Universidad Nacional Autónoma de México

Facultad de Ciencias

Data science based models for evaluating the performance of microgrids with distributed generation and energy storage

T E S I S

que para obtener el título de

Físico

p r e s e n t a

Eduardo Rafael Martínez Pérez

Director de tesis:

Dr. Miguel Robles Pérez

Ciudad Universitaria, Ciudad de México, 2022

1. Datos del alumno

   Martinez
   Pérez
   Eduardo Rafael
   +52 5518172673
   Universidad Nacional Autónoma de México
   Facultad de Ciencias
   Físico
   314557234

2. Datos del tutor

   Dr.
   Robles
   Pérez
   Miguel

3. Datos del presidente

   Dr.
   del Río
   Portilla
   Jesús Antonio

4. Datos del vocal

   Dra.
   Arriola
   Ríos
   Verónica Esther

5. Datos del suplente

   PhD.
   Alvarado
   Reyes
   José Manuel

6. Datos del suplente

   PhD.
   Santana
   Rodríguez
   Guillermo

7. Datos del trabajo escrito

   Data science based models for evaluating the performance of microgrids with distributed generation and energy storage.
   55 pp
   2022

Índice general

Contents

Acknowledgments

Agradezco a mi padre Eduardo y a mi madre Francis por la vida, el sustento y su amor. A mis abuelos Flora y Ulises por su crianza. A mi hermana Bárbara por su amor y compañía desde el día en que nací. A mi hermano Diego por sus oídos, consejos y los grandes momentos juntos. A Marco por acompañar y apoyar a la mujer que me dio la vida. A mi familia en general por la unión y solidaridad.

Agradezco a Alina por su apoyo, acompañamiento, amor y tolerancia, así como a toda su familia.

Agradezco a mis maestros Rodolfo y Carmen por su guía y enseñanzas sobre la Naturaleza, la Humanidad y las Estrellas. Al Dr. Miguel Robles por recibirme en su grupo de investigación y presentarme una rama del conocimiento que terminó por definir mis intereses profesionales.

A mis amigos y amigas Eduardo, Daniel, Andrea, Daniela y Susana, por las tardes tanto amenas como extenuantes de estudio y trabajo.

A la UNAM y a la Facultad de Ciencias por la esperanza, la formación integral y los recursos brindados.

Al instituto de Energías Renovables por abrirme las puertas a nuevas posibilidades.

A todas las personas que han contribuído a mi educación a lo largo de toda mi vida. A la sociedad mexicana, cuyas contribuciones posibilitaron mis estudios y posibilitan los de tantos más.

A todo el equipo de Grid Singularity por la oportunidad de desarrollarme profesionalmente en una rama del conocimiento y tecnología que me apasiona.

Al Universo por la oportunidad de estar en este momento colaborando por el progreso humano.

I'm grateful to my father Eduardo and my mother Francis for life, support and love. My grandparents Flora and Ulises for raising us. My sister Bárbara for her love and accompaniment from the day I was born. My brother Diego for our great moments together. To Marco for his caring ans support to the woman who brought me to the world. To all of my family for our unity and solidarity.

I'm also grateful to Alina for her support, accompaniment, love and tolerance, as well as to all of her family.

I'm grateful to my teachers Rodolfo and Carmen for their guidance and teachings about Nature, Humanity and Cosmos. And to PhD. Miguel Robles for receiving me in the research team and for introducing me to a branch of knowledge which ended up defining my professional interests.

Also to my friends Eduardo, Daniel, Andrea, Daniela and Susana, for all those good and stressful evenings studying or doing homework.

I'm grateful to UNAM and the Science Faculty, for the integral formation, and provided resources.

To the Renewable Energy Institute for opening the doors allowing new possibilities.

To all people who have contributed to my education during my whole life. To the Mexican society, whose contributions make from study a reality for me and many more Mexicans.

To all of Grid Singularity team for giving me the opportunity to build myself professionally on a field of study I'm passionate about.

To the Universe for the opportunity to co-work towards the progress of humanity.

Abstract

Español

En los últimos años, el sector de la energía ha ido dirigiendose hacia la adopción de tecnologías digitales. Varias industrias las han incorporado para tener una representación virtual de los objetos físicos. En el ámbito del sector energético, ya se han desarrollado modelos virtuales de módulos fotovoltaicos (PV), baterías o incluso generadores de energía eólica, que se han integrado en diferentes aplicaciones de computadora [16]. Una simulación computacional es un buen primer paso para construir una representación digital de un determinado sistema. La información utilizada para la simulación pueden ser puntos de datos reales (que representan el estado del sistema físico desplegado) o datos estimados obtenidos de bases de datos disponibles y reunidos para simular el sistema.

Además de la digitalización de las industrias, debido a la democratización del acceso a tecnologías fotovoltaicas, también ha aumentado la producción distribuída de energía, lo que plantea retos y oportunidades potenciales para mejorar la resiliencia de los sistemas de energía eléctrica existentes, así como para lograr una transición energética socialmente justa [1, 2, 22].

La producción de energía distribuida, los sistemas de almacenamiento de alta capacidad y las soluciones de software modernas permiten construir sistemas a pequeña escala similares a una red eléctrica conformada por consumidores y productores de energía, que comparten localmente los recursos energéticos, también conocidos como Microrredes o Comunidades Energéticas. Algunos consideran que las Comunidades Energéticas son el futuro del sistema energético [4]. Por lo tanto, una mejor comprensión de estos sistemas puede contribuir a desarrollar estrategias de optimización particulares.

Esta tesis presenta un estudio del comportamiento estadístico y del rendimiento de un grupo de comunidades energéticas *simuladas* (microrredes) a una escala nacional. Esto se hace a través de un programa de computadora que imita un sistema de microrredes evaluando la dinámica de la generación de energía. El programa se desarrolló para este trabajo. Se simularon aproximadamente 10.000 comunidades energéticas utilizando diferentes ubicaciones y un número diferente de activos (módulos fotovoltaicos [PV] o sistemas de almacenamiento de energía [ESS]).

El trabajo explora el límite físico del escenario hipotético en el que los recursos se utilizan al máximo para una estrategia de carga-descarga particular que busca reducir picos de demanda. Por lo tanto, diferentes itinerarios de carga-descarga de los ESSs o estrategias de comercio colaborativo conducirán probablemente a resultados diferentes.

Con los datos obtenidos de las simulaciones, se realiza un análisis cualitativo y cuantitativo para evaluar el impacto que tendría una determinada configuración del Sistema de Activos Comunitarios (CAS, por sus siglas en inglés) en una comunidad con necesidades energéticas específicas. El análisis conduce a la definición de un modelo matemático que describe el rendimiento de un CAS de características particulares sobre una comunidad energética. Este último puede utilizarse para estudiar analíticamente el sistema y medir su rendimiento con funciones matemáticas que representan a los indicadores de desempeño pertinentes.

El modelo puede utilizarse eventualmente para evaluar y determinar la configuración más eficiente de un CAS para una comunidad concreta (en términos de coste-beneficio), optimizando el número de activos fotovoltaicos y de ESS que debería tener. Sin embargo, este trabajo explora exclusivamente la perspectiva energética, dejando las métricas económicas de los mercados locales para futuras investigaciones.

El análisis y la modelización propuestos pueden reproducirse para caracterizar un CAS con cualquier tipo de algoritmo de control para las baterías o de transacciones de energía entre los miembros de la comunidad. Como se ha mencionado, el modelo también puede ampliarse para incluir el componente económico para evaluar mejor (y por tanto optimizar) el ahorro energético y los beneficios monetarios.

Uno de los principales logros de este proyecto es el desarrollo de una herramienta de cómputo científico para simular microrredes. El pograma de computadora contiene un algoritmo de reducción de picos para los ESSs. El algoritmo mencionado también representa un logro, al igual que la metodología utilizada para analizar los datos obtenidos de la simulación. Por último, otros logros significativos son el análisis visual que se presenta y el modelo matemático que describe el rendimiento del sistema.

Se analizan y presentan los resultados de las simulaciones, proporcionando una visión de cómo los indicadores de desempeño se correlacionan con el requerimiento de energía de una Comunidad Energética de 10 casas. Los resultados muestran que la ubicación geográfica se correlaciona principalmente con el requerimiento de energía de la comunidad. Esto se debe, en parte, a los diferentes aparatos eléctricos que se utilizan en zonas climáticas con condiciones locales estacionales variadas. Por lo tanto, la ubicación influye en las necesidades energéticas de una comunidad y, de forma indirecta, en el rendimiento de un CAS de tamaño determinado.

El modelo matemático puede utilizarse para evaluar y comparar cuantitativamente el rendimiento de CAS de diferente tamaño y determinar la cantidad de activos más apropiada, tomando como referencia los requerimientos de energía de la comunidad.

English

In recent years, the energy sector has been facing a transformation into the adoption of digital technologies. Several industries have incorporated them to have a virtual representation of physical objects. In the scope of the energy sector, virtual models of photovoltaic modules (PVs), batteries, or even wind power generators have already been developed and integrated into different software solutions [16]. A computational simulation is a good first step to building a digital representation of a particular system. The information used for the simulation can be either real data points (representing the state of the existing deployed physical system) or estimated data obtained from available databases and put together to build a representation of the simulated system.

Apart from the digitalization of the industries, due to the democratization of PV technologies, distributed energy production has also increased, presenting challenges and potential opportunities to improve the resiliency of existing electric power systems and a social-just energy transition [1, 2, 22].

Distributed energy production, high-capacity storage systems, and modern software solutions enable the possibility of building small-scale grid-like systems conformed by energy consumers and producers, which locally share the energy resources, also known as Microgrids or Energy Communities. Some consider Energy Communities to be the future of the energy system [4]. Therefore, a better understanding of such systems can contribute to developing particular optimization strategies.

This thesis presents a study of the statistical behavior and performance of *simulated* energy communities (microgrids) at a country scale level. This is done via software that mimics a microgrid system by evaluating the dynamics of the energy generation. The simulation program was also developed for this work. Approximately 10,000 energy communities were simulated using different locations and a different number of assets (photovoltaic modules [PVs] or Energy Storage Systems [ESSs]).

The work explores the physical boundary of the hypothetical scenario in which the resources are used as much as possible for a peak shaving/shifting charge-discharge battery strategy. Therefore, different ESS charge-discharge itineraries or collaborative trading strategies will likely lead to different results.

With the data obtained from the simulations, a qualitative and quantitative analysis is made to assess the impact a particular Community Assets System (CAS) configuration would have on a community with specific energy requirements. The analysis leads to the definition of a mathematical prediction model, which describes the performance of a CAS of particular characteristics on an energy community. The latter can be used to analyze further and measure the system's behavior and performance with relevant Key Performance Indicators analytically.

The model can eventually be used to evaluate and determine a CAS's most efficient cost-benefit configuration for a particular community by optimizing the number of PV and ESS assets it should have. However, this work exclusively explores the energetic perspective, leaving the local markets' economic metrics for further research.

The proposed analysis and modeling can be replicated to characterize any CAS with any type of trading algorithm or an alternative ESS strategy. As mentioned, the model can also be expanded to include the economic component to evaluate better (and therefore optimize) energy savings and monetary benefits.

One of the main achievements of this project is developing a scientific software tool to simulate microgrids. The software has an integrated ESSs peak shaving/shifting management algorithm. It also represents an achievement, as does the methodology used to analyze the resulting simulation data. Lastly, another significant achievement is the visual analysis and the mathematical model for its performance description modeling.

The simulations' results are analyzed and presented, providing insight into how the KPIs correlate with the energy requirement of a 10-Load Energy Community. Results show that the geographic location mainly correlates with the community's energy requirements. This is partly because different electric appliances are used in climate zones with variate average local seasonal conditions. Therefore, the location influences a community's energy requirement and indirectly impacts the performance under a CAS of a particular size.

The prediction model can be used to evaluate and quantitatively compare the performance between different CAS sizes and determine the most appropriate size based on particular performance benchmarks and energy requirements of the community under study.

Chapter 1

Introduction

1.1 The Energy Sector

The energy industry moves the economy and the modern world. Almost every activity in our daily routines depends on electricity coming into our homes and workspaces. All industries rely on the availability of energetic goods, such as oil, gas, and carbon. The upside of these resources is that they provide a continuous power supply; as for the downside, the air we breathe is also continuously polluted.

Recently, renewable energy sources have had an enormous boost. Big solar power plants and wind farms have been built. They are already used to supply energy to various groups of people around the world [21].

The electricity markets have two main actors: the producers and the consumers. One provides the resources to the other, who pays for them. In recent years, the availability of more affordable Photovoltaic Modules (PVs) has allowed consumers to turn themselves into a hybrid of consumer and producer (prosumer). Also, integrating Energy Storage Systems (ESSs) or batteries increases a community's self-consumption rate and, consecutively, its self-sufficiency. This corresponds to a distributed way of producing the energy consumed by our civilization.

The current scheme in many places has been so far centralized in big power plants that produce electricity for many. The power is later distributed to the final consumer through the transmission lines, a process in which a significant energy loss occurs. Furthermore, several difficulties are present in managing the power distribution, such as congestions on the grid nodes, due to the physical limitations of the infrastructure.

The power delivery depends on many factors before it reaches the end-user, which causes it to be a low-resilient system, with a propensity to bring power outages on high demands or system failures. In the decentralized scheme, Households can now turn themselves into actors in a Local Energy Market (LEM), where energy is produced on-site, managed, and traded locally.

1.1.1 Decentralized Energy Resources and the Smart Grid

Electric circuits have two main types of elements: active and passive components. Active components are those that supply energy to the circuit. On the other hand, passive components can only receive energy, which could be either dissipated or stored in some manner. Diodes and capacitors are examples of active and passive components, respectively. At a grid scale, power generators are active components of the grid, while power-consuming devices are passive components. Modern high-storage-capacity batteries can be considered active and passive, depending on whether they charge or discharge.

Decentralized Energy Resources or DERs are those made available in a distributed manner. Two of the most used technologies are the PVs and ESSs. Managing the energy resources optimally is a vital piece of the work to make the most out of this kind of system.

The way energy is managed in such a scheme tends to be digital. This requires software development to couple the physical assets with a computer-like brain/engine capable of performing specific actions based on a pre-configured behavior or a reactive configuration. The number of Internet of Things (IoT) devices that can connect to the internet and perform actions based on cloud applications is increasing [9]. PVs and ESSs, among other technological appliances, are no exception. However, having power devices in a grid-connected facility requires the whole system to be managed accordingly, with a digital, intelligent brain.

A Smart Energy Management Systems (SEMS) aims to provide an integrated solution for the opportune management of energy-related IoT assets [5]. Participants of the electricity markets and governments are investing in the digitalization of the energy sector by implementing such management tools in the transmission, distribution, and consumption processes. Some of which aim to be energy trading systems, Load balancing agents, or dashboards that facilitate data management for such purposes. As an instance of a recent SEMS application, the growing electric vehicle (EV) industry has innovated around the EVs charging optimization by managing charge times and speeds intelligently for large vehicle fleets in different parts of the world. Also, in recent years the ideas and implementations of microgrids or Energy Communities (EC) have gained traction. Several technological solutions to implement and manage EC are currently being developed in what appears as a race to gain market share in the future of a renewed energy sector, betting on future regulations that will allow this.

1.1.2 Local Energy Markets

An energy market can exist wherever there are a producer and a consumer. On the centralized scheme, energy is traded on the wholesale electricity market, which is only accessible by a selected group of companies. These companies are energy producers and retailers who buy and sell electricity in bulks.

On the contrary, the spread and adoption of DERs technologies have directed the innovation into developing tools to enable and manage local markets, which are part of a more extensive hierarchical market structure. A community-level energy market can be created for households to trade in and share resources. It has been proven in theory and practice that trading in Local Energy Markets (LEMs) increases a community's self-consumption; in other words, it helps optimize the usage of locally produced energy [13].

1.2 Microgrids or Energy Communities

An Energy Community or microgrid can be considered as a set of consumers, producers, and possibly prosumers who can interact with each other by exchanging energy resources. In a real case scenario, a house can be considered a microgrid, as it is undoubtedly an energy consumer. If we add to this house a PV panel, then this two-element system would also be a microgrid, as the house Load interacts with the PV (by consuming the energy it produces). At the same time, the PV interacts with the house Load by transferring the energy the house needs. Then, if this house had a battery, this element could act at different times as a consumer and as a sort of producer (from the perspective that it can inject energy back into the system). This three- element system would also be a microgrid: the PV could then decide (be configured) to exchange with the house Load, the battery, or both. When there is no PV production, the battery could exchange energy with the Load from its internal storage, and so on.

Such objects (Loads, PVs, batteries, wind turbines, and so on) are sometimes called *assets*. Between every asset on the grid/community, the possibility to exchange with whoever else is participating enables several markets where energy can be bought and sold. This democratizes access to the energy markets, as there would be no need to be a large producer to participate.

On [22], Thombs discusses four plausible outcomes for the future energy system, classified according to two bipolar paradigms: democratic vs. monopolistic and centric vs. decentralized. He thinks that the most socio-ecological and sustainable configuration is necessarily democratic and formed by an adequate combination of centric and decentralized schemes. Local Energy Communities can contribute to the democratization of future energy systems. However, planning and regulation should be made carefully to ensure that nobody is excluded from the well-being and benefits (social, economic, and ecological) that could surge from these new structures, achieving an indeed just energy transition.

Going back to the example, in a city, a microgrid could look like a set of buildings that may or may not have several assets that could also participate in the exchange. For instance, a school, a six-apartment residential building with PV generation, and a small house with a battery system or an electric automobile (which has a battery and thus could be considered an asset) could form a local microgrid or energy community. At the simplest, this example reduces to a group of buildings with a particular Load consumption pattern, which could also have energy production and storage capability. Each of those buildings could have any kind of DER asset, such as PV, ESS, a small wind power generator, etc. Due to their specific energy production and consumption patterns, self- consumption will vary. Each household's energy surplus can be traded within the community, for which another member may be willing to pay. This idea is illustrated in figure 1.1

There are several Key Performance Indicators (KPIs) used to characterize an energy community's performance. In section 3, they will be further discussed and analyzed. One of the most relevant KPIs is the self-sufficiency, which measures the autonomy level of a community. A community that is entirely grid-dependent has a 0% self-sufficiency rate. A completely independent community has a 100% self-sufficiency rate. Everything in between is partially dependent and probably corresponds to the prosumer type of a LEM participant. In a community-level energy market, the resources of the houses which participate in it are shared, obtaining a general profit for the community and higher levels of self-sufficiency from the wider grid.

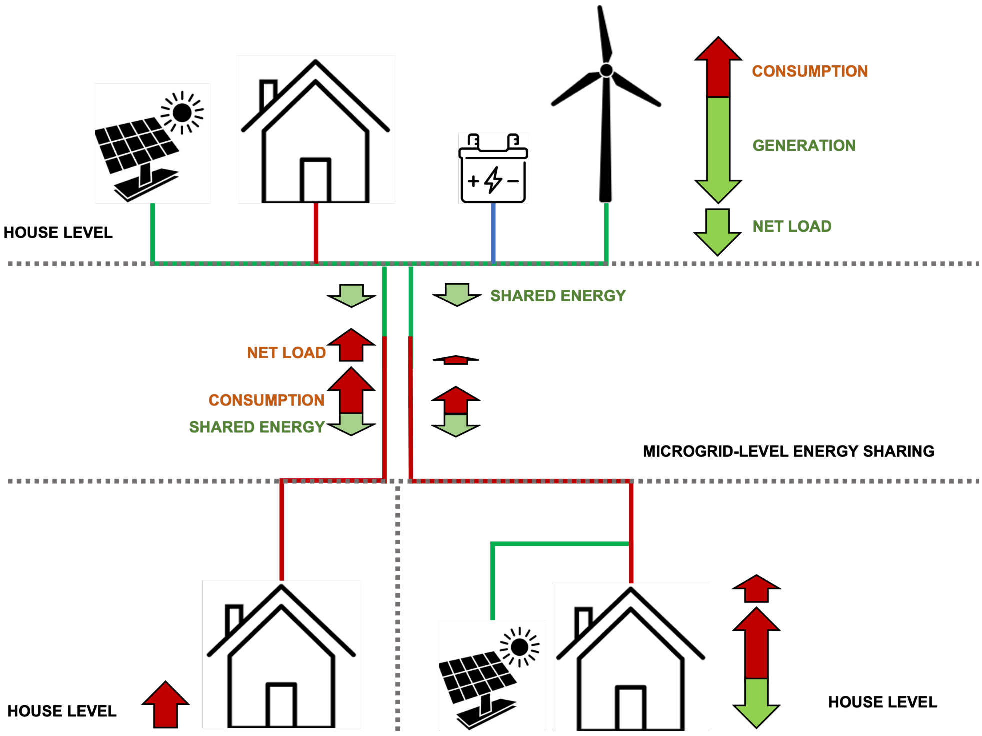


Figure 1.1: Small 3-house community in which locally produced energy surplus is shared between its members.

1.3 Additional Context

This section presents additional information on the relation between power and energy to support the model in chapter 3. A brief description of the databases and a simplified explanation of data processing techniques are also given. The next chapter will dive deep into the models.

1.3.1 Energy and Power

Power systems can have a historical record of the power demand or generation. This data is shaped as a time series, in which a value for power corresponds to every time slot. Power is measured in Watts and is a measure of the amount of energy being transferred every second. The total energy (E) consumed on a time interval for a particular power Load profile ($P(t)$) is:

$$E = \int_{t_1}^{t_2} P(t)dt. \tag{1.1}$$

For discrete systems (the time series data is a discrete curve), the integral is approximated by a Riemann sum of constant increments Δt. The data processing described in the following chapter was done with 15-minute-resolution time series.

By calculating the numeric integral for such signals, we can measure relevant quantities such as the produced energy, and the total consumption, among others, for a particular time interval. This work presents the analysis for a one-year time interval.

1.3.2 Databases details

The following two sections describe at a high level how the databases are conformed and some characteristics of each of them. This work uses data from the US mainly because of data availability. Although there are more data sources from different countries, at the time research on available data was conducted, the following databases were the most adequate for our objectives.

OpenEI's Residential Energy Consumption Database

Open Energy Initiative (OpenEI) is a funded North American organization that provides public data sets, aiming to provide the raw materials with which energy systems modeling research is conducted [15]. The database used is called *Commercial and Residential Hourly Load Profiles for all TMY3 Locations in the United States.* Its documentation describes the considerations that were made when simulating the Load profiles.

As a general description, the consumption profiles are modeled based on climatic zones. For each climatic zone shown in figure 1.2, different combinations of commodities characteristics are used. For instance, for one climate zone, heating could be done by local natural gas combustion, and for another, by using electricity. Also, the average house uses different electric devices for each climate zone, resulting in different consumption habits and electric energy requirements.

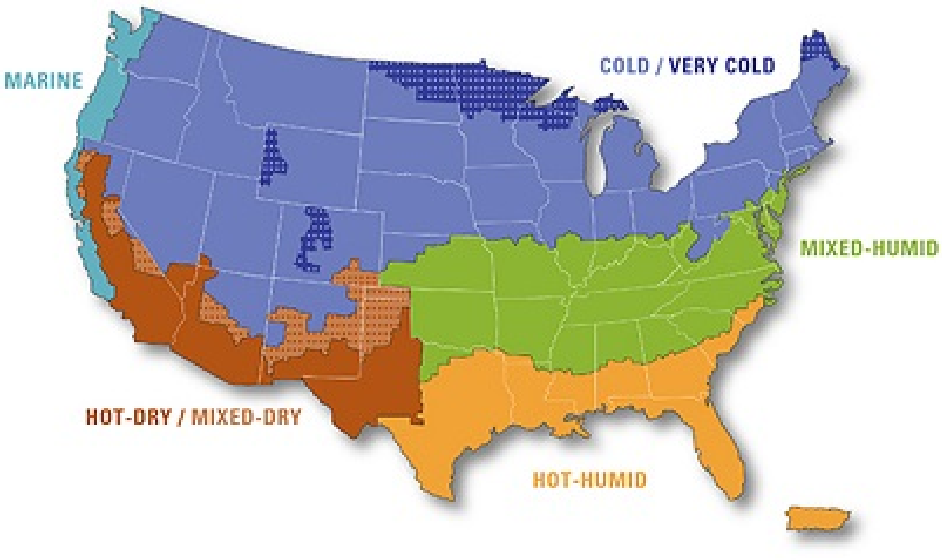


Figure 1.2: USA Climate Map.[12]

The energy requirements of a community depend on the houses' electric appliances and the Load level for each household (which can be base, low, or high and correlates with the socioeconomic factor). This work only considers power consumption and discards other types of energy consumption, such as local fossil fuel combustion.

Power consumption habits also correlate with weather conditions. TMY stands for Typical Meteorological Year and represents the average climatic conditions at a particular location, based on many years of historical data. Based on this, the power demand throughout the year was simulated by considering typical meteorological conditions. Although the database provides commercial facilities' energy consumption information, only residential facilities are used to simulate the energy communities.

NREL's PV power generation plants

The National Renewable Energy Laboratory (NREL) published a one-year PV generation data set for 2006, with actual weather data measurements. The most significant part of the database consists of UPV-type plants, which have single-axis sun tracking technology. The other part corresponds to DPV-type plants, whose inclination is fixed and tilted depending on the latitude to maximize the solar resource. UPV has more simulated plants with greater geographical variability; therefore, that subset was used.

1.3.3 Data science technical terms and models

For all the data processing, some popular data analysis python libraries were used (pandas, NumPy, Scikit Learn, among others). They allow fast table operations for large data sets and other data analysis tools.

K-means is a basic machine-learning-like algorithm with which a set of data points can be segmented into K different groups, based on a particular characteristic element [8]. For this work, this technique was applied to segment the available plants into smaller groups spread across different regions.

1.4 Applications

Physical assets with micro-controllers can nowadays be monitored and controlled constantly with manual and automated pre-configurations. The number of start-ups within the energy sector is increasing, and some have grown already to a company level. Some of them have developed different software applications which aim to build what is referred to as a digital twin of a physical system. Energy heat networks are also optimized by using measurements and applying fluid dynamics and thermodynamics. Some of the mentioned tools are being developed and sold as digital services and intelligence.

Grid Singularity's approach is an open-source energy Exchange, which provides the infrastructure needed to support real-time LEMs, energy trading, and managing the market strategy for different hierarchical market designs. It includes support for Grid Operators and Aggregators, who are supposed to distribute power through transmission lines and manage the physical assets. The Exchange executes Peer to Peer (P2P) transactions after a bid-offer matching process and summarizes all relevant market information and a community's KPIs.

Recent studies have concluded that hierarchical LEM market structures and strategies increase local trading and local consumption within communities [13]. Nowadays, energy communities are a reality in some cities, and technology is constantly being developed to make them more accessible.

In African countries, microgrids managed by a SEMS have been developed [20], increasing the community's self-consumption significantly, therefore providing better life quality to underdeveloped communities. Energy communities are an excellent example of the idea that teamwork achieves greater things than individuality.

1.5 Objectives

The objectives of this work focus on understanding a microgrid as a complex system, by using software tools and available data and further study it with mathematical tools. The following list describes the particular objectives this work addresses:

- Define and understand the state of a microgrid and the way it behaves over time by creating a simulation of microgrids with local PV generation and ESSs, to characterize it with Key Performance Indicators (KPIs).
- Identify and contribute to the development of scientific-purpose software tools to simulate a microgrid with local generation and energy storage.
- Use public databases of PV production and power Loads time series distributed in different locations to use them as input for the microgrid simulation.
- Fit a mathematical model which represents the correlations between the system's KPIs accurately.

In chapter 2, a description of the developed simulation tool is presented, beginning with the data processing modules in sections 2.1 to 2.3, continuing with the approach followed to simulate a battery, and finalizing with how everything is put together to build the house-simulation and community-simulation models in section 2.4. In this same chapter, the second and third objectives are addressed.

In section 3.1, an analysis of the impact that PV and ESS assets have on a single house Load is portrayed, allowing us to gain qualitative and quantitative insight into the simulated system, with the performance indicators introduced in section 3.1.2. This section sets the ground for the first listed objective.

The system's performance analysis is further studied in section 3.2, this time for 10-house communities in different geographic regions with varying numbers of assets for each configuration. The correlation between the performance indicators is studied in section 3.2.5 through a mathematical model based on the simulation results, which addresses the fourth objective. The general discussion of the results is presented in section 3.3. Chapter 4 presents the work's conclusions, and further work is proposed.

Chapter 2

Methodology

A microgrid can be simulated by grouping houses, for which a data-based model for a home should be created beforehand. The basic information needed to build such a digital object is the electric signals corresponding to the power Load and power generation.

By using Object Oriented Programming (OOP), a class for each kind of asset can be defined (Load, PV and ESS). On top of that, a house class, and on top of them all, the community class containing all of the other objects. Figure 2.1 illustrates the simulated system.

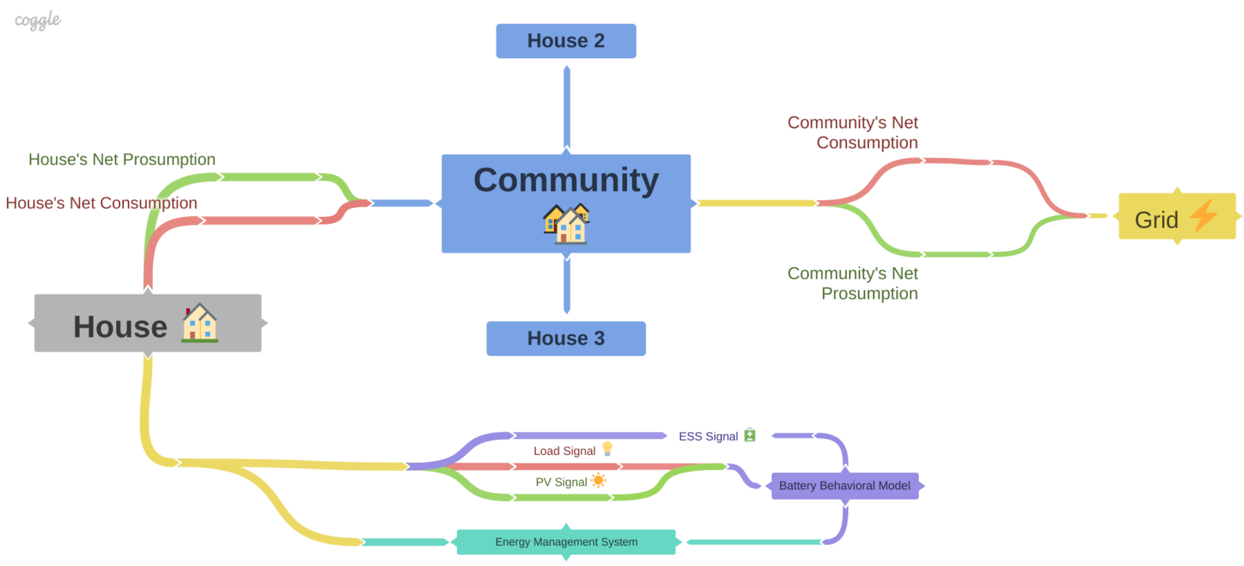


Figure 2.1: Simulated System.

Data has to be supplied to build each house object with an associated set of power signals. Figure 2.2 shows the software's structure for generating the simulation's input signals. These are obtained by using the time series from the databases and processing them into the required pieces of information.

Several parameters can be specified on the software according to the simulation necessities. For example, the PVs' peak generation capacity and the ESS storage capacity. Those parameters are fixed at the moment of building the community class instance.

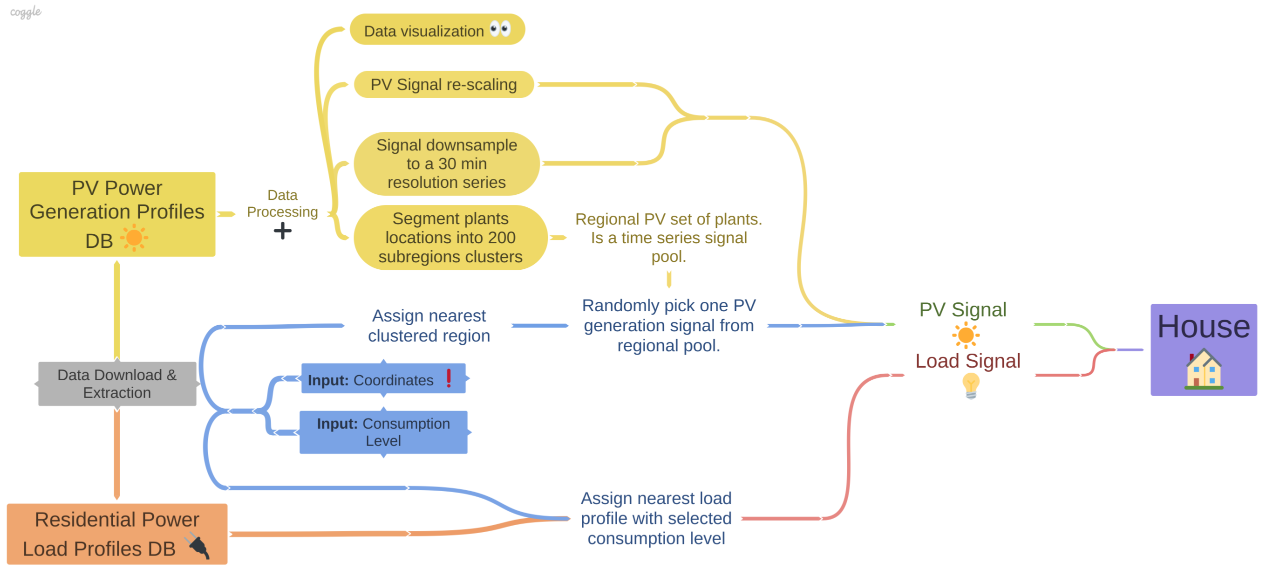


Figure 2.2: House model diagram.

2.1 Data Download and Pre-Processing

Data is directly downloaded from NREL's and OpenEI's websites and stored locally. Downloaded files consist of .zip and .csv files, ordered in a working directory. Each database has a dedicated script to download, extract, and arrange data files. Once data is downloaded and arranged, tables are slightly processed to be able to read from them. Metadata (such as the location and power plant's capacity) is stored in separate tables.

2.2 Data Processing

Before data can be used in the actual simulations it needs to be processed first. The following three sections briefly describe different data processes applied to the signals to further use them.

2.2.1 PV plants regions clustering

All the plants on NREL's database are segmented into K groups and labeled accordingly using scikitlearn's k-means algorithm [8]. This clustering was based on the plants' coordinates, which caused an indirect segmentation of the US territory (see figure 2.7).

The distribution of the clusters depends on the distribution of the PV plants in the database. After the clustering process, the geographic center of these locations is obtained by averaging latitudes and longitudes.

OpenEI's database does not include coordinates for the consumption profiles data points. Instead, it provides the name of the nearest meteorological station to the site. Such names are joined with an auxiliary database that contains the stations' names and coordinates [25]. Finally, one power Load signal and one power generation signal are taken, coupled, and assigned to the specified coordinates. Figure 2.2 shows diagrammatically how the processing occurs within the overall simulation's flow.

2.2.2 PV signal re-scaling

NREL's database includes relatively large power plants with several MW of peak installed power. To obtain a signal corresponding to a residential size PV module, time series are re-scaled according to a polynomial fit fixed to be 0 at the axis intersection (see figure 2.3). With this fit, the power generation time series can be re-scaled by multiplying the signal times a re-scale factor obtained from the curve fit.

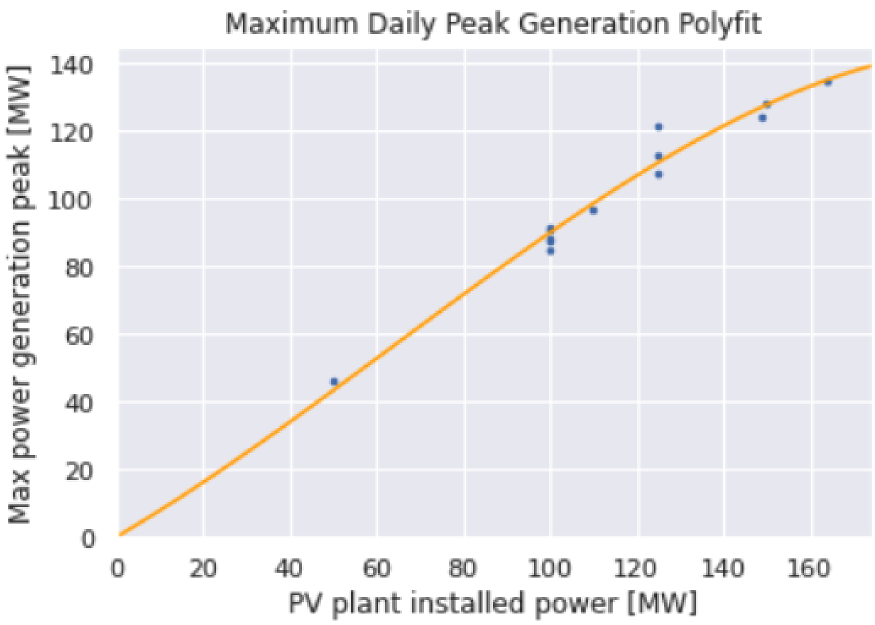


Figure 2.3: 3rd-degree polynomial fit using peak generation values.

Figure 2.4 compares two histograms corresponding to the power generation signal before and after the rescaling. It can be observed that the frequency pattern is the same, but the values on the green histogram are considerably smaller.

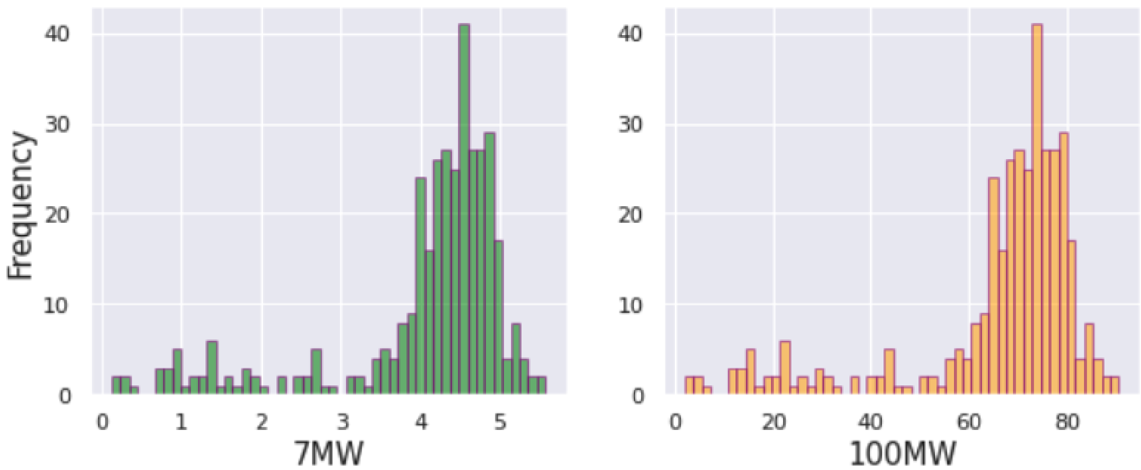


Figure 2.4: Histogram plots of parent and re-scaled child signal.

2.2.3 Load and PV signals coupling

To couple the signals, the values must have the same temporal resolution to be able to operate between the signals. As NREL's data comes with a 5-minute resolution and OpenEI's in an hourly one, the time series are down-sampled and up-sampled, respectively, to be transformed into two 30-minute-resolution compatible signals.

2.3 Data visualization

The following sections present a visual overview of the PV plants database and the clustering procedure described in previous sections. The PV plant distribution determines the regions where a house can be simulated, mainly because of the available data.

2.3.1 Plants & consumers data distribution

Any available signal in the database can be represented with a point on a map. Expressed in those terms, each database provides a different distribution of points according to their available data. The intersection of both distributions represents the regions where a simulation is feasible, considering that there is data for consumption and generation profiles.

Figure 2.5 shows the DPV-type generation plants. The color of the point represents the plant's peak generation capacity. Figure 2.6 shows the UPV-type generation plants.

Note how the UPV subset is slightly more populated than the DPV plants set. The presented simulation and analysis were done using the UPV plants database's subset.

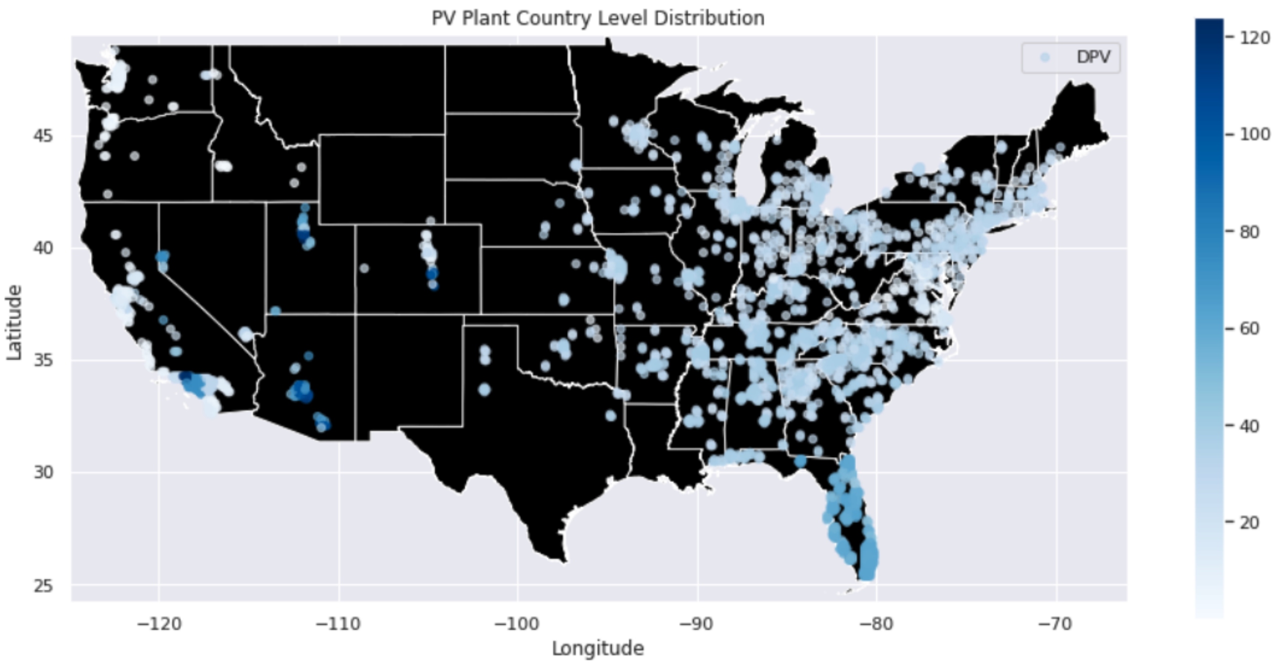


Figure 2.5: DPV PV plants on NREL's database.

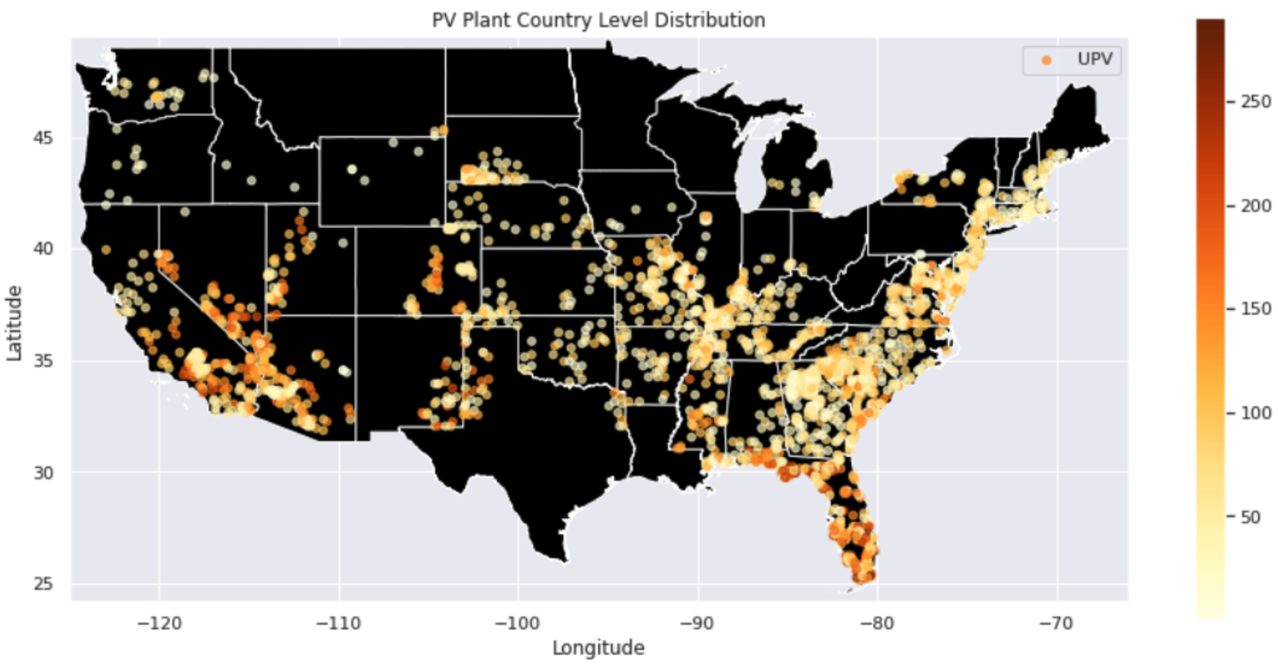


Figure 2.6: UPV PV plants on NREL's database.

2.3.2 Plants clustering

The developed software's data visualization module also permits seeing how the plants are segmented. The clustering aims to group plants from nearby regions, which should have similar weather conditions. Then, when a simulated house is built near a cluster, one of the grouped plants (PV signals) is assigned to that house randomly. Then, if more homes were created in the same location, they would have similar weather-wise signals but would not be identical. In other words, if I have solar generation, I expect my neighbor's generation profile to be similar to mine, but not exactly the same. The clustering addresses this situation.

Figure 2.7 shows the regional segmentation for a total of 200 clusters. That number was chosen so that each plant from a segment is contained within a relatively small-sized region, therefore having a similar amount of radiation. Based on the map visualization, a trial and error criterion was used to decide this number. Each cluster is colored differently from the ones around them. Also, a small black circle is located in the region's geographical center.

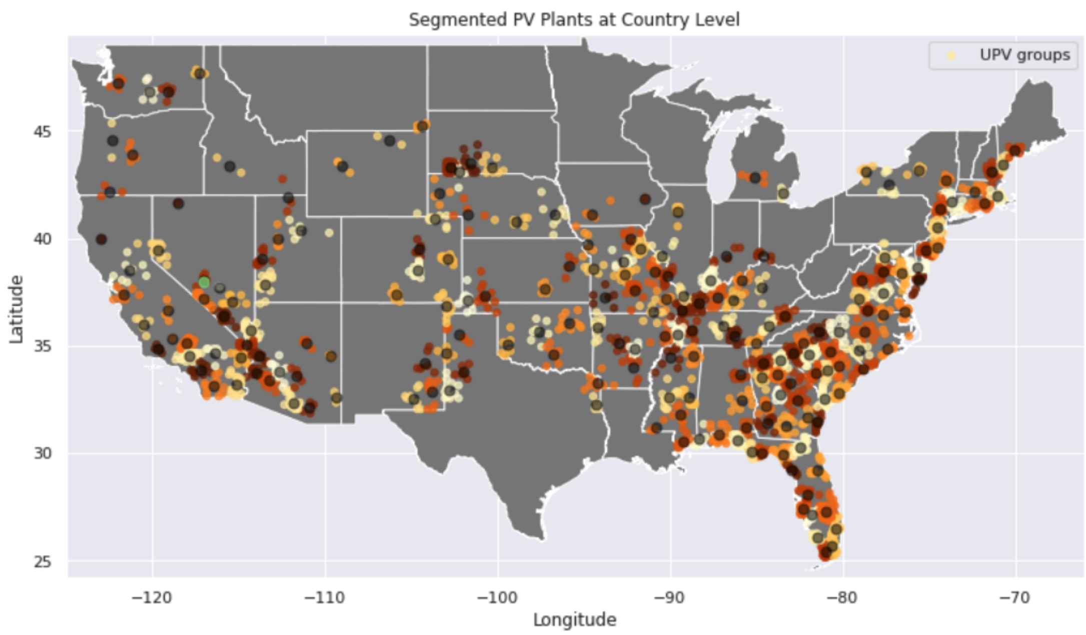


Figure 2.7: UPV type PV plants regional segmentation for $K = 200$ clusters.

2.4 Systems Simulation Modeling

The following sections present the battery charge-discharge strategy (which will be called Battery Behavioral Model)and the energy community models. Afterward, these are described. The community model is illustrated in figure 2.1 Note that the used model has ten houses instead of three, as the figure suggests.

2.4.1 Battery Behavioral Model

A Battery Behavioral Model (BBM) should be understood as a set of rules under which an ESS asset behaves. The presented BBM model controls the power a battery withdraws from or exports to the wider grid at a particular time.

This section is about the description of the BBM and other considerations. The algorithm uses as input a 1-day PV and Load. The model aims to use as much solar energy as possible and reduce the evening peak demand Load.

As seen in figure 2.8, typical Load profile has a maximum during evening hours. PV energy is generated from the beginning of the day until the first hours of the afternoon, peaking around 12 pm. The time during which there is PV generation (sunlight) depends on the latitude in which the producer is located. However, regardless of the latitude, all PV-Load pairs present a very similar pattern.

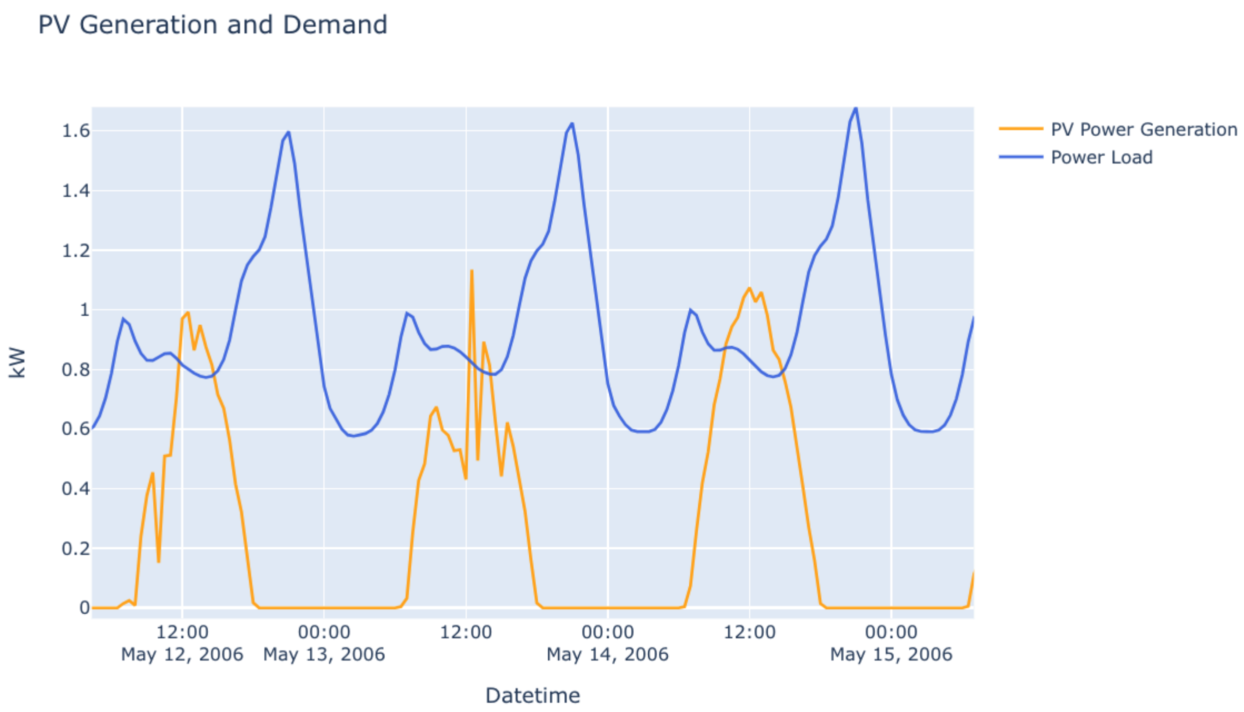


Figure 2.8: 3-day PV-Load coupled signal.

A use case for an energy storage system (ESS) arises when trying to satisfy the evening demand with renewable (solar) resources. As peaking times for both profiles do not match up, if a battery is charged during the generation period and discharged during consumption peaking hours, the stress over the grid system can be significantly reduced, contributing to the whole system resiliency, also called *flexibility*.

The model was constructed by setting the rules to tell the ESS under which conditions to charge and to discharge based on a given pair of Load and PV profiles. The charging and discharging algorithms are described in the following sections. A use case of the developed algorithm was presented in [24].

ESS specifications

The following variables were introduced to characterize the ESS's manufacturing constraints:

- Battery charge state: C.
- Maximum amount of charge supported by the system: C_{max}.
- Minimum amount of charge allowed: C_{min}.
- Battery charge-discharge rate state: B.
- Battery charge rate state: B^+.
- Battery discharge rate state: B^-.
- Maximum charge-discharge rate: B_{max}.

The charge state (C) is related to the charge-discharge rate (B) according to equation 2.1

$$C_i = C_{i-1} + rB_{i-1}, \tag{2.1}$$

where r is the resolution of the time series expressed in hours. In this case $r = 0.5$.

Determination of when to charge

To determine when the battery has to charge and discharge is equivalent to specifying $B(t_i)$ for $i = 0, 1, ..., N$, where N is the number of periods in which the day is split.
In order to achieve this, the following rules were codified:

Def: Function which estimates the time window in which there is sunlight.

Being t_i the ith time slot of a given day and G_i the PV power generated at t_i time of the day, let

$$f(G) \to (t_a, t_b) \tag{2.2}$$

be a function of the PV profile that returns an estimate of the time interval at which there is sunlight (t_a, t_b); where $t_a < t_b$ and $G_{a-3} = G_{a-2} = G_{a-1} = G_{b+1} = G_{b+2} = G_{b+3} = 0$ and $G_a, G_b > 0$. If no such points exist, t_a defaults to 07:30:00 am, and t_b to 17:30:00 pm.

The base charging strategy is to base $B^+(t)$ on $G(t)$'s behavior, as the battery is expected to charge faster if more PV energy is present. However, it is not entirely possible to set $B^+(t) == G(t)$, because of the restriction B_{max} on the battery. As the model was initially intended to work with forecasted profiles, the strategy was set to consider potential deviations from actual values. Therefore it would not be an option even for $G(t) < B_{max} \forall t$.

The way by which the possible deviations are managed is by setting $B^+(t) == \eta G(t)$, where η is defined as follows:

The Riemann sum for $G(t)$ is calculated up for a time $t_a < t_x < t_b$, obtaining the total energy produced during the time interval (t_a, t_x). Assuming the time is equally partitioned, it can be considered that $t_i - t_{i-1} = r, \forall i$, where r is the resolution of the time series, thus having

$$I(t_x) = \Sigma_{t_i=t_a}^{t_i=t_x} rG(t_i).$$

Then, considering $I \neq 0$, η is defined as

$$\eta(t_x) = \begin{cases} \frac{C_{max}}{I(t_x)} & , \quad C_{max} \leq I \\ \\ 0.95 & , \quad C_{max} > I. \end{cases} \tag{2.3}$$

This would ensure that if $B^+(t) = \eta(t_x)G(t)$, the battery will be fully charged by time t_x if $B^+(t) < B_{max}$ for $t < t_x$. However, if there is a time for which $\eta(t_x)G(t) > B_{max}$, then $B^+(t) = B_{max}$, and in these cases, the battery would be slightly under full capacity by the time t_x. A potential improvement of the model would be to define a modulating function $m(t)$ to ensure that $G(t) > B^+(t) = m(t)\eta(t_x)G(t) > B_{max}$ always holds, and the battery reaches full capacity at time t_x. However this is let for further development.

A graph of a sample $\eta(t_x)G(t)$ profile, for a certain day, is shown in figure 2.9.

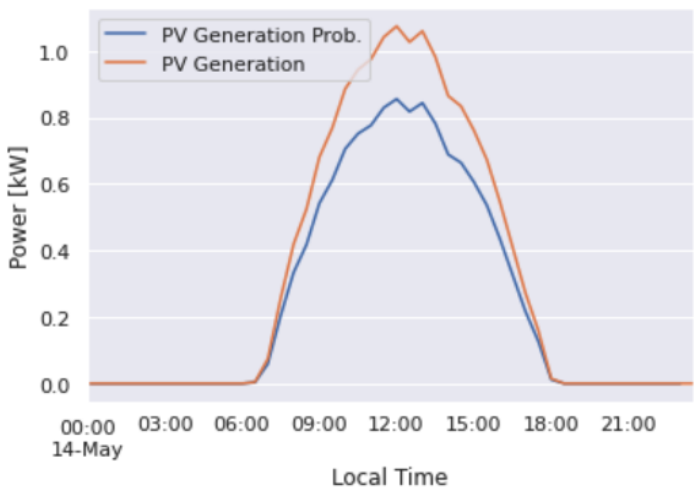


Figure 2.9: Orange: $G(t)$, Blue: $\eta(t_x)G(t)$.

Determination of when to discharge

At this stage, the rules previously described help simulate how the battery gets fully charged. Next, the rules defining how this charge is used are presented.

The developed strategy aims for evening-peak-Load reduction. The more the peak is reduced, the better the itinerary. $B^{-}(t)$ is based on the Load profile to flatten out the peak as much as possible. This BBM can be used for purposes other than microgrid simulations. For instance, the algorithm can calculate a future schedule if forecast consumption signals are provided.

An offset A is calculated in a way that, for a certain time interval (t_c, t_d), the following holds:

$$I(t_x) = \Sigma_{t_i=t_c}^{t_i=t_d} r(L(t_i) - A) = C_{c-1} \approx C_{max}, \tag{2.4}$$

is the Load profile function. The above condition considers C_{c-1}, which is the latest charge state value right before the beginning of the discharging phase.

The latter being considered,

$$B^{-}(t, t_c, t_d) = \begin{cases} A(t_c, t_d) - L(t) & , \quad 0 \leq L(t) - A(t_c, t_d) \leq B_{max} \\ B_{max} & , \quad L(t) - A(t_c, t_d) \geq B_{max} \\ 0 & , \quad L(t) - A(t_c, t_d) \leq 0 \end{cases} \tag{2.5}$$

is used as the discharge rate function. To find an appropriate (t_c, t_d) pair, the two following procedures are followed:

Firstly, the *average value* for $L(t)$ is calculated for the second half of the day (time interval $(t_{N/2}, t_N)$). Then, for each time step in the interval, the deviation from the average value is calculated, and if smaller than a certain threshold, the time slot is stored in a list.

Then, t_m is calculated as the midpoint between the earliest stored time and the oldest one. The time point is rounded up according to the time step (resolution) of the time series. Finally, t_c and t_d are calculated as $t_c = t_m - 3h$ and $t_c = t_m + 3h$.

Figure 2.10 illustrates the whole charging strategy model for an entire day. The PV and Load profiles are plotted together with the battery's charge state and charge/discharge rate functions. Also, the offset which satisfies equation 2.4 is represented by the black curve.

Note how the charge rate curve is always under the PV profile curve. Also, the charge state reaches a maximum at a particular hour, then discharges completely. And finally, the discharging phase of the charge/discharge profile curve corresponds with the region above the offset of the Load profile.

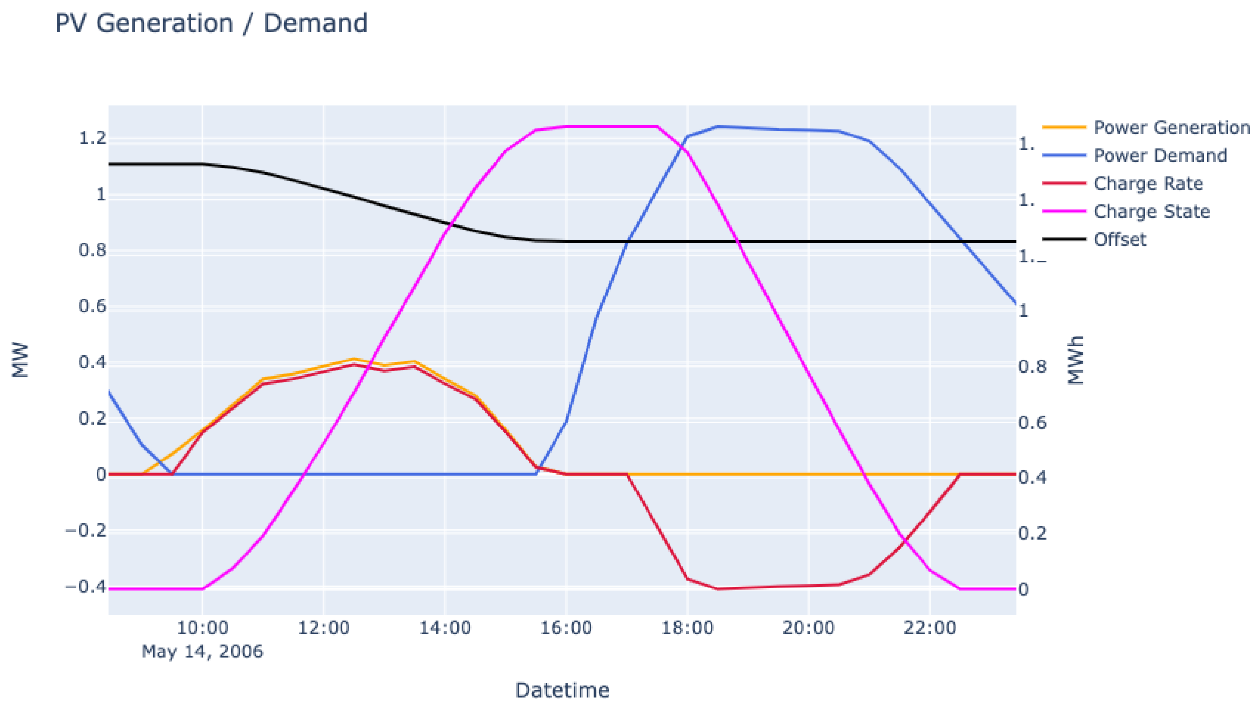


Figure 2.10: Key curves describing the battery charge-discharge strategy/model.

2.4.2 Microgrid/Energy Community Model

The first of the following sections describes a hypothetical situation to explain how a microgrid on a typical day could look. On it, the existence of different levels at which a microgrid can be established is emphasized. This level gradation is equivalent to the hierarchical LEMs described by Grid Singularity [17]. The second section mentions some additional considerations for the energy community simulation.

The model

With the latter in mind, the developed energy community model consists of N buildings (which have a residential Load profile), from which $m \leq N$ can have PV generation. Technically, it would be possible to have a situation where a building had storage capacity without necessarily being a producer. However, it is common to see ESS in facilities that also use some kind of energy production system. Considering this, there is also the possibility of having a total of p ESS distributed in some of the N buildings, where $p \leq m$.

Each of the created buildings is coupled to its holding community, as illustrated in figure 2.1. Buildings are built so that one-third of the houses are low-level consumers, another third are base-level consumers, and the last third are high-level consumers.

The specifications for the installed assets are outlined in table 2.1. These are the same across the different consumption levels buildings to simulate a situation where a fixed-sized system is to be evaluated in different geographic regions.

Load Level	ESS Specs		PV Specs
	C_{max} kWh	B_{max} kW	Installed Peak Load kW
LOW	7.0	3.5	7.0
BASE	7.0	3.5	7.0
HIGH	7.0	3.5	7.0

Table 2.1: PV & ESS sizing based on Load level.

The chosen capacities were based on available commercial PV and ESS assets. The analysis introduced in the next chapter does not explore the scenario in which assets are owned by the community.

2.5 Data and Models' Limitations

In this section, limitations corresponding to data are briefly mentioned. Also, more comments on the constraints of the models introduced in this section (BBM, Energy Community) are presented.

Data

The first limitation the work faces is concerning data availability. A known issue is that more real data sets need to be made available or created. Many open data initiatives create or provide valuable data resources to further develop these technologies [6, 19]. Countries that lead these efforts are located in the European Union [10] and the United States [15]. Other initiatives are focused on the African power systems [3], or other under-development countries.

The data used for energy generation was created with weather measurements and radiance levels for 2006. Meanwhile, the household Load profiles have been modeled with different average weather data and socioeconomic features. Load and PV time series are commonly correlated because both patterns are weather dependent. As the data comes from different sources, it is improbable that the consumption patterns and the generation have this natural correlation for short time intervals. This sets the simulation slightly away from a more realistic situation. Although the data sources are different, at least a seasonal correlation should be present to some degree because the data is representative of the same place.

Also, both data sets contain different data points in different locations. The data processing explained in chapter 2 aims to match Loads and generation curves as well as possible (geographic-wise). Yet, it is far from perfect.

Battery Behavioral Model

A whole branch is open concerning the management of the ESSs assets. The developed Battery Behavioral Model aims to use exactly one full charge per day. The battery use could be maximized by adjusting the model to allow the battery to charge and discharge more times a day.

Energy Community Model

The implemented model is very restrictive in forcing a house to possess its assets (house's assets) and allowing at most one of each kind. Also, the possibility for the community as a whole to own assets is not supported. If this is implemented, more realistic and variate CAS scenarios could be within the scope. The system under study is complex and has many variables that can all have different impacts. All the elements someone could think of can be appended to the presented models and would require a considerable amount of analysis.

Chapter 3

Results

Results are divided into two main parts: an analysis of a single *1-house* community and its assets in a particular geographic region. Secondly, a study of 200 communities across the United States is presented. Each simulation considers a Community Assets System (CAS) with different characteristics in terms of the amounts of each asset type.

The first approach focuses on a community's Net Load yearly profile, as well as the statistic distributions which reflect the different states (with respect to the grid) in which the community has been (e.g., average power Load(s) during the whole year). Also, a comparison between PV, PV & ESS, and no assets configurations is presented. This analysis pretends to describe the community's annual behavior. It compares the state in which the community is with a particular asset configuration and another by varying N_{PV} and N_{ESS}.

In the second approach, 200 10-household communities are created in random locations where data exists on the database for all combinations of PVs and ESSs assets. Then, the correlation between relevant KPIs (calculated during the simulation) and the Communities' Energy Requirement (E) is explored. A mathematical model which describes how the system reacts to the presence of a particular Community Assets System (CAS) configuration is proposed.

The mentioned model can be used to gain insight into how the CAS performs with a community of a particular energy requirement and a particular combination of PVs and ESSs. Different KPIs provide information that reflects the annual behavior of the communities and can be used to find an optimal CAS for a particular objective (e.g., maximize the saved energy volume, maximize saved/earned money, maximize self-sufficiency, and so on). The energy markets component can be integrated into the model by implementing some grid fee mechanisms and strategies, as Grid Singularity proposes [17].

The same methodology followed in this second approach can be replicated with a more sophisticated and realistic simulation engine, such as the Grid Singularity Exchange [17], to observe how the model would change under particular trading conditions or preferences, and quantify the impact such conditions would generate.

A geographic exploration is made to see how the communities would hypothetically perform when distributed in the country. These visualizations are combined with annual net consumption and self-consumption data. In this part of the analysis, it will be noted that there are geographic regions with particular average energy requirements.

3.1 Individual Energy Community Statistical Analysis

In this section, a comparison between different simulation scenarios is addressed. The simulation is made for a single community, which base case considers a single Load. Other scenarios explore both the PV and PV + ESS cases. These three cases are compared with one another. Their yearly statistics are evaluated to identify each asset's impact on the house.

Although it could be argued that a single house does not form itself a community, the term *community* appears on several occasions during this section to refer to such a scenario. If we think of the 1-house community as a *zero level community* or something of the sort, the use of the term might be justified and, therefore, would not be incorrect. It also could be thought of as if a local energy market is inside the house for assets to trade in.

The following analysis is made to gain insight into how a 1-Load system (1 house) behaves in the cases in which PV and ESS assets are also part of the system. For each of these cases, different situations could occur. For instance, energy imports from the wider grid or energy exports due to an impossibility to store the generated energy surplus and, therefore, the impossibility to consume it. These cases will be discussed with sample one-day signals.

The following analysis was done for a simulated community on (35.23, -80.70) in Charlotte, North Carolina. The location for this community has been arbitrarily selected from the regions present on the database and is only illustrative. The following analysis can be replicated for different locations as the results will likely vary slightly.

3.1.1 Base cases description and their relation with the grid

The base cases are the ones on which an asset is introduced at a house level. In the following section, the *Load* case refers to the case in which a house is only a consumer (i.e., has no ESS nor PV generation). The *Load + PV* case represents how the system would behave being a prosumer without an ESS. And so will *Load + PV + ESS* case be the situation with PV generation, storage capacity, and consumption.

The following sections are a comparative analysis based on each case's statistical behavior. Also, a visual inspection of the comparison is presented. The graphics show how each asset affects the house's behavior at the community level.

Case 1 (base case): 1 Load

The base and most simple configuration of such a system consists of a single Load, representing a regular grid user who consumes energy based on a specific behavior pattern. Usually, the consumption pattern shows a progressive increase of the Load, reaching its peak during the afternoon. These patterns may vary for other weather or social conditions (such as seasonal temperature changes or the type of person living in the building, among other variables).

It has also been observed that there are two peaking moments for some consumption habits: one in the morning and the other in the evening. Figure 3.1 shows the Load profile for a BASE consumption level house in the mentioned region.

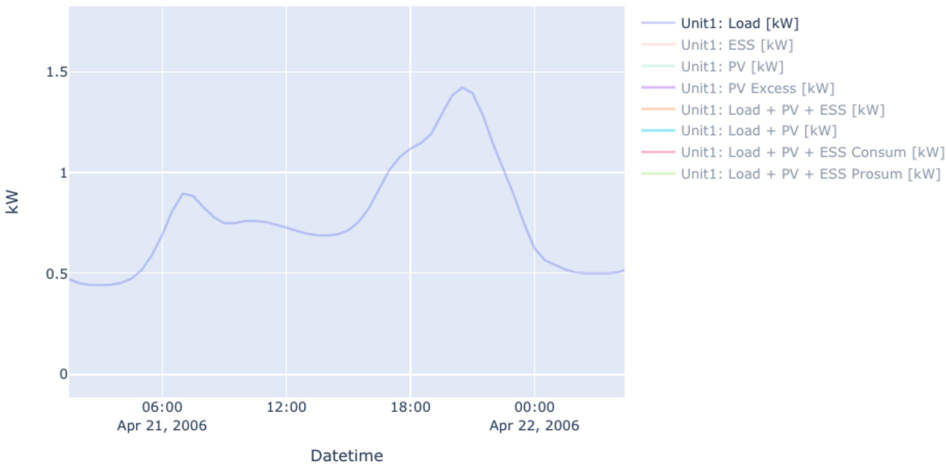


Figure 3.1: Load profile for a single house near Charlotte, North Carolina.

This particular day illustrates the two-peak condition described earlier. However, the evening's peak tends to be greater for most of the year, so it represents a more significant opportunity to increase the grid's flexibility by *shaving* it. For this case, the *1-house-community*'s Net Load is the same as the Load asset profile.

Case 2: 1 Load + 1 PV

This system configuration represents a community participant who can generate energy through a photovoltaic installation. This energy is consumed at the moment it is generated. If more power is produced than consumed, an energy surplus exists, which is then exported to the grid (because there is no other house or asset to share the energy with).

Figure 3.2 shows the base Load and PV profiles and resulting signal or Net Load/demand profile for a sample house in North Carolina. Positive values represent energy consumption, and negative values represent energy production. Note that when the energy is generated, it is very likely (depending on the solar radiation and peak installed power) that there will be an energy surplus. Also, note how the peak generation and peak Load do not happen at the same time interval. This prevents the system from having better use of the generated energy, provoking a lower self-consumption rate and self-sufficiency.

Case 3: 1 Load + 1 PV + 1 ESS

In this third case, an ESS is introduced to address the peaks mismatch problem. The algorithm aims to shift the energy from when it is generated to be of use when it is most needed.

Figure 3.3 shows this site's base Load, PV, and ESS profiles. As previously exposed, the BBM focuses on reducing the evening Load peak.

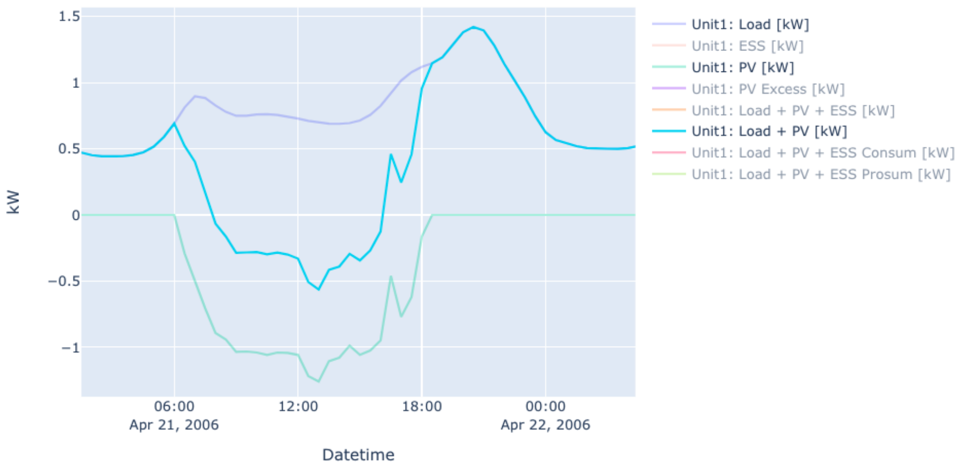


Figure 3.2: Load and PV profiles for a single house near Charlotte, North Carolina.

In the figure, the interactions between the different profiles can be observed: the battery charges during PV generation and discharges during the evening Load peak, the PV generation charges up the battery, and the PV excess reduces the overall demand accordingly when the PV energy is produced.

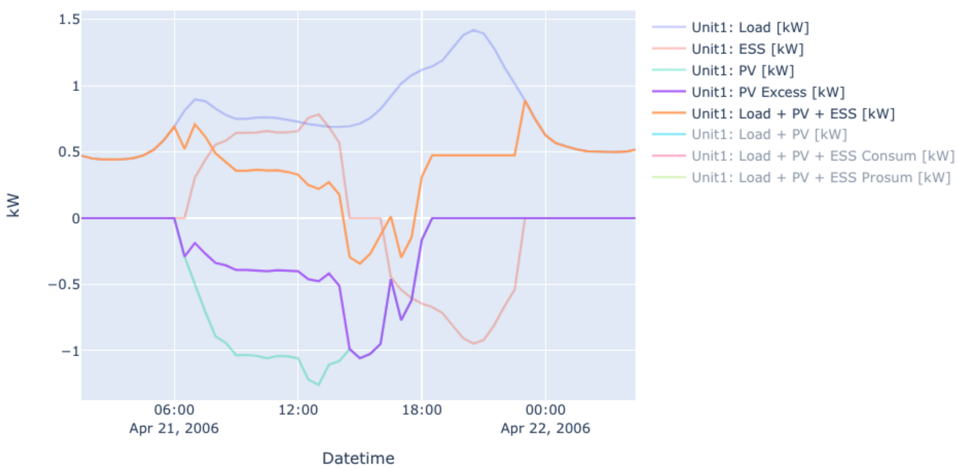


Figure 3.3: Load, PV, and ESS profiles for a single house near Charlotte, North Carolina.

It can also be observed from the sample day in figure 3.3 that if the battery would have charged at a slower rate during sunny hours, the surplus peak around 15:00 wouldn't exist. This would imply a higher self-consumption and self-sufficiency. This is a clear example that optimizing the BBMs is a crucial factor that directly impacts the performance of a CAS. The development of more refined BBMs is the subject of further work.

Note that figure 3.3 also shows the net profile of the house (orange line) by summing up all of the different assets' profiles on the system. Based on this behavior, it can be said that the Net Load of the system is a linear combination of all of the assets' profiles, which would also apply to larger communities and not only to this *1-house-community* case.

In general, the orange curve can be mathematically described as:

$$\text{NetLoad}(t) = \sum_{i=0}^{N_{Loads}} \text{Load}_i(t) + \sum_{j=0}^{N_{PVs}} \text{PV}_j(t) + \sum_{k=0}^{N_{ESSs}} \text{ESS}_k(t), \tag{3.1}$$

where N_{Loads}, N_{PV}, N_{ESS} are the total number of each Load, PV, and ESS assets, respectively, and $\text{Load}_i(t)$, $\text{PV}_j(t)$, and $\text{ESS}_k(t)$ are the particular Load, PV, and ESS profiles for the i-th, j-th, and k-th assets of each asset type.

Note that the $\text{NetLoad}(t)$ might also be negative for some cases. This is due (as shown in figure 3.3 [purple line]) to an excess of PV energy that can't be stored due to a lack of space on the ESS and a lack of Load demand. As there are no other units in the community, there is no further opportunity to share this energy surplus with another house or asset other than the grid. If such a case happened, the generated energy could be consumed in a higher percentage by the community depending on the number of Loads, amount of storage capacity, and PV generation capacity installed by the community members.

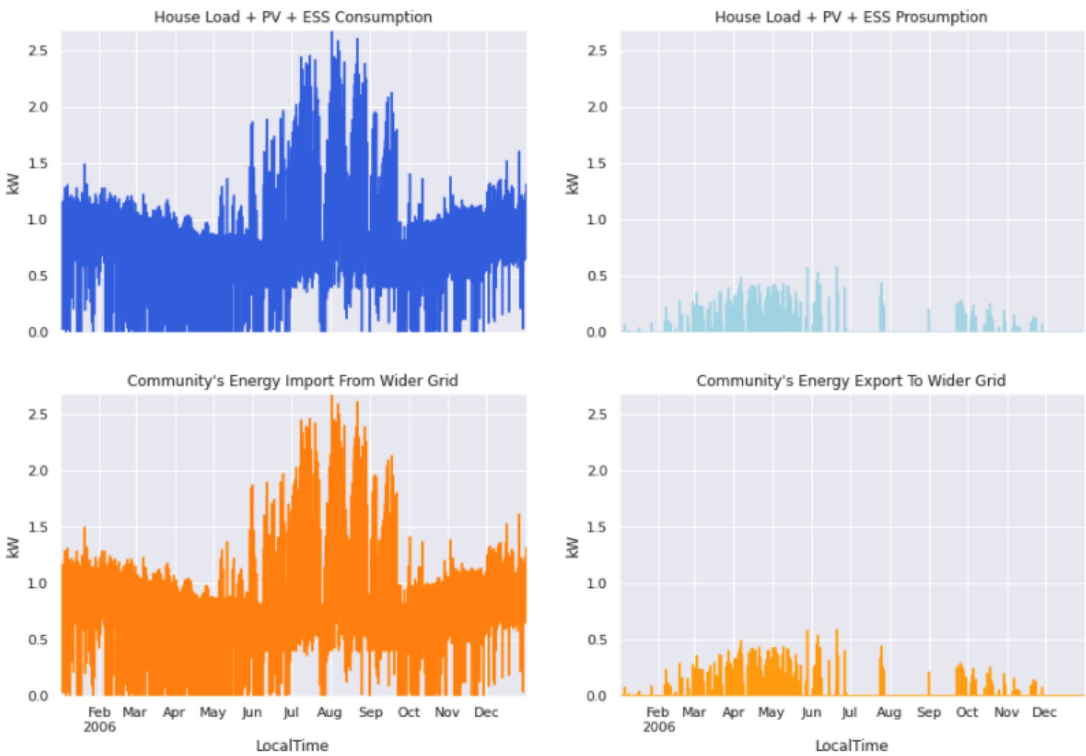


Figure 3.4: Grid Export/Import Interpretation Equivalency for the Community's Net Load Profile.

As it will be further explained, these parameters also influence a particular community's *self-consumption* and *self-sufficiency*. The following section proposes a model to mathematically describe how different variables (E, N_{PVs}, N_{ESSs}) affect the KPIs.

For the sake of completeness of this section, figure 3.4 depicts how, for a whole year, the negative values of the Net Load profile can be interpreted as the power that the community exports to the grid, as well as the positive ones, are considered as the power that the community needs to import. These are also named the community's *consumption* and *prosumption* profiles.

Figure 3.4 shows a correspondence between the house's net profile and the grid's profile. For larger communities, the correspondence holds, but not with a single house net profile, but with the community's net profile, or in other words, the sum of all the Load, PV, and ESS assets' profiles on the community (see equation 3.1).

3.1.2 KPIs description: a tool for cases and scenarios comparison

In summary, three main cases have been discussed so far. Each of which, for this particular *1-house-community* scenario, has a specific Net Load yearly profile. Figure 3.5 shows the Net Load for the *1-house-community* for each discussed case.

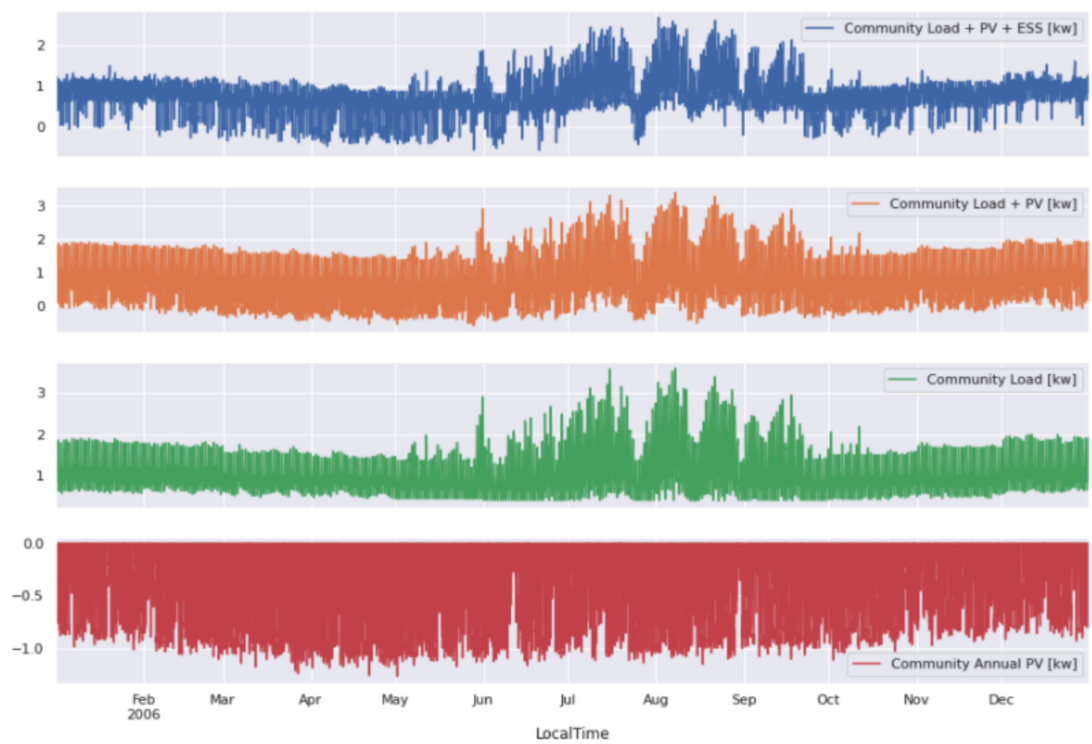


Figure 3.5: Annual PV Generation and Net Load Profile for the Load, Load + PV, and Load + PV + ESS Cases.

On these graphs, seasonal behavior can be observed. Particularly summer months tend to have a higher Load demand, most likely due to HVAC systems to cool the facilities. Different locations can produce different patterns for the community's Net Load due to a dependency on the community's consumption habits during the day (or season) and the number of electric assets inside the facility (Load level / socioeconomic). To analyze the whole year quantitatively, histograms for each time series are shown in figure 3.6.

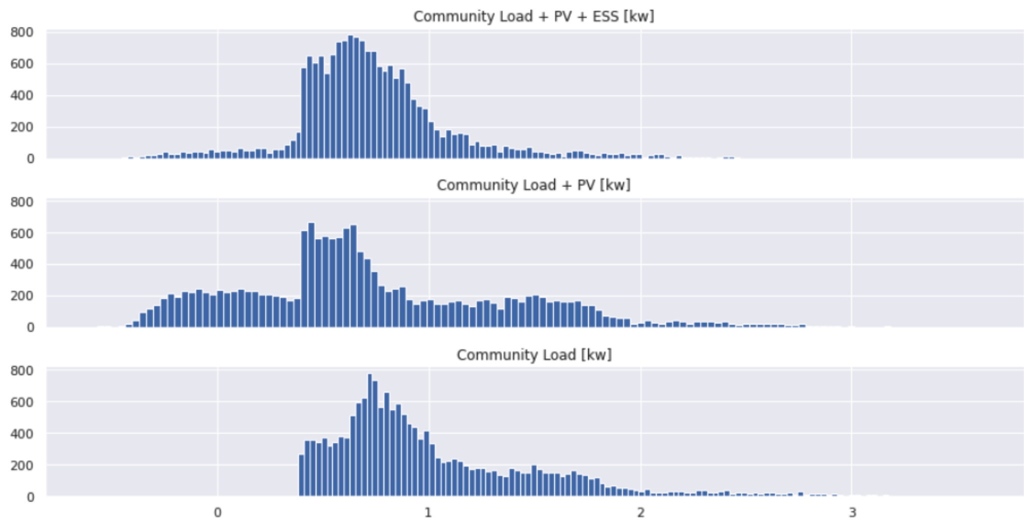


Figure 3.6: Community's Net Load annual histogram for studied cases. For the *Load* case, only positive values for the Net Load are present, representing that energy is being imported from the grid the whole year. Furthermore, there is a minimum (positive) value, which means that the house represented by the Load profile will always require at least that amount of power. For the *Load + PV* case, there are several occurrences in which the community (house) is also exporting energy (negative values). These represent the times of the day when there is a surplus of PV energy. Also, the distribution is wider as there are times when power is exported, and also high Load peaks still occur. However, for the *Load + PV + ESS* case, it can be observed that the number of times at which the house is exporting energy is significantly reduced, implying that the PV energy use rate is improved due to the ESS system (this idea is further discussed in the following section). Also, the times the house has a semi-high Net Load are reduced. The distribution is more concentrated around the annual average value, representing greater stability on the community's Net Load, as it can be located more precisely. In other words, the community's average demand would be lower, and at the same time, the actual Net Load will more likely be near the average value.

For a more quantitative comparison of the three cases, some key performance indicators (KPIs) were used to measure the community's behavior:

- **Average Power Consumption [kW]:** Annual average power that the community takes from the grid.
- **Average Power Prosumption [kW]:** Annual average power that the community exports to the grid.
- **Energy Requirement [kWh]:** Total energy usage within a community.
- **Energy Consumption [kWh]:** Annual energy imported from the grid.
- **Energy Prosumption [kWh]:** Annual energy exported to the grid.
- **PV Energy Production [kWh]:** Annual PV energy generated.
- **Self Consumption Rate:** Total Self Consumed Energy compared with the total Energy Requirement.

- **PV Self Consumption Rate:** Self Consumption Rate shares attributable to the PV Asset(s).
- **ESS Self Consumption Rate:** Self Consumption Rate shares attributable to the ESS Asset(s).
- **Self Sufficiency:** Represents the percentage of the total consumed energy provided by the community and measures how much it relies on the power supplied by the grid.
- **Optimal Self Sufficiency:** Represents the self-sufficiency the system would have if all of the produced energy were consumed (100% self-consumption rate).

Table 3.1 contains each of the mentioned KPIs for the 1 house scenario, as well as total annualized values from which they're calculated. The numbers in the table confirm how adding a single 3.5 kWp PV asset causes the average power consumption to decrease and also makes it more likely to be around the average value (see *Stdev Power Consumption [kW]* KPI in table). They also show that having an ESS causes a decrease in the total *energy-prosumption*, which translates to a greater *self-consumption-rate*.

	Load	**Load + PV**	**Load + PV + ESS**
Average Power Consumption [kW]	2.22	1.84	1.61
Average Power Prosumption [kW]	0	-0.46	-0.21
Stdev Power Consumption [kW]	1.08	1.2	0.89
Stdev Power Prosumption [kW]	0	0.32	0.17
Annual Energy Consumption [kWh]	19,424.89	14,106.76	13,665.84
Annual Energy Prosumption [kWh]	0	494.87	53.95
PV Energy Production [kWh]	0	5,813.01	5,813.01
Self-consumption Rate [%]	0	91.49	99.07
Self-sufficiency [%]	0	27.38	29.65
Optimal Self-sufficiency [%]	0	29.93	29.93

Table 3.1: KPIs for a 1-house community located in Charlotte, North Carolina.

It can also be said that for this particular location, self-sufficiency is almost equal to the optimal value, considering the total energy generated by the house. Therefore, this specific house produces nearly a third of the energy it needs. This does not necessarily hold for another home in this location because the KPIs reported on the table also depend on the consumption habits. Therefore, a different Load pattern would produce slightly different KPIs. The scenario discussed represents a community with a one-third *Production to Consumption* ratio (PtC). This represents, in that sense, a community without enough energy resources to be independent of the grid (self-sufficiency < 100%). The situation where self-sufficiency = 100%, if possible, is only viable if $energyproduction >= energyrequirement$. Some researchers have called this ratio between consumption and production the *Production to Consumption ratio* (PtC rate) [13].

A new question now comes into the discussion: *How different would these KPIs be for an analogous community in a different location?*. The first part of the answer to this question will be discussed in section 3.2.

3.1.3 Influence of PV and ESS assets on a house (community) net load: a Community Asset System Analysis

Apart from considering the numeric values of each Net Load profile distribution, visually exploring how the ESSs and the PVs (Asset System) influence the community's Net Load through time, the following figures 3.7, 3.8 and 3.9 compare by pairs the Net Load for all three cases. The point's color represents the hour of the day. The mentioned figures relate to figure 3.6 because the latter could be considered the distribution of the projection of the points on the axis.

Both figures 3.7 and 3.8 visually inspect how the introduction of a PV or an ESS asset affects the system compared to the system without it. The points that form the identity line represent the moments in the simulation at which there was no difference between the cases on the vertical axis and those on the horizontal component of the graph. The highest end of the line represents the moments at which the Net Load is maximum for both cases. The points below the line are the moments at which the case represented by the vertical axis had a lower net demand than the horizontal one. Finally, the points above the line are the moments at which the vertical case had a higher net demand than the horizontal one.

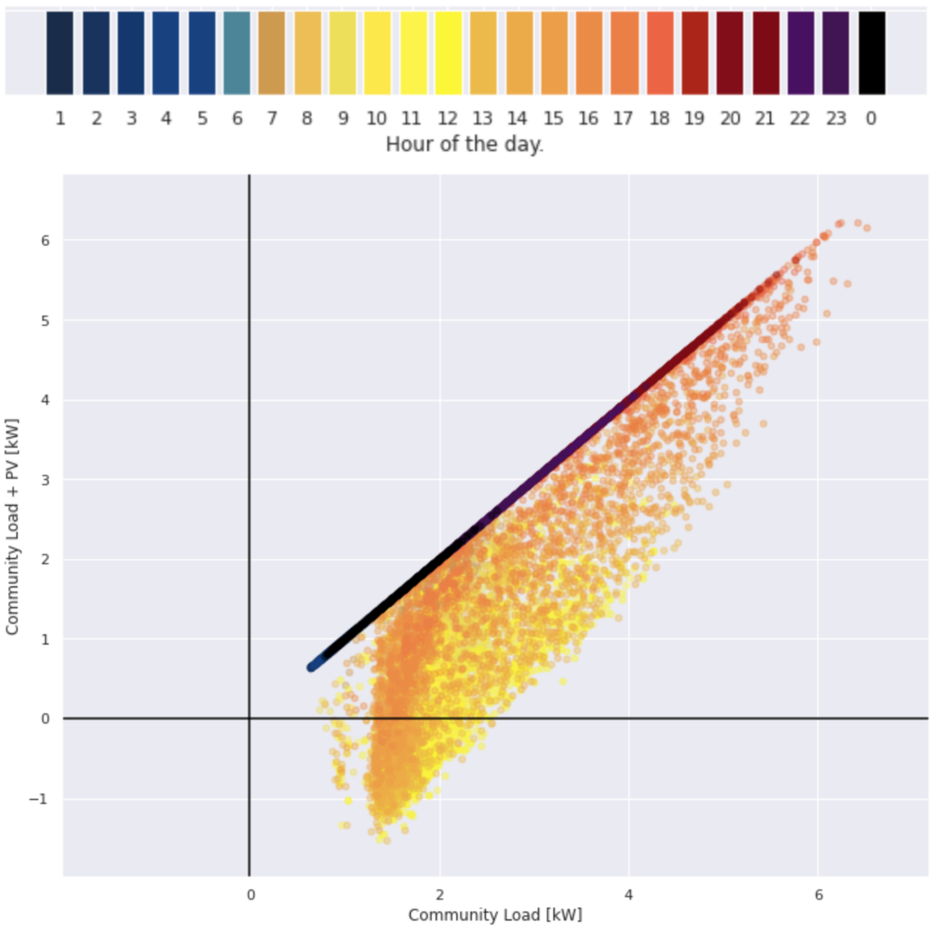


Figure 3.7: Load + PV vs. Load case comparison.

Figure 3.7 shows that for sunny hours, the Net Load for the Load + PV case is lower than the case without the PV asset. In the case for which an ESS is introduced, figure 3.8 shows that the Load is relatively higher for sunny hours, as a portion of the solar energy charges the battery. However, in this case, it can also be observed that the Net Load is lower for the evening hours than for the case without the ESS system. Note how some points in the high end of the line were *drawn down*, representing that there are lower peaks (at peak hours) for the vertical case.

Also, note how although many points (darker points) were also drawn below the line, these do not correspond to the moments at which the grid had the maximum demand (upper right end of the line), implying that the BBM could be adapted (at least for this particular location) to improve it's performance and peak reduction capability.

Further research could explore how these patterns look under the presence of one or more features such as Grid Fees according to a Local Energy Market (LEM) model, which is also supported by the Grid Singularity Exchange [17], or even *peer to peer* (P2P) trading conditions [11].

The approach followed in this work could be further used as a base methodology to evaluate the Community Asset System (CAS) from the perspective of a whole community and to predict, based on data, the performance they could have elsewhere. A similar evaluation system could be developed to characterize the P2P trading evolution over time under certain restrictions (A Community Trading Profile Analysis). Some primary restrictions could be the different fees that may apply to the traded energy, battery system specifications, or geographical location of the community (as will be next outlined for the Community Asset System). Some studies also consider a closeness rate indicator to measure how far away are the trading partners from one another.

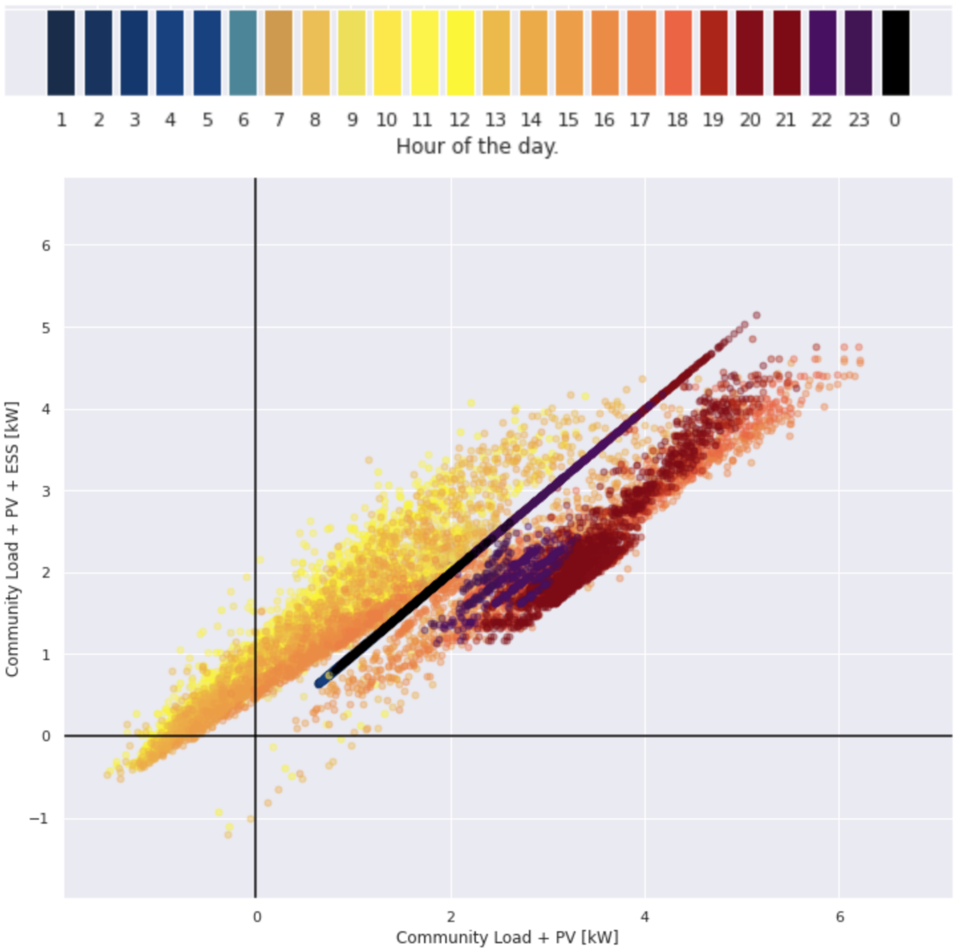


Figure 3.8: Load + PV vs. Load + PV + ESS case comparison.

In figure 3.9, the overall impact of having both PV and ESS assets can be observed. Overall, during the sunny hours and battery discharging period, the system presents a lower net demand and smaller Load peaks.

This particular visual exploration can be further used to intuitively understand the impact a specific PV or ESS asset(s) would have on a community. We'd expect a similar pattern for each of them. However, they would differentiate from one another depending on the community's consumption habits, the total solar radiation at the site, the size of the PV modules, and total storage capacity.

As mentioned before, changing other variables, such as the ESS specs would most likely change the point distributions. In this work, however, all the variables are fixed except the location.

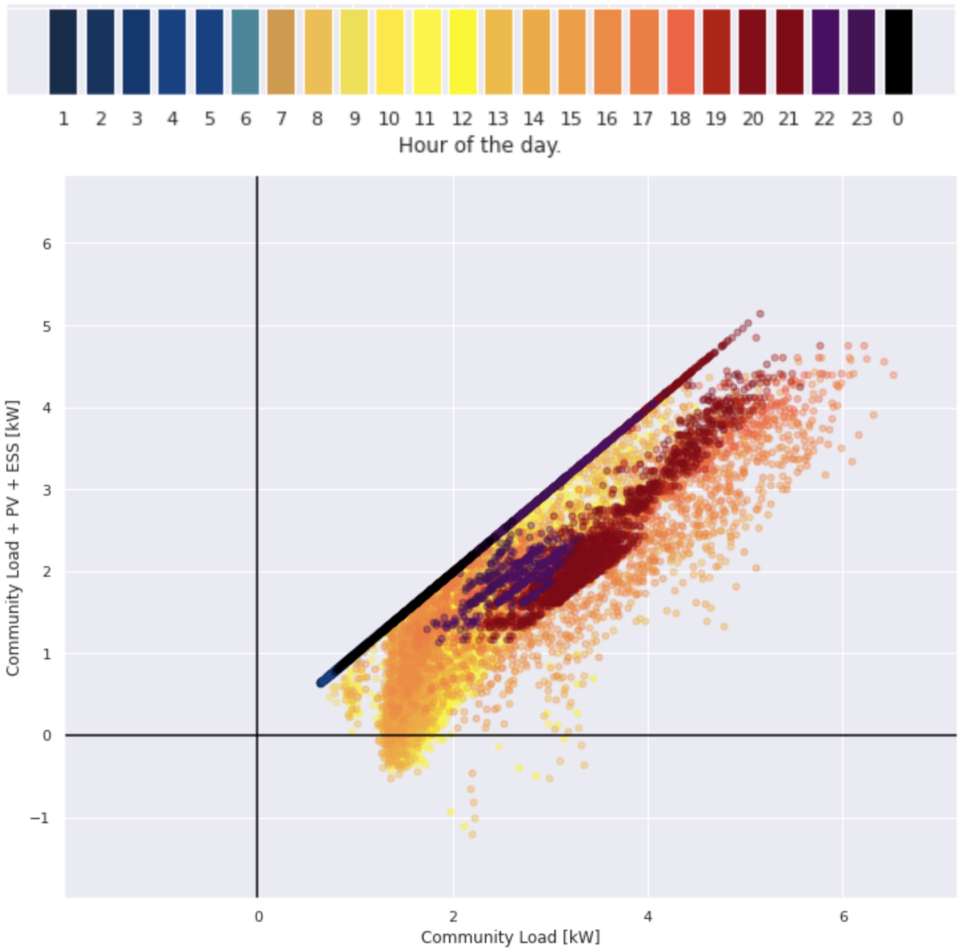


Figure 3.9: Load vs. Load + PV + ESS case comparison.

Another additional element to be considered is the BBM. Different charge-discharge strategies could lead to other point distributions. Further work could be dedicated to looking for optimal charging strategies (or in-house-level trading plans).

3.1.4 Replicating the 1-house community model in different regions

For this part of the study, various locations were chosen based on the climatic areas described in the consumption profiles database (see chapter 2). For each climate zone, different building energy models were used by the database's creators [14], according to climatic regions mentioned in [12].

The same *1 Load + 1 PV + 1 ESS* scenario was reproduced in (27.93, -82.39) Tampa, Florida, and in (37.81, -122.16) San Francisco, California. With identical ESS and PV characteristics (3.5 kWp and 7kWh capacity). Tables 3.2 and 3.3 contain the KPIs calculated for each community.

Their corresponding distributions can be seen in figures A.1 and A.2. The impact of the assets on the communities' Net Load distribution can also be noted by observing the decrease of both the average consumed (and prosumed) power and the standard deviation.

	Load	**Load + PV**	**Load + PV + ESS**
Average Power Consumption [kW]	1.66	1.42	1.18
Average Power Prosumption [kW]	0.00	-0.60	-0.31
Stdev Power Consumption [kW]	0.67	0.80	0.56
Stdev Power Prosumption [kW]	0.00	0.41	0.29
Annual Energy Consumption [kWh]	14,511.21	10,503.18	9,800.48
Annual Energy Prosumption [kWh]	0.00	832.34	129.64
PV Energy Production [kWh]	0.00	4,840.37	4,840.37
Self-consumption Rate [%]	0.00	82.80	97.32
Self-sufficiency [%]	0.00	27.62	32.46
Optimal Self-sufficiency [%]	0.00	33.36	33.36

Table 3.2: KPI's for 1-house community located in Tampa, Florida.

Each created community has a lower total annual energy consumption than the previous one. This can be attributed to the Load consumption level (socioeconomic factors), the consumption behaviors molded by the climatic changes between each zone or the human consumption habits considering the use of different power-consuming devices.

Although both Net Load distributions vary between communities, the patterns share the same basic behavior: PV energy production during the day and peaking consuming hours at some point of the morning or evening.

	Load	**Load + PV**	**Load + PV + ESS**
Average Power Consumption [kW]	0.86	0.83	0.50
Average Power Prosumption [kW]	0.00	-1.00	-0.56
Stdev Power Consumption [kW]	0.33	0.44	0.23
Stdev Power Prosumption [kW]	0.00	0.51	0.35
Annual Energy Consumption [kWh]	7,550.49	4,725.10	3,114.17
Annual Energy Prosumption [kWh]	0.00	3,050.79	1,439.86
PV Energy Production [kWh]	0.00	5,876.18	5,876.18
Self-consumption Rate [%]	0.00	48.08	75.50
Self-sufficiency [%]	0.00	37.42	58.76
Optimal Self-sufficiency [%]	0.00	77.83	77.83

Table 3.3: KPI's for 1-house community located in San Francisco, California.

For the community with the lowest annual energy consumption, the energy produced by the same 3.5 kWp PV asset makes greater self-sufficiency, as there is less energy need than in the other cases. However, self-consumption does not necessarily increase proportionally.

There is a certain point at which the ESS gets saturated (due to physical capacity and/or the algorithm's characteristics) and is therefore unable to use more of the produced energy. This is the energy that is eventually prosumed to the grid. The greater the energy prosumption is, the more significant the gap between the actual and the optimal self-sufficiency will be.

$$\Delta S = \text{OptimalSelfSufficiency} - \text{SelfSufficiency}. \tag{3.2}$$

The performance of the Community Asset System (particularly the ESSs) can be evaluated by relating it with the *S gap* size (see equation 3.2). Note that the presence of the ESS helps reduce the *S gap* for all three locations. A Battery Behavioral Model (BBM) would be better (in terms of self- consumption optimization) if the gap were shorter for the same conditions. Please find in appendix A a comparison between the 2D points distribution plots of each community with further comments about their similarities and differences, together with the Net Load profile annual histogram.

3.2 A Country-Level Annual Data Analysis: Multi-Asset Scenarios

This section describes the results of simulating 200 10-Load Energy Communities at a country level. This same simulation was executed for different Community Asset System configurations regarding the number of PVs end ESSs. More precisely, for the set of points $\{(n_{pv}, n_{ess}) : n_{pv}, n_{ess} = 0, 1, 2, ..., 10 \ \ \& \ \ n_{ess} \leq n_{pv}\}$. Giving a total of 50 different scenarios.

Relevant KPIs are calculated for each of the 200 communities with a particular CAS. In particular, the presented method aims to describe the *community-energy-net-consumption* (CNC) by subtracting the *self-consumed-energy* (SCE) from the *community-energy-requirement* (E):

$$CNC = E - SCE. \tag{3.3}$$

If SCE were maximized, we would have a community that takes the most out of the energy it produces (i.e., consumes all of it). Thus, we can deduce that SCE is caped by *community-energy-production.* In this section, more work on the description of the SCE KPI is presented, together with a mathematical model. The section describes how the *self-consumption* varies with respect to the energy a particular community requires (E) and the amount of PVs and ESSs the 10-Load community has among its members.

3.2.1 Model's Base Functions outline

The following lines present the essential functions (in correspondence with power-Load signals) from which the KPIs can be derived. To begin with, we have three main types of signals for an arbitrary energy community: one for a Load asset type, one for a PV asset type, and one for the ESS asset type. All of which are power time series.

Let's define those signals as $L_i(t), PV_j(t)$, from which an additional $ESS_{ij}(T)$ time series is derived after processing the previous signals with the battery charge-discharge strategy algorithm. Samples of these signals can be observed in figures 3.1, 3.2, and 3.3, respectively.

Then, for each community, the sum of all the signals grouped by asset type results in a general signal. Remembering equation 3.1, we have that:

$$C_{Load}(t) = \sum_{i=0}^{N_{Loads}} \mathrm{Load}_i(t), \tag{3.4}$$

$$C_{PV}(t) = \sum_{j=0}^{N_{PVs}} \mathrm{PV}_j(t), \tag{3.5}$$

and

$$C_{ESS}(t) = \sum_{k=0}^{N_{ESSs}} \mathrm{ESS}_k(t). \tag{3.6}$$

Remembering that every general signal (meaning the sum of each member's x-asset-type signal) can be analyzed from the sum of its components, let there be two more functions defined as

$$Consumption(f) = \left\{ \begin{array}{ll} f, & \text{if } f \geq 0 \\ 0, & \text{if } f < 0 \end{array} \right\} \tag{3.7}$$

and

$$Prosumption(f) = \left\{ \begin{array}{ll} -f, & \text{if } -f \geq 0 \\ 0, & \text{if } -f < 0 \end{array} \right\}, \tag{3.8}$$

which split a particular signal into its consumption and prosumption components.

With these functions, the KPIs can be calculated as follows:

Community's Energy Requirements: E

$$E = \int_{\text{year}} Consumption\left(C_{Load}(t)\right) dt. \tag{3.9}$$

Community's Net Consumption: CNC

This KPI is case-dependent. The case is marked on the subtext. We'd then have:

$$CNC_L = \int_{\text{year}} Consumption\left(C_{Load}(t)\right) dt, \tag{3.10}$$

$$CNC_{L+PV} = \int_{\text{year}} Consumption\left(C_{Load}(t) + C_{PV}(t)\right) dt, \tag{3.11}$$

or

$$CNC_{L+PV+ESS} = \int_{\text{year}} Consumption\left(C_{Load}(t) + C_{PV}(t) + C_{ESS}(t)\right) dt. \tag{3.12}$$

Community's Energy Production: b

$$b = \int_{\text{year}} Prosumption\left(C_{PV}(t)\right) dt. \tag{3.13}$$

Community's Net Prosumption: CNP

This is also a case-dependent KPI, thus having:

$$CNP_L = \int_{\text{year}} Prosumption\left(C_{Load}(t)\right) dt = 0, \tag{3.14}$$

$$CNP_{L+PV} = \int_{\text{year}} Prosumption\left(C_{Load}(t) + C_{PV}(t)\right) dt, \tag{3.15}$$

or

$$CNP_{L+PV+ESS} = \int_{\text{year}} Prosumption\left(C_{Load}(t) + C_{PV}(t) + C_{ESS}(t)\right) dt. \tag{3.16}$$

Battery Usage: $B_{\textbf{usage}}$

$$B_{\text{usage}} = \int_{\text{year}} Consumption\left(C_{ESS}(t)\right) dt \tag{3.17}$$

Based on these base expressions, other KPIs can be calculated.

3.2.2 KPIs relations to Base Functions

To begin with, the *self-consumption* (SC) can be expressed in terms of the difference between the *energy-requirements* of a community and the *community-net-consumption*:

$$SC = E - CNC_{L+PV+ESS}. \tag{3.18}$$

It is possible to express (separate) the self-consumption in terms of the energy savings each asset provides as:

$$SC_{PV} = E - CNC_{L+PV} \tag{3.19}$$

and

$$SC_{ESS} = CNC_{L+PV} - CNC_{L+PV+ESS}. \tag{3.20}$$

Note that the expression above is always larger than or equal to zero. This is because the net consumption will necessarily have to be lower in the case where a battery is present. That is the primary purpose of a battery system.

The self-consumed energy attributed to the effect of the ESSs should be understood in this context as an extra self-consumption added by the ESSs. This means that what is referred to as the SC_{ESS} is not equal to the total amount of energy that flowed through the battery.

Then for the *self-sufficiency* (SS) the following equivalent expressions:

$$SS = \frac{SC}{E}100\%,$$

$$SS = \frac{SC_{PV} + SC_{ESS}}{E}100\%, \tag{3.21}$$

and

$$SS = SS_{PV} + SS_{ESS}.$$

For the *optimal-self-sufficiency* (OSS):

$$OSS = \left\{ \begin{array}{ll} \frac{b}{E}100\%, & \text{if } \frac{b}{E}100\% \leq 100\% \\ 100\%, & \text{if } \frac{b}{E}100\% > 100\% \end{array} \right\}. \tag{3.22}$$

Note that it is capped by 100%.

Finally, for the difference between the optimal and actual *self-sufficiency* we have:

$$\Delta S = OSS - SS. \tag{3.23}$$

Note how if we know the base functions $L(t)$, $PV(t)$, and $ESS(t)$, then we can calculate all of the KPIs for any community or Local Energy Market (LEM). However, another possibility to know what the $CNC_{L+PV+ESS}$ is for a community with certain energy requirements is to model it based on the behavior of all the simulated scenarios. The latter will be explored in the following subsections.

3.2.3 Simulation's Specifications

The following specifications were considered at the moment of creation of each of the communities:

(i) Fixed number of Loads for all of the explored scenarios. 10 Load profiles from the database for each community.

(ii) Each PV asset was considered to have a 3.5 kWp generation capacity. However, afterward, it was decided to increase the generation capacity to achieve communities with Production to Consumption (PtC) ratios ≤ 1, which was achieved with a 7kWp generation capacity per PV module. The reason for this is that the equations proposed in section 3.2.5 to model the performance, are functions of this ratio.

(iii) Each ESS asset has a storage capacity of 7MWh with a maximum charge-discharge rate of 3 kW.

(iv) The proportion of a community's High, Base, or Low consumption level profiles is a third each. This means a community will have, on average, three homes for each consumption level.

The variables in this analysis are the total amount of installed PV assets (or, similarly, the total amount of PV power installed) and the number of installed ESS assets (similar to the total storage capacity).

It is worth mentioning that the regions available to create communities are a subset of the whole country, derived from the *intersection* of both databases (generation and residential Load profiles). Although the Load profiles database is widely spread through all climate zones, the PV generation database has very little data on some areas of the country (see Figure 2.7).

That said, figure 3.10 shows how the empty regions in the previous map replicate in this one. The figure shows the yearly energy consumption for 200 communities with 10 Loads and non-Asset Systems, as described in the figure's title. These regional holes will also play a role in the PV generation geographical distribution, as will be later discussed.

3.2.4 Simulated scenarios' geographic KPIs visualization

Each scenario will be identified by a vector like $(N_{PVs}, N_{ESSs}) = (7, 5)$, for instance. It should be kept in mind that all of them have 10-Loads.

It has already been mentioned that each scenario comes in place when a variation of the total number of PV or ESS assets is introduced. Another restriction over the system, which appears because of the simulation setup algorithm, is that the number of ESS assets (l), the number of PV assets (m), and the number of Loads (n), hold the relation $l \leq m \leq n$.

The following sections will show data representations for some limit (N_{PVs}, N_{ESSs}) scenarios. For instance (0, 0) represents a no production and no storage scenario, while (10, 10) represents the scenario with maximum storage and generation peak power. The scenario (10, 0) represents a community with maximum generation capacity, but without storage availability.

Scenario 1: (0, 0)

Let's begin by discussing figure 3.10, the community's annual consumption, which simultaneously represents the *community-energy-requirement* (E). This is the only relevant KPI for this particular case because all of the others are zero due to the lack of PVs and ESSs assets. The color bar illustrates the Energy spectrum in which 10-Load communities' energy consumption level is situated for the US modeled households.

Note that the *annual-energy-requirement* correlates strongly with the climate zones used to model the Load profiles (see figure 1.2). The south-eastern communities with the highest annual consumption probably correspond to the *hot-humid* climate zone. The intermediate consumer types would be located in the *mixed-humid* region. The *hot-dry* area presents as well an almost average consumption. The *cold* part of the country would be the next on the list, appearing to have just a slightly higher consumption than the communities on the west coast, in the *marine* climate zone. It will be shown in the following scenarios that energy consumption presents a high correlation with most of the KPIs.

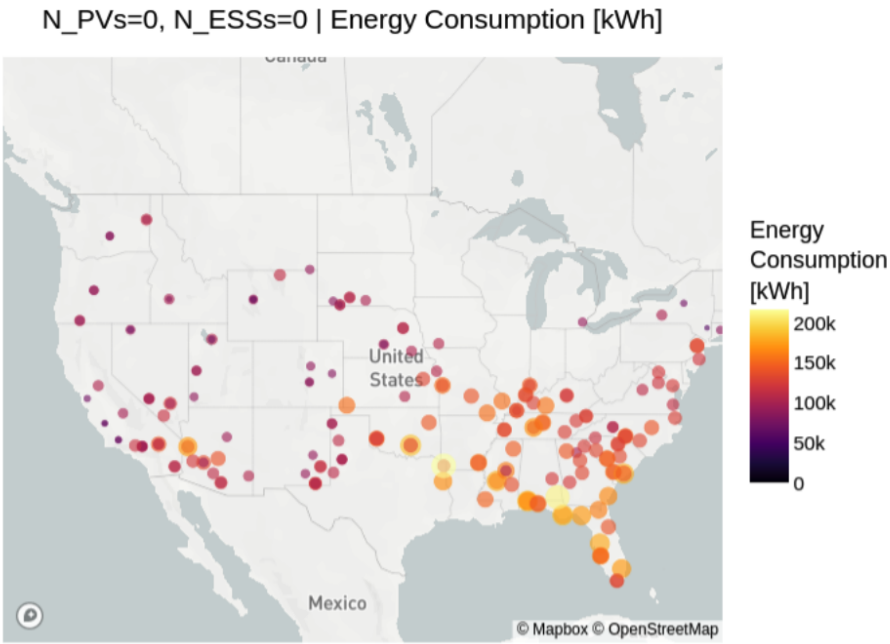
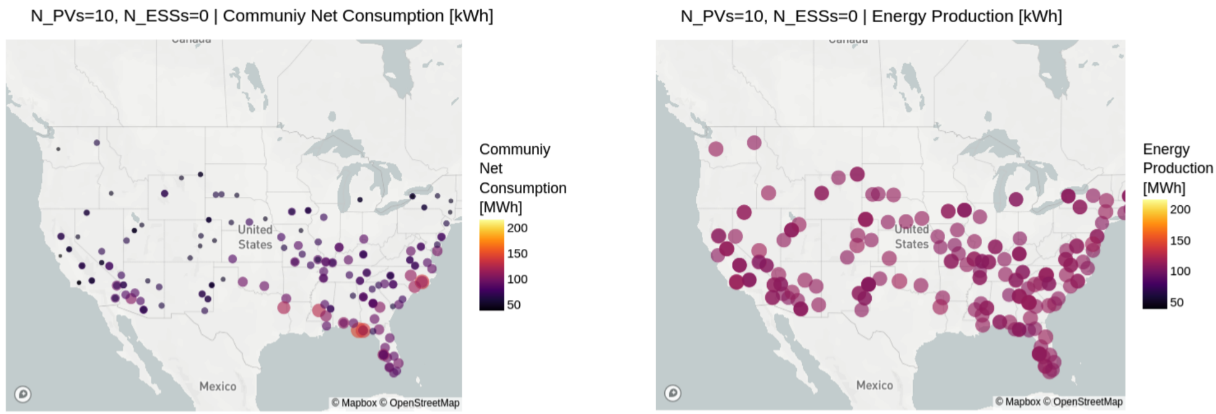


Figure 3.10: Regional Communities Energy Requirements Distribution at country level for scenario (0, 0).

Scenario 2: (10, 0)

In figure 3.11a, the same information as in the previous scenario is shown but now includes the PV component and the KPIs derived from that. The consumption level has decreased significantly because the community uses a portion of the PV energy produced.

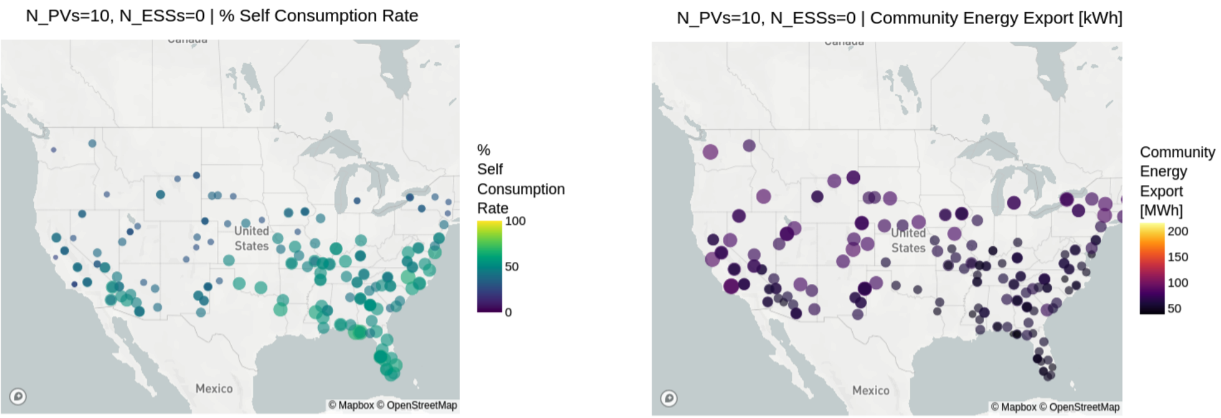


(a) Annual Energy Net Consumption (MWh) **(b)** Annual Energy Production (MWh)

Figure 3.11: Annual Energy Consumption and Production KPIs for scenario (10, 0). Color & size scale with energy.

The average annual consumption has been reduced by about 50 MWh compared to the first scenario. Not all of the produced energy is used, however.

This affirmation is supported by figure 3.12a, which shows that self-consumption is now greater than zero but less than 100% for most communities. Also, for the communities with the lowest self-consumption rates, figure 3.12b shows that they have the highest yearly prosumption rates.



(a) % Self Consumption

(b) Community Net Prosumption (MWh)

Figure 3.12: Self Consumption (a) & Community Net Prosumption (b) for scenario (10, 0).

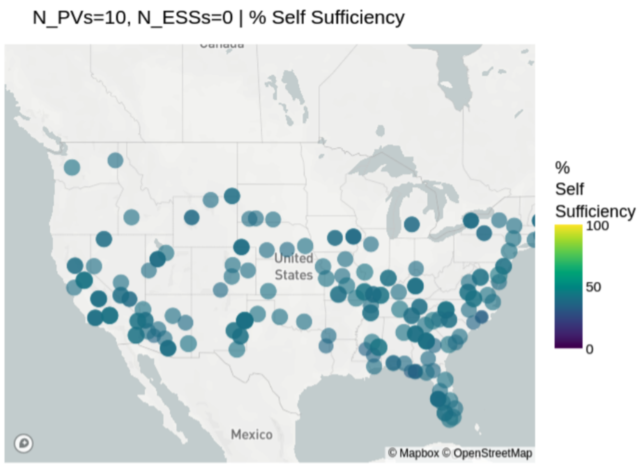


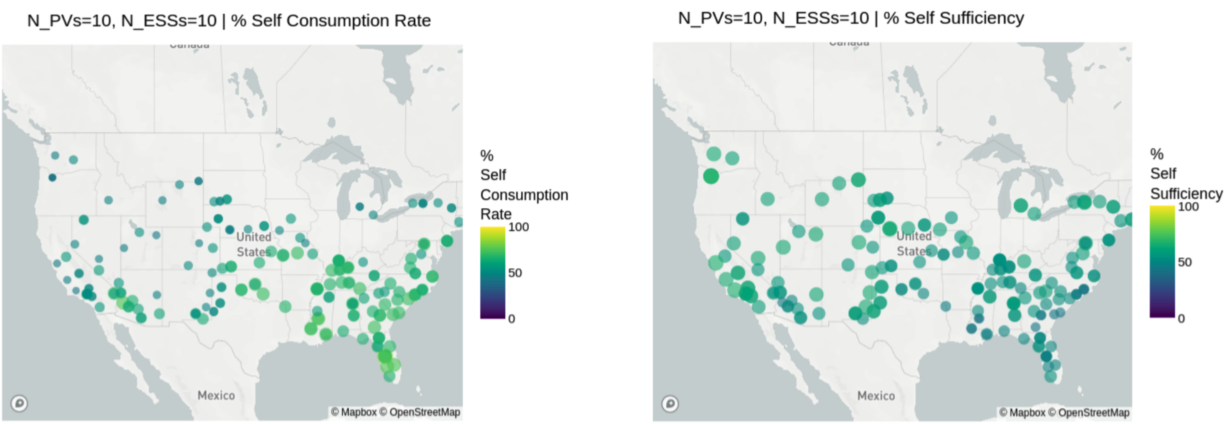
Figure 3.13: Communities' % Self Sufficiency for scenario (10, 0).

Note that for this scenario, the self-sufficiency is entirely attributable to the PV assets, therefore equivalent to the PV component of the self-sufficiency). Also, figure 3.11b shows that the PV energy production is very average throughout the territory without any evident geographic correlation.

Distributions based on climatic zones are generally observed for all KPIs, except for the PV generation. According to [7], we could expect a slight dependency on the generation based on latitude and a more noticeable correlation with climatic zones. A plausible explanation for this could be data processing. In particular, the signals coupling procedure (explained in section 3), in combination with the clustering applied to the available PV plants, the total energy production distribution is softened across the geographic region.

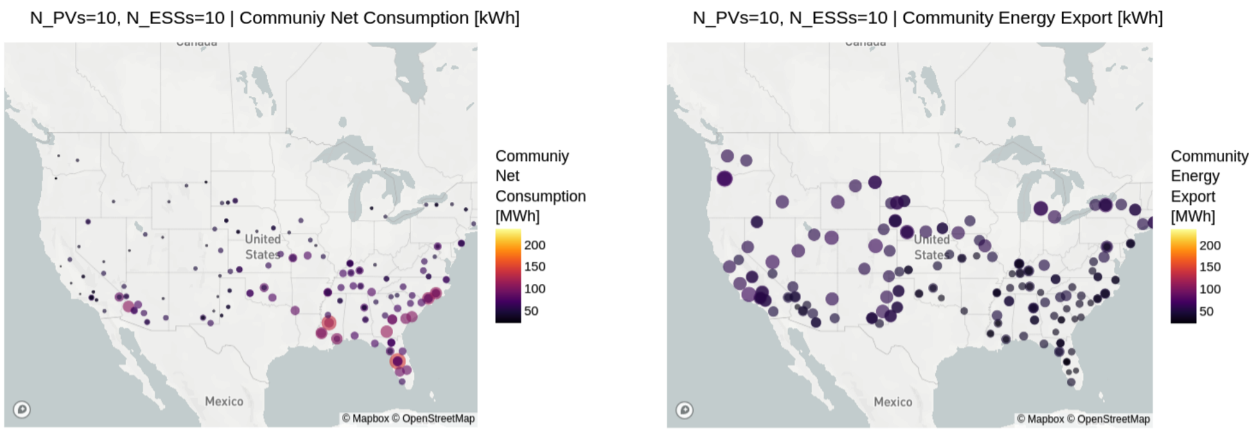
Scenario 3: (10, 10)

In this limit case scenario, the maximum amount of PV generation and storage capacity are present for each community. Figures 3.14 and 3.15 show the KPIs on the maps. Note that for this scenario, the ESSs cause the self-consumption rate and self-sufficiency to increase for some communities. Note also that the yearly prosumption has been drastically reduced due to the ESS assets.



(a) % Self Consumption Rate **(b)** % Self Sufficiency

Figure 3.14: Self Consumption Rate & Sufficiency KPIs for scenario (10, 10).



(a) Annual Community Net Consumption (kWh) **(b)** Annual Community Net Prosumption (kWh)

Figure 3.15: Annual Community Net Consumption & Prosumption for scenario (10, 10).

Note that this scenario produces the communities with the highest self-consumption rates (favored by the amount of ESS assets), the highest self-sufficiency rate (favored by the number of both PV and ESS assets), and the lowest yearly energy consumption.

These map plots show qualitatively two main things: The main correlation encountered is that each region will have different average energy requirements for the location variation. Secondly, the other KPIs are correlated geographically based on a similar pattern, a consequence of a direct correlation with the energy requirements. The latter assertion motivated the following and final analysis.

3.2.5 Correlation of the basic community's KPIs with its Energy Requirements

Based on the calculations shown in section 3.2.2, it is known that a community's KPIs can be deduced by knowing the SC_{PV} and the SC_{ESS}. Therefore, two general mathematical functions that reflect the behavior of the mentioned KPIs when varying E, N_{PV}, and N_{ESS} are proposed.

These equations are proposed as a result of the data behavior observations. In particular, asymptotic behavior for both small and large PtC (Production to Consumption $[E/b]$) ratios was observed, for which the fits were based on sigmoidal functions, which portrays this pattern. In section 3.3, more insight into these equations is provided.

For the SC_{PV}, the following expression is proposed:

$$SC_{PV} = 2b\left[\frac{1}{1+\exp\left(-k_1\frac{E}{b}\right)} - \frac{1}{2}\right]. \tag{3.24}$$

And for the SC_{ESS}:

$$SC_{ESS} = B\left[1 - \frac{1}{1+\exp\left(-k_2\frac{E-E_0}{b}\right)}\right]. \tag{3.25}$$

As for the parameters, b and B hold a proportional relation with N_{PV} and N_{ESS}, respectively, as described below:

$$b \approx \epsilon_1 N_{PV}\left[n_{pv}\right] PP\left[\frac{MWp}{n_{pv}}\right] AVGSunLightHours\left[\frac{h}{day}\right] 365\left[\frac{day}{year}\right],$$

where PP is the peak installed power generation for each PV module, and the *AVGSunLightHours* term should be a function of the Latitude and solar radiance for particular coordinates. ϵ_1 can be considered as an efficiency factor or something of the sort.

As for B, it can be approximated as:

$$B \approx \epsilon_2 N_{ESS}\left[n_{ess}\right] StorageCapacity\left[\frac{storageMWh}{n_{ess}}\right] N_{\text{FullChargesPerDay}}\left[\frac{h}{day}\right] 365\left[\frac{day}{year}\right].$$

Both b and B have units of $[MWh/year]$. Equivalently, ϵ_2 would represent an efficiency factor for the ESS usage yet to determine.

The other parameters $k1$, $k2$, and $E0$, are to be fit for each scenario, which produces a set of points that enable the study of the parameter's correlation with the variables. Further work can ultimately determine a relation between the coefficients and the used variables (E, N_{PV}, N_{ESS}).

Figures 3.16a, and 3.16b show the point distributions and the proposed fit for SC_{PV} and SC_{ESS}. Each fit produces a set of adjusted parameters, which can be further analyzed to express in terms of the asset scenario.

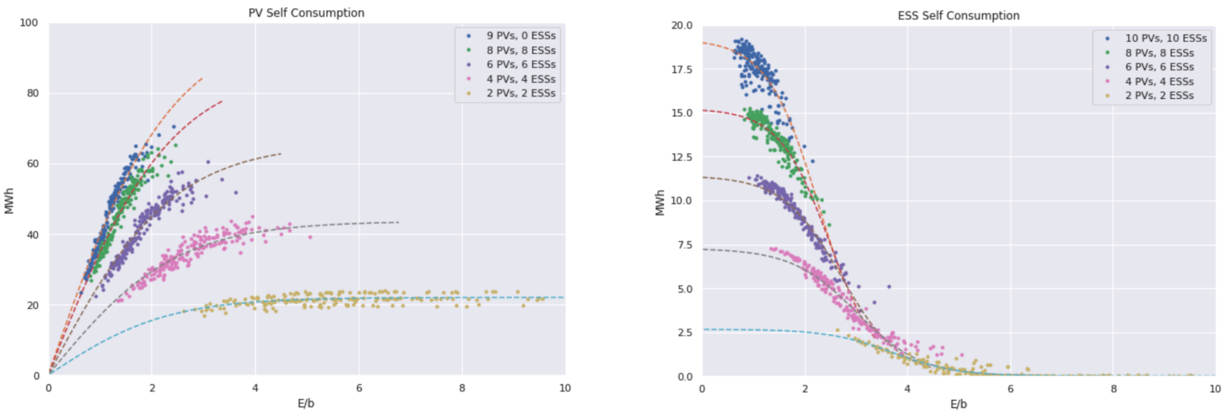
As discussed before, all of the other KPIs curves can be obtained from those two curves. Figures 3.17a, 3.19a, 3.19b, 3.20 show the points distribution with the fit curves referenced before. Also, for the coefficient adjustment, correlations with the amount of PVs and ESSs are shown in figures 3.21a, 3.21b and 3.22.

3.3 Discussion

This section will discuss the graphs from the above figures according to their analysis and intermediate steps to compute the model fitting properly: First, self-consumption and self-sufficiency KPIs are analyzed. Then, a coefficient visual inspection is presented. Finally, the bond between the CAS performance (measured by the KPIs) and the geographic location is clarified and discussed.

Self Consumption

Figures 3.16a and 3.16b present two different dependency patterns concerning the energy requirements. By substituting $x = E/b$, the referenced figures were obtained.



(a) PV Self Consumption vs E

(b) ESS Self Consumption vs E

Figure 3.16: Energy Self Consumption: PV and ESS components.

For the SC_{PV}, it is observed that it tends to reach a saturation, which coincides with the amount of produced energy b. However, how this saturation occurs is more relevant. For low $\frac{E}{b}$ values, a linear increment is observed. As the values grow, the growth rate starts to drop until the function does not grow anymore.

For the SC_{ESS}, it is observed that when the energy requirement increments, the additional energy savings shares provided by the ESSs become more and more insignificant. This is mainly because households that consume more energy will consume the produced energy more efficiently. Also, low values tend to a saturation point, low-capped by the energy storage capability.

With a first derivative analysis for equations 3.24 and 3.25, together with the boundary conditions, it can be deduced that the grow/decrease rates for the proposed functions are:

$$\frac{\partial SC_{PV}}{\partial E}(E = 0) = \frac{k_1}{2} \tag{3.26}$$

and

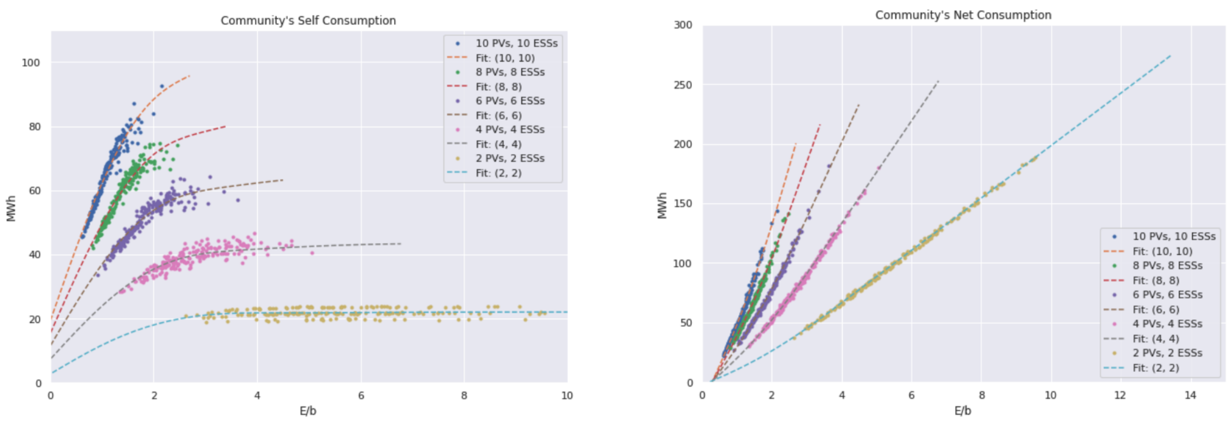
$$\frac{\partial SC_{ESS}}{\partial E}(E = E_0) = -\frac{Bk_2}{4b}. \tag{3.27}$$

Therefore, those coefficients are an indicator of how 'fast' these curves tend to their saturation states. Coefficient E_0 is the energy requirement for which the ESS system provides energy savings equivalent to half its storage capacity.

All of the fit parameters have correlations with N_{PV} or N_{ESS}. Figures 3.21 and 3.22 show that the dependency is much stronger with N_{PV}. Further work on these correlations needs to be done to determine a general expression for each.

Figure 3.17a shows how the batteries' effect becomes visible because low energy requirements are closer to the saturation value than without them. To observe this effect more clearly, compare figure 3.17a with figure 3.16a and note how the points are closer to the horizontal asymptote.

Having a complete expression for the self-consumption, it is possible to calculate the *community-net-consumption* directly. This function fit provides the possibility to estimate the values for communities with PtC ratios that were not necessarily represented in the simulations.



(a) Total Self Consumption vs. E

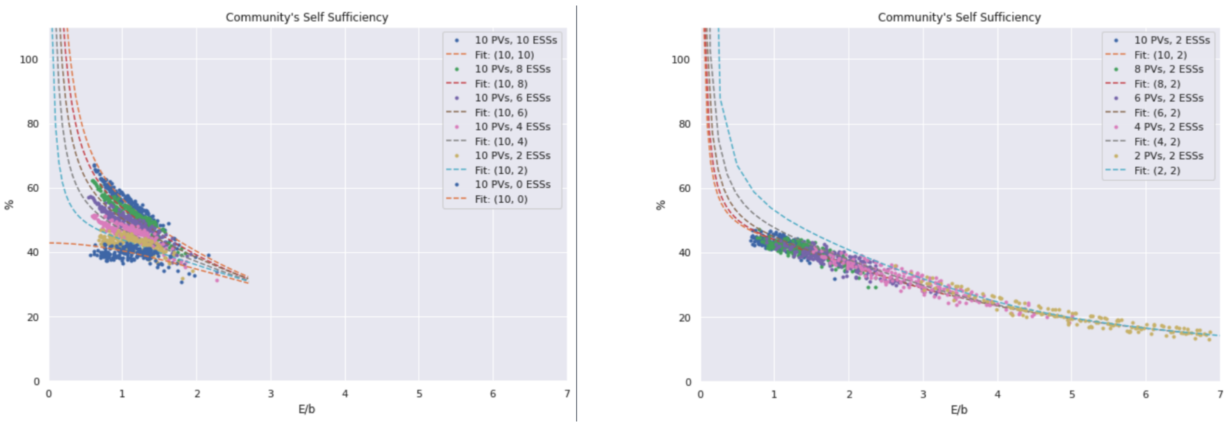
(b) Community's Net Consumption vs. E

Figure 3.17: Energy Self Consumption and its impact on the Community's Net Consumption.

It is yet to be proven that the fit is still good for even lower energy requirements. Those energy gaps can be filled by recreating the simulations so that communities with lower PtC ratios are created. This can be achieved by increasing the generation capacity, for instance. This is also potential future work.

Self-Sufficiency and the ΔS Gap

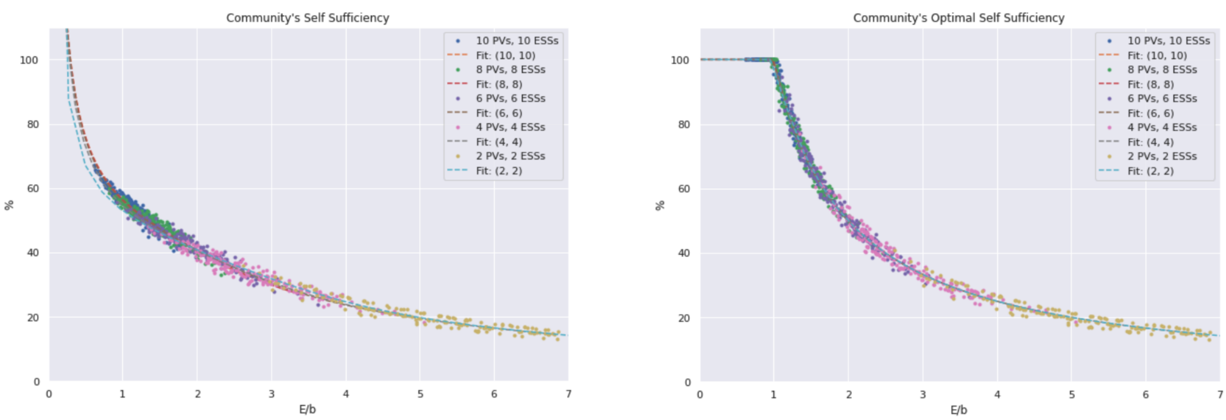
Figure 3.19 show both the actual *self-sufficiency*, and the *optimal-self-sufficiency*. According to the used fits, a theoretical point exists at which a community could be 100% self-sufficient with a particular CAS. After all, the curve still has a $1/E$ dependency (see equation 3.21).



(a) **(b)**

Figure 3.18: Self-Sufficiency for different U-StP ratios (Unitary Storage to Production ratio).

It is yet to be proven that the fit is good for lower energy values. One possibility is that the self-sufficiency obtained from the simulation results (the actual self-sufficiency such a system would have) ultimately reaches a saturation point in correspondence with the physical limitations of the CAS and the management strategy (Battery Behavioral Model) that will not necessarily be 100%. One more time, further work could address this. Determining the performance limits of the CAS would make it easier to measure the CAS performance directly, as it could serve as an optimization baseline.



(a) Self Sufficiency vs. E **(b)** Optimal Self Sufficiency vs. E

Figure 3.19: Actual/Optimal Self Sufficiency.

Let's stay for now with ΔS, which is a comparison between the actual *self-sufficiency* and the hypothetical of consuming all of the produced resources (*optimal-self-sufficiency*). Figure 3.20 shows ΔS for two different Unitary Storage to Production (U-StP) ratios: 1 and 1/2 (these represent a situation where there is a battery for each PV, and where there is a battery for each pair of PVs, respectively). Note how the curve with the highest U-StP ratio has lower ΔS values, meaning that the self-sufficiency is closer to the optimal, or in short, that there is more self-consumption.

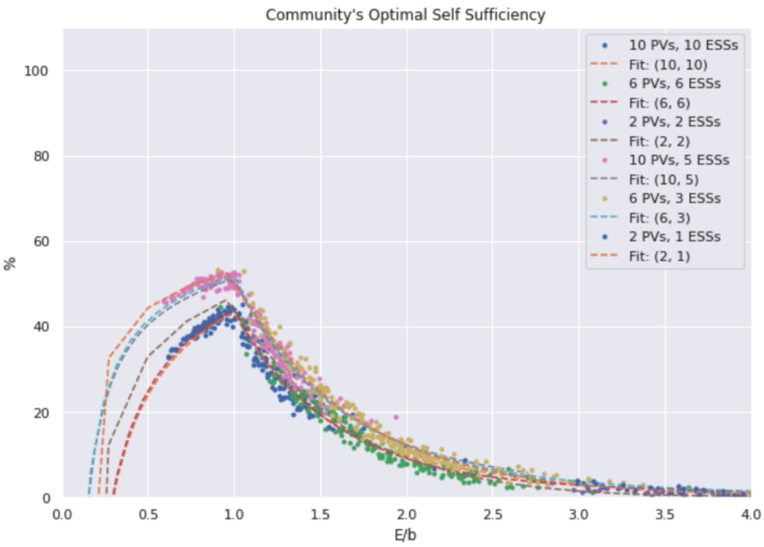


Figure 3.20: ΔS gap vs. E

Figure 3.18, shows how the greater the U-StP ratio is, the greater is also the self-sufficiency. However, note that for large PtC ratios (E/b values), the differences between schemes with different U-StP ratios decrease, because, as said before, communities with higher energy requirements (Consumption) have greater chances to completely use the PV resources without the need of storage.

As for how much the self-sufficiency increases for different CAS configurations, it will depend on the individual asset's characteristics because that is what determines how much a single PV panel produces and how much an ESS stores in a whole year. The answer to this question could be found by manually calculating with the proposed fit model without the need to run more simulations.

Adjusted Parameters: k_1, k_2, and E_0

For each CAS scenario, values for k_1, k_2, and E_0 were calculated so that the error is minimized for the fit functions introduced in equations 3.24 and 3.25. That produced three lists of values for each one of the fit parameters. The variables are N_{PV} and N_{ESS}. Figures 3.21 and 3.22 show the correlations between the parameters and the variables.

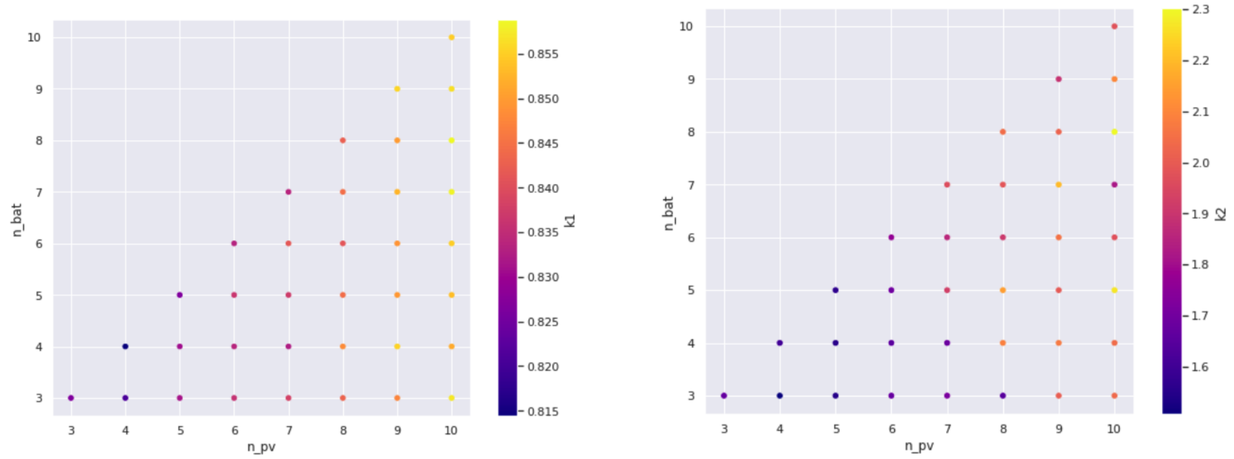
At first glance, $k1$ and $E0$ can be proposed as:

$$k_1 \approx \alpha_1 N_{PV} = \alpha_1^* b \tag{3.28}$$

and

$$E_0 \approx \alpha_2 N_{PV} = \alpha_2^* b. \tag{3.29}$$

However, the correlations for k_2 are not yet that clear. It could be that the correlations are not precisely linear. Finding a general expression for k_1, k_2 and E_0 is left out as future work.



(a) k1 **(b)** k2

Figure 3.21: k1 and k2 coefficients correlations with N_{PV} and N_{ESS}.

If we suppose the latter equations are valid for the moment, making the connection with equations 3.26 and 3.27 reveal two things, the first being that the more PVs a CAS has, the higher the rate with which the self-consumption increases. Secondly, the amount of PV assets also impacts E_0 linearly. As a reminder, E_0 is the energy requirement at which a community's ESS provides extra energy savings (*self-consumption*) equivalent to half of the storage capacity. Therefore, the larger E_0 is, the most chances the ESSs have to provide savings to the community.

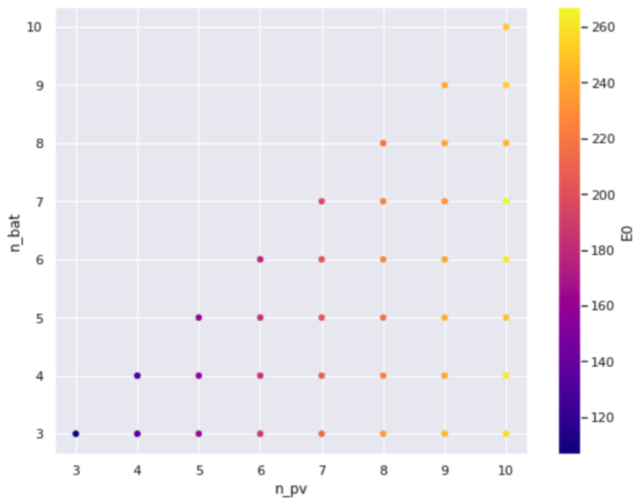


Figure 3.22: E0 coefficient correlations with NPV and N_{ESS}.

Geographic Correlation

The variable from which a geographic correlation has been observed (and from which all of the KPIs inherit their correlation) is the energy requirements.

However, suppose it is assumed that the average household behavior (Load consumption basic pattern) persists across regions. In that case, the model could be said to represent the CAS's performance for any 10-Load-Multi-Asset CAS in the US, based only on its energy requirements. As already discussed, these vary from region to region. As a first guess, the location does not have a direct impact on a community's self-consumption, but an indirect one as:

$$E = E(Lat, Long).$$

Thus, we have:

$$SC = SC(E) = SC\left(E(Lat, Long)\right) = SC(Lat, Long).$$

In a more strict sense, the produced energy b should also have a geographic correlation attributable to different climate zones. Therefore, in a more refined model, SC would directly relate to the community's location. In this work, the fluctuations around the b mean value are attributable to the mentioned geographic correlation, possibly with other different aspects.

Limitations of the Self Consumption Prediction Model

The model for SC_{PV} and SC_{ESS} were fit based on the observations. However, only a particular energy requirement spectrum was explored, being determined by the household Load profiles database. The presented work leaves the low energy requirement range unexplored and is yet to determine whether the proposed fits reproduce the simulation behavior for those values. Therefore, the model corresponds primarily with the explored energy spectrum (approximately between 50 and 200 $\frac{MWh}{year}$).

Chapter 4

Conclusions

In this chapter, conclusions regarding the work's limitations and simulation results are given. Also, further work possibilities are mentioned, together with further ideas about microgrids or local energy communities.

4.1 General Conclusions

The energy industry is one of the largest and most important industries worldwide and continuously renovating towards digital technologies. Currently, huge efforts are being made to integrate renewable resources into the energy supply chain. In recent years, distributed generation has been proven to be effective in increasing a community's resiliency, as well as its *self-sufficiency* against the wider grid. This leads to incorporating asset systems (PVs, ESSs, EVs, Wind Generators, heat pumps, among others) into communities, creating the possibility of Local Energy Markets in which the asset owners can actively participate. The amount and type of assets a community has and how they are managed determine the system's performance.

For all of the above reasons, having a better understanding of Energy-Communities (microgrids) gains more importance each time. A mathematical model allows one to analyze such a system from the abstract mathematical point of view. Therefore, the analysis of these systems would not entirely depend on simulations alone, as they could be translated and processed algebraically. This would help when trying to respond to particular questions regarding the systems. Also, understanding the model's behavior by analyzing it via these functions can be used to find optimal configurations for a CAS (asset-wise and strategy-wise).

This work presents a methodology to analyze and mathematically represent an Energy Community applied in a country-level use case. The simulation shows that the main factor influenced by the region is the household's consumption levels and habits. In a much less significant way, the energy generation is also region-dependent.

The study also provides evidence that using batteries enables better use of the produced energy. The proposed fits can be used to determine analytically how much the savings would be incremented.

It should be kept in mind that different BBMs based on various market strategies could produce different levels of self-sufficiency without necessarily increasing the number of PV and ESS assets. The presented results correspond to an *Evening's Load Peak Shaving - Battery Behavioral Model*.

The proposed approach can serve to characterize different types of CAS. To do that, all the statistical analysis and simulation that has been carried out on this document can be repeated for different systems configurations varying possibly the following parameters:

- Community's PV max production capacity (number of PVs or changing the max generation capacity for single PV assets)
- Community's maximum battery storage potential (number of ESSs or varying the storage capacity for single ESS assets)
- Community's total numbers of Loads
- Battery Behavioral Model

With a more sophisticated simulation engine, such as the *Grid Singularity Exchange* that espouses an agent-based hierarchical, bottom-up market design [17], other variables such as an asset's trading strategies, dynamic grid tariffs, or even the user's trading partner preferences could be taken into consideration.

4.2 Future Work

In microgrid simulations in correspondence with the existing system, many variables are in play simultaneously. Understanding how the system behaves concerning as many variables as possible is needed to understand the system thoroughly.

Therefore, future work is required to study the system changes with other variables, such as the number of Loads. Also, different system sizing approaches, P2P trading KPIs, and new BBMs with different optimization strategies could be considered. Some sample BBMs strategies could be: *multiple peak modulation/shaving*, *maximization of self consumption rate*, *maximization of self sufficiency rate*, *maximize earnings*, among others.

Currently, digital marketplaces are being developed to sell and buy such intelligent management strategies [6] [19]. The development and delivery of optimized BBMs and Energy Management Strategies could be adapted for potential business models.

Of interest is also to develop a mathematical representation of the market model in a similar way as was done for the power signals. The market models would include information about the trading peers' digital identity, the energy volume traded, and the rate it was exchanged for.

Relevant economic KPIs can be studied similarly as presented in this thesis to understand and optimize the community's energy and monetary performance. For that, the further development of the mathematical expressions which describe the KPIs would provide a more robust description of the performance, which implies that the correlations between the variables and KPIs are at some point understood.

Appendix A

Regional comparison for the 1-house community scenario

As mentioned in section 3.1, the same scenario was reproduced for locations in Tampa and San Francisco. Figures A.1 and A.2 show the Net Load histograms for communities in both cities.

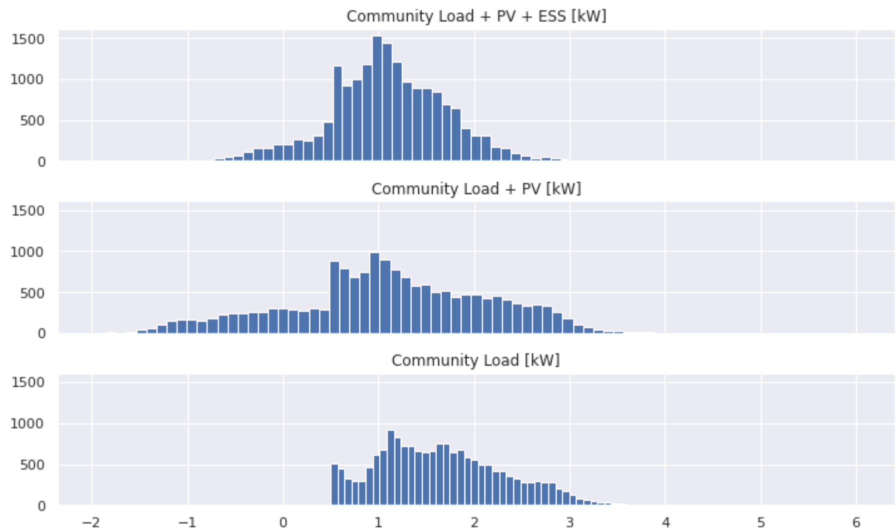


Figure A.1: Power histograms for a 1-house community in Tampa, Florida.

As shown in tables 3.2 and 3.3, San Francisco Community's yearly energy consumption is the smallest of the three exemplified communities. At the same time, it has a similar PV energy annual production rate, therefore having a greater opportunity to satisfy its energy needs with higher self-sufficiency. The unused self-produced energy appears higher than for the Tampa and Charlotte communities, based on the several occurrences at which energy is exported to the grid. Note also how the Load peaks are very effectively reduced.

Figures A.3 and A.4 help to observe how the asset system acts over the community's Net Load profile. Note how for figure A.4 the BBM discharges the battery at the peak Loads very effectively.

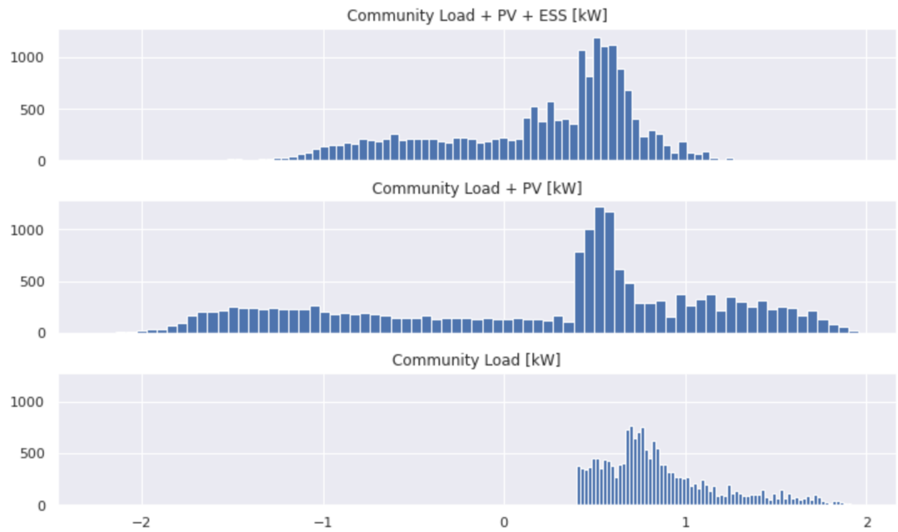


Figure A.2: Power histograms for a 1-house community in San Francisco, California.

In figure A.3, the discharge hours do not coincide with the same precision level from San Francisco. Also, the highest peaks occur at some few morning intervals, which correspond most likely to cold winter mornings.

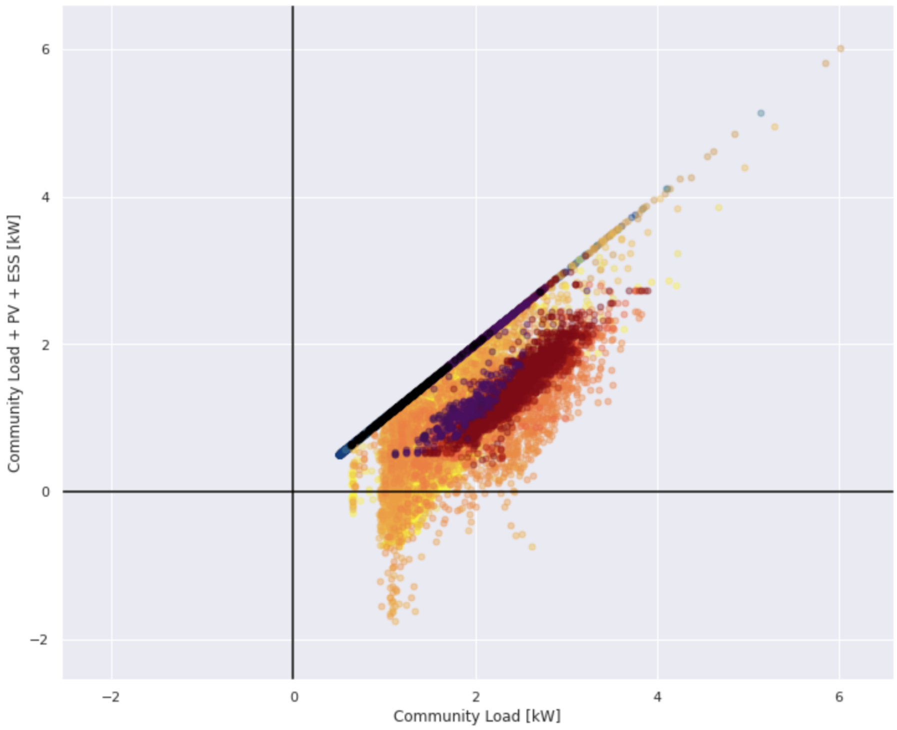


Figure A.3: Cases net load comparison. Tampa community.

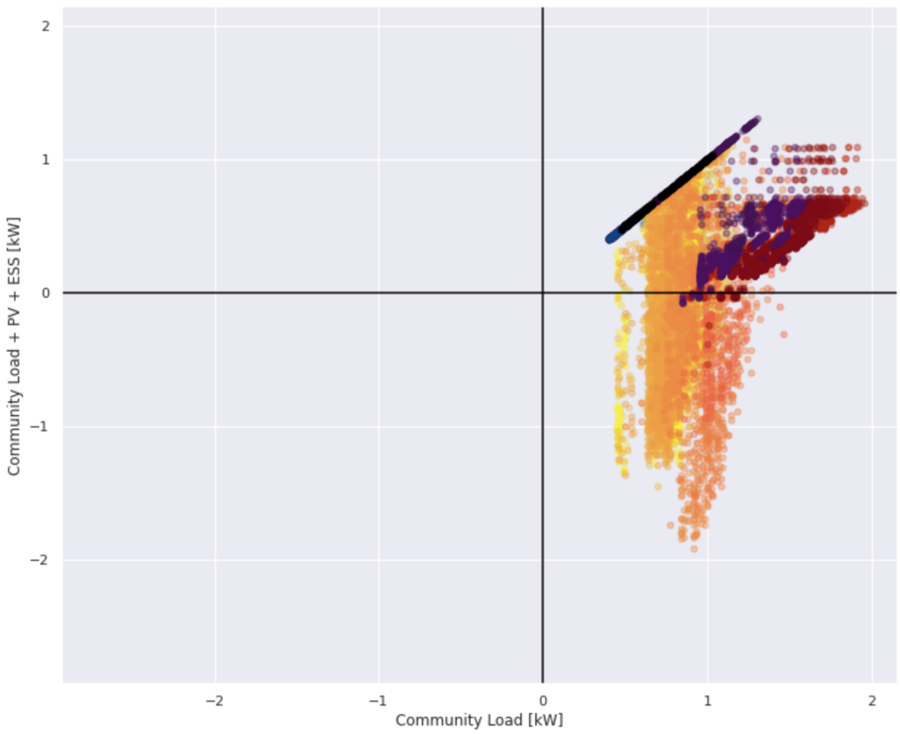


Figure A.4: Cases Net Load comparison. San Francisco community.

Appendix B

Further ideas on the Energy Community Math Model

B.0.1 An extension for the Energy Community model

With regards to point (iv) on this section's introduction, it can be theorized that for a given group of communities, a certain finite number of consumer types will appear. Based on which, a community's net load (Great Load)can be modeled as follows:

$$GLoad(t) \approx \sum_{\text{Consumer Type}} n(ConsumerType) L_{\text{Consumer Type}}(t) \tag{B.1}$$

Where L_i is the curve which describes the i-th consumer type on the community, and $n(i)$ a function which returns the total number of consumers of a certain (i-th) consumer type. A method to determine L_i for all consumer types could be worked upon in future research.

If the same idea is extended to PV curve types (perhaps determined by peak power, PV panel age, efficiency, or other variables) and ESS behavioral model strategies (as well depending on the algorithm, storage capacity, and max charge-discharge power among other possible dependencies), equation 3.1 can be rewritten as:

$$GProfile(t) = GLoad(t) + GPV(t) + GESS(t) \tag{B.2}$$

where

$$GPV(t) \approx \sum_{\text{PV Type}} m(\text{PV Type}) P_{\text{PV Type}}(t) \tag{B.3}$$

and

$$GESS(t) \approx \sum_{\text{ESS BM}} l(\text{ESS BM}) B_{\text{ESS BM}}(t) \tag{B.4}$$

are the Great PV and Great ESS profiles, which correspond to the community's PV Asset System's and the ESS Asset System's overall behavior. And n, m and l functions return the number of assets of such a kind.

Under the scope of this model, the current study uses the load profiles corresponding to a Low, Base and High load level from the residential load profiles database as as $L_{\text{Consumer Type}}(t)$ functions (3 in total, as was already mentioned). Analogously, the $P_{\text{PV Type}}(t)$ functions are a representation of the available PV profiles within the region determined by the segmentation procedure, whose outcome is visible in figure 2.7 by using the solar plants database. Finally, $B_{\text{ESS BM}}(t)$ would represent a "peak shaving" charge-discharge power schedule determined by a Behavioral Model algorithm, however the possibility to use several within the same community exists. In this case, $B_{\text{ESS BM}}(t)$ is also dependent of $P_{\text{PV Type}}(t)$ and $L_{\text{Consumer Type}}(t)$, therefore $B_{\text{ESS BM}}(t, P, L)$ However a different $B_{\text{ESS BM2}}(t)$ could have only a dependency with time, or it's dependencies could be others, as for instance the market energy prices.

Note how a person, company or institution could have control over the ESS power curve by programming a particular algorithm based on the Consumer Type and PV Type signals [18]. Hence, this would be an *easy controllable* variable. PV Type signals would not be really controllable, however, they are highly predictable based on weather conditions [23]. Load profiles are also predictable based on historical data. However this tend to have higher error margins.

Further work could be the development of a load forecasting model based on historic data, and schedule labels to better predict the consumption and then apply the ESS behavioral models to it. Another future line of work could be to include in this model a set of functions for $EV_{\text{EV User Type}}(t)$ which would represent an electric vehicle's charging (and possibly discharging) cycle's pattern.

Also, an analogy of this could be applied to the price at which the energy is sold, and study not the community's net load but the fluctuations of the energy prices. The "base" functions can be also predefined to behave in a certain way under certain conditions, and would be analogous to the ESS BM meaning that its performance could be tracked as it is done on this work for the communities net load.

Referencias

[1] K. Anderson, K. Burman, T. Simpkins, E. Helson, and L. Lisell. New york solar smart dg hub-resilient solar project: Economic and resiliency impact of pv and storage on new york critical infrastructure. 2016.

[2] K. H. Anderson, N. A. DiOrio, D. S. Cutler, and R. S. Butt. Increasing resiliency through renewable energy microgrids. *International Journal of Energy Sector Management*, 2(2), 8 2017.

[3] T. Brown, J. Hörsch, and D. Schlachtberger. Pypsa: Python for power system analysis. *Journal of Open Research Software*, 6(1):4, 2018.

[4] A. Caramizaru and A. Uihlein. *Energy communities: an overview of energy and social innovation*, volume 30083. Publications Office of the European Union Luxembourg, 2020.

[5] C. Chen, S. Duan, T. Cai, B. Liu, and G. Hu. Smart energy management system for optimal microgrid economic operation. *IET Renewable Power Generation*, 5(3):258, 2011.

[6] R. Energy. Energy data map. Available at https://energydatamap.com/.

[7] W. B. Group. Esmap. global solar atlas 2.0 technical report. 2019.

[8] J. A. Hartigan and M. A. Wong. Algorithm as 136: A k-means clustering algorithm. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 28(1):100–108, 1979.

[9] N. Hossein Motlagh, M. Mohammadrezaei, J. Hunt, and B. Zakeri. Internet of things (iot) and the energy sector. *Energies*, 13(2):494, Jan 2020.

[10] O. E. M. Initiative. Open energy modelling initiative's manifesto. Available at https://openmod-initiative.org/manifesto.html.

[11] E. Mengelkamp, B. Notheisen, C. Beer, D. Dauer, and C. Weinhardt. A blockchain-based smart grid: Towards sustainable local energy markets. *Computer Science - Research and Development*, 33(1-2):207–214, 2017.

[12] U. O. of Energy Efficiency & Renewable Energy. Building america climate-specific guidance. Available at https://www.energy.gov/eere/buildings/building-america-climate-specific-guidance.

[13] G. C. Okwuibe, A. S. Gazafroudi, S. Hambridge, C. Dietrich, A. Trbovich, M. Shafie-khah, P. Tzscheutschler, and T. Hamacher. Evaluation of hierarchical, multi-agent, community-based, local energy markets based on key performance indicators. *Energies*, 15(10):3575, 2022.

[14] C. N. Ong, Sean. Commercial and residential hourly load profiles for all tmy3 locations in the united states. *OpenEI*, 11 2014. Available at https://data.openei.org/submissions/153.

[15] OpenEI. About open energy information. Available at https://openei.org/wiki/OpenEI:About.

[16] F. Pimenta, J. Pacheco, C. M. Branco, C. M. Teixeira, and F. Magalhães. Development of a digital twin of an onshore wind turbine using monitoring data. *Journal of Physics: Conference Series*, 1618(2):022065, 2020.

[17] G. Singularity. Grid singularity's technical approach. *Medium.* Available at https://gridsingularity.github.io/gsy-e/technical-approach/.

[18] G. Singularity. Introduction to grid singularity exchange software development kit. *Medium.* Documentation available at https://energydatamap.com/.

[19] G. Singularity. Grid singularity and rebase energy awarded 2021 ai4cities grant. *Grid Singularity - Medium*, May 2021.

[20] B. Soltowski, D. Campos-Gaona, S. Strachan, and O. Anaya-Lara. Bottom-up electrification introducing new smart grids architecture—concept based on feasibility studies conducted in rwanda. *Energies*, 12(12), 2019.

[21] A. D. The International Renewable Energy Agency (IRENA). Renewable energy statistics 2022. 2022.

[22] R. P. Thombs. When democracy meets energy transitions: A typology of social power and energy system scale. *Energy Research Social Science*, 52:159–168, 2019.

[23] M. Tovar, M. Robles, and F. Rashid. Pv power prediction, using cnn-lstm hybrid neural network model. case of study: Temixco-morelos, méxico. *Energies*, 13(24), 2020.

[24] M. A. Tovar Rosas, M. Robles Pérez, and E. R. Martínez Pérez. Itineraries for charging and discharging a bess using energy predictions based on a cnn-lstm neural network model in bcs, mexico. *Renewable Energy*, 188:1141–1165, 2022.

[25] G. Weather. Master location identifier database (mlid) - standard version. Dec 2010. Documentation available at http://www.weathergraphics.com/identifiers/.