JOURNAL
OF
APPLIED
QUANTITATIVE
METHODS

Quantitative Methods Inquires

META-HEURISTICS FOR RENEWABLE ENERGY RESIDENTIAL COMMUNITIES DETECTION

Giacomo di Tollo[a], Graziella Pacelli[a], Oliver Toth[b], Salvatore Vergine[a]

[c]*Marche Polytechnic University, Faculty of Economics, Piazza Roma 22, 60121 Ancona, Italy*

[b]*University of Luxembourg, Department of Computer Science, 2 Av. de l'Universite, 4365 Belval Esch-sur-Alzette, Luxembourg*

Abstract

Metaheuristics are high-level heuristic strategies that guide the behaviour of subordinated heuristics to solve optimization problems. This paper proposes a population-based metaheuristic approach for the detection of cost-efficient renewable-based residential energy communities. The problem is formulated as an optimization model minimizing total operational costs through coordinated photovoltaic generation and energy storage systems. Particle Swarm Optimization (PSO) is employed to efficiently explore the large combinatorial space of candidate communities. Results on a residential case study show that PSO achieves near-optimal solutions comparable to Monte Carlo-based benchmarks, while reducing computational time.

Keywords: metaheuristics; energy communities; renewable energy.

1. Introduction

Renewable energy is gaining more and more importance for different classes of stakeholders such civil society, industry, policy makers, and academics. The deployment of many publicly funded initiatives that encourage the Twin transition (i.e., integrating green and digital transition) witness this aspect at the policy makers stage, whilst many firms have introduced ESG-related concepts to (also) drive productivity and profitability. From the policy makers perspective, this led to an increased attention to address the climate change, to enhance energy security, and to support sustainable economic development. Many contributions highlight the fact that large-scale deployment of renewable energy technologies (i.e., solar, wind, and hydroelectric power) is essential to achieve substantial reductions in greenhouse gas emissions and to meet international climate targets [19, 9]. Moreover, renewables sources are increasingly recognized for their potential to foster technological innovation, to reduce dependence on fossil fuel imports, and to promote resilient energy systems [16]. Also renewable energy related aspects are gaining attention, such as the relation between declining costs of renewable technologies and competitiveness, making renewable sources a key pillar of future energy transitions across both developed and emerging economies [10].

In this context, a key concept is given by energy communities, that can be defined as networks where people have ownership (or can influence the management) of renewable energy sources [17, 18, 14]. The literature analyzes both the technology used by energy communities and the increased efficiency triggered by their deployment. As for the first aspect, these communities are reported to use different technologies for generating, storing, sharing, and managing renewable energy: these include distributed sources, solar photovoltaic systems, and small-scale wind turbines, often combined with storage systems to mitigate intermittency and enable energy autonomy. In this context, Microgrids, Smart metering, and Information and Communication Technologies (ICT) support efficient local distribution, monitoring, and technical integration [7]: advanced energy system modelling and optimisation tools help design community operations, incorporating diverse business models and objectives [1].

As for the second aspect, energy communities that embed renewable energy sources have demonstrated significant potential in enhancing overall energy efficiency within local systems, leading to optimized resource utilization and reduced transmission losses [13]; Moreover, also non-financial outcomes have been investigated, since they can offer affordable energy and im-plement energy efficiency measures targeted at vulnerable households, hence mitigating energy poverty and fostering inclusive participation [8]: empirical studies further indicate that their im-plementation can lead of up a 75% cost reductions through dynamic local balancing of generation and consumption[12]. Of course, the two afore mentioned aspects are intertwined: for instance, blockchain-based platforms are developed to facilitate robust peer-to-peer energy trading, allowing members to directly exchange surplus energy [2].

Anyhow, investigating the effectiveness of such communities triggers the need to resort to multi-objective optimization since many aspects have to be taken into account to determine the optimal community configuration. From a methodological perspective, this problem naturally leads to large-scale combinatorial optimization problems, where exact methods rapidly become computationally infeasible.

In this contribution, we introduce a Meta-heuriostic approach to build energy communities that rely on renewable energy (i.e., solar energy) and compare it with a Montecarlo simulation, in order to show the robustness of the approach and to outline the metaheuristics' improved performances about execution time. Metaheuristics are high-level strategies that coordinate the action of subordinated heuristics to find near-optimal solutions of an optimization problem: they are used when the space to be explored is too huge for being dealt with by an exact methods, and they do not guarantee to find the optimal solution, but a near-optimal one within a reasonable amount of time. This will be useful in our case-study, since an exact solution would lead to the combinatorial explosion, being unable to find the optimal solution in a given time. Please notice that, given the increasing importance witnessed by the private residential sector for the use of solar systems, we are tackling residential units only.

The paper is structured as follows: section 2 will discuss data at hand; section 3 will describe the optimization model, and section 4 will outline the experimental analysis, before concluding in section 5.

2. Our Data

Our set of data comes from observations regarding different consumers' categories, i.e., resi-dential units, small offices, large offices, warehouses, secondary schools, hotels, restaurants and retail strip malls. The afore mentioned categories feature a great variability in demand profiles within the member of temptative energy communities: this can be seen from figure 1, that shows the main statistics of the electricity demand of buildings belonging to Residential Units, Hospitals, Small Offices, and Schools.

In what follows we focus on residential buildings, whose analysis is important to develop effective and sustainable energy management strategies and policies, on top of more and more informed infrastructure planning.[4, 6] The analysis of the other typologies can be found in the appendix.

Our data come from the Open Energy Data Initiative (OEDI) platform, which provides end-use load profiles for the U.S. building stock obtained via calibrated and validated physics-based simulations for multiple residential and commercial building types and end uses, across all U.S.

JOURNAL
OF
APPLIED
QUANTITATIVE
METHODS

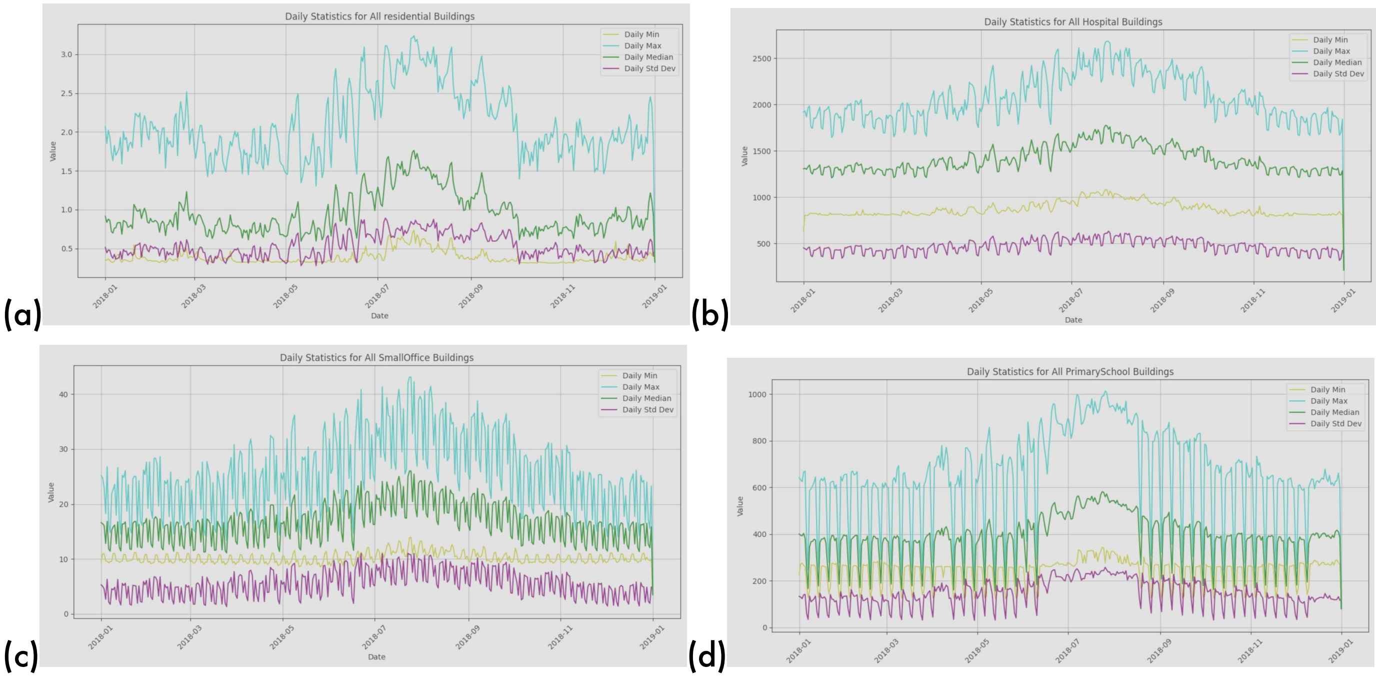


Figure 1: Main statistics of electricity demand of residential buildings(a), hospitals (b), small offices (c), and schools (d)

climate regions. In addition, the dataset includes historical weather information (e.g., humidity, pressure, wind speed and direction, temperature, and solar radiation), which is required to generate consistent PV production profiles [20].

As for residential buildings, we dispose, for each buildings, the following features: Square Footage, i.e., the total floor area of the residential building; Number of Bedrooms, i.e., the total number of bedrooms in the house; Number of Occupants, i.e., the total number of individuals resid-ing in the building; Bath Fan Energy, i.e., the energy consumed by bathroom exhaust fans; Ceiling Fan Energy, i.e., the energy consumed by ceiling fans used for ventilation or cooling; Clothes Dryer Energy, i.e., the energy consumed by the clothes dryer in the residential unit; Clothes Washer Energy, i.e., the energy consumed by the washing machine; Cooking Range Energy, i.e., the energy consumed by the cooking range or stove; Cooling Energy, i.e., the energy consumed by cooling systems; Dishwasher Energy, i.e., the energy consumed by the dishwashers. Total Electricity Energy, i.e., the total electricity consumed by the residential buildings; Exterior Lighting Energy, i.e., the energy consumed by outdoor lighting systems;

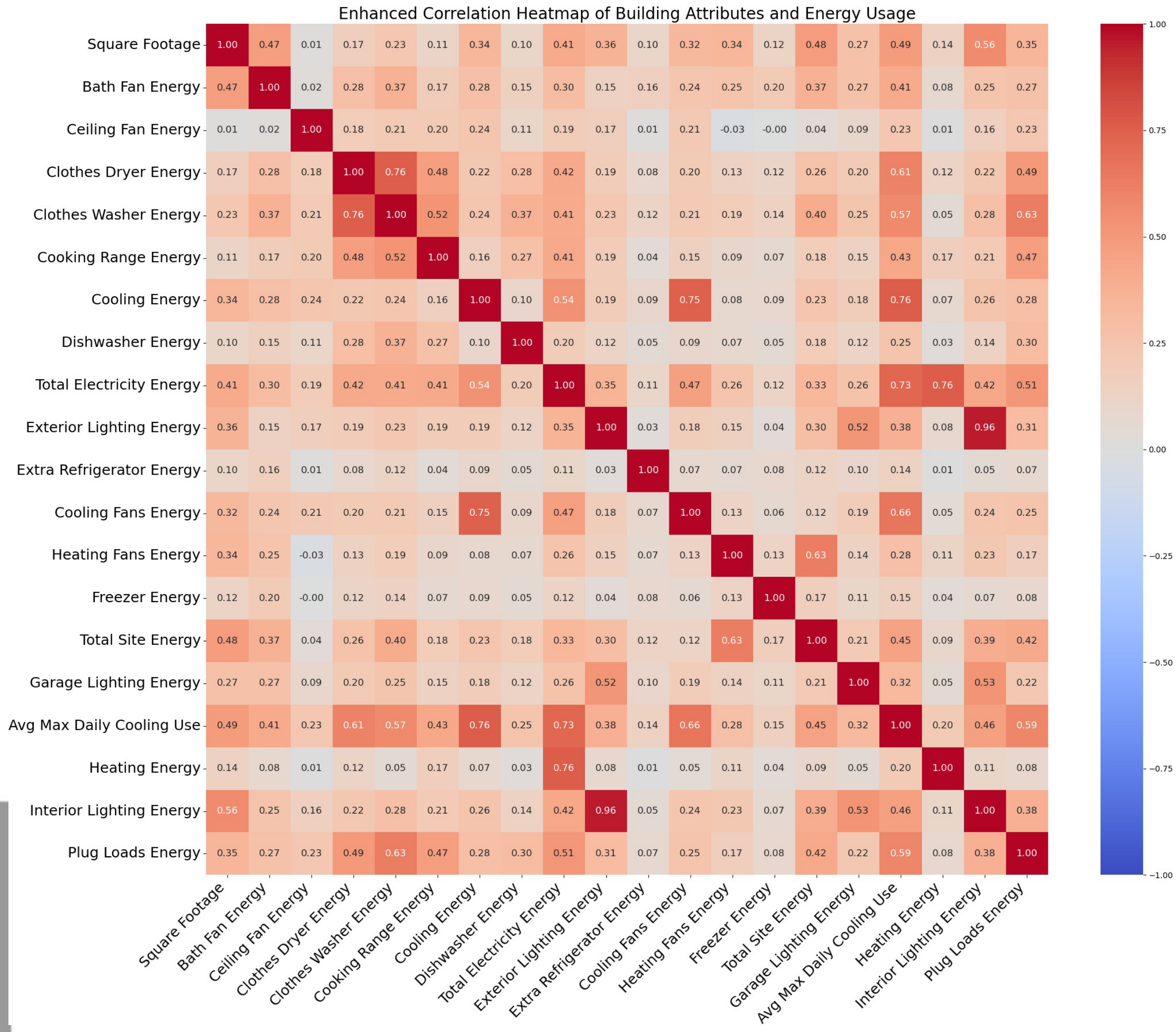
Extra Refrigerator Energy, i.e., the en-ergy consumed by an additional refrigerator; Cooling Fans Energy, i.e., the energy consumed by fans used for cooling purposes; Heating Fans Energy, i.e., the energy consumed by fans used for heating purposes; Freezer Energy, i.e., the energy consumed by standalone freezers; Total Site Energy, i.e., the total energy consumed on-site, including electricity and other sources; Garage Lighting Energy, i.e., the energy consumed by lighting systems in the garage; Avg Max Daily Cooling Use, i.e., the average maximum cooling energy usage per day; Heating Energy, i.e., the energy consumed by heating systems such as furnaces or heat pumps; Interior Lighting Energy, i.e., the energy consumed by lighting systems inside the house; and Plug Loads Energy, i.e., the energy consumed by appliances and electronics connected to electrical outlets.

As a first analysis, we have performed a correlation analysis amongst the different variables over all buildings, whose outcome is reported on the correlation matrix shown in figure 2.

As expected, we found a (moderate) positive correlation between **Square Footage** and **To-tal Electricity Energy** (approximately 0.41), as well as between, **Total Site Energy** and **Square Footage** (approximately 0.48), stressing out that the building size ranges amongst the determinants of the overall energy demand. It can also be seen that there exists a strong posi-tive correlation between **Cooling Energy** and **Total Electricity Energy** (approximately 0.75), showing the energy-intensive nature of air conditioning ; on the contrary, **Heating Energy** ex-hibits a weaker correlation with both **Square Footage** (0.34) and **Total Electricity Energy** (0.30), showing a smaller contribution to overall energy consumption, potentially due to regional climate factors.

Plug Loads Energy and **Interior Lighting Energy** exhibit strong positive correlations with **Total Electricity Energy** (approximately 0.51 and 0.96, respectively). These findings suggest that electrical appliances and lighting are among the dominant electricity consumers in residential buildings. **Dishwasher Energy** and **Clothes Washer Energy** show moderate correlations with **Total Electricity Energy** (around 0.27–0.30), indicating their moderate contribution to overall energy usage. **Cooling Fans Energy** and **Heating Fans Energy** correlate positively with **Total Electricity Energy** (0.47 and 0.45, respectively), showing their significant role in supporting HVAC systems. **Exterior Lighting Energy** displays a moderate correlation with **Total Electricity Energy** (0.35), which reflects its reliance on electrical systems, albeit less than interior lighting.

Figure 2: Correlation matrix between the attributes of residential buildings.



Extra Refrigerator Energy shows a weaker correlation with **Total Electricity Energy** (0.10), indicating that additional refrigeration units have a smaller impact compared to other factors like cooling or lighting. **Garage Lighting Energy** correlates weakly with **Square Footage** (0.27) and moderately with **Total Electricity Energy** (0.22). This suggests a localized impact, likely dependent on garage-specific activities.

3. Our Model

In this contribution, we aim at identifying cost-efficient renewable-based residential energy communities by minimizing the total operational costs through the coordinated use of photovoltaic (PV) generation and an energy storage system (ESS). In particular, the optimization framework is designed to properly exploit the energy produced by the community's PV system and to manage the ESS so as to store the excess PV generation, assuming equal charging and discharging efficiencies. The total cost of the community includes the electricity purchasing costs from the external grid, together with the operating and maintenance costs of the PV system and the charging/discharging costs of the ESS. Energy export to the main grid is not considered, since all surplus PV production is absorbed by the ESS, whose capacity is treated as a decision variable of the optimization problem. This assumption is consistent with small-scale residential communi-ties where grid export is limited or economically unattractive. This modeling choice promotes the full self-consumption of renewable energy and supports the development of energy communities aligned with the net-zero emission objective. We consider energy communities composed of N residential units, indicated with $i \in \{1, 2, \ldots, N\}$, and we model the system dynamics over hourly time steps $t \in \{1, 2, \ldots, 24\}$. For a given commu-nity size N , the objective is to minimize the total operational cost of the community, defined as the sum of grid electricity costs, ESS costs, and PV generation costs [5]:

$$\min \sum_{t=1}^{24} \left(\sum_{i=1}^{N} C_{i,t}^{Grid} + C_t^{ESS} + C_t^{PV} \right), \tag{3.1}$$

where $C_{i,t}^{Grid}$ denotes the cost of electricity purchased from the external grid by unit i at time t, while C^{ESS} and C^{PV} represent the operational costs associated with the ESS and the PV generation, respectively. These costs depend linearly on the amount of energy drawn from the grid, the energy charged into and discharged from the battery, and the energy produced by the PV system. The ESS is modeled as a battery device that allows the community to temporally shift energy by storing electricity when available and supply it when needed. Let SOC_t denote the state of charge of the ESS at hour t, while E^{cESS} and E^{dESS} represent the energy charged and discharged from the ESS during hour t, respectively. The ESS operation is subject to the following constraints:

$$SOC_{min} \leq SOC_t \leq SOC_{max}, \tag{3.2}$$

$$0 \leq E_t^{cESS} \leq 0.8 \cdot SOC_{max}, \tag{3.3}$$

$$0 \leq E_t^{dESS} \leq 0.8 \cdot SOC_{max}, \tag{3.4}$$

$$SOC_t^{Loss} = \eta_{Loss} \cdot SOC_t, \tag{3.5}$$

$$SOC_t = SOC_{t-1} + E_t^{cESS} \cdot \eta_c - \frac{E_t^{dESS}}{\eta_d} - SOC_t^{Loss}. \tag{3.6}$$

In this formulation, SOC_{max} is a decision variable of the optimization problem representing the maximum storage capacity of the battery, SOC^{Loss} denotes the self-discharge (loss) coefficient of the ESS, and n_c and n_d and are the charge and discharge efficiency parameters, respectively.

Finally, for a fixed number of units N, the system must satisfy the hourly energy balance constraint:

$$\sum_{i=1}^{N} E_{i,t}^{Load} = \sum_{i=1}^{N} E_{i,t}^{Grid} + E_t^{PV} - E_t^{cESS} + E_t^{dESS}, \tag{3.7}$$

where $E_{i,t}^{Load}$ denotes the electricity demand of unit i at time t. This constraint ensures that, at each hour, the total energy demand of the community is met through a combination of grid electricity, photovoltaic generation, and energy storage operations.

4. Our Experimental Setting

As stated in the Introduction, we want to introduce a meta-heuristic approach to the problem identified by [4], and compare it with a Monte Carlo simulation, in order to show the robustness of the approach and to outline the metaheuristics' improved performances about execution time. This section compares the two approaches, by introducing the Montecarlo simulation in section 4.1, and the metaheuristic approach in section 4.2.

4.1. Monte Carlo Simulation

A widely adopted methodology to evaluate the robustness of energy community sizing and operation problems under uncertainty consists in combining the underlying optimization model with a Monte Carlo (MC) simulation scheme. In [5], this approach is employed to jointly capture the stochastic variability of electricity demand profiles and weather-dependent photovoltaic (PV) generation.

The MC framework is structured as follows. At each simulation run, a day is randomly sampled from the reference year, thereby fixing the corresponding solar irradiation profile and the associated PV production. Then, an incremental energy community is constructed by progres-sively aggregating individual load profiles until a target community size N is reached. For each value of N , the resulting optimization problem is solved, and the relevant performance indicators are recorded. Repeating this procedure over a large number of independent replications (typi-cally 1000) allows for the computation of averaged outcomes and statistically robust estimation of system performance. While conceptually straightforward and effective in accounting for uncertainty, this MC-based strategy becomes computationally demanding as the community size increases, since it requires solving the optimization problem a very large number of times. Empirical evidence from the computational experiments on residential energy communities reported in [5] shows that the time required to complete the MC routine grows sharply with N . For small communities, the com-putational effort is limited to a few minutes (approximately 4.10 minutes for N = 1 and about 10.05 minutes for N = 9), but it reaches about 16.42 minutes at N = 10 and about 49.33 minutes at N = 15. For larger communities, the total runtime exceeds one hour (approximately 1h04m at N = 16, about 1h49m at N = 24, and about 1h56m at N = 25) and exceeds two hours for $N \geq 26$ (around 2h02m at N = 26-27, peaking at about 2h15m at N = 28, and remaining above 2h13m at N = 30).

For these reasons, in this work, we propose a meta-heuristic approach aimed at efficiently exploring the large configuration space of candidate communities while preserving robustness and significantly reducing execution times when scaling up the number of units.

To facilitate reproducibility and enable a direct quantitative comparison with our proposed meta-heuristic approach, table 2 reports the numerical results of the residential case study re-ported in [5] and the average computational time required by the Monte Carlo procedure for each selected community size. Specifically, the table collects the values obtained under a fixed PV capacity setting (Panel A) and under a variable PV capacity setting (Panel B), for increasing community size $N \in \{1, 5, 10, 15, 20, 25, 30\}$. IC_N denotes the individual cost of the energy community, defined as the total optimized community cost divided by the number of units N . The quantity SC_N represents the average single-unit cost obtained when the N units are optimized independently. The indicator PS_N denotes the percentage cost savings achieved by the member within the community with respect to the disaggregated configuration. Finally, $IESS_N$ and IPV_N indicate the individual energy storage system (ESS) capacity and the individual photovoltaic (PV) capacity, respectively, obtained by dividing the optimal community-level capacities by the number of units.

The numerical results confirm the main trends discussed in [5] for the residential case study. Under fixed PV capacity (Panel A), individual costs IC_N exhibit a monotonic decrease with the community size, together with increasing percentage cost savings PS_N , which stabilize around 13% for large N. At the same time, the individual ESS capacity $IESS_N$ decreases as N grows, highlighting the capital-efficiency gains induced by aggregation. When the PV capacity is left as a decision variable (Panel B), individual costs are systemati-cally lower than in the fixed-PV case. This effect is accompanied by significantly smaller percentage cost savings, which remain below 4% even for large communities. Moreover, the variable-PV setting leads to larger individual ESS and PV capacities.

Table 1: Residential case-study results reported in [5] for selected community sizes N.

N	1	5	10	15	20	25	30
Panel A: Fixed PV capacity							
IC_N [$]	5.7050	5.2268	5.0686	5.0342	5.0142	5.0227	5.0186
SC_N [$]	5.7050	5.7051	5.6263	5.6240	5.6214	5.6419	5.6478
PS_N [%]	0.0000	10.2564	11.9212	12.4935	12.7813	12.9694	13.1560
$IESS_N$ [kW]	2.6645	1.5411	1.3695	1.2803	1.2376	1.2091	1.1913
Panel B: Variable PV capacity							
IC_N [$]	4.8220	4.7604	4.6882	4.6770	4.6749	4.6864	4.6902
SC_N [$]	4.8220	4.8720	4.8174	4.8123	4.8132	4.8283	4.8336
PS_N [%]	0.0000	2.5305	2.9522	3.0905	3.1374	3.2166	3.2433
$IESS_N$ [kW]	3.6107	3.2009	2.9683	2.9153	2.8874	2.8647	2.8495
IPV_N [kW]	2.4310	2.4405	2.3831	2.3704	2.3606	2.3707	2.3768
Time [min]	4.10	6.01	16.42	49.33	97.62	116.12	133.88

4.2. Meta-heuristic approach

Meta-heuristics are general strategies that guide the action of subordinated heuristics, and can be partitioned into trajectory-based and population-based methods: in the first method, the exploration can be seen as a trajectory over the search landscape; in the latter, as the evolution of a discrete points set. In this contribution we have resorted to the latter class of methods, and implemented Particle Swarm Optimization to solve the problem devised in section 3. Particle Swarm Optimization (PSO), introduced by [11], is a population-based metaheuristic algorithm that gets inspiration from the coordinated movement and information sharing observed in flocks of birds or schools of fish: a swarm of particles explores the search space, with each particle rep-resenting a potential solution. Particles update their positions and velocities iteratively based on their individual best-known position (personal best) and the swarm's global best-known position (global best), enabling collective intelligence to guide the search toward optimal regions. A key strength of PSO lies in its effective mechanism for balancing exploitation (i.e., intensification) to find well performing solutions, and exploration (i.e., diversification), to refine promising areas for higher precision: this adaptive trade-off is facilitated by parameters such as inertia weight and acceleration coefficients, allowing the algorithm to efficiently handle multimodal, non-convex, and high-dimensional optimization landscapes. Several extensions and hybrid variants of PSO have been successfully proposed to address constrained and real-world optimization problems. In particular, hybrid PSO-based approaches have proven effective in financial applications, such as portfolio selection under complex cost structures and constraints, where adaptive penalty mecha-nisms are employed to handle infeasibility and improve convergence properties [3]. These features make PSO particularly effective for complex real-world problems, including those prevalent in finance, such as portfolio selection, risk management, and market forecasting, where objective functions often involve nonlinearity, constraints, and uncertainty.

In our implementation, we have considered a community of N buildings (as defined in Table 2), in which the fitness function is given by Equation 3.1. Each particle in the swarm contains the indicator of a candidate building, and at each iteration, each particle's fitness is evaluated. We track each particle's personal best position (*pbest*) and the global best position (*gbest*), cor-responding to the highest fitness value found so far. After updating, constraints are enforced to ensure feasibility, and this process is iterated for a fixed number of generations or until conver-gence, yielding the final optimal community. More details on the PSO can be sent upon request to the authors.

We can see that our PSO-based meta-heuristic approach leads to comparable results with regard to the Monte Carlo simulation, but a much lower computational time. Indeed, for small-sized instances ($N = 1$, $N = 5$) the results are comparable also with regard to the computational time: this is confirmed by a comparison of the results based on a Wilcoxon signed rank test (tested against a 0.05 significance level), which led us to reject the null hypothesis of difference between the two approaches over the two small sized instances, but not on the others.

5. Conclusions

In this contribution we address the optimization problem investigated in [5] by proposing a metaheuristic-based approach for detecting cost-efficient renewable-based residential energy communities. Specifically, we consider the optimization model for energy community with PV generation and ESS, and we aim at minimizing operational costs by promoting PV self-consumption through storage, with the ESS capacity treated as a decision variable.

Table 2: Residential case-study results obtained by our PSO approach for selected community sizes N .

N	1	5	10	15	20	25	30
Panel A: Fixed PV capacity							
IC_N [$]	5.7050	5.2267	5.5371	4.9914	5.1413	5.0714	5.007
SC_N [$]	5.7054	5.7055	5.6391	5.7194	5.4120	5.6920	5.6570
PS_N [%]	0.0000	10.4733	12.0041	13.4415	12.8178	12.9000	13.1492
$IESS_N$ [kW]	2.6647	1.5674	1.3651	1.0051	1.0615	1.3413	1.1922
Panel B: Variable PV capacity							
IC_N [$]	4.8218	4.7609	4.7251	4.6507	4.5001	4.6023	4.6900
SC_N [$]	4.8223	4.8717	4.6144	4.8375	4.9948	4.9438	4.8609
PS_N [%]	0.0000	2.5304	3.042	3.4162	3.5129	3.3402	3.1945
$IESS_N$ [kW]	3.6107	3.2012	2.8814	3.0143	2.979	2.5790	2.7302
IPV_N [kW]	2.4314	2.4438	2.3914	2.4961	2.5061	2.2987	2.2518
Time [min]	4.32	4.52	5.39	8.25	10.81	13.53	14.07

To deal with the intrinsic combinatorial nature of community detection and the computational burden arising from simulation-based approaches, we compared a standard Monte Carlo routine, which requires solving a large number of optimization scenarios, with a population-based meta-heuristic strategy based on PSO. The computational results on the residential case study highlight that the proposed PSO approach preserves the main economic trends observed under the Monte Carlo benchmark (i.e., cost reductions induced by aggregation), while providing a substantial reduction in computational time as the community size increases. In particular, for the largest community sizes considered, the PSO routine achieves near-comparable performance indicators with execution times that remain within minutes, as opposed to the significantly longer runtime required by the Monte Carlo procedure.

Overall, these findings suggest that PSO-based metaheuristics represent a viable and scalable alternative for energy community detection problems, enabling an efficient exploration of large configuration spaces while maintaining solution quality. Future research directions include extending the framework to multi-objective formulations (e.g., jointly optimizing costs, emissions, and equity metrics), incorporating additional sources of uncertainty, and accounting for energy export and market participation mechanisms.

6. Acknowledgments

The authors acknowledge financial support from the European Union - NextGenerationEU program, Missione 4 Componente 1, CUP D53D23006470006, MUR PRIN 2022 n. 2022ETEHRM "Stochastic models and techniques for the management of wind farms and power systems" by the Italian Ministero dell'Universitá e della Ricerca.

References

[1] Barabino, E., Fioriti, D., Guerrazzi, E., Mariuzzo, I., Poli, D., Raugi, M., Rezaei, E., Schito, E., and Thomopulos, D. (2023). Energy communities: A review on trends, energy system mod-elling, business models, and optimisation objectives. Sustainable Energy, Grids and Networks, 36:101187.

[2] Bhavana, G. B., Anand, R., Ramprabhakar, J., Meena, V. P., Jadoun, V. K., and Benedetto, F. (2024). Applications of blockchain technology in peer-to-peer energy markets and green hydrogen supply chains: a topical review. Scientific Reports, 14(21954).

[3] Corazza, M., Di Tollo, G., Fasano, G., and Pesenti, R. (2021). A novel hybrid pso-based meta-heuristic for costly portfolio selection problems. Annals of Operations Research, 304(1):109–137.

[4] De Blasis, R., Pacelli, G., and Vergine, S. (2024a). Energy community with shared photovoltaic and storage systems: Influence of power demand in cost optimization. Applied

Stochastic Models in Business and Industry, pages 1–23.

[5] De Blasis, R., Pacelli, G., and Vergine, S. (2024b). Energy community with shared photovoltaic and storage systems: influence of power demand in cost optimization. Applied Stochastic Models in Business and Industry.

[6] De Blasis, R., Pacelli, G., and Vergine, S. (2026). Investment valuation of photovoltaic and energy storage systems for diverse energy communities: A real option approach. Applied Energy, 404:127185.

[7] Gjorgievski, V. Z., Cundeva, S., and Georghiou, G. E. (2021). Social arrangements, technical designs and impacts of energy communities: A review. Renewable Energy, 169:1138–1156.

[8] Hanke, F., Guyet, R., and Feenstra, M. (2021). Do renewable energy communities deliver energy justice? exploring insights from 71 european cases. Energy Research & Social Science, 80:102244.

[9] International Energy Agency (2021). Net zero by 2050: A roadmap for the global energy sector. Technical report, IEA, Paris.

[10] International Renewable Energy Agency (IRENA) (2023). World energy transitions outlook 2023: 1.5°c pathway. Technical report, International Renewable Energy Agency, Abu Dhabi.

[11] Kennedy, J. and Eberhart, R. (1995). Particle swarm optimization. In Proceedings of ICNN'95-international conference on neural networks, volume 4, pages 1942–1948. IEEE.

[12] Lezama, F., Soares, J., Hernandez-Leal, P., and Kaisers, M. (2019). Local energy markets: Paving the path toward fully transactive energy systems. IEEE Transactions on Power Systems, 34(5):4081–4088.

[13] Lode, M. L., Te Boveldt, G., and Coosemans, T. (2022). A transition perspective on energy communities: A systematic literature review and research agenda. Renewable and Sustainable Energy Reviews, 163:112479.

[14] Minuto, F. D. and Lanzini, A. (2022). Energy-sharing mechanisms for energy community

members under different asset ownership schemes and user demand profiles. Renewable and Sustainable Energy Reviews, 168:112859.

[15] Parker, A., Horsey, H., Dahlhausen, M., Praprost, M., CaraDonna, C., LeBar, A., and Klun, L. (2023). Comstock reference documentation: Version 1. Technical report, National Renewable Energy Laboratory, Golden, CO. NREL/TP-5500-83819.

[16] REN21 (Renewable Energy Policy Network for the 21st Century) (2023). Renewables 2023 global status report collection. Technical report, REN21, Paris, France.

[17] Rescoop.eu (2022). Community energy: A practical guide to reclaiming power.

[18] Roberts, B. P. and Sandberg, C. (2011). The role of energy storage in development of smart grids. Proceedings of the IEEE, 99(6):1139–1144.

[19] Shukla, P. R., Skea, J., Slade, R., Al Khourdajie, A., van Diemen, R., McCollum, D., Pathak, M., Some, S., Vyas, P., Fradera, R., et al. (2022). Climate change 2022: Mitiga-tion of climate change. Contribution of working group III to the sixth assessment report of the Intergovernmental Panel on Climate Change, 10:9781009157926.

[20] Wilson, E., Parker, A., Fontanini, A., Present, E., Reyna, J., Adhikari, R., Bianchi, C., CaraDonna, C., Dahlhausen, M., Kim, J., LeBar, A., Liu, L., Praprost, M., White, P., Zhang, L., DeWitt, P., Merket, N., Speake, A., Hong, T., Li, H., Mims Frick, N., Wang, Z., Blair, A., Horsey, H., Roberts, D., Trenbath, K., Adekanye, O., Bonnema, E., El Kontar, R., Gonzalez, J., Horowitz, S., Jones, D., Muehleisen, R., Platthotam, S., Reynolds, M., Robertson, J., Sayers, K., and Li, Q. (2021). End-use load profiles for the U.S. building stock. Data set NREL/TP-5500-78192, National Renewable Energy Laboratory (NREL), Golden, CO, USA.

1. Appendix A. Energy conventions and standards in our set of data

Our datasets also include information about the energy conventions applied to various build-ings, where conventions may be instantiated to:

- **ComStock**, which is a highly granular, bottom-up model using multiple data sources, statistical sampling methods, and advanced building energy simulations to estimate the end user hourly energy consumption [15];
- The Database of Energy Efficiency Resources which is referred to as **DEER**;
- The U.S. Department of Energy, which is referred to as **DOE**:.

Depending on their usage, we map all energy conventions to the corresponding standards, that may be instantiated to *Older* (i.e., conventions established before the widespread adoption of modern energy efficiency practices), *M id−range*, (i.e., the ones implemented during a transitional period, developed with stricter efficiency requirements, but still far from the higher efficiency levels mandated by modern standards), and *Modern* (see table A.3 and figure A.3).

This information is also a key feature to measure how the different standards perform in different climate zones: this test would be helpful to assess whether modern standards show an energy savingover all instances: Figg. A.4 and A.5 are helpful to this extent.

The climate zone encoding refers to the American Society of Heating, Refrigerating and Air-Conditioning Engineers (ASHRAE) standard. The encoding *6A* refers to very cold winters, mild summers, high humidity and *6B* to very cold winters, mild summers, low humidity.

Based on the two charts we can observe that in different climate zones the electricity usage of buildings with different standards might not always reflect the intuition where we believe that the modern standards are always using the lowest electricity. In Fig. A.5 it is showcased that the mid-range standards produce lower electricity consumption for the commercial buildings opposed to Fig.A.4.

Table A.3: Convention Mapping

Convention	Standard Type
ComStock DOE Ref Pre-1980	Older Standards
ComStock DOE Ref 1980-2004	Older Standards
ComStock DEER Pre-1975	Older Standards
ComStock DEER 1985	Older Standards
ComStock DEER 1996	Mid-Range Standards
ComStock DEER 2003	Mid-Range Standards
ComStock DEER 2007	Mid-Range Standards
ComStock 90.1-2004	Mid-Range Standards
ComStock 90.1-2007	Mid-Range Standards
ComStock DEER 2011	Modern Standards
ComStock DEER 2014	Modern Standards
ComStock DEER 2015	Modern Standards
ComStock DEER 2017	Modern Standards
ComStock 90.1-2010	Modern Standards
ComStock 90.1-2013	Modern Standards

2. Appendix B. Commercial buildings

As for the commercial buildings, the data are obtained from the OEDI platform [20] partitions buildings belonging to the geographical area of Los Angeles (California) in the following classes:

- LargeOffice, i.e., offices featuring substantial square footage, e.g., corporate headquarters or large-scale administrative facilities.
- PrimarySchool, i.e., primary educational facilities that accommodate young students and host classrooms, administrative offices, cafeterias and gyms, just to name a few.
- Outpatient, i.e., healthcare facilities that provide non-overnight medical services, e.g., diag-nostic imaging and consultations.

- RetailStandalone, i.e., individual retail outlets that are not part of larger shopping centers.

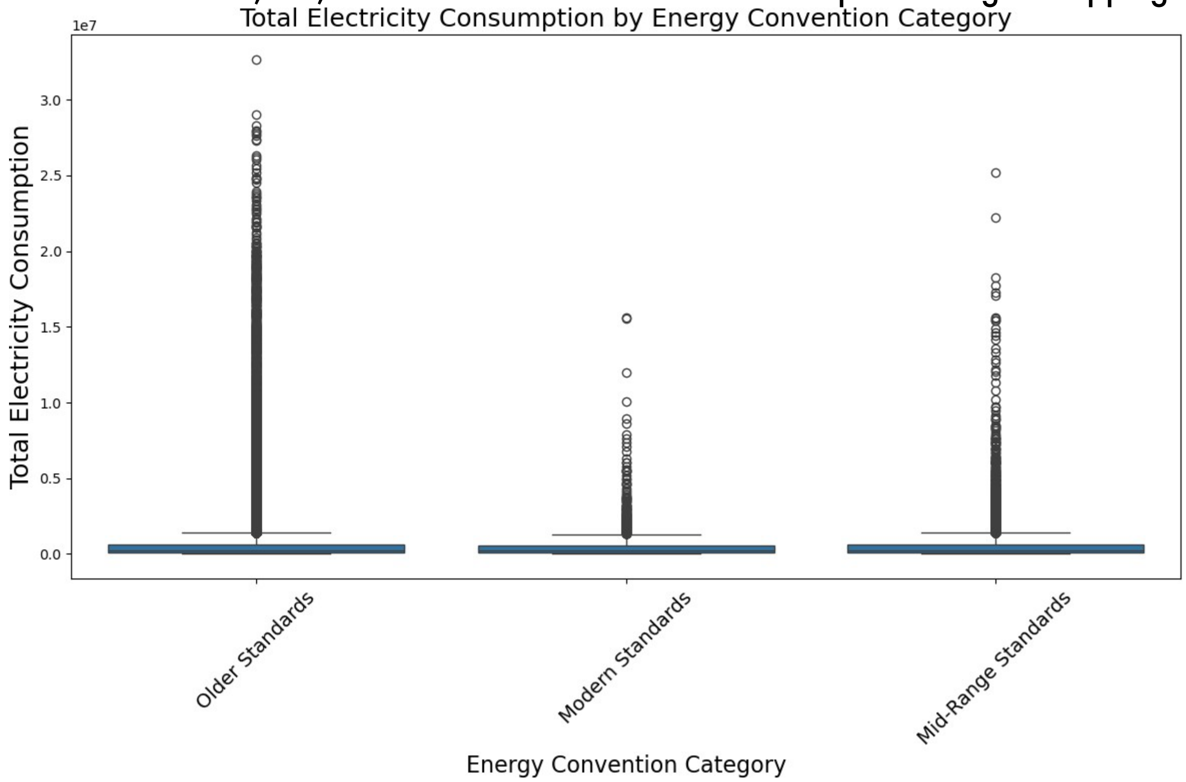


Figure A.3: *Total electricity consumption of all buildings, partitioned by each category of standards: the box plots are shown to assess the electricity consumption spread, highlighting the median, inter quartile range, and outliers for each group. Buildings with modern standards exhibits a more compact range of energy consumption, suggesting improved efficiency compared to older and mid-range standards.*

- MediumOffice, i.e., mid-sized office buildings, commonly used by businesses, law firms, or consultancies.
- QuickServiceRestaurant, i.e., fast-food facilities focused on high-speed service and limited on-site customers.
- FullServiceRestaurant, i.e., restaurants and related businesses offering service and an exten-sive menu.
- SmallHotel, i.e., small-scale hotels with limited rooms and amenities.
- Hospital, i.e., large healthcare facilities providing inpatient and outpatient services,

emer-gency care, and specialized treatments.

- SmallOffice, i.e., small office buildings housing independent professionals or small businesses.

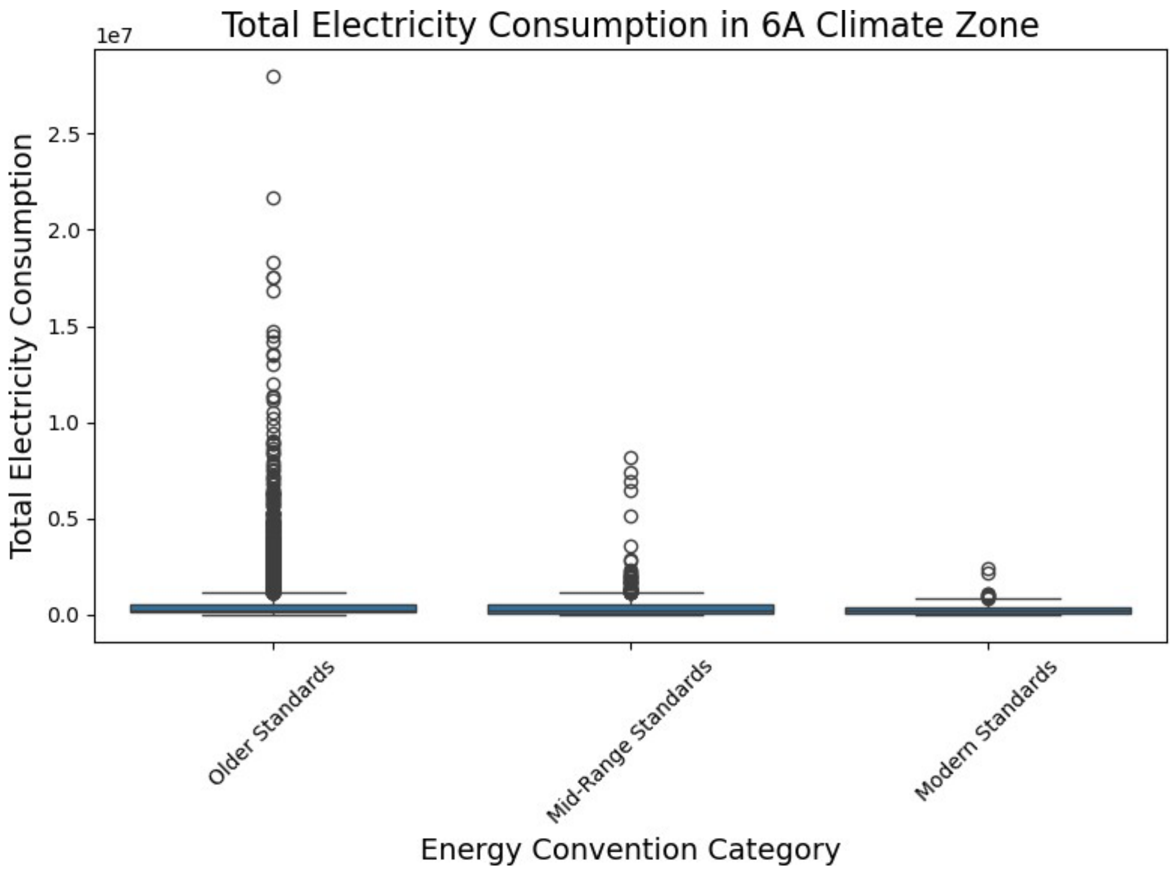


Figure A.4: *Different climate zone: Electricity usage of commercial buildings sectioned by standards in climate zone 6A*. This plot was made by ranging buildings based on their energy convention standards and showing there electricity consumption in every climate zone.

- Warehouse: Refers to buildings primarily used for storage and distribution of goods. Energy usage depends on heating, cooling (if climate-controlled), and lighting.

- LargeHotel, i.e., large-scale hotels offering extensive amenities such as conference rooms, restaurants, gyms, and spas.

- SecondarySchool, i.e., middle and high schools that provide education for teenagers, also hosting classrooms, laboratories, libraries, and sports facilities.

- RetailStripmall, i.e., clusters of retail stores that share a single building structure and park-ing area.

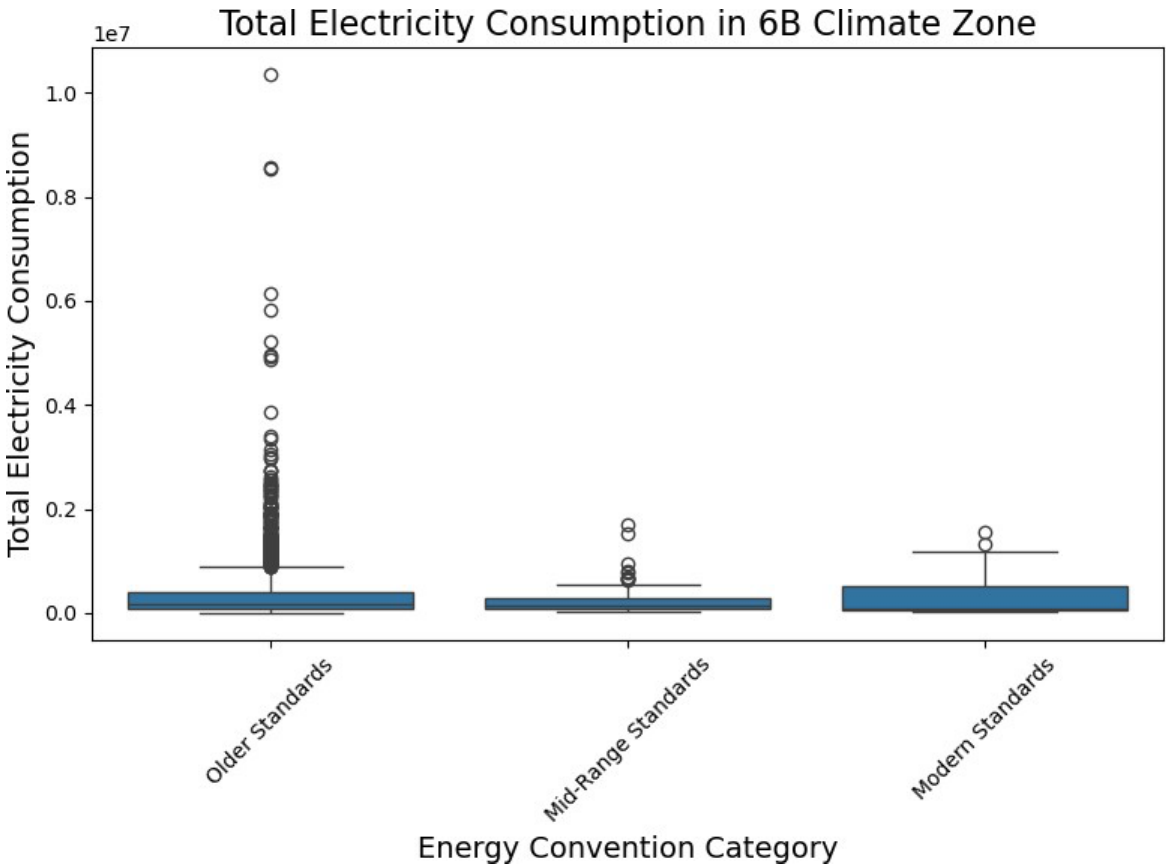


Figure A.5: *Different climate zone: Electricity usage of commercial buildings sectioned by standards in climate zone 6B*

For each building, we dispose of data about their square footage, energy usage conventions, and operational features such as: Square Feet (i.e., the total floor area of the building), Cooling Temperature Delta (i.e., the difference in cooling target temperature settings, measured in Fahren-heit degrees), the Heating Temperature Delta (as before but relative to heating), Number of floors in the building, Total Electricity Consumption of the building over a specified period, Total Site Energy Consumption, including all energy sources, Natural Gas Consumption, Cooling Energy (i.e., the energy consumed by cooling systems in the building), Eeating Energy (as before but rel-ative to heating), Interior Lighting Energy (i.e., the energy consumed by interior lighting systems within the building), Exterior Lighting Energy (i.e., the energy consumed by exterior lighting systems outside the building), Fans Energy (i.e., energy consumed by fans used in HVAC systems or other ventilation systems), Interior Equipment Energy (i.e., the energy consumed by inte-rior equipment such as appliances and office equipment), Refrigeration Energy, Water Systems Energy, Pumps Energy, Heat Recovery Energy (i.e., the energy consumed or saved by heat recovery systems in the building).

Similarly to what introduced on the main text, we have performed a correlation analysis amongst the different variables over all buildings, whose outcome is reported on the correlation matrix shown in figure B.6. As expected, we have found a strong positive correlations between **Total Electricity Consumption** and both the buildings' **Square Feet** (0.79) and **Total Site Energy Consumption** (0.73).

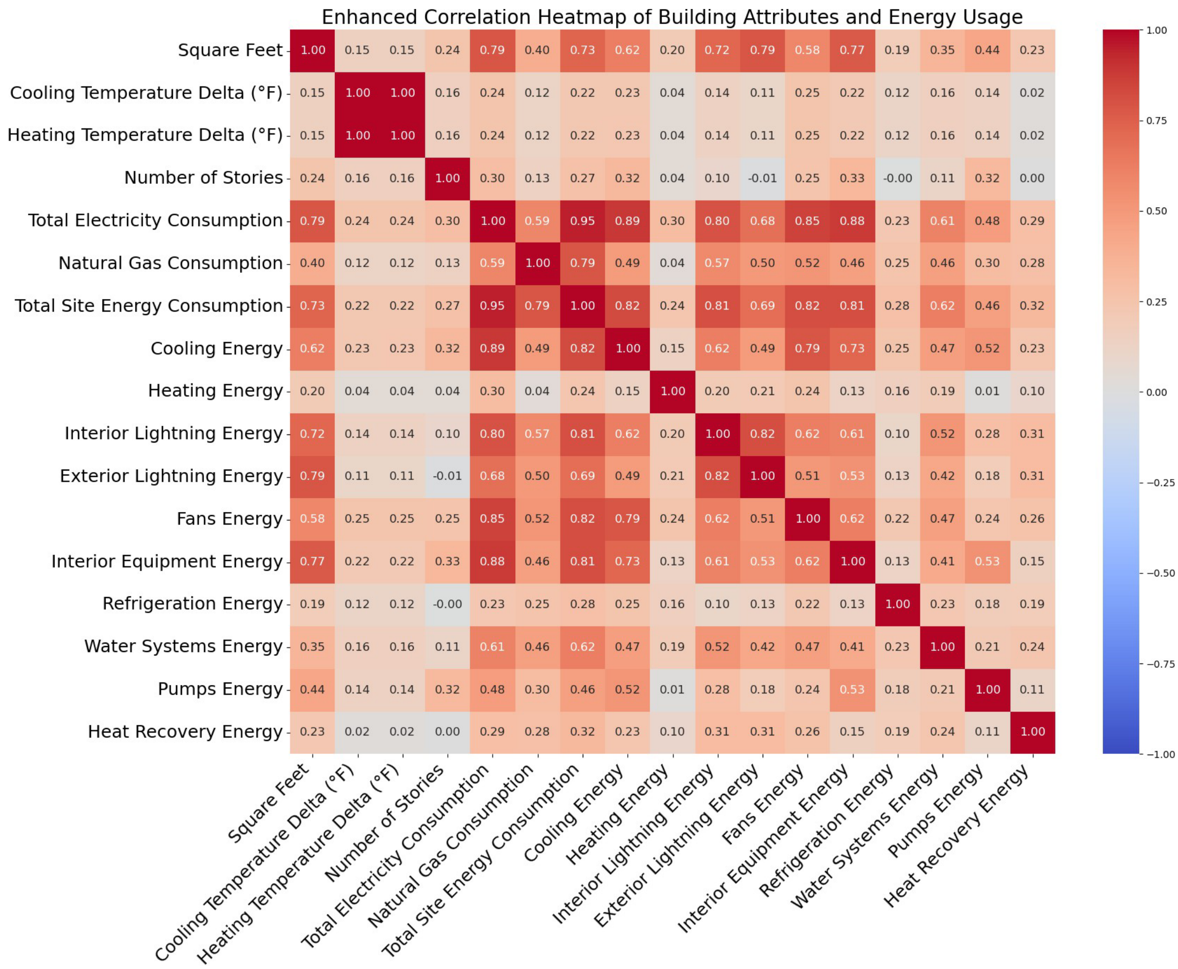


Figure B.6: Correlation matrix amongst all commercial buildings' operational features.

More interestingly, we remark a moderate correlation between **Natural Gas Consumption** and the building size (0.40), suggesting that larger buildings still rely on natural gas systems. We also remark that cooling systems and lighting highly contribute to larger buildings' overall energy demand, as witnessed by significant positive correlations between (**Cooling Energy – Interior Lighting Energy**) and (**Square Feet–Total Electricity Con-sumption**). Interestingly, Cooling and heating temperature deltas show low (albeit positive) correlations with **Total Electricity Consumption** and **Total Site Energy Consumption**, suggesting non-linear behaviours of these variables with respect to the energy consumption pat-terns: a further investigation of these aspects is left for further works.

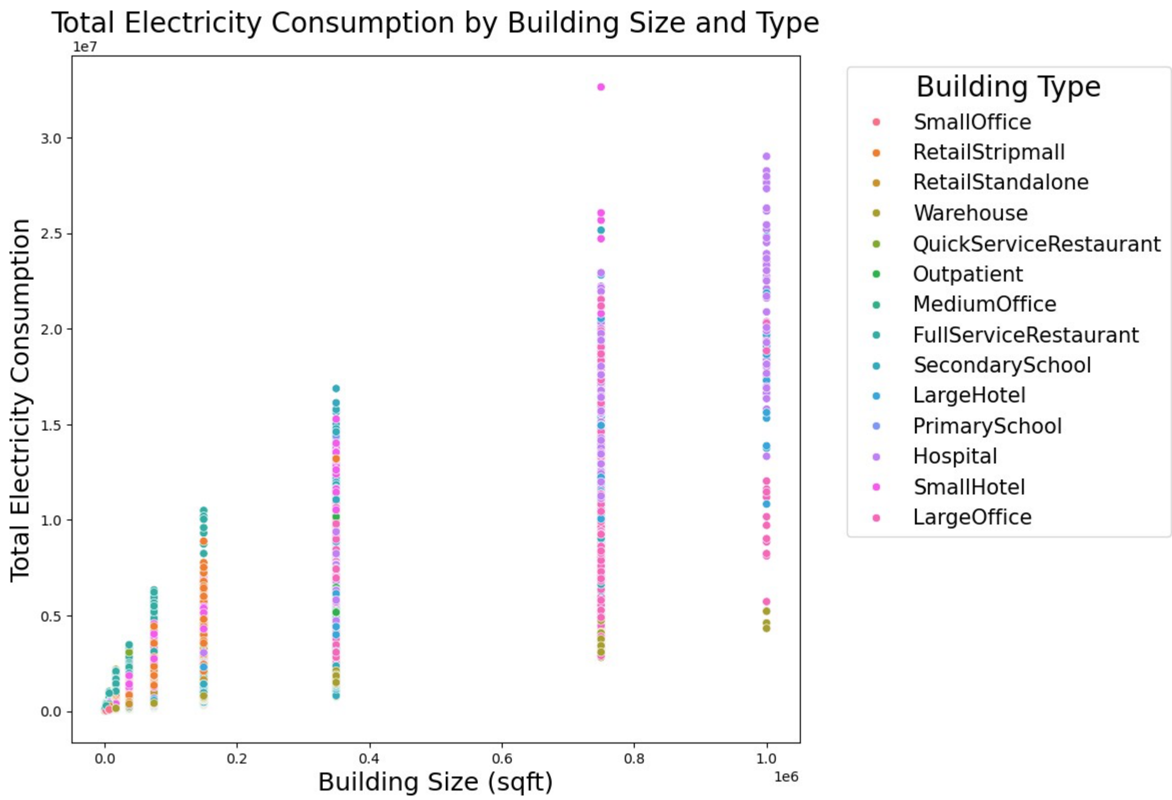


Figure B.7: *Electricity usage of commercial buildings with different sizes. The commercial buildings are partitioned in the building types shown in the legend.*

Furthermore, the **Fans Energy** (0.58) and **Interior Equipment Energy** (0.77) highlight a notable correlation to elec-tricity consumption, especially in larger buildings. As expected, we have found strong positive correlations between the building size and the energy usage (a linear regression between these two components would lead to an R^2 of 0.62). Figure B.7 shows the energy consumption of the buildings taken into account, that are partitioned in the building types shown in the legend. Overall, we may assert that Hospitals and Large Offices shows the biggest portion of energy consumption (see figure B.8), but, for ease of understanding, if we partition the buildings into **small-sized** (up to 200, 000 sqft), **medium-sized** (between 200, 000 sqft and 400, 000 sqft) and **large-sized** (above 600, 000 sqft), we remark that in the first class the highest energy consumption is mostly associated with full-service restaurants and medium offices, secondary schools and small hotels emerge as the primary contributors to energy consumption of the second class, whilst large offices and hospitals show the energy usage, with occasional contributions from large hotels.

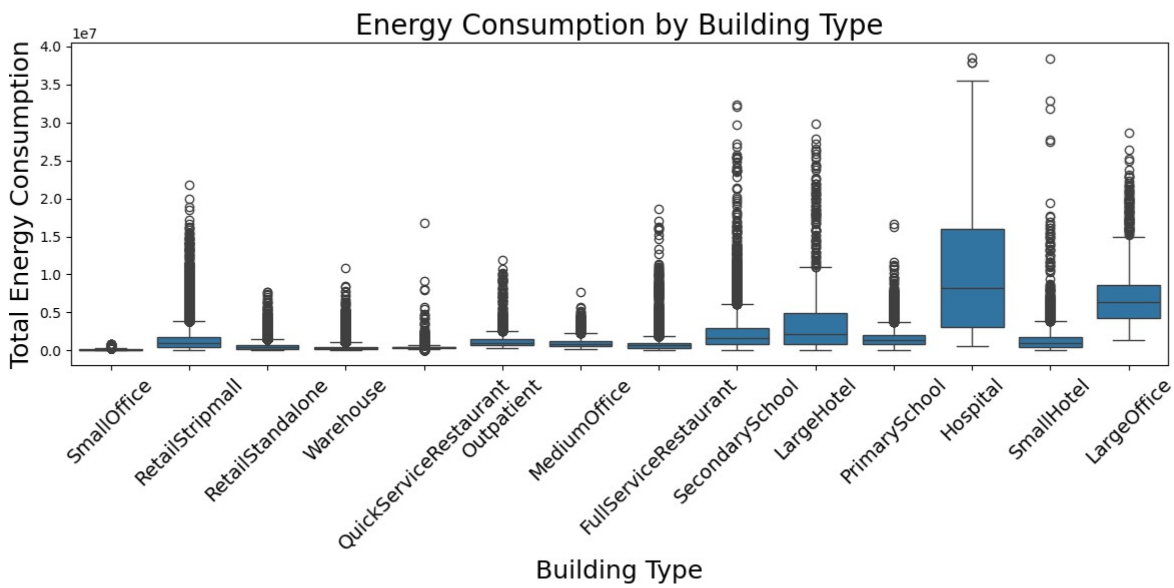


Figure B.8: *Electricity usage of commercial buildings partitioned by building type.* The plot was made by catego-rizing the buildings by their type and then creating a box plot for their energy usage.

JOURNAL
OF
APPLIED
QUANTITATIVE
METHODS

JOURNAL
OF
APPLIED
QUANTITATIVE
METHODS