Politecnico di Torino

Politecnico di Torino

Master's degree in Energy and Nuclear Engineering

Master's degree Thesis
Academic Year 2021/2022
Graduation Session: July 2022

Development of AI-based models for the management of energy communities

Supervisor:
Prof. Alfonso CAPOZZOLI

Candidate:
Laura MODICA

Co-supervisors:
Ing. Antonio GALLO
Dott. Marco Savino PISCITELLI

The term "Energy", when associated with energy communities, has a double meaning: on the one hand it refers to clean energy, that is, to the renewable sources scattered throughout nature which enable us to light and heat our homes; on the other hand, it refers to the energy of people who voluntarily decide to come together to make a difference, to cooperate.

Abstract

The transition towards a new sustainable development paradigm and the promotion of carbon neutrality by 2050 envisages the enforcement of several energy-saving measures in the buildings sector, which accounts for more than a third of total energy consumption. The distributed nature of renewable energy resources is the key to promote local energy production in buildings. Unfortunately, most of the renewable energy sources, employed in the residential field, are powered by solar or wind energy, marked by a non-programmable nature. This could lead to mismatches between energy production and consumption, undermining the stability of power grids. In order to overcome these problems, a new model is increasingly being developed based on Energy Communities (ECs). These take the form of a group of electric and thermal utilities which share the same power generation systems and are able to consume a good deal of self-produced energy. Coordinated management of buildings belonging to an EC is one of the key requirements for making an interaction with the power grid effective and impactful. This can be achieved through the deployment, at different scales, of Internet of Thing-based technologies, which enable long-term monitoring data. The subsequent development of a data-driven Virtual Simulation Environment (VSE) allows us to understand the role played by different integrated energy technologies at different scales, from individual buildings to entire aggregates. The results achievable through this model are helpful in identifying the best combination of energy systems to meet multiple objectives which include environmental, energy and economic ones through the assessment of Key Performance Indicators.

This work set out to make an existing VSE, enforced on a district scale, more flexible and suitable for simulating real-world contexts. In detail the current research

suggests the training and the validation of Artificial Neural Networks applied on a building scale, with the aim of simulating the thermal dynamics of multiple units belonging to different ECs. The performances of these data-driven supervised learning models are then compared whit those of grey-box models developed by other researchers from data coming through the same measurement program. Given the regulatory development drivers based on economic enhancements and consequent reductions in the cost of purchasing electricity from the grid, we then deployed models capable of describing the operation of a Micro Combined Heat and Power (μ-CHP) system and reconstructing the energy demand for charging electric vehicles. Three different scenarios were thus investigated, with the aim of meeting the energy needs of members and optimally managing energy flows with respect to the power grid. The results emerging from these analyses show advantages when energy production takes place on site through a centralized μ-CHP plant, sized to meet the entire EC's requirements. Due the increasingly ambitious decarbonization goals, electric vehicle offer a set of important advantages in terms of sustainable mobility. However, emissions related to power generation to recharge these units should not be overlooked. Meeting their electricity demand through a μ-CHP system results in lower costs than relying only on the power grid, especially in high-cost periods. In addition to this, that strategy could result in a reduction of grid dependency and of greenhouse gas emissions where biofuels are available at affordable costs.

List of Contents

List of Figures

List of Tables

List of Acronyms

AC	Alternating Current
ANN	Artificial Neural Network
BESS	Battery Energy Storage System
CEC	Citizens' Energy Communities
CHP	Combined Heat and Power
COP	Coefficient Of Performance
CSC	Collective Self-Consumption
CSV	Comma Separated Value
CTES	Cold Thermal Energy Storage
CV-RMSE	Coefficient of Variation of Root Mean Square Error
DC	Direct Current
DHN	District Heating Network
DHW	Domestic Hot Water
DLR	Deep Learning Reinforcement
DR	Demand Response
DSM	Demand Side Management
DYD	Donate Your Data
EC	Energy Community
EER	Energy Efficiency Ratio
EH	Electric Heater
ELF	Electrical Load Factor
ESS	Energy Storage System

EV	Electric vehicle
FF	Flexibility Factor
GB	Gas Boiler
GWP	Global Warming Potential
HP	Heat Pump
HTES	Hot Thermal Energy Storage
HVAC	Heating, Ventilation and Air Conditioning
IEM	Internal Market in Electricity Directive
JSON	JavaScript Object Notation
k-NN	k-Nearest Neighbors
KPI	Key Performance Indicator
LSTM	Long Short-Term Memory
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
MAS	Multi-Agent System
MGT	Micro Gas Turbine
MSE	Mean Square Error
PAR	Peak-to-Average Ratio
PV	Photovoltaic
RBC	Rule-Based Control
REC	Renewable Energy Community
RED II	Renewable Energy Directive II
RMSE	Root Mean Square Error
RNN	Recurrent Neural Network
SCI	Self-Consumption Index
SoC	State of Charge
SSI	Self-Sufficiency Index
TES	Thermal Energy Storage
VSE	Virtual Simulation Environment

1 Introduction

This chapter traces the historical evolution of relevant regulatory frameworks and the strategies adopted at national and international levels in the field of energy and climate change. As part of the latest strategies issued by the European Union, ECs are explored in detail, addressing the pertinent legal structures and the role of digitalization, with a focus on Italy. Thus, a state-of-the-art analysis regarding VSEs and an in-depth look at methods for comparing the performance of different ECs and buildings will be conducted.

1.1 The history and regulatory framework of Energy Communities

Over the past two decades, the evolutionary dynamics related to the national and international energy system have been strongly influenced by *anthropogenic climate change*. This phenomenon is largely driven by atmospheric emissions of climate-altering such as carbon dioxide CO_2, nitrogen monoxide N_2O, methane CH_4, hydrofluorocarbons that comes from human activities. Scientists claim that human activity has resulted in a 47% increase in the amount of carbon dioxide in the atmosphere compared to levels observed in 1750 [1], mainly due to energy generation. To address this phenomenon, the programmatic pathway developed at the European level involves driving the *energy transition* through the reduction of greenhouse gas emissions released into the atmosphere, the use of renewable energy to meet final consumption and the promotion of tools to improve energy efficiency. These key aspects have been the cornerstone for the organization of several meetings at the international level, which have included the enforcement of regulatory frameworks at different stages. Since ‘70s, researchers have been aware about the

impact of anthropogenic emissions, but it was only in 1997 when governments decided to sign the Kyoto Protocol. It was an international agreement that came into effectiveness in 2005 and imposed an obligation to reduce greenhouse gas emissions between 2008 and 2012 by no less than 5% from the level recorded in 1990 [2].

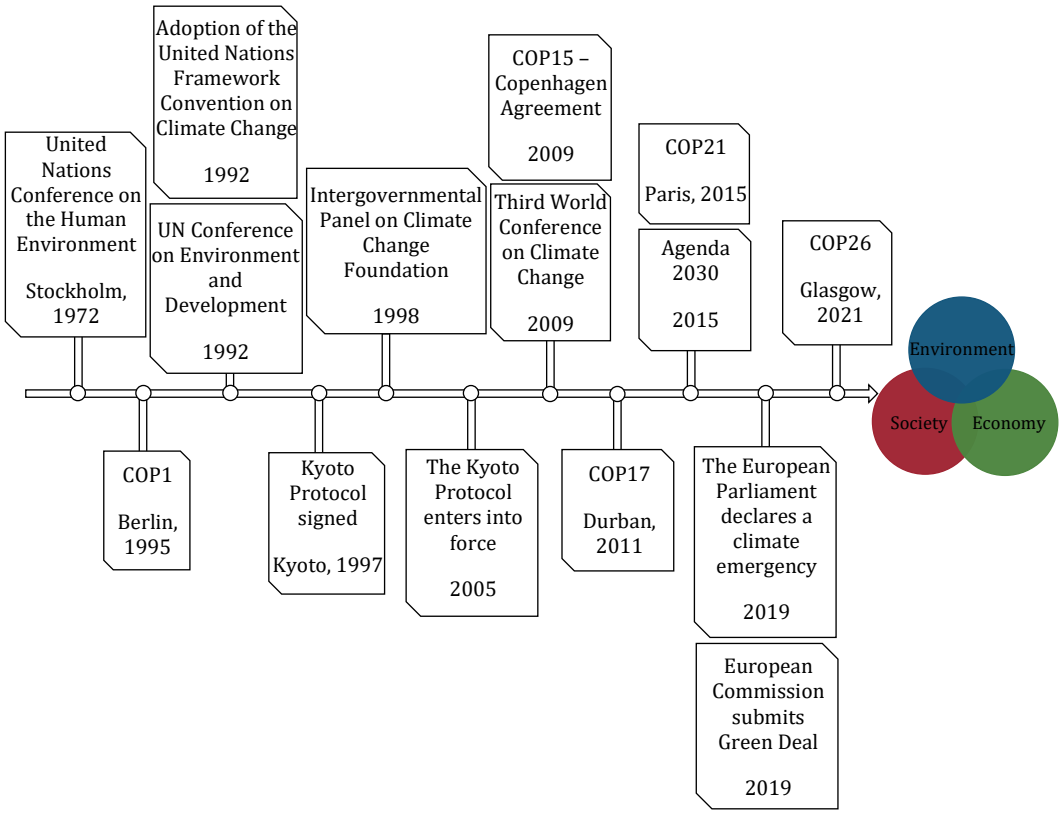


Figure 1.1: The stages of approach towards sustainable development.

In the following years, the targets set by the international community were continuously raised. During COP15[1] in 2009, the Copenhagen Agreement was signed to support limiting global average temperature rise to less than 2 °C. In 2015, during COP21 in Paris, member states signed an agreement to keep the global average temperature increase to well to less than 2 °C and, if possible, within 1.5 °C. To achieve the goal of *carbon neutrality* by 2050, each member state must also develop a Nationally Determined Contribution [3], which is an action plan describing the actions it intends to take to reduce emissions and climate impacts, updated every five years.

[1] COP stands for Conference Of Parties. This term refers to the annual meeting of the member countries that signed the Framework Convention on Climate Change on 9th May 1992, an international agreement on the environment which subsequently entered into force on 21st March 1994.

At COP26 in Glasgow, one of the most relevant provisions was to limit temperature increase to an overall level of less than 1.5 °C compared to pre-industrial era. In Figure 1.1, the timeline of the main international meetings is reported.

The European Union is playing a leading role in global efforts to prevent or minimize the negative effects of these phenomena, by encouraging and supporting the national initiatives of member states, as the first international organization to formally commit to achieve carbon neutrality by 2050. The Clean Energy for all Europeans Package [4] was launched at the community level in November 2016. It consists of a set of policies that incorporates eight legislative acts, which include key actions to be taken between 2021 and 2030. This plan requires all member states to commit by 2030 to cutting emissions by 40% or more (compared to the level recorded in 1990), to increase and promote renewable energy production by at least 32% of total final consumption and to improve energy efficiency in final consumption by no less than 32.5%. More recently, the European Union set out an even more ambitious strategy to achieve emission reduction targets, proposing a new action plan to cut them by 55% before 2030. The Clean Energy for all Europeans Package encompasses several European Directives, including the Renewable Energy Directive II (Directive 2018/2001/EU RED II) [5] and the Electricity Internal Market Directive (Directive 2019/944/EU IEM) [6], on which the organizational structure of ECs is based. Specifically, the first one introduces definitions of *Collective Self-Consumption* (CSC) and *Renewable Energy Community* (REC), while the latter defines the *Citizens' Energy Communities* (CEC). Articles 2 and 21 of the RED II [5] describe a CSC scheme, whose members are at least two consumers who produce renewable energy to meet their consumption needs and who are part of the same building. Shared renewable energy can be produced by facilities owned or operated by third parties who are not considered self-consumers of renewable energy. Articles 2 and 22 of the same Directive define the REC as "A legal entity which, following applicable national law, participates voluntarily, is autonomous and controlled by shareholders or members located in the proximity of renewable energy projects. Shareholders can be natural persons, local authorities or Small and Medium-sized Enterprises. Renewable energy communities have the power to take part in all relevant electricity markets" [5]. This scheme aims at bringing environmental, economic and social benefits to the community through the production and sharing of renewable energy. One or more private companies may belong to the REC, as long as this does not represent a major

commercial or professional activity. The IEM Directive [6], in Articles 2 and 16, analyses the CEC with reference to "A legal entity that takes part voluntarily and is controlled by natural persons, local authorities, members or partners of small businesses. It can cooperate in the generation, also from renewable sources, in distribution, consumption, storage, supply of energy and energy efficiency services". It also considers the CEC as a solution which has the potential to support technological developments, such as Smart Grids and integrated demand management. The main difference between the two proposed schemes is that the REC is based on the integration of an ever-increasing renewable energy sources to produce electricity and heat, as well as the use of biofuels. On the other hand, the CEC only considers electricity and addresses issues related to power distribution grid. In addition to this, the use of renewable energy is not mandatory, but the main goal is to develop equal and non-competitive conditions as market players. There are no geographical limitations for CECs, so users do not need to be close to the energy project for participating as members. Conversely, RECs are based on precise territorial proximity requirements. Both types of ECs require legal entities to act as coordinators and effective control from specific actors.

RECs are one of the most sustainable forms of renewable energy development, as they allow for a reduction in land use since the plants are installed on spaces already used for other purposes. Consumers involved in groups and communities can be made more aware of their consumption, resulting in cost control and activation of energy efficiency measures. Therefore, social value is an essential feature assumed by ECs, as they can create shared value where they are deployed, empowering citizens and local communities to perform as effective actors in the energy system, while promoting the development of skills, technologies and professionalism. Finally, they fit in as an important tool to eradicate energy poverty by including those users that are the most vulnerable. Thanks to these benefits, some EC-related projects are developing at an increasing rate, also through the evolution of the regulatory framework that provides for the delivery of economic subsidies. Around 3500 ECs are already operational in European countries, mainly concentrated in North-western Europe [7]. Most of the ECs are developed in Germany, followed by Denmark while Italy has a lower advancement of these entities than even other less developed countries from an energy point of view.

	Germany	Denmark	Netherlands	UK	Sweden	France	Belgium Poland Spain	taly
No. ECs	1750	700	500	431	200	70	34	20

Table 1.1: Number of active ECs in some European countries.

In our country, the Integrated National Energy and Climate Plan [8] set up the lines of action, which promote energy security and energy market development to lead the transition. Its introduction into the Italian regulatory framework took place in conjunction with the partial transposition of the RED II through Ministerial Decree 12/2019 Milleproroghe. Pursuant to Article 42bis of the aforementioned Ministerial Decree, an experimental framework has been launched regarding CSC and RECs, legally recognized as of 2020 [9]. This Ministerial Decree, converted by Law No. 8/2020 of 28th February 2020, provided that members of ECs may also be simple consumers, i.e., not owning a production plant, as long as the point of withdrawal from the power grid belongs to the same secondary cabin and the energy is produced by a plant of a maximum capacity of 200 kW. In the face of this partial transposition in Italy, ARERA[2] has issued Resolution 4th August 2020 318/2020/R/eel and Ministerial Decree 16th September 2020. Two different draft laws have been proposed in order to address two issues. The first one concerns a regulatory procedure which identifies the different actors and their relationship and indicates the practices that encourage self-consumption. The second one refers to an act of ministerial competence, which addresses the issue of incentives in detail. The regulatory model introduced by the Authority takes the form of a *virtual energy sharing*. It consists of an hourly balance between the energy which is fed into the grid by community renewable energy source and the energy which is withdrawn from the power grid by the members within the same community. The lower of the two quantities, hour by hour, stands for shared energy among the members. This model favours the self-consumption without infrastructural interventions, as no new metering equipment is to be installed and the use of the public power grid is envisaged – as a consequence, the development of a private power grid for electricity exchange is not contemplated –. There are two types of remuneration provided to members belonging to the CSC scheme or REC on the share of energy produced and shared. The first concerns a refreshment linked to the

[2] ARERA is the Italian Regulatory Authority for Energy, Networks and Environment.

decommitment of the power grid portion affected by this scheme – provided for by the Authority – calculated in relation to the variable share of power grid charges. Another one concerns an incentive issued by the Energy Services Manager[3], deriving from the Ministerial Decree of 16th September 2020. Through the entry into force of Legislative Decree No. 199 of 15th December 2021 [10], the RED II has been fully transposed. The novelties introduced by this action included the increase in the maximum power for individual plant which can be incentivized, the electrical perimeter from the secondary cabin to the primary cabin and the incentivization of CSC within 30% of the existing power of the total related to plants held by the EC or in its availability.

Several projects related to both CSC and RECs are active in many Italian cities, while CECs have not been established. Currently, applications concern different contexts: residential and pertaining to Small and Medium-sized Enterprises. Figure 1.2 and Figure 1.3 depict the distribution of ECs in relation to type and energy source used for local energy production.

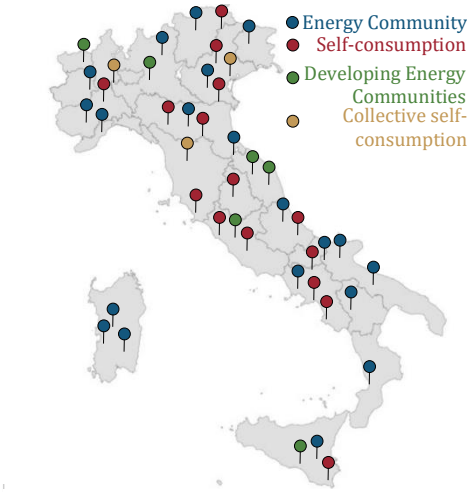


Figure 1.2: Distribution of the types of ECs in the national territory.

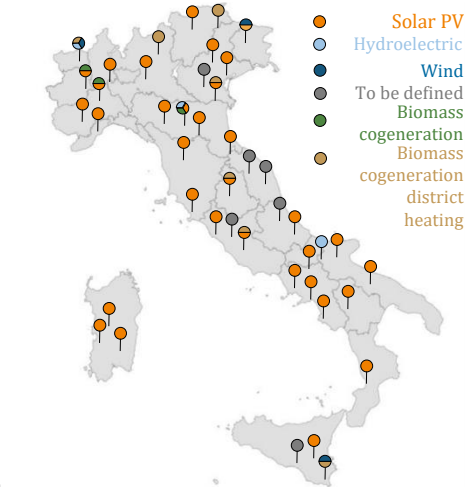


Figure 1.3: Distribution of energy resources of ECs in the national territory.

There are twelve RECs built at national level, while CSC schemes have been realized by four organizations. The predominantly used energy sources are related to solar

[3] GSE is the Italian Energy Services Manager.

and hydropower, but in some cases other sources such as biomass[4] or biogas[5] are facilitated. As a matter of fact, the first ECs in Italy were born before the enactment of the RED II, mainly in the North of the country. In 2018, the approval of a law by the Piedmont Region played a fundamental role as it was the first legislative project regarding ECs [11].

Following the damage caused by COVID-19, several countries, including those belonging to the European Union, have announced extraordinary measures to support their economies. For this reason, a significant portion of their own funds will be dedicated to energy transition. This included, at the national level, the design of the Italian National Recovery and Resilience Plan. Under Mission 2 of this programme, "Green Revolution and Ecological Transition", the development of small-scale distributed systems and the promotion of ECs is planned, for a total investment of 2.2 Mld€ within Component 2 "Renewable energy, hydrogen, network and sustainable mobility" [12]. This action could expand the experimental activities on ECs already started with the RED II and are anticipated to develop projects in areas where the greatest territorial and social impact are expected. The plan also provides support for smaller towns – with fewer than 5000 inhabitants – which are often at risk of depopulation.

Regulatory frameworks at European and national level require the introduction of a variety of measures to meet market needs in line with the evolution of renewable and more sustainable alternatives to fossil fuels which could put power grids into trouble. The maintenance of the instantaneous balance between supply and demand is ensured through generation managed mainly by thermal power plants. The increasing mix of generation from intermittent renewable energy sources and the gradually clear expansion in peak load due to a continuous intensification of electrification in buildings require a complete reorganization of the energy infrastructure. There are two main problems, related to ECs, which are addressed by our study. The first one concerns the need for energy flexibility of buildings, to ensure that the power grid act adequately under conditions of sudden changes in demand, taking into account the necessary balances requirements and resilience. The second

[4] Alpine Energy Community of Tirano, district heating of Dobbiaco-San Candido, Azienda Agricola Valier, Società agricola Fattoria Lucciano, Società Agricola podere Vallescura.
[5] Pinerolese Energy Community, Green energy community GECO, Cooperativa Agricola Speranza.

one involves the difficulty of developing district-scale environments that can reproduce real-world contexts. To address the first issue, the individual building has the potential to become an active system that can meet the requirement for energy flexibility, defined as "the ability to manage its demand and generation according to local climate conditions, user needs and power grid requirements (penalty signal) without jeopardizing technical and comfort constraints" [13].

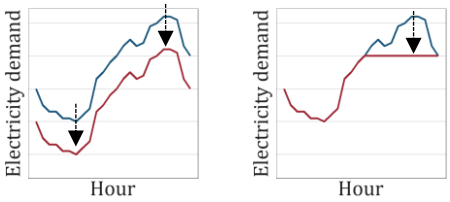
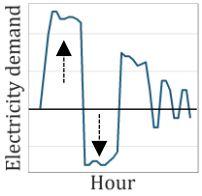
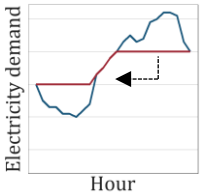
Peak Clipping and Energy Efficiency: It consist in lowering the peak of demand or total demand.	
	Figure 1.4: Peak clipping and energy efficiency scheme [14].
Valley Filling: It consist in increasing the load in period when there are no peaks in demand.	
	Figure 1.5: Valley filling scheme [14].
Load shifting: It consist of a shift of the load from a peak period to a downstream period.	
	Figure 1.6: Load shifting scheme [14].

Table 1.2: Strategies to achieve energy flexibility.

The energy flexibility of the building enables Demand Side Management (DSM) and load control strategies and is an alternative to Supply Side Management. According to this management scheme, high operating costs due to supply-side resource management can be reduced. In addition to this, power purchase costs can be moderated as fewer energy reserves are used to balance peak situations. Specifically, DSM consists of a set of actions which affect users consumption of primary energy and includes various activities that can be introduced to modify energy demand over time [15] and exhibit flexibility, such as those described in the Table 1.2. Often, energy flexibility in buildings can be managed by decoupling energy

production from consumption through storage systems. There are several ways to strengthen this property, but the most interesting are the following ones:

- **Flexible loads:** they include all schedulable utilities, whose loads can be shifted over time according to Demand Response (DR) policies. Controllable equipment in a building includes Heating, Ventilation and Air Conditioning (HVAC) systems or household appliances.
- **Hot Thermal Energy Storage (HTES) and Cold Thermal Energy Storage (CTES):** these are set up as storage technologies which typically contain hot water and chilled water, respectively. They are capable of storing and releasing energy, weighing more or less flexibility on the power grid depending on their size [16].
- **Building envelope:** the thermal capacity of the envelope components provides additional opportunities for thermal storage. Opaque envelope is increasingly using materials which exploit the principle of phase change accumulation[6], while this technology for transparent envelopes is still under development.
- **Battery Energy Storage Systems (BESSs):** they are set up as storage systems capable of electrical energy buffering. When connected to a local power generation system, they can be leveraged to increase self-consumption by charging when production exceeds consumption. Alternatively, they can be charged when power grid electricity prices are low and discharged during peak periods. Electric vehicles (EVs) can also be configured as systems for storing electricity which can be shared by community members when needed.
- **Power generation systems and energy carriers:** Two types of power generation systems may be involved, namely renewable sources or CHP generation systems.

Regardless of how energy flexibility is defined, it can be expected that it will not become a constant value over time since it is based on interactions with the building's interior and exterior environment [17]. The management of a district of buildings, each able to exercise its energy flexibility, determines greater advantages

[6] Phase change materials (PCM) are materials with high storage capacity and exploit the latent part of the phase change.

towards the network than a single building, whose impact is certainly low. However, it is necessary for them to be served by the same aggregator, as otherwise there could be negative impacts in terms of power grid stability, such as the phenomenon of only shifting the peak of demand, without reducing it [18].

In recent decades, with the development of building-integrated monitoring infrastructures, Information and Communication Technologies are emerging as an opportunity to improve the management of various energy systems. Advanced control strategies are being developed in this context and involve complex decision-making processes which operate the energy systems according to a given set of decision variables. In this way, multi-objective functions can be met, such as maintaining comfort conditions during periods of occupancy while minimizing energy consumption and maximizing self-consumption. It is also possible to foresee external events such as changes in energy prices or changes in weather and climate conditions. The complexity involved in the energy management of a group of buildings requires considerable effort in order to develop an advanced control activity such as the one just described. Appropriate strategies for controlling the energy systems of multiple buildings are being tested, but their practical application in the field calls for the development of an environment at the district level, which has important cost consequences. To solve this additional problem, the design and validation of a VSE allows these issues to be overcome, ensuring time management of different influencing factors such as the effect of employment, as long as this archetype is provided with a character of flexibility and adaptation to variable contexts. A VSE can consist of models, which are usually classified in the following way according to their structure [19]:

- Calibrated or “white box” simulation method;
- Empirical or “black box” method;
- “Grey box” method.

White-box models – also known as engineering or physical models – are based on a detailed description of the system and require more knowledge of its structure. They basically consist of applying the conservation equations of mass, energy and momentum, the solutions to which can be obtained with extremely detailed input information. For this reason, modelling is time consuming and requires high

computational power, so they are not suitable for analysing the operation of the system in real time. In addition to this, these models are unable to their own to account for social or environmental factors since they are entirely based on physical equations. As a matter of fact, modules based on data-driven models are often included to evaluate these factors. However, they are able to describe system behaviour without measurement data. Over time, they have been increasingly implemented in various simulation software such as TRNSYS, EnergyPlus, Modelica, IDA and Dest.

Black-box models take the form of data-driven ones and are based on measured time series data. They do not rely on complete physical knowledge of the system and do not require detailed input parameters. Moreover, they are computationally light and are able to incorporate socioeconomic and environmental factors. In addition to this, their accuracy makes them suitable for real-time monitoring because the deployment is carried out at a low computational cost. On the other hand, the need for a wide availability of historical data and the need for high-quality data to learn the relationships that bind the functional parameters of one or more systems are major disadvantages. As a matter of fact, missed data, outliers or inconsistencies can lead to the development of an unclassified model. Some black-box models lack interpretability because the relationships between input and output are so grafted, complex and non-linear that it is really difficult to interpret them physically.

Grey-box models respect the significance of the physical phenomenon under study because they involve the use of algebraic equations. For this reason, they are easier to interpret than black-box ones. A common disadvantage with black-box models is their difficulty in generalization and scalability. These models are also more computationally efficient than white-box ones because parameters are identified through a process drive by measured data.

For analysing complex building dynamics, black-box and grey-box models are preferable because they allow for fewer input parameters than white-box ones, although they are often less accurate.

1.2 State of the art

ECs have attracted considerable interest in recent years, both at the regulatory and research levels. As a matter of fact, these can provide significant

benefits from an energy point of view, as shown by the study proposed by Ceglia et al. [20]. They quantified the difference in energy flows to and from the power grid that occurs in the case of ECs compared to individual buildings. The EC examined involved two office buildings located in Naples. Each was equipped with a photovoltaic (PV) system and EV charging stations. The results that emerged from their analysis showed significant energy sharing, which was substituted to electricity grid supply. Due to the non-programmable character of renewable energy sources, storage systems are playing an increasingly important role in residential settings. Bartolini et al. [21] analysed the operation of storage units in combination with PV systems. Through two scenarios, they identified ways to manage CSC, reducing grid dependence. Further analysis involved the use of a controllable cogeneration plant to better manage energy flows.

Development of VSEs, both at individual building level and at the district level, has been a widespread topic in the literature. H. Sen et al. [22] have evolved a Nearly Zero Energy Community virtual environment in Modelica. The proposed framework overcomes the difficulties of modeling the interactions between different heterogeneous energy systems in this building type through an integration of interdependencies. In the VSE, three types of buildings (residentials, retail buildings and offices) were considered, each equipped with a PV system for a total peak power of 60 kW. Auxiliary batteries are available in each building, as well as water-to-air Heat Pumps (HPs). Domestic Hot Water (DHW) demand is met by a water tank in which the HP coil is immersed. All these energy systems were shaped using a hybrid model, based on both calibrated simulation and empirical approach. Their interaction with the power grid was looked at later. Results obtained from the application of the model to a community located in Florida showed a good ability to reconstruct net energy demand, but a poor capacity to predict PV production. At the district level, an interesting framework is CityLearn [23] based on OpenAI Gym Environment, a standardized interface used for simulations aimed at control. This VSE is used for comparison of DLR algorithms applied in DR tasks. It contains many models of energy systems such as DHW production ones, HPs, Electric Heaters (EHs), Gas Boilers (GB) and solar systems. In contrast, the heat load is pre-calculated. The same standardized interface has been used by Z. Wang et al. [24] to develop a VSE called "*AlphaBuilding ResCommunity*", exploitable for training and validation of control algorithms applied to Thermostatically Controlled Load. The authors apply a grey-box 1R1C model to

simulate the variation of indoor air temperature of buildings belonging to aggregate. Considering the difficulty of tuning algorithms at the level of individual buildings in a neighbourhood setting and the complexity of simulating energy systems based on physical approach, this study focuses on the considerable effort to scale algorithms on different systems than the one used for their validation. AlphaBuilding ResCommunity tries to overcome these issues as it is built on the basis of parameters evaluated from measured data for a vastness of buildings with different uses. However, one limitation concerns not taking into view the HVAC system, but only its on/off status.

The intermittent nature of small-scale renewable power generation technologies increasingly integrated into buildings, albeit with appropriate integrated storage systems, can place constraints on flexibility and energy self-sufficiency. To partially overcome these issues, further adoption of controllable distributed generation systems may be effective. For example, cogeneration plants generate electricity and heat in the form of hot water, superheated water or steam. At the technological level, for small-scale applications, Internal Combustion Engine, steam engines, Micro Gas Turbine (MGT), fuel cells and Stirling Engine [25] stand out. The application of such type of distributed generation systems at the residential level has become increasingly important in recent years, as both thermal and electrical loads can be met using a single system. In the case of excess generation, energy flows can be exchanged with the District Heating Network (DHN) and with the power grid, respectively. M. Pirouti et al. [26] in their study described the potential of residential-scale distributed CHP systems to produce heat and electricity, following the considerable primary energy savings achievable. Instead A. Olympios et al. [27] argued that cogeneration technologies and HP systems play a key role in achieving expected decarbonisation targets. In their analysis, they compared three different design scales involving cogeneration systems serving a district of the Isle of Dogs in London. The economic analysis and the study of the amount of climate-changing emissions allowed them to identify the optimal configuration. In [28] a strategy involving a group of buildings is introduced, based on the possibility of increasing self-sufficiency from a thermal point of view by integrating micro-cogeneration plants. The types of building suitable for the installation of these plant systems and the optimal management strategies for these, to have an impact at the district level, were examined. Instead, an application case was presented by G. Prinsloo et al. [29], who

designed a smart microgrid capable of managing the production of thermal and electrical carriers for members belonging to a rural village in South Africa through an integrated CHP system. Feedback related to heat and electricity utilization paradigms made this solution effective. On the other hand, some studies have been carried out to highlight the drawbacks related to the adoption of such systems. For example R. Rodríguez et al. [30] analysed the amount of investment costs depending on the chosen technology and looked at fundamental research to reduce these costs and achieve grid parity. Considerable efforts are being made in this direction. As a matter of fact, H. Ren et al. [31] have proposed models to reduce the energy costs of a μ-CHP installed at residential level, taking into account optimal sizing of storage systems and integration of back-up boilers.

Increased electricity generation within ECs satisfied much of the power grid requirements, while also bringing benefits to consumers. Electrification of transport is one of the key strategies for smart cities, required due to the urgency of developing decarbonization actions. In addition to this, EVs are progressively seen as systems which offer flexibility, because the intrinsic energy storage can be effectively harnessed in specific situations, avoiding withdrawal from the power grid. Proper planning of EV charging strategies must be provided for, so as to avoid impacts rather than benefits on the power grid. Numerous research works have been established on this topic, for example J. Wang et al. [32] developed four EV charging scenarios in order to shift the load from peak hours to off-peak hours. Through the simulation of such scenarios by integrating numerous vehicles, the authors demonstrated that a flexible charging strategy – involves through DR ones – can be useful in order to mitigate the characteristic valleys of the load profile during the night hours. This strategy also significantly reduces system operating costs. A. Foley et al. [33] analysed two charging scenarios by assessing a peak profile and an off-peak profile for more than 200000 EVs in the territory of Ireland. Off-peak charging is significantly more beneficial than peak ones, leading to a 10% increase in renewable energy used in transport and a significant reduction in CO_2 emissions. C. K. Ekman [34] has built a model from historical data concerning the combining of energy sector of wind energy and charge of about 500000 EVs in Denmark, to apply several strategies in order to assess the impact of charging on energy availability. The author showed that charging vehicles at night results in a reduction of excess wind energy sold to the power grid without causing changes in installed capacity or energy purchased from the power

grid. EVs bring significant benefits, such as reduced emissions related primarily to the use of renewable, non-fossil energy carriers. However, a large-scale development of these systems could burn the power grid. W. Schill and C. Gerbaulet [35] have proposed a study on the impacts that the ever-increasing deployment of EVs may lead to on the Germany's energy system. The development of a model which optimize energy flows and vehicle charging schedules allows for load profiles with less effects. Integrating constraints related to minimum load, related costs or the amount of generation from the plants allows realistic estimates of optimal charging patterns.

A VSE can be based on an inverse modelling approach, due to the wide availability of monitoring data. Data-driven methods, which can assess building energy consumption, include grey-box models and black-box ones. The former of these maintain the physical significance of the studied object and require the measured data to determine representative parameters. The latter of these, on the other hand, require both input and output data to determine the parameters of the relationship between them. K. Arendt et al. [36] conducted an analysis comparing white-box, grey-box and black-box models developed to simulate the indoor air temperature of a university building. They showed that the black-box models exhibit better performance than the other two in the case of long forecasts, provided that adequate monitoring data are available. In addition to this, it was found that grey-box models are able to predict indoor air temperature for short time horizons with high accuracy. S. Royer et al. [37] proposed a black-box model to simulate the thermal dynamics of a building located in southern France. The validated model showed a scalable nature. As a matter of fact, it revealed excellent prediction capabilities of indoor air temperature for the same building placed in three different American cities. The same model, applied on another type of building with another heating system placed in the same French city as the starting one, showed good accuracy.

Artificial Neural Networks (ANNs) are among the most widely used black-box models for forecasting the energy consumption of an individual building or a cluster. Pinto et al. [38] developed a new energy management strategy applied to a cluster of four commercial buildings, each equipped with a HP, TES and EH. Developing of a model of each building using EnergyPlus allowed them to train Long Short-Term Memory (LSTM) neural networks, which are capable of learning long-term dependencies and predicting the change in indoor air temperature. This model was then coupled with a VSE for control-related purposes. Gokhale et al. [39] proposed a

Physics-informed neural networks approach to simulate the thermal dynamics of building and integrate control actions which can reduce energy costs and keep indoor air temperature within acceptable limits. This model requires the introduction of prior knowledge of neural network structure, with the aim of avoiding problems of generalization. The authors show that training this type of model requires less data than other architectures and their application has exhibited excellent performance for long predictions. Various types of ANN can be used to simulate or predict building behaviour and, in this regard, the literature is showing increasing interest. For example, Di Natale et al. [40], in their paper "Physically Consistent Neural Networks for building thermal modelling: theory and analysis", presented a new physics-based neural network architecture which requires only past operational data as inputs to provide the thermal dynamics of individual buildings.

From the emerging results in literature, a systematic analysis of individual buildings and aggregate data is shown to be essential. In the past, the availability of energy consumption information for individual buildings – from utility bills or meter readings – has made it possible to fully understand the relationships between the amount of energy efficiency and the characteristics related to the building component. Recently, the integration of measurement systems at different scales has enabled increasingly detailed data suitable for benchmarking analysis. In this view, the comparative analysis of the energy operation of buildings in the EC is relevant because it can act as a mechanism to compare the performance of several similar systems. This can lead to informing stakeholders of any deviations from baseline performance target and, where appropriate, motivating and supporting them to carry out improvement and optimization activities. Generally, it is usual to distinguish three types of benchmarking activities applied in the energy context:

- **Internal benchmarking:** This process involves evaluating the expected energy consumption of the system against influencing factors and using it as a reference to characterize performance and/or energy efficiency. The baseline value can be identified in relation to historical or expected performance of the same system.
- **External benchmarking:** This process involves classifying one or more systems from an energy perspective, against the historical or current average

performance of its peers[7], so that any need for intervention can be characterized.

- **Engineering Model-Based benchmarking:** This process involves the construction of an archetype, a system appropriately modelled and designed through dynamic thermo-energetic simulation codes. This modelled system acts as a baseline against which to report the energy performance of the analysed system.

Adequate comparative analysis can be achieved through the application of combined energy performance indicators, also referred to as Key Performance Indicators (KPIs). These ones can be applied at different levels of aggregation, from a single device level to a single building level and then extended to a group of buildings. Several ways have been developed for this purpose, some of which are briefly described below. A. Cielo et al. [41], for example, proposed a method of sizing a solar PV system equipped with a storage system, based on an optimization process obtained through two key independent performance indices in the field of self-sufficiency and self-consumption. By varying the size of the solar PV system and the electrochemical storage system, the optimal one could be identified. Á. Corredera et al. [42] developed an automated system which, from continuously recorded information, can automatically calculate performance indicators. Internal benchmarking analysis allows reporting any anomalies in thermal energy delivery by a HP and/or a biomass boiler. The system is suitable for buildings and also includes the ability to manage on-site renewable energy production, energy consumption and occupant comfort control. This can lead to significant annual savings. Other studies [43] have suggested the exploitation of key factors for the design phase of a microgrid. These have included environmental, technical and occupant satisfaction-related performance indicators. External benchmarking has highlighted the importance of exploiting locally produced renewable energy, the ability to deal with emergency situations and the opportunity to reduce transmission losses through the power grid, as well as emissions. Another type of approach at the level of building groups was adopted by J.L. Hernández et al. [44] under which an indicator-based sustainability protocol has been designed. This also includes assessments of energy efficiency,

[7] When applied in the field of buildings, this activity consists in identifying certain properties such as the same intended use, the same dimensions, the same operating conditions.

economic, social and environmental impacts. In particular, sustainability plays an important role along with the design of Smart Cities and energy efficiency and the proposed sustainability protocol make it possible to overcome the complexities and costs associated with environmental assessment standards – such as Leadership in Energy and Environmental Design –. With a view to an ever-greater interaction of the building with the power grid, KPIs representing the energy flexibility of building systems have recently become widespread. Their proper application allows an optimization of the energy flexibility related to the building, also enabling a minimization of energy costs, according to Clauß et al. [16].

1.3 Contribution of this work

This thesis work is part of a project related to the development of a simulation platform for ECs. The existing VSE used has the main purpose of improving energy management on a district scale. In its initial configuration, this archetype provided that a well-defined fraction of buildings would be equipped with a PV system and attached BESSs to partially meet the electricity needs of the community. Temperature control of the individual buildings, all of which were assumed to be single thermal zones, was given to air-to-air HPs that in some cases could manage the charge of TES, if any. DHW production, on the other hand, was through separate generating systems, which could be an EH or a GB depending on the building under consideration. One of the main contributions involved the development of alternative models to grey-box ones, through the design and the validation of black-box models suitable to predict the change in indoor air temperature of building which might belong to an EC. This makes it possible to consider their thermal dynamics on a district scale with high accuracy and acceptable computational effort. So, in the first part of this study we will show the steps required for the development of LSTM neural network models and the results, in performance terms, obtained. Subsequently, the research aims to assess the different technological solutions available to ECs to meet the needs of their members and to manage the energy flows. To be more specific, a model is developed to simulate the operation of a micro-cogeneration plant, a system based on the combined production of thermal energy and electricity. On the other hand, the growing interest in reducing climate-changing emissions has led to a proliferation of new systems, including EVs, also capable of providing services to the power grid. Hence, a model is proposed which could reconstruct the electric power demand from any EVs belonging

to the members of the community. Three potentially realistic energy scenarios were therefore investigated. All of them are based on the same reference configuration of the VSE, corresponding to the initial set-up before the integration of the modules we developed. However, these scenarios differ in the type of energy systems involved. A first scenario provides for the integration of a MGT plant to cope with increased thermal and electrical loads of community members. A second scenario analyses in economic, environmental and especially energy terms the impacts brought about by the introduction of consumer units such as EVs as the percentage of owner-users changes. Therefore, this scenario will include more analyses. Finally, the last scenario incorporates the previously sized micro-cogeneration plant in the basic configuration of the VSE and takes on the presence of EVs as energy demand systems among community members. The data obtained from the different simulated scenarios were aggregated, historicized and used to evaluate some useful parameters for benchmarking activity. So we have compared the aforementioned KPIs with those related to the different community members. The advantage of the methodology proposed in this study is that it can be easily adapted to look at different ECs as long as measured or simulated data related to building specifications are available. The major contributions of this work can be summarized as follows:

- Training and validation of LSTM neural networks to predict the indoor air temperature of individual buildings.
- Development of first-order models for a DHN supplied by a MGT and to simulate the occurrence of EVs in the EC.
- Development of different energy scenarios for ECs which are based on current trends and performance analysis considering KPIs for individual and aggregated buildings.

In detail, the thesis consists of six parts: *chapter 2* analyses data analytics-based algorithms such as ANN. An in-depth study of the structure of an LSTM neural network architecture and an assessment of the mathematical model which describes it is carried out. In addition to this, the same section describes the mathematical models related to the operational state of the micro-cogeneration plant and to the demand for electricity to recharge EVs owned by community members. Attention is also given to Rules-Based Control (RBC) policies, appropriately chosen to burden the

power grid as flexibly as possible. *Chapter 3* describes the structure of the VSE used in this research and the assumptions on which it is formulated. Next, it focuses on an analysis of the parameters to be provided as input to initialize the simulations, a description of the thermal and electrical demand models and the various energy systems and a presentation of the control strategies adopted. In addition to this, the different energy scenarios investigated will be mentioned. *Chapter 4* is devoted to our case study, while *chapter 5* deals with the presentation of the results obtained under the different scenarios, an evaluation of them and other critical insights. Finally, *chapter 6* focuses on conclusions describing potential future developments and possible improvements.

2 Methodology

This chapter will be devoted to Artificial Neural Networks (ANNs) and the theoretical background on which the Long Short-Term Memory (LSTM) network architecture is built. Then, an analysis of the hyperparameters and the main error metrics used to obtain detailed training information of an ANN will be carried out. Next, the focus will be on describing models related to energy systems. To be more specific power generation technologies supplied by renewable energy sources, such as solar PV or wind power, are marked by a random character. In addition to this, stochastic behavioural patterns of occupants, variations in the price of energy purchased on the power grid and intrinsic constraints on power distribution systems make complex the problem of determining appropriate strategies to meet multi-objective functions. To partially overcome these issues, further integration of controllable generation and consumption systems may be effective. In this direction, we will introduce a model of a micro-cogeneration system serving the entire EC and a model of electricity consumption for charging EVs. The parameters and assumptions underlying the formulation of such models will be also analysed in detail.

2.1 Fundamentals of Artificial Neural Networks

ANNs are advanced machine learning algorithms, known for predictive analytics. They are qualifying as computational models based on mathematical entities – also called *artificial neurons* – which are made up by several layers of cells interconnected with each other. Within machine learning, ANNs are ranked as data-driven models based upon supervised learning techniques, where input and output are known from a set of data. The key aim in the training phase is to find the value of those parameters which provide the relationship between input and output data.

Such connection can then be used with input data, where the boundary conditions are similar to those of the training phase, in order to predict output values. ANNs are broken down into two broad categories:

- **Feed-forward** [45]**:** it is a structure within which neural connections are not based on loops, so information is moved only in the direction from input to output.
- **Recurrent Neural Network (RNN)** [46]**:** it is a structure within which connections between neurons may or may not be directed. Being loop-based, the advantage of the RNN architecture is the ability to leverage a memory layer which provides time information to an input layer. This makes RNNs highly enforced to predictive modelling of data sequences.

LSTM neural networks belong to the category of RNNs. They are configured as networks with *infinite impulse response*, in which there are states of storage – also called *gated states* – able of containing data derived from feedback trials. LSTM neural networks are architectures first introduced by S. Hochreiter and J. Schmidhuber [47] in 1997, primarily developed to remove long-term dependency problems. As a matter of fact, these models are based on two sequential phases: training and testing. During the training phase, the network learns the relationship that binds inputs and outputs. When this phase is over, the model has learned from the inputs to estimate the outputs. Therefore, during the testing phase, its accuracy at predicting the output is proven. Usually the first phase of training takes place through the gradient descent method, which consists in determining the local minimum of the error function in an N-dimensional space. However, in traditional schemes, the error gradient disappears as it propagates over time. This does not contribute to a good learning phase. The parameters in the first layers of the network are subject to a small gradient update and the effect of this phenomenon is that longer sequences are ruled by short-term memory. A LSTM neural network is organized in concatenated layers, each of which consists of four levels that interact with each other according to a certain logic, unlike traditional neural networks which involves only one level per layer. The main component of a LSTM neural network layer is the cell state which is represented in Figure 2.1 by the horizontal line placed in the upper portion of the diagram.

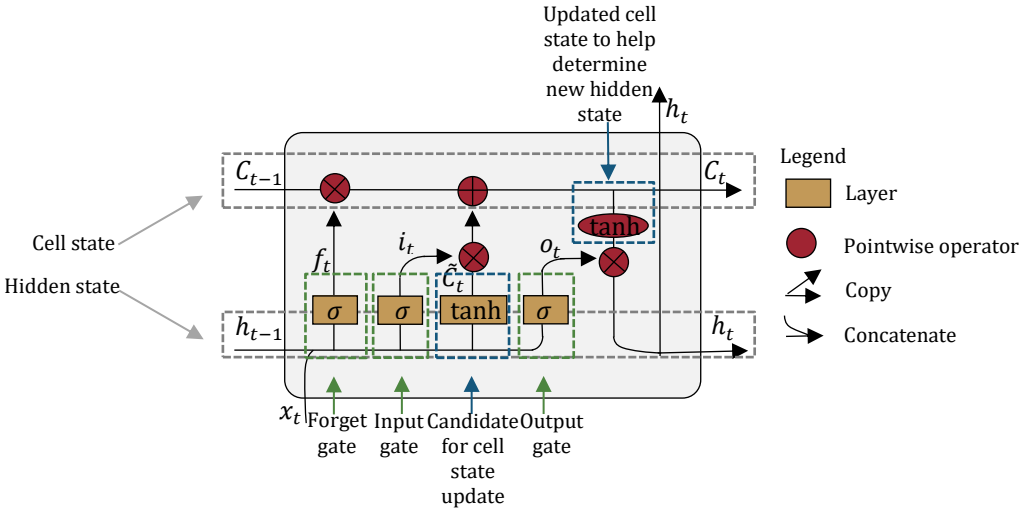


Figure 2.1: Structure of a layer belonging to an LSTM neural network for time series forecasting.

This runs within the entire module, can have linear interactions with other components and has the function of moving data vectors – containing information – from the previous layer to the next one. Is regarded as the level that is given the task of keeping long-term memory and detecting long-term relationships. It is possible to provide or subtract information to the cell state through structures called *gates*. A gate is a means of moving information and consists of a sigmoid layer and a pointwise multiplication operator. The function built into a gate layer can vary between zero and one and is used to determine the fraction of information to be retained when crossing the gate. A value of zero means that no information will be transferred through the gate, while a value of one reveal that all information will be moved. Conventionally, a LSTM neural network consists of three gates, which are designed to control and protect the cell state. Gates can therefore be seen as different neural networks which learn what is important information to keep or forget during training. All gates stick to the hidden state, which is a state that is accountable for preserving short-term memory.

The mathematical model provides a step, within each LSTM layer, which is able to choose which information is relevant and which is negligible to the cell state. The *forget gate layer* makes this judgment. More in detail, it analyses the vector obtained from the combination of the previous hidden state information h_{t-1} and the information of the new current input x_t. At the output of the sigmoid function, a worth anywhere between zero and one is given for each number contained in the cell state C_{t-1}. A value of zero means to delete that information, while a value of one indicates

that the information must be fully retained. This process is expressed by the following equation [47], [48]:

$$f_t = \sigma_g(W_f x_t + U_f h_{t-1} + b_f) \tag{2.1}$$

in which $f_t \in (0,1)^h$ is the activation vector of the forget state, σ_g represents the sigmoid function, $W \in R^{h \times d}$ accounts a matrix of recurring weights, $U \in R^{h \times d}$ provides a matrix of weights to be supplied in input and $b \in R^h$ is a vector of bias parameters. Finally, $x_t \in R^d$ and $h_t \in (-1,1)^h$ are respectively an input vector to the LSTM layer and an hidden layer vector, even said LSTM output vector[8].

The second step involves to identifying the information to be stored within the cell state. This process is based on two phases: the first one is that the *input gate layer* – which is a sigmoid layer – defines the data to be updated. Next, a *tanh layer* calculates a vector of new data C_t which should be added to the cell state. The results of both of these procedures are combined and provided to the cell state. The two phases are described in the following equations:

$$i_t = \sigma_g(W_i x_t + U_i h_{t-1} + b_i) \tag{2.2}$$

$$\tilde{C}_t = \sigma_c(W_c x_t + U_c h_{t-1} + b_c) \tag{2.3}$$

in which $i_t \in (0,1)^h$ is the activation vector of the input gate, $\tilde{C}_t \in (-1,1)^h$ marks an activation vector of the input state, while σ_c is the hyperbolic tangent function.

The cell state C_{t-1} is then upgraded to a new cell state C_t. To do this, multiply the cell state C_{t-1} for the output of the forget gate layer f_t and add the product of the sigmoid output i_t with the new cell state C_t, as we can see from the Equation 2.4.

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t \tag{2.4}$$

where $C_t \in R^h$ is a vector, as well as C_{t-1}. The outcome depends on the cell state but does not match C_t. In particular, to determine the output, an additional sigmoid layer must be used in order to find out which information within the cell state must be given. Therefore, the cell state data is subjected to a tanh operation, so as to get results

[8] The *h* and *d* superscripts stand for the number of input features and the number of hidden layers, respectively.

in the range of -1 to 1, and the outcome achieved is multiplied by the output of the sigmoid layer mentioned above. All this is defined by the equations below.

$$o_t = \sigma_g(W_o x_t + U_o h_{t-1} + b_o) \tag{2.5}$$

$$h_t = o_t \cdot \sigma_h(C_t) \tag{2.6}$$

where $o_t \in (0,1)^h$ is the activation vector of the output gate and σ_h can be a hyperbolic function or a linear function of the cell C_t. The hidden state allows us to obtain the expected result since information from previous inputs is available within it and is used for forecasting activities.

2.1.1 Hyperparameters and error metrics

The design of an optimal LSTM model requires a careful selection of hyperparameters. Among the most important ones, which need to be considered for an LSTM-based model, are:

- **Batch size:** part of the training dataset which contains a specific number of records.
- **Epochs:** number of times the neural network completely examines the training dataset. Since learning is based on an iterative process, it is necessary to provide the neural network several times with the training dataset to update the weights of all neurons.
- **Learning rate:** rate at which a neural network must leave out past information acquired and replace it with others. The higher this hyperparameter, the more the network could change its mind quickly.
- **Number of hidden layers:** levels of neurons to be inserted between the input layer and the output layer. As the number of hidden layers increases, the computational cost grows.
- **Number of hidden units:** neurons belonging to each hidden layer. It is advisable to set a number of neurons between those of the input layer and those of the output layer.

LSTM neural networks can be used for both regressive[9] and classification[10] problems. If we focus on the regressive ones, the supervised model is built by splitting the dataset into three parts: training set, validation set and testing set. The highest fraction of the data available is used for the training phase, while small portions of datasets are used for the validation and testing process. During the training phase, the model tries to find the minimum of a cost function. This is usually based on Mean Square Error (MSE), expressed as follows:

$$\mathrm{MSE} = \frac{1}{\mathrm{N}} \sum_{\mathrm{i}=1}^{\mathrm{N}} |\mathrm{y_i} - \hat{\mathrm{y}}_\mathrm{i}|^2 \tag{2.7}$$

where y_i is real value while $\hat{y}_i$ is the predicted value. The MSE is a statistical indicator very sensitive to any outliers in the dataset[11]. One of the goals of the LSTM neural networks is to adjust the weights of each layer so as to minimize the MSE.

Once the training is completed, the testing dataset is used to evaluate the performance of the model. Specifically, input data from the testing dataset are provided to the neural network, which will estimate one or more output values. These are then compared with the actual output values, from which the MSE and also other error metrics, can be determined:

- **Mean Absolute Error (MAE):** it is the average value of the difference between the predicted output and the actual output

$$\mathrm{MAE} = \frac{1}{\mathrm{N}} \sum_{\mathrm{i}=1}^{\mathrm{N}} |\mathrm{y_i} - \hat{\mathrm{y}}_\mathrm{i}| \tag{2.8}$$

  An advantage of this metric is that it is not much affected by the presence of outliers. A disadvantage is related to the fact that it retains the same unit of measurement as the variable to be predicted. This is not very suitable when a comparison between different models needs to be made.
- **Mean Absolute Percentage Error (MAPE):** it is the average of the sum of all relative errors

[9] Regressive problems predict numerical output.
[10] Classification problems model the belonging of one or more records to a specific class, so they predict a categorical output.
[11] For this reason, before the *Knowledge discovery* phase, a *data pre-processing* and *data-segmentation* phase is developed.

$$\mathrm{MAPE} = \frac{1}{N}\sum_{i=1}^{N}\frac{|y_i - \hat{y}_i|}{y_I} \tag{2.9}$$

The advantage of this error metric is that it is a percentage value, useful for developing model comparisons. A drawback is related to its mathematical formulation, as if the actual output is zero the MAPE goes to infinity[12].

- **Root Mean Square Error (RMSE):** it is the square root of MSE

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}|y_i - \hat{y}_i|^2} \tag{2.10}$$

- **Coefficient of Variation of RMSE (CV-RMSE):** it is determined by dividing the RMSE with the mean of the variable to be predicted

$$\mathrm{CV-RMSE} = \frac{\mathrm{RMSE}}{\bar{y}} \tag{2.11}$$

The advantage of this metric is that it is a dimensionless value.

It is important to avoid the overfitting phenomenon, with respect to which the model shows excellent performance in training, but bad performance in testing. Finally, one of the limitations of these and other ANNs concerns the difficulty of understanding whether training has actually effective [49].

2.2 Micro-cogeneration system

Buildings demand for electricity and thermal energy is generally met through separate generation and supply. In most RECs, electricity is generated by PV or wind systems arranged on site, while thermal energy comes from boilers or systems which use electricity directly at the end user. Electricity generation from PV systems peaks during the daylight hours of the day is often higher than user demand, but it requires a power grid offtake during the rest of the day. To address these issues, an alternative model to that traditionally used involves μ-CHP plants, highly controllable units able to combine electricity production with thermal energy production, starting from the conversion of the same primary energy source. Combined production allows significant fuel conversion efficiency, resulting in lower fuel supply costs and greenhouse gas emissions compared to separate production. Typically, a μ-CHP plant

[12] This is a frequent condition when applied to energy systems.

has a maximum generation capacity of less than 200 kWe, so it is a small power plant which may or may not be connected to the power distribution grid. In these applications, the thermal energy produced is often higher than the electrical energy and this aspect plays a role of particular importance since a large part of the buildings consumption is attributable to domestic heating, carried out through DHNs or to domestic cooling through absorption systems. μ-CHP systems are averagely deployed at the residential level due to significant daily load fluctuation, costs required to expand the DHNs and difficulties in obtaining authorizations. Despite this, one among the innovative μ-CHP technologies used for domestic purposes is the MGT. A MGT is a small power unit with a compressor, a turbine and an alternator set on the same shaft. Combustion air enters a radial compressor operating at a very high speed, after passing through a filter, so as to increase its pressure and temperature. Subsequently, in a combustion chamber the air is mixed with the fuel and combustion occurs. The flue gases resulting from this process pass through a centripetal turbine, inside which the gases expand. The mechanical work is then converted into electrical energy by means of a generator mounted on the turbine shaft, without the need for deceleration components. The work cycle is almost always regenerative, which is, the heat recovery involves the high-temperature flue gas leaving the turbine, that is used to preheat the air entering the combustion chamber through an air-to-air heat recovery unit and then to exchange heat with the water supplying the thermal utilities by passing through an additional heat exchanger. The advantage of these plant systems over other systems of similar size lies in the limitation of noise and vibrations and also in the absence of an auxiliary heat-disposal system. To this, they do not require excessive maintenance and have a high service life. The disadvantage of these systems lies mainly in the costs.

At the regulatory level, European Directive 2004/8/EC [50] implemented through Legislative Decree 8 February 2007, No. 20 defines that a μ-CHP system can be considered high efficiency if the Primary Energy Saving index takes a positive value. This factor is driven by the ratio of the combined production of a μ-CHP plant versus separate production through standard technologies. In addition to this, it is also necessary to meet the minimum threshold for overall efficiency of the cogeneration (equal to 75% for the MGT). In this regard, it is the Energy Services Manager which recognizes compliance with the technical conditions of high-efficiency cogeneration, the benefits of which are many. These include priority

dispatch of the electricity produced, tax breaks regarding the excise duty of the methane gas used, right to obtain energy efficiency certificates, right to access to the mechanism of *Exchange on the Spot* of electricity produced and exemption from payment of general system charges.

The integration of a MGT, able to meet the electrical and thermal demand of users belonging to the EC considering its interaction with renewable energy systems, electrical and thermal storage could be useful to increase self-sufficiency and independence from the electricity grid. The *simulation parameters* used to derive the μ-CHP model in the VSE are as follows:

- **Electricity price for CHP:** valorisation of electricity fed into the power grid for dedicated Energy Services Manager withdrawal or for sale to the market in $\frac{\$}{\text{kWh}}$.
- **Gas price for CHP:** price of natural gas used as a primary energy source by CHP in $\frac{\$}{\text{kWh}}$.

Instead, the *parameters* of the *energy system* defined in input are:

- **Electrical efficiency $\boldsymbol{\eta_{el,CHP}}$**: electrical efficiency of the MGT, determined in relation to the outdoor temperature.
- **Thermal efficiency $\boldsymbol{\eta_{th,CHP}}$**: thermal efficiency of the MGT, determined in relation to the outdoor temperature.
- **Global efficiency $\boldsymbol{\eta_{g,CHP}}$**: global efficiency of the MGT, determined in relation to the outdoor temperature.
- **Heat exchange efficiency:** efficiency associated with the heat exchanger which separates the DHN with the user.
- **Distribution efficiency:** efficiency associated with heat dissipations occurring in the distribution network.
- **Maximum thermal power:** maximum heat exchangeable power for each building.

The sizing of a MGT to serve the entire EC comes from the analysis of the electrical load profile of the aggregate of buildings belonging to the same community. This is determined through the following equation:

$$E_{f,CHP}^{nominal} = \frac{\alpha \cdot E_{el}^{peak}}{\eta_{el,CHP}^{nominal}} \quad (2.12)$$

where α is a dimensionless coefficient chosen based on the duration curve of the electrical load, $E_{f,CHP}^{nominal}$ represents the energy supplied by the fuel, E_{el}^{peak} represents the electrical absorption under peak conditions and finally $\eta_{el,CHP}^{nominal}$ represents the electrical efficiency, under nominal operating conditions, of the μ-CHP. The mathematical model is shown below:

$$E_{el,CHP} = E_{f,CHP} \cdot \eta_{el,CHP} \quad (2.13)$$

$$E_{th,CHP} = E_{f,CHP} \cdot \eta_{th,CHP} \quad (2.14)$$

in which $E_{el,CHP}$ and $E_{th,CHP}$ represent the electrical and thermal energy respectively produced by the plant and dependent on the energy supplied by the fuel $E_{f,CHP}$. $\eta_{el,CHP}$ and $\eta_{th,CHP}$ represent the electrical efficiency and thermal efficiency of the system, respectively. These two parameters are closely linked to each other, since through the application of the first law of thermodynamics it is possible to determine an overall efficiency of the cogeneration plant $\eta_{g,CHP}$ (also called *Utilization Index*):

$$\eta_{g,CHP} = \frac{E_{el,CHP} + E_{th,CHP}}{E_{f,CHP}} \quad (2.15)$$

In initializing the input parameters, however, it must be considered that the electrical and thermal conversion efficiencies of a MGT are not held constant, but depend on the outdoor temperature. As a matter of fact, its increase results in a reduction of both the electrical power and the electrical efficiency of the system, because of the change in air density. To take account of this phenomenon, the variation in electrical efficiency with temperature was considered according to the relationship provided by Y. Hwang in his study [51], the trend of which is shown in Figure 2.2. While the electrical efficiency decreases with increasing temperature, the thermal efficiency rises. It was also assumed that the power grid losses linked to the electric carrier are zero, since centralized power generation from large thermal power plants and subsequent transmission and distribution is absent. On the other

hand, the thermal energy delivered to the utility differs from the thermal energy produced by the micro-cogeneration plant, since a distribution efficiency due to thermal dissipation which may occur during transport must be considered.

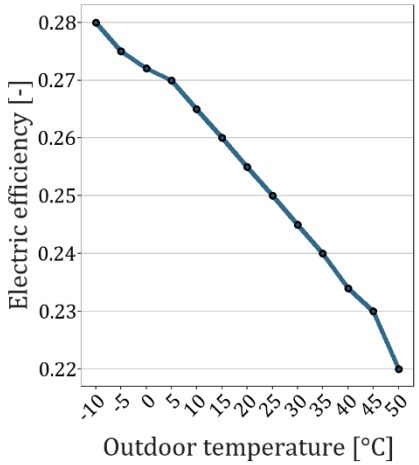


Figure 2.2: Trend of the electrical efficiency of the MGT as the outdoor temperature changes.

A heat exchange efficiency related to the heat exchangers integrated in each building and interfaced with the DHN must also be taken into account. In more detail, most small DHN models applied to ECs involve considering the layout, but in this one the goal is precisely to descend from this characteristic. So it is planned to implement a heat exchanger model for each building and then through efficiency parameters it is possible to identify the amount of thermal energy reaching the end user.

2.2.1 Baseline control strategies

In this work, we assumed a MGT connected to the electrical distribution grid to decrease the net electricity demand of the EC and thus serve a DHN to which all buildings are connected. This energy system operates only during the heating period, otherwise much of the thermal energy produced would be completely dissipated. An analysis of the optimal synergies between the power grid and the EC identified the time periods with respect to which it is convenient to use this facility, in order to meet the demand compared to the purchase from the power grid. More in depth, the study of D. Li et al. [52] was taken as a reference because it incorporates a control algorithm which identifies a morning (from 6:30 to 9:00) and afternoon (from 16:30 to 22:00) system turn-on as optimal. In addition to this, its operation during the night period is

not assumed since the electricity demand of the utilities is low and it is cheaper to draw from the power grid, by virtue of the reduced purchase price. In this way, the annual operating hours considered are therefore 1332 and the MGT turns on and keeps the load constant when the production of energy from the PV system is not sufficient to meet the load of the community. Moreover, an additional assumption considered is that the μ-CHP system does not modulate its output power, but it works under nominal conditions during the entire ignition period. Specifically, the thermal energy produced by the system is made available to the various buildings through a DHN, which totally or partially replaces the action of the HP and TES, and in the case of excess is dispersed into the environment. Electricity, on the other hand, is fed into the power grid, so that buildings belonging to the EC can purchase it, paying for the amount consumed on their bill. However, the advantage of this mechanism is related to the fact that at the regulatory level there is an enhancement of shared energy, which is the share of energy produced and consumed on site.

2.3 Electric vehicles

EVs are gaining more and more interest because of their low emissions advantageous for achieving the goal of decarbonisation in the transport sector as well. Charging an EV can be done in three different modes: *conductive charging*, which involves the use of a cable, a plug and a connector, *inductive charging*, which is done through a magnetic field that allows the transfer of energy between a fixed underground winding and a winding allocated inside the vehicle, or *battery swap*, which consists of a change of battery in special stations. Conductive charging is the one conventionally used and can be either slow or fast. Typically, slow charging involves a power grid connection through a connector and can be in alternating or single-phase or three-phase current. Based on the battery capacity of the vehicles, such recharging can take more than 6-8 hours. Fast charging is the one that is being developed more and more and takes from 15 to 30 minutes. Since the energy which is stored in the battery is in DC, it is needed to have a battery charger to do the conversion work from the AC coming from the power grid. Conductive charging is done through special charging stations, which often have an impact on the electrical power engaged by the user. Given the ever-increasing penetration of such technologies to achieve emission reductions in transportation, an adequate knowledge of the electricity demand profile related to vehicle charging can be useful

to better manage any simultaneous loads. For this reason, it is critical to set *simulation parameters* required to reproduce the electrical absorption profile of these consumption units:

- **μ_s gaussian distribution:** mean of the gaussian distribution related to the charging start time.
- **σ_s gaussian distribution:** standard deviation of the gaussian distribution related to the charging start times.
- **μ_{SoC} gaussian distribution:** mean of the gaussian distribution related to the State of Charge (SoC) at the instant of arrival.
- **σ_{SoC} gaussian distribution:** standard deviation of the gaussian distribution related to the SoC at the instant of arrival.

The power committed for EVs charging depends on a series of information, which must be provided as inputs in terms of *energy system parameters*:

- **EVs capacity $C_{ev_{BESS}}$:** BESS capacity of the single EV.
- **Number of EVs:** fraction of EVs compared to the total number of buildings belonging to the EC.
- **EV charging efficiency $\eta_{EV_{charge}}$:** battery charging efficiency of the EV.
- **EV Depth of Discharge DoD:** EV BESS discharge depth.
- **EV BESS Round-trip efficiency $\eta_{rte_{BESS}}$:** round-trip efficiency for EV BESS.
- **EV BESS max charge/discharge power:** maximum charging/discharge power of the BESS of the EV.
- **EV BESS converter efficiency η_{BESS}:** efficiency of the AC/DC converter connected to the BESS of the EV.

The model used to evaluate electric consumption profiles for charging vehicles was set by D. Zou et al. [53]. A moment of charging begins concentrated in the late afternoon, after users return from work, and described by a Gaussian distribution defined at times according to the following expressions:

$$f_s(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(x+24-\mu_s)^2}{2\sigma_s^2}\right) \qquad 0 \leq x \leq \mu_s - 12 \qquad (2.16)$$

$$f_s(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(x-\mu_s)^2}{2\sigma_s^2}\right) \qquad \mu_s - 12 \leq x \leq 24 \qquad (2.17)$$

where x represents the charging start time, while μ_s and σ_s are the mean and standard deviation, respectively. It was also assumed that the time for EVs to stay at the charging station is about 14 hours (i.e. all night), so the end of charging assumes a Gaussian distribution which has this deviation from the start time of charging.

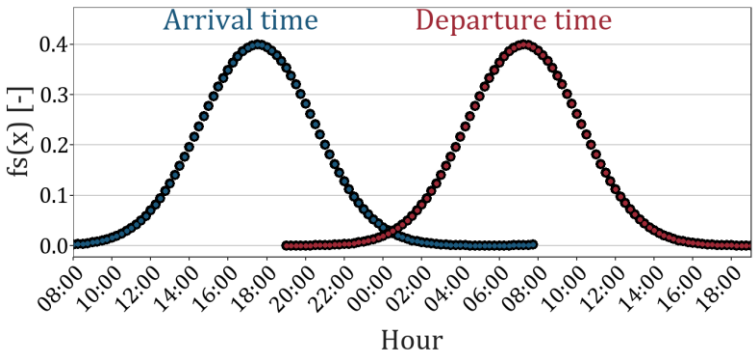


Figure 2.3: Gaussian distributions of the arrival instant (in blue) and the departure instant (in red).

The SoC of EVs at the instant of arrival and start of charging is also determined from a Gaussian distribution, according to the following equation:

$$f_{SoC}(x) = \frac{1}{\sigma_{SoC}\sqrt{2\pi}} \exp\left[-\frac{(x-\mu_{SoC})^2}{2\sigma_{SoC}^2}\right] \qquad (2.18)$$

here x represents the remaining charge, while μ_{SoC} and σ_{SoC} are the mean and standard deviation of the distribution, respectively.

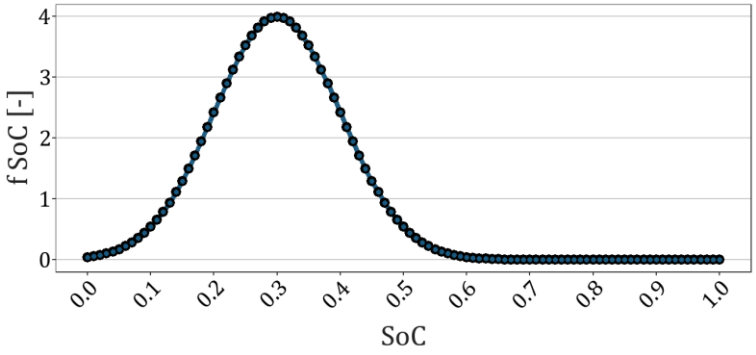


Figure 2.4: Gaussian distribution of the SoC of the EV at arrival instant.

Energy consumption related to charging EVs belonging to community members is determined as follows:

$$\Delta_{SoC} = 1 - f_{SoC} \tag{2.19}$$

$$E_{el,EV} = \Delta_{SoC} \cdot \frac{C_{ev_{BESS}}}{\eta_{EV_{charge}}} \tag{2.20}$$

From the EVs charging profile and battery capacity, the electric load of each building can be redefined with respect to the base configuration, according to the following expression:

$$E_{\text{electricity consumption}_{building}} = E_{el,HVAC} + E_{el,DHW} + E_{el,appliances} + E_{el,EV} \tag{2.21}$$

2.3.1 Baseline control strategies

The EV model was completed by implementing an appropriate RBC strategy. Specifically, a gradual recharging is planned throughout the time the EV is at the charging station, with the aim of reaching a maximum SoC at the instant the vehicle leaves the charging station.

3 Virtual Simulation Environment

This chapter reviews the structure of the VSE, the modules which is composed by and the parameters required to initialize a simulation. Next, we will focus on the mathematical models used both to define the thermal load, the electrical load and the gas load in residential buildings.

3.1 Description of the Virtual Simulation Environment

Recently, various advanced control strategies are applied to building energy systems with the aim of coordinating them optimally. Several research activities are taking place in this field and require the construction of standardized simulation environments, so as to train controllers according to different approaches and to compare different algorithms.

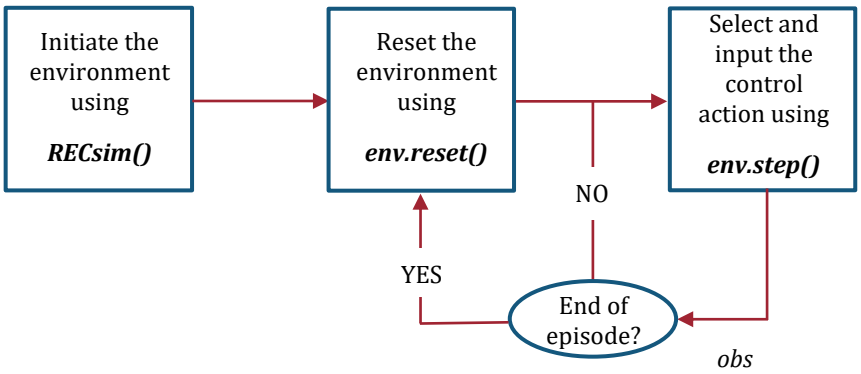


Figure 3.1: Structure of the VSE.

The VSE used for this study is fully implemented in Python and it employs OpenAI Gym, a usable interface for comparing various control algorithms. The structure of the VSE is as follows: initialization is performed by invoking a *RECsim()* Python class and providing input parameters. At the beginning of a new episode, the *env.reset()* function is called and then the controller deploys the control actions in the environment through the function *env.step()*. The virtual environment performs a simulation by providing a tuple consisting of the *s* state, the *r* reward, the *d* end-of-episode information and a set of comments *i* which the environment wishes to provide to the controller through the *env.obs()* function. Based on this data, the controller performs new control actions for the next step and it provides them to the simulation environment.

Several parameters must be defined at the beginning of each simulation, which have been grouped into three categories: *simulation parameters*, *building parameters* and *energy system parameters*. Specifically, the *simulation parameters* used to derive the model are the following ones:

- **Sample size:** number of buildings to be simulated.
- **Step size:** time step size in minutes.
- **Simulation horizon:** tuple which holds the start and the end dates of the simulation time interval.
- **Location:** pvlib item which concern the location.
- **Weather data:** CSV file with five climate variables inside. These relate to global horizontal radiation, direct normal irradiation, direct horizontal irradiation, wind speed and outdoor air temperature.
- **HVAC mode:** operational status of the HVAC system which can be “heating” or “cooling”.
- **Model noise μ:** standard deviation of model noise, due to a variety factors not explicitly modelled such as occupant behaviour or related to model approximations.
- **Measurement error:** standard deviation of the measurement error defined in °C, linked to the possible poor accuracy of the measuring sensor.
- **Cost weight:** JSON file which concern the weights for each reward term.
- **Pricing scheme:** JSON file which concern different Time of Use tariff configuration.

- **Gas price:** purchase price of natural gas.
- **GWP_{coal}:** amount in kg of CO_2 emitted per unit of electricity produced from coal.
- **$GWP_{natural\ gas}$:** amount in kg of CO_2 emitted per unit of electricity produced from natural gas.
- **$GWP_{petroleum}$:** amount in kg of CO_2 emitted per unit of electricity produced from petroleum.

To characterize the thermal loads of each building, it is necessary to initialize specific parameters, connoted as *building parameters* and described in detail below:

- **Thermal time constant τ:** tuple containing mean and standard deviation of the thermal time constant.
- **Equivalent heat gain temperature $T_{HG,eq}$:** tuple which contains mean and standard deviation of the equivalent heat gain temperature, that includes solar heat gains and internal heat gains. It is configured as a variable dependent on the time step.
- **R/C ratio:** ratio of thermal resistance R to thermal capacity C.
- **Heat gain ratio HG_{ratio}:** ratio between internal heat gain and solar heat gain. This variable is included to divide the two types of heat gains and depends on the thermophysical characteristics of the opaque and transparent envelope, the transparent opaque ratio and varies depending on the building.
- **Internal heat gain method:** method which can be "DOE" or "ECOBEE", within which the occupation, plug load and lighting schedules are specified.
- **Floor area:** tuple which holds the average value and standard deviation of floor area of the buildings.
- **Temperature set-point T_{sp}:** tuple which holds the average value and standard deviation of the internal set-point temperature
- **Temperature range T_{range}:** tuple which holds the average value and standard deviation of the acceptable indoor air temperature range from the set-point temperature. The lower level of acceptability is $T_{sp} - \frac{T_{range}}{2}$, while the upper level of acceptability is defined ad $T_{sp} + \frac{T_{range}}{2}$.

In order to simulate *energy system*, appropriate relevant *parameters* must be defined to characterize their operation:

- **Supply heating $T_{supply\ heating}$ and cooling temperature $T_{supply\ cooling}$:** tuple which holds the average value and standard deviation of the heating and cooling temperature for thermal generation plants.
- **Design heating $T_{design\ heating}$ and cooling temperature $T_{design\ cooling}$:** design outdoor air temperatures during the heating season and the cooling period, respectively.
- **PV penetration:** percentage of buildings belonging to the EC provided with PV system.
- **PV module:** PV module technical specifications taken from the SANDIA database.
- **Inverter:** inverter technical specifications taken from the SAPM database.
- **Depth of Discharge DoD:** Depth of Discharge of BESS.
- **BESS round-trip efficiency $\eta_{rte,\ BESS}$:** round-trip efficiency of BESS.
- **BESS max charge/discharge power:** maximum charging/discharging power of the BESS. It is expressed as a function of the capacity of the BESS and when a value of 1 indicates that the BESS is loaded/unloaded in an hour.
- **BESS converter efficiency η_{BESS}:** efficiency of the DC/DC drive connected to the BESS.
- **HTES/CTES penetration:** percentage of buildings belonging to the EC provided with HTES/CTES.
- **Hot/Cold water storage sizing factor:** coefficient for sizing HTES/CTES as a function of HVAC heat/cool capacity.
- **HTES/CTES round-trip efficiency $\eta_{rte,TESS}$:** round-trip efficiency of HTES/CTES.
- **HTES/CTES loss coefficient:** parameter between 0 and 1 which defines HTES/CTES losses.
- **Gas fired/electrical heater ratio:** ratio between the number of buildings with GB and the number of buildings with EH.
- **Gas fired boiler efficiency η_{gas}:** GB efficiency.
- **Electric water heater efficiency η_{el}:** EH efficiency.

- **Daily DHW demand:** tuple which holds shape and magnitude of the distribution of average annual demand of DHW in litres per building.
- **DHW volume per draw:** tuple which holds the average value and standard deviation of the volume of DHW taken at each event expressed in litres/min.
- **Appliances consumption:** tuple which holds the average value and standard deviation of the household appliances daily consumption.

3.2 Building energy demand

This work uses the Ecobee Donate Your Data (DYD) dataset [54] from the Ecobee DYD program. The experimental campaign was promoted by the home automation company Ecobee, which manufactures several technologies, including smart thermostats [55]. Various users can take part in the program by sharing data anonymously through smart thermostats installed at their buildings, in order to support research activities for a sustainable future. The monitoring campaign, which is still active today, includes information from more than 100000 buildings located mainly in North America.

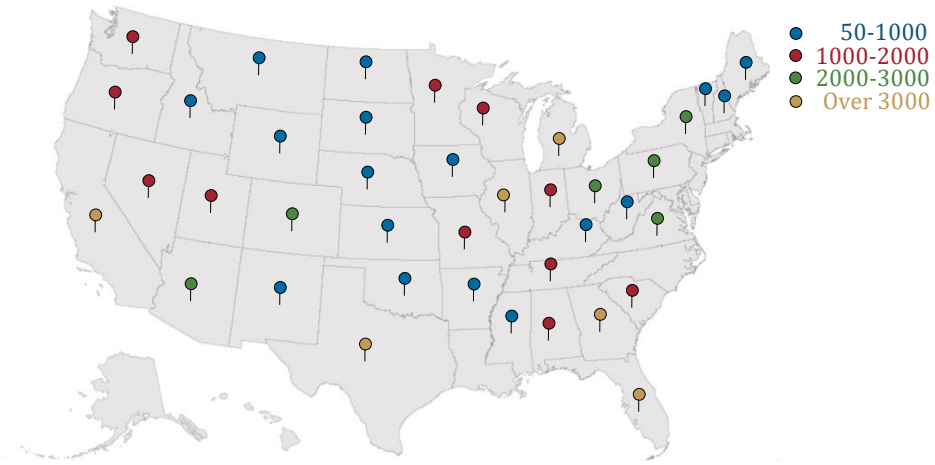


Figure 3.2: Number of buildings by location belonging to the DYD Ecobee dataset.

From the dataset containing the time series of some parameters for each building taking part in the DYD program, Z. Wang et al. [24] constructed a reduced-model 1R1C (or grey-box) of the thermal properties of each building. In detail, their analysis is applied to each building made it possible to evaluate the constant of thermal time τ_i and an equivalent temperature related to the internal and solar heat gains $T_{HG,eq,i}$, from which were defined, for the i-th building, the thermal resistance R_i

and the thermal capacity C_i. Therefore, frequency distributions of these two parameters are constructed for buildings belonging to the same state, from which the mean and standard deviation are derived. These data are used to describe the thermal dynamics of buildings in a simplified way. On the other hand, the electrical load of each building can be determined from three components using the following mathematical expression:

$$E_{\text{electricity consumption,building}} = E_{\text{el,HVAC}} + E_{\text{el,DHW}} + E_{\text{el,appliances}} \quad (3.1)$$

where $E_{\text{el,HVAC}}$ represents the electrical consumption of the HP or HVAC system and is configured as a thermal sensitive electrical load, while $E_{\text{el,DHW}}$ and $E_{\text{el,appliances}}$ refer to the electrical energy consumption for DHW production by EH and household appliances, respectively, and both are set up as non-thermal sensitive electrical load.

3.2.1 Building thermal demand

For each building – taken as a single thermal zone – the indoor air temperature T_{in} describes thermal comfort conditions. It is a reference point for allowing the thermal mass of the building, to be used as a flexible lever for managing integrated plant systems. The following equation represents the thermal model:

$$T_{\text{in},t+1} = e^{-\frac{\Delta t}{\tau}} \cdot T_{\text{in},t} + \left(1 - e^{-\frac{\Delta t}{\tau}}\right)\left(T_{\text{out}} + HG_{\text{int}} + HG_{\text{sol}} + R \cdot \mu + R \cdot E_{\text{th,load}}\right) + \epsilon \quad (3.2)$$

where HG_{int} and HG_{sol} are the equivalent temperatures associated with internal heat gains and solar heat gains, respectively. R is the thermal resistance determined by the thermal time constant τ and the parameter RC_{ratio}:

$$R = \sqrt{\tau \cdot RC_{\text{ratio}}} \quad (3.3)$$

Therefore, the thermal capacity of buildings is worth:

$$C = \frac{RC_{\text{ratio}}}{R} \quad (3.4)$$

The HG_{int} is established through the input building parameters as follow:

$$HG_{\text{int}} = T_{\text{HG,eq}} \cdot HG_{\text{ratio}} \cdot HG_{\text{int,sched}} \quad (3.5)$$

where $HG_{int,sched}$ represents an hourly schedule which includes occupancy, lighting and other miscellaneous loads.

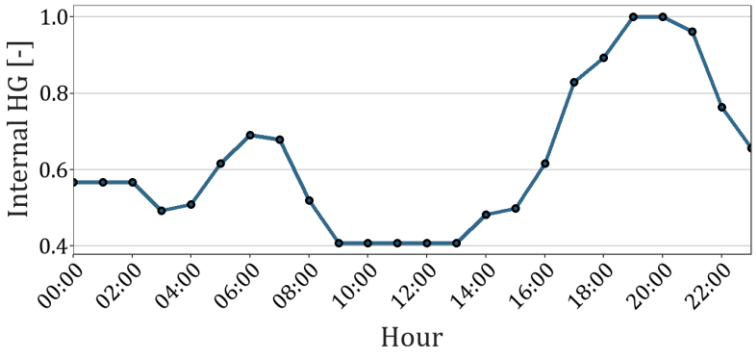


Figure 3.3: Internal heat gain schedule for a typical day.

On the other hand, HG_{sol} is expressed as:

$$HG_{sol} = T_{HG,eq} \cdot (1 - HG_{ratio}) \cdot HG_{sol,sched} \tag{3.6}$$

where $HG_{sol,sched}$ is the schedule set by the United States Department of Energy for the reference building.

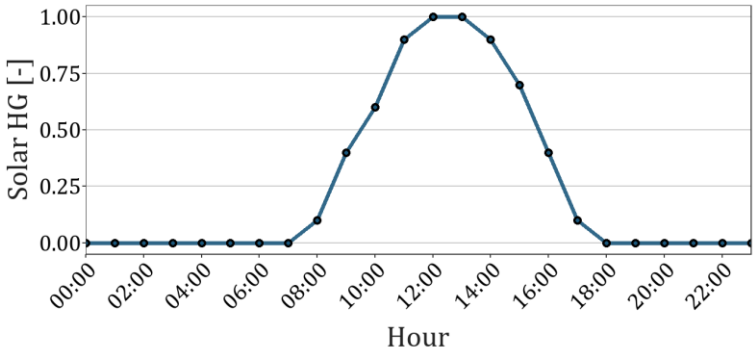


Figure 3.4: Solar heat gain schedule for a typical day.

So, the heat load of the building in winter ca be established as follows:

$$T_{eq\ heating} = T_{sp} - T_{design\ heating} \tag{3.7}$$

$$Q_{heating} = \frac{T_{eq\ heating}}{R} \tag{3.8}$$

Whereas the cooling load of the building in the summer state is calculated through the following equations:

$$T_{eq\ cooling} = T_{sp} - T_{design\ cooling} - HG_{int} - HG_{sol} \tag{3.9}$$

$$Q_{cooling} = \frac{T_{eq\ cooling}}{R} \tag{3.10}$$

3.2.2 Energy demand for the production of Domestic Hot Water

Domestic Hot Water (DHW) demand accounts for a significant part of energy consumption in residential buildings. It is not uniform during the day, as it depends on user habits, the season, the climate in which the building is located and occupancy. A number of studies in the literature show how knowledge of DHW consumption profiles can help develop appropriate control strategies to promote energy saving. The statistical model, called DHWcalc, proposed by U. Jordan et al. [56] allows the generation of consumption profiles through probability functions. Specifically, the mathematical formulation of the model is as follows:

$$p(t) = p_{day}(t) \cdot p_{weekday}(t) \cdot p_{season}(t) \tag{3.11}$$

where $p_{day}(t)$, $p_{weekday}(t)$ and $p_{season}(t)$ represent the daily consumption profile described by a Gaussian distribution, the consumption profile on weekdays and weekends and the seasonal variation of consumption in a year, respectively. From research of E. Fuentes et al. [57] and R. Hendron et al. [58], a typical daily profile and the relationship between weekends and working days are derived.

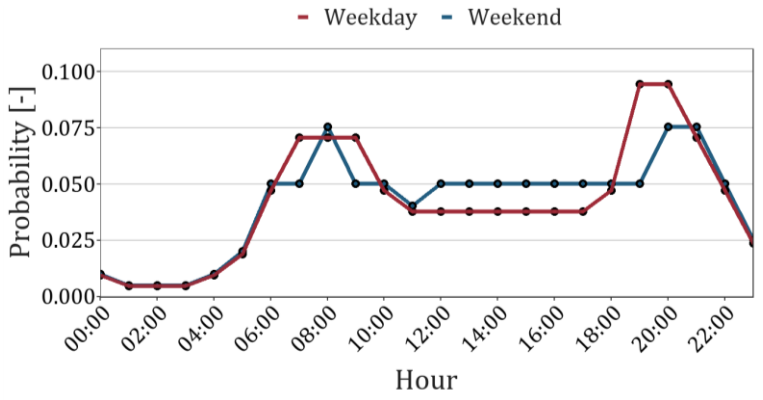


Figure 3.5: Typical daily weekday and weekend profile.

A 20% increase in consumption is considered at the weekend. To this, the DHWcalc model represents annual seasonal fluctuations through a sine function. The thermal demand for DHW at a given instant can be determined using the following equation:

$$E_{DHW}(t) = p(t) \cdot \rho \cdot cp \cdot V \cdot (T_{mains} - T_{DHW}) \quad (3.12)$$

where V represents the amount of DHW per event, provided as input as an energy system parameter. T_{mains} and T_{DHW} are the temperature of water coming from the aqueduct and the flow temperature of DHW, respectively.

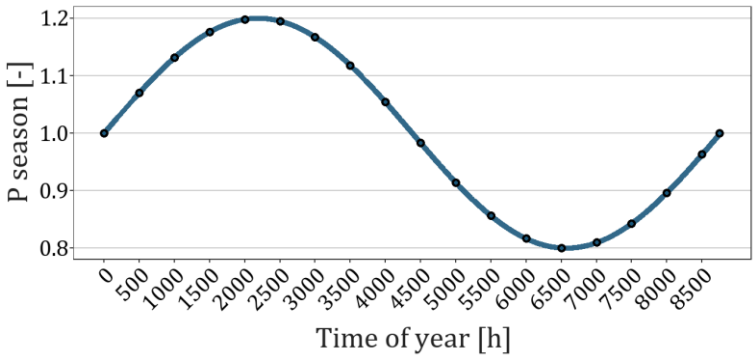


Figure 3.6: Trend of the annual seasonal variation.

In this analysis, these parameters are fixed and assumed constant because the effects of seasonal changes are already included in the probability function.

3.2.3 Energy demand for household appliances

An electrical consumption model of household electrical appliances was implemented in the VSE, which is equivalent to one third of a building's total electrical energy demand [59] due to increasingly significant efficiency improvements. The main household appliances which contribute to energy demand in residential areas are televisions, refrigerators, freezers, washing machines, dryers, dishwashers. While specific consumption patterns are available depending on the type of appliance dictated by relevant regulations [60], a single daily consumption profile was assumed for all buildings analysed proposed by the study of Pratt et al.'s [58]. The daily consumption associated with household appliances was determined by Monte Carlo estimation of the area subtended by the daily profile and then distributed over the day. This estimation provides a scattering of results with respect to the same overall

average load, so the actual estimated area varies every day for all buildings, which means that the building will not have the same average load every day.

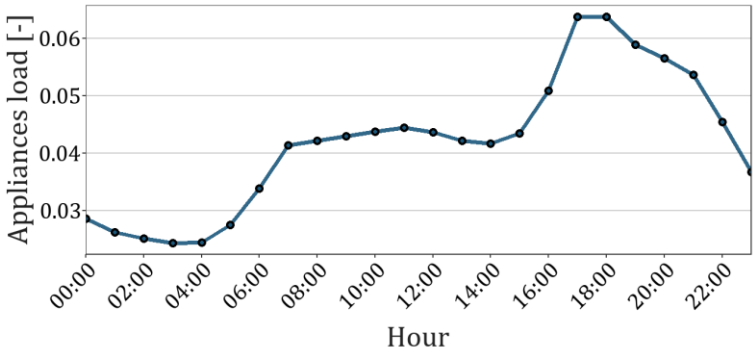


Figure 3.7: Total combined residential equipment schedule for a typical day.

In order to optimally estimate the electrical load, the household appliances consumption is an input data, provided among the energy system parameters.

3.3 Modeling of the energy systems

One of the objectives of the study is to investigate the potential impact of different energy generation and storage systems installed in variable context, mainly characterized by different climatic zones and environmental parameters. The VSE involved the introduction of models related to different distributed energy conversion, storage and generation technologies able to meet the needs of each building.

3.3.1 Thermal energy production: HVAC

To achieve a realistic and adequate analysis, it is important to build a model of the energy system which provides temperature control within individual buildings. To this end, the integration of an air-to-air vapor compression HP system was planned: it can be powered electrically by a PV system – if available – or by the power grid. In order to obtain the electrical absorption data, this system was sized according to the maximum winter heat load and the minimum summer cooling load so that it would meet heating and cooling needs, respectively, even under the most severe conditions. The operating parameters depend on the external temperature conditions, so the constructed model took into account the strong variability of the

system's operating modes as a function of these changes during the simulation period. More specifically, Coefficient Of Performance (COP) and Energy Efficiency Ratio (EER) represent the efficiency of converting electrical energy into thermal energy and are defined as a function of the difference between the temperature of the external environment T_{amb} and the heat transfer fluid temperature to the condenser $T_{supply\ heating}$ and evaporator $T_{supply\ cooling}$ respectively, through the report proposed by J. Vivian et al. [61]:

$$COP = 6.81 - 0.121\ (T_{supply\ heating} - T_{amb}) + 0.00063\ (T_{supply\ heating} - T_{amb})^2 \quad (3.13)$$

$$EER = 6.81 - 0.121\ (T_{amb} - T_{supply\ cooling}) + 0.00063\ (T_{amb} - T_{supply\ cooling})^2 \quad (3.14)$$

Graphs depicting the COP and the EER trends are shown in Figure 3.8 and Figure 3.9, as the outdoor air temperature changes.

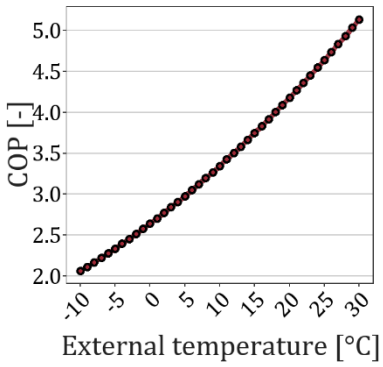


Figure 3.8: COP heat pump trend with condenser temperature equal to 45 °C.

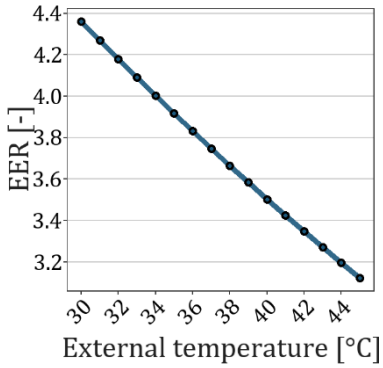


Figure 3.9: EER heat pump trend with evaporator temperature equal to 7 °C.

The final electricity consumption of the air-to-air HP is determined in winter and summer trim according the following equations:

$$E_{el,HVAC} = \frac{Q_h}{COP} \quad (3.15)$$

$$E_{el,HVAC} = \frac{Q_c}{EER} \quad (3.16)$$

The HP in winter trim can supply thermal energy Q_h to the building and charge the thermal storage, which can be seen as a flywheel capable of avoiding continuous

switching, on and off, of the system. Likewise, in summer mode the HP provides cooling energy Q_c to the building and is able to charge the cold storage. The higher the value of these performance parameters, the lower the power consumption to provide these services.

3.3.2 Domestic Hot Water production

HPs can provide advantageous solutions when combined with other power generation systems. In some cases, the air-to-air HP alone may not be able to cope with building load due to the outdoor temperatures, so there may be significant advantages when supplemented with a traditional GB or EH. Depending on the configuration chosen, it is possible that the two plant systems will produce in parallel by distributing the heat load or there may be specific priority rules concerning their operation. In the VSE, the HP was assumed to meet the heat load due to air conditioning, while a GB or an EH, depending on the technology installed in the specific building, meets the demand for DHW. Both are assumed to be non-flexible loads, so generation corresponds to consumption in each simulation episode. For a building served by an EH, the electricity consumption can be mathematically determined by applying the following equation:

$$E_{el,DHW} = \frac{\rho \cdot cp \cdot V \cdot (T_{mains} - T_{DHW})}{\eta_{el}} \quad (3.17)$$

In the case of DHW production by GB, the gas load is described by a different conversion efficiency:

$$E_{gas,DHW} = \frac{\rho \cdot cp \cdot V \cdot (T_{mains} - T_{DHW})}{\eta_{gas}} \quad (3.18)$$

For a consistent assessment of the electrical load and the gas load, the ratio of the number of buildings with GB and the number of buildings with EH was included among the energy system parameters.

3.3.3 Distributed Energy Resources

The electricity needs of community members can be met by integrating locally available power generation systems, such as distributed PV systems. Of particular importance is the association of this plant system with a HP, as it can result in an

increased share of self-consumption and savings in utility bills, by virtue of covering part or all of the HP's electricity needs. The model of PV system was built through a Python library called "pvlib". For good integration with the electrical load, the model exploits typical operating parameters, usually provided in the manufacturer's data sheet. In detail, information is required for such a module, which concerns the characteristics of the PV module, the attributes of the inverter and an indication of the percentage of buildings equipped with a PV system.

3.3.4 Energy Storage Systems

Renewable energy production facilities generally feature non-programmable generation, so integrating them with Energy Storage Systems (ESS) – like electric batteries or sensible heat storage – allows for maximum utilization of on-site generation potential and maximizes flexibility, self-consumption and self-sufficiency. Electrical storage technologies can be of different types, each with different discharge times, capacities and power ratings. The operation of a storage system can be described using the SoC model, according to which E_{ESS} defines the amount of energy exchanged between the storage and the electricity/thermal user and ΔE_{ESS} is the change in the amount of energy available in the storage:

$$\Delta E_{ESS} = E_{ESS} \cdot \eta_{charge} \quad \text{when} \quad E_{ESS} > 0 \tag{3.19}$$

$$\Delta E_{ESS} = \frac{E_{ESS}}{\eta_{discharge}} \quad \text{when} \quad E_{ESS} < 0 \tag{3.20}$$

where η_{charge} is the charge efficiency and $\eta_{discharge}$ is the discharge efficiency. A storage system typically has technical safety constraints, defined by the amount of energy exchanged and the SoC, which are required to preserve their long-term operation. Especially during the charging and discharging phases, the following conditions must be observed, which are fundamental to ensure that these phases are not too quick:

$$E_{ESS,min} < E_{ESS} < E_{ESS,max} \tag{3.21}$$

It is also important to observe the SoC limits defined by the manufacturer to keep the battery in good condition over time:

$$SoC_{min} < SoC < SoC_{max} \quad (3.22)$$

A BESS is a widespread technology capable of storing electrical energy over a shorter or longer period of time and available in different sizes. This system is not expected to exchange electricity with the power grid, but must be charged exclusively by the PV arrays or on-site generation facilities:

$$|E_{BESS}| < |E_{PV}| - |E_{el,load}| \quad \text{if} \quad E_{BESS} > 0 \quad (3.23)$$

The BESS cannot feed electricity into the power grid, but its discharge is expected only if the PV system or on-site generation system cannot meet the building's electricity needs:

$$|E_{BESS}| < |E_{el,load}| \quad \text{if} \quad E_{BESS} < 0 \quad (3.24)$$

In the BESS model it is necessary to incorporate, among the input parameters, the Depth of Discharge, the Round-Trip Efficiency, the ratio of Maximum charge/discharge power and also an efficiency of the DC/DC converter connected to the BESS.

Thermal Energy Storage (TES) has increasingly attracted interest, given the large integration of renewable source systems in recent years, as it allows for indirect storage of electrical energy through the use of HPs. The thermal energy produced by these technologies can be stored and then be harnessed again for the production, through heat engines, of electricity. Applications involving low-temperature sensitive thermal storage systems are well established by virtue of the almost complete electrification of air conditioning systems. CTES are also increasingly being integrated into more applications and used for summer air conditioning. For this study, it is assumed that HTES and CTES can only be charged by the HP:

$$|E_{TES}| < |E_{HP}| - |E_{th,load}| \quad \text{if} \quad E_{TES} > 0 \quad (3.25)$$

And they can be discharged only when the HP cannot directly meet the building's heat load:

$$|E_{TES}| < |E_{th,load}| \quad \text{if} \quad E_{TES} < 0 \quad (3.26)$$

In this way, these systems allow the electrical load to be manages, since they can determine its peak variation, thereby improving the building's flexibility with respect

to the power grid. Regarding the input parameters, the model requires the percentage of buildings which have a HTES/CTES, a coefficient for sizing HTES/CTES as a function of HVAC thermal capacity, the Round-Trip Efficiency of HTES/CTES and finally a parameter between 0 and 1 which defines the losses of these system.

3.4 Control strategies

In order to reduce CO_2 emissions mainly related to building heating, an optimal solution is therefore to replace traditional thermal energy production systems (e.g. boilers) with HPs. This makes it possible to switch from thermal consumption to electrical consumption, which can be met by renewable energy sources. To optimally manage energy production and utilization, a MAS-based framework integrating three different independent controllers was developed. These operate according to RBC strategies, designed to exploit the flexibility of the building while reducing energy costs. The indoor air temperature is controlled by imposing its value within an acceptable range determined by the comfort conditions of the occupants. For this reason, in the winter set-up the heat load is met by an air-to-air HP or by a HTES, when the indoor air temperature falls below the lower limit of the acceptable range. Therefore, it increases until it reaches the upper limit of the acceptability range. A similar strategy is implemented in summer set-up, based on reversing the mode of operation of the HP. TES is managed according to a peak shifting strategy, as a matter of fact, under low price and zero SoC conditions, the HP provides full storage charging. However, when the electricity price is high – that is, during peak periods – the system is charged only when there is a thermal load in the building and the SoC is greater than zero. As far as the BESS is concerned, a control strategy widely used in literature was adopted. It is planned to be charged when there is a surplus of production from PV compared to the building load, otherwise it is discharged. In case the surplus is more than the system charge, then the overproduction is fed into the power grid. On the other hand, if the electricity produced by PV is not enough to fully charge the system, then the completion of charging is done by taking from the power grid.

3.5 Key Performance Indicators

The potential for large-scale energy flexibility, if cost-effective, enables supporting stakeholders to provide services to the power grid. This feature is also increasingly being analysed because of changes in energy prices. As a matter of fact, shifts in demand related to prices, self-production and self-consumption are among the most widely used KPIs for quantifying energy flexibility and interaction with the power grid. Given the wide variety of buildings, these metrics must be suitable for evaluating a large number of technologies by including building-specific influencing factors. These indicators can be defined as absolute values (e.g. electrical or thermal load, energy generated from renewable sources, electricity prices taken from the power grid) or as relative ones (e.g. load-to-average factor), calculated with reference to specific scenarios.

Several authors have proposed formulas suitable for assessing individual building KPI, which are subsequently extended to the EC level. KPIs currently applied to EC buildings are the Self-Consumption Index (SCI) and the Self-Sufficiency Index (SSI). These indicators are useful because they are easily applicable to different building stocks and also meet the needs described by the benchmarking analysis. The SCI [62] represents the fraction of self-consumed electricity E_{sh} of the total electricity produced on site P_{local} and is mathematically determined according to the following equation:

$$SCI = \frac{\sum_{k=1}^{8760} E_{sh}(t_k)}{\sum_{k=1}^{8760} P_{local}(t_k)\ \Delta t} \tag{3.27}$$

Building with values of this indicator close to zero purchase most of their electricity from the power grid due to the small size of the renewable energy systems placed on site (e.g. PV, ESS, CHP). The SSI [62] represents the share of electricity self-consumption E_{sh} in total electricity consumption E_{total} and is calculated using the following equation:

$$SSI = \frac{\sum_{k=1}^{8760} E_{sh}(t_k)}{\sum_{k=1}^{8760} E_{total}(t_k)\ \Delta t} \tag{3.28}$$

An index value close to one indicates that the entire amount of energy consumed is supplied by renewable energy sources located on site and is therefore characterized by a high degree of self-sufficiency. An index value close to zero indicates that most of

the electricity produced on site is sold to the power grid. Like the SCI, the SSI depends on the size of renewable energy systems and the type of systems available. These indicators can be used for a variety of purposes. For example, A. Cielo et al. [41] used these two indices to build a scale model of an EC consisting of on-site renewable energy generation facilities (PV and BESS). Further integration of economic indicators allowed them to identify the optimal configuration. G. Pinto et al. [18] also examined the SSI to measure the effects of different control strategies in the integration of renewable electricity. G. Mutani et al. [63] in their study examined the design stages of a REC and then applied these models to the reality of Villar Pellice (Turin). The evaluation of these indicators is crucial because a Law of the Piedmont Region assumes a minimum level of 70% SCI and the choice of interventions to be undertaken must meet this need.

On the other hand, energy flexibility performance indicators are related to time, cost and efficiency. They are widely used because they can be integrated with RBC paradigms, with the aim of minimizing operating costs while accounting for energy price fluctuations [16][13]. According to [16], the Flexibility Factor (FF) is mathematically defined as a function of the electrical energy consumed in peak hours $E_{el_{peak}}$ and in off-peak hours $E_{el_{off\ peak}}$:

$$FF = \frac{E_{el_{off-peak}} - E_{el_{peak}}}{E_{el_{off-peak}} + E_{el_{peak}}} \tag{3.29}$$

Again, different authors use this metric for different purposes. For example, G. Pinto et al. [18] used this indicator to quantify the performance of a controller.

Information about the shape of the load over time takes on a relevant role for comparison activity. In this regard, another metric related to building belonging to the EC can be evaluated and defined as Peak-to-Average Ratio (PAR):

$$PAR = \frac{\sum E_{peak,i}}{\sum E_{av,i}} \cdot \frac{1}{D} \tag{3.30}$$

where $E_{peak,i}$ represents the peak electrical power on the i-th day, $E_{av,i}$ represents the average power on the i-th day and D is the total number of days.

[13] Since it is not easy to combine all these properties into a compact metric, different definitions are given in [16], [18], [74]–[76].

To best describe the shape of the load, an additional indicator called Electrical Load Factor (ELF) can be determined as follow:

$$\mathrm{ELF} = \frac{\sum \mathrm{E_{net,av}}}{\mathrm{E_{peak}}} \tag{3.31}$$

where $\mathrm{E_{net,av}}$ indicates the average net load.

Within ECs, the purchase of electricity from the power grid is not excluded. While a good share of this comes from renewable energy sources, a small fraction comes from non-renewable energy sources. An assessment of CO_2 emissions for the production of electricity purchased from the power grid is a good indicator which allows to compare different ECs. In this regard, it is possible to assess the effect of climate-altering substances emitted into the atmosphere, depending on the energy carrier used for energy production:

$$\mathrm{GHG} = \mathrm{GWP_{coal}} \cdot \mathrm{E_{el_{coal}}} + \mathrm{GWP_{natural\ gas}} \cdot \mathrm{E_{el_{natural\ gas}}} + \mathrm{GWP_{petroleum}} \cdot \mathrm{E_{el_{petroleum}}} \tag{3.32}$$

where $\mathrm{GWP_{coal}}$, $\mathrm{GWP_{natural\ gas}}$ and $\mathrm{GWP_{petroleum}}$ were provided as input simulation parameters, while the data of electricity produced and presumably taken from the different energy sources are obtained from a database, which provides hourly electric grid monitor [64].

Performance indicators available in the literature can be distinguished into classes according to load shifting capacity, energy performance, flexibility and economic aspects. In this study, are not analyse cost indicators, which are usually evaluated during the design phase, as done in the study [63] or in the study [18] for the comparison of different Deep Reinforcement Learning architectures. As a matter of fact, the analysis reported aims to integrate the information collected from the KPI submitted and acts both as a mechanism for comparison among different ECs and as a basis for their design.

4 Case study and implementation

This chapter introduces our case study and all the parameters which are initialized to perform the simulations. In more detail, the first part will describe the structure of the ANN, which were trained from monitoring data of some buildings located in North America. Then the modules of the VSE and their configuration will be analysed in more detail with reference to this work. The last paragraphs will introduce the various, potentially realistic energy scenarios that will be investigated in the case study.

4.1 Design of Long Short-Term Memory neural network

N. Luo and T. Hong [65] extracted a subset of data for buildings located in California, Illinois, New York and Texas from the full dataset of the DYD program. These data allowed us to construct ANN models for one of the four North American states whose data were made available, in order to describe the thermal dynamics of buildings. The algorithms thus performed are useful for applying advanced control strategies at the district level. California is the location taken into account in our analyses, for which monitoring data are available at the individual building level. The proposed methodological framework employs LSTM neural network, an archetype widely adopted for building control purposes, since it can process entire sequences of data. In this study, the development of models of this type of neural network – aimed at predicting the change in indoor air temperature of buildings – is proposed. The results, with regard to model performance, will therefore be compared with those

of the reduced-order models developed by Z. Wang et al. in their study [24], evolved from the data inferred by the same monitoring campaign.

The collected data were organized into 12 CSV files, each of which contains one month of data and have a sampling rate of five minutes. The data gathered for each building covers a time horizon from 1st January 2017 to 31st December 2017. These include:

- **DateTime:** date and time when the observation was made.
- **Event:** conditions which replace the default schedule, such as *demand response* or *vacation.*
- **Schedule:** conditions set by the occupant, such as *home* or *sleep* or *away.*
- **Control temperature:** average indoor air temperature measured by sensors and based on user-defined comfort programs.
- **Heating set-point temperature:** indoor air temperature set-point during the heating season.
- **Cooling set-point temperature:** indoor air temperature set-point during the cooling season.
- **Relative humidity of indoor air.**
- **Auxiliary heat:** operating time of any heat source other than a HP.
- **Cooling component:** operating time of any cooling system.
- **Heating component:** operating time of any HP.
- **Fan:** operating time of fan.
- **Thermostat temperature:** indoor air temperature measured by remote thermostats.
- **Thermostat motion:** motion detection by remote sensors.

Time series of outdoor air temperature and outdoor relative humidity, taken from the nearest weather station to each building, are also available.

To achieve good performance of LSTM neural network model, adequate input parameters must be provided. In more detail, in order to predict the change in indoor air temperature during the winter months and the summer ones, the final set of features is composed by:

- **Thermostat temperature.**
- **Outdoor air temperature.**
- **Occupancy:** it is defined as a binary variable, where 0 indicates no occupants, while 1 stands for room occupancy detection.
- **Normalized heating power.**
- **Normalized cooling power.**
- **Cyclic time features:** they are referred to the time of day, day of the week and day of the year.
- **Solar irradiance.**

Other variables, such as the relative humidity of the indoor air or the relative humidity of the outdoor air, have a slight effect on the thermal dynamics of the building and for that reason were removed. Also the features involving heating set-point temperature and cooling set-point temperature were excluded because they take an almost constant trend during the respective seasons. Finally, control temperature is not used as an input for each LSTM neural network because it is a variable linked to room occupancy, which was not established according to a fixed procedure. In this way, all developed neural network structures are similar and include an input layer consisting of six nodes and an output layer made up of only one node.

Data analysis requires a preliminary process involving measured data. In more detail, these come from the real world and may be characterized by missing values, outliers and/or inconsistencies. A data pre-processing phase is shown to be essential to convert raw data into higher quality data, useful to make suitable supervised learning models. An analysis of the dataset, obtained by merging the data from the different CSV files, does not reveal records with outliers, but some observations with inconsistencies and missing values are stated. In detail, records with a *Thermostat Temperature* below 20 °F or above 120 °F emerged and were replaced with missing values as meaningless. LSTM neural networks are able to exploit information contained in rather long time series. Despite this, to avoid numerical instability problems and to manage the weights of each module, it is good practice to sample the dataset in small time sequences called chunks. To avert information leakage, the k-NN[14] method is applied to each chunk. This algorithm is

[14] The k-NN algorithm is widely applied in the case where there are missing values in the dataset. It involves detecting the K nearest samples of data in the dataset through the

enforced only in the case where missing values are current for a time horizon of less than five hours, while for larger time frames, records are expected to be deleted since their reconstruction would not be easy. In the case where a chunk contains enough information to be useful for training the neural network, it is retained. Conversely, if the time horizon of the data belonging to the sequence is less than that used for prediction added to the forecast time step, then it is discarded since it does not contain enough information which can be useful for model development.

To achieve good performance of the neural network during the training phase, it is advisable to address the seasonality of indoor air temperature with respect to time variables. Two periodicities emerged from the time series related to measured air temperature, concerning the time of day and the day of the week. An additional periodicity is taken regarding the day of the year, although multiple years of data are not available. This assumption makes the model more scalable to different contexts. Therefore sine-cosine transformations for time of day, day of the week and day of the year are applied on the rearranged data with an hourly sampling rate. These transformations take the form shown below:

$$\mathrm{x_{sin}} = \sin\left(\frac{2\pi \mathrm{x}}{\max(\mathrm{x})}\right) \tag{4.1}$$

$$\mathrm{x_{cos}} = \cos\left(\frac{2\pi \mathrm{x}}{\max(\mathrm{x})}\right) \tag{4.2}$$

where x_{sin} and x_{cos} are the cyclic features of variable x. Through this application, each time observation takes on a continuous distribution. In addition to this, since the measurement campaign did not include the collection of energy consumption data – which play an important role in explaining the thermal dynamics of the building – parameters belonging to the meta data were used. In more detail, the heating and cooling power levels available in each building are employed to establish the normalized heating power to be provided in winter and the normalized cooling power to be subtracted in summer. As a matter of fact, the winter heating load of the building can be met by the HVAC system and/or an EH. In the case where the two plant systems cannot work simultaneously, then an operation of one of them under nominal conditions was assumed. Otherwise, it has been employed that the HVAC system runs at nominal operating conditions, while the EH works at 50% over its nominal size

application of a Euclidean distance function. These are then used to predict the value of the missing record.

during the on-time period. In addition to this, during the cooling season, it was assumed that the HVAC system always performs under nominal conditions during the running time frames. Finally, the number of remote sensors detecting motion, which are derived from the meta data, are also used as useful information to define room occupancy schedules. Another key aspect in assessing the thermal dynamics of buildings concerns incident solar irradiance, which is absorbed by opaque and transparent envelope components and stored as heat, subsequently transferred to the indoor air. It is a relevant variable because it provides a contribution to the change in sensible heat load. In absence of this information in the available data, California has been partitioned into sixteen climate zones, following the segmentation suggested in [66]. Next for each building, belonging to the meta data referred to this state, was assigned the corresponding climate zone and to each of these was matched a time series of solar irradiance given for the same year as the data historicized in the dataset. To obtain improved model results, the datasets were also sampled by making a distinction between the winter period – which runs from October to March – from the summer period – which runs from April to September –.

Data scaling is a mandatory process for the application of supervised learning techniques based on neural network algorithms. Several scaling methods can be adopted – including standardization or rescaling on the maximum – but the one widely used to deal with energy data is minimum-maximum scaling, expressed by Equation 4.3.

$$\bar{y} = \frac{y - \min(y)}{\max(y) - \min(y)} \tag{4.3}$$

where $\bar{y}$ is the scaled variable and y is the variable to be scaled. Because algorithms such as neural networks do not learn according to the physical meaningfulness of a process, but perform comparative analyses between data and rely on grafted, complex, nonlinear mathematical relationships, this step is critical.

Neural networks are supervised learning algorithms in which, with respect to the dataset, the input and output variables are known. From our dataset, we developed a matrix containing only the input variables X and a vector carrying the only output variable y, corresponding to indoor air temperature. So, the new dataset containing the matrix X and vector y are then partitioned into two different datasets. One of these, containing multiple records, is used for learning the neural network. The

second one, on the other hand, is applied for the performance testing phase of the model, which is useful to understand how good its training was. The models, developed entirely in Python using the Pytorch framework, are trained for a forecast horizon of one hour based on a history of the previous twelve hours. The ADAptive Moment estimation optimizer, used in place of the classical gradient descent method, updates the weights of individual modules in the network with the goal of minimizing the MSE. The LSTM neural network model requires the setting of hyperparameters before the actual learning phase. The proper selection of these markedly affects the performance of the neural network. Since this stage is regarded as an art and it is very difficult to establish the optimum because of the complexity of the models, Optuna [67], a tool which automates the search for appropriate hyperparameters for the neural network, is used in this study.

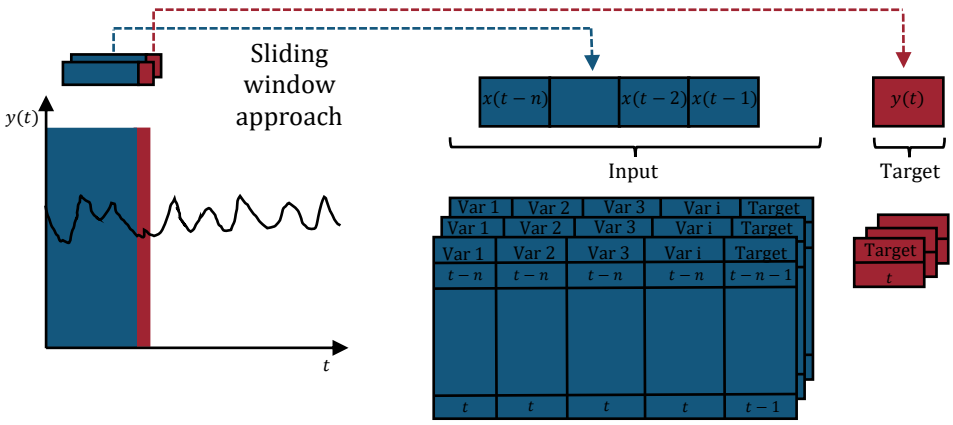


Figure 4.1: Sliding window approach.

In the model training and deployment phases, the use of an unlocked window of fixed length based on a series-to-supervised approach is planned, as explained in Figure 4.1. For this reason, the variable to be predicted will have a time delay of one hour relative to the other input features. Thus, LSTM neural network models trained in this way will provide an hourly prediction – according to a one step ahead scheme – of indoor air temperature based on the previous twelve hours of data, using an open loop controller. This means that once the forecast is made, it is not used at the next time step as input to the neural network, but only measured or hypothesized real variables are expected to be exploited.

4.2 Implementation of the energy scenarios

The test case selected for this study concerns residential complexes consisting of 50 buildings belonging to Texas, California and New York – North American locations for which actual monitoring data were made available for buildings belonging to them – in order to represent an EC. To this, different technological configurations capable of satisfying the energy needs of the three different ECs are analysed, identifying the optimal one which may be able to offer services to the external power grid. The proposed methodology is therefore based on an analysis in heating mode – from 10 January to 10 February – with a time-step simulation of 5 minutes. The chosen time horizon allows a complete assessment of the energy contribution of all systems envisaged at the community level, which would not be easy to develop in the summer trim – since one of the energy technologies analysed operates only during the heating period –. The chosen simulation time-step allow to obtain a remarkable precision of the simulation, but has the disadvantage of being computationally expensive.

The meteorological data were evaluated from Python's pvlib module, which provides a Typical Weather Year based on location. It contains information on outdoor air temperature, relative humidity, solar irradiance and wind speed. In all buildings belonging to the EC, the energy demand for space heating and space cooling is determined in such a way as to maintain the indoor air temperature within an acceptable range of variation, established from Fanger's comfort theory. In particular, the indoor air temperature must be maintained between 21.88 °C and 23.75 °C in heating season and between 23.13 °C and 24.32 °C during the cooling season. The design outdoor temperature were set to be the same for all simulated locations and corresponding to -8.3 °C for the winter regime and 33 °C for the summer regime, while the supply heating temperature was assumed to be 45 °C and the supply cooling temperature was set at 7 °C. The COP and EER of air-to-air HP systems were not assumed constant, but their dependence on outdoor temperature and supply temperature was considered.

DHW demand is met by either a GB or an EH, depending on the building. The ratio between buildings equipped with GB and those equipped with EH was set at 0.2. As stated by E. Vakkilainen [68], the GB chosen has a reference efficiency of 94%, while the EH has been hired with an efficiency of 98%. The daily DHW consumption

is 2.3 l with a standard deviation of 44.8 l, while the design volumetric flow rate is 20 l/event with standard deviation 5 l/event. The production of DHW by EH makes a contribution to the electrical load of the building, as well as the consumption by household appliances, assumed to be 4.5 kWh with standard deviation of 110 kWh.

The PV modules considered in the study were taken from the one available in the Sandia database, named Canadian Solar CS5P 220M 2009 [69]. Since the generation of electricity takes place in Direct Current (DC) from the PV panels and this is not exploitable for domestic purposes, it is necessary to integrate a PV inverter which performs the function of converting DC to AC. The selected inverter belongs to the technologies available in the SAPM database and is named ABB MICRO 0 25 I OUTD US 208 208V.C. Considering that not all buildings may have at their disposal PV systems, a penetration rate of 75% was assumed and kept fixed for all simulations. Information on climate data such as outdoor air temperature, solar irradiance and wind speed, made it possible to determine the power produced by the PV system throughout the simulation time horizon.

The technical parameters of BESS used for this research were derived from the analysis of a series of technical data sheets on commercial technologies. Among them, a typical Depth of Discharge of 60% was taken.

Parameter	Value
Round-trip efficiency	0.98
Maximum charging power	1 C
Maximum discharging power	½ C
DC/DC converter efficiency	98%
Discharge depth	60%

Table 4.1: BESS parameters.

In addition to this, the system exchanges power with a DC bus by means of a DC/DC drive, whose efficiency value is estimated at 98%. The maximum charging and discharging power was defined in proportion to the capacity and was set to 1 C and ½ C, respectively. Other simulation parameters referred to this energy system are defined in Table 4.1. HTES and CTES technologies were not provided for all buildings, but only for 70% of the simulated total. To model these systems, some parameters

described in Table 4.2 were set. A Hot/Cold water storage sizing factor was assumed to be 2, while a loss coefficient was set at 98% to account for any heat loss.

Parameter	Value
Hot/Cold water storage sizing factor	2
HTES/CTES Round-trip efficiency	0.98
HTES/CTES loss coefficient	0.98

Table 4.2: HTES and CTES parameters.

Finally, additional parameters useful for simulations are defined in Table 4.3. These include model noise and measurement error, as well as information related to the thermophysical characteristics of buildings.

Parameter	Value
Model noise	0.002
Measurement error	0.1
R/C ratio	(0.4, 0.01)
Heat gain ratio	(0.3, 0.01)
Internal heat gain method	DOE

Table 4.3: Other simulation parameters.

In order to derive thermal resistance value and heat gains, it is necessary to provide as input the parameters of thermal time constant and Equivalent heat gain temperature, shown in Table 4.4 for the three North American states taken into account in the current case study.

Parameter	Value for Texas	Value for California	Value for New York
τ	(19.5, 9.9) h	(15.3, 5.8) h	(17.3, 9.2) h
$T_{HG,eq}$	(12.4, 7.0) °C	(11.4, 5.7) °C	(11.4, 6.0) °C

Table 4.4: Other building parameters.

The EC can supply electricity to the power grid when there is a surplus of production over consumption or it can withdraw from the power grid in deficit conditions. The cost of electricity taken withdrawal varies depending on the tariff,

which can be monorary, biorary or triorary depending on the building. Since a large part of the grid electricity is generated from fossil sources, an appropriate analysis of the amount of CO_2 equivalent emitted for the share of energy withdrawn is appropriate. Specifically, emissions per electric kWh produced are determined in terms of CO_2 intensity, defining the specific emissions to be 1.01 kg/kWh for coal, 0.41 kg/kWh for natural gas and 0.97 kg/kWh for oil [70]. The CO_2 emissions for each state depend not only on energy consumption, but also on when energy is consumed. As a matter of fact, to calculate this, electricity production data for a whole year from a database [70] were used, which differ by location. Finally, the natural GB feedstock cost was assumed to be 0.093 $/kWh[15] as required by the Authority for domestic customers [71].

With respect to local decentralized power generation technology, the use of a MGT allows for reduced dependence on the power grid. For the integration of such system, the set-up of the VSE requires additional input parameters, including the enhancement of the electricity fed into the power grid by this technology, fixed at 0.054 $/kWh. The chosen MGT has an overall efficiency under ISO conditions at 15 °C of 85% and an electrical efficiency of 26%, as described by R. Boukhanouf [72] and Y. Hwang [51]. Table 4.5 summarises the parameters related to this technology and its integration with a DHN. The sensible effectiveness of the counterflow plate heat exchanger placed at the level of each building served by the DHN was set at 80%, as defined by Y. Luo et al. [73]. A maximum capacity of each substation equipped with heat exchangers serving the individual building was also assumed to be 10 kW.

Parameter	Value
Electrical efficiency	0.26
Thermal efficiency	0.57
Overall efficiency	0.85
Heat exchange efficiency	0.97

Table 4.5: MGT parameters.

[15] This value was derived from the data referring to domestic users in the north-western area (Valle d'Aosta, Piedmont, Liguria) for the period between 1st October 2021 and 31st December 2021.

To ensure a certain energy flexibility, consumption management is particularly important within the ECs. The number of EVs is increasing significantly and, in the near future, a reduction in the costs associated with their purchase is expected, which could lead to their strong expansion. Proper management of the electrical absorption linked to their charging is essential to avoid burdening the power grid. To simulate the EVs charging network, a battery capacity of 90 kWh was set, in accordance with the license plate data of the main vehicles available in the market. Other technical parameters used in the simulation have been summarised in Table 4.6 and concern charging efficiency and other factors related to EV batteries performance.

Parameter	Value
EV capacity	90 kWh
EV charging efficiency	0.98
EV Depth of Discharge	0.98
EV BESS Round-trip efficiency	0.98
EV BESS max charge power	1
EV BESS max discharge power	½
EV BESS converter efficiency	0.98

Table 4.6: EV parameters.

With the objective of evaluating the arrival time and the SoC of the simulated EVs, a number of parameters related to the model used must be provided as input. These were taken from the study by D. Zou et al. [53] and presented in Table 4.7.

Parameter	Value
μ_s normal distribution	17.53
σ_s normal distribution	2.96
μ_{SoC} normal distribution	0.3
σ_{SoC} normal distribution	0.1

Table 4.7: Other EV parameters.

So far, we have analysed the VSE and presented the data to be provided as input to its. In the following, we will instead move towards describing the proposed scenarios.

4.2.1 First scenario

The first scenario focuses on the optimal management of energy streams, which result from the combination of intermittent renewable energy source systems and a micro-cogeneration plant. This is able to meet the needs in periods of energy deficit and the cost of purchasing electricity from the power grid high. In more detail, PV generation matches the electricity demand of community members during daylight hours and, in the case of any deficit, the withdrawal from the power grid is expected. In the early morning and late afternoon the MGT is activated, according to a properly defined RBC strategy, and helps to balance the electrical load. During the night period, on the other hand, it is necessary to withdraw from the local power grid since it is not economically viable to use the micro-cogeneration system. The design of the MGT was based on an analysis of the electrical and thermal load profiles of the entire EC over a time horizon ranging from 1st November to 31st March[16]. Sometimes, the thermal energy produced by this system is able to cover the building's thermal demand. This allows the power supply to the HPs to be cut off during the time frames when the MGT is in action and the local power generation to be surplus to the electrical load, if we consider HPs peak electrical draw for sizing and operation of the system under nominal conditions. On the other hand, one of the objectives of the EC is to change self-sufficiency and self-consumption, so as to reduce withdrawal peaks from the distribution power grid. For this purpose, the size of the micro-cogeneration plant must be determined on the basis of a realistic assessment of electricity and thermal energy demand, thus taking into account only a small fraction of absorption by HPs, related to thermal storage management. This consideration stems from the fact that a share of the buildings heat load is offset through the thermal output of the MGT. To be more precise, the heat output is distributed equally to all buildings belonging to the EC. For this reason, it may happen that in some cases the HP remains off because the building heat load is fully met by the micro-cogeneration system, during its operating time horizon. On the other hand, for other buildings it may be the case that a turn-on of the HP is required to compensate for the share of heat load not satisfied by the MGT. In order to assess the electrical absorption fraction by the HPs, in an initial simulation the size of this plant system was set with reference to the peak electrical load of the entire EC occurring in the base case. A subsequent sensitivity

[16] This time horizon has been assumed as the heating period, the regime in which the μ-CHP plant works.

analysis of the results obtained identified the fraction of electrical absorption attributable to the air-to-air HP systems in buildings, as the size of the micro-cogeneration system changes. Through this information, a careful evaluation of the electrical load duration curves of the EC was carried out – referring limitedly to the operating hours of the MGT – and helped us determine the optimal size

4.2.2 Second scenario

The second scenario examines the energy, economic and environmental impacts involved in the introduction of consumption units such as EVs. The results of this scenario include two different analyses, which stand out in relation to the percentage of users who own EVs. A first analysis, which is assumed to be more realistic, predicts a penetration rate of such consumption units of 50%. A second one, which we can take into account as an ideal one, concerns a penetration rate of 100%.

4.2.3 Third scenario

Over the past decade, activities to encourage the introduction of EVs have intensified to an extreme degree. The increasing development of charging infrastructure has led to a great consensus in the adoption of such consumption units. The consequences associated with this phenomenon include a significant increase in electrical absorption, emission of climate-altering substances for the production of the electricity useful for charging and costs. With the aim of overcoming these problems, the third scenario proposes a strategy that can complement the micro-cogeneration plant sized in the first scenario, previously described, to also partially meet the demand from EVs. This strategy could be useful to minimize the share of energy exported to the power grid and reduce electricity withdrawal when the cost is high. The MGT was sized with reference to a percentage of the peak electrical load of the distributed HPs and, when turned on, it works under nominal operating conditions. Since the EVs arrive at the building charging stations in the late afternoon (around 18:00) and leave in the morning (around 8:00), it is possible to supply some of the electricity generated locally by the MGT to meet the electrical absorption linked to the charging of batteries of the vehicles. A reduction in the withdrawal of electricity from the power grid resulting from this strategy, could certainly lead to a reduction in purchase costs and a reduction in the CO_2 emitted for its generation. On the other

hand, the transfer of excess energy produced by the MGT is subject to economic valorisation, so it might not seem economically viable to take advantage of the μ-CHP plant to meet the additional load related to EVs. Economic, energy and environmental analyses were used to assess the potential benefits that could be achieved.

5 Results and discussion

This chapter provides the results of the methodologies that have been described in the previous chapters. More specifically, the LSTM neural networks models were applied to some buildings placed in California. Next, the three different scenarios are proposed regarding the various energy technologies which can meet the needs of a real EC, achieving the goal of harnessing most of the production from local, and sometimes uncontrollable, energy sources to reduce the load on the distribution power grid. The results obtained through the analysis of the different scenarios are presented for three locations in North America. A series of graphs and tables will make it possible to evaluate the various configurations and identify the optimal one considering energy, economic and environmental objectives.

5.1 Hyperparameter tuning and deployment

Simulations are applied to different buildings, with the aim of validating neural network models, which can later be used for control purposes. A cluster of buildings located in California were selected for this case study, all of which are detached. The dataset was split in summer and winter datasets to better characterize the behaviour of the buildings, and then, 70% of the records in the source dataset were used for the training phase from the data. The remaining percentage was exploited to evaluate the goodness of the models. Before the stage of actual supervised learning of LSTM neural networks, suitable hyperparameters had to be fixed (see section 2.1.1). Through Optuna [67], the appropriate ones were identified for each building examined, requiring the objective function to minimize the RMSE. These are shown for two buildings for winter (W) and summer (S) in Table 5.1.

Hyperparameter	Building 1 W	Building 2 W	Building 1 S	Building 2 S
Batch size	64	64	64	64
Lookback	12	12	12	12
Learning rate	0.0029	0.0027	0.0019	0.0044

Table 5.1: Hyperparameters of LSTM neural networks.

The number of hidden layers was set at 2 for all models, as was the number of hidden units of 256 and the amount of epochs equivalent to 50. The batch size was not set constant for the training phase of all neural networks, although for the buildings examined it is the same. As a matter of fact, configurations in which this hyperparameter tool a value of 128 were also investigated, but in all cases they did not prove adequate with respect the objective function. The results obtained during the testing phase of the LSTM neural network models are depicted for the same buildings in the Figure 5.1, which show a comparison of the trends of real variable and predicted one.

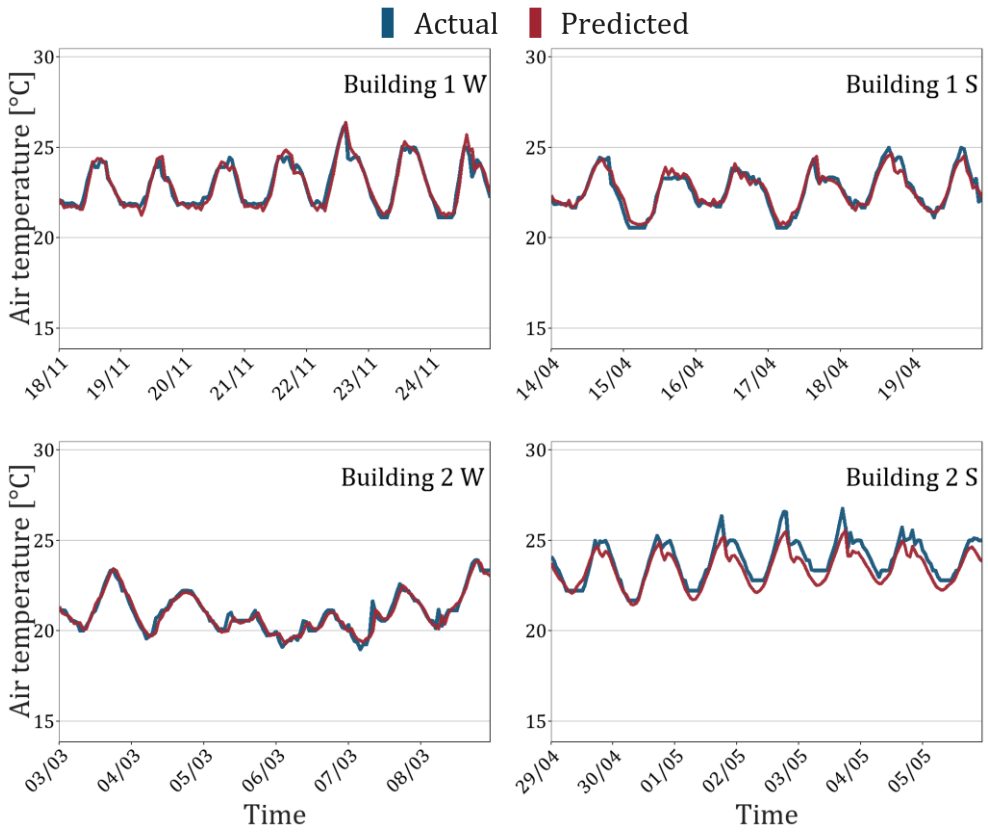


Figure 5.1: Comparison of actual and predicted indoor air temperature.

The indoor air temperature ranges between 21 °C and 26 °C during the winter testing period for *Building 1*, while it fluctuates between 20 °C and 27 °C for the summer time. A similar behavior occurs in *Building 2*, where between February and March the temperature varies between 19 °C and 24 °C, while during May there is a change between 19 °C and 28 °C. The prediction time horizon varies for the two buildings in both seasons due to the exclusion of some records from the source dataset marked by missing values. Although a strict continuous recurrence is not observed in the indoor air temperature variation during the testing days, the algorithms are able to realistically predict it in all cases.

With the aim of evaluating the performance of these models following training, three different error metrics were calculated employing the testing dataset: RMSE, MAE and R^2. As listed in Table 5.2, in all situations the RMSE is less than 0.60 °C. The low MAE value of less than 0.50 °C also shows that the modeling was able to understand the dynamics under study.

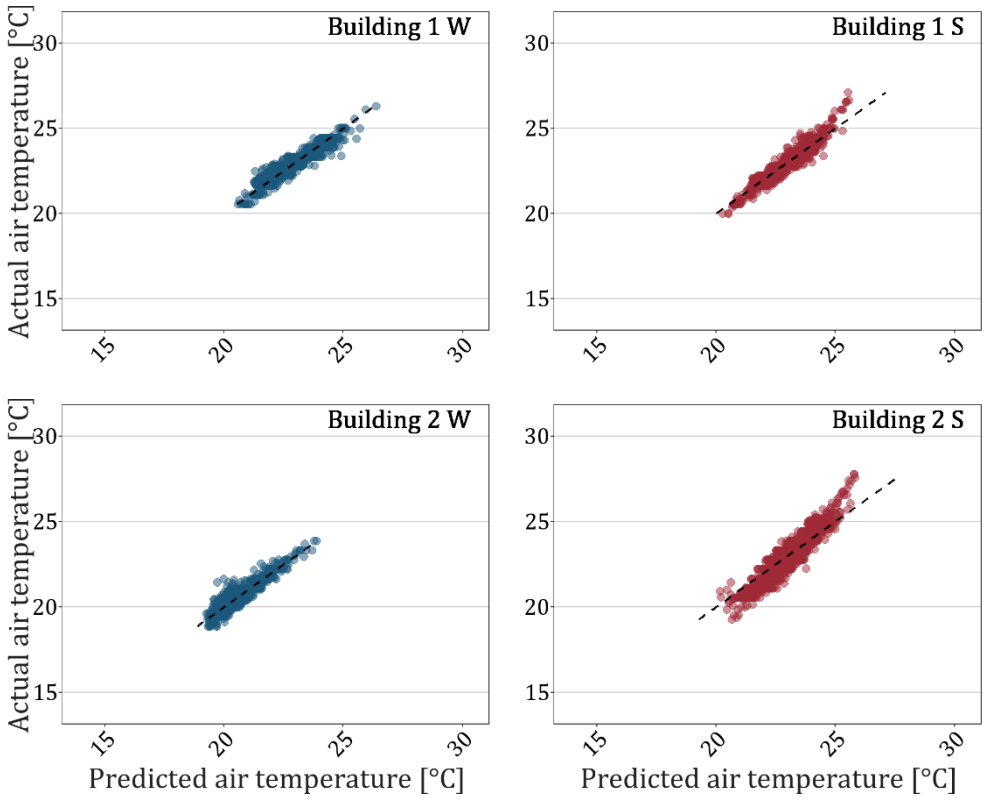


Figure 5.2: Correlation between actual and predicted indoor air temperature.

Parameter	Building 1 W	Building 2 W	Building 1 S	Building 2 S
RMSE	0.32 °C	0.31 °C	0.33 °C	0.60 °C
MAE	0.25 °C	0.24 °C	0.26 °C	0.50 °C
R^2	0.90 °C	0.85 °C	0.93 °C	0.89 °C

Table 5.2: Error metrics for LSTM neural networks of the buildings surveyed.

In addition to this, the high value taken by R^2 suggests that the fitted regression line is near all the points standing for records, as we can see from Figure 5.2. The error occurring in all cases is pseudo-normal distribution, as shown by the Figure 5.3. This is evidence of the high accuracy of neural networks of predicting indoor air temperature.

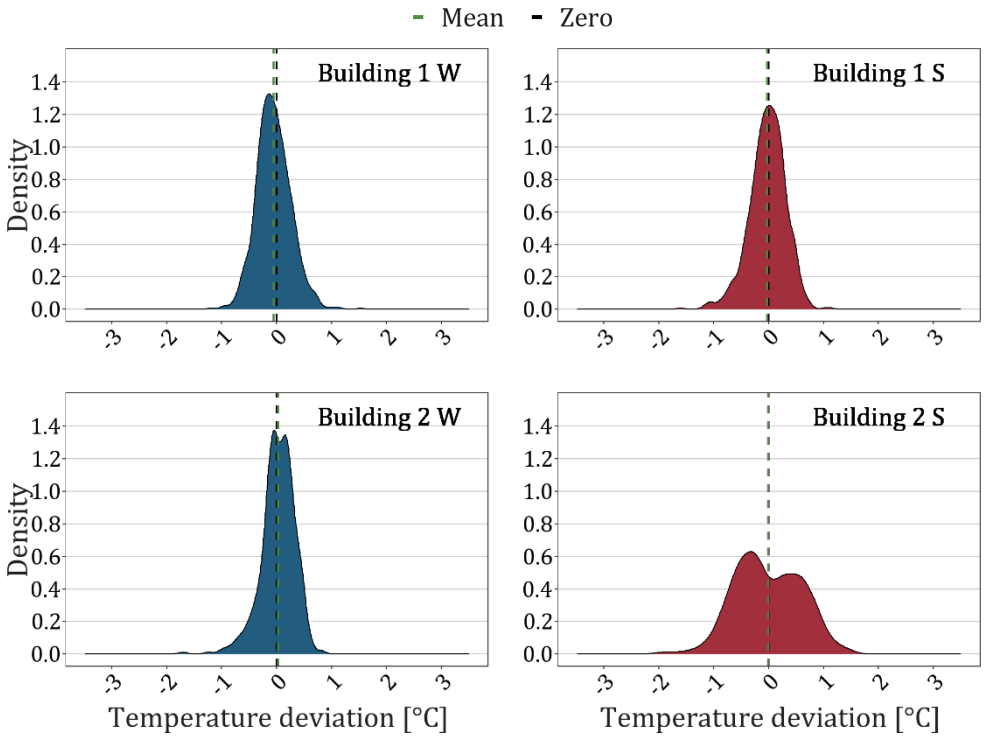


Figure 5.3: Error distributions between predicted and actual indoor air temperature.

Slight shifts to the right of the error distributions for data about *Building 1 S* and *Building 2 W* mean that the models have a tendency to overestimate indoor air temperature. On the other hand, the leftward shifts of the same kinds of distributions for data regarding other two suggest a slight underestimation of the same variable compared to the true value. Other models trained and tested also showed symmetric distributions with mean very close to zero.

5.1.1 Model comparison

To evaluate the performance of their model applied to each building in the North America measurement dataset, Wang et al. [24] assessed the value of R^2 and RMSE for each application. To compare these models with neural networks we developed for buildings belonging to the state of California, the values of the error metrics were sampled. Compared to the total number of buildings in the starting datasets, the authors built their model for only a fraction of these. By looking only at the buildings for which values of thermal time constant and equivalent temperature related to internal and solar heat gains are available, it is possible to extract their metrics and compare them with those of our models. To make a consistent comparison, we recalculated the error metrics considering both the predictions obtained for the winter season and those related to the summer season. All this is shown in Figure 5.4.

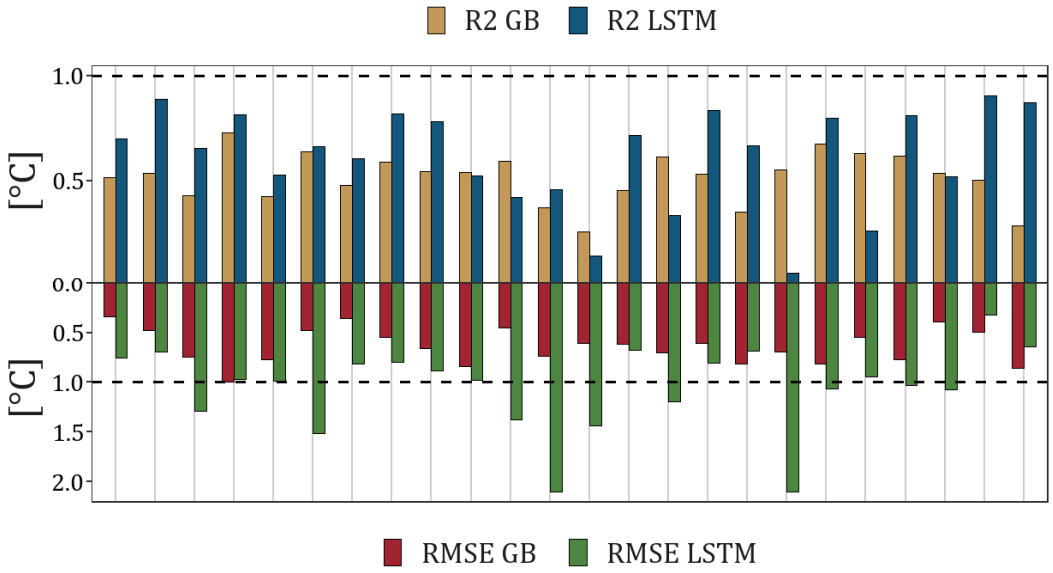


Figure 5.4: Model comparison (GB refers to the grey-box models developed by Z. Wang et al. [24]. LSTM, on the other hand, refers to the models we implemented).

In all cases, the R^2 value of the neural network models is higher than that of the grey-box models developed for the same buildings. Conversely, it often happens that the RMSE of the grey-box models is higher than that of our models. Generally, it is found that the models proposed by Wang et al. [24] in their study has better error metrics than ours. This could be due to some factors affecting the thermal dynamics of the building which were not taken into account by us. The authors probably employed additional information from the measurement campaign, since there are

multiple datasets coming from it, and also data referring to longer time horizons. An improvement of our ANN models could be useful to obtain similar or higher quality results than grey-box models. Our models can understand the wide variability between minimum and maximum, but with worse average detail. Nevertheless, our average accuracy is lower. It is also true that for only 24 out of 50 buildings under analysis we know the error metrics of the models developed by Z. Wang et al. [24]. All of these have an RMSE of less than 1 °C. In our case, on the other hand, 28 buildings compared to the total investigated have an RMSE lower than the target value. However, given the better performance of the models proposed by the other authors, this will be used for simulation of the different scenarios.

5.2 First scenario

The *first scenario* describes the procedure of sizing a MGT according to a centralized scheme and analyses its integration with renewable energy generation systems arranged at the level of each building. All results submitted are derived from a comparison with the *baseline scenario*[17], which is taken as a reference. The duration curves of load are determined by arranging the electrical powers in descending order according to the 1332 annual operating hours of the plant and are shown in Figure 5.5. Since peak electrical load values occur for only a few hours compared to the total operating time of the plant, this would result in a considerable share of electricity fed into the power grid in the case of its operation at full cogeneration regime under rated operating conditions. Although, the feed-in of energy to the power grid is submitted to economic enhancement from the Energy Services Manager, it is more convenient to share energy within the community since it is possible to take advantage of greater benefits related to the promotion of alternative generation solutions. In addition to this, adjusting the system to meet loads lower than the peak value is not cost-effective as this results in a significant reduction of expected performance. To overcome all these issues, the optimal plant size was determined as a function of electrical output, which we conventionally set at three-quarters of the operating hours. For those hours, during the heating season, when production is in deficit with respect to the electrical

[17] The baseline scenario involves the inclusion of a number of energy conversion technologies that can meet the needs of each building, including an air-to-air HP serving the thermal zones for controlling their temperature and, depending on the building, an EH or GB are used for the production of DHW. In addition, electricity and heat storage technologies are integrated. For a set percentage of buildings, PV systems are arranged.

load, the withdrawal from the power grid is provided (blue band of Figure 5.5). For the remaining time frame, the micro-cogeneration plant is able to meet the electrical load of the EC (red band of Figure 5.5) and, in the case of surplus production, the transfer to the power grid is envisaged (grey band of Figure 5.5).

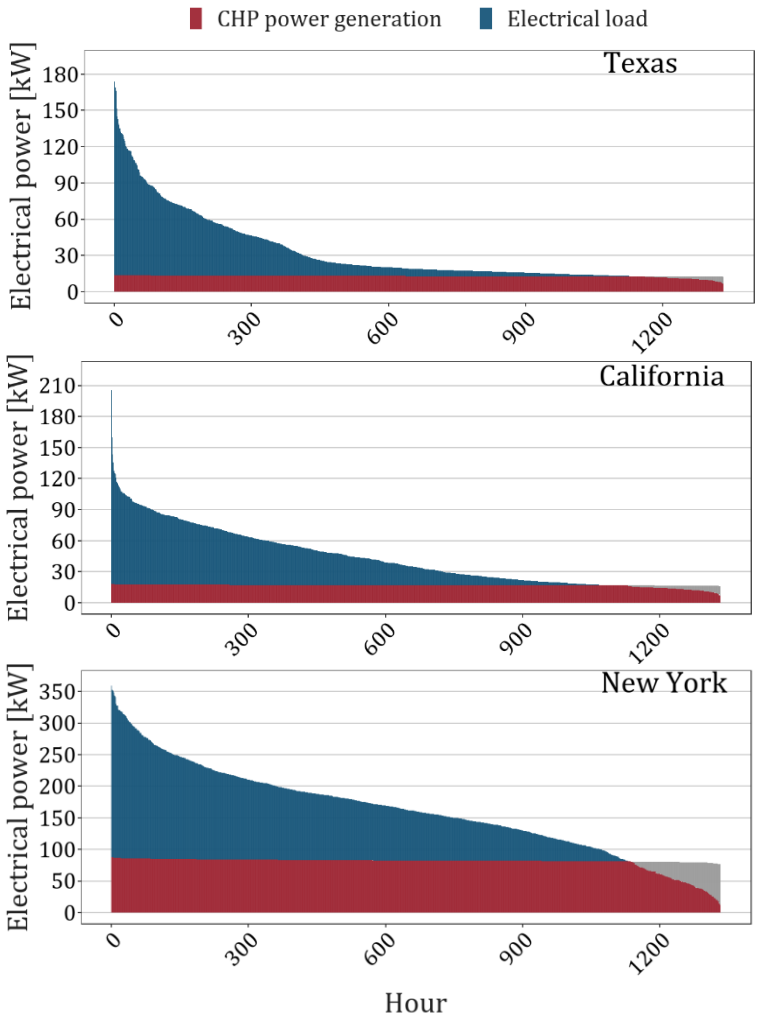


Figure 5.5: Electrical load duration curve (blue) and electrical generation curve (red) for the cluster of buildings belonging to the three different ECs.

Therefore, this analysis allowed defining the size of the micro-cogeneration plant subsequently expressed through a percentage of the electrical load rated sum of air-to-air HPs. This, as a matter of fact, is regarded as a benchmark for the whole community since represent the highest consumption contribution for each building in the basic configuration. Specifically, the size of the MGT was set at 6% of the peak electrical load of HPs for Texas, 8% for California and finally for New York the optimal size is 31%. The micro-cogeneration plant sized in this way allows an average of

12.2% of the heat load of all building to be met for Texas, 15.7% for California and 15.8% for New York. The supplementary heat load is thus offset by the air-to-air HPs arranged at the individual building level. A properly sized μ-CHP system can realize potential energy savings and a subsequent decrease in energy withdrawal from the power grid. These results are shown in Figure 5.6, through the distribution of the change in net electrical load referred to the three different communities. The net electrical load, defined as the difference between electricity consumption and local generation, ranges from -58 kWh to 240 kWh in the basic configuration for the ECs analysed.

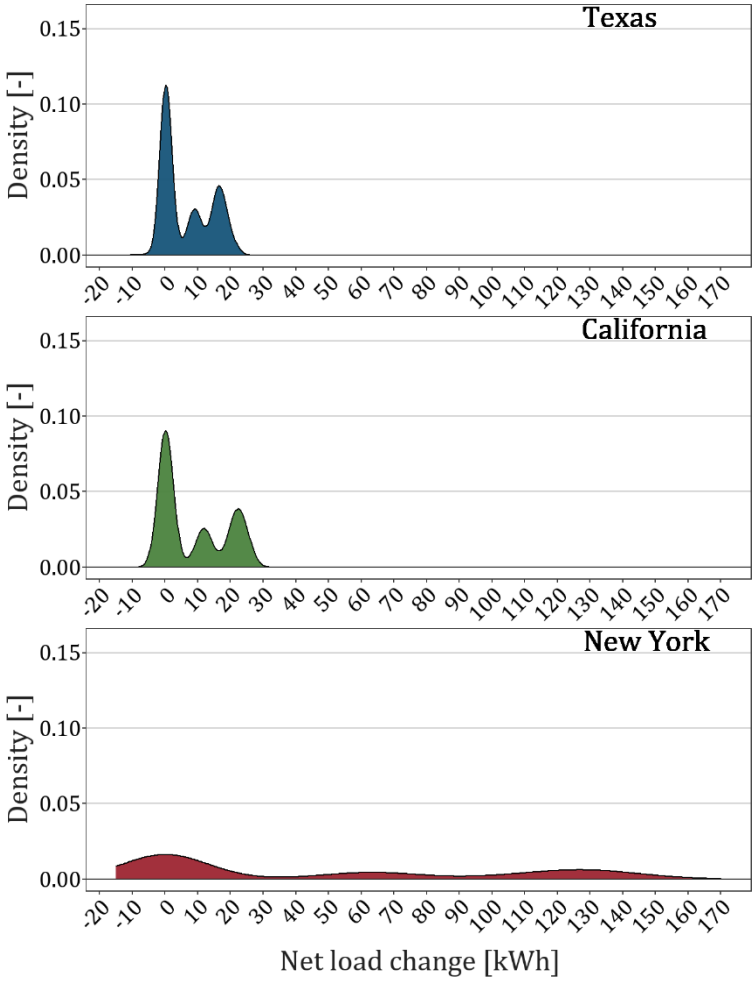


Figure 5.6: Density distribution of net electric load difference for the cluster of building belonging to the three different ECs.

Based on these data obtained from the simulations, we can infer a percentage reduction in the net electrical load of entire communities – resulting from the introduction of the MGT – equal to 23.6% for Texas, 22.2% for California and 27.4%

for New York. This strategy makes it possible to reduce electricity withdrawal from the power grid by an average of 24%, since the net electrical load is lowered. Therefore, together with the energy savings achieved, there is a reduction in the purchase cost of electricity, as well as a benefit related to economic enhancement for the use of locally arranged technologies and an economic compensation for electricity fed into the power grid. Although it is not evident from Figure 5.6, operation under nominal conditions of the μ-CHP system during the established time intervals results in a peak in the net electrical load distribution that is lower than in the base case, as on-site generation increases for the same amount of electric consumption by community members. In addition to this, an analysis of the electrical load distribution, when the μ-CHP system is operating, shows no peaks at high electric load values. This is related to the fact that during the night-time period, when no local generation system is serving the buildings and the net load matches with the members electrical load, consumption is low. The cost of taking electricity from the power grid for the entire EC is reduced in average by 27.5%. However, it is critically important to observe the cost increase incurred at the community level to procure the natural gas feedstock for the micro-cogeneration plant.

A lower share of electricity withdrawn corresponds to a lower emission of climate-altering substances produced for its generation. Although the MGT is fossil fuel consuming, an overall reduction in CO_2 emission to the atmosphere is observed. Analyses have shown a net reduction ranging from 0.1% to 7.3% depending on the EC observed. Regarding the individual units belonging to the ECs, Figure 5.7 shows the reduction in electricity purchase costs compared to the CO_2 savings resulting from the adoption of this plant system.

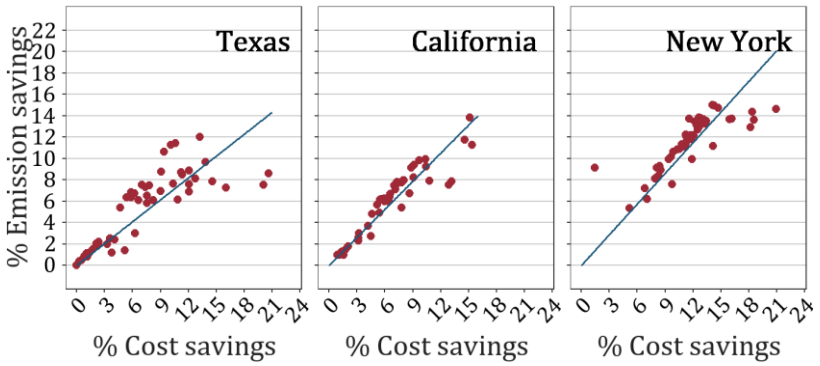


Figure 5.7: Percentage reduction in electricity costs compared with the reduction of CO_2 emitted for the cluster of building belonging to the three different ECs.

A linear relationship is observed between the reduction in electricity cost and the corresponding CO_2 emission for each building. As to this matter, more detailed data are given in Table 5.3, depending on the location.

Parameter	**Texas**	**California**	**New York**
Cost percentage reduction	8.6%	6.6%	11.7%
CO_2 percentage reduction	5.7%	6.0%	11.5%

Table 5.3: Average percentage cost reduction and CO_2 emissions for each building belonging to the three ECs.

Indeed, assuming a virtual sharing scheme, it is as if the individual building is not attributed the share of CO_2 emitted for energy production from the MGT serving the entire EC. In truth, the additional share of climate-altering emissions linked to the use of this plant system is actually considered as attributed to the entire community.

5.3 Second scenario

The *second scenario* assesses the impact of consumption units such as EVs in terms of electrical absorption, increased emission of greenhouse gases due to surplus generation and increased costs of energy purchase. Two different analyses were conducted, considering two fractions of users owning these consumption units corresponding to 50% and 100% of the members belonging to the EC, respectively. All the results obtained were compared with the *baseline scenario*.

Electricity consumption by each of the three ECs examined is between 9.9 kWh and 239 kWh in the base case. The percentage increase in electricity uptake ranges between 34% and 142% assuming that half of the members of each community have an EV and between 65% and 273% assuming that all community members have one. These results are shown in Figure 5.8, which appear the distribution of total electrical load for the different communities under analysis. The median of the distributions, represented by the vertical bar, takes on an ever-increasing value as the percentage of EVs in use by community members increases. The peak value of the distribution also gradually shifts towards higher electrical loads. The increase in electricity drawn from the power grid also leads to a growth in the emission of climate-altering substances linked to its generation.

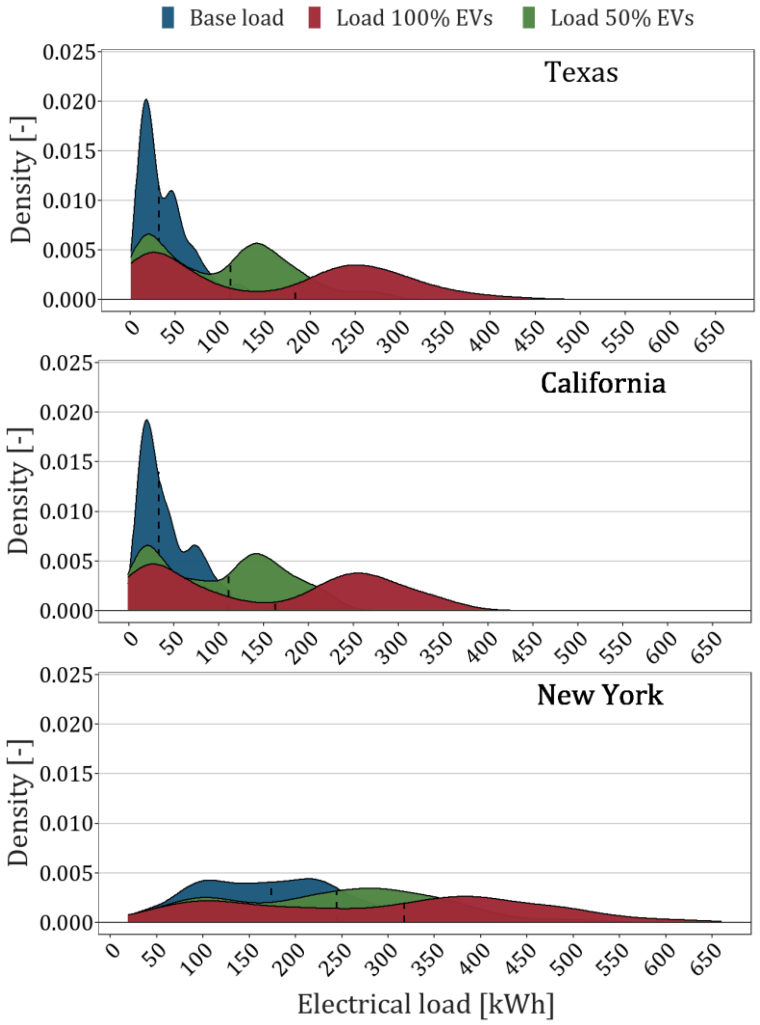


Figure 5.8: Density distribution of electrical load for the cluster of buildings belonging to the three different ECs.

Specifically, in the event that 50% of the members of the EC are equipped with an EV, the overall emission of climate-altering substances increases by an average of 115%, whereas if all members own an EV, the increase is worth about 224%. The assumption behind these considerations does not take into account the reduction in CO_2 emitted due to the non-use of fossil fuels by EVs. Figure 5.9 shows the distribution of CO_2 emissions among buildings belonging to the EC as the percentage of EVs changes. Table 5.4 gives the relative increase in CO_2 emitted for each location, both in the case were 50% of the members belonging to the EC own an EV and in the case where all members are equipped with EVs.

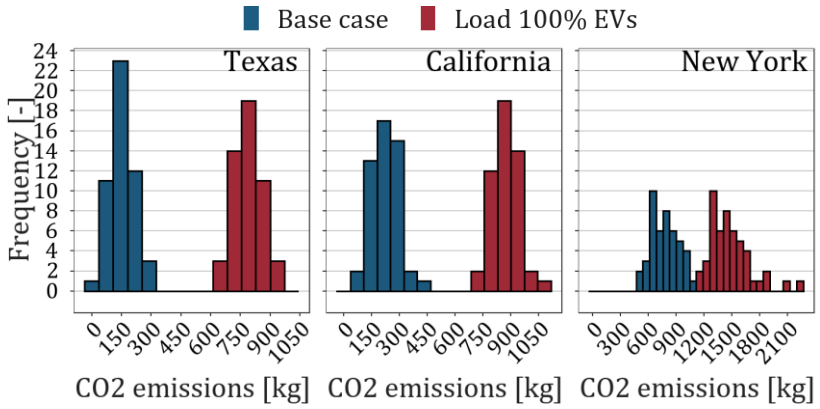


Figure 5.9: Frequency distribution of CO_2 emissions for the cluster of buildings belonging to the three different ECs.

The percentage emission increase values differ by locality based on the energy source used for electric generation. To be more specific, in California and New York most electricity is generated from natural gas, followed by coal. In the other state, however, there is a fraction of generation coming from renewable sources for which the emission of climate-altering substances is zero, including hydroelectric and wind power.

Parameter	**Texas**	**California**	**New York**
CO_2 percentage increase – 50% EV	162%	150%	35%
CO_2 percentage increase – 100% EV	317%	289%	67%

Table 5.4: Percentage increase in CO_2 emission.

Finally, the higher emissions and higher baseline costs related to the EC located in New York result in a lower percentage change in these variables than in the other ECs.

5.4 Third scenario

The *third scenario* is built on the basis of the combination of the two previously described above and aims to evaluate the effect in terms of energy costs and climate-changing emissions of the use of the micro-cogeneration system also for the fulfilment of the electrical load related to EVs.

As a matter of fact, we have carried out an appropriate analysis of the reduction in the costs of withdrawing electricity from the power grid and the saving of CO_2 emitted into the atmosphere, which can be achieved by using the micro-cogeneration system also to meet the load of EVs compared to selling the surplus electricity to the power grid and the simultaneous purchase from EVs. Figure 5.10 shows the results obtained by considering the two different penetration rates of EVs within the buildings of the ECs examined.

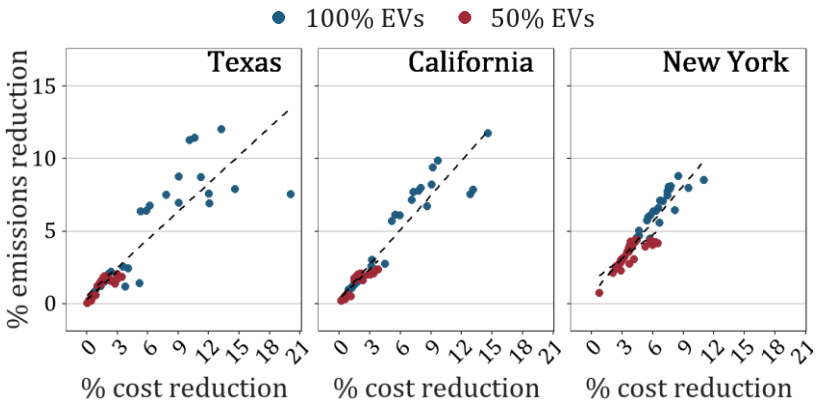


Figure 5.10: Percentage reduction in electricity costs compared with the reduction of CO_2 emitted for two different penetration rates of EVs by ECs.

Compared with the case of purchasing the entire share of electricity to meet the buildings load from the power grid, the configuration which involves the use of the MGT for this purpose as well yields significant economic and environmental benefits for each building belonging to the EC, described in percentage terms in the Table 5.5. Both the percentage reduction in CO_2 emitted and the percentage reduction in cost are greater in the case of 50% of EVs. This is justified by lower energy withdrawal from the power grid due to lower consumption for the same local production.

Parameter	**Texas**	**California**	**New York**
Cost percentage decrease - 50% EV	4.2%	3.8%	8.9%
CO_2 percentage decrease - 50% EV	3.3%	3.3%	8.6%
Cost percentage decrease - 100% EV	1.3%	1.6%	6.6%
CO_2 percentage decrease - 100% EV	1.2%	1.4%	6.2%

Table 5.5: Reduction of electricity costs and CO_2 emitted for two EV scenarios.

In addition to this, compared with an average reduction of 14.7% in the case of half of user-owned EVs and 10.8% in the case of 100% EVs in the electricity-related costs of the entire EC, there is an average reduction of 2.8% in the case of 50% of EVs and 2.2% in the case of 100% of EVs in the emissions of climate-changing substances.

5.5 Key Performance Indicators

KPIs for assessing the impact of new integrated energy technologies and comparing different proposed scenarios include SCI and SSI. In more detail, SCI indicates the share of energy generated and consumed directly locally, while SSI refers to the demand for energy met by on-site arranged generation systems compared to total demand. Figure 5.11 shows a comparison of the value of SCI and SSI for the three scenarios analysed, referring to the three ECs and their buildings.

Compared with the base case, the presence of the micro-cogeneration plant significantly affects these two indicators, resulting in an increase since the sharing of energy within the EC can increase in conditions of greater generation. In the case where EVs are also included as consumption units of the EC, the SCI undergoes a further increase compared to the other two scenarios as the sharing of energy produced on site is maximized. The percentage increase in SSI in the presence of the MGT is 210% on average, while the increase in SCI is 385%.

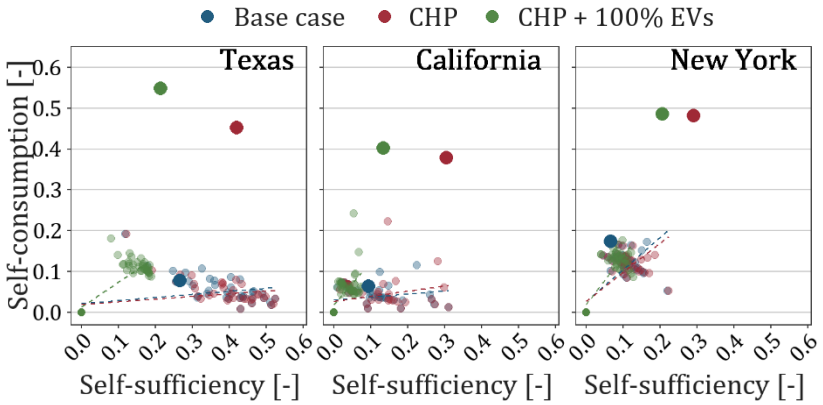


Figure 5.11: SCI vs SSI for individual buildings (smaller dots) and the entire EC (larger dots) with reference to the three different scenarios.

As regard to individual buildings, we can count the FF. This indicator is related to the electrical energy consumed during peak hours and during off-peak hours.

Boxplots related to the three scenarios referring to the buildings belonging to the community are represented in Figure 5.12.

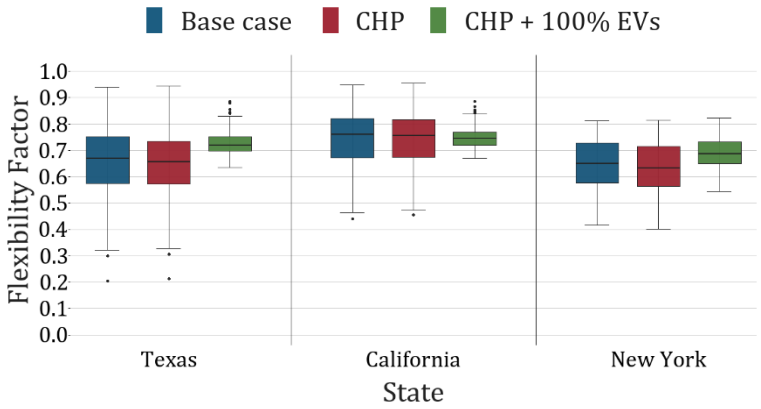


Figure 5.12: FF of buildings for three different scenarios and for three ECs.

The lower and upper part of the box represent the first and third quartile, respectively, while the middle line refers to the median of the data. On average, a reduction in this indicator is observed by moving from the first scenario to the third scenario, as there is a reduction in the electrical load $E_{el_{off\ peak}}$. Certainly, the high values of the FF are linked to a considerable contribution from the PV system able to produce a good share of electricity during peak hours.

5.6 Discussion

The results obtained from the LSTM neural network models and the three simulated scenarios will be discussed below in order to identify some key points of the conducted analyses.

5.6.1 Critical remarks on the results of Long Short-Term Memory networks

One of the main purposes of the tool employed is to develop LSTM neural network models which can predict the indoor air temperature changes of buildings located in a North American state. The data deriving from this our investigation can be effectively applied within the VSE, to estimate energy consumption rates related to indoor air temperature control and apply control strategies which can satisfy multi-objective functions. Preliminary analyses conducted on the datasets showed that making as input data such as indoor air temperature, outdoor air temperature,

occupancy schedule, normalized heat output, cyclic temporal variables and solar irradiance can be useful in predicting indoor air temperature with acceptable accuracy.

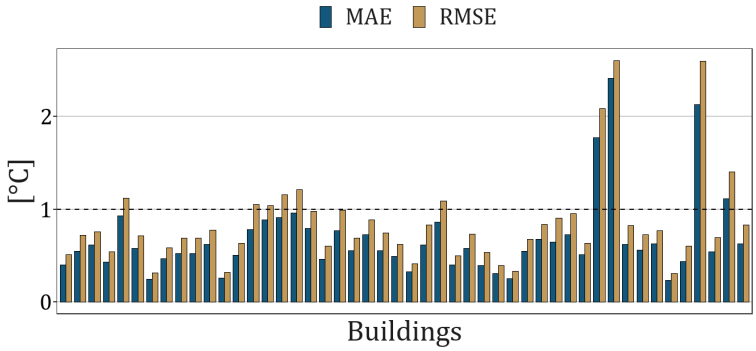


Figure 5.13: Comparison of error metrics related to all analysed buildings for winter months.

With respect to all neural networks tested following training based on data from the winter months, for thirty-eight of these the error metrics demonstrated good model performance in the testing stage. Figure 5.13 compares the MAE and the RMSE error metrics for all buildings examined for the heating season. An average RMSE of 0.85 °C is observed. As a maximum threshold for our analysis, we assumed acceptable only models with this index less than 1 °C. This comparative analysis was also carried out for the simulated buildings during the summer season, and the results are shown in Figure 5.14. Compared with the winter case, in the summer time frame fewer models – equal to twenty-seven – are able to predict the outdoor air temperature with high accuracy. In this case, the mean value assumed by RMSE is 1.10 °C, which is higher than the winter instance.

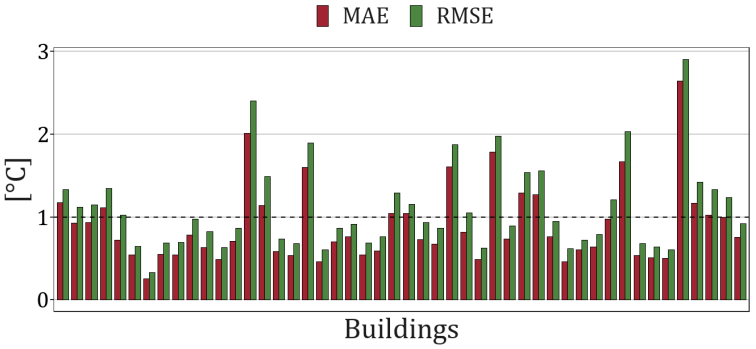


Figure 5.14: Comparison of error metrics related to all analysed buildings for summer months.

This study revealed shortcomings in the measurement campaign from which the data used for training LSTM neural networks are derived, linked, for example, to the absence of monitoring data related to the operational status of energy systems. This greatly constrains the development of good models, since the feature to be predicted is strongly influenced by these data. For this reason, in order to obtain adequate performances, additional explanatory variables affecting the variation of indoor air temperatures, such as normalized power values, were included. The limited availability of observations is one of the reasons why neural network training was effective for some buildings and not for others. In addition to this, a tendency to track temperature evolutions was shown for all our models, but a bad prediction was obtained due to the failure to replicate peak or valley values of indoor air temperature.

5.6.2 Critical remarks on the results of the three scenarios

The development of different energy scenarios has the purpose of describing ways to manage energy flows, providing a scalable and adaptable tool in different contexts for different activities, including the design phase. The configuration of energy systems in the first scenario implies advantages both from an energy point of view and from economic and environmental ones, with reference to the withdrawal of energy from the power grid. The option of the second scenario is shown to be hardly feasible at community level, because of the high shares of electricity to be taken from the power grid, to which high purchasing costs and emissions of climate-changing substances are connected. The current outlook predicts an increasing adoption of these consumer units. The electrical absorption peaks which occur on the power grid for their charging could lead to a violation of resilience and stability requirements, therefore a management of simultaneous loads is advisable. Finally, the configuration envisaged in the third scenario acts as a good compromise between the shares of electricity sold to the power grid under the first scenario and the need to purchase electricity from the power grid under the second scenario. Table 5.6 gives a summary of the main energy, economic and environmental parameters analysed, referring to the three scenarios. In more detail, we define at the community level the net electricity load, the cost of purchasing electricity, the consumption of natural gas and its cost and finally the share of CO_2 emitted, determined by including both the

emissions related to the withdrawal of energy from the power grid and the emissions due to the combustion of natural gas in the μ-CHP plant.

Parameter		**Texas**	**California**	**New York**
Net electrical load [kWh]	Base case	20 872	28 687	127 508
	First scenario	15 917	22 316	92 628
	Second scenario	105 602	113 500	212 350
	Third scenario	100 684	107 125	177 471
Electricity cost [$]	Base case	2 772	3 450	16 922
	First scenario	2 108	2 506	11 641
	Second scenario	12 695	13 898	27 404
	Third scenario	11 965	12 914	22 060
Consumption of natural gas [kWh]	Base case	828	828	828
	First scenario	15 068	19 218	84 678
	Second scenario	828	828	828
	Third scenario	15 068	19 218	84 678
Natura gas cost [$]	Base case	69	69	69
	First scenario	999	1 270	5 544
	Second scenario	69	69	69
	Third scenario	999	1 270	5 544
CO_2 emissions [kg]	Base case	10 023	11 224	48 918
	First scenario	10 012	10 871	45 346
	Second scenario	41 786	43 655	81 433
	Third scenario	41 531	43 183	77 663

Table 5.6: Relevant parameters obtained for three ECs.

In the second and third scenarios, the reduction in CO_2 emissions due to the adoption of EVs instead of combustion engines is not taken into account. The

integration of the micro-cogeneration plant affects the self-sufficiency and power grid independence of the EC. As a matter of fact, the data for the first and third scenarios attest to a reduction in net electricity, which is also followed by a reduction in electricity costs taken from the power grid. The latter result is linked to the fact that, although electricity exported to the power grid is subject to pricing related to on-site exchange, its purchase from the power grid has a higher cost. From an environmental point of view, the micro-cogeneration system uses a fossil fuel so, for the same electricity demand, the reduction in emissions from the entire EC could be justified by generating the electricity taken from the power grid derived by a good share of fossil sources, marked by higher carbon intensity. A more detailed analysis of the energy sources used for power generation suggests that the deviation of climate-changing emissions in all North American states surveyed is justified by an energy mix, during the simulation period, in which coal is one of the largest energy supply, followed by natural gas and then renewable energy sources. It is also true that the simultaneous generation of heat and electricity has an impact in terms of reducing emissions compared to separate generation, since the two carriers are produced from the same primary energy. Therefore, while there is a reduction in the cost of purchasing electricity and in climate-changing emissions, an increase in overall energy costs occurred due to the purchase of a high share of natural gas.The first scenario and the third scenario demonstrate that a multi-energy approach can result in significant advantages in terms of energy flows management. Specifically, the first scenario minimizes dependence on the local power grid by leveraging on-site generation systems and ensuring a peak-load free profile. In contrast to this, the third scenario takes advantage of local generation systems when economically viable, but this still results in a high net electrical load. The μ-CHP system, sized in the first scenario, is able to meet a very small fraction of the electrical absorption related to EVs, as can be seen from Figure 5.15, in which the electrical load duration curves are shown considering two different penetration rates of these consumption units among the members of the ECs examined.

In the event that ECs had a well-defined percentage of EVs, it might be useful to find the size of the μ-CHP plant by considering their electric consumption. Even in the case of increasing the size of the MGT, given an operation under nominal conditions, there would be a significant transfer of electricity to the power grid. This is linked to the characteristic duration curve which has an elbow point, due to a shift

from high electric consumption shares related to vehicles batteries charging to low levels related to the base electrical load of the EC.

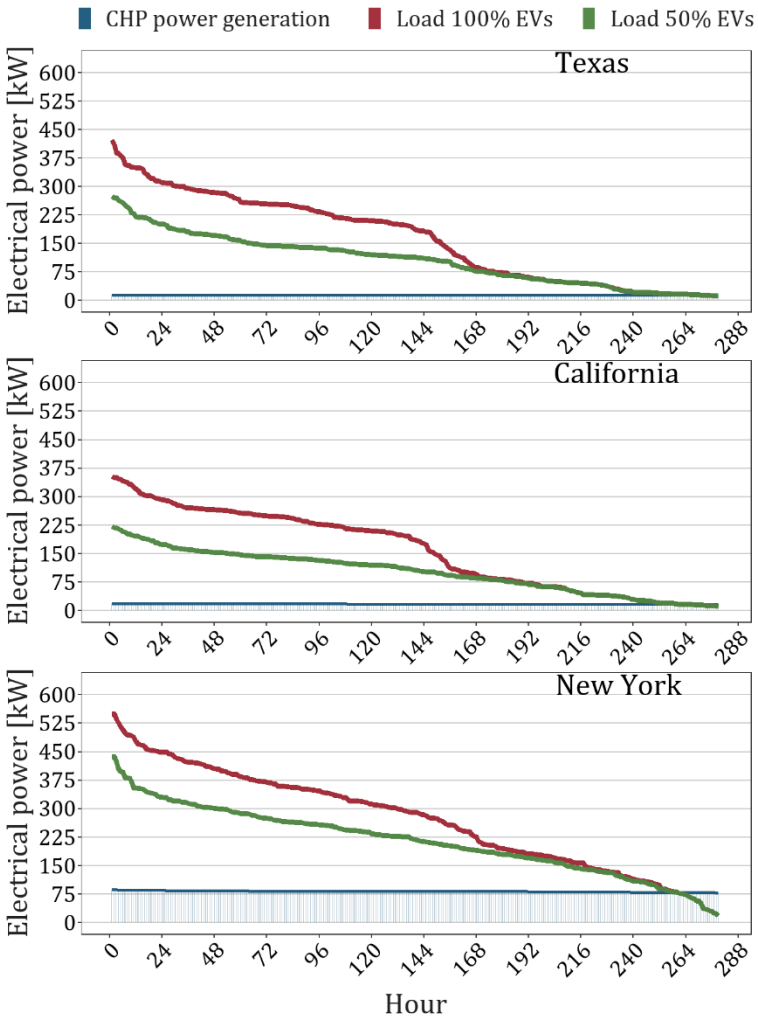


Figure 5.15: Electric load duration curve for different percentages of EVs (red and green) and electric generation (blue) for the cluster of buildings belonging to the three different ECs.

In addition to this, the operating period of the μ-CHP plant is not equal to the time horizon of EVs charging. In the face of a possible change in the size of the MGT, this phenomenon results in an important share of electricity supply to the power grid. In order to maximize energy sharing among members of the EC and take advantage of economic enhancement, it might be profitable to combine the operating hours of the μ-CHP plant with the charging hours of EVs. Since the arrival instant and the departure instant of EVs are not always the same, it is advisable to adopt an appropriate advanced control strategy that combines the two time horizons, so as to

result in significant benefits in terms of energy flow management. Controlling the simultaneity of electrical loads will be a major challenge to allow for the ever-increasing proliferation of electrical consumer units and avoid infrastructure interventions, which could require really large investments.

6 Conclusion

The introduction of ECs leads to new possibilities in the area of coordinated building energy management. VSEs are useful tools for identifying advanced control strategies that can manage the operational status of the variety of energy systems operating on a district scale. This thesis work employed an existing VSE and the main objective was to enhance its ability to simulate different real-world energy contexts. In the starting set-up of the VSE, local power generation occurred only with distributed PV systems available in only a fraction of buildings belonging to the EC. The disposition of BESSs was configured as a means of providing energy flexibility to the power grid. To ensure comfortable conditions inside the buildings, air-to-air HPs were set up that in some cases could handle the charging of TES. DHW production, on the other hand, was done with separate generation systems. With the aim of obtaining accurate prediction of the evolution of air temperature inside buildings and considering their thermal dynamics on a district scale, LSTM neural networks were formulated. In addition to this, to increase self-sufficiency and self-consumption at the local level while minimizing impacts on the power grid, a MGT operation model was proposed. This allows optimal management of a building district's energy flows. Finally, to take into account their increasing proliferation, a model capable of reconstructing energy demand from EVs was introduced. Their inclusion in the VSE made it possible to simulate absolutely real contexts.

The development of neural network models within a VSE has the final goal of estimating the energy consumption of a cluster of buildings. The main purpose of deploying such models is to compare them with existing grey-box models developed for the same buildings, in order to evaluate the performance of both and to identify those that most likely describe the thermal energy dynamics of multiple buildings,

which can take part of a district. Another goal is to make models which are computationally light and scalable to different contexts, suitable for control purposes.

The LSTM neural networks were trained and validated using time series of variables measured in a multiplicity of buildings belonging to the areas of a North American state. After a first phase of data pre-processing, which is useful to remove inconsistencies and reconstruct missing values, the seasonality of the variable to be predicted as a function of time data was addressed. This is required to obtain good results from the ANN downstream of the training phase. In addition to this, the absence of monitored variables related to the operation of energy systems showed itself as a strong constraint of the measurement campaign. As a matter of fact, such information greatly affects the variation of indoor air temperature. To overcome this problem, a methodology capable of reconstructing time series of values related to normalized heating power and normalized cooling power was formulated. An additional explanatory variable was embedded in the dataset derived from the Ecobee DYD program, which concerns solar irradiance as a function of the climatic zone to which the individual building belongs. Finally, since the indoor air temperature change from the winter season to the summer one – for example, the set-point value shifts – an additional sampling of the datasets was shown to be substantial, with the aim of splitting the winter months' time series and the summer ones. So, if we want simulate thermal dynamics of buildings belonging to the same North America state for an entire year, it is required to use two LSTM neural network models, as proven by the case study submitted. Since these algorithms, based on supervised learning techniques, develop mathematical comparative analyses, a minimum-maximum rescaling was taken as optimal to provide data between 0 and 1 to the layers. Here, we found that one of the main problems encountered in the model development of a LSTM neural network is the choice of hyperparameters, since setting them incorrectly could lead to disastrous results. For this reason, a driven research of these factors was carried out through the use of Optuna tool [67]. During the training phase, the ADAptive Moment estimation optimizer was used in order to minimize the MSE. During the deployment phase, neural networks are able to predict a forward step from the data on the twelve backward steps embedded within an unanchored moving window. In the current state, re-training of the models according to a fixed time schedule is not planned, although it might be advisable. Therefore, a multiplicity of buildings belonging to California state were selected. For each of these,

an LSTM neural network was trained for both the winter and summer months. Compared with all the models developed, just under 80% of them showed acceptable performances in winter testing, while about 56% are able to predict the change in indoor air temperature with high accuracy during summer testing. These rates were established downstream of the assessment of RMSE – for which a value far below 1 °C is taken as reasonable – and of MAE. The results of models found to be acceptable were subsequently compared with those derived from the grey-box model proposed by Wang et al. [24], through some error metrics. With regard to RMSE, emerging evidence from this analysis suggests that a LSTM models takes slightly worse performance than the grey-box ones, while the value of R^2 exhibits the opposite behavior – that is, for our models it is higher than for grey-box ones –. Applying the same methodology to similar units with different locations is one of the main advantages of the proposed algorithms.

Next, the focus was on a study of the energy flows affecting an EC. To this, with the aim of reducing power grid dependence, an analysis of the operation of multiple energy systems available within ECs was carried out, in order to describe the benefits and issues to be addressed for their effective integration in real-world settings. Each of the ECs – located in Texas, California and New York, respectively – consists of fifty buildings and the planned simulation time horizon is from 10th January to 10th February, so analyse the heating season in which the μ-CHP system is operating. In the base case, an air-to-air HP system controls the air temperature inside the buildings, so as to keep it within an acceptable range. TESs are arranged in some buildings and have the function of storing heat under specific conditions defined by RBC strategies, so as to better manage energy flows. In addition to this, depending on the building, an EH or a GB provides for DHW demand. A PV system is able to generate electricity that can be directly exploited on site and a BESS stores electricity in case of surplus production. Three different scenarios were implemented and compared against the base case in order to identify the configuration that can better manage energy flows. In the first stage, a centralized μ-CHP system is placed to serve the entire EC, so as to mainly meet the electricity deficit which occurs when the PV system is not producing and the price of electricity from the power grid is high. The MGT is managed according to the most common control strategy available in the literature, which is to turn it on only at well-defined time intervals. In a second step, the energy, economic and environmental effects of consumption units such as EVs is analysed. At

the energy level, it is crucial to assess the electrical consumption of such systems to avoid impacts and provide benefits to the power grid, so an instant of start charging and an instant of end charging must be identified. These data were determined from a model available in the literature. The combination of the different simulated scenarios has allowed, through KPIs, to highlight a number of issues related to the introduction of EVs, as their charging has a significant impact on the power grid. Finally, the third scenario emerges as a potential compromise solution between the two seen before. It is clear, however, that the micro-cogeneration system is particularly undersized to meet the electrical load of ECs whose members own EVs. Nevertheless, increasing the capacity rating of the MGT requires careful preliminary economic analysis, because of the high costs of this technology and the significant payback time. In the first and third scenarios, community-level economic benefits for the purchase of electricity and also reductions in climate-changing emissions are observable. The latter results is in compliance with the sustainability principles established by the legislative framework. However, there is a significant increase in natural gas purchase costs, which results in an increase of overall cost incurred to meet the community's energy needs. ECs are growing realities, but they require careful management of energy flows to achieve the goals behind their development.

In future work, regarding the training and validation of LSTM neural network models, a comparison of the models presented in this work with ones obtained by simulating a closed-loop controller – that is relies on the predicted values to predict future values – could be carried out. In addition to this, since information regarding the floor area or age of each building are available, neural network models applied at the district level could be developed based on the present structure of the algorithm. On the other hand, as far as VSE is concerned, this could be used to manage the production and consumption of energy systems in a different way. For example, at present, all buildings belonging to the EC are connected to the DHN and are equipped with an air-to-air HP which control indoor air temperature. Depending on the needs of the individual building belonging to the cluster, a choice could be made to have HP technology or to be served only by the DHN. Another development could concern the use the μ-CHP plant for DHW production as well. Given the scalable nature of the VSE, this could also be used for groups of buildings belonging to other locations, as long as monitoring data are available.

Bibliography and sitography

[1] 'Climate Change 2021 – Le basi fisico-scientifiche: i messaggi principali del rapporto', *IPCC – Focal Point Italia*, https://ipccitalia.cmcc.it/messaggi-chiave-ar6-wg1/

[2] Conference of the parties, 'Kyoto Protocol to the United Nations framework convention on climate change', (Dec. 1997), https://unfccc.int/resource/docs/convkp/kpeng.pdf

[3] UNFCCC, 'Nationally Determined Contributions (NDCs).', https://unfccc.int/process-and-meetings/the-paris-agreement/nationally-determined-contributions-ndcs/nationally-determined-contributions-ndcs

[4] European Commission, 'Clean energy for all Europeans package', (May 22, 2019), https://ec.europa.eu/info/news/clean-energy-all-europeans-package-completed-good-consumers-good-growth-and-jobs-and-good-planet-2019-may-22_en

[5] Consiglio dell'Unione Europea, 'Direttiva (UE) 2018/2001 del Parlamento Europeo e del Consiglio dell'11 dicembre 2018 sulla promozione dell'uso dell'energia da fonti rinnovabili', (2018), https://eur-lex.europa.eu/legal-content/IT/TXT/PDF/?uri=CELEX:32018L2001

[6] Parlamento Europeo e Consiglio dell'Unione Europea, 'Direttiva (UE) 2019/944 del Parlamento Europeo e del Consiglio del 5 giugno 2019 relativa a norme comuni per il mercato interno dell'energia elettrica e che modifica la

direttiva 2012/27/UE', (Jun. 2019), https://eur-lex.europa.eu/legal-content/IT/TXT/PDF/?uri=CELEX:32019L0944

[7] A. Caramizaru and A. Uihlein, 'Energy communities: an overview of energy and social innovation', *EUR 30083 EN European Commission*, (Mar. 2020), doi: 10.2760/180576.

[8] MISE, 'Piano Nazionale Integrato per l'Energia e il Clima', (2019), https://www.mise.gov.it/index.php/it/energia/energia-e-clima-2030

[9] ENEA, 'La comunità energetica - vademecum 2021', (2021), https://www.enea.it/it/seguici/pubblicazioni/pdf-volumi/2021/opuscolo-comunita-energetica.pdf

[10] 'Decreto Legislativo 8 novembre 2021, n. 199', (2021), https://www.gazzettaufficiale.it/atto/serie_generale/caricaDettaglioAtto/originario?atto.dataPubblicazioneGazzetta=2021-11-30&atto.codiceRedazionale=21G00214

[11] GECO & ENEA, 'Le comunità energetiche in Italia', 1–36 (Oct. 2020), doi: 10.12910/DOC2020-012.

[12] 'Piano Nazionale di Ripresa e Resilienza', 119–156 (2021).

[13] S. Ø. Jensen *et al.*, 'IEA EBC Annex 67 Energy Flexible Buildings', *Energy and Buildings* **155**, 25–34 (2017), doi: https://doi.org/10.1016/j.enbuild.2017.08.044.

[14] B. P. Esther & K. S. Kumar, 'A survey on residential Demand Side Management architecture, approaches, optimization models and methods', *Renewable and Sustainable Energy Reviews* **59**, 342–351 (Jun. 2016), doi: 10.1016/J.RSER.2015.12.282.

[15] P. Warren, 'A review of demand-side management policy in the UK', *Renewable and Sustainable Energy Reviews* **29**, 941–951 (2014), doi: 10.1016/J.RSER.2013.09.009.

[16] J. Clauß, C. Finck, P. Vogler-Finck & P. Beagon, 'Control strategies for building energy systems to unlock demand side flexibility – A review', *Proceedings of the 15th IBPSA Conference San Francisco*, (2017), https://www.researchgate.net/publication/324877113_Control_strategies_for_building_energy_systems_to_unlock_demand_side_flexibility_-_A_review

[17] H. Johra, A. Marszal-Pomianowska, J. R. Ellingsgaard & M. Liu, 'Building energy flexibility: A sensitivity analysis and key performance indicator comparison', *Journal of Physics: Conference Series* **1343(1)**, 012064 (Nov. 2019), doi: 10.1088/1742-6596/1343/1/012064.

[18] G. Pinto, A. Kathirgamanathan, E. Mangina, D. P. Finn & A. Capozzoli, 'Enhancing energy management in grid-interactive buildings: A comparison among cooperative and coordinated architectures', *Applied Energy* **310(1)**, 118497 (Mar. 2022), doi: 10.1016/J.APENERGY.2021.118497.

[19] F. Amara, K. Agbossou, A. Cardenas, Y. Dubé & S. Kelouwani, 'Comparison and Simulation of Building Thermal Models for Effective Energy Management', *Smart Grid and Renewable Energy* **06(04)**, 95–112 (2015), doi: 10.4236/SGRE.2015.64009.

[20] F. Ceglia, E. Marrasso, C. Roselli & M. Sasso, 'Small Renewable Energy Community: The Role of Energy and Environmental Indicators for Power Grid', *Sustainability 2021* **13(4)**, 2137 (Feb. 2021), doi: 10.3390/SU13042137.

[21] A. Bartolini, F. Carducci, C. B. Muñoz & G. Comodi, 'Energy storage and multi energy systems in local energy communities with high renewable energy penetration', *Renewable Energy* **159**, 595–609 (Oct. 2020), doi: 10.1016/J.RENENE.2020.05.131.

[22] S. Huang *et al.*, 'An open-source virtual testbed for a real Net-Zero Energy Community', *Sustainable Cities and Society* **75**, 103255 (Dec. 2021), doi: 10.1016/J.SCS.2021.103255.

[23] J. R. Vázquez-Canteli, J. Kämpf, G. Henze & Z. Nagy, 'CityLearn v1.0: An OpenAI gym environment for demand response with deep reinforcement learning',

BuildSys 2019 - Proceedings of the 6th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation, 356–357 (Nov. 2019), doi: 10.1145/3360322.3360998.

[24] Z. Wang, B. Chen, H. Li & T. Hong, 'AlphaBuilding ResCommunity: A multi-agent virtual testbed for community-level load coordination', *Advances in Applied Energy* **4**, 100061 (Nov. 2021), doi: 10.1016/J.ADAPEN.2021.100061.

[25] M. Bianchi, P. R. Spina, G. Tomassetti, D. Forni & E. Ferrero, 'Le tecnologie innovative ed efficienti nei sistemi di generazione in assetto co-trigenerativo e nei sistemi integrati con unità a pompa di calore nelle applicazioni industriali e del terziario'.

[26] M. Pirouti, A. Bagdanavicius, J. Ekanayake, J. Wu & N. Jenkins, 'Energy consumption and economic analyses of a district heating network', *Energy* **57**, 149–159 (Aug. 2013), doi: 10.1016/J.ENERGY.2013.01.065.

[27] A. v. Olympios, A. M. Pantaleo, P. Sapin & C. N. Markides, 'On the value of combined heat and power (CHP) systems and heat pumps in centralised and distributed heating systems: Lessons from multi-fidelity modelling approaches', *Applied Energy* **274**, (Sep. 2020), doi: 10.1016/J.APENERGY.2020.115261.

[28] Å. Nystedt, J. Shemeikka & K. Klobut, 'Case analyses of heat trading between buildings connected by a district heating network', *Energy Conversion and Management* **47(20)**, 3652–3658 (Dec. 2006), doi: 10.1016/J.ENCONMAN.2006.02.030.

[29] G. Prinsloo & R. Dobson, 'Combined solar heat and power with microgrid storage and layered smartgrid control toward supplying off-grid rural villages', *Energy Science & Engineering* **3(2)**, 135–144 (Mar. 2015), doi: 10.1002/ESE3.57.

[30] L. Romero Rodríguez, J. M. Salmerón Lissén, J. Sánchez Ramos, E. Á. Rodríguez Jara & S. Álvarez Domínguez, 'Analysis of the economic feasibility and reduction of a building's energy consumption and emissions when integrating

hybrid solar thermal/PV/micro-CHP systems', *Applied Energy* **165**, 828–838 (Mar. 2016), doi: 10.1016/J.APENERGY.2015.12.080.

[31] H. Ren, W. Gao & Y. Ruan, 'Optimal sizing for residential CHP system', *Applied Thermal Engineering* **28(5–6)**, 514–523 (Apr. 2008), doi: 10.1016/J.APPLTHERMALENG.2007.05.001.

[32] J. Wang, C. Liu, D. Ton, Y. Zhou, J. Kim & A. Vyas, 'Impact of plug-in hybrid electric vehicles on power systems with demand response and wind power', *Energy Policy* **39(7)**, 4016–4021 (Jul. 2011), doi: 10.1016/J.ENPOL.2011.01.042.

[33] A. Foley, B. Tyther, P. Calnan & B. Ó Gallachóir, 'Impacts of Electric Vehicle charging under electricity market operations', *Applied Energy* **101**, 93–102 (2013), doi: 10.1016/J.APENERGY.2012.06.052.

[34] C. K. Ekman, 'On the synergy between large electric vehicle fleet and high wind penetration - An analysis of the Danish case', *Renewable Energy* **36(2),** 546–553 (Feb. 2011), doi: 10.1016/J.RENENE.2010.08.001.

[35] W. P. Schill & C. Gerbaulet, 'Power system impacts of electric vehicles in Germany: Charging with coal or renewables?', *Applied Energy* **156**, 185–196 (Oct. 2015), doi: 10.1016/J.APENERGY.2015.07.012.

[36] K. Arendt, M. Jradi, H.R. Shaker & C.T. Veje, 'Comparative Analysis of White-, Gray-and Black-box Models for Thermal Simulation of Indoor Environment: Teaching Building Case Study', *2018 Building Performance Modeling Conference and SimBuild co-organized by ASHRAE and IBPSA-USA*, 173–180 (2018), https://www.ashrae.org/File%20Library/Conferences/Specialty%20Conferences/2018%20Building%20Performa

[37] S. Royer, S. Thil, T. Talbert & M. Polit, 'Black-box modeling of buildings thermal behavior using system identification', *IFAC Proceedings Volumes* **47(3)**, 10850–10855 (2014), doi: 10.3182/20140824-6-ZA-1003.01519.

[38] G. Pinto, D. Deltetto & A. Capozzoli, 'Data-driven district energy management with surrogate models and deep reinforcement learning', *Applied Energy* **304**, 117642 (Dec. 2021), doi: 10.1016/J.APENERGY.2021.117642.

[39] G. Gokhale, B. Claessens & C. Develder, 'Physics informed neural networks for control oriented thermal modeling of buildings', *Applied Energy* **314**, 118852 (May 2022), doi: 10.1016/J.APENERGY.2022.118852.

[40] L. di Natale, B. Svetozarevic, P. Heer & C. N. Jones, 'Physically Consistent Neural Networks for building thermal modeling: theory and analysis', 1–19 (Dec. 2021), http://arxiv.org/abs/2112.03212.

[41] A. Cielo, P. Margiaria, P. Lazzeroni, I. Mariuzzo & M. Repetto, 'Renewable Energy Communities business models under the 2020 Italian regulation', *Journal of Cleaner Production* **316**, 128217 (Sep. 2021), doi: 10.1016/J.JCLEPRO.2021.128217.

[42] Á. Corredera, A. MacÍa, R. Sanz & J. L. Hernández, 'An automated monitoring system for surveillance and KPI calculation', *EESMS 2016 - 2016 IEEE Workshop on Environmental, Energy, and Structural Monitoring Systems, Proceedings*, 1–5 (Jun. 2016), doi: 10.1109/EESMS.2016.7504806.

[43] D. Chan, M. Cameron & Y. Yoon, 'Implementationof micro energy grid: A case study of a sustainable community in China', *Energy and Buildings* **139**, 719–731 (Mar. 2017), doi: 10.1016/J.ENBUILD.2017.01.055.

[44] J. L. Hernández, A. Macía, A. Vasallo, E. Vallejo, C. Zubia & C. Criado, 'Application of a KPI-Driven Protocol for Sustainability Assessment', *Proceedings* **1(698)**, (Nov. 2017), doi: 10.3390/proceedings1070698.

[45] N. Couellan, 'Probabilistic robustness estimates for feed-forward neural networks', *Neural Networks* **142(113)**, 138–147 (Oct. 2021), doi: 10.1016/J.NEUNET.2021.04.037.

[46] A. Sherstinsky, 'Fundamentals of Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM) network', *Physica D: Nonlinear Phenomena* **404(8)**, 132306 (Mar. 2020), doi: 10.1016/J.PHYSD.2019.132306.

[47] S. Hochreiter & J. Schmidhuber, 'Long Short-Term Memory', *Neural Computation* **9(8)**, 1735–1780 (Nov. 1997), doi: 10.1162/NECO.1997.9.8.1735.

[48] X. Song *et al.*, 'Time-series well performance prediction based on Long Short-Term Memory (LSTM) neural network model', *Journal of Petroleum Science and Engineering* **186**, 106682 (Mar. 2020), doi: 10.1016/J.PETROL.2019.106682.

[49] 'Advantages and Disadvantages of Artificial Neural Networks | Asquero.' https://www.asquero.com/article/advantages-and-disadvantages-of-artificial-neural-networks/

[50] Parlamento Europeo e Consiglio dell'Unione Europea, 'Direttiva 2004/8/CE del Parlamento Europeo e del consiglio dell'11 febbraio 2004 sulla promozione della cogenerazione basata su una domanda di calore utile nel mercato interno dell'energia e che modifica la direttiva 92/42/CEE', (Feb. 2004),https://eur-lex.europa.eu/legal-content/IT/TXT/PDF/?uri=CELEX:32004L0008&from=EN

[51] Y. Hwang, 'Potential energy benefits of integrated refrigeration system with microturbine and absorption chiller', *International Journal of Refrigeration* **27(8)**, 816–829 (Dec. 2004), doi: 10.1016/J.IJREFRIG.2004.01.007.

[52] D. Li, X. Xu, D. Yu, M. Dong & H. Liu, 'Rule Based Coordinated Control of Domestic Combined Micro-CHP and Energy Storage System for Optimal Daily Cost', *Applied Sciences 2018* **8(1)**, (Dec. 2017), doi: 10.3390/APP8010008.

[53] D. Zou, D. Meng, Y. Dai, S. Lu & H. Xie, 'Optimal Charging Strategy of Electric Vehicles with Consideration of Battery Storage', *E3S Web of Conferences* **236**, 02015 (Feb. 2021), doi: 10.1051/E3SCONF/202123602015.

[54] 'Share your anonymized data for good | Donate Your Data | Smart home devices and thermostats | ecobee', https://www.ecobee.com/donate-your-data/

[55] 'Smart Thermostats & Smart Home Devices | ecobee', https://www.ecobee.com/en-us/

[56] U. Jordan & K. Vajen, 'DHWcalc: program to generate Domestic Hot Water profiles with statistical means for user defined conditions', *Proc Sol WorldCongr 2005*, (2005), www.solar.uni-kassel.de.

[57] E. Fuentes, L. Arce & J. Salom, 'A review of domestic hot water consumption profiles for application in systems and buildings energy performance analysis', *Renewable and Sustainable Energy Reviews* **81**, 1530–1547 (Jan. 2018), doi: 10.1016/j.rser.2017.05.229.

[58] R. Hendron & C. Engebrecht, 'Building America Research Benchmark Definition: Updated December 2009', (2010), http://www.osti.gov/bridge

[59] IEA, 'Appliances and Equipment', (2021), https://prod.iea.org/reports/appliances-and-equipment

[60] G. Azzolini *et al.*, 'Aggiornamento delle metodologie per la stima dei consumi energetici nel settore residenziale e analisi dei dati del database delle detrazioni fiscali mediante la metodologia sviluppata', *ENEA*, 8–15 (2018).

[61] J. Vivian, E. Prataviera, F. Cunsolo & M. Pau, 'Demand Side Management of a pool of air source heat pumps for space heating and domestic hot water production in a residential district', *Energy Conversion and Management* **225**, 113457 (Dec. 2020), doi: 10.1016/J.ENCONMAN.2020.113457.

[62] G. Masson -Iea, P. Jose, I. Briano & M. Jesus Baez -Creara, 'A methodology for the analysis of PV self-consumption policies'.

[63] G. Mutani, S. Santantonio & S. Beltramino, 'Indicators and representation tools to measure the technical-economic feasibility of a renewable energy community. The case study of villar pellice (italy)', *International Journal of Sustainable Development and Planning* **16(1)**, 1–11 (Feb. 2021), doi: 10.18280/ijsdp.160101.

[64] U.S. Energy Information Administration, 'Real-time Operating Grid', https://www.eia.gov/electricity/gridmonitor/dashboard/electric_overview/US48/US48

[65] ‘Benchmark Datasets for Buildings.’, https://bbd.labworks.org/ds/bbd/ecobee

[66] ‘EZ Building Climate Zone Finder’, https://caenergy.maps.arcgis.com/apps/webappviewer/index.html?id=5cfefd9798214bea91cc4fddaa7e643f

[67] ‘Optuna - A hyperparameter optimization framework’, https://optuna.org/

[68] E. K. Vakkilainen, ‘Boiler Processes’, *Steam Generation from Biomass*, 57–86 (Dec. 2017), doi: 10.1016/B978-0-12-804389-9.00003-4.

[69] ‘Canadian Solar CS5P-220M [2009]’, https://pvfree.herokuapp.com/pvmodules/1160/

[70] ‘Electricity Data - U.S. Energy Information Administration (EIA)’, https://www.eia.gov/electricity/data.php

[71] ‘ARERA - Condizioni economiche per i clienti del mercato tutelato’, https://www.arera.it/it/dati/condec.htm

[72] R. Boukhanouf, ‘Small combined heat and power (CHP) systems for commercial buildings and institutions’, *Small and Micro Combined Heat and Power (CHP) Systems: Advanced Design, Performance, Materials and Applications*, 365–394 (Apr. 2011), doi: 10.1533/9780857092755.3.365.

[73] Y. Luo, Y. Shi & N. Cai, ‘Distributed hybrid system and prospect of the future Energy Internet’, *Hybrid Systems and Multi-energy Networks for the Future Energy Internet*, 9–39 (Jan. 2021), doi: 10.1016/B978-0-12-819184-2.00002-X.

[74] A. G. Marijuán, G. Etminan & S. Möller, ‘Smart Cities information system - Key Performance Indicator guide’, (Feb. 2017).

[75] S. Verbeke, D. Aerts, G. Reynders, Y. Ma & P. Waide, ‘Final report on the technical support to the development of a smart readiness indicator for buildings’, *European Commission*, (Jun. 2020), doi: 10.2833/41100.

[76] G. Airò Farulla *et al.*, 'A Review of Key Performance Indicators for Building Flexibility Quantification to Support the Clean Energy Transition', *Energies 2021* **14(18)**, 5676 (Sep. 2021), doi: 10.3390/EN14185676.

Acknowledgements

A chiusura di questo lavoro di tesi desidero ringraziare le persone che mi hanno sostenuta e hanno svolto un ruolo determinante per la sua realizzazione. I miei ringraziamenti sono innanzitutto rivolti al Prof. Alfonso Capozzoli per avermi dato la possibilità di svolgere l'intera attività presso il laboratorio BAEDA, costituito da un gruppo di lavoro esperto, dinamico e attivo. L'argomento di questo lavoro è stato decisamente stimolante ed è stato sviluppato efficacemente in un contesto in cui conoscenza e cooperazione sono gli elementi cardine. Ringrazio inoltre Antonio Gallo, Marco Savino Piscitelli e tutti gli altri membri del laboratorio BAEDA per l'interesse provato nei confronti del mio lavoro e per la sensibilizzazione che mi hanno offerto riguardo questo tema.

Un ringraziamento speciale va ai miei genitori, fonte di sostegno e di coraggio. La passione che mi hanno trasmessa per lo studio e i loro sacrifici sono stati fondamentali per completare il mio percorso e raggiungere questo traguardo. Devo molto anche ai miei fratelli, Enrico e Michele, che hanno sempre cercato di insegnarmi ad affrontare qualunque situazione si presentasse e a gestire le mie paure.

Un immenso grazie va a Giovanni, la persona che mi ha incoraggiata sin dall'inizio di questo percorso di studi. Ha trovato sempre le parole giuste per superare i momenti di difficoltà e per mostrate la sua vicinanza. Il suo supporto è stato sostanziale per sostenere alcune mie scelte e la sua presenza fondamentale per distrarmi da brutti momenti. Che questo sia solo una delle tante soddisfazioni che spero di condividere con lui!

Grazie a Vanessa, coinquilina e amica sempre presente. La condivisione dei momenti trascorsi durante questi cinque anni, ci ha fatto ritrovare la familiarità che non abbiamo mai voluto perdere.

Un ringraziamento a Nicoletta, Alessandro, Veronica, Sara e Nino, Rosetta e Giovanni per essere stati sempre presenti durante questo percorso. Le nostre lunghe discussioni sono state una vera fonte di insegnamento.
In questo momento di gioia, un pensiero particolare va ai miei nonni che sarebbero sicuramente felici di festeggiare con me questo importante traguardo.
Infine, vorrei dedicare questo momento a me stessa. La mia determinazione è stata capace di rispondere ad ansie e momenti difficili con sempre più impegno. Ho imparato molto da questo percorso e sicuramente, dopo 5 anni, ne esco come una persona quasi completamente diversa. Che questo traguardo sia solo un punto di partenza!