions promoted a more efficient use of local resources.

The resource efficiency achieved through peer-to-peer reduces the environmental footprint, as it allows avoiding overusing material and energy resources to cover energy needs. Another benefit of resource efficiency is reduced total costs for the community. The EC-B strategy realised cost savings of 32.2% and 10.2% compared to the BAU and ID strategies, respectively. Interestingly, these reductions are comparable to those achieved with the EC+ strategy. This suggests that collaboration within the same building was the primary driver of capacity and cost reductions while trading between buildings yielded only marginal benefits.

Table 7
Technology and costs results for the investment strategies considered.

	BAU	ID	EC-B	EC+
Solar (kW)	–	408	337	332
Batteries (kWh)	–	55	0	0
Total costs (€)	220 288	166 240	149 307	148 352
Investment costs (€)	–	30 322	21 438	21 150
Operational costs (€)	220 288	135 917	127 869	127 202

Table 8
Summary of the neighbourhood's operational results for the four cases.

	BAU	ID	EC-B	EC+
PV electricity consumption (MWh)	–	442	524	528
Discharge consumption (MWh)	–	41	–	–
Imports from the grid (MWh)	1 238	754	714	710
Curtailment (MWh)	–	829	556	538

Several factors determine the benefit of collaboration among buildings over only per-building sharing. First, collaboration among buildings has the potential to provide sufficient rooftop space for increased solar production. This offers the possibility to meet the energy demands of buildings with inadequate space and subsequently reduce their operational costs. Second, the community can lower the total investment costs by sharing surplus electricity from production or storage technologies.

In the EC-B scenario, every building maximised its solar installation without utilising its entire rooftop space. Given that all buildings have similar solar production profiles due to geographical closeness, expanding solar production is not advantageous. Essentially, installing more panels on unused spaces in other buildings would be analogous to simply expanding production on each individual building's rooftop.

Thus, the primary cause of benefiting from adopting the EC+ strategy arises from sharing surplus electricity to reduce investment needs. Yet, this becomes a marginal gain due to a strong positive correlation between energy surplus and deficits among the buildings. Fig. 4 illustrates how buildings generally exhibit similar patterns of surplus and deficit electricity generation. This implies a positive correlation between buildings, which in turn limits the opportunities for peer-to-peer trading among them.

7.2. Operational decisions

Table 8 details the various electricity sources utilised to satisfy the neighbourhood's annual electricity demand. These include electricity produced by solar panels, power discharged from batteries, and grid imports. The table also presents the amount of curtailed power from solar panels.

In the BAU scenario, as anticipated, the demand was met by importing directly from the grid due to the absence of alternative energy

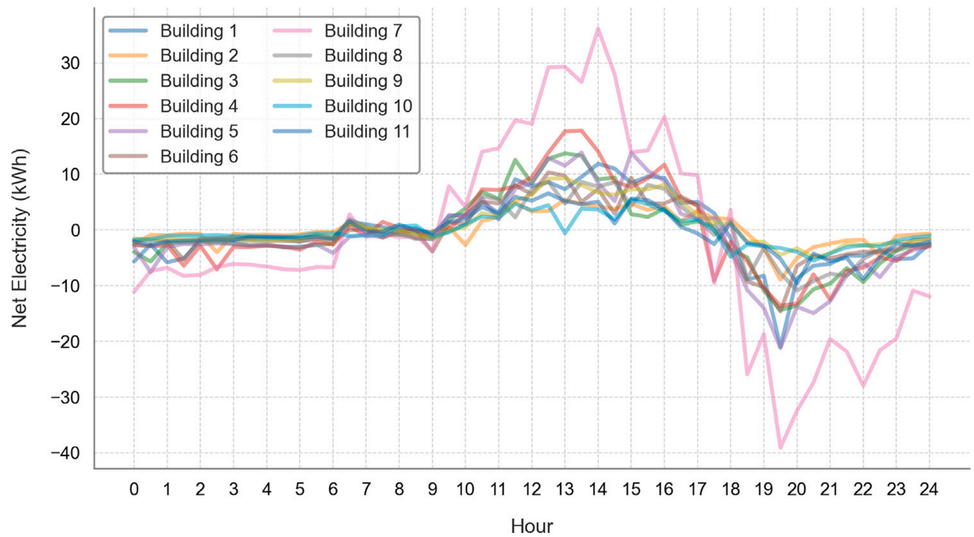


Fig. 4. Net electricity of each building in the EC-B case for a representative day in May. Positive value indicates surplus and negative deficits.

sources. The ID scenario showcased the highest electricity generation from solar panels, attributed to its larger capacity installed. However, a significant 65% of this energy was curtailed, even though the dwellings invested in storage solutions. The curtailment rates were reduced in the EC-B and EC+ scenarios, dropping to 51% and 50%, respectively. These results underline that the flexibility delivered by peer-to-peer trading is considerably more efficient for the neighbourhood than individually installing storage devices.

Also, peer-to-peer bolstered the community's autarky, which is evident from the lower volumes of grid imports and higher amounts of local electricity consumed in the EC-B and EC+ scenarios. It is worth noting that the operational results between these two scenarios were nearly identical, explained by their similar investment decisions.

7.3. Individual electricity bills

The individual electricity bills of households in the EC-B and EC+ strategies were determined using the Nested Shapley Value. However, such computation was unnecessary for the BAU and ID scenarios since they lacked collaboration between participants.

Fig. 5a depicts the distribution of savings of the household per investment strategy compared to the BAU case. The EC+ strategy delivered the largest median (286 €/y), minimum (100 €/y) and maximum (567 €/y) absolute savings. However, these were almost equivalent to those obtained in the EC-B strategy, with a median electricity bill of only 3 €/y lower. Hence, expanding peer-to-peer transactions between buildings did not bring a substantial average reduction of costs to each dwelling. This is expected given their similar strategic and operational results.

Expanding on this analysis, Fig. 5b presents the division of savings between high- and low-demand electricity consumers, expressed as a percentage of their respective electricity bills. Interestingly, the EC+ strategy results in more evenly distributed bill reductions between these two groups than the EC-B strategy. This suggests that the neighbourhood-wide community fosters a more equitable distribution of savings than the per-building approach.

Within each consumer group, however, there is significant variability in savings. Interestingly, the similar range of savings across these consumer groups indicates that total energy consumption is not the primary factor affecting the relative benefits a household can gain from joining an EC. Instead, the key variable appears to be the alignment between solar production at the community level and the consumption patterns. This is demonstrated in Fig. 6, which displays the hour-by-hour correlation between the solar generation and load profiles for four specific dwellings. These selected dwellings represent the extreme cases of bill reduction (maximum and minimum) within each consumption group in the EC+ scenario. A negative correlation signifies that periods of high solar production coincide with low energy consumption, while a positive indicates that both are high or low simultaneously. Notably, dwellings with greater bill reductions (dwellings 119 and 85) generally exhibit more positive correlations throughout the day than those with smaller reductions (dwellings 208 and 66). Therefore, it is primarily the households' load patterns that determine the relative savings, while their total consumption only influences its absolute value.

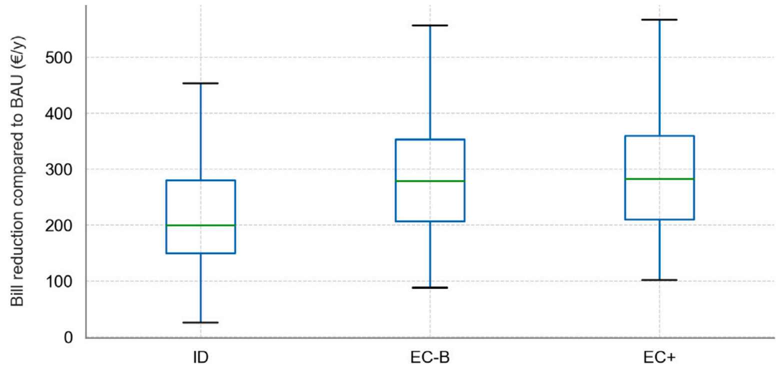
Moreover, a comparison between the annual cost attributed to each household in the EC+ strategy with those obtained applying the BAU (in blue), ID (in yellow) and EC-B (in red) strategies is depicted in Fig. 7. Points situated below the 45-degree dashed line signify households that got a lower electricity bill in the EC+ than in the alternative strategies. Conversely, points above this line indicate higher costs in the EC+ scenario.

When applying the EC+ strategy, all households had lower electricity bills than in the BAU and ID strategies. In contrast, this unanimous result is not observed when comparing the EC+ and EC-B strategies. Despite most households lowering their cost by expanding the community to all buildings, 111 dwellings got better electricity bills when the ECs were adopted per building.

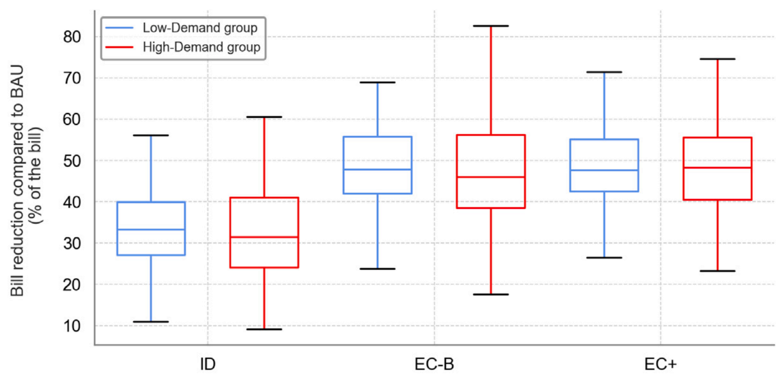
7.4. Decision-making: economic and governance aspects

The previous section compared the costs incurred by individual residents under each strategy, providing insights into their willingness to support one strategy over the other. However, the varied economic preferences among households and governance structures within the neighbourhood would both determine the decision on which strategy to adopt. A governance structure encompasses the processes and practices adopted for decision-making. As such, it ensures a decision is legitimate.

To illustrate the connection between governance structures and individual preferences in decision-making processes, Fig. 8 illustrates a simple example. In Building 1 (B1), residents unanimously prefer forming their own EC as it yields higher benefits to all members. If there is a decision-making framework where each building can decide its preferred strategy, building 1 will opt for the EC-B. On the contrary, if the governance structure is designed such that the decision needs to be made at the community level, the majority of residents within the neighbourhood (B1, B2, and B3) would vote for the EC+ strategy.



(a) Aggregated savings



(b) Savings for high- and low-intensive consumers

Fig. 5. Distribution of savings achieved in the ID, EC-B and EC+ strategies compared to the BAU strategy.

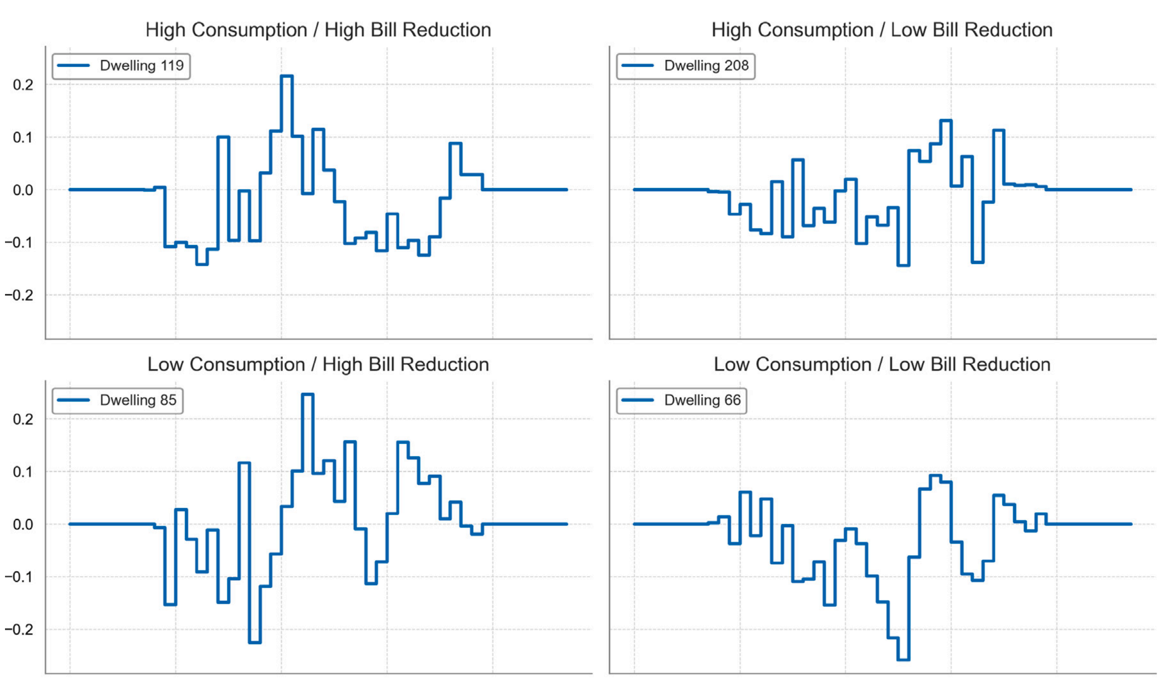


Fig. 6. Hourly average correlation between solar production and load demand of four selected households based on yearly observations.

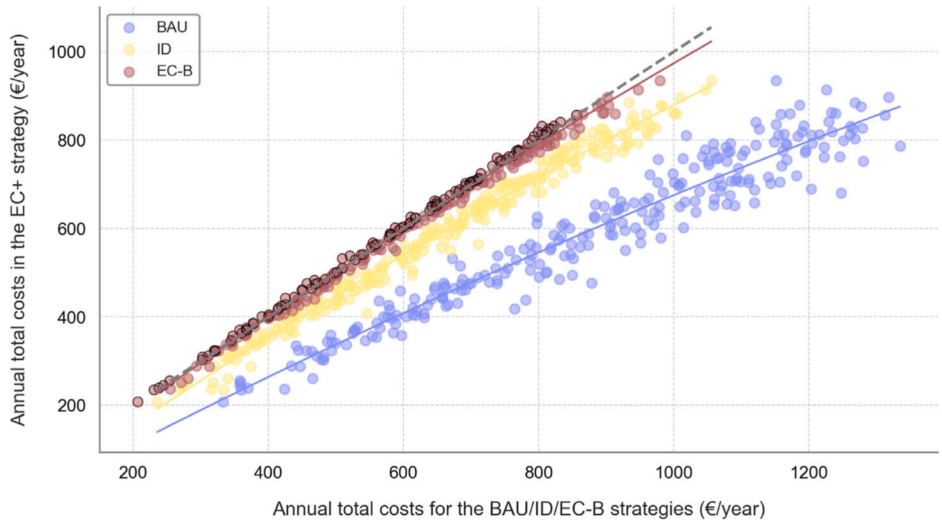


Fig. 7. Scatter plot of the annual electricity bills for the 251 households. The x-axis captures the bill for the BAU, ID and EC-B scenarios, while the y-axis shows those for the EC+ scenario.

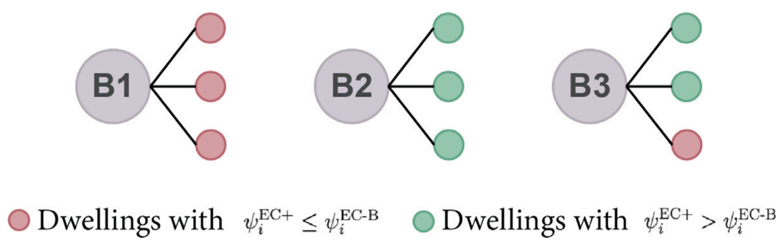


Fig. 8. Example of a neighbourhood with three buildings, each with three residents. Here, ψ_i^{EC+} and ψ_i^{EC-B} indicate the Nested Shapley value associated to resident i in the EC+ and EC-B strategies, respectively.

Then, Building 1 would need to accept, independently of their own preferences.

As presented in the example, an optional governance structure that can rule the decision-making process within multi-dwelling neighbourhoods is *neighbourhood-wide democratic voting*. In this approach, all households in the community have one voting right to decide whether to adopt the EC-B or EC+ strategy. Hence, each dwelling will vote for the strategy that delivers the highest cost reduction following Eq. (24).

$$\textbf{Strategy}(\mathcal{N}) = \begin{cases} \text{EC+} & \text{if } \sum_{i\in\mathcal{N}} \mathbb{1}\{\psi_i^{\text{EC+}} \leq \psi_i^{\text{EC-B}}\} \\ & \quad \geq \sum_{i\in\mathcal{N}} \mathbb{1}\{\psi_i^{\text{EC+}} > \psi_i^{\text{EC-B}}\} \\ \text{EC-B} & \text{otherwise} \end{cases} \tag{24}$$

with ψ_i^{EC+} and ψ_i^{EC-B} indicating the electricity bill of agent i under the EC+ and EC-B strategy.

Another alternative is a *per-building democratic voting* where each building decides its strategy as follows:

$$\textbf{Strategy}(\mathbf{B}) = \begin{cases} \text{EC+} & \text{if } \sum_{i\in B} \mathbb{1}\{\psi_i^{\text{EC+}} \leq \psi_i^{\text{EC-B}}\} \\ & \quad \geq \sum_{i\in B} \mathbb{1}\{\psi_i^{\text{EC+}} > \psi_i^{\text{EC-B}}\} \\ \text{EC-B} & \text{otherwise} \end{cases}$$

In the case study, the neighbourhood-wide democratic voting would ensure the adoption of the EC+ strategy, given its broad support among most dwellings (see Fig. 5). Nonetheless, when adopting the per-building democratic voting, buildings 6, 9 and 10 would choose the EC-B strategy, yielding a different investment strategy selection compared to neighbourhood-wide voting.

Therefore, both the economic aspect of individual preferences and the governance structure are pivotal factors in the decision-making process regarding the investment strategy to adopt.

While this analysis may resonate with the concept of stability in cooperative game theory – traditionally applied to assess the formation of subcoalitions within a single game – the focus here is different. Instead, this section compares the decision-making processes under two distinct investment strategies, the EC-B and EC+, by analysing how different governance structures and economic outcomes influence individual preferences. This analysis does not delve into the stability of the coalition due to the payoff but rather investigates the preferences across two strategic scenarios, highlighting the critical role of governance and economic factors.

7.5. Computational details of the case study

The Nested Shapley Value method was applied to compute the electricity bills of the 251 agents in the EC+ case. The ordered tree was constructed following the k-means algorithm considering the attributes of total annual electricity consumption and mean average peak load. The resulting ordered tree had 67 nodes and 4 layers (see Fig. 9). The parent nodes preceding the leaf nodes were associated with a maximum of $p = 10$ households in order to ease the computation of the Nodal Shapley Value.

The method was also applied in some buildings in the EC-B case, as some contain more than 10 apartments. The BAU and ID cases did not require computing the Nested Shapley value since they do not consider coalitions within the neighbourhood.

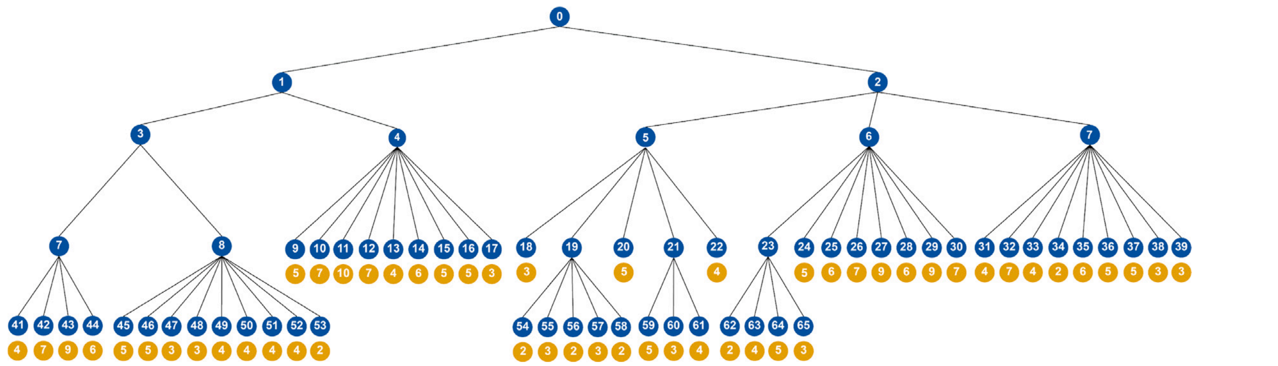


Fig. 9. Resulting ordered tree for the EC+ case. Nodes are indicated in blue, and instead of displaying individual leaf nodes, the number of leaf nodes connected to each parent is shown in yellow.

The Nested Shapley value algorithm, the k-means method and the joint investment model were developed in Python 3.10. The Pyomo package [37,38] and Gurobi Solver v9.5.2 were selected to implement and solve the optimisation model. All computations were performed on a computer cluster with CPU 2x 2.3 GHz Intel E5-2670v3 (12 core) and 64 Gb RAM.

To perform the EC+ and the EC-B cases, 5453 and 5074 instances of the model were solved, respectively. These are significantly lower numbers than the 2^{251} and $1.46 \cdot 10^{20}$ instances required for computing the true Shapley value in both cases. The reduction in the number of instances was accompanied by parallel computing to speed up the computation of the characteristic value of each sub-coalition.

8. Conclusions

This paper proposes the Nested Shapley value as an approximation method of the Shapley value, designed to deal with coalitions with a large number of players. The method was applied to analyse the decision-making process of multi-dwelling ECs with four potential joint investment strategies. To do so, an investment optimisation model tailored to multi-dwelling buildings was formulated. The paper also explores the integration of the Nested Shapley value with governance structures to better understand joint investment decision-making in ECs.

By clustering agents, the Nested Shapley value method significantly reduces the computational complexity, making it feasible for value distribution in large-scale coalitions. Mathematically, the Nested Shapley Value satisfies three fairness axioms: efficiency, null player axioms, and symmetry (when properly formulated clusters). The only axiom it does not meet is linearity.

The results obtained for the symmetry axiom and the linearity axiom give clear instructions on utilising the proposed tool. Namely, the symmetry property prescribes that the clusters need to be defined to be internally homogeneous, whereas the lack of the linearity property prescribes that, when considering multiple projects, they must be evaluated on a one-by-one basis rather than considering them all together. The method is applicable to all large-scale cooperative games and is not confined to value allocation in ECs. However, the weaker version of the symmetry property presents the challenge of identifying the specific attributes of agents that influence their contributions to the community. This task may not be straightforward across different types of games.

Moreover, the study applies the Nested Shapley value to allocate costs in ECs with joint investments, given that they often comprise a large number of residents, and fair value distribution is crucial for their long-term stability. By computing the Nested Shapley value, the paper explores how individual payoffs might affect the acceptance and legitimacy of investment strategies within ECs. It was demonstrated that quantifying individual payoffs, along with considering the governance structure of the community, can be used to assess the likelihood of adopting different investment strategies. For the case study, it was shown that adopting neighbourhood-wide democratic voting was essential for ensuring the implementation of a joint investment strategy that leads to the most resource-efficient community with the lowest variability of benefits for residents.

However, there are challenges to adopting the Nested Shapley value method in ECs: first, its mathematical complexity can hinder understandability, potentially affecting residents' trust in the method; second, the distribution of costs post-operations may limit end users' ability to adapt to more favourable economic outcomes on time; and finally, regulatory constraints may inhibit the formation of collective metering and billing systems required for ECs. Furthermore, although outside the scope of this paper, the practical application of large energy communities is likely to imply additional coordination burdens such as administrative and transaction costs which may require careful assessment.

Future studies should qualitatively examine the real-world applicability of allocation methods like the Nested Shapley value to dive into the current status and possible solutions to the barriers mentioned. Additionally, exploring the use of the Nested Shapley Value in contexts beyond ECs could provide valuable insights, particularly concerning the construction of ordered trees that maintain the symmetry property.

CRediT authorship contribution statement

Raquel Alonso Pedrero: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. **Paolo Pisciella:** Conceptualization, Formal analysis, Methodology, Supervision, Writing – original draft, Writing – review & editing. **Pedro Crespo del Granado:** Funding acquisition, Project administration, Resources, Supervision.

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

We provided a link for a repository in the paper.

Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work, the author(s) used ChatGPT in order to check for any grammar mistakes to guarantee the proper readability of the paper. Any content was generated by the AI technology, but it was employed as a support tool for grammar checking. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the publication.

Acknowledgements

The authors are grateful for the funding provided by the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No. 869918.

Appendix A. Nomenclature

A.1. Theory and Nested Shapley value

See Table A.9.

A.2. Joint-investment model

See Table A.10.

Appendix B. Closed-formulation of number of computations for a symmetric tree

From the definition of the Nested Shapley value, the number of computations is

$$\text{computations} = \sum_{n \in F \setminus F^s} (2^{|F_n^C|} - 1) \tag{B.1}$$

Assuming a symmetric tree where all the nodes have an identical number of children nodes c such that $c = |F_n^C|$ for all $n \in F \setminus F^s$, the computational load simplifies to:

$$\text{computations} = A(2^c - 1) \tag{B.2}$$

The constant A is defined as the geometric series $\sum_{l=1}^{q-1} n^{l-1}$ which can be expressed in a closed-form. The derivation is as follows:

$$A = n^0 + n^1 + n^2 + \cdots + n^{q-2} \tag{B.3}$$

$$A \cdot n = n + n^2 + n^3 + \cdots + n^{q-1} \tag{B.4}$$

$$A \cdot n - A = n^{q-1} - 1 \tag{B.5}$$

$$A(n-1) = n^{q-1} - 1 \tag{B.6}$$

$$A = \frac{n^{q-1} - 1}{n-1} \tag{B.7}$$

Upon substituting Eq. (B.7) into Eq. (B.2), the computations requirement for computing the Nested Shapley values of agents allocated in a symmetric tree is obtained:

$$\text{computations} = \frac{n^{q-1} - 1}{n-1}(2^c - 1) \tag{B.8}$$

Appendix C. Accuracy comparison

To examine the computational accuracy of the Nested Shapley value, the following section presents two numerical examples comparing the proposed method with the true Shapley value and the stratified expected value from Cremers et al. [26].

In the first example, every agent contributes the same in generation and storage technologies, while the second assumes that agents are not identical in the resources they contribute. The optimisation model for computing the characteristic functions is the one presented in this paper without the investment decisions, which are assumed to be fixed. Hence, only the operational part is considered. Both examples are applied to a community of 12 agents with a time horizon of one day.

The results of the Nested Shapley value are presented for different ordered tree structures with $q = 3$. Table C.11 shows the characteristics of the ordered trees used in these two examples. The results for both examples are presented in Table C.12 and Table C.13.

In scenarios with centralised resources, the stratified expected value method provides a closer approximation to the true Shapley value than the Nested Shapley approach does. To quantify this, the maximum relative difference from the true Shapley value is 16% for the stratified expected value, compared to a significantly higher 98% for the Nested Shapley. Conversely, in communities with decentralised resources, the Nested Shapley value outperforms when utilising a tree with 8 nodes in the second layer. Not only does it yield more accurate results, but it also reduces the computational effort required.

The stratified expected value, as described in the work by Cremers et al. [26], is particularly effective for communities where each agent's contribution to generation and storage capacity is identical. This method selects a "representative" subcoalition for each Shapley value stratum based on average demands. In such cases, the representative subcoalition accurately reflects the collective benefits of the agents within each stratum.

However, the situation becomes more complex when agents contribute differently in generation and storage capacities. In the second example, the representative coalition was generated, allocating the generation and storage capacities based on the community's average, which parallels the demand handling of the stratified expected value method. It should be noted that this approach of handling the capacity of assets is not discussed in [26], as the authors do not provide guidance on how to apply the method for scenarios where agent contributions vary.

Appendix D. Properties of the Nested Shapley value

The Shapley Value is the only payment rule to satisfy four important properties: Efficiency, Symmetry, Linearity and Null player. The Nested Shapley Value introduced in this paper fully satisfies Efficiency and the Null player properties, while it does not satisfy the Linearity property. The Symmetry property only holds if clustering is performed by gathering equivalent actors in the same clusters in every child node of the nesting tree beside the leaf node. In the following appendix, the considered properties are formally presented.

Table A.9
Mathematical notation for the indices, sets and parameters in the cost allocation methods.

Indices and sets	
$i, j \in \mathcal{N}$	Agents
a, b	Partitions
n, m, s	Nodes
l	Layers in the tree from 1 to q
P	Set of partitions of $\mathcal{N}$
F	Set of nodes
C_n	Coalition of agents associated to node n
F_n^C	Set of children nodes of node n
F_n^P	Set (singleton) of parent nodes to node n
F_i	Set of nodes associated with agent i
F^s	Set of nodes with a single agent
F^l	Set of nodes in layer l
B	Set of agents in building B
Parameters	
p	Number of partitions
N	Number of players
M	Number of nodes
q	Number of layers in the nodal tree
ϕ_i	Shapley value for agent i
ϕ_i^O	Owen value for agent i
ϕ_n^*	Nodal Shapley Value computed using the value function $v(C_m)$ with $n \in F_m^C$
ψ_n^*	Corrected Shapley value
ψ_i	Nested Shapley value of agent i
ρ_i	Silhouette coefficient for agent i
σ_p	Average silhouette coefficient for partition p
e_B	Preference between strategies at building B

Table A.10
Mathematical notation for the indices, sets and parameters in the cost allocation methods.

Indices and sets	
$h \in \mathcal{H}$	Dwellings in the coalition
$g \in G$	Generation technologies
$s \in S$	Storage technologies
$i \in I$	Season
$t \in T$	Timestep (e.g., hour, half-hour)
Parameters	
H	Number of dwellings in the coalition
a_g and a_s	Annuities for generation and storage technologies
α_i	Seasonal scaling factor
λ	Constant of proportionality in ESS (kW/kWh)
β	Conversion factor from power to energy (kW/kWh). It would be calculated based on the time period over which the power is applied
ψ	Grid losses factor (kWh/kWh)
L_{ith}	Electricity demand (kWh)
G_{git}	Generation profile (kWh/kW per timestep)
c_g^P and c_s^P	Investment cost, power related (€/kW)
c_g^I	Instalment cost (€/kW)
c_s^C	Investment cost, capacity related (€/kWh)
c_s^{MC}	Maintenance costs, capacity related (€/kWh)
c_s^{MP}	Maintenance costs, power related (€/kW)
c^G	Volumetric grid tariff (€/kWh)
c_{it}^{Spot}	Spot price (€/kWh)
c^{Tax}	Tax cost (€/kWh)
$\overline{G}_g$	Maximum capacity for generation technologies based on shared areas (kW)
$\overline{S}_s$	Maximum capacity of ESS (kWh/dwelling)
SOC_{si}	Initial state of charge in season i (kWh)
η^C and η^D	Charging and discharging efficiency factors
Variables	
x_g^{PG} and x_s^{PG}	Generation capacity and storage technologies power rates, respectively (kW)
x_s^C	Capacity of storage technologies (kWh)
y_{ith}^G	Electricity withdrawal of each dwelling from the main grid (kWh)
y_{it}^{GA}	Electricity withdrawal of common areas from the main grid (kWh)
y_{git}^G	Electricity delivered from shared areas to all the coalition (kWh)
y_{ith}^A	Electricity withdrawal of each dwelling from common areas (kWh)
y_{it}^{EA}	Electricity withdrawal of the entire community from common areas (kWh)
y_{sit}^C and y_{sit}^D	Electricity charge and discharge of ESS (kWh)
y_{sit}^{SOC}	State of charge of the batteries (kWh)
b_{sit}	Binary variable for controlling the charge and discharge of ESS

Table C.11
Features of the ordered trees considered for the two examples.

Tree	# nodes 2nd layer	# nodes 3rd layer	Computations
2-X	2	12	161
3-X	3	12	72
4-X	4	12	59
6-X	6	10	83
8-X	8	8	268

Proposition 1. *The Nested Shapley Value satisfies the efficiency property:*

$$\sum_{i \in \mathcal{N}} \psi_i = v(\mathcal{N})$$

Proof. By definition,

$$\sum_i \psi_i = \sum_{i=1}^{N} \sum_{n \in F^q \cap F_i} \psi_n^* = \sum_{i=1}^{N} \sum_{n \in F^q \cap F_i} \sum_{m \in F_n^P} \frac{\phi_n^*}{v(C_m)} \psi_m^*$$

also it should be considered that ϕ_n^* with $n \in F_m^C$ is the plain Shapley Value for the children nodes n to node m. The last expression can be

Table C.12
Results for community with centralised resources.

Agent	2-X	3-X	4-X	6-X	8-X	Stratified	Shapley
1	−0.3	−2.0	−2.0	0.1	−0.1	−3.9	−4.0
2	−0.4	−6.7	−8.0	−9.2	−9.1	−6.7	−6.8
3	2.5	10.3	10.0	12.4	13.4	13.4	13.3
4	1.1	7.2	7.0	12.1	11.2	11.9	11.1
5	−0.1	−2.6	−2.5	−0.2	−5.3	−5.3	−5.3
6	−5.6	−5.5	−4.3	−4.5	−4.5	−3.6	−3.7
7	−0.9	−6.7	−8.0	−9.2	−9.1	−11.8	−12.0
8	2.9	6.5	6.3	−3.2	1.5	4.4	5.2
9	7.9	7.7	7.9	8.2	8.2	6.8	6.5
10	30.0	29.6	28.7	31.3	31.5	31.3	31.5
11	18.2	18.0	17.7	18.2	18.3	20.6	20.5
12	40.2	39.6	42.6	39.2	39.2	38.3	39.0

rewritten as

$$\sum_{i=1}^{N} \sum_{n \in F^q} \sum_{m \in F^P} \frac{\phi_n^*}{v(C_m)} \psi_m^* \mathbb{1}\{F^q \cap F_i\} \mathbb{1}\{F_n^P\}$$

where the product of the indicator functions assumes value 1 if a node in the last layer (i.e. a node not having children nodes) and its parent

Table C.13
Results for community with decentralised resources.

Agent	2-X	3-X	4-X	6-X	8-X	Stratified	Shapley
1	3.9	−5.5	−5.5	−3.0	4.2	12.4	7.9
2	−1.5	7.9	7.9	7.2	6.4	8.4	3.4
3	29.1	23.2	23.2	24.0	31.3	32.4	30.6
4	55.4	58.6	58.6	55.5	51.1	45.1	51.1
5	26.6	28.6	28.6	26.9	24.0	20.8	23.7
6	13.1	13.0	38.9	15.8	15.1	13.6	11.8
7	13.1	10.6	10.6	9.7	8.6	10.7	12.0
8	10.8	15.6	15.6	19.0	16.9	13.7	15.6
9	7.3	7.2	43.3	3.0	2.8	5.8	5.1
10	32.0	31.6	13.0	32.8	32.6	31.2	31.3
11	39.4	38.9	31.6	37.9	36.3	41.1	37.5
12	43.8	43.3	7.2	44.0	43.6	37.8	42.7

node are selected. The expression can be easily rewritten as

$$\sum_{m\in F^P}\sum_{n\in F^q}\sum_{i=1}^{N}\frac{\phi_n^*}{v(C_m)}\psi_m^*\mathbb{1}\{F^q\cap F_i\}\mathbb{1}\{F_n^P\}$$

Let us break down the analysis between the nodes belonging to the last layer and nodes belonging to previous layers. Assuming that there are q layers for the nesting of the formula, every node in the last layer is associated with exactly one player, while in previous layers, a node could contain several players. By definition of the indicator function, the expression for the last layer can be rewritten as

$$\sum_{m\in F^{q-1}}\sum_{n\in F_m^C}\frac{\phi_n^*}{v(C_m)}\psi_m^*$$

with $\sum_{n\in F_m^C}\phi_n^* = v(C_m)$ by definition of ϕ_n^* as the Shapley value of the nodes children to node m, whose value function is $v(C_m)$. This leads to the new expression

$$\sum_{i=1}^{N}\psi_i = \sum_{m\in F^{q-1}}\psi_m^* \tag{D.1}$$

where m belongs to layer $q-1$. This expression can be restated as

$$\sum_{i=1}^{N}\psi_i = \sum_{m\in F^{q-1}}\psi_m^* = \sum_{m\in F^{q-1}}\sum_{\overline{m}\in F_m^P}\frac{\phi_m^*}{v(C_{\overline{m}})}\psi_{\overline{m}}^* = \sum_{\overline{m}\in F^{q-2}}\sum_{m\in F_{\overline{m}}^C}\frac{\phi_m^*}{v(C_{\overline{m}})}\psi_{\overline{m}}^*$$

which, by the same argument used above, leads to

$$\sum_{i=1}^{N}\psi_i = \sum_{\overline{m}\in F^{q-2}}\psi_{\overline{m}}^* \tag{D.2}$$

which is a recursive update to (D.1). By recursion, the expression leads back to the first node, and the expression becomes

$$\sum_{i=1}^{N}\psi_i = \sum_{m\in F^1}\psi_m^* = \psi_1^* = v(\mathcal{N}) \tag{D.3}$$

Proposition 2. *The Nested Shapley Value satisfies the symmetry property within nests. Namely, if $v(S\cup\{i\}) = v(S\cup\{j\})$ for every $S\subset\mathcal{N}\setminus\{i,j\}$ and i and j are assigned to the same cluster in every layer of the nesting tree, then*

$$\psi_i = \psi_j$$

Proof. Let us assume that the last parent node in the nesting tree contains coalition S and $i,j\in S$. Then for every subcoalition of S (let us say S_1) it is obtained that $v(S_1\cup\{i\}) = v(S_1\cup\{j\})$. This implies that the Nodal Shapley Value for l_i and l_j satisfies the property $\phi_{l_i}^* = \phi_{l_j}^*$, with l_i and l_j corresponding to the leaf nodes mapped to agents i and j. The reason is that $\phi_{l_i}^*$ and $\phi_{l_j}^*$ are simply the Shapley Values assigned to the children nodes to node n such that $l_i, l_j \in F_n^C$. The Nested Shapley Value is obtained for agent i and j is obtained respectively by the formulas

$$\psi_i = \psi_{l_i}^* = \frac{\phi_{l_i}^*}{v(C_n)}\psi_n^* \tag{D.4}$$

and

$$\psi_j = \psi_{l_j}^* = \frac{\phi_{l_j}^*}{v(C_n)}\psi_n^* \tag{D.5}$$

with $\phi_{l_i}^* = \phi_{l_j}^*$. This completes the proof.

Table D.14
Characteristic function values for the example following Algorithm 1.

Parent node	Nodes	Agents	v	w	z
–	{1}	{1, 2, 3, 4}	2000	1400	3400
1	{2}	{2, 3}	1000	700	1700
1	{3}	{1}	200	100	300
1	{4}	{4}	800	500	1300
1	{2, 3}	{1, 2, 3}	1200	800	2000
1	{2, 4}	{2, 3, 4}	1800	1200	3000
1	{3, 4}	{1, 4}	1000	600	1600
2	{5}	{2}	400	300	700
2	{6}	{3}	600	400	1000

Table D.15
Nested Shapley Values for the three games considered in the example and the sum of games v and w.

	v	w	z	$v+w$
ψ_1	200	133.33	333.33	333.33
ψ_2	400	314.29	713.73	714.29
ψ_3	600	419.05	1019.61	1019.05
ψ_4	800	533.33	1333.33	1333.33

Proposition 3. *The Nested Shapley Value satisfies the null player property. Namely, if a player $i\in\mathcal{N}$ is such that $v(S\cup\{i\}) = v(S)$ for each $S\subseteq\mathcal{N}\setminus\{i\}$ then $\psi_i = 0$*

Proof. In the last layer of the nesting tree, the uncorrected Shapley Value assigned to the leaf node l_i, mapped to player i, is $\phi_{l_i}^* = 0$ because i is a Null player and $\phi_{l_i}^*$ is computed as an ordinary Shapley Value. Since the Nested Shapley Value for player i is given by

$$\psi_i = \psi_{l_i}^* = \frac{\phi_{l_i}^*}{v(C_n)}\psi_n^* \tag{D.6}$$

where n is the parent node of l_i. Therefore, the property is satisfied.

Remark 1. The Nested Shapley Value does not satisfy the Linearity property, which states

$$\psi_i(v+w) = \psi_i(v) + \psi_i(w) \tag{D.7}$$

This non-compliance can be demonstrated through an example. Assume the coalition of agents $i = 1,2,3,4$ organised in an ordered tree as illustrated in Fig. 1 (Section 4).

Table D.14 shows the parent node, the coalition of nodes and agents and their respective values for games v, w and $z = v + w$.

The results of calculating the Nested Shapley values for each agent under the three different games are presented in Table D.15. It can be observed that the linearity property is not met given that the sum of the Nested Shapley values of games v and w are not equal to the values in game z.

Appendix E. Algorithm k-means for the ordered tree

The algorithm should be initialised with $l = 1$ and $F^1 = \mathcal{N}$ (see Algorithm 2).

Algorithm 2 RecurssiveNodes

```
Inputs:
    l, p̄, F^l
Assume F^l is an ordered set of nodes with last node m
for n ∈ F^l do
    Initialize:
    F_n^C = ∅, F^{l+1} = ∅
    if | C_n |> p̄ then
        compute p* using the silhouette method
        compute P using the k-means with p = p*
        for p = 1, ..., p* do
            m = m + 1
            C_m = P_p
            F_n^C = F_n^C ∪ {m}
            F^{l+1} = F^{l+1} ∪ {m}
            F_m^P = {n}
        end for
    else
        for i ∈ C_n do
            m = m + 1
            C_m = {i}
            F_n^C = F_n^C ∪ {m}
            F^{l+1} = F^{l+1} ∪ {m}
            F_m^P = {n}
        end for
    end if
end for
if ∃ n ∈ F^{l+1} such that |C_n| > 1 then
    RecurssiveNodes(l + 1, p̄, F^{l+1})
end if
```

References

[1] Caramizaru E, Uihlein A. Energy communities: an overview of energy and social innovation. Report EUR 30083 EN, Publications Office of the European Union; 2020, http://dx.doi.org/10.2760/180576.

[2] Chan G, Evans I, Grimley M, Ihde B, Mazumder P. Design choices and equity implications of community shared solar. Electr J 2017;30(9):37–41. http://dx.doi.org/10.1016/j.tej.2017.10.006.

[3] Reis IFG, Gonçalves I, Lopes MAR, Henggeler Antunes C. Business models for energy communities: A review of key issues and trends. Renew Sustain Energy Rev 2021;144. http://dx.doi.org/10.1016/j.rser.2021.111013, Export Date: 03 May 2023; Cited By: 64.

[4] Cohen JJ, Azarova V, Kollmann A, Reichl J. Preferences for community renewable energy investments in Europe. Energy Econ 2021;100:105386. http://dx.doi.org/10.1016/j.eneco.2021.105386.

[5] Koirala B, Hakvoort R. Chapter 18 - integrated community-based energy systems: Aligning technology, incentives, and regulations. In: Innovation and disruption at the grid's edge. Academic Press; 2017, p. 363–87. http://dx.doi.org/10.1016/B978-0-12-811758-3.00018-8.

[6] Li N, Hakvoort RA, Lukszo Z. Cost allocation in integrated community energy systems - A review. Renew Sustain Energy Rev 2021;144:111001. http://dx.doi.org/10.1016/j.rser.2021.111001.

[7] Shapley LS. Notes on the N-person Game. Rand Corporation; 1951.

[8] Sousa J, Lagarto J, Camus C, Viveiros C, Barata F, Silva P, Alegria R, Paraíba O. Renewable energy communities optimal design supported by an optimization model for investment in PV/wind capacity and renewable electricity sharing. Energy 2023;283:128464. http://dx.doi.org/10.1016/j.energy.2023.128464.

[9] Fleischhacker A, Lettner G, Schwabeneder D, Auer H. Portfolio optimization of energy communities to meet reductions in costs and emissions. Energy 2019;173:1092–105. http://dx.doi.org/10.1016/j.energy.2019.02.104.

[10] Jin X, Wu Q, Jia H. Local flexibility markets: Literature review on concepts, models and clearing methods. Appl Energy 2020;261:114387. http://dx.doi.org/10.1016/j.apenergy.2019.114387.

[11] Khorasany M, Mishra Y, Ledwich G. Market framework for local energy trading: a review of potential designs and market clearing approaches. IET Gener Transm Distrib 2018;12(22):5899–908. http://dx.doi.org/10.1049/iet-gtd.2018.5309.

[12] Massol O, Tchung-Ming S. Cooperation among liquefied natural gas suppliers: Is rationalization the sole objective? Energy Econ 2010;32(4):933–47. http://dx.doi.org/10.1016/j.eneco.2010.02.008.

[13] Faria VAD, de Queiroz AR, Lima LMM, Lima JWM. Cooperative game theory and last addition method in the allocation of firm energy rights. Appl Energy 2018;226:905–15. http://dx.doi.org/10.1016/j.apenergy.2018.06.065.

[14] Junqueira M, Costa LCd, Barroso LA, Oliveira GC, Thome LM, Pereira MV. An Aumann-Shapley approach to allocate transmission service cost among network users in electricity markets. IEEE Trans Power Syst 2007;22(4):1532–46. http://dx.doi.org/10.1109/TPWRS.2007.907133.

[15] Saad W, Han Z, Poor HV, Basar T. Game-theoretic methods for the smart grid: An overview of microgrid systems, demand-side management, and smart grid communications. IEEE Signal Process Mag 2012;29(5):86–105. http://dx.doi.org/10.1109/MSP.2012.2186410.

[16] Saad W, Han Z, Poor HV. Coalitional game theory for cooperative micro-grid distribution networks. In: 2011 IEEE international conference on communications workshops. ICC, 2011, p. 1–5. http://dx.doi.org/10.1109/iccw.2011.5963577.

[17] Tveita E, Löschenbrand M, Bjarghov S, Farahmand H. Comparison of cost allocation strategies among prosumers and consumers in a cooperative game. In: 2018 international conference on smart energy systems and technologies. SEST, 2018, p. 1–6. http://dx.doi.org/10.1109/SEST.2018.8495690.

[18] Han L, Morstyn T, McCulloch M. Constructing prosumer coalitions for energy cost savings using cooperative game theory. In: 2018 power systems computation conference. PSCC, 2018, p. 1–7. http://dx.doi.org/10.23919/PSCC.2018.8443054.

[19] Abada I, Ehrenmann A, Lambin X. On the viability of energy communities. Energy J 2020;41(1).

[20] Fleischhacker A, Corinaldesi C, Lettner G, Auer H, Botterud A. Stabilizing energy communities through energy pricing or PV expansion. IEEE Trans Smart Grid 2022;13(1):728–37. http://dx.doi.org/10.1109/tsg.2021.3121283.

[21] Cheng Y, Fan S, Ni J, Sun S, Wang R, Li G, Xin Z, Ai Q. An innovative profit allocation to distributed energy resources integrated into virtual power plant. In: International conference on renewable power generation. RPG 2015, 2015, p. 1–6. http://dx.doi.org/10.1049/cp.2015.0547.

[22] Gjorgievski VZ, Cundeva S, Markovska N, Georghiou GE. Virtual net-billing: A fair energy sharing method for collective self-consumption. Energy 2022;254. http://dx.doi.org/10.1016/j.energy.2022.124246.

[23] Tushar W, Yuen C, Mohsenian-Rad H, Saha T, Poor HV, Wood KL. Transforming energy networks via peer-to-peer energy trading: The potential of game-theoretic approaches. IEEE Signal Process Mag 2018;35(4):90–111. http://dx.doi.org/10.1109/MSP.2018.2818327.

[24] Norbu S, Couraud B, Robu V, Andoni M, Flynn D. Modelling the redistribution of benefits from joint investments in community energy projects. Appl Energy 2021;287:116575. http://dx.doi.org/10.1016/j.apenergy.2021.116575.

[25] Lee W, Xiang L, Schober R, Wong VWS. Direct electricity trading in smart grid: A coalitional game analysis. IEEE J Sel Areas Commun 2014;32(7):1398–411. http://dx.doi.org/10.1109/jsac.2014.2332112.

[26] Cremers S, Robu V, Zhang P, Andoni M, Norbu S, Flynn D. Efficient methods for approximating the Shapley value for asset sharing in energy communities. Appl Energy 2023;331:120328. http://dx.doi.org/10.1016/j.apenergy.2022.120328.

[27] Moulin H. Cooperative microeconomics: A game-theoretic introduction, vol. 313, Princeton University Press; 1995.

[28] Owen G, Winter E. The multilinear extension and the coalition structure value. Games Econom Behav 1992;4(4):582–7.

[29] Giménez JM, Puente MA. A new procedure to calculate the owen value. In: ICORES 2017 - proceedings of the 6th international conference on operations research and enterprise systems. Vol. 2017-January, 2017, p. 228–33. http://dx.doi.org/10.5220/0006113702280233, Export Date: 29 June 2023; Cited By: 3.

[30] MacQueen J. Some methods for classification and analysis of multivariate observations. In: Proceedings of the fifth Berkeley symposium on mathematical statistics and probability. Vol. 1, Oakland, CA, USA; 1967, p. 281–97.

[31] Eurostat. Electricity prices for household consumers - bi-annual data. 2023.

[32] Andresen I, Healey Trulsrud T, Finocchiaro L, Nocente A, Tamm M, Ortiz J, Salom J, Magyari A, Hoes-van Oeffelen L, Borsboom W, Kornaat W, Gaitani N. Design and performance predictions of plus energy neighbourhoods – Case studies of demonstration projects in four different European climates. Energy Build 2022;274:112447. http://dx.doi.org/10.1016/j.enbuild.2022.112447.

[33] Staffell I, Pfenninger S. Using bias-corrected reanalysis to simulate current and future wind power output. Energy 2016;114:1224–39. http://dx.doi.org/10.1016/j.energy.2016.08.068.

[34] Pfenninger S, Staffell I. Long-term patterns of European PV output using 30 years of validated hourly reanalysis and satellite data. Energy 2016;114:1251–65. http://dx.doi.org/10.1016/j.energy.2016.08.060.

[35] ENTSO-E. Transparency platform API. 2023.

[36] Pflugradt N, Stenzel P, Kotzur L, Stolten D. LoadProfileGenerator: An agent-based behavior simulation for generating residential load profiles. J Open Source Softw 2022;7(71):3574. http://dx.doi.org/10.21105/joss.03574.

[37] Hart WE, Laird CD, Watson J-P, Woodruff DL, Hackebeil GA, Nicholson BL, Siirola JD. 2nd ed.. Pyomo - Optimization modeling in Python, vol. 67, Springer Science & Business Media; 2017, http://dx.doi.org/10.1007/978-3-319-58821-6.

[38] Hart WE, Watson J-P, Woodruff DL. Pyomo: modeling and solving mathematical programs in Python. Math Program Comput 2011;3(3):219–60. http://dx.doi.org/10.1007/s12532-011-0026-8.