Master Thesis

MSc Energy for Smart Cities

Analysis of Capabilities of Machine Learning for Local Energy Communities to Provide Flexibility to the Grid

Report

Autor: Valeriia Maksimovich
Supervisor: Andreas Sumper
Co-supervisor: Alejandro Hernandez Mateus

Call: October 2021

ETSEIB

Escola Tècnica Superior
d’Enginyeria Industrial de Barcelona

UPC

Table of content

List of Tables

List of Figures

Abbreviations

VPP – Virtual Power Plant

LEC – Local energy community

DG – Distributed generation

GHG – Greenhouse gas emission

DNI – Direct Normal Irradiation

AI – Artificial intelligence

ML – Machine Learning

DL – Deep Learning

LSTM – Short-Term Memory

RNN – Recurrent Neural Network

P2P – Peer-To-Peer Energy Trading

CEC – Citizen Energy Community

DSO – Distribution System Operator

MS – Member States

SM – Smart Meters

ESS – Energy Storage Systems

EV – Electrical Vehicle

PV – Photovoltaic

PMU – Phasor Measurement Unit

HEMS – Home Energy Management System

DS – Distribution System

LTLF – Long Term Load Forecasting

STLF – Short Term Load Forecasting

NN – Neural Network

TS – Transmission System

TSO – Transmission System Operator

V2G – Vehicle-To-Grid

V2X – Vehicle-To-X

BM – Business Model

MAE – Mean Absolute Error

MSE – Mean Squared Error

RMSE – Root Mean Absolute Error

R2 – R-squared

1. Introduction

Energy market design is changing worldwide. Small-scale low carbon electricity generation or so-called Distributed Generation made it possible for neighboring citizens not only to jointly own and operate microgeneration or storage facilities but also be actively involved in the energy market by selling the excess energy and earn a profit. The thesis is investigating the concept of Local Energy Communities from the current regulatory framework and technical point of view mainly assessing capabilities to be flexible on the energy market, meaning delivering or consuming electricity for maintaining the generation-consumption balance and the required grid frequency.

Nowadays, thanks to smart meters deployment and sensors' measuring capabilities, the ability to gather data from customers up to the service provider have disrupted the electricity sectors, with the opening of new services and markets. This makes it possible to operate with the energy data more freely and frequently than before.

Combining the disruption with new energy data and legislation enabling energy communities operation, this thesis assesses the possibility to make the dispatch and flexibility provision as automatic and “smart” as possible with the help of Artificial Intelligence and Machine Learning techniques. More precisely, the Master’s Thesis is aiming to answer the following questions regarding Local Energy Communities (LECs):

1. What are LECs and what are the positive and negative aspects of LECs' existence?

2. From Local Energy Community to Smart Local Energy Community - Can ML techniques support LEC to automatically dispatch/ feed energy from/to LEC? How can flexibility be used in the context of LEC?

3. What are the market structures and business models for LEC integration? Which energy market players participate in LEC business area?

The thesis is organized as follows: The first part of the thesis is a theoretical part, where the literature review was done. After the general definition and legal introduction in Section 1, 2.1, 2.2 and 2.3. Different advantages and disadvantages of LECs are reviewed in Sections 2.3 and 2.4. The following section of this thesis presents a brief review of the literature on Big Data foundations and techniques (Section 4). The second part of the thesis is a practical experiment where the author works with a dataset from real households, performs basic data visualization tasks, and performs machine learning-based generation forecasting to evaluate flexibility. The methodology and results are explained in Sections 5 and 6. The last subsection of the given thesis compares different market models of LEC in different countries (Section 7). Main contributions, conclusions, and future work are discussed in Section 8. A representative list of references is provided at the end of the thesis.

2. Local Energy Communities – overview

2.1. Legal framework and definition of LECs in the EU

The latest EU energy-market related legislation, which consists of mainly two documents: Regulation on the internal market for electricity (E-Regulation, 2019/943 of the European Parliament) [2] and Directive on common rules for the internal market in electricity (E-Directive, 2019/944/EU of the European Parliament) [1], introduce new actors and define their rights and obligations as well as define the legal framework for their operation to ensure fair treatment at the energy market. These new energy market participants are the following: Aggregators, Active customers, Renewable energy communities, and Citizen energy communities (CEC), Local energy communities. In this thesis, we are interested in the latter and we consider Local Energy Community and Citizen Energy Community the same.

Important points regarding LEC operation extracted from Directive 2019/944/EU of the European Parliament in the next paragraph:

1. Ownership structure (recital 46)– shared between stakeholders within the community.

2. New roles, rights, and responsibilities (recital 43) – local participation of citizens in the energy markets now allows them to have a stake in the activities such as production, distribution, flexibility provision, energy trading, which brings certain responsibilities to maintain the well-functioning of the grid (see Figure 1).

3. DSO status of CEC (recital 47)– if CEC becomes a DSO or so-called closed-DSO it should be treated as a DSO and be subjected to the obligations related to it. As regards ownership of the distribution infrastructure, CEC can own, purchase, build, lease the network.

4. Free entrance and leaving (recital 43)– all citizens have the right to freely join or leave the community without losing access to the network operated by this community.

5. Conflict of interest avoidance (recital 44)– decision-making power should not belong to the community members whose main economic activity is within the energy sector.

6. CEC legal status (recital 44)– the Directive leaves the possibility for a CEC to choose their legal status, which can be an association, a cooperative, a partnership, a non-profit organization, or a small or medium-sized enterprise.

7. Purpose-driven – CECs are supposed to be driven by a purpose other than making a profit, such as environmental, bill reduction, community engagement, innovation, etc.

8. Electricity sharing (recital 46)– CEC cannot be discriminated against and subjected to network charges and extra levies.

9. Market access (Art. 16, 3a) –Third-party access to the TS and DS systems for the consumers.

10. Cross-border trading participation rights (Art. 16, 2a).

11. Imbalances caused in the networks (Art. 16, 3c)– CEC are financially responsible for the imbalances they cause.

To summarize, the roles of LECs are extremely broad. The roles range both from typical activities for an energy cooperative as well generating new business models. The research done by [36] shows that the LECs can be engaged in several activities, such as renewable energy generation (solar, wind, hydro), supply (resale of electricity), distribution (ownership of the management of a distribution grid), consumption and sharing, energy efficiency services (flexibility, energy monitoring, etc.), electro-mobility (car sharing, carpooling), and other (consulting services). For more information, the reader refers to Section 7. The emerging roles of LECs are visualized in Figure 1.

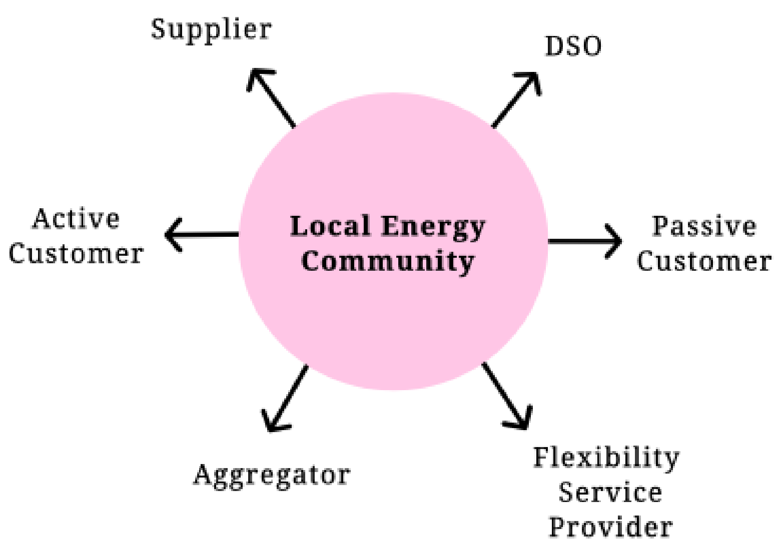


Figure 1 Emerging roles of LEC in accordance with the Directive 2019/944/EU of the European Parliament

2.2. Nomenclature

In the last decade, a number of new energy market players arose. It is essential to point out the differences and have a clear distinction among similar concepts from the outset. There are similar but slightly different citizens cooperation, such as Distributed Generation (DG), embedded generation, microgrids, virtual power plants (VPPs), Citizen Energy Communities and local energy communities. In this thesis, we explore the latter concept.

There is no universal definition of a LEC due to the variety of different arrangements and stakeholders involved. Among LEC, the following terms that represent the same concept were found in the literature: Renewable Community, Community Energy, Energy Cooperative. These terms are used to define social groups at the local level that generate and distribute renewable energy, holding high degrees of ownership of a LEC project, as well as collective benefits [12]. Energy communities involve two main dimensions: process and outcome, where process concerns with who a project is developed and run by, who is involved and who has influence. The outcome dimension describes who the project is for, who it is that benefits particularly in economic or social terms. In LECs, the process is open and participatory, and the outcome is local and collective [12].

In other words, LECs are legal entities that are commonly owned by voluntary participants, and they are a smaller-sized representation of a large electricity grid, which performs the same activities, such as generation, consumption, small-scale distribution, storage, and supply. The main difference however is that the main goal of such communities is other apart from gaining profit, it can be environmental (reducing greenhouse gasses (GHG) and other emissions), social (such as energy security, namely independence from the common energy system, which tariffs are often high and not transparent) or economical (reducing the bill size), however, the latter is not the main goal. The last objective can be achieved by providing flexibility to the main grid and being remunerated according to the tariffs, therefore generating new revenue streams by selling flexibility services.

2.3. Advantages of LECs

In this section, the research was performed to find the main benefits that local energy communities bring to their members and the energy system. Benefits were classified into four major groups: financial, social, environmental, and technological. The classification was mainly adapted from the review done by [20], which was conducted in the UK, the USA, and Germany. The following sections offer a brief overview of each of them.

2.3.1. Environmental benefits

- Lifestyle change - Increased importance of final consumers

Participation in the LEC requires more meaningful consumption and therefore behavioral change.

LEC participation influences people's lifestyle choices and helps to develop a more sustainable attitude. This means that people involved in LEC activities are generally more receptive to ethical and environmental commitment and question their behavior concerning energy consumption [20].

- Increasing RE generation

The nature of distributed generation (small-scale, on-site) forces unconventional generation. Usually, it means building RE (or at least low carbon) power plants on the consumer's premises. The range of technologies varies from solar, biomass, micro-hydro, or wind turbines to fuel cells and storage technologies.

2.3.2. Social benefits

- Increased transparency on the energy market

In the past, the energy generation process was not transparent and the only information which was available were the yearly or monthly energy bill which was the number of total kWh used – without any information of its origin or detailed day-to-day consumption. LEC members now will have access and therefore a better understanding of their generation, storage and energy flows.

- Education and acceptance

Since community members will be having the generation assets and the part of the grid (depending on the arrangement), they will have to have both a general understanding of the technical knowledge of the ES and basics of energy business functioning. The following will be needed: formation of a better understanding of RE technologies, the whole generation system, from the technical details of PV and AC/DC transformation to transmission of energy and storage. But the educational benefits go beyond the technical field. A common benefit found is energy-saving behavior, often combined with general awareness-raising for issues connected with energy consumption, for instance, climate change [20]. All such knowledge, collaboration, and ownership structure lead to overall Community sense empowerment.

2.3.3. Financial benefits

- Flexibility from LECs

Sources of flexibility are generation from RESs, stored energy from energy storage systems (ESS), demand-side management (DSM), energy trading, etc. A more detailed working principle of flexibility provision is described in Chapter 5 as the thesis practical part is focused on flexibility forecasting using real household electricity data.

- P2P energy trading

Peer-to-peer energy (P2P) trading is an innovative pricing and trading concept in the electricity market, which enables trade energy between two parties without a mediator which can increase the energy price for the buyer. The trading occurs via a platform – an online marketplace. The trading can be done automatically with the help of machine learning when one party has excess energy and the other requires energy. P2P trading can be realized between two LECs and can be a potential financial source for community members.

The trading can be done between different households which are members of one community as well as between different local energy communities. The EU Directive 2019/944/EU of the European Parliament sets new recommendations and obstacles which arise when such trading takes place: *Consumers should be able to consume, store and sell self-generated electricity to the market and to participate in all electricity markets by providing flexibility to the system, for instance through energy storage, such as storage using electric vehicles, through demand response, or energy efficiency schemes. Recent technology developments will help those activities in the future [1].*

- Benefits from ownership and investments

Economic value can come from different sources, such as investment programs made by private or public companies as well as local, national or international institutions and organizations. Looking at the examples of LECs ownership, it often occurs that a LEC can be owned by a property owner (building company) [4].

2.4. Disadvantages of LECs

The disadvantages were classified into three major categories: Institutional, Organizational, and Behavioral (social) barriers. The classification was based on [21]. There are more barriers that LEC stakeholders can face; however, it is out of the scope of this thesis. This section aims to give a brief overview. For more detailed information, the reader is addressed to [20] and [21]. The description is presented in the following sections.

2.4.1. Institutional barriers

- Lack of political support

Consistent governmental programmers, support, and subsidies are needed for the successful creation functioning of LECs. Unstable political conditions and poor governance might lead to project failures. Political instability results in frequent changes in schemes and incentives, introducing ambiguity in project planning and forecasting [22]. Lack of uniform nationwide policies and communication among different institutions limits private sector participation and investments in LEC projects [21]. Governmental financial institutions also play a significant role in the development of LECs; credit, and loans system must function well.

- Complex policies and market rules

Ambiguous national and local policies might bring confusion for potential members of LECs and can be considered as a barrier to entry. Clear and uniform and clear national policy is needed. Moreover, conflict of interest between local and national governments should not take place. Bureaucracy and approval processes should be efficient and should be as clear as possible.

2.4.2. Organizational barriers

- Protection and safety

The energy data are sensitive, and they are generated and consequently stored on a cloud or a server. As with all power plants, the electricity data should be well protected against cyber-attacks.

- Lack of long-term or initial funding

This barrier includes insufficient findings or compensations mechanisms for flexibility provision. Needless to say, that like with every power plant installation project, LEC projects usually face a number of economic and financial barriers, especially during project initiation, such projects require a high upfront cost. The communities as such don't have paying capacity; therefore, they should be subsidized or invested. These include high start-up costs, no start-up funding, and low investment and subsidies [21]. The highest cost is due to power plan planning and installation, equipment purchasing (PV panel, inverter, cables, fuel cell, etc.). The high initial investment can be considered as a large barrier to entry.

- Lack of professional support

This barrier is due to a lack of engineering knowledge among the general public. Complex engineering projects will need to be outsourced or sponsored by the government. It is expected that the number of trained people and businesses will arise therefore creating employment placement.

2.4.3. Social barriers

- Skepticism due to lack of knowledge and increased responsibility

The lack of knowledge on the energy market side or the technological side can lead to overall skepticism and fear of participating in LEC. Uncertainty and fear of responsibility due to the novelty of the concept will also increase the unwillingness to participate. Another reason for this barrier is asymmetric information that is commonly occurred in the energy market when less information is available to the end-user. This is due to the nature of energy markets where the data access available only for one party (DSO, TSO, or generation).

3. Flexibility – general overview

With high penetration of renewables as well as high usage of electrical vehicles (EV) in the power system (PS), problematics of maintaining supply-balance and system frequency became crucial. One of the ways to achieve the above-mentioned goals is to perform demand response (DR) or Demand Side Management (DSM) programs to flexible loads – this can be done in the form of active load control or load shedding as well as in the form of passive encouraging (e.g., thought price signals) energy end-users to shift their load when it is needed.

Reference [9] states the following: DSM's main advantage is that it is less expensive to intelligently influence a load than to build a new power plant or install some electric storage device. Concepts of DSM and flexibility, in general, are correlated and seeks for same results – **modification of the customer electrical load**. The concept of flexibility best described in [10]: Flexibility on demand-side is defined as the capability of consumption modification in response to control signals. Possible sources of those control signals may be external market signals to the smart meter or internal control signals from the home energy management system (EMS).

As you can see from Figure 2 from the International Energy Agency (IEA), 20% of all final energy consumption made in the EU in 2018 was from the residential sector. It is assumed that this is the reason flexibility in the domestic sector and DSM programs gained attention in the last 5 years. It is a useful source for balancing the grid.

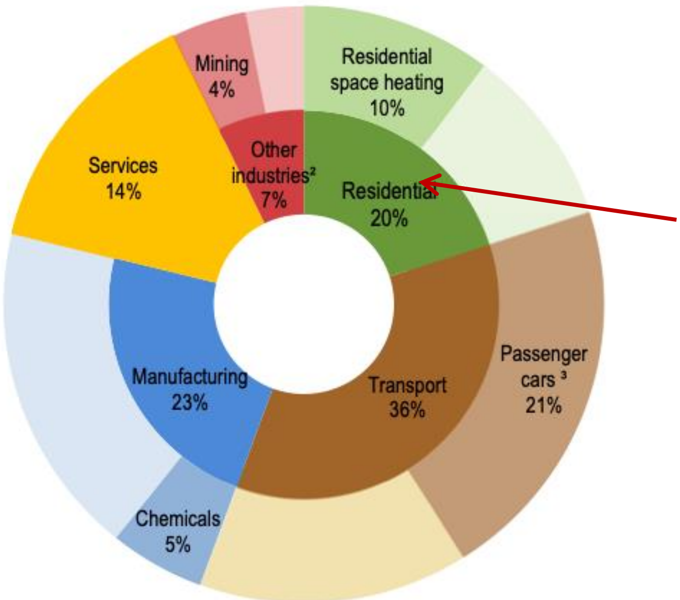


Figure 2 Final energy consumption by sector in 2018 in IEA1[11]. Notes: 1. Refers to 2018 data for sixteen IEA countries for which data are available for most end uses Australia, Belgium, Canada, Czech Republic, Finland, France, Germany, Hungary, Italy, Japan.

3.1. Flexibility and LECs

To ensure overall system stability and back up intermediate renewable energy sources (RESs), the overall electricity system requires flexibility in both energy generation and consumption sides. On the consumption side, technically, LECs are considered as one entity – a microgrid, which allows operating with it as with one large consumer or generation plant (depends on the availability of generation).

LECs can benefit from load **forecasting** to support decision-making regarding flexibility provided to the grid/from the grid. If a high degree of forecasting is achieved, LECs can sell future flexibility to aggregators or take part in Demand Response (DR) programs from utilities and therefore generate revenues. Load and generation forecasting might be economically beneficial for both generators and consumers [3]. Load (consumption as a generation) forecasting is a widely studied and important problem since it can help all energy market players to forecast future demand.

It is important to set up a clear goal and outcome of the data analytics processes prior to developing algorithms and establish the whole ecosystem with the data flow. The goal of load forecasting in the context of flexibility provision is the following: To make the forecast as precise as possible to successfully provide contracted flexibility and generate profit.

4. Machine Learning foundations and techniques

4.1. Data Science and Machine Learning Foundations

There are several data science techniques that can be used to analyze large time-series measurement data. For instance, statistical analysis techniques, artificial intelligence (AI), machine learning (ML), deep learning (DL), and many others [3]. Each of them is used for its specific purpose as well as to solve similar problems. In this section, a basic description of each of the data science techniques in the domain of the power grid is offered.

The goal of data science is to extract value from data. The steps of the data management lifecycle are demonstrated in Figure 3

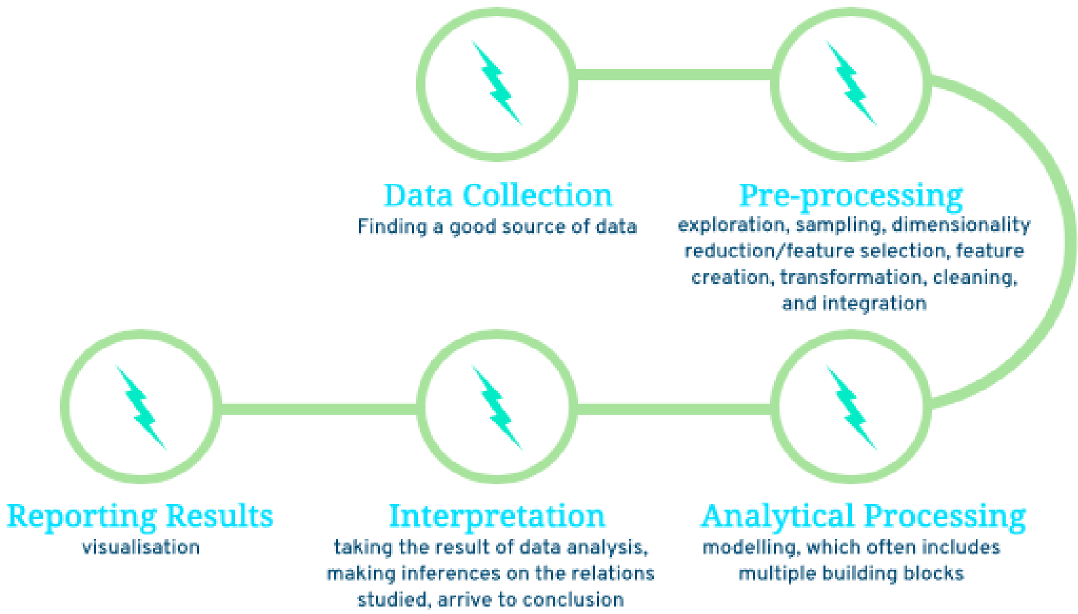


Figure 3 Main steps of the data management life cycle, source [3]

Approaching the techniques such as AI, ML, and DL, Figure 4 proves how subsections of AI are organized. AI is a broader term while ML and DL are the methods to build such an artificial system. AI can be viewed as science whose prerogative is to build smart programs that are independent and do not require human intervention. ML uses algorithms and datasets to construct models and learn by themselves according to the specified rules. DL is a specific type of ML which is more complex and has multilayers and classifies information in a more sophisticated way. The way in which they differ is in how each algorithm learns. DL automates much of the feature extraction piece of the process, eliminating some of the manual human intervention required and enabling the use of larger data sets [42].

All three methods find their applications in the field of electricity data analysis. Examples of specific applications are found in Section 4.2.

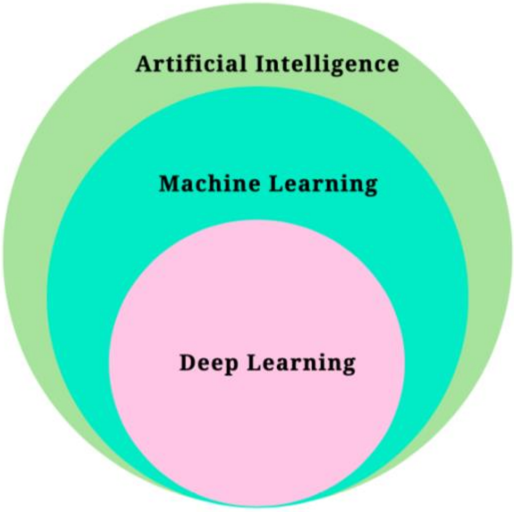


Figure 4 Differences between Artificial intelligence, Machine learning, and Deep learning

4.2. Big data and LECs

Data in the context of electricity grids play a key role. The energy data in the residential settings might be generated from smart meters or other intelligent appliances, sensors, and home energy management systems (HEMS). This allows a higher degree of digitalization and automatic decision-making. Data related to LECs are heterogeneous and come from different (external and internal) sources. The data can be both spatial (related to space) and temporal (time-series data). All the data we can use will be combined to extract as much information as possible – so-called **data fusion** [3]. There can be different sources of the data that can be combined (fused) in the context of LEC to make better sense of the data. We divide them into two categories.

Internal origin (asset level):

1. Consumption and generation data from SM
2. Sensors in the generation assets renewable and not renewable
3. Different consumption, motion, humidity, temperature sensors
4. Smart home devices, HEMS

External origin:

1. Environmental data from weather stations – temperature, humidity, solar irradiation, wind

speed, etc.

2. Grid data from phasor measurement unit (PMU) in a distribution feeder – voltage, current, phase, frequency.

3. Open Data about the electricity markets – different providers tariffs

The formats of data used in data analytics are different, and they can be numerical (measurements, numbers), categorical (text), or visual (images). It is important to mention that the computing power needed for forecasting combining different databases might be high in large LECs.

In the context of the energy system, ML techniques can bring value and help to extract some insights into the data. There are several possible tasks we can perform using ML in the area of LECs and microgrids. Machine learning objectives are often grouped into descriptive tasks and predictive tasks. Descriptive tasks aim to discover interpretable patterns that describe past data, and predictive tasks are those where the goal is to identify patterns observed in training data in order to estimate future predictions of risks and other outcomes [3]. Some examples of predictive and descriptive tasks are given below:

Predictive tasks examples:

1. Load forecasting and flexibility forecasting for LECs (short-, medium- and long- term) – RES generation forecasting for the purpose of local energy trading, bidding, flexibility provision

2. Electricity prices forecasting – the same methodology as before.

3. Predictive maintenance - Extensive deployment of SMs, sensors and monitoring technologies are used for reliability assessment of system equipment over time and to optimize the maintenance plans accordingly [41].

4. Congestions and outages identification and prediction – the data from SM in use with ML can detect deviation and anomalies to identify congestions in the power grid. The knowledge from historical data can be utilized to issue predictions of weather-related transmission outages few hours ahead [3].

Descriptive tasks examples:

1. DER Analytics - It is crucial for utilities to have correct information related to DERs at the distribution circuit and behind-the-meter (BTM) [3].

2. Generation data and Customer load profiling and classification and clustering – (clustering based on the historical demand)

3. Grid Modeling – SM data can be used in conjunction with substation data to obtain refined models for distribution planning, or to obtain insight into specific modeling problems [3].

4.3. Electrical load forecasting theory

In this thesis, we will focus mainly on household electricity load forecasting. Electricity demand forecasting is a predictive analytics task, and it is considered an essential tool to gain an understanding of future demand [23]. With the deregulation of the electricity generation and distribution sector as well as with increased RES use the importance of load and demand forecasting increased. With supply and demand fluctuating and the changes of weather conditions and energy prices increasing by a factor of ten or more during peak situations, load forecasting is vitally important for utilities. Short-term load forecasting can help to estimate load flows and to make decisions that can prevent overloading [24].

There are several traditional methods for load forecasting. They range from simple and intuitive to very complex models. They are usually classified into five categories, based on [25]:

1. Subjective – this method uses basic and intuitive and commercial knowledge without taking into consideration past trends

2. Univariate – this method takes into consideration past information using a single variable.

3. Multivariate – this is a multivariable method, which takes into account other variables (weather, prices) which usually affect consumption and generation

4. End-use – this method is a bottom-up approach; it looks in detail at the consumption side and which it consists of.

5. Combination of the above

Enterprises are increasingly moving towards the use of advanced data science techniques to forecast customer load and demand. In general, customer demand is modeled as sequential data of customer demands over time. Hence, the main forecasting problem can be formulated as a time series forecasting problem [23]. In this thesis, the **"end use approach"** for load forecasting will be used, meaning we will analyze the consumption/generation data of the end-users.

The topics of short- and long-term electricity load forecasting (STLF and LTLF) are currently in the attention of researchers around the world. There are several methods of **deep learning** that can be used for load prediction. We have selected two different DL techniques, namely long short-term memory (LSTM) and recurrent neural network (RNN), which are the most popular deep learning methods [55]. The study results in [55], on how LSTM and simple RNN based model can forecast STLF, demonstrate how LSTM outperformed other deep learning algorithms resulting in high accuracy of forecasting.

4.3.1. Machine Learning techniques for energy consumption forecasting

There are many advanced data science techniques that can be used for our purpose of domestic load prediction. Generally, time series forecasting techniques fall into the two main categories of statistical and computational intelligence methods. The classification is based on the study done in [23].

Statistical intelligence methods are commonly used; one of the most common is ARIMA. It supposes that the time series has only linear components. However, most real-world time series data consist as well of nonlinear components. To deal with non-linearity, several techniques are used: the autoregressive conditional heteroscedastic (ARCH) model, general autoregressive conditional heteroscedastic (GARCH).

Computational intelligence techniques are also frequently used for the problem of time series prediction. These methods are artificial neural networks (ANN), support vector machine (SVM), K-nearest neighbors (KNN), and adaptive neuro-fuzzy inference system (ANFIS) [23].

4.3.2. Simple RNN model

Another machine learning technique that was chosen and proved its suitability for electricity load prediction is a Simple Recurrent Neural Networks (RNN) model. Recurrent Neural Networks containing feedback and allowing previous information to be stored. Therefore, it resembles a chain of information connected with the feedback of each element. This is better explained with visualization. In Figure 5 below, a fragment of a neural network *A* takes an input value *xt* and returns the value *ht*. This feedback loop allows information to be transferred from one step of the network to another. Therefore, RNN is a type of ML that works as a loop.

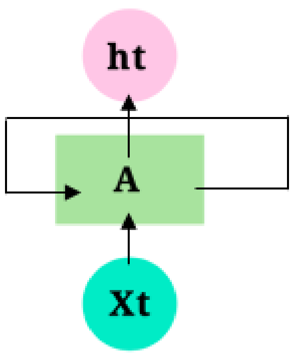


Figure 5 Schematic model of Recurrent neural networks with feedback.

The fact that RNNs resemble a chain suggests that they are closely related to sequences and lists. Over the past few years, RNN has been applied with incredible success to a range of applications: speech recognition, language modeling, translation, image recognition [27]. Because of the list-nature of RNN networks, they can model temporal dependencies and are especially suitable for the prediction of sequence data [28].

To explain the working principle, the author adopts the example from [34]. In the example in Figure 6, the final output is a 2D tensor of shape (*timesteps, output_features*), where each timestep is the output of the loop at time *t*. Each timestep *t* in the output contains information about timesteps *0* to *t* in the input sequence. For this reason, the full sequence of outputs is not needed; instead only the last output (*output_t*) is needed since it already contains the information about the entire sequence [34].

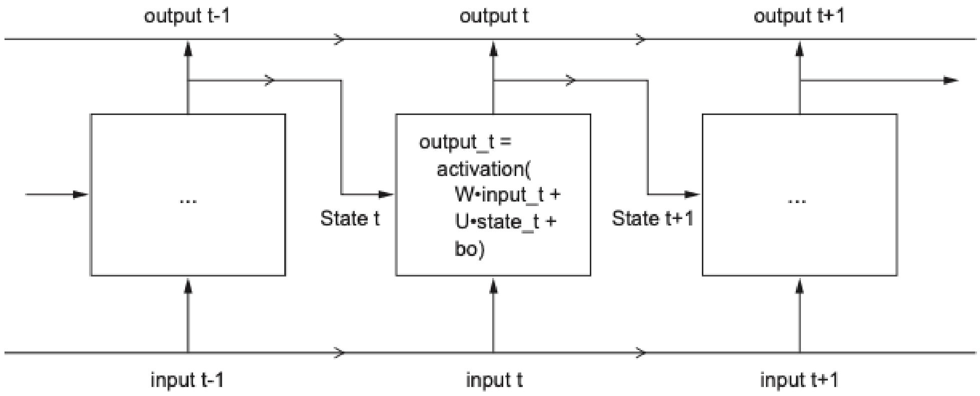


Figure 6 A simple RNN model, adapted from [34]

4.3.3. LSTM model

Long short-term memory (LSTM) is a special type of RNN architecture. The source [29] is stating that LSTM is promising that has the capabilities to learn the long-term dependencies. [29] LSTM perfectly solves a number of different tasks and is widely used nowadays. LSTM method is wildly used to perform machine learning tasks, such as image and speech recognition, translation, etc. One of the features is that LSTM has a great capability in remembering data for a long-term time and therefore they are very suitable for forecasting time series data.

The LSTM network structure differs from the conventional perceptron architecture as it has a **cell** and **gates** which control the flow of information. Specifically, the LSTM has an input gate, and a forget gate, an internal state (cell memory), and an output gate [26].

The comparison of several state-of-the-art time series forecasting techniques was made in the study [23] demonstrating the substantially better performance of the LSTM network. The reader is addressed to the above-mentioned research for a more detailed explanation. LSTM techniques with the main advantage being the capability to capture nonlinear patterns in time series data.

The working principle (adapted from [34]) of the LSTM is similar to the simple RNN with the difference being the following: LSTM adds a way to carry information across many timesteps. (see Figure 7) It is an added data flow that carries information across timesteps. Call its values at different timesteps *Ct*, where *C* stands for *carrying*. The impact of *Ct* is the following: it is combined with the input connection and the recurrent connection (via a dense transformation), and it affects the state being sent to the next timestep (via an activation function). Therefore, the carry dataflow is a way to modulate the next output and the next state [34].

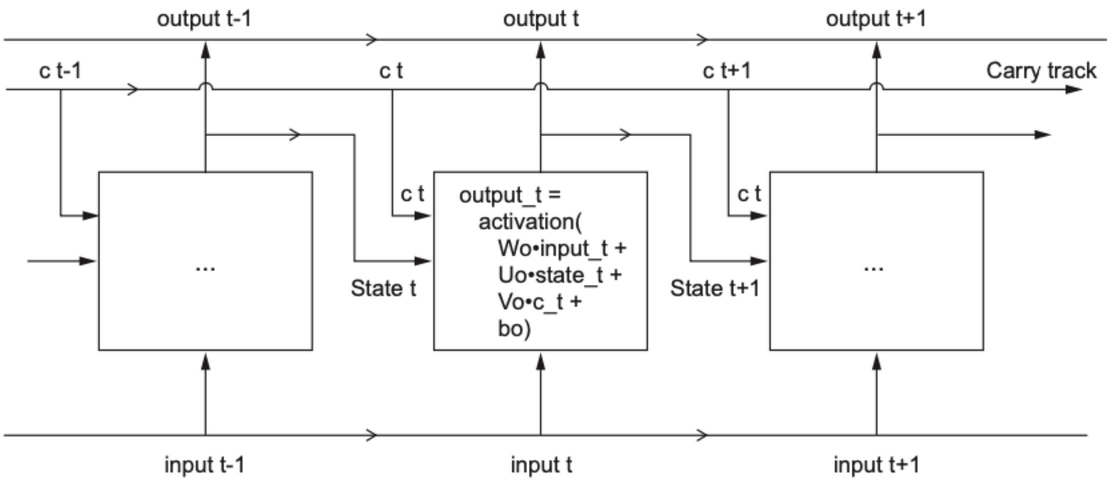


Figure 7 An LSTM model, taken from [34]

The main difference between RNN and LSTM is that LSTM can store long-term time dependence information and can map inputs and outputs accordingly [26].

5. Smart Local Energy Community – flexibility provision experiment

5.1. Dataset description

For the purpose of the data analysis and forecasting, the dataset of 25 individual houses' electricity consumption will be used in the thesis. The dataset was obtained with a license from Pecan Street Inc [5]. It is a project initiated in the US. It is a project which collects the energy data from households, and it is one of the largest sources of disaggregated customer energy and water data in form of a database. There are electrical load data for hundreds of individual households across the US. The project allows researchers and academics to receive access to their database and make use of it for research purposes.

As part of the research, the author received access to the dataset license and metadata. In the received data, there are static datasets available from twenty-five houses in California, New York, and Austin, 75 different houses in total. The time-resolution available ranges are 1-sec, 1-min, and 15-min. For the purpose of the experiment, we have selected the 15 min dataset from New York. The observations of power consumption within each household were collected every 15 minutes. Therefore, it can be calculated that there are 96 (15-min) samples per day. It is a multivariable time-series dataset that describes 6-months of electricity consumption and generation from 25 different houses. The data was collected between 1st of May 2019 until 31st of October 2019 and therefore represent the summer consumption pattern, which is usually lower than the winter consumption pattern, see Figure 8.

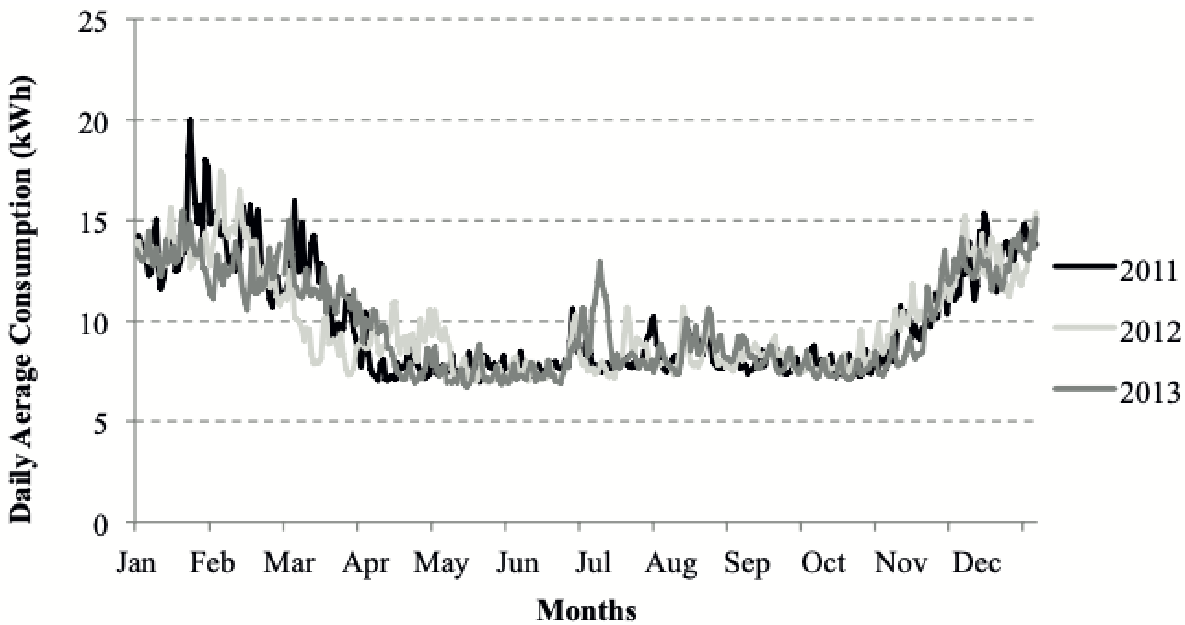


Figure 8 Daily average consumption, an example of seasonal energy demand, taken

from the research [18].

The reason behind choosing New York as a location for the experiment is that NY city can serve as a representative example for several other cities in the world. New York has an average monthly Direct Normal Irradiation (DNI) of 1491.5 kilowatt hours per square meter per day (kWh/m2/day) which is comparable with the cities like Barcelona – 1739.9 kWh/m2/day, Istanbul 1378.0 kWh/m2/day, Florence 1492.9 kWh/m2/day, Washington 1560.0 kWh/m2/day, Yantai City 1142.5 kWh/m2/day, Seoul 1229.0 kWh/m2/day, Hiroshima 1236.0 kWh/m2/day, etc. [43].

The chosen dataset is suitable for the purpose of the experiment because it is possible to simulate a local energy community of 25 different types of houses. The data provided includes detailed and precise generation and consumption information. Table 1 points out the metadata:

Table 1 Metadata description: different household consumption and generation information from pecan street database

Types	**Information**
Building information:	Customer unique ID, building type, house construction year, square footage
Generation and consumption:	PV, PV panel direction, PV size, battery, Electrical vehicles, individual rooms consumption, total grid consumption
Consumption: energy-intensive appliances:	Energy-intensive appliances: freezer, washing machine, electric heater, swimming pool or jacuzzi and electric water heater, kitchen appliances (oven, icemaker, microwave)

The goals of the practical part of the thesis are:

- Forecast grid consumption for each of 25 houses - separately
- Forecast consumption patterns of electrical appliances for each house - separately

Data sources used for analysis:

- Historical households’ consumption and generation data
- Meta-data: type of households and most energy-intensive devices used in the building

5.2. Data analytics

In this section, we perform data visualization for some attributes of the dataset that will later be used for forecasting. Firstly, in the graph below (Figure 9), aggregated average consumption of twenty-five houses is shown.

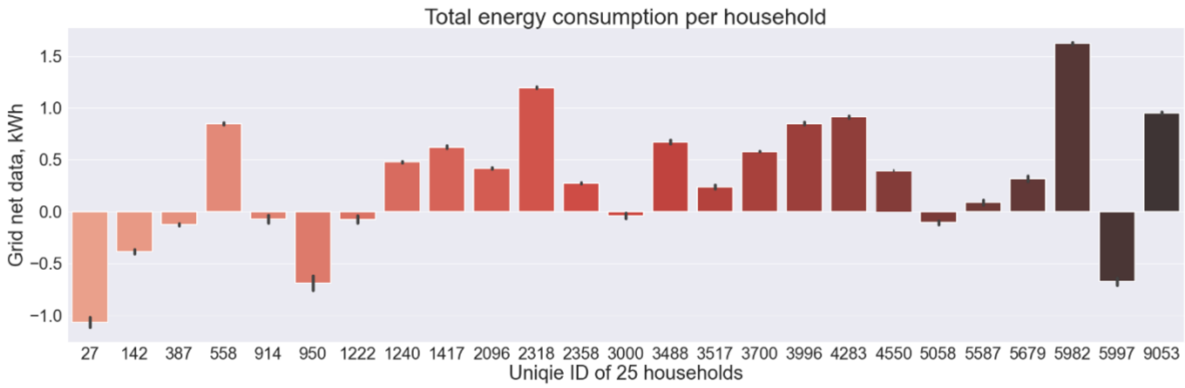


Figure 9 Aggregated average daily profiles of twenty-five houses in NY, data taken from Pecan Street [5]

Secondly, the aggregated PV generation of each house is illustrated in Figure 10. The time range is the maximum that was available in the dataset, which is 6-months from May until October (summer season). It is visible that all the houses follow the same electricity generation trend, which might be affected by weather conditions (sunny or cloudy days), day/night picks, etc. As was described before, there are twenty-five houses in the dataset, each house has a data ID which can be seen at the right corner of the graph.

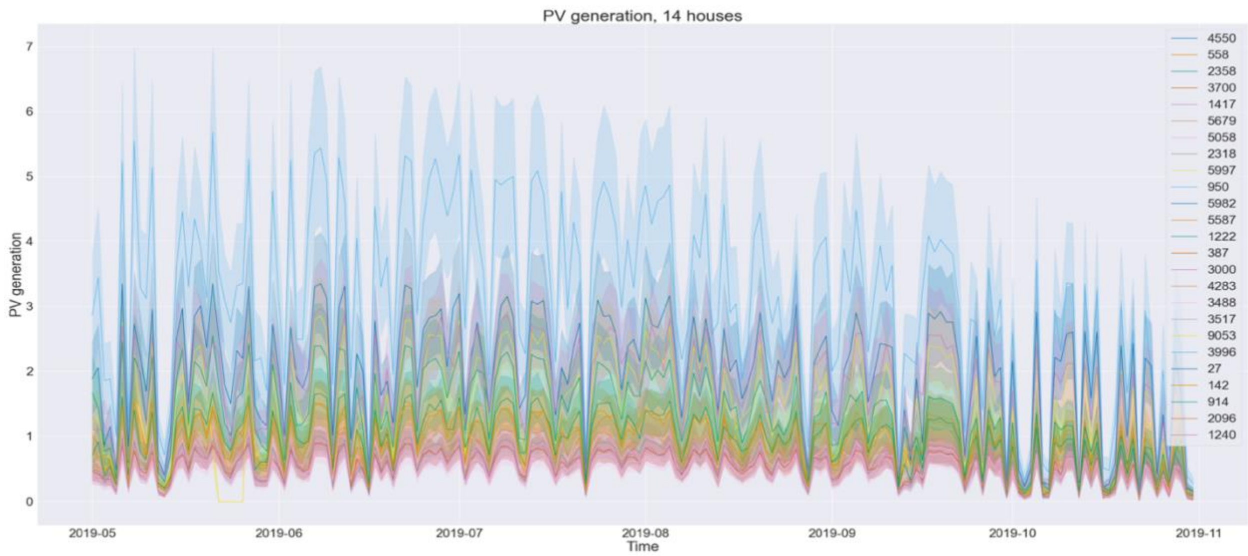
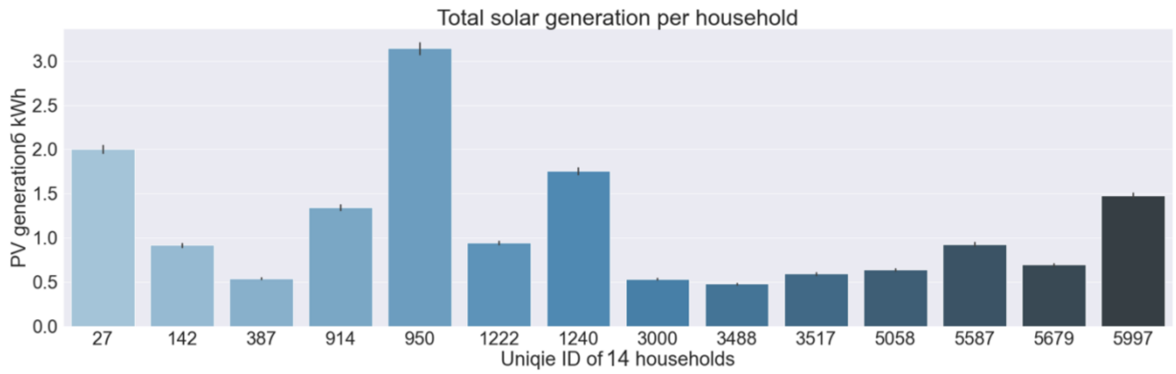


Figure 10 Aggregated 6-months PV generation profiles of twenty-five houses in NY, data taken from Pecan Street

As was mentioned before, several houses from 25 houses have photovoltaic (PV) panels and electric vehicles (EVs), generation and consumption, respectively. From the available data and metadata, it was identified that 14 out of 25 houses have PV panels, and 5 houses are with EVs. At the same time, 4 of the houses have both solar PV and an EV. Figure 11 a and b reveal the aggregated energy generation from PV (a) and aggregated energy consumption (b) that EV charges use.



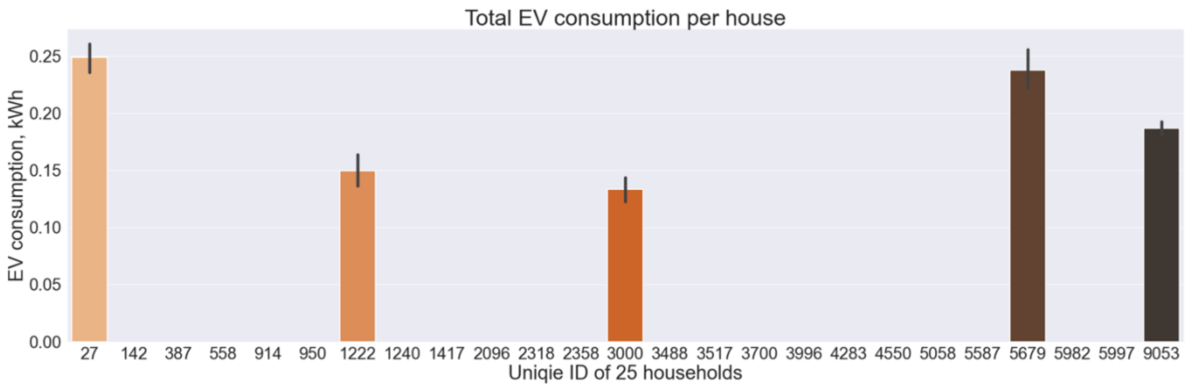


Figure 11a - Total PV generation by a house (above), 11b - Total EV charging consumption by a house (below), data taken from Pecan Street

The monthly consumption trend is shown in Figure 12 of one representative house № 914 in the period from 9.07.2019 until 9.08.2019. The repeating trend from day to day is shown with the positive values (taking electricity from the grid) to negative values (feeding the electricity grid with the generated electricity. The EV size and PV size are unknown. The given house is a single-family house located in the city of Ithaca, NY state. From the big appliances, it has one EV, clothes washer, dryer and jacuzzi. The house also has PV panels (the size is unknown).

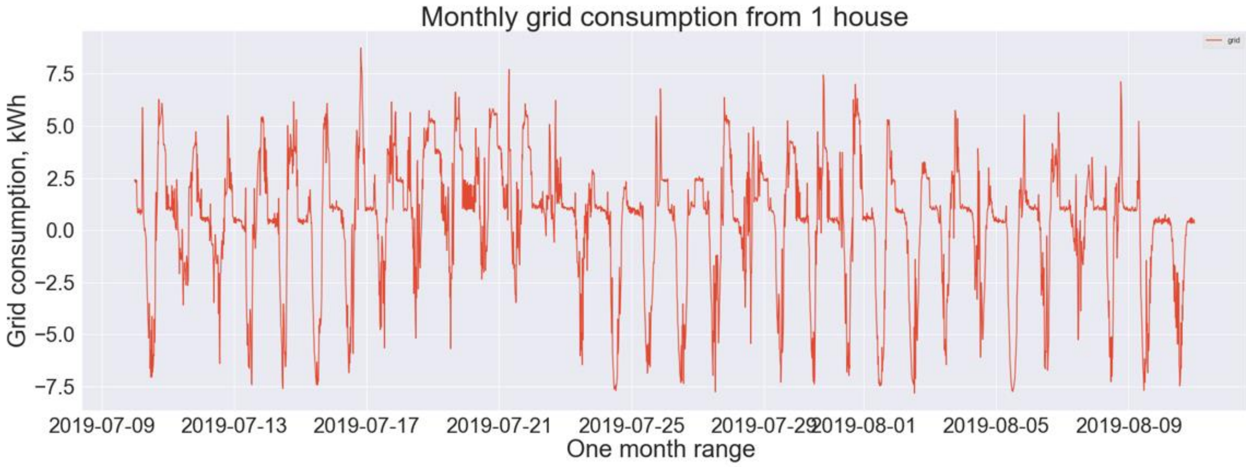


Figure 12 Monthly consumption from one house in NY, data taken from Pecan Street

Ithaca is a large town in the US, with coordinates: latitude 42.443962 and longitude -76.501884. Ithaca has average yearly Direct Normal Irradiation (DNI) of 1253.4 kilowatt hours per square meter (kWh/m2) and daily GHI of 3.434 kWh/m2/day which is lower than the average DNI in the state of New York. The average temperature in Ithaca is 9.8 °C. The monthly averages of DNI for the city of Ithaca are given in Figure 13. The highest DNI is in July, the month that we are taking for the analysis of the load pattern.

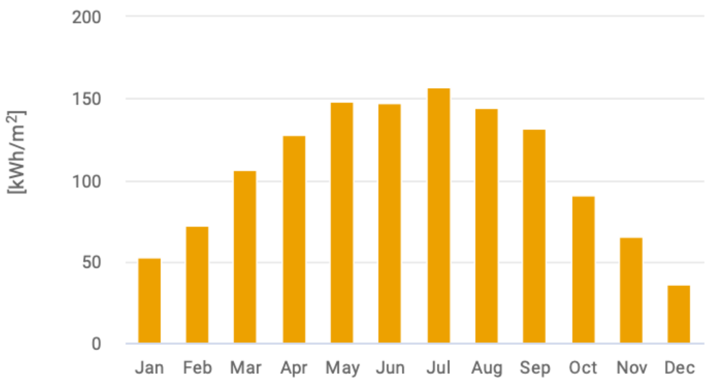


Figure 13 Monthly average Direct Normal Irradiation (DNI) in the city of Ithaca, data taken from [43]

Figure 14 shows the daily household grid consumption from the representative house № 914. It resembles the well-known “duck’s curve”. It is clear that the aggregated load is rather consistent, and its daily pattern is highly noticeable. The feature of the duck’s curve is that it has two peaks: in the morning and in the evening as well as one big drop during midday, following the renewable energy sources production. The drop is caused by the large power input from PV panels. Variable generation resources significantly reduce the load on conventional generators during the day but not during the night; a surge in generation demand may occur at the time of sunrise and sunset.

This trend might be seen in Figure 14 [35].

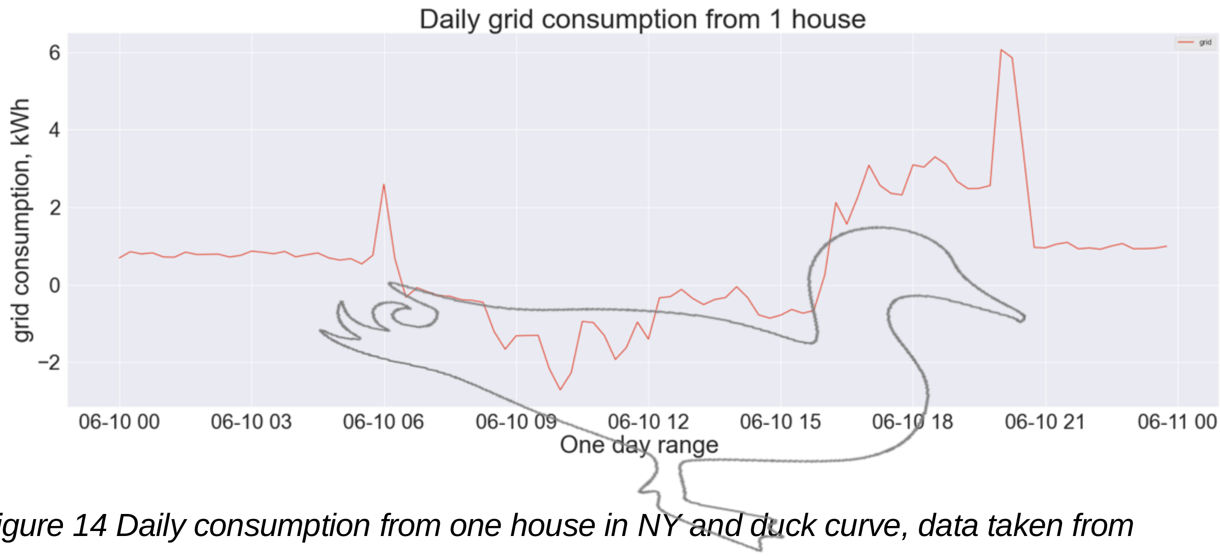


Figure 14 Daily consumption from one house in NY and duck curve, data taken from Pecan Street

5.3. Machine Learning training and forecasting experiment set up

There is a wide range of statistical analysis tools, such as MATLAB, R, Microsoft Excel, Python (Pandas), or Tableau. They are used to draw meaningful interpretations, visualize or compare trends as well as report of the research findings. In this thesis, the energy analytics and predictions were made in Python programming language with the help of libraries used for machine learning and data analytics, such as pandas and Keras. Anaconda Jupyter Notebook which is a web-based data analysis environment was used. The example of how to train neural networks (NN) models are explained in the following sections.

5.3.1. Training the LSTM Network

Commonly, the process of building an ML algorithm starts with the development of a simple network, either directly applying architectures that have already been successfully used to solve this type of problem or using hyperparameters that have already given useful results before. In this thesis, the author decided to follow this approach and choose as a reference the following tutorial [30] which focuses on a forecasting model construction using LSTM and RNN models. Below, the training process will be explained in detail on the example of the LSTM model.

5.3.1.1. Preparing data

Before the start of building and training the model, it is important to perform **data clearing** and **data normalization.**

Data Clearing. If the time series contains noisy and missing values, the noisy values are smoothed, and the missing values are replaced by the appropriate method [23].

Data Normalization. Normalization is the scaling of data from the original range so that all values are in the range 0 to 1. Normalization requires knowledge or estimation of the minimum and maximum observed values [33]. An example of normalization code is presented below:

```
In [163]:
def normalize_data(df):
    scaler = sklearn.preprocessing.MinMaxScaler()
    df['grid']=scaler.fit_transform(df['grid'].values.reshape(-1,1))
    return df

df_norm = normalize_data(df)
df_norm.shape

Out[163]: (17664, 1)
```

Figure 15 Data Normalization example using MinMaxScaler function, based on the tutorial [30]

Splitting the data. To achieve successful time series forecasting we need to configure an LSTM network. First, we will split the total household consumption data into two parts: **a training set and a test set.** Since the dataset is 6-month long, we take the first three months for the training set, and the remaining 3 months of data is used for the test set.

5.3.1.2. LSTM architecture building

There are several **hyperparameters** that influence the quality of the forecasting result. Adjusting these hyperparameters is the process of modifying network components which helps us to achieve maximum performance and accuracy of the model. Several examples are presented below:

Table 2 Hyperparameters chosen for the LSTM model training

Hyperparameter name	Chosen type/value	Explanation
Optimizer (solver)	adam	The optimizer is responsible for the minimization of the objective function of the NN [31]. Adam optimization is a stochastic gradient descent method that is based on adaptive estimation of the first-order and second-order moments [32].
Loss	MSE	The purpose of loss functions is to compute the quantity that a model should seek to minimize during training. Regression loss function was chosen - mean_squared_error. This type of loss

		computes the mean of squares of errors between labels and predictions [32].
Number of epochs	10	Number of training iterations
Batch size (normalization)	1000	Determines how often the network weights are updated. Batch normalization is done before entering each layer.
Dropout	0,15	Slows down learning with regularization methods to avoid overfitting for NN.
Activation	tanh	Layer activation function. Hyperbolic tangent activation function was chosen.

5.3.1.3. Forecasted Results

The results of the forecasting are illustrated in Figure 16 and Figure 17 below. The visualized consumption and generation curve is for the 6-month range (05.2019-11.2019) of the house № 3000. From Figure 16 is can be observed that the predicted curve follows the shape of the actual power consumption curve and correctly identifies the peaks. From the visual interpretation, it is noticeable that a simple RNN model made forecasting with higher accuracy; in the following subsection we will prove it quantitatively with the help of accuracy metrics.

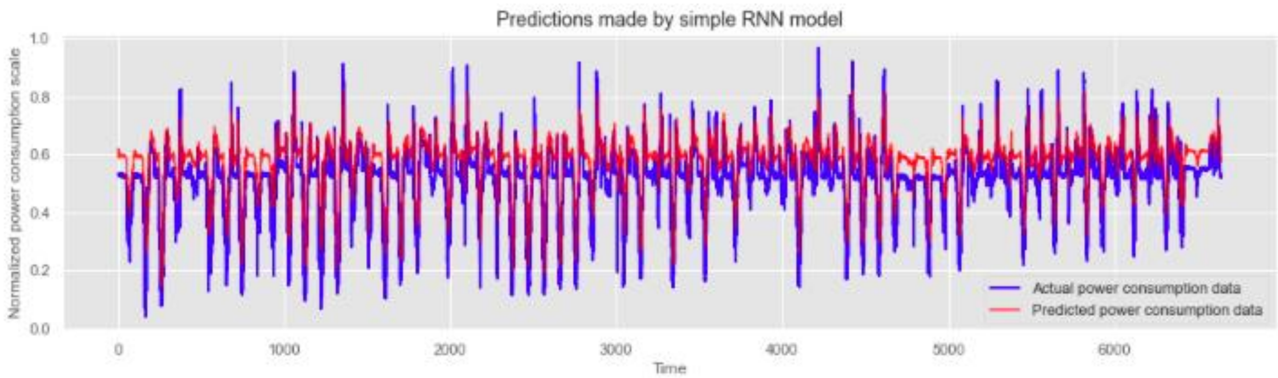


Figure 16 Predictions made by simple RNN model

Figure 17 demonstrates predictions made by the LSTM model. The accuracy is not high; however, as with the earlier Figure 16, we can see that the predictions curve (red) resembles the shape of the real electricity consumption curve (blue).

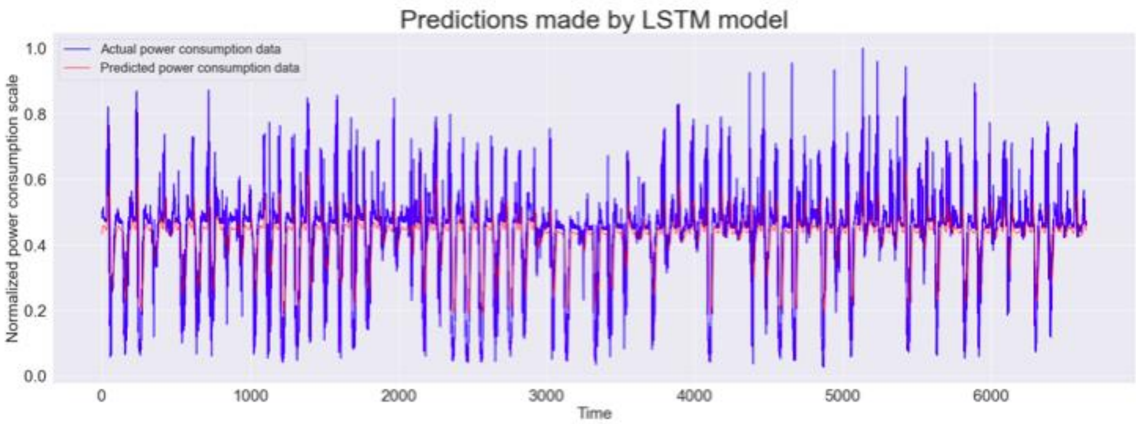


Figure 17 Predictions made by LSTM model

5.3.1.4. The quality of forecasted results

After the forecasting was done, it is necessary to evaluate the model accuracy. There are several performance measures parameters used to assess the accuracy of time series data forecasting using Python, called metrics. The most common performance parameters are: Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Absolute Error (RMSE), R-squared (R^2).

- **MAE** is an average of the forecast positive error values.
- **MSE** is an average of the squared forecast error values. Squaring the forecast error values forces them to be positive; it also has the effect of putting more weight on large errors [44]. MSE of zero means no error occurred.
- **RMSE** is calculated as square root of the MSE. RMSE of zero means no error.
- **R^2** is the square of the Correlation Coefficient (R) and it is calculated by the sum of squared of prediction error divided by the total sum of the square which replaces the calculated prediction with mean. R^2 is between 0 to 1. The larger R^2 is the higher the correlation between actual value and predicted values. Smaller the R^2 poorer the model.

In the forecasting code, R^2 metric was used to evaluate accuracy of both LSTM and RNN models. The predictions presented the following results:

- R^2 Score of RNN model = 0.543942483197771,
- R^2 Score of LSTM model = 0.4096853402745917

Therefore, in the case of 6-month one house load forecasting, **RNN model performs slightly better.**

6. Flexibility provision use case

6.1. Twenty-five houses flexibility analysis

In this section, an analysis of available data on appliance consumption was performed. In Table 3, all appliances that appeared in the database from Pecan street were listed. Peak power consumption values from all analyzed twenty-five houses were identified from the database as well as a particular house № 3000 was taken to perform further analysis.

Table 3 Appliances information from the Pecan street dataset of 25 houses from New York.

Measured Appliances	Number of houses with the appliance	Peak power consumption [kW] – all 25 houses (from the dataset)	Peak power consumption [kW] – house № 3000 (from the dataset)	Average power consumption [kW] (from the dataset) – all 25 houses	Typical range (min-max) of power consumption [kW], taken from [8]	Potential Flexibility and demand response programs
EV charging	14	7,035	3,872 car 1 2,735 car 2	0,192	2 - 7	Yes[1]
Washing machine	16	1,015	0,910	0,0043	0,5	Yes
Clothes dryer	17	6,004	0,815	0,0665	1 - 4	Yes
Jacuzzi	3	1,87	-	0,096	3 - 7	Yes
Dish washer	12	1,374	1,342	0,0142	1,2 -1,5	Yes
Air Compressor	11	3,046	2,224	0,146	-	Yes
Kitchen Appliances (incl. light)	18	1,496	0,857	0,0357	-	No
Electric	11	14,039	0,288	0,043	2 - 5	Yes[2]

heater (heating or radiant floor heating)						
Water heater	10	13,099	-	0,201	6,6 – 8,8	Yes
Pump (used to circulate water in a hydronic heating or cooling system)	5	0,61	0,259	0,081	0,2 – 0,4	Yes
Living room (incl. light)	6	4,65	1,215	0,075	-	No
Bedroom Appliances (Inc. Light)	4	1,436	1,402	0,029	-	No
Furnace and air handler	11	1,196	-	0,084	1 – cooker 0,02 –hood	Rather Not
Freezer	7	1,25	0,606	0,085	0,3 – 0,4	No
Microwave	3	1,006	-	0,005	0,6 – 1,7	No
Refrigerator	12	1,213	-	0,058	0,15 – 0,4	No
Oven	6	3,913	-	0,015	2,1	Yes
Cooktop or cooktop and oven	10	4,37	-	0,037	1,4 – 1,8 (w/o oven)	Rather not
Garage Appliances (incl. light)	7	1,501	-	0,052	-	No

Note: *1,2 – we send the signal to decide when to charge

For the purpose of analysis, the peak power consumption values of all available appliance's from 25 houses for 6 months were taken. The proportion of each appliance consumption is presented in Figure 18. 20% of overall electricity household consumption comes from heating elements.

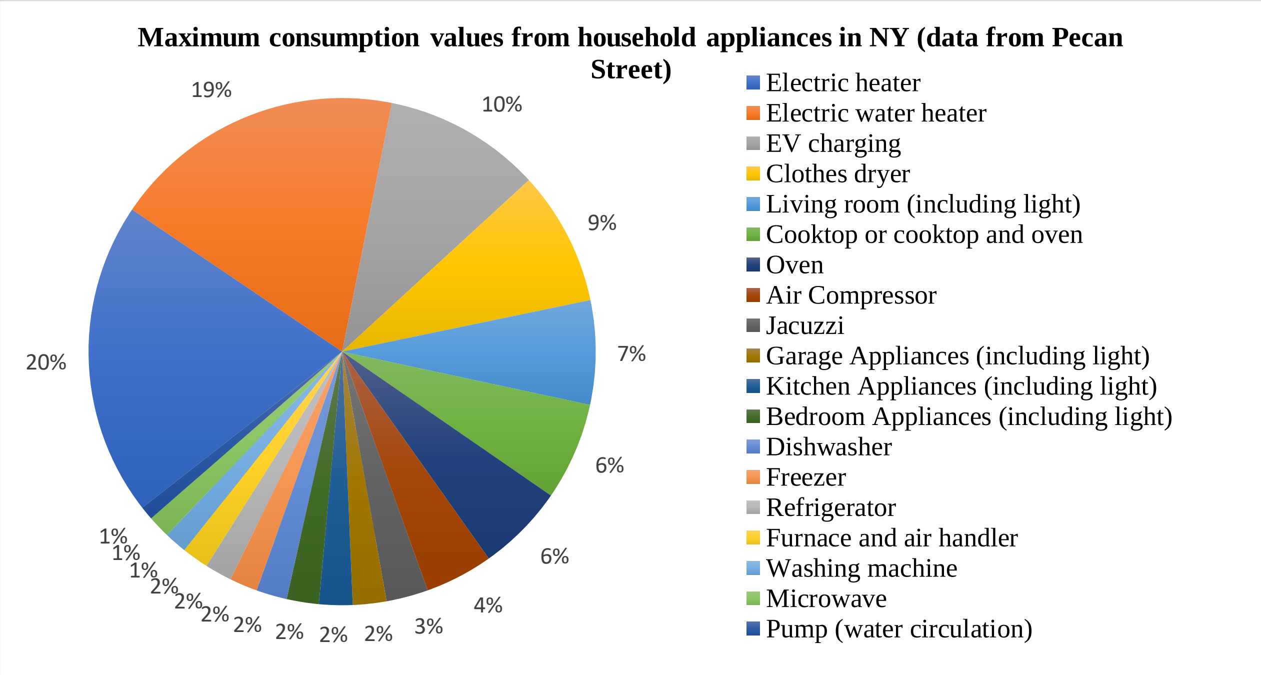


Figure 18 Percentage of consumption of all measured appliances in all 25 houses in New York (Pecan Street Database)

6.2. One house flexibility analysis

Analyzing the capability of available devices to supply flexibility, it is essential to analyze the consumption patterns of energy-intensive applicants. For illustration, we take a stand-alone electric space heater consumption profile. In Figure 19 (top) the seasonality is evident; since the measurement was done from May until November, we can see that the heating requirements drop with temperature rise in June and it restarts again in October. While the heating season is October Figure 19 (middle), it is clear from the peaks that the heating element was switched on every day. During the heating season in the middle of October, the heating element was switched on several times during the day, as visible from Figure 19 (bottom).

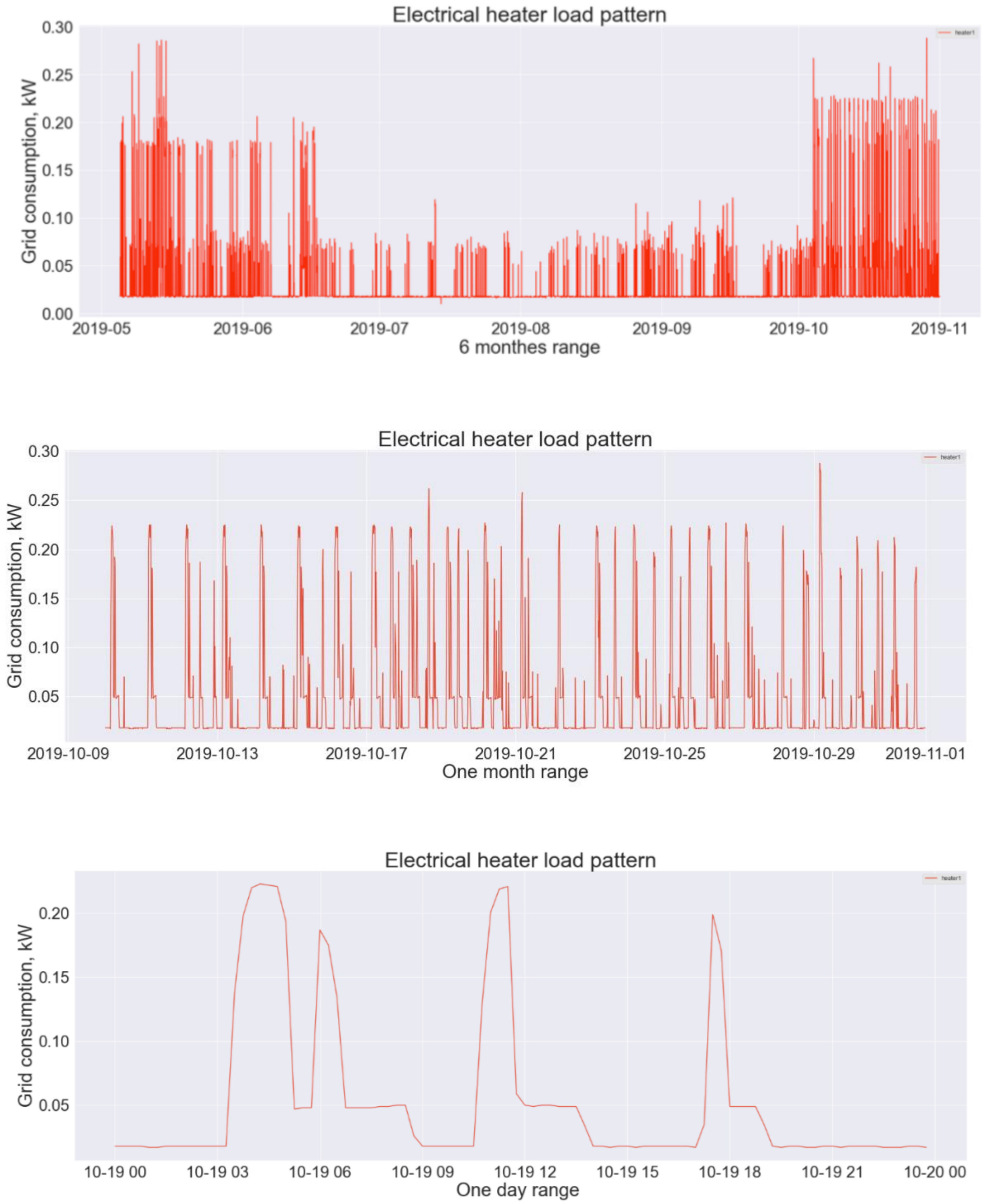


Figure 19 Electric heater consumption, from top to bottom: 6-months range, 1-month range, 1-day range, data taken from Pecan street New York database

The second most energy-intensive consumption appliance (19%, Figure 18) of overall electricity consumption is the electric water heater. To be able to use flexibility from the water heating, research [14] proposes to use, so-called domestic hot water buffers (DHW buffers) which are water heaters that can preheat and store heat in a water buffer. Such buffer allows the decoupling of hot water production and consumption which offers opportunities to create flexibility.

Not all electrical appliances that we use in our houses can deliver flexibility to the same extent since the switching off and on time as well as consumption patterns are significantly different. In the thesis, with inspiration taken from [17], we group all devices into two categories, into appliances which consumption can be shifted to a different period – **demand shifting devices** (as well called: demand shifting, schedulable or postponable appliances), and devices with real-time control capabilities (using power controllers), so-called **real-time devices**. As described in [17]:

- Schedulable appliances: fully flexible and can be turned on at a later time when the real-time electricity price is reasonable, e.g., washing machine, dishwasher, and air conditioning

- Real-time devices: have a low degree of flexibility depending upon the basic needs and consumers' priority, e.g., lighting devices, computers, and televisions can be the candidates of this category.

Looking at the consumption from house № 3000, Figure 20, the largest share of power consumption comes from the EV charging station (40%); the reason for that is the two EVs that the house possesses. This is a potential source of flexibility – the appliance that can be controlled externally. The study done in [14] describes the process of making use of EV chargers for flexibility purposes. The user sets the expected time of departure and charging time through a linear portal site. DR control system is able to then interrupt the charging process when needed.

Another large consumption device is an air compressor (13%). Air compressors convert electrical energy into potential energy store in a pressurized gas. Air compressors compress the gas under high pressure. This device falls into the category of postponable appliances.

The third share of the chart belongs to the living room (8%), and bedroom (8%). Estimation of appliances that can be used in these types of rooms: TV, music system, lighting system or wi-fi router are real-time appliances with a lower degree of flexibility.

In the third and fourth positions, there are useful appliances from the flexibility point of view: dishwasher (8%) a washing machine (5%), respectively. Usually, another appliance that accompanies the two is an electric clothes dryer. All three so-called wet appliances. Wet appliances are postponable. Postponing can be achieved either through direct engagement with end-users, by asking them to use their devices at a particular hour, or indirectly through automatic switching on and off signals (using power controllers). This can be done during the low tariff times or at night (threshold can be set), which can be equal to the times when the electricity is the cheapest and there is a need from the energy supplier to consume the electricity. This process of demand-side management is called **demand shifting.** [15] explains it as following: if demand is needed to be lower tomorrow, intelligent consumers can plan ahead and—if their process allows it—do their tasks earlier or later. Processes that can be shifted typically belong to one of the following categories:

- Inert thermal processes (heating, cooling)
- Inert diffusion processes (ventilation, irrigation, etc.)
- Mass transport (pumps with tanks, conveyor belts, etc.)
- Logistics (schedules, dependencies, lunch breaks, etc.) [25]

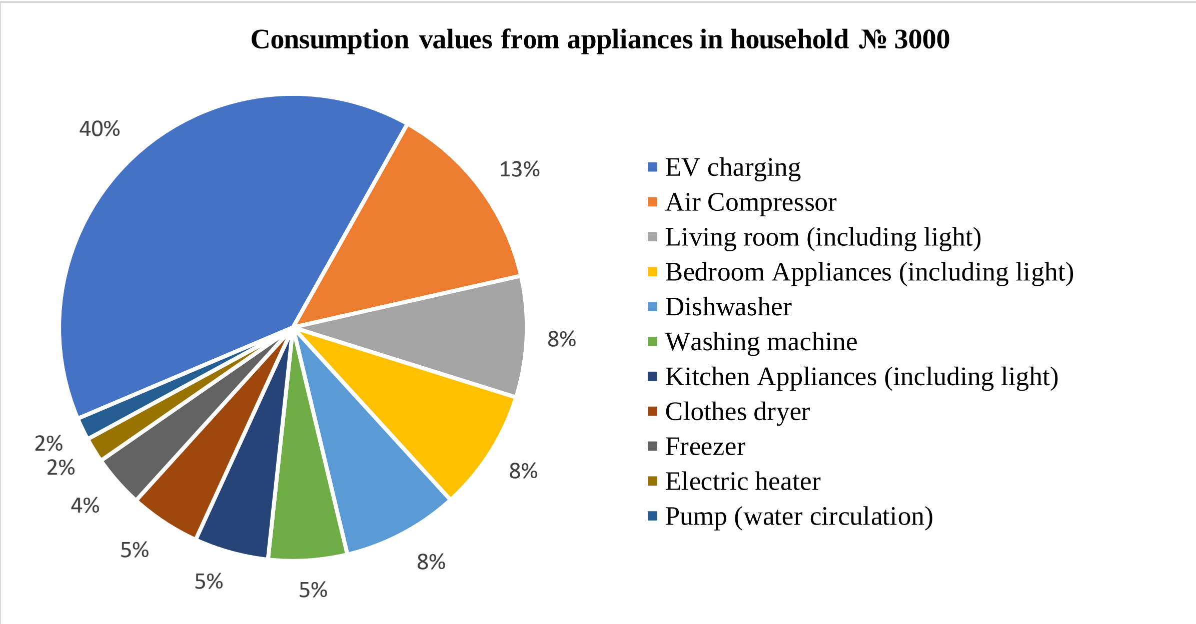


Figure 20 Percentage of consumption of all measured appliances in house № 3000 in New York (Pecan Street Database)

6.3. Consumption forecasting of selected appliances

In this section, several selected houses and selected appliances were chosen to perform the load forecasting. During the forecasting, we have in mind the simplified model of a possible flexibility exchange market structure, which is shown in Figure 21. The model stands for the situation when the users register voluntarily for participation in flexibility provision, and they receive monetary compensation if they change their baseline profile. Adapted with modifications from [16].

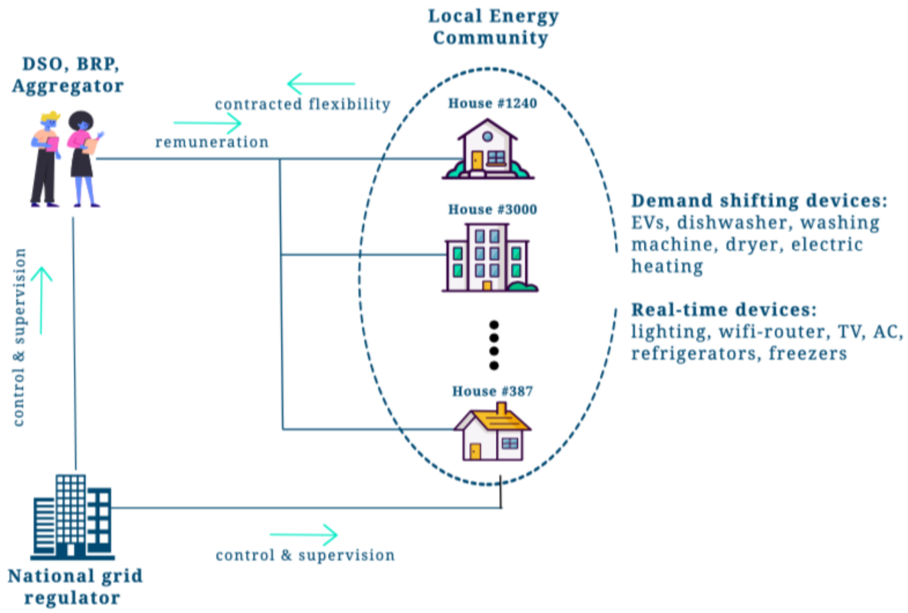


Figure 21 Possible flexibility exchange model in which DSO, BRP, or an aggregator can be in control of the management of devices with DR capabilities [16].

For the forecasting purposes, we analyze two houses, №3000 and №1222, which both have PV generation and one or two EVs and then a range of different large-consumption devices. Figure 22 demonstrates available appliances from selected houses from the dataset:

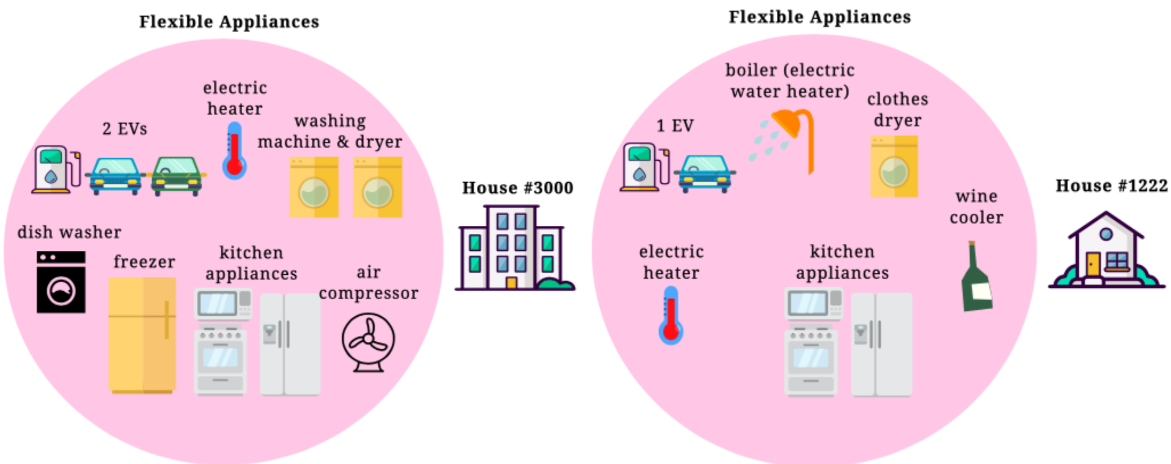


Figure 22 Measured flexible appliances in the house №3000 and №1222 – both with PV generation

6.3.1. Forecasted Results

In the code, the author tried to visualize and forecast all the available appliances; however, for the purpose of the example, only one appliance the **electrical dryer** load pattern from the house **№3000** was visualized for 6-month (Figure 23), 1-month (Figure 24) and 1-day time-range (Figure 25)

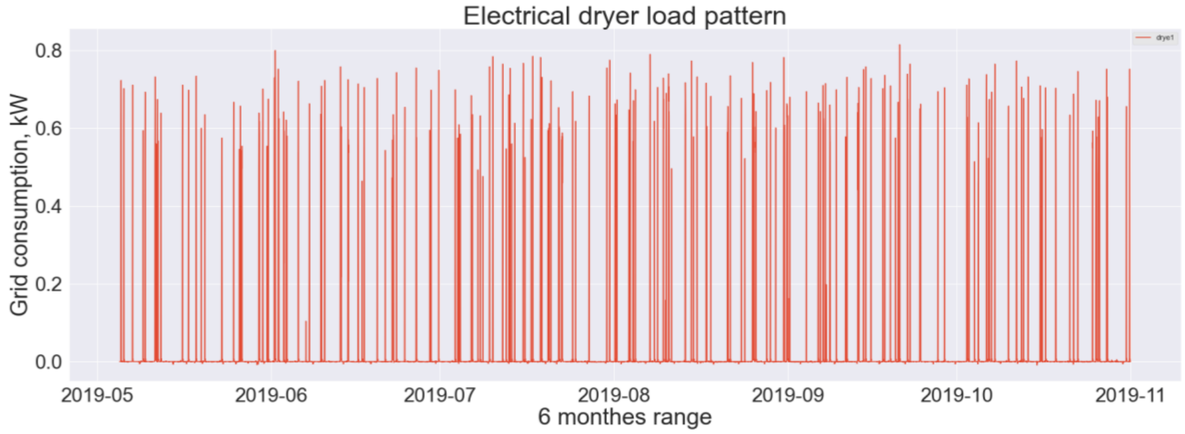


Figure 23 Electricity data visualization for an electrical dryer, house № 3000, 6-month range

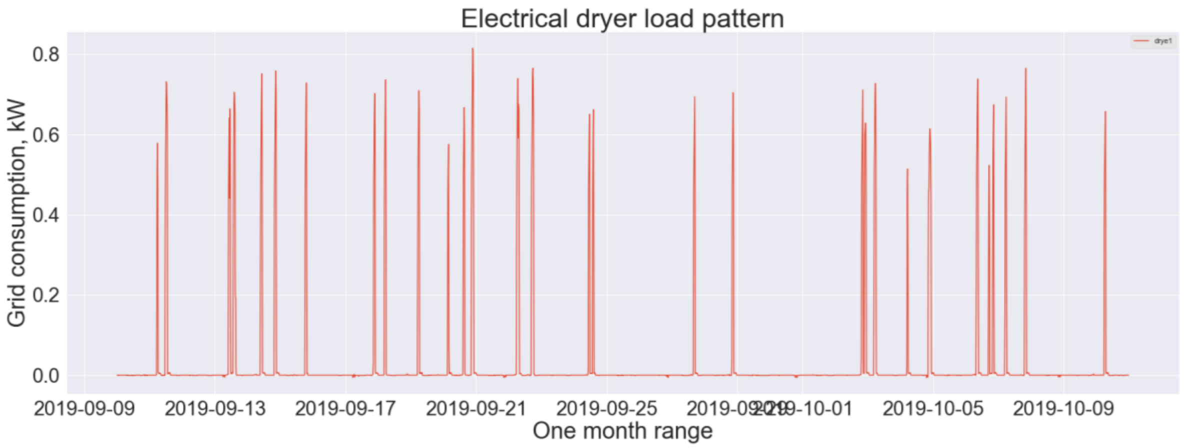


Figure 24 Electricity data visualization for an electrical dryer, house № 3000, 1-month range

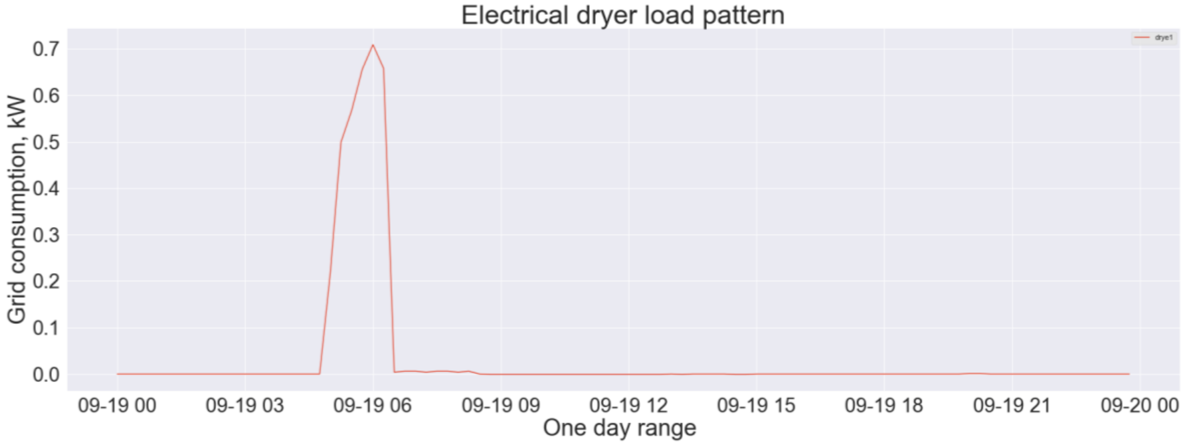


Figure 25 Electricity data visualization for an electrical dryer, house № 3000, 1-day range

As for the earlier forecasting, two main ML techniques were used: LSTM and simple RNN. Here we forecast 6-month (05.2019 - 11.2019) long consumption of an individual appliance consumption pattern; we again take one house № 3000 and forecast the load of an electric clothes dryer. In this case, from the visual assessment we can make an assumption that **the LSTM model performed better than simple RNN** since as can be seen from Figure 26 the forecasting (red curve) does not reach negative values (Figure 27). The model can perform better since the forecasting peaks almost twice smaller than the real consumption; therefore, the LSTM model can be trained better, or some hyperparameters should be adjusted to achieve better results.

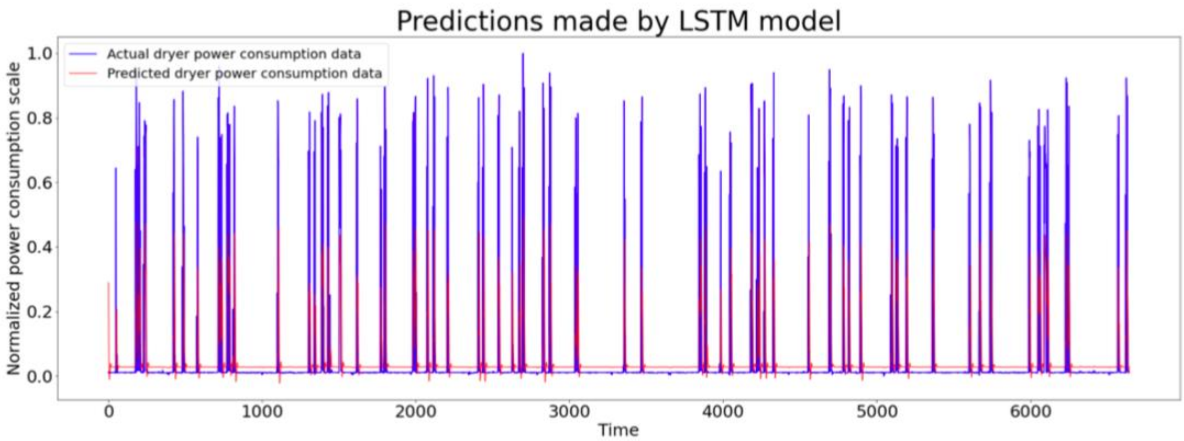


Figure 26 Predictions made by LSTM model (electrical clothes dryer example)

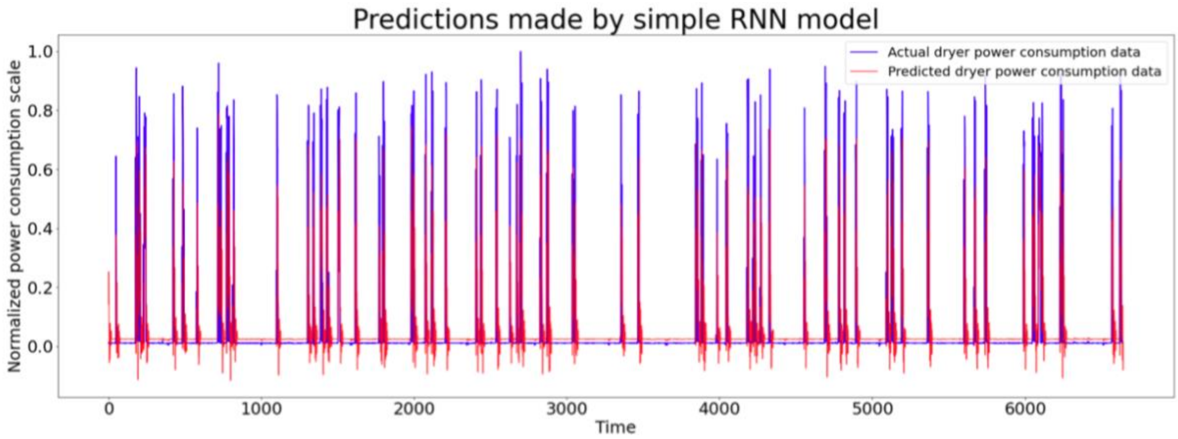


Figure 27 Predictions made by a simple RNN model (electrical clothes dryer example)

Indeed, looking at the performance (forecasting accuracy) parameter square of the Correlation Coefficient - R^2, for more information please read subsection 5.3.1.4, we can conclude that the **LSTM model performed better than the simple RNN model** since the R^2 Score of LSTM model is closer to zero. We can see that the LSTM score of R^2 is significantly lower, meaning that the model makes forecasting with fewer errors. The values of R^2 scores are:

- R^2 Score of LSTM model = 0.2654782470015554

- R^2 Score of RNN model = 0.5272295667696455

6.3.2. Flexibility forecasting conclusions

After the forecasting experiment using advanced methods of Machine Learning, it can be concluded that both methods (RNN and LSTM) are suitable for consumption/generation forecasting both for individual appliances, all appliances for one house and accumulative consumption/generation of a local energy community.

7. Business Models associated with LECs

As was pointed out before, however worth mentioning, energy communities are not profit-driven entities, which bring different value-added services (VAS); they, have an increased sense of local community and collective ownership, they bring social benefits. VAS that LECs provide are aiming at meeting customer needs (lowering energy bills, optimizing energy use, being a prosumer, choosing a specific energy mix, etc.), or services primarily meeting energy grid needs (articulating DER that provide energy, capacity, regulation, and/or other services to the power system) [37]. There are different BM archetypes that exist both theoretically and as a real-world installation; this, chapter discusses the main types, taking into account different objectives and stakeholders involved.

All business models presented below are based on the shared economy principles (when the profit is shared among stakeholders equally). This arrangement builds the required trust bridges between all the stakeholders, and very particular with the prosumers. Depending on the particular services that LECs offer the business model differs and the stakeholders involved are different. There is a large number of business models (BMs) that are designed to materialize on several services, while maximizing LEC’s full potential.

7.1. Interaction of LECs with other energy market players

The following section aims at setting up the relationships between LECs and different stakeholders (LECs business partners) operating on the energy markets: DSOs and TSOs, aggregators, EV charger owners, generation assets owners, and finally other LECs and microgrids. Since a LEC cannot be owned by any energy market player, the interactions which will arise between LECs and energy companies are important to investigate. Different BMs and revenue stream structures are investigated in relation to each of the entities.

7.1.1. LECs and DSOs

According to the Directive 2019/944/EU, DSOs in the EU cannot own LEC; therefore, it should treat LECs as a competitor, partner, and consumer at the same time. There are different possibilities of using the distribution grid by LEC, and it can lease, build, or buy the networks. Figure 28 depicts the main needs in the context of a DSO and LEC cooperation.

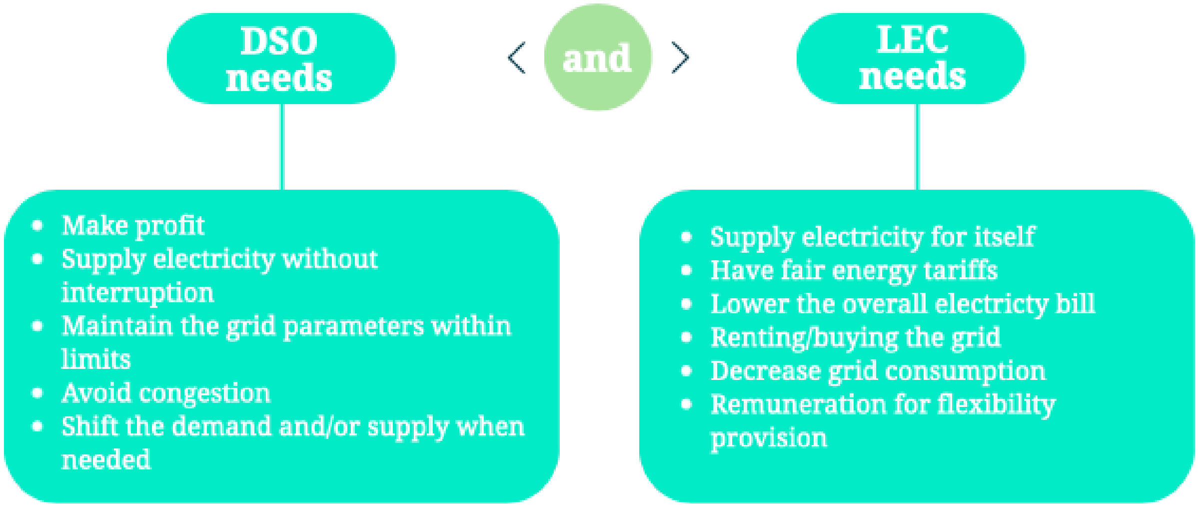


Figure 28 DSO and LECs relationship

DSO is indirectly taking part in most of the BM presented in this chapter since LEC is not always completely isolated from the main grid. Several business models associated with DSO and a LEC are proposed in [37]:

1. Utilities can provide DER as a service in exchange for a fixed monthly fee. It could make it more affordable to all customers and facilitate geographical optimization of RES deployment through the grid. The main utility's role is that of a market facilitator and operator.

2. Another example of a business model is peer-to-peer energy services transactions within the LEC or between different LECs, typically supported by blockchain technology and done automatically.

7.1.2. LECs and aggregators

By far one of the most important LEC partners is an aggregator. The aggregator's goal is to collect and trade available flexibility from different generation and consumption points. LEC is a reliable source for flexibility since it can shift the demand up or down. This thesis is focusing on assessing flexibility provision from LECs either directly to a DSO or indirectly to an aggregator company.

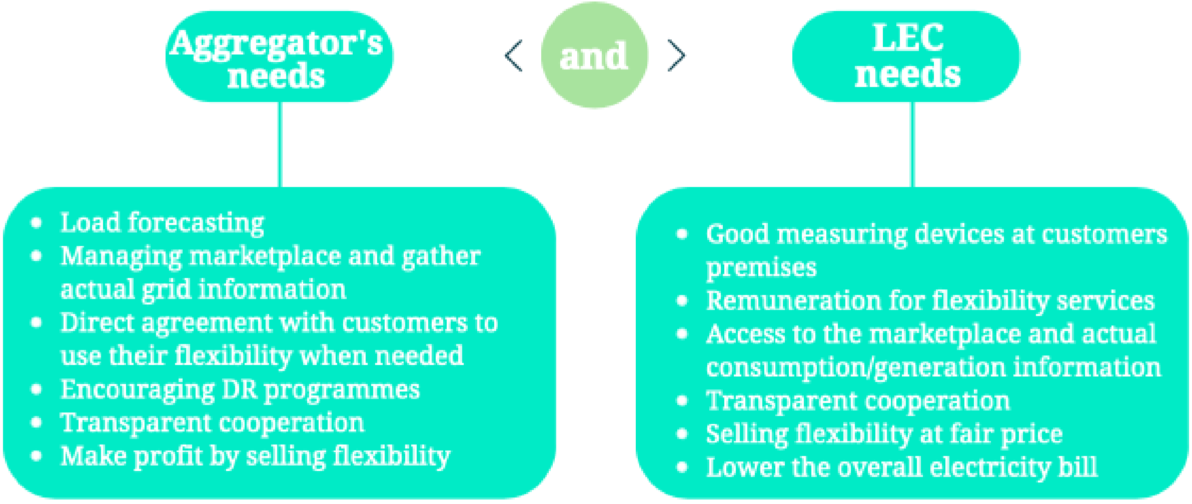


Figure 29 Aggregators and LECs relationship

Business models associated with an aggregator:

1. Community flexibility aggregation. LEC can collaborate with an aggregator, monetizing on the flexibility provision. The aggregator can offer flexibility services directly to the DSO or as an electricity retailer through the different markets available (frequency regulation, reserves and other ancillary services) so that this can balance its portfolio and therefore avoid deviation penalties [37]. Nowadays it is no longer difficult to access flexibility markets for small-scale generation, since European directives advising to reduce minimum generation requirements and bureaucracy associated with that. Residential demand flexibility has become commercially profitable for LECs.

One of the BM arrangements is: community aggregators may be created to operate at a local level and the flexibility collected is grouped by a larger aggregator. Alternatively, community aggregators can also operate directly at the PS level, provided they are able to meet the required conditions. Bilateral contracts should be signed between community aggregators and LECs through which LECs commit to deliver fixed amounts of flexibility by changing energy consumption patterns, benefiting from reduced electricity bills, and using electricity more efficiently [40].

7.1.3. LECs and electrical vehicles

Smart mobility and EVs represent a good asset for solid business models. The new BM arise after the transition from the current context of large amounts of individually-owned and inefficiently-operated EVs, to small amounts of smartly and system-integrated, collectively-owned EVs. By deploying a LEC which brings together users and other public transport infrastructure allows a more satisfactory coverage of the mobility services demand while reducing the energy requirements and emission implications associated with the manufacture of all these individually-owned vehicles [38].

Some business models from this sector are:

1. A LEC can monetize on participation in such an organic mobility network by providing demand response through Vehicle to Grid (V2G), Vehicle to Building (V2B), Vehicle to Home (V2H), Vehicle to Vehicle (V2V), or Vehicle to Load (V2L). V2G working principle is: if an EV owner is participating in V2G program, he or she is able to sell electricity to the grid during hours when the car is not in use (or high tariff times), and to charge the car during low tariff times. It will also be possible to connect cars with V2G technology to a house and use them as an uninterruptible power supply, in this case it is called Vehicle to Building (V2B).

2. Another BM is vehicle sharing solutions, carpooling. E-mobility cooperatives are created by engaging shareholders (households, SMEs, public entities, social and technical entrepreneurs, etc.) to provide community public transportation, car-sharing or car-pooling services. A working example of such a BM is a Spanish mobility cooperative called Som Mobilitat, which provides rental service of electric cars, with EVs which can be either owned by the cooperative or by individuals, enterprises and public institutions [40].

3. Smart charging. In energy communities with high shares of EV (like in the case with the houses from Pecan Street data), smart charging schemes can be designed to schedule load operation to off-peak times or when local energy generation is available, thus optimizing the utilization of local resources and flattening demand peaks [39].

7.1.4. LECs and other energy market participants

A LEC can sell on different market types as freely as other parties. If LEC members are not willing to participate in day-to-day trading, they can sign a contract with an electricity retailer in which an electricity retailer can trade on behalf of a LEC.

1. Local energy markets (LEM) is another BM that can bring together different energy market participants – consumers, prosumers, LECs, energy traders, etc. In LEM, trading conditions, i.e. pricing, can be directly negotiated among market participants (prosumers, consumers, LECs,), allowing LECs to select to whom they sell their electricity and consumers to choose the market participant they buy their this electricity from, at the same time as they know how it is generated [40].

LEM is established to promote P2P energy exchanges either in a fully decentralized way, allowing community members to freely negotiate with each other, or more centrally, through intermediate entities [40].

8. Conclusions and future work

8.1. Main contributions from the thesis

This thesis researched novel energy market participants - Local Energy Communities both theoretically and practically. As regards the theoretical part, the author tried to look at the problematic of LECs from different perspectives: the extensive literature review, legal framework in EU, advantages and disadvantages of LECs, current definition of LECs, their new roles, possibilities, opportunities and obligations, legal framework of LEC. The research was done on the relationships between LECs and their business partners, as well as selected business cases were described. The particular part of the thesis was dedicated to the flexibility provision: state of the art, flexibility provision methods, appliances behavior and classification, as well as description and consumption curve visualization of the appliances from analyzed dataset.

The main goal of the practical part was to investigate to which extent LECs can be actively involved in the flexibility market. Machine learning forecasting capabilities were used in order to enable LECs to forecast their future collective as well as individual load. The ultimate goal then is to be able to operate on the local flexibility market. For flexibility provision, LECs are then financially remunerated as well as the electricity use becomes more efficient.

The main contributions of the thesis are summarized below:

- DATA ANALYTICS AND VISUALISATION. The topic of flexibility provisions from LECs was extensively studied both theoretically and practically. The practical part consists of two main tasks. The first being to perform basic data analytics, visualization, comparison and interpretation, and the second is electricity load forecasting using ML techniques. In order to achieve these goals, the author made use of the dataset, which was requested by the Pecan Street company. The analyzed data consisted of electricity flow recordings of 25 different houses from New York. The data were very precise and allowed the author to perform data visualization and comparison both for individual houses with detailed electricity consumption from different appliances and to perform collective analysis and comparison.

- LOAD FORECASTING USING ML. As regards the second goal of load forecasting, the author made forecasting of overall electricity consumption and generation of one representative house number 3000 during the period of 6 month period (Figure 16 and Figure 17). Another forecasting that was performed used one energy-intensive household device - electrical clothes dryer from the house the number 3000 (Figure 26 and Figure 27); forecasting was done for 6 month period as well. The process of ML model creation is composed of data pre-processing and interpretation, training the model, adjusting hyperparameters, and finally results in interpretation.

- TECHNIQUES USED FOR FORECASTING AND THEIR ACCURACY. Two main ML techniques were used for load forecasting, namely: LSTM and simple RNN. A theoretical introduction of both techniques was done. In the first case or overall house energy load forecasting, the simple **RNN model performed slightly better than LSTM model** (assessment metrics - performance parameter R^2), however for the second case of individual devices consumption forecasting, the picture was reverted, and **LSTM showed better final results than RNN** (again assessment metrics - performance parameter R^2).

- FLEXIBILITY ANALYSIS. The analysis of the electrical appliances' behavior and classifications according to their operation mode, showed us that ML forecasting methods and data analysis methods could be used both to forecast and make sense of the consumption and/or generation data.

- BUSINESS MODELS REVIEW. Relationships between LECs and other energy markets participants (DSOs and TSOs, aggregators, EV charger owners, generation assets owners, other LECs and microgrids) were described. The most promising potential business models were identified and described paying attention to their challenges and development strategies. The identified business models have different objectives and stakeholders involved.

8.2. Future work

This thesis focus field was flexibility provision from LECs. In the future, the research and experiments can be done on the following topics:

1. An extensive financial analysis on potential energy and monetary savings both for LECs, aggregators or DSOs in the context of LECs operating on the flexibility market and aggregating community flexibility.

2. Software and hardware model development and testing for automatic flexibility provision from LECs. Automatic dispatch, decision-making algorithms based on load forecasting.

3. Interviews with existing LECs on the topic of issues associated with their operation in the context of flexibility.

Appendix A - Temporary planning and costs

The budget of this project includes only the time and personnel devoted to the project. The laptop and internet connection is included in the salary of the personnel.

In the process of writing the thesis, no laboratory equipment or paid computer software was used. Therefore, the main component of potential total costs of the thesis is the working hours of personnel. The personnel cost breakdown is done accounting to the time devoted by the author to develop the thesis (1 in total), as well as the time of external consultants (2 in total)

In order to calculate the thesis author costs, we should make a few assumption, and the following approach was taken. The Master's Thesis at UPC - Universitat Politècnica de Catalunya at the faculty ETSEIB is equal to 30 ECTS (European Credit Transfer and Accumulation System). According to the European Commission [19], one ECTS equal to about 28 hours of study. That makes 840 hours in total. The thesis was written during January-June 2021. The thesis was written in Spain, where the salary for a junior engineer position ranges between 22-27 €/. According to glassdoor.com, in Barcelona, this amount is equal to 23 €/h (with VAT).

Regarding the external consultants, who can be considered superior energy engineers with an average salary of 45€/h. Taking into account the information above, Table 4 and Table 5 can be proposed:

Table 4 Number of hours dedicated to Thesis development by the author

Months	**Thesis Author Working hours**	**Consultants Working hours**
January	152	2
February	152	3
March	176	30
April	160	30
May	160	15
Jun	40	10
Total	**840**	**89**

Table 5 Total project cost

Thesis author cost, before VAT reduction	**External consultants cost, before VAT reduction**	**Total , before VAT reduction**	**VAT (21%)**
840 * 23 = 19 320 €	89 * 45 = 4 005 €	23 325 €	4 898.25 €

Appendix B - Environmental Impact

This thesis had no significant impact associated with the environment. However, there are several aspects we can consider as a potential thread to the environment: electricity usage to charge personal computer, electricity needed to run servers at which coding for machine learning was trained, since the model uses significant amount of energy to perform the machine learning tasks.

Therefore, to summarize, energy consumption sources that were used during the thesis are: personal computer, cloud storage and cloud-based open-source coding software. Therefore, it is clear that machine learning has carbon footprint. The Thesis was written on MacBook Air (13-inch, Early 2015) with processor 1,6 GHz Dual-Core Intel Core i5, which average power consumption under load is equal approximately to 22 W. Knowing this we multiply this number with number of hours used to code, verify and run the training of machine learning algorithms in Python using Jupyter Notebook, which was equal to approximately to 100h. Then, using the source [45] on electricity emissions factors, which in Spain is equal to 0.309 kgCO2e per kWh, it is considered to be the average electricity emissions factor in Europe. Multiplying all three numbers gives as:

$$\text{Power consumption} * \text{Time} * \text{Carbon Factor in Spain}$$

$$22W * 100h = 22\ \text{kWh} * 0.309\ \text{kg eq.CO2/kWh} = 6.8\ \text{kg eq.CO2}$$

Therefore, the CO2 emissions associated with the training of the machine learning model for this thesis is equal to **6.8 kg eq. CO2**. It is expressed in carbon dioxide equivalent which is usually used to describe different greenhouse gases with a common unit (CO2e).

However, the environmental damage (or CO2 produced during the electricity generation to power the servers and charge the computer) can be compensated by the rational use of energy with the help of load and generation forecasting that was proposed in this thesis.

Acknowledgments

Throughout the writing of this Master's Thesis, I have received great support, assistance, and help. Therefore, I would like to express my deep gratitude to all those who took part in the development, research, preparation and review of my Master's Thesis.

First of all, let me express my gratitude to my supervisor Professor Andreas Sumper for the opportunity to work on the topic that interests me and for his valuable pieces of advice throughout the thesis preparation.

Secondly, I would like to express my sincere gratitude to my scientific advisor Alejandro Hernandez Matheus for your time, insightful feedback and help at all stages of the thesis (including the coding part). You provided me with the tools that I needed to choose the right direction and successfully completed my thesis.

I would like to sincerely thank the leading organization – European Institute of Innovation and Technology (EIT Innoenergy) for the financial support of my journey in the Energy for Smart Cities Master's Program. Without this support, I would not be writing this Thesis.

I would like to thank my mother Tatiana Maksimovich, sister Zhanna Maksimovich and father Evgeny Maksimovich for their mental support during the whole master's program – this would be impossible without them.

Bibliography

[1] European Commission, "Directive (EU) 2019/944 of the European Parliament and of the Council of 5 June 2019 on common rules for the internal market for electricity and amending" Directive 2012/27/EU (Text with EEA relevance.), Pub. L. No. 32019L0944,L 158.

[2] European Commission, "Regulation (EU) 2019/943 of 5 June 2019 on the internal market for electricity", Official Journal of the European Union, vol. 62, no. L 158, pp. 54–124, 2019.

[3] Mladen Kezunovic, Pierre Pinson, Zoran Obradovic, Santiago Grijalva, Tao Hong, Ricardo Bessa, "Big data analytics for future electricity grids", Electric Power Systems Research, Volume 189, 2020,106788, ISSN 0378-7796, https://doi.org/10.1016/j.epsr.2020.106788.

[4] NVE (Norges vassdrags- og energidirektorat) (2019): "Descriptive study of Local Energy Communities". A report commissioned by NVE. THEMA Consulting Group Multiconsult Norge AS. Ekstern Rapport Nr 1/2019.

[5] Pecan Street Inc. 2021. PECAN STREET – Pecan Street Inc. [online] Available: <https://www.pecanstreet.org> [Accessed 6 March 2021].

[6] JRC Photovoltaic Geographical Information System (PVGIS) - European Commission. [online] Available: <https://re.jrc.ec.europa.eu/pvg_tools/en/#DR> [Accessed 8 March 2021].

[7] W. Kong, Z. Y. Dong, Y. Jia, D. J. Hill, Y. Xu, and Y. Zhang, “Short-term residential load forecasting based on LSTM recurrent neural network,” IEEE Transactions on Smart Grid, vol. 10, no. 1, pp. 841–851, 2019.

[8] Daftlogic.com. 2021. Power Consumption of Typical Household Appliances. [online] Available at: <https://www.daftlogic.com/information-appliance-power-consumption.htm> [Accessed 11 March 2021].

[9] Peter Palensky and Dietmar Dietrich. 2011. "Demand Side Management: Demand Response, Intelligent Energy Systems, and Smart Loads." IEEE Transactions on Industrial Informatics 7, 3 (2011), 381–388.

[10] R. Ahmadiahangar, T. Häring, A. Rosin, T. Korõtko and J. Martins, "Residential Load Forecasting for Flexibility Prediction Using Machine Learning-Based Regression

Model," 2019 IEEE International Conference on Environment and Electrical Engineering and 2019 IEEE Industrial and Commercial Power Systems Europe (EEEIC / I&CPS Europe), Genova, Italy, 2019, pp. 1-4, doi: 10.1109/EEEIC.2019.8783634.

[11] IEA, Largest end-uses of energy by sector in selected IEA countries, 2018, IEA, Paris. [online], Available at: https://www.iea.org/data-and-statistics/charts/largest-end-uses-of-energy-by-sector-in-selected-iea-countries-2018, [Accessed 23 March 2021].

[12] Walker, Gordon & Devine-Wright, Patrick. (2008). "Community Renewable Energy: What Should It Mean? " Energy Policy. 36. 497-500. 10.1016/j.enpol.2007.10.019.

[13] Susana Soeiro, Marta Ferreira Dias, "Community renewable energy: Benefits and drivers", Energy Reports, Volume 6, Supplement 8, 2020, Pages 134-140, ISSN 2352-4847, https://www.sciencedirect.com/science/article/pii/S2352484720315122

[14] R. D'hulst, W. Labeeuw, B. Beusen, S. Claessens, G. Deconinck, K. Vanthournout, "Demand response flexibility and flexibility potential of residential smart appliances: Experiences from large pilot test in Belgium", Applied Energy, Volume 155, 2015, Pages 79-90, ISSN 0306-2619, https://doi.org/10.1016/j.apenergy.2015.05.101.

[15] P. Palensky and D. Dietrich, "Demand Side Management: Demand Response, Intelligent Energy Systems, and Smart Loads," in IEEE Transactions on Industrial Informatics, vol. 7, no. 3, pp. 381-388, Aug. 2011, doi: 10.1109/TII.2011.2158841.

[16] Fernando Lezama, Joao Soares, Bruno Canizes, Zita Vale, "Flexibility management model of home appliances to support DSO requests in smart grids, Sustainable Cities and Society", Volume 55, 2020, 102048, ISSN 2210-6707, https://doi.org/10.1016/j.scs.2020.102048.

[17] Basit, G. A. S. Sidhu, A. Mahmood and F. Gao, "Efficient and Autonomous Energy Management Techniques for the Future Smart Homes," in IEEE Transactions on Smart Grid, vol. 8, no. 2, pp. 917-926, March 2017, doi: 10.1109/TSG.2015.2504560.

[18] Gouveia, João & Seixas, Julia & Luo, Shiming & Bilo, Nuno & Valentim, António. (2015). "Understanding electricity consumption patterns in households through data fusion of smart meters and door-to-door surveys". Available at: https://www.researchgate.net/publication/277879669_Understanding_electricity_consumption_patterns_in_households_through_data_fusion_of_smart_meters_and_door_to_door_surveys

[19] Education and training. (2021, January 26). European Commission. [online]. Available at: https://ec.europa.eu/education/resources-and-tools/european-credit-transfer-and-accumulation-system-ects_en, [Accessed 23 April 2021].

[20] Vasco Brummer, "Community energy – benefits and barriers: A comparative literature review of Community Energy in the UK, Germany and the USA, the benefits it provides for society and the barriers it faces", Renewable and Sustainable Energy Reviews, Volume 94, 2018, Pages 187-196, ISSN 1364-0321, https://www.sciencedirect.com/science/article/pii/S1364032118304507

[21] Harshit Vallecha, Debraj Bhattacharjee, John Kalu Osiri, Prabha Bhola, "Evaluation of barriers and enablers through integrative multicriteria decision mapping: Developing sustainable community energy in Indian context", Renewable and Sustainable Energy Reviews, Volume 138, 2021, 110565, ISSN 1364-0321, https://www.sciencedirect.com/science/article/pii/S1364032120308492

[22] Morgan MG, Zerriffi H. "The regulatory environment for small independent microgrid companies. " Electr J 2002;15:52–7. https://doi.org/10.1016/S1040-6190(02) 00385-8

[23] Hossein Abbasimehr, Mostafa Shabani, Mohsen Yousefi, "An optimized model using LSTM network for demand forecasting, Computers & Industrial Engineering", Volume 143,2020, 106435, ISSN 0360-8352, https://www.sciencedirect.com/science/article/pii/S0360835220301698

[24] Eugene A. Feinberg, Dora Genethliou, "Load Forecasting", Applied Mathematics for Power Systems, State University of New York, Stony Broo

[25] Dr. Chandrasekhar Reddy Atla "Load Forecasting: Methods & Techniques PRDC" Available at: https://www.saarcenergy.org/wp-content/uploads/2018/05/Load-Forecasting-Methods-and-Techniques-by-Dr.-Chandrasekhar-Reddy-Atla-Dr.-Chandras.pdf

[26] Greff, K., Srivastava, R. K., Koutník, J., Steunebrink, B. R., & Schmidhuber, J. (2017). LSTM: A Search Space Odyssey. "IEEE Transactions on Neural Networks and Learning Systems", 28(10), 2222–2232. https://doi.org/10.1109/TNNLS.2016.2582924.

[27] Andrej Karpathy, "The Unreasonable Effectiveness of Recurrent Neural Networks, " May 21, 2015. [online], Available at: http://karpathy.github.io/2015/05/21/rnn-effectiveness/ [Accessed 30 April 2021].

[28] Parmezan, A. R. S., Souza, V. M., & Batista, G. E. (2019). "Evaluation of statistical and machine learning models for time series prediction: Identifying the state-of-the-art and the best conditions for the use of each model. " Information Sciences, 484, 302–337.

[29] Shahzad Muzaffar, Afshin Afshari, "Short-Term Load Forecasts Using LSTM Networks" Energy Procedia, Volume 158, 2019, Pages 2922-2927, ISSN 1876-6102, https://www.sciencedirect.com/science/article/pii/S1876610219310008

[30] Mallam, S. (2019). Hourly energy consumption time series RNN, LSTM. [online] Avaliable at: https://www.kaggle.com/msripooja/hourly-energy-consumption-time-series-rnn-lstm [Accessed 25 February 2021]

[31] Reimers, N., & Gurevych, I. (2017). “Optimal hyperparameters for deep LSTM-networks for sequence labeling tasks“. arXiv preprint arXiv:1707.06799.

[32] Keras, [online], Available at: https://keras.io/api/losses/, [Accessed 1 April 2021].

[33] Jason Brownlee, "How to Scale Data for Long Short-Term Memory Networks in Python", on July 7, 2017 in Deep Learning for Time Series. Available at: https://machinelearningmastery.com/how-to-scale-data-for-long-short-term-memory-networks-in-python/

[34] Chollet, François, "Deep Learning with Python", Publisher: Manning Shelter Island, NY, 2018, Manning Publications Co., ISBN: 9781617294433.

[35] Burnett, Michael, "Energy Storage and the California, Duck Curve", June 1, 2016 Submitted as coursework for PH240, Stanford University, 2015

[36] Caramizaru, A. and Uihlein, A., “Energy communities: an overview of energy and social innovation“, EUR 30083 EN, Publications Office of the European Union, Luxembourg, 2020, ISBN 978-92-76-10713-2, doi:10.2760/180576, JRC119433.

[37] Casals, X, Sanmartí, M, Salom J, “Smart Energy Communities: Insights into its structure and latent business models”, June 2019, Institut Català d’Energia -Institut de Recerca en Energia de Catalunya (IREC), available at: icaen.gencat.cat/publicacions/

[38] Garcia-Casals X., “A transition analysis for a 1.5C Climate Vision“, 2017b Available at: http://xavigarciacasals.blogspot.com.es/2017/08/a-transition-analysis-for-15c-climate.html

[39] Brown D, Hall S, Davis ME. “Prosumers in the post subsidy era: an exploration of new prosumer business models in the UK“. Energy Pol 2019;135:110984. https://doi.org/10.1016/j.enpol.2019.110984

[40] Inês F.G. Reis, Ivo Gonçalves, Marta A.R. Lopes, Carlos Henggeler Antunes, “Business models for energy communities: A review of key issues and trends“, Renewable and Sustainable Energy Reviews, Volume 144, 2021, 111013, ISSN 1364-0321, https://doi.org/10.1016/j.rser.2021.111013.

[41] Q. Yan, T. Dokic, M. Kezunovic, “Predicting impact of weather caused blackouts on electricity customers based on risk assessment“, IEEE Power and Energy Society General Meeting, Boston, MA, 2016.

ETSEIB

[42] Eda Kavlakoglu, Program Manager, AI vs. Machine Learning vs. Deep Learning vs. Neural Networks: What's the Difference?, [online] Available at: https://www.ibm.com/cloud/blog/ai-vs-machine-learning-vs-deep-learning-vs-neural-networks [Accessed 3 May 2021]

[43] Solargis. (n.d.). Global solar atlas. [online] Available at: https://globalsolaratlas.info/map?c=37.987504,-80.145264,6 [Accessed 10 May 2021]

[44] Jason Brownlee. "Time Series Forecasting Performance Measures With Python in Time Series" (2017, February 1). Machine Learning Mastery, [online]. Available at: https://machinelearningmastery.com/time-series-forecasting-performance-measures-with-python/ [Accessed 12 June 2021].

[45] Carbon Footprint Ltd report, Country Specific Electricity Factors from August 2018 https://www.carbonfootprint.com/docs/2018_8_electricity_factors_august_2018_-_online_sources.pdf [Accessed 18 June 2021].