Analysis of Load Profile Generation Methods and Their Effect on the Results of Energy Optimization Models

Mohamad Wehbe

Thesis to obtain the Master of Science Degree in

Energy Engineering and Management

Supervisors: Prof. Hugo Gabriel Valente Morais
Dr. Eva Schischke

Examination Committee

Chairperson: Prof. Francisco Manuel da Silva Lemos

Supervisor: Prof. Hugo Gabriel Valente Morais

Member of the Committee: Prof. Luis Gomes

December 2023

I declare that this document is an original work of my own authorship and that it fulfils all the requirements of the Code of Conduct and Good Practices of the *Universidade de Lisboa*.

To my family and friends that were always supportive.

Acknowledgements

This Master's Thesis is dedicated to my family and friends who have been a pillar of strength and support during both challenging and rewarding times.

To my friends, both those miles away and those who made a foreign land feel like home, I deeply value your unwavering companionship. Over the past two years, during moments of adversity, your presence has been instrumental in fostering my resilience.

I would like to express my sincere gratitude to my supervisors, Eva Schichke, Sebastian Glombik, and Hugo Morais. Your guidance, constructive feedback, and encouragement have been invaluable in this journey.

My appreciation extends to my classmates in the SELECT and InnoEnergy program for making this a joyful ride and to the organizers of the program for providing us with such a great opportunity.

Abstract

This study presents an analysis of the effects of different load profile generation methods on the results of energy optimization models. Load profile generation methods are used to synthesize demand load data, which represents the energy consumption patterns of end-users over time. The study breaks down the analysis into two parts: First an analysis of load profile generators is done through various methods that aim to characterize and benchmark their performance. The second part analyses energy model optimization under different load profile generation methods as inputs. The study uses an open-source energy system modelling framework (OEMOF) to optimize a multi-energy system under various scenarios. The study evaluates the performance and robustness of each method, uncovers their underlying assumptions, and biases, and explains how they could affect the results of energy model optimization. The paper ends with some best practices for working with Load Profile Generators.

Keywords

Energy Optimization, Energy Modelling, Demand Profiles, Energy System Modelling Frameworks, Load Profile Generators

Resumo

Este estudo apresenta uma análise dos efeitos de diferentes métodos de geração de perfil de carga nos resultados de modelos de otimização energética. Os métodos de geração de perfil de carga são usados para sintetizar dados de carga de demanda, que representam os padrões de consumo de energia dos usuários finais ao longo do tempo. O estudo divide a análise em duas partes: primeiro, uma análise dos geradores de perfil de carga é realizada por meio de vários métodos que visam caracterizar e avaliar seu desempenho. A segunda parte analisa a otimização de modelos de energia sob diferentes métodos de geração de perfil de carga como entradas. O estudo utiliza uma estrutura de modelagem de sistema de energia de código aberto (OEMOF) para otimizar um sistema multienergia em vários cenários. O estudo avalia o desempenho e a robustez de cada método, revela suas suposições subjacentes e viéses, e explica como eles podem afetar os resultados da otimização de modelos de energia. O artigo conclui com algumas boas práticas para trabalhar com Geradores de Perfil de Carga.

Palavras-chave

Otimização energética, modelação energética, perfis de procura, estruturas de modelação de sistemas energéticos, geradores de perfis de carga

List of Figures

List of Tables

List of Abbreviations

ALPG	Artificial load profile generator
DSM	Demand side management
ESMFs	Energy system modelling frameworks
GEneSYS-MOD	Global energy system models
LP	Linear programming
LPG	Load profile generator
MILP	Mixed integer linear programming
NILM	Non-intrusive load monitoring
OEMOF	Open energy modelling framework
SLP	Standardized load profile

List of Software

Python	Python is a high-level, general-purpose programming language.
demandlib.	The first tested artificial load profile generator. Main use case is for household electricity and heat that is based on the python language.
Loadprofilegenerator	The second tested artificial load generator. It is based on C# that is used to generate load profiles for households for electricity, gas, hot water and cold water.
synPRO	The third artificial load profile generator. It can create electric, heat, water, and EV load profiles for residential and commercial use.
Districtgenerator	The fourth artificial load profile generator. It is python based and can produce electric and heat load profiles for residential households.
OEMOF	A modular open-source framework for modelling energy supply systems.

Chapter 1

Introduction

1.1 Background

An energy system encompasses the entire process chain, from the extraction of primary energy sources to the utilization of final energy, delivering services and goods to end-users [1]. To gain a deeper understanding of these intricate systems, they can be modelled, creating a computational representation of their physical and operational characteristics. This modelling allows for the simulation and analysis of system behaviour under various conditions, providing insights into system responses and potential areas of improvement. As the global demand for energy continues to rise, there's an increasing emphasis on using these models for optimization. Energy system optimization, given its inherent complexity, requires sophisticated strategies, advanced technologies, and a comprehensive understanding of system dynamics. Through optimization, the balance of efficiency, sustainability, and resilience can be achieved, ensuring that energy systems meet the evolving demands of our modern world.

There are many different Energy system modelling frameworks (ESMFs) available, each with its own assumptions, methods, and features. Some commonly used frameworks include Open Energy Modelling Framework (OEMOF), urbs, and Global Energy System Model (GENeSYS-MOD). A deeper dive on the model used for this study, OEMOF, will be conducted in the next chapter. To compare and evaluate these ESMFs, it is important to have transparent and open access to their source codes and documentation [2] Open-source energy system modelling frameworks (OS-ESMFs) enable such transparency and facilitate the assessment of their quality and suitability for different optimization techniques.

Assumptions play a pivotal role in energy system modelling and optimization. The accuracy and reliability of a model's output are intrinsically tied to the quality and validity of its input assumptions [3]. Among these inputs, demand load data stands out as particularly critical [4]. Demand load data represents the energy consumption patterns of end-users over time. For instance, underestimating peak demand can result in inadequate infrastructure planning, while overestimating can lead to oversizing and unnecessary costs and inefficiencies. Furthermore, the source and methodology used to derive this demand load data can introduce biases and uncertainties.

There are several methods to obtain this data. Direct measurements using smart meters or sensors provide real-time insights but might be limited in scope or raise privacy concerns. Deploying smart meters is also costly. Historical data, while valuable, might not always reflect future scenarios, especially in rapidly changing environments or in the face of significant policy or technological shifts. Another issue is the lack of availability of data and tools necessary to analyse the collected historical data[5]. A final method, surveys and statistical extrapolations, while useful, have notable limitations. Firstly, conducting and analysing surveys is time-intensive, often not aligning with the swift decision-making needs of energy planning. Secondly, biases can be introduced through survey design, question phrasing, and the administration process. Such biases can skew load data, impacting the accuracy of energy system models.

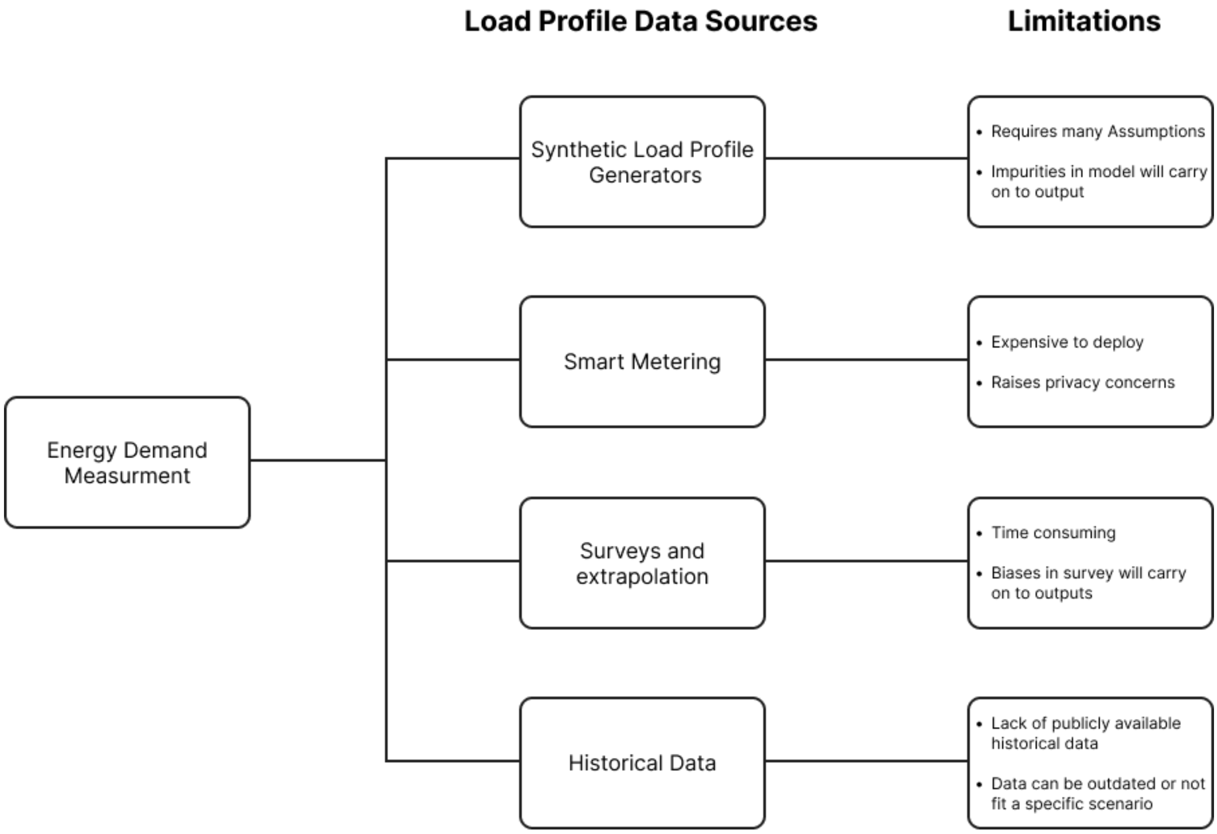

Figure 1.1 Energy demand measurement methods and their respective limitations.

Given these complexities in obtaining and validating load data, there's a growing interest in synthetic methods of data generation, leading us to the realm of Load Profile Generators (LPGs).

LPGs offer a more dynamic approach to generating demand load data. Instead of relying solely on historical records or direct measurements, LPGs synthesize load profiles based on a combination of factors, including user behaviour, appliance efficiency, weather patterns, and more. This synthetic generation allows for the creation of diverse and adaptable load profiles that can cater to a wide range of scenarios, from the introduction of new energy-efficient technologies to shifts in consumer behaviour due to policy changes.

However, the efficacy of an LPG is deeply rooted in the assumptions and methodologies it employs. Different LPGs might prioritize certain factors over others, leading to variations in the generated profiles. For example, an LPG focusing on residential loads might give more weight to behavioural patterns, while another targeting industrial loads might emphasize operational schedules and machinery efficiency.

The choice of an LPG and its underlying assumptions can have profound implications for energy system modelling. An accurate and representative load profile can lead to optimization results that are more aligned with real-world outcomes, ensuring that energy systems are designed and operated efficiently. On the other hand, relying on an LPG with misaligned assumptions can result in models that, while mathematically sound, might not reflect the true complexities and nuances of real-world energy consumption.

While LPGs present a promising solution to the challenges of obtaining accurate load data, it's imperative to approach their use with a critical eye, evaluating their methodologies and ensuring that the generated profiles truly resonate with the specific needs and realities of the energy system being modelled.

1.2 Motivation

The driving forces behind this research can be categorized into two primary aspects: internal and external motivations. The internal motivation delves into the technical and methodological challenges faced by professionals in the field, while the external motivation explores the broader societal, economic, and environmental implications of the study. Together, these motivations provide a holistic perspective on the significance and timeliness of the research undertaken.

For scientists and engineers dedicated to modelling and optimizing energy systems, the quality and accuracy of input data are of utmost importance. Precise energy demand data is a cornerstone for producing reliable and actionable results in energy system models. However, a persistent challenge is the lack of specific, granular demand data that mirrors specific real-world scenarios.

To bridge this data gap, LPGs have emerged as invaluable tools, creating synthetic demand profiles. But the use of LPGs introduces a crucial consideration: the assumptions made about the inputs that feed into these generators. Whether it's about user behaviour, appliance efficiency, climatic conditions, or other factors, these assumptions can significantly shape the synthetic profiles generated by LPGs. This generated synthetic load profile then might hold biases that are fed into the optimization model and may or may not influence the optimization results. The ripple effect of these input assumptions on the final energy system models and their optimization results is a critical area that isn't fully explored.

From an internal standpoint, the motivation is to delve into these input assumptions, understanding their nuances and implications. This research seeks to dissect the various inputs to LPGs, shedding light on how these foundational assumptions influence the outcomes of energy system models and their subsequent optimization.

Shifting our perspective to the external landscape, energy systems stand as the backbone of modern societies, fuelling industries, homes, and transportation. As global challenges like climate change, urbanization, and economic shifts intensify, the demand for efficient and sustainable energy systems grows ever more pressing. Accurate energy system modelling and optimization influence not just technical outcomes but also broader policy decisions, economic investments, and environmental impacts. Delving into the intricacies of Load Profile Generators and their input assumptions is more than a scientific pursuit; it's a commitment to ensuring our energy systems are resilient, sustainable, and attuned to the evolving demands of society.

1.3 Problem Statement and Objectives.

The modelling and optimization of energy systems play a pivotal role in shaping sustainable energy landscapes. Specifically, households within the European Union present a unique set of challenges and opportunities in this domain. The challenges mainly stem from a lack of load profile data availability for researchers [6]. Within this context, LPGs have emerged as crucial tools for synthesizing demand profiles. However, a significant gap persists in understanding how the assumptions behind these LPGs, and the choice of one LPG over another, influence the outcomes of energy models and their optimization results. One paper however, does explore the effects of data sampling frequency on the resultant load profile features [7]. Concentrating on a specific type of energy model, such as residential systems, allows for a more precise and in-depth analysis. Yet, the overarching challenge remains: Without a comprehensive grasp of these effects, there's a risk that energy system models might yield results that, while mathematically sound, may not align with real-world scenarios, potentially leading to suboptimal decisions in system design and operation.

Given the identified problem, this research aims to:

1. How do the different LPG tools compare in terms of their required input data, methodologies, outputs, and applicability?

2. How do input assumptions impact the synthetic demand profiles produced by LPGs, and how do these profiles compare to historical data?

3. What is the extent of the influence of LPG input assumptions on the results of energy system models, particularly in their optimization outcomes?

4. What are the best practices for utilizing LPGs in energy system modelling to ensure that the generated profiles closely mirror real-world scenarios? .

Chapter 2

State of the Art and Technology

2.1 LPG tools

LPGs are computational tools designed to simulate and generate electricity consumption patterns, commonly referred to as load profiles, for various types of consumers, ranging from individual households to entire communities or regions. These profiles represent the temporal variation of electrical load over a specified period. LPGs play a crucial role in energy system modelling, and demand-side management studies, as they provide insights into the daily routines, preferences, and energy consumption behaviours of end-users. Accurate representation of these consumption patterns is essential for optimizing energy distribution, ensuring grid stability, and designing effective energy-saving strategies. The development and application of LPGs have gained significant attention in recent years, especially with the increasing integration of renewable energy sources and the need for more flexible and responsive energy systems [4], [8].

2.1.1 Demandlib

Overview:

Demandlib [9] is a Python package that can create load profiles for electricity and heat demand based on standardized methods from BDEW (German Association of Energy and Water Industries) [10]. It can be used to model different sectors, such as households, businesses, industries, and agriculture, and different energy types, such as electricity, heat, and water. It can also generate customized profiles by adjusting parameters such as annual demand, building class, temperature, and holidays. Demandlib is based on python and can be installed through GitHub.

Table 2.1 SLP types from BDEW used in demandlib for generating electricity and heat load profiles.

type	description	explanation
G0	General trade/business/commerce	Weighted cluster of profiles G1-G6
G1	Business on weekdays 8 a.m. - 6 p.m.	e.g. offices, doctors’ surgeries, workshops, administrative facilities
G2	Businesses with heavy to predominant consumption in the evening hours	e.g. sports clubs, fitness studios, evening restaurants
G3	Continuous business	e.g. cold stores, pumps, sewage treatment plants
G4	Shop/barber shop	
G5	Bakery with bakery	
G6	Weekend operation	e.g. cinemas
G7	Mobile phone transmitter station	continuous band load profile
L0	General farms	Weighted cluster of profiles L1 and L2
L1	Farms with dairy farming/part-time livestock farming	
L2	Other farms	
H0/H0_dyn	Household/dynamic household	

Inputs:

To generate a load profile using demandlib, the minimum inputs required are the year of the demand series, the annual demand of the sector in kWh, and the type of the standardized load profile (SLP) according to BDEW classification [11]. The BDEW classification distinguishes between different sectors and energy types, such as households, businesses, industries, agriculture, heat, and electricity. Table 2.1 shows the SLP types and their descriptions.

There are some additional inputs that can enhance the accuracy of the results. The first one is the building class according to BDEW. The different classes are presented in table 2.2. The building classes are used to specify the building type and location, which affects the heat demand profiles. Note that they are different from the profile types which are presented in table 2.1 and are used to specify the energy type and sector, which affects both the heat and electrical demand profiles.

Table 2.2 Building classes from BDEW used in demandlib for generating heat demand profiles.

Abbreviation	Full Name	Abbreviation	Full Name	Abbreviation	Full Name
EFH	Single Family House	MFH	Multi Family House	GMK	Metal and Automotive
GHA	Retail and Wholesale	GKO	Local Authorities, Credit Institutions, and Insurance Companies	GBD	Other Operational Services
GGA	Restaurants	GBH	Accommodation	GWA	Laundries, Dry Cleaning
GGB	Horticulture	GBA	Bakery	GPD	Paper and Printing
GMF	Household-like Business Enterprises	GHD	Total Load Profile Business/Commerce/Services		

More additional inputs include the time series of the temperature data, the wind class of the building location, the holidays of the region, and the seasons of the region. demandlib has built-in holiday data for Germany that can be used if no other holiday data is provided. The temperature data can be obtained from external sources such as weather databases or APIs. The building class and the wind class are parameters that affect the heat demand profiles. The seasons are used to define the summer, winter, and transition periods for different regions.

Algorithm:

Demandlib [12] employs a variety of methods and algorithms tailored to the specific sector and energy type for generating load profiles.

For electrical profiles, Demandlib follows a different methodology. These profiles are based on BDEW standards and represent electricity consumption patterns in Germany. The BDEW profiles used have a time resolution of 15 mins. The generation of these profiles involves factors dependent on the SLP type, the day, and the hour, with data sourced from BDEW datasets. For household profiles, demandlib includes a dynamization function to account for seasonal variations in electricity consumption.

Demandlib operates on certain assumptions based on BDEW methods. It assumes that the generated profiles are representative of consumption patterns in Germany. These profiles are based on historical measurements and may not reflect future consumption trends. The profiles are normalized to an annual demand, requiring users to scale them for specific scenarios.

Outputs:

Demandlib produces a pandas DataFrame that represents the hourly load profile in MWh/h for the specified sector and energy type.

For electrical profiles, the output contexts are diverse. They encompass residential households, commercial entities such as shops, bakeries, and cinemas, agricultural operations like dairy farming, and specific industrial units like mobile phone transmitter stations.

The energy type for the output can be either heat or electricity. Demandlib has conceptualized water profiles, but they are not currently implemented.

The default granularity of the output is hourly. However, users can adjust this granularity using pandas methods for resampling or aggregation to suit their needs.

2.1.2 Districtgenerator

Overview:

Districtgenerator is another open-source python-based tool on GitHub that is under current development [13]. Its primary purpose is to generate building-specific thermal, electrical, and occupancy profiles for residential districts. It integrates two open-source tools, TEASER [14], and Richardson.py [15]. It classifies buildings using the TABULA archetype approach which is a methodology for classifying residential buildings across Europe based on their energy-related characteristics [16].

Inputs:

The following are the data requirements for generating results:

Geographical and Climatic Data: This data encompasses essential geographical and climatic specifics about the district's location, which are pivotal as the energy demand of a district can significantly vary based on its geographical position and prevailing climatic conditions. The parameters include:

- Location: Denotes the city or region of the district. The tool primarily caters to sites within Germany. The tool utilizes the TABULA approach, which offers both a national definition tailored to specific countries and a common definition for cross-country comparisons.
- Climate Zone: Each city or region is associated with a specific climate zone, which provides a standardized categorization based on the region's typical weather patterns, influencing the energy demand patterns.
- Altitude: Represents the height above sea level for the specified location. Altitude can influence temperature and, consequently, heating, or cooling demands.

Temporal Data: Temporal data allows users to define the granularity of the generated load profiles. The granularity is pivotal as it determines the resolution at which energy demands are captured. The primary parameter here is:

- Time Resolution: Users can set the granularity for the load profiles. For instance, a resolution can capture energy demand patterns every 15 minutes or at broader intervals.

Building-specific Data: This data offers insights into the specific characteristics of the buildings within the district, and it is based on TABULA archetype approach. Information about the parameters can be obtained from the TABULA website. The parameters are:

- Building Types: Categorizes buildings based on their structure and usage, such as Single-Family House or Multi-Family House.
- Construction Year: Indicates the year when the building was constructed, which can influence its energy efficiency and demand patterns.
- Retrofit Level: Specifies if the building has undergone any energy efficiency retrofitting.
- Living Space Size: Represents the total area of the building, influencing its energy demand.

Algorithm:

Building Typology with TABULA: The districtgenerator uses the TABULA typology to define building characteristics. This typology provides a structured approach to categorize buildings based on various attributes such as construction year, building type, and heating system. The tool then uses this typology to generate representative load profiles for different building categories.

Integration with TEASER: TEASER, an open framework for urban energy modelling, plays a pivotal role in the districtgenerator's algorithm. TEASER's primary function within the districtgenerator is to provide a detailed thermal model of the buildings. [17]. This network is then used to simulate the building's thermal behavior in response to external conditions, such as weather variations.

Load Profile Generation with richardson.py: The richardson.py tool is designed to generate electrical and lighting load profiles as well as occupancy based on statistical data and typical consumption patterns [18]. Its methodology is rooted in analyzing historical consumption data, identifying patterns, and then using these patterns to generate representative load profiles for different scenarios. Within the districtgenerator, the richardson.py methods are employed to produce electrical load profiles that are representative of typical consumption behaviors, considering daily and seasonal variations.

Customization and Adjustments: The districtgenerator allows for adjustments and customizations in the load profile generation process. Users can modify specific parameters, such as indoor temperature settings, to influence the resulting load profiles. Additionally, the tool provides options to adjust the granularity of the output, allowing users to obtain profiles at different time resolutions.

Outputs:

After executing district generation, users will obtain building-specific profiles in the .csv format. The results include profiles for space heating demand, domestic hot water demand, electricity demand for lighting and electric household devices, number of persons present, and internal gains from persons, lighting, and electric household devices. All values are provided in Watts and adhere to the required time resolution.

2.1.3 LoadProfileGenerator

Overview:

The LoadProfileGenerator [19] is a specialized tool designed for the creation of load profiles detailing energy and water consumption in households. The tool's foundation is built upon a behaviour model which posits that individuals are driven by intrinsic desires, influencing their daily activities and, consequently, their consumption patterns [20].

LoadProfileGenerator offers a high degree of flexibility and customization. It provides modellers with prebuilt templates for various household elements, while also allowing the creation of custom elements from scratch. This adaptability ensures that the tool can cater to diverse scenarios, making it suitable for modelling households across different regions or contexts.

The tool has a free user-friendly interface, extensive documentation, and a website with information on the tool’s usage. The source code is also available on GitHub and written in C#.

Inputs:

Modelling of the households in LoadProfileGenerator requires modelling different elements within each household. These elements are required to create the decision process, which is then used to calculate the load profiles. Figure 2.1 shows the minimum required elements for modelling a decision process that will lead to the creation of a load profile.

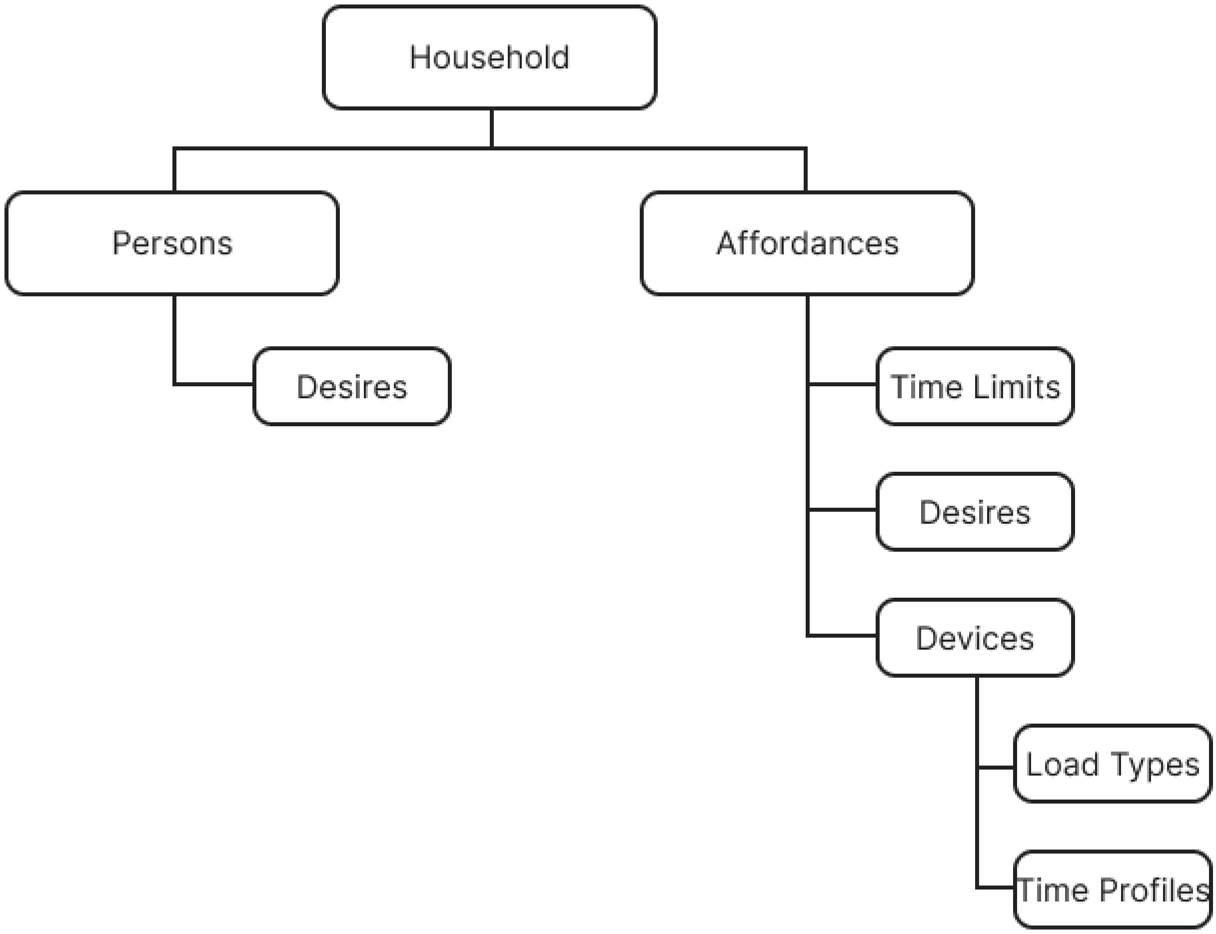

Figure 2.1 Minimum required elements for modeling a decision process

Desires: Desires form the foundational element for modelling a household. They represent the intrinsic needs or wants of the household members. These desires drive their daily activities and behaviours, influencing their energy and water consumption patterns. For instance, the desire to eat might prompt cooking activities, while the desire for comfort might lead to adjusting the thermostat.

Person: This aspect delves into the unique characteristics or profiles of individual members of the household. Each person's behaviour, preferences, daily routines, and even their profession or age can influence the household's overall consumption patterns. For example, a person working from home might have different energy consumption patterns compared to someone who goes to an office.

Load Types: Load types refer to different categories of energy or water consumption within a household. This could encompass various needs such as electricity, lighting, heating, cooling, or appliance usage. Each load type would have its distinct consumption pattern and would be influenced by the activities of the household members.

Devices: This section focuses on the myriad devices or appliances within a household. From kitchen appliances like ovens and refrigerators to entertainment systems and water heaters, each device contributes uniquely to the household's energy, heating, cooling, and water footprint. The frequency, duration, and manner of device usage play a crucial role in shaping the household's load profile.

Time Profile: The time profile captures the temporal patterns of consumption. It provides insights into when certain devices are predominantly used or when specific activities, like cooking, bathing, or laundry, typically occur. This temporal mapping helps in understanding peak consumption times and potential energy-saving opportunities.

Time Limits: Time limits set specific boundaries or restrictions on when certain activities can or cannot happen. These constraints reflect real-world scenarios and habits. For instance, certain noisy activities might be restricted during nighttime hours, or specific devices might be programmed to operate during off-peak times.

Affordances: In the context of household modelling, affordances refer to the potential actions, behaviours, or activities that the household environment permits or encourages. It's about understanding what's possible within the given setup of the household. For example, a household with a home gym might have exercise-related energy consumption, while one with a garden might have specific water usage patterns.

By considering such a comprehensive range of inputs, the LoadProfileGenerator offers a detailed and nuanced approach to modelling individual households, aiming to create accurate and representative load profiles.

Algorithms:

The interaction between the elements in the LoadProfileGenerator's behavioural model is intricate. The desires of the simulated inhabitants influence their choice of activities. These activities are then linked to specific devices in the household that can fulfil them. For example:

- An inhabitant might have a desire for entertainment.
- This desire could lead them to consider watching TV as an activity.
- The TV, as a device, offers the activity of watching, which the inhabitant might choose based on their current desires and other influencing factors.

The model also considers external factors, such as time constraints, to further refine activity selection. For instance, certain activities might only be available or preferred during specific times of the day.

LoadProfileGenerator tool is also based on richardson.py and makes several foundational assumptions that influence the generated load profiles. One of the primary assumptions is the needs-based approach to human behaviour. The tool assumes that human actions within a household are primarily driven by their needs or desires, such as hunger or the need for entertainment. This approach might not capture all nuances of human behaviour but provides a structured way to simulate typical household activities. Another assumption is the deterministic order in which people choose activities. While this might not always reflect real-life randomness, it provides a consistent basis for simulation. Furthermore, the tool assumes that all affordances (potential activities) are always bound to devices, even in cases where this might not make intuitive sense. For instance, an activity like a weekend trip might not be directly tied to a household device but is still modelled within the tool's framework.

Outputs:

The time resolutions of the internal calculation and the result files are independent of each other, and both are freely configurable. Internally, the simulation is performed with 1 min or better. For the output files, any time resolution can be selected, if it is an integer multiple of the internal resolution. If desired by the user, extensive evaluations can be generated, which evaluate the generated load profile according to a variety of criteria. Output files are generated for both the overall household profile and the individual appliance profiles. The output files include csv files for household and appliance energy expenditure, generated charts, and a generated report.

Validation:

LoadProfileGenerator validates the generated load profiles with two different approaches. The first approach focused on modelling and analysing the results of a single household. The second approach modelled an entire community and compared the results with literature:

1. Single Household Analysis:

   Objective: Compare a single household's specific characteristics with plausibility criteria and characteristics derived from literature.

   Methods of comparison:

   Activity Analysis: Examine the generated activities of each resident within the household. This provides insights into the daily routines and preferences of the household members.

   Comparison with Reference Values: Characteristics of the generated profiles are compared to established reference values. Two primary references used are:

   - Wo Bleibt Die Zeit [21]: This literature analyses how the German population utilizes their time, providing a benchmark for activity-based energy consumption [21].
   - Erhebung Wo im Haushalt bleibt der Strom: This source offers energy usage statistics based on household size, serving as a comparative metric for energy consumption patterns[22].

Daily Consumption Analysis: The average daily consumption of each household is analysed and compared to literature values of expected daily consumption for similar households.

Energy Consumption Share Analysis: The share of energy consumption (autonomous vs. behavioural, time spent per activity) is analysed and compared between the generated profiles and literature values.

Results:

Activity Analysis: The grid diagrams and visualizations reveal patterns related to various activities, such as sleeping, working, leisure activities, and more. These patterns provide insights into the daily routines and preferences of the household members.

Comparison with Reference Values: The specific household with the selected appliances is slightly below average in terms of energy consumption. This is attributed to the consistent use of energy-saving devices, such as LED light bulbs, set under "Energy saving devices."

Daily Consumption Analysis: There is no significantly increased evening peak in winter, as might be expected from the sinusoidal curve of the H0 profile. This discrepancy is attributed to the modelling of only a household (not an entire house) and the use of LEDs for lighting. Also, since the entire house infrastructure, such as circulation pumps, electric water heating, and heating systems, is not included in the simulation, this affects the overall energy consumption patterns.

Energy Consumption Share Analysis: The integration of a photovoltaic system in the modelled household affects the load. The PV profile was scaled so that the annual electricity generation corresponds to the electricity demand of the household. Initial impressions suggest that a significant portion of the energy consumption can be covered by the PV system. However, detailed calculations reveal that the grid load is significantly larger than initially assumed. The annual values indicate that about 1/3 self-consumption can be achieved, which aligns with literature values. This serves as an indication of the plausibility of the generated load profile. The modelling of electricity consumption for light appears to be accurate. However, there's a noted weakness in the modelling, where the light is always switched off when leaving the room, leading to modelling artifacts. The effects of this are minimal due to the low power consumption for lighting.

2. Community-Wide Analysis:

   Objective: Compare the load profile of an entire community to community load profiles sourced from literature.

Methods of comparison:

Peak Power comparison: The peak power values for communities of varying sizes are compared with values from literature. This helps in understanding the scalability and accuracy of the generated profiles.

load profile shape analysis: The overall shape and pattern of the generated load profile are compared to established literature profiles, such as the H0 profile. This ensures that the generated profiles mimic real-world energy consumption patterns at a community level.

Results:

Peak Power Comparison: The curve of the 100 randomly generated households is observed to be partly above the reference curve. In contrast, the curve representing only the predefined households is just below the reference curve. The similarity with the reference is evident, suggesting that the generated profiles closely match real-world data. Overall, the peak load, depending on the parameters of the settlement being generated, is either just below or just above the simultaneity curve. This indicates that the generated profiles represent reality well.

Load Profile Shape Analysis: The self-consumption and grid load/feed-in were calculated for each minute of the year. The grid load appears significantly larger than initially assumed. The annual values align with literature values, suggesting that about 1/3 self-consumption can be achieved. This serves as an indication of the plausibility of the generated load profile.

2.1.4 synPRO

Overview:

synPRO [23] is a tool designed to generate electric load profiles of a single household, a commercial zone, or a community with multiple commercial and residential zones [24]. The loads have a minimum time resolution of 1 minute, specifically tailored for German households. The tool uses a bottom-up stochastic model to model electric, heating, domestic hot water, and electric car load profiles. To simulate a load profile users can use the tool online. Unlike the other tools that were covered, the tool's source code is not published online.

Inputs:

The following is the required data for generating load profiles:

Weather dataset: there are two available datasets here. The first is ERA5 [25] which is an hourly historical real measurement dataset. The second dataset is TRY[26]. This dataset is based on the periods 1988-2007 and contains 15 different zones within Germany. Uploading another weather dataset to be used in the simulation is currently not possible.

Location: The name of the city is required to determine the public holidays and the proper weather file within the selected dataset.

Simulation year: Only used to select the appropriate weather dataset file and create the appropriate time index.

Aggregation levels: 4 different aggregations are available. At the neighbourhood level, the data provided is limited to aggregated time series for three distinct profile types: electricity, heating, and tap water heating. Progressing to the building level, alongside the aggregated district data, time series for the three profile types are also stored specific to individual buildings. In the third aggregation option, the zone level, electrical time series are recorded for individual commercial main zones or apartments, though it's noteworthy that thermal time series are exclusively available for select buildings. Finally, at the device or sub-zone level, the data provided becomes even more detailed. Here, in addition to the information available at the building and zone levels, electrical time series for individual devices within residential settings or specific sub-zones within commercial zones are provided.

Net usable area: This usable area of each building is required for thermal simulation of any building and for electrical simulation of the commercial zones. It does not affect electrical simulation of residential households.

Share of each zone: The total area share of each zone needs to be specified.

Commercial zone details: If a commercial zone is included, for each zone several details need to be specified. The first detail is the zone type according to either the SIA [27] or the TEK [28] standard. Next, the light, device, and ventilation efficiency need to be specified (old, standard, or efficient).

Residential zone details: If a residential zone is included, for each zone several details need to be specified. The Socio-economic factor of the house needs to be chosen from a list. This specifies the time spent on each activity and is based on the Harmonised European Time of Use Surveys (HETUS) [29]. Next the number of residents needs to be specified. There is a maximum limit of four residents. The efficiency of devices needs to be specified (old, standard, efficient). Finally, the devices available in the house needs to be selected. If this information is unavailable, there is an option for a randomly composed set of devices based on statistical socio-economic assumptions.

Algorithm:

The model's foundation is rooted in the understanding that daily routines and appliance usage in households are influenced by socio-economic factors such as family status, working patterns, age, housing type, and family situation. To capture this diversity, the model distinguishes between different household classes. In contrast to LoadProfileGenerator which relies on richardson.py to model behaviour, synPRO uses a stochastic approach based on sampling from probability distribution.

Like LoadProfileGenerator, SynPRO assumes that domestic electricity use is caused by the operation of technical appliances and categorizes them as user dependant and use independent. To model the appliance operation (load trace) the tool uses data appliance data sourced from the German segment of the Harmonized European Time Use Survey (G-HETUS).

Outputs:

The output files are csv files and can have a temporal resolution of 1 minute, 15minutes, or 1 hour. The files vary according to the selected aggregation levels.

Validation:

The paper tries to validate the tools results by comparing simulations belonging to houses with different socio-economic features to a dataset of 430 smart-meter measured load profiles from project Intelliekon. The households from Intelliekon [30] were measured between December 2009 and November 2010.

The results of the validation are as follows:

The annual demand of the simulated and real household data are very similar. The difference in the yearly mean electricity demand across all groups are below 5.2%.

When comparing monthly demand, the measured data shows that the highest variation between any simulated and real household monthly demand is 13.2%. SynPRO slightly underestimates the electricity usage in summer and winter, and in slightly overestimates the usage in fall and spring.

When daily demand is compared, the relative error is between 8.8% and 16.1%. For the daily demand, a correlation analysis was done and shows a correlation between 0.9 and 0.98 for all cases.

For hourly demand, the comparison shows a lowest relative error of 5.8% and a highest relative error of 16.9%. A duration curve shows that the load profiles compared are almost identical. In particular,

2.2 The OEMOF Model

Open Energy Modelling Framework (OEMOF) [31] is an optimization model based on linear programming (LP) and mixed integer linear programming (MILP). It can be used to build an energy system model. The tool is open-source [32] and based on the python coding language. When building OEMOF, the goal was to create a flexible open-source tool that is capable of mathematically mapping the multidimensional interdependencies within the energy industry [33]. One thing that stands out in OEMOF is the great variety of different models that can be built and research questions that can be answered.

2.2.1 Creating a Model

The tool can be used as a model generator. It contains a model library with different energy components or constituents (e.g., a demand source like a house, or a storage system like a utility scale battery) that can be combined to create a model. The structure of the models is modular meaning that each component corresponds to one module that is described by a system of linear or mixed integer linear equations. Furthermore, each type of component has a set of parameters that describe it.

The foundational components of the generic modelling approach in OEMOF are illustrated in figure 2.2. At its core there are two integral levels: Generic classes and graph theory.

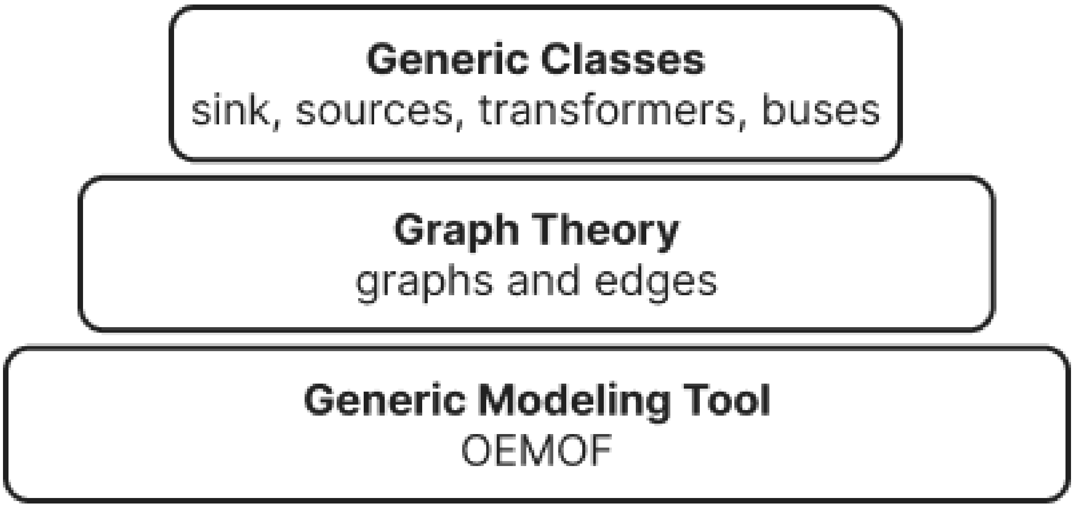

Figure 2.2 The Generic Modeling approach of OEMOF

On the surface, the different components in the energy library are divided into sets of generic classes. The following classes are defined:

- Source: A source is a component that can provide energy (has one output). For example, a source component can be used to model a wind farm, diesel engine, or electricity imports.
- Sink: A sink is a component that consumes energy (has one input). For example, an electric vehicle, a household, or a whole city can be modelled using a sink component.
- Transformer: A component that converts the energy type or value. For example, an electrolyser that “transforms” electricity to hydrogen or a transmission line that introduces losses to the total

energy transmitted. A transformer can have many inputs and outputs. The inputs and outputs values will be different if the conversion parameter is not zero.

- Bus: A component that can have multiple inputs and outputs. The input and output values must stay balanced at every timestep.

Some components in OEMOF's model library do not belong to any of the mentioned generic classes. One example is a general storage system.

Modelling in OEMOF is based on graph theory. An OEMOF model is represented by edges (also called flows) and nodes. A node is any modelled component (source, sink, transformer, bus). It is a very flexible element that can represent any component within an energy system. An edge maps the movement or flow of energy within the model.

Let us go through an example of a simple model that can be built in OEMOF. We can take an example of a household that has a solar panel and is connected to the grid. Let us assume that we want to understand how much electricity will this house consume from the grid, and what solar capacity is required if we are trying to minimize financial cost. To build the model a source can be modelled as the solar panel. The source's parameters could be a time-series representing the unit energy the panels produce over a year. The nominal value which dictates the size of the solar panel would be the unknown variable. To model the grid imports, we can use another source component but use different parameters. this time we can use variable cost as a parameter to represent the cost of energy at a specific time period. to model the household, we can use a sink. The sink will have a time-series and a nominal value to represent its annual load profile. To bring the model together we can add an electric bus and edges as shown in figure 2.3.

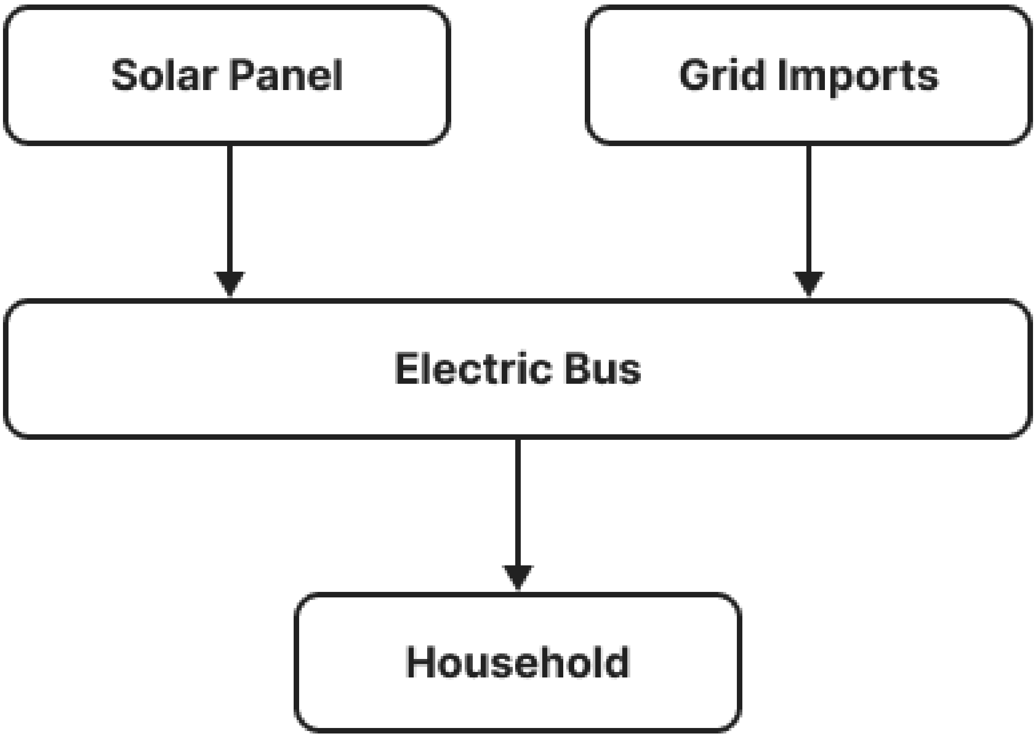

Figure 2.3 Edge and flow figure of generic household optimization problem.

What is left now is to optimize the model and find the amount of energy to consume from the appropriate solar panel size and hence the solar and grid imports for each time-step. The next section dives deeper into OEMOF's optimization methodology. At the end of the section, the simple example will be revisited to clarify how this model could be optimised for cheaper electricity consumption.

2.2.2 Optimizing a Model

Once the model is built, it can then be optimized using (MI)LP such as cbc solver or gurobi. Mathematical optimization problems can in general be described in the following forms:

Linear Programming

1. Objective Function: Linear programming involves optimizing a linear objective function which can be represented as: $Z = c_1x_1 + c_2x_2 + \cdots + c_nx_n$ where Z is the objective function to be maximized or minimized, c_I are the coefficients, and x_i the decision variables.
2. Constraints: The decision variables are subject to a set of linear constraints that can be written as:

$$Z = c_1x_1 + c_2x_2 + \cdots + c_nx_n$$

$$a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \leq b_1$$

$$a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \leq b_1$$

$$\vdots$$

$$a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \leq b_m$$

where a_ij are the coefficients of the constraints, and b_i are the constants on the right-hand side of the inequalities.

Mixed Integer Linear Programming (MILP):

1. Objective Function: Just like LP, MILP has a linear objective function.
2. Constraints: The constraints are also linear, similar to LP.
3. Integer and Continuous Variables: Unlike LP where all decision variables are continuous, in MILP, some or all the decision variables are required to take integer values. Let x_i be integer for $i \in I$ where I is the set of indices for which x_i are integer.

We call a solution feasible if it satisfies all the constraints. The optimal solution is the feasible solution that optimizes the objective function.

OEMOF can determine the lowest feasible solution for the objective function that has been created through modelling. The objective function native to OEMOF is the minimization of either financial or environmental costs. A generic OEMOF optimization equation is presented. this equation represents the objective function at a specific timestep where $flow$ represents the flow of energy or cost and c represents a weighting factor such as cost of emission.

$$y = \sum_t flow1_t \cdot c1_t + flow2_t \cdot c2_t + \cdots \rightarrow \text{Min.}$$

The two types of models that can be created and that lead to different optimization processes are a dispatch and an investment model. If a component is modelled as a dispatch model, the function optimizes its operation (operation optimisation). If a component is modelled as an investment model, the function optimizes its cost (design optimization).

In our household example we need is to understand how much electricity imports we require from the grid, which is an operation optimization problem. we also need to decide the size of the solar panel, which is a design optimization problem. For this example, the following objective function describes this model:

$$min\colon \sum_{pv} x_{\text{pv}}^{\text{capacity}} \cdot c_{\text{pv}}^{\text{capacity cost}} + \sum_{g,t} x_{\text{g}}^{\text{imports}}(t) \cdot c_{\text{g}}^{\text{marginal cost}}$$

The first part of the equation represents the cost of the solar panels, the second part of the equation represents the cost of grid imports.

In OEMOF, some constraints are already included in the framework (for example, no negative energy production) while others can be set by parameters (for example, the max capacity of a solar panels on a roof). Our modelled example can now be solved with the help of two tools. A mathematical model formulation tool called pyomo, and an optimizer, cbc solver or gurobi. we will dive deeper into the programming aspects of OEMOF in the next section.

ESyOpT

ESyOpT is a package that introduces economic and technological context around basic OEMOF components [34]. It introduces more input parameters to each component and contributes more elements such as investment costs of a storage system or maximum power of a solar panel to the objective function and constraints.

Chapter 3

Methodology

3.1 Research Philosophy

The problem being solved is the uncertainty that energy modelling experts face concerning the effects of their assumptions on the overall results of their study. To shed some light on this issue this study aims to understand how the underlying assumptions of energy system modelling from the load profile perspective affects the result of model optimization. As stated in the objective, it is important to understand any biases introduced by the LPG tools that show up in the load profiles and how these biases propagate to the optimization results. Once understood, some clarifications and recommendations can be made to help researchers better understand the impact of their load profile modelling assumptions on the overall optimization results.

Two testing methods are available. A black box method, and a white box method. In the context of this study a white box method would involve analysing the internal workings, algorithms, and processes of the LPG tools and the modelling framework. This would entail a deep dive into the mathematical formulations, algorithms, and data structures used by these tools. On the other hand, the black box method involves treating the LPG tools and optimization models as a closed system, where the focus is primarily on the inputs and outputs without considering the internal processes. By providing a set of standardized inputs and analysing the resulting load profiles and optimization results, we can infer the biases and assumptions made by the tool. For the purpose of this study the black box approach will be pursued since this method is less intensive and instead, it relies on empirical testing and observational analysis. By concentrating on observable outcomes, the black box method emphasizes empirical results. This can lead to more tangible and actionable insights, especially when the goal is to understand the impact of assumptions on optimization results.

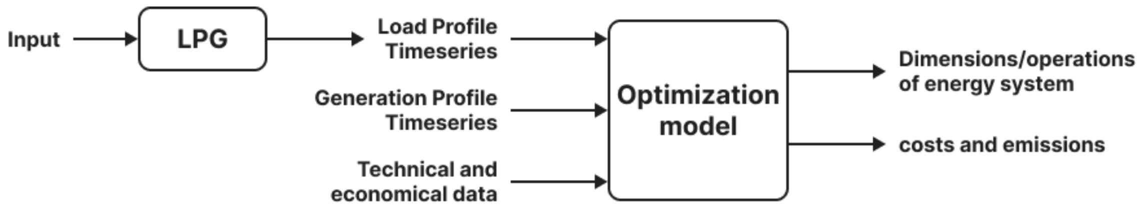

Figure 3.1 Process flow of energy system optimization model with focus on load profile timeseries simulation using LPG.

For this study, which aims to discern the effects of energy system modelling assumptions on optimization results, an inductive reasoning approach is most fitting. This approach, rooted in specific observations, aligns with the study's exploratory nature and its empirical emphasis, especially given the use of the black box method. Inductive reasoning allows the research to detect patterns and draw broader generalizations based on these observations, offering flexibility and the potential to develop new theories grounded in the data. In contrast to deductive reasoning, which tests a predefined hypothesis against observations, inductive reasoning starts from specific cases and moves towards broader insights, making it more suitable for this study's objectives. After we clarified the philosophy, we will move into describing the research design and justifying the decisions that were taken.

Having established the philosophical underpinnings of our study, it's crucial to translate this foundation into a practical and coherent research design. This design will serve as the blueprint for our investigation, ensuring that our approach is systematic and aligned with our objectives. As we delve into the research design, we'll outline the specific methods and strategies that will guide our exploration.

3.2 Research Design

The study focuses on modelling energy systems in the residential household sector. The objectives are as follows:

1. How do different LPG tools compare in terms of their required input data, methodologies, outputs, and applicability?

2. How do input assumptions influence the synthetic demand profiles produced by LPGs, and how do these profiles align with historical data?

3. To what extent do LPG input assumptions affect the results of energy system models, especially in their optimization outcomes?

4. What are the best practices for employing LPGs in energy system modelling to ensure the produced profiles closely reflect real-world scenarios?

Objective 1 was addressed in chapter 2. To address the remaining objectives, the study is structured into two primary sections.

The initial section centres on the second objective. This involves an analysis of simulated load profiles. Each tool is employed to craft a simulated load profile that closely matches historical data from various sources. Once established, these profiles are compared with the historical datasets to compare accuracy. However, validating the accuracy of a household load profile is complex. A direct comparison of generated load profiles with real-time household data for each timestep might not provide significant insights. Even with similar modelled activities, slight deviations in the timing of appliance use can lead to discrepancies between the profile and the real data. Merely contrasting the annual or monthly totals also doesn't provide a comprehensive understanding of the simulated load profile's precision [20]. Thus, various visual and statistical techniques are employed to assess load profile accuracy, with some techniques being adaptations from diverse studies. The analysis is segmented into two subsections each with its own methods:

- Analysis of load profiles
    - Visual/graphic analysis
        - Load profile characterisation method

        - Relative heat maps method
    - Statistical analysis
        - Annual energy demand
        - Monthly energy demand

In the initial subsection load profiles undergo two different methods of visual comparison. The first method of visual analysis, the load profile characterisation method, focuses on clustering and comparing load profiles according to some studied parameters. This visual analysis approach predominantly stems from a study on electric load profile scrutiny from the Lawrence Berkley National Laboratory [35] and another exploring patterns and variability of electric load profiles from the same institution [36].

For the load profile characterisation method, the parameters that were used to characterize the load profile are listed in table 3.1. Before diving into the parameters, it is important to note that the data studied was aggregated on an hourly basis, and clustered by season, and weekdays/weekends. The clustering function used can be found in Annexe A.

Table 3.1 Parameters studied in the visual analysis.

Parameter	Description	Source
Near-Peak Load	highest hourly load of daily profile	[35]
Near-base Load	Lowest hourly load of daily profile	[35]
Rise start time	The latest time in the morning when the load is less than: base load + 0.05 * (peak load - base load)	[36]
Fall finish time	The earliest time in the afternoon when the load is less than: base load + 0.05 * (peak load – base load)	[36]

When calculating the near-peak load a 97.5 percentile of the daily load was considered in the Price, 2010 [35] study. This reason was to avoid outliers and to use a more stable value that is representative of the daily peak loads. This is appropriate since the study used a 15-min granularity. However, since this study dealt with hourly granularity, the near-peak load was taken as the highest hourly load of the daily profile. This method still avoids outliers and uses stable values since the values are averaged hourly values, grouped by season (an average value of 91 days). This averaging and grouping filters out the outliers and absolute peak values and provides a true near-peak value. To test the validity of this assumption, max hourly values were compared to average values and the average values were found to be reasonably smaller than the max values. A similar approach was adopted for the near-base load. Two more parameters outlined in table 3.1 were picked for the analysis to understand the rise start time and fall finish time as well as the steepness of the rise/fall to get an idea of how much time is spent at high and low loads.

The second visual analysis method, the relative heatmaps method, is used as a way to visualize the load profile difference between the actual and simulated dataset. The heatmap is created by subtracting the timesteps of the actual and simulated load profiles. As mentioned earlier, a direct statistical comparison between each actual and simulated timestep will not lead to useful results, nevertheless, visualising the difference can show in which general time of the year or day the simulated load profile overestimated or underestimated energy demands.

The visual analysis was conducted with the objective of characterizing the load shape. Not only was this approach useful for the second objective, but also, it was used to understand objective three. Through this characterization, distinctions in load characteristics were drawn. Once the load characteristics were established, the results of the optimization model were subsequently compared. Any disparities observed between the simulated and actual load profiles were then traced back to these load characteristics. Consequently, a comprehensive comparison of the load profile characterization was undertaken to identify potential biases present within the load profile.

In the second subsection to answer objective 2 a statistical analysis was carried out. The statistical analysis also focused on indicators that were used in study papers with the aim of validating synthetic load profiles. [20], [24], [37].

The following indicators will be studied:

- Yearly energy demand
- Monthly energy demand
- Frequency analysis

The yearly and monthly energy demand can give a comparative overview of the synthetic demand profiles generated by LPGs in relation to historical data. By comparing the yearly and monthly figures between synthetic and actual demand profiles, we can identify the efficacy of the LPGs in replicating real-world demand patterns across different timescales.

The frequency analysis, in this context, focuses on comparing the occurrence rates of specific power levels throughout the hours of the year. It provides insights into how often certain demand levels appear in both synthetic and historical datasets. This analysis is crucial because, while average values might align, the distribution of power levels and their frequencies can offer a more granular understanding of the profiles' similarities and differences. For instance, if certain power levels recur frequently in historical data but are underrepresented in synthetic profiles, it might indicate areas where the LPGs could be refined for better accuracy.

Section 4.2 is dedicated to solving objective 3 which aims to understand the effects that the simulated load profiles had on the optimization results. In accordance with the black box and inductive approach, a common energy model was built and fed both the real and simulated load profiles. More details about the energy model can be found in section 3.2.3. The model created contains two different optimization problems. A design optimization problem and an operation optimization problem. By building a model that solves these problems, a comprehensive approach to understanding the influence of LPG input assumptions on various facets of energy system modelling was achieved.

The design optimization problem focuses on the configuration and sizing of different energy system components. The key metric here is total system cost. This helps to determine the optimal infrastructure needed to cater to the energy demand represented by the load profiles. By feeding both real and simulated profiles into this model, we can assess whether different demand patterns lead to different system configurations or sizes.

On the other hand, the operation optimization problem deals with the day-to-day functioning of the energy system. It evaluates the best operational strategies, like when to store energy or when to release it, based on the demand patterns. By comparing the operational strategies derived from real versus simulated load profiles, we can understand the nuances of daily operations influenced by the LPG assumptions.

Using this approach, an in-depth understanding was gained regarding how synthetic load profiles, influenced by LPG assumptions, are impacted in terms of design and operational decisions within energy system models. Through rigorous testing and comparative analysis, the importance of precise demand representation in achieving optimal and efficient energy system modelling was highlighted.

3.2.1 Data Collection and tools

The data used across the whole study was consistent. To simulate the load profile four tools that were described in chapter 2 were used. These tools were used due to their full or partial focus on load profile generation of residential households in Germany and for the purpose of use in energy system modelling. Three more notable publicly available load profile generator tools that focus on residential context exist that have not been studied in this paper. The first one is the Artificial Load Profile Generator (ALPG). This tool was not used since its main purpose is to benchmark different demand side management (DSM) approaches. Because of this, the tool only produces static load profile components and flexible load profile components such as washing machines are not simulated to allow for DSM algorithms to decide on the scheduling of these flexible devices [38]. The second tool ANTgen was designed for use with non-intrusive load monitoring (NILM) tools. Due to the scope of this study the tool was not included. The final unused LPG was RAMP[39].

For the purpose of this study, the simulated load profiles were compared against metered load profiles. A total of 16 publicly available residential building datasets are available online [5]. Since all four LPG tools are focused on German residential household the datasets that were picked had to be in Germany or a location were residents had similar energy use behaviour. Furthermore, since two LPG tools, LoadProfileGenerator and synPRO require a full or partial information of the used appliances for modelling, and all four tools require information about the household residents, one of the selected datasets need to have metadata that provide this information so the effects of these variables can be analysed.

The first dataset that was used is UKdale. This dataset provides building and appliance-level metered electricity from five households in the UK. The sampling resolution is six seconds. It was picked since there was a lack of appliance level metered residential load profiles in Germany. The UK also has a similar GDP per capita to Germany [40]. The annual household consumption in the UK and Germany are also similar[41].

From this dataset, house 1’s 2016 data was picked since it’s the most recent clean full year data in the dataset. The house was built in 1905 but had several energy improvements made. It contains a solar-thermal system. Heating uses natural gas. The house contains four occupants, two adults and two children. The appliances available in the house are listed in table 3.2.

Table 3.2 UKdale's house 1 appliances

Channel Number	Appliance	Channel Number	Appliance
1	aggregate	28	subwoofer livingroom
2	boiler	29	Living room lamp tv
3	Solar thermal pump	30	DAB radio livingroom
4	laptop	31	kitchen lamp 2
5	Washing machine	32	Kitchen phone & stereo
6	dishwasher	33	utilityrm lamp
7	tv	34	samsung charger
8	Kitchen lights	35	bedroom d lamp
9	HTPC	36	coffee machine
10	kettle	37	kitchen radio
11	toaster	38	bedroom chargers
12	fridge	39	hair dryer
13	microwave	40	straighteners
14	LCD office	41	iron
15	hifi office	42	gas oven
16	bread maker	43	data logger pc
17	amp livingroom	44	child’s table lamp
18	adsl router	45	child’s ds lamp
19	Livingroom s lamp	46	baby monitor tx
20	soldering iron	47	battery charger
21	USB hub	48	office lamp 1
22	Hoover	49	office lamp 2
23	kitchen dt lamp	50	office lamp 3
24	bedroom ds lamp	51	office pc
25	lighting circuit	52	office fan
26	livingroom_s_lamp2	53	LED printer
27	iPad charger		

The second dataset used is the WPuQ dataset. This dataset contains household electric and heat pump load of 38 households measured in Hamelin in Lower Saxony, Germany. For this study, only the household electricity consumption data is relevant. Of the 38 houses a total of 21 houses will be used in this analysis. The used dataset IDs are listed in table 3.3. For the visual analysis and the analysis of the optimization results the focus was only on five datasets. To choose these datasets, the houses were arranged in order of annual energy expenditure and the five median houses were picked. They are bolded in table 3.3.

Table 3.3 id of WpuQ datasets used in this study

WPuQ households										
3	4	5	7	9	12	14	16	18	19	20
21	22	27	28	29	32	34	36	38	39	

The datasets used in the study were cleaned and processed using python and stored locally.

3.2.2 LPG Simulation

The goal of the simulation was to create datasets that would resemble the UKdale and WPuQ datasets as much as possible. The UKdale set provided metadata that allowed for a more detailed load profile simulation design. However, the WPuQ dataset provided only generic information about the households. What was know is that all the households are single-family households located in a district near Hamelin, Lower Saxony, Germany. The houses were built in the late 90s and early 2000s. Although this would affect the accuracy of the models, the lack of information reflects the reality of many researchers building energy models as metadata is rarely available.

Demandlib

To simulate the UKdale house1 load profile the inputs used are presented in table 3.4.

Table 3.3 Demandlib inputs for simulating the UKdale house 1 load profile.

Parameter	Value	Comments
Holidays	United Kingdom	Demandlib uses a tool “workaledar” to simulate holidays.
Year	2016	
Annual electricity demand	3517.79	This value is in kWh. The UKdale load profile was measured to obtain this value.
Resample	Hourly	
Location	Postdam	Closest temperature profile

Although researchers usually do not have access to a specific household's electricity demand, they usually are able to obtain average values from literature to use them in when building their models.

To simulate the WPuQ profiles the inputs used are presented in table 3.5.

Table 3.4 Demandlib inputs for simulating the WPuQ load profile.

Parameter	Value	Comments
Holidays	Germany	Demandlib uses a tool "workaledar" to simulate holidays.
Year	2020	
Annual electricity demand	2786.66	This value is in kWh. This value is the annual electricity demand of the median house in the dataset
Resample	Hourly	
Location	Kassel	Closest temperature profile

Districtgenerator

To recreate house 1 using Districtgenerator the inputs used are summarized in table 3.6.

Table 3.5 Districtgenerator inputs for simulating house 1 load profile.

Parameter	Value	Comments
Building type	Terraced house	According to metadata from house 1 [42]
Retrofit level	2	According to TABULA and the available metadata on house 1, level 2 is appropriate [43]
Area	105	This is the reference area for a terraced house in the UK [43]
Occupants	4	Randomly generated by tool
Building year	1902	According to the metadata provide.

The inputs shown in this table are the inputs relevant for the electrical load profile. Since Districtgenerator was built with assumptions from TABULA in mind, referring to the TABULA tool was appropriate to decide on the assumptions for simulation. With house 1, the metadata made it clear what values had to be used. The only assumed value here was the area. Since it was not available in the metadata, the value assumed was obtained from the reference data of a terraced house built in 1902 in the UK.

Table 3.6 Districgenerators inputs for simulating the WPuQ load profile.

Parameter	Value	Comments
Building type	Single family house	According to metadata from the study [44]
Retrofit level	0	No metadata on retrofit was provided.
Area	122	This is the reference area for a single family house built in 2000 in Germany [43]
Occupants	2	Randomly generated by tool
Building year	2000	According to metadata from the study [44]

From the metadata shared by the WpuQ study [44], the building type and building year can be inferred. It was also mentioned that the average number of occupants among all the households was 2.38, hence rounding the value to 2 for the occupant input. This parameter, however, is a randomly generated value between 2 and 4 so the simulation had to be repeated until the value 2 is reached. The area was taken from the TABULA database as every house type has a reference value for area. Since the retrofit level was not given, it was assumed to be 0, which means that no retrofitting was done since the house was built. This assumption was made since these houses were built relatively recently.

LoadProfileGenerator

Although the user can create their own components in LoadProfileGenerator, it is time consuming and might take days to simulate a single household [20]. In this study, a house was created by putting together the relevant and available prebuilt components and introducing some modifications. The inputs used to create house 1 are summarized in table 3.8.

Table 3.7 LoadProfileGenerator inputs for simulating the house 1 load profile.

Parameters	Values	Comments
Temperature Profile	Berlin	Closest temperature profile to London from the available locations.
Geographic Location	Berlin	Big city, similar to London
House Type	HT20	Single family house with no heating or cooling
Modular household	CHR45	Changed energy intensity to energy saving, switched one

		occupant, added an infant, and and modified traits.
Energy Intensity	Energy saving, but prefer measured devices if available	

To build the house 1 model, a new house was created. The chosen house type, which determines the house infrastructure and autonomous devices, was HT20, a single-family house with no heating or cooling. This was picked since the focus of the study was only on electricity. No modifications were made to the house type. Concerning the modular household, which determines the inhabitants and their activities, no prebuilt components closely resembled the inhabitants of house 1 so some modifications had to be Introduced. The modular household selected was CHR45, family with one child, 1 at work, 1 at home. This was chosen because, looking at house 1's load profile, it can be inferred that during the day atleast one occupant is still home. Since house 1 includes one four-year-old and one two-year-old, they were included to the simulation and their activities were modeled. One default 16-year-old inhabitant was removed from the CH45 house.

Table 3.8 LoadProfileGenerator inputs for simulating the WPuQ load profile.

Parameters	Values	Comments
Temperature Profile	Berlin	Closest temperature profile to Hamelin.
Geographic location	Kassel	Closest available city to Hamelin
House type	HT20	Single family house with no heating or cooling
Modular household	CHR39	Couple 30-64 years with work
Energy Intensity	Energy saving	Since the buildings are relatively new

For simulating the WpuQ load profile, the temperature profile selected was berlin. The closes prebuilt geographic location available was Kassel which is around 100 km away. The selected house type remained HT20 since the data covers single family households and the heating and cooling measurements were neglected. Concerning the household, since the average number of household inhabitants in the paper was 2.38, the house was modeled to have two occupants. The occupants picked both had standard employee shifts. Since the houses tested are relatively new and had heat pumps (not included in the studied metered data), it is assumed that the houses and appliances are of low energy intensity.

SynPRO

Table 3.9 SynPRO inputs for simulating the house 1 load profile.

Parameters	Values	Comments
Profile type	Electric	
Weather dataset	ERA5	The option with required simulation year
Location	Berlin	Closest temperature profile to Hamelin.
Temporal resolution	1 hour	
Simulation year	2016	
Socio-economic factor	family	
Number of residents	4	
Device Efficiency	standard	
Net Usable Area	105	According to values from TABULA [43]

Tables 3.10 and 3.11 present the input parameters set for the synPRO simulation tool to simulate electric load profiles.

Table 3.10 describes the simulation for House 1. The location chosen for this house is Berlin, primarily due to its close temperature profile similarity to Hamelin. The simulation year identified for House 1 is 2016, and the house is presumed to accommodate four residents.

Table 3.11 focuses on WPuQ. For this simulation, the location selected is Kassel, also because of its temperature proximity to Hamelin. Similar to the other simulations it was assumed that the house has two occupants.

Table 3.10 SynPRO inputs for simulating the WPuQ load profile.

Parameters	Values	Comments

Profile type	Electric	
Weather dataset	ERA5	The option with required simulation year
Location	Kassel	Closest temperature profile to Hamelin.
Temporal resolution	1 hour	
Simulation year	2020	
Socio-economic factor	family	
Number of residents	2	
Device Efficiency	standard	
Net Usable Area	122	According to values from TABULA [43]

3.2.3 Energy Modelling and Optimization

An analysis was conducted on an energy model to examine the influence of LPG input assumptions on its performance and optimization results. This model was developed using the OEMOF framework.

The subsequent sections detail the structure of the energy model and its core components. The chosen parameters for these components will be outlined and discussed, highlighting their significance in the modelling process.

Lastly, the optimization metrics used to evaluate the model's outcomes will be presented. These metrics provide a standard for assessing the effectiveness of the model in different scenarios.

Through this structured examination, the study aims to provide a clear understanding of the role of LPG input assumptions in energy system modelling and optimization.

The energy model built using OEMOF was based on a hypothetical scenario of a residential household in Europe. The model that was built had to have both the design and operation optimization problems that OEMOF can solve. With this in mind, the following model was created:

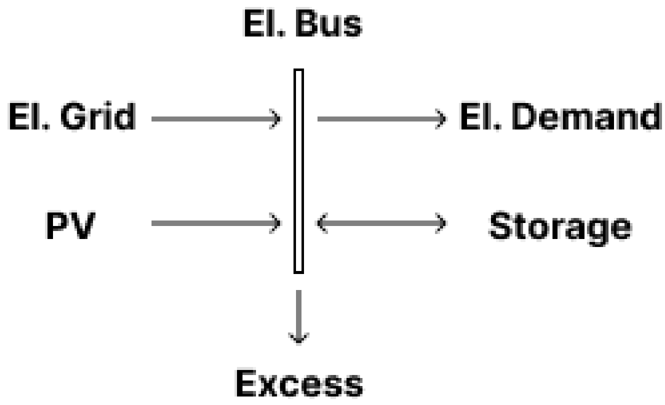

Figure 3.2 OEMOF Energy Model

The energy system that was modeled is a grid connected residential household. The system can potentially have PV modules and energy storage installed. Only the household's electricity demand was modeled. The heat and water demands were neglected in this model.

To create this energy model six components were included in the system. Since the electrical energy is being studied, the first required component is the electrical bus. This component is modeled as a bus component in OEMOF.

The second component is the electric grid. This was created as an OEMOF.solph source component. This component contains only one variable, the variable cost. During the time of modeling, the cost of electricity in Germany was fluctuating around 0.3 €/kWh.

The third component is the electrical demand of the household. This was modeled as a sink component. The household electrical demand flow is the main variable that was tested in this study. Here, fixed electrical demand was used. The fixed electrical demand used was the actual and simulated load profiles.

An excess component was required to be included for the electricity bus to allow overproduction. The excess component was modeled using a sink component and has no purpose other than receiving overproduced electricity.

The last two components included are components from the ESyOpT package. The first ESyOpT component is the PV plant. For this, many parameters were required. The first parameter assumed was the PV capex. The value assumed was 1392 €/kW [45]. This number can be broken down into three separate components, the module (412.2 €/kW) the structural balance of system (86.2 €/kW) and the electrical balance of system (181.6 €/kW). The OPEX was assumed to be 26 €/kWp/yr [46]. The installation and inverter costs were assumed to be 0 as they were accounted for in the PV CAPEX. The lifetime of the PV and inverter modules were assumed to be 30 and 15 years respectively according to the study. To allow the full freedom in the optimization process no constraints were put on the area or power availability of the PV. Finally, the same PV normalized timeseries output was obtained internally and used across all models for fair comparison of the results. The PV component was set for design optimization.

The final component, also an ESyOpT component, is the battery storage. The parameters selected were also found in literature, but also some assumptions were made. The investment costs were broken down between the battery and the inverter. The battery CAPEX was assumed to be 356 €/kWh broken down into 248.5 and 107.5 €/kWh from the battery and electrical balance of system respectively [47]. From the same source, the inverter CAPEX was found to be 478.6 €/kW. Another source bundled the CAPEX of the battery and inverter system and found a price of 850 €/kWh [46]. Under these numbers however, the optimization model decided to not include a battery in the simulation. For this reason, it was assumed that the inverter + battery CAPEX was 500 €/kWh instead and with no OPEX cost. This was done since it is important for this study that the battery was included in the simulation to be able to study the operational optimization decision according to the different load profile inputs, and not just the design optimization decisions. Finally, for both the PV and energy storage component, a weighted average cost of capital of 0.034 was selected as literature pointed to values being between 0.03 and 0.04 [48]. The code for this model can be found in Annexe A.

The optimization model created has the following simplified objective function:

$$\min: \sum_{pv} x_{pv}^{\text{capacity}} \cdot c_{pv}^{\text{capacity cost}} + \sum_{s} x_{s}^{\text{capacity}} \cdot c_{s}^{\text{capacity cost}}$$

$$+ \sum_{s,t} x_{s}^{\text{flow}}(t) \cdot c_{s}^{\text{marginal cost}} + \sum_{g,t} x_{g}^{\text{imports}}(t) \cdot c_{g}^{\text{marginal cost}}$$

This objective function is composed of four segments. The first segment represents the investment cost of the PV system (design optimization problem). The second segment represents the investment cost of the storage system (design optimization problem). The third segment represents the operational cost of the storage system (operation optimization problem). Finally, the fourth segment represents the cost of the grid imports (operation optimization problem). The goal of the optimization was to minimize these costs that are associated with the energy system.

Chapter 4

Results

4.1 Analysis of Load Profiles

The following section addresses our second objective, which is centred on the analysis of simulated load profiles. Employing a range of tools, we aim to generate profiles that align with historical data. Our methodology combines both visual and statistical techniques, many informed by existing research. The section is structured into two main subsections: visual analysis, which includes the load profile characterisation and relative heatmaps methods, and statistical analysis, focusing on annual and monthly energy demands. This framework ensures a detailed assessment of the accuracy of the simulated load profiles.

4.1.1 Visual Analysis

In the visual analysis the presented graphs depict the average power load profiles of a household, segregated into two distinct categories: weekends and weekdays. Each graph further distinguishes the load profiles based on the seasons: Spring/Fall, Summer, and Winter. The table of power values for each of these figures can be found in Annexe B. First the actual load profile, UKdale dataset of house 1, will be analysed, followed by the simulations of house 1 made by the four tools demandlib, Districtgenerator, LoadProfileGenerator, and synPRO.

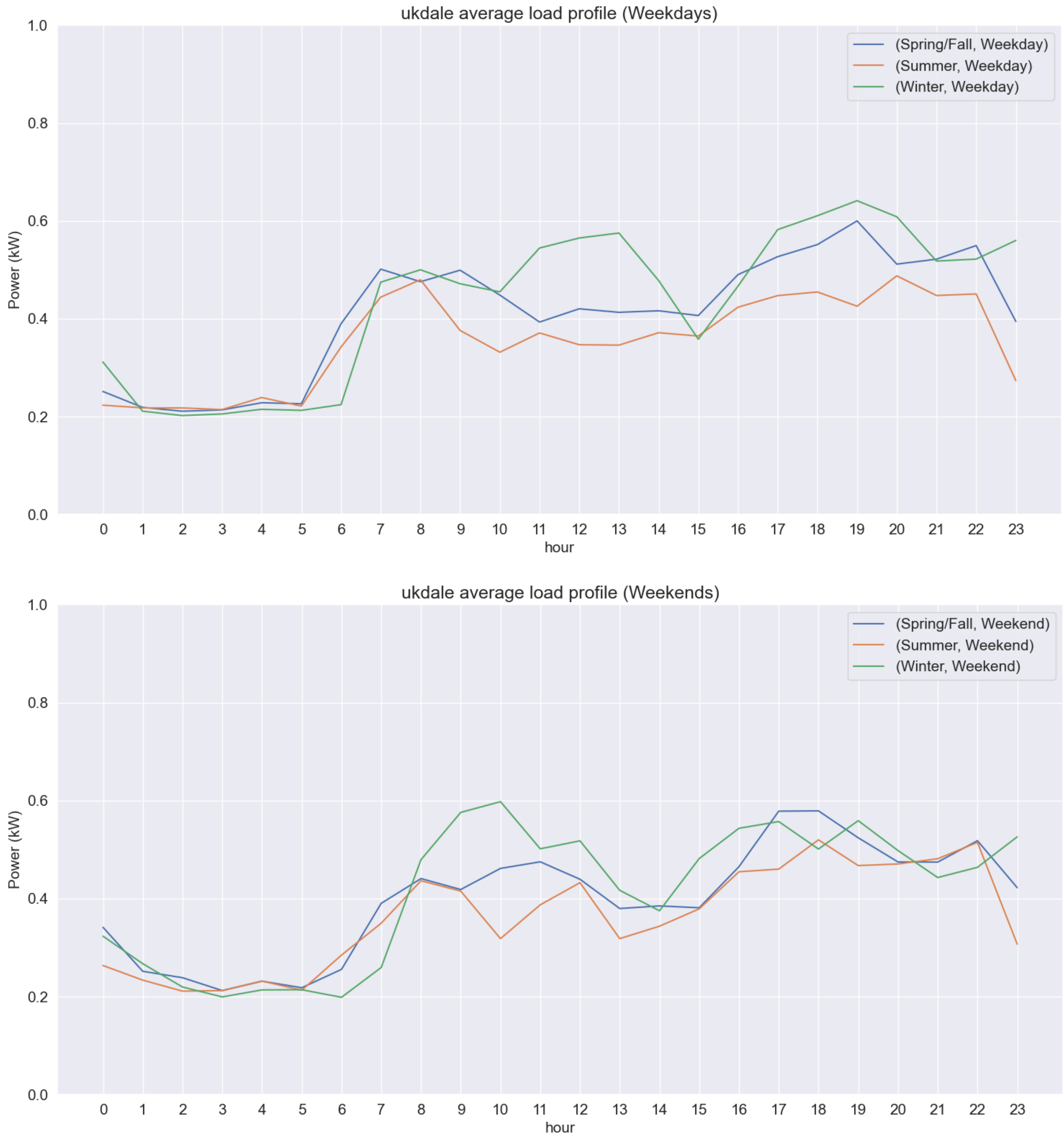

Figure 4.1 UKdale clustered load profiles

UKdale

The load profile of the UKdale Both on weekends and weekdays, a consistent base load is observed between 0.2 kW to 0.3 kW during or slightly past midnight until 5 in the summer and spring/fall and until 6 am in the winter. The only difference between the weekdays and weekends is the start of the base load at night. On the weekends the base load is approached slightly later at night.

The early peak power consumption across weekdays/weekends and the different season is inconsistent. In general, a 0.5kW morning peak will occur around 8 in the summer and spring/fall, but slightly later than usual in winter and at a higher power of around 0.6 kW. A pronounced revival of the early peak happens during midday especially during winter. Looking at the individual device profiles of house 1, this revival can be attributed to the use of kitchen appliances perhaps by the stay-at-home occupant.

The late peak happens around 18 on the weekends and slightly later at around 19 on the weekdays.

This is because one of the occupants comes back home late from work on weekdays. The intensity of the peak is around 0.6 kW, matching the morning winter peak. Except in summer where its at 0.5 kW, similar to the summer and spring/fall early peaks. After a slight dip in the late peak, a renewed jump can be observed. The individual device load profiles shows that this jump is caused due to the use of the dishwasher appliance.

During weekends, a gradual increase in power load is seen from 6 to the early peak, post which there's a decline until the evening. Notably, the Winter curve's rise is more steeply inclined than its counterparts. In the summer and spring/fall, the power drops steeply past 22. In the winter, the power rises instead.

Winter consistently exhibits higher power consumption than the other seasons, especially during peak hours. Summer sees the least power consumption around midday during weekends, possibly due to reduced indoor activities. It is however important to note that in early June and late July, there were periods where the house was unoccupied for several days. Hence, the reduced load can be partially attributed to the periods of inoccupation as the values are averaged.

The Spring/Fall curve serves as an intermediary between the Summer and Winter curves, exhibiting characteristics of both, but not as pronounced.

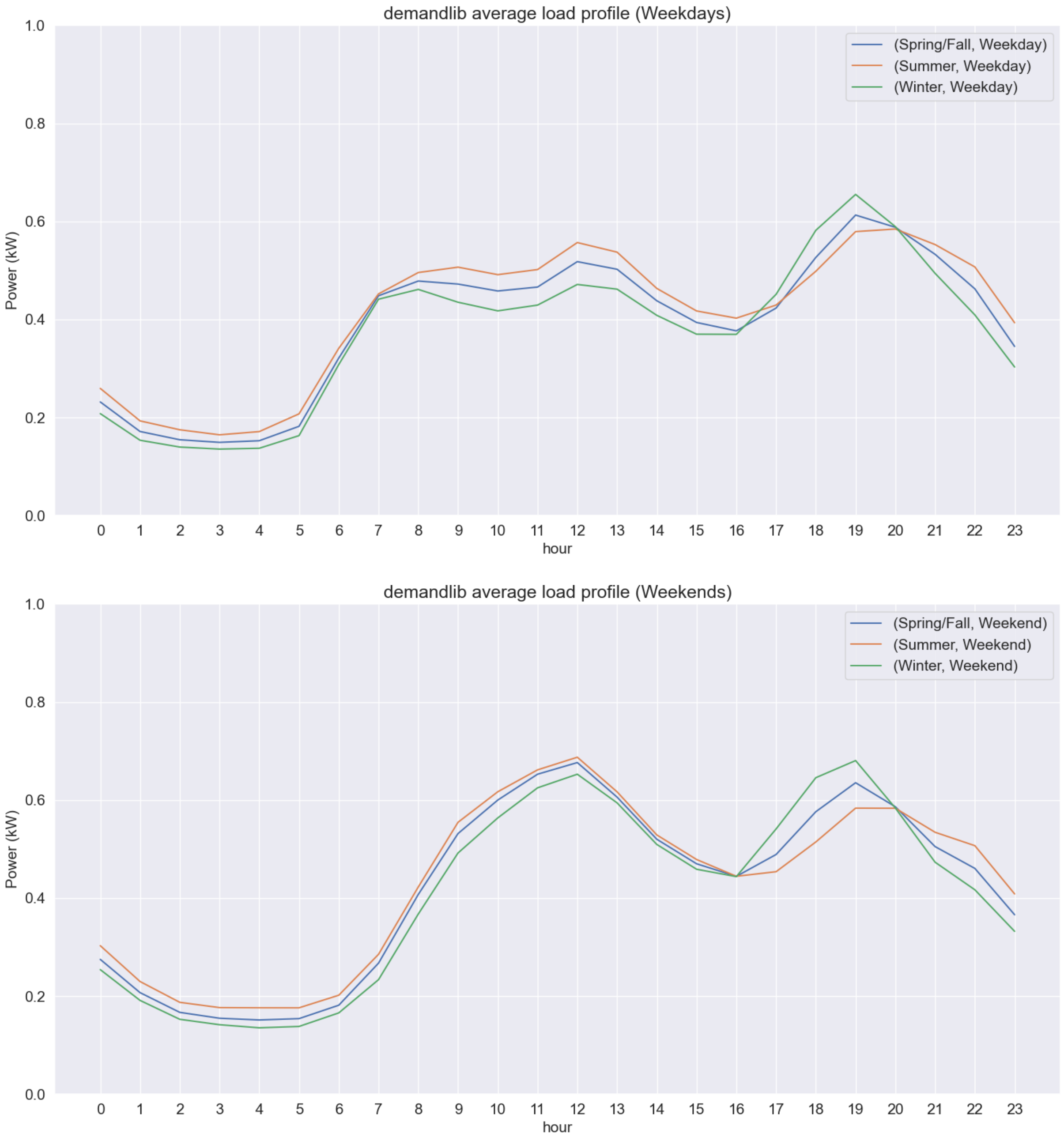

Figure 4.2 Demandlib clustered load profiles

Demandlib

Moving on to the simulated data, the next four tools were used to simulate the UKdale house 1 dataset.

The load profile produced by demandlib is smooth and consistent. The base load all year round was at the lowest around 0.14 kW during winter and at the highest 0.17 kW during summer. The base load is sustained for around 6 hours in weekends and around 5 hours in weekdays.

The weekend load rise starts at around 6 and the weekday load rise starts at 5. The load rise is much steeper during the weekday. An early semi-peak happens during the weekday at around 7 after which a real early peak happens during midday. On the weekend, since the rise is less pronounced, there is no early semi-peak and only one midday peak at 12. On the weekends the midday peak is more intense.

The late peak happens at 19 consistently across the year. During the winter, the peak power is slightly around 0.67 kW and is at its highest. The summer late power peak is the lowest at around 0. But is longer lived than the late peak in winter and spring/fall. after 19 power decreases consistently.

Winter load profile power is consistently the lowest except during the late peak when it becomes the highest.

Districtgenerator

The hourly load profile generated by Districtgenerator is characterised by many peaks and inconsistencies across.

One notable characteristic of the Districtgenerator load profile is the small midnight peak at 1 during the weekends and the slightly smaller one during the weekdays. The base load is short lived, and unstable. The load profile power rises steeply at 5. During weekdays an early peak is reached at 8. During weekends this peak is inconsistent. During spring/fall and winter it happens around 8, but during summer, two early peaks can be seen. The intensity of the peaks also varies. The highest early peaks of 0.9 kW happen during the spring/fall on weekends. During the weekdays, the intensity is more consistent at around 0.65 kW.

Winter afternoons and evenings are characterised by two peaks, at 16 and 21. In the winter these peaks are more intense than the earlier ones. The highest peak happens on the weekends in winter at 21. This peak is around 1 kW. After 21 electricity consumption declines but is nowhere near the base load. All load profiles sustain high load for a long time.

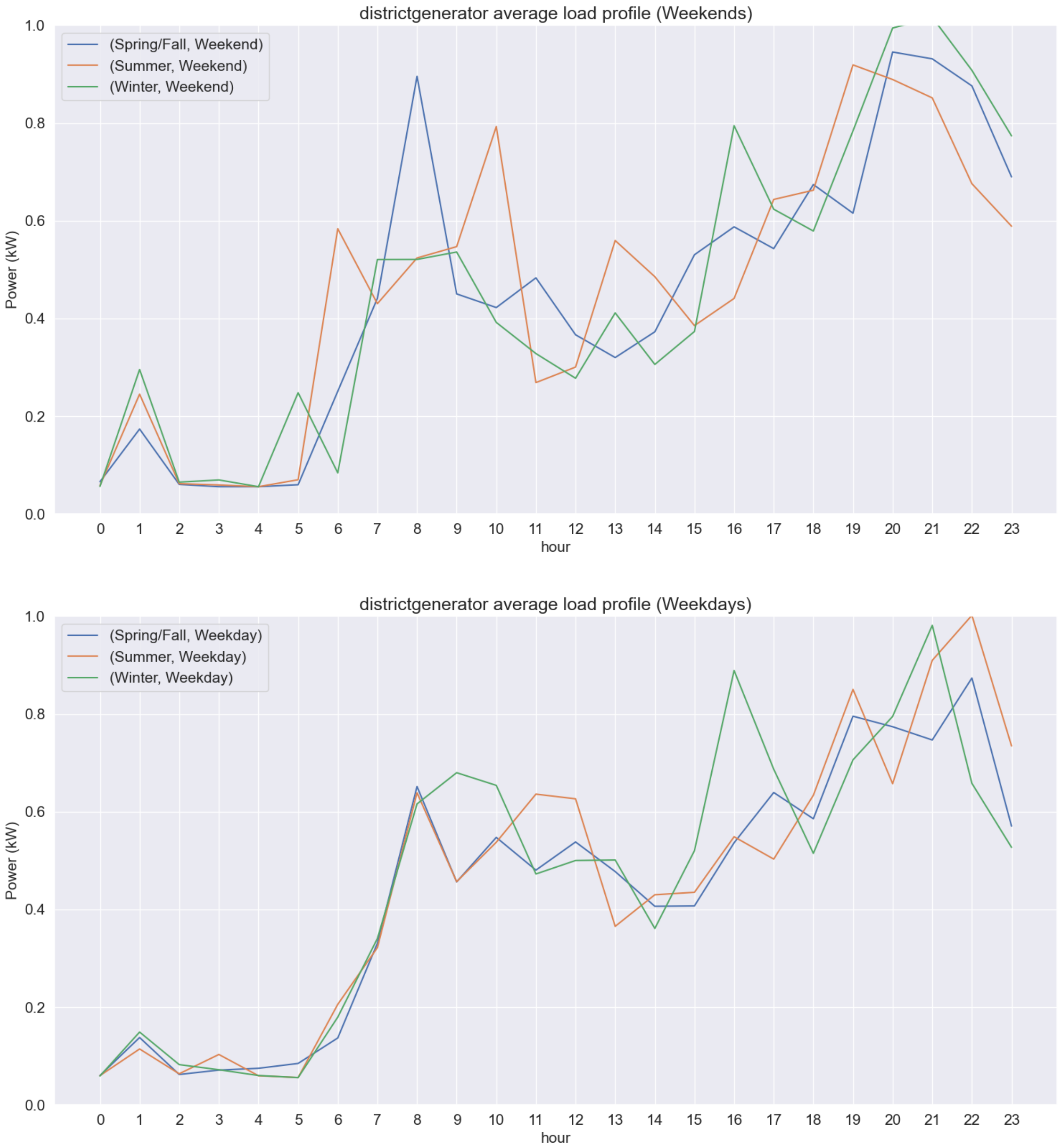

Figure 4.3 Districtgenerator clustered load profile

LoadProfileGenerator

The load profile of this tool is consistent in shape and differs mainly in intensity. It is characterised by huge peaks and rapid rises and falls. The base load is stable, lasting for 5 hours between 0 and 5 at an intensity of around 0.11 kW all year long. Note that the y-axis was extended for this tool's load profile due to its very intense midday peak.

Starting from 5 am, there's a pronounced surge in power consumption for all profiles. There are weak peaks in the early morning that are dwarfed by the huge midday peak at 11 across all load profiles. Themost significant spike is observed in the summer profile, peaking sharply during weekdays and weekends at 2.2 kW and 2.0 kW respectively. In comparison, both the Spring/Fall and Winter profiles exhibit more moderated peak which vary between 1.25kW and 1.5kW.

Post the midday peak, all profiles witness a rapid decline returning to around 0.5 kW within 2 hours. Despite having the highest peak in the summer, the summer load profiles dip below the other seasons in the afternoon.

Between 16 and 17 all profiles exhibit an increase in consumption, albeit less pronounced than the morning surge. The evening peak is reached at 18 on the weekdays and 19 on the weekends. Concerning the intensity of the peaks, the winter and spring/fall maintain an peak between 0.6 kW and 0.75 kW while the summer is higher during the weekdays at 0.9 kW and lower during the weekends at 0.57 kW.

After the late peak, energy use rapidly declines until it converges towards the base load by midnight..

The Summer season consistently manifests the highest and most abrupt peaks, indicative of intense short-term power requirements. Perhaps due to the randomly selected outdoor appliances despite this being an apartment. The profiles however are very similar all yearlong except during the summer peaks which are considerably higher than the other peaks.

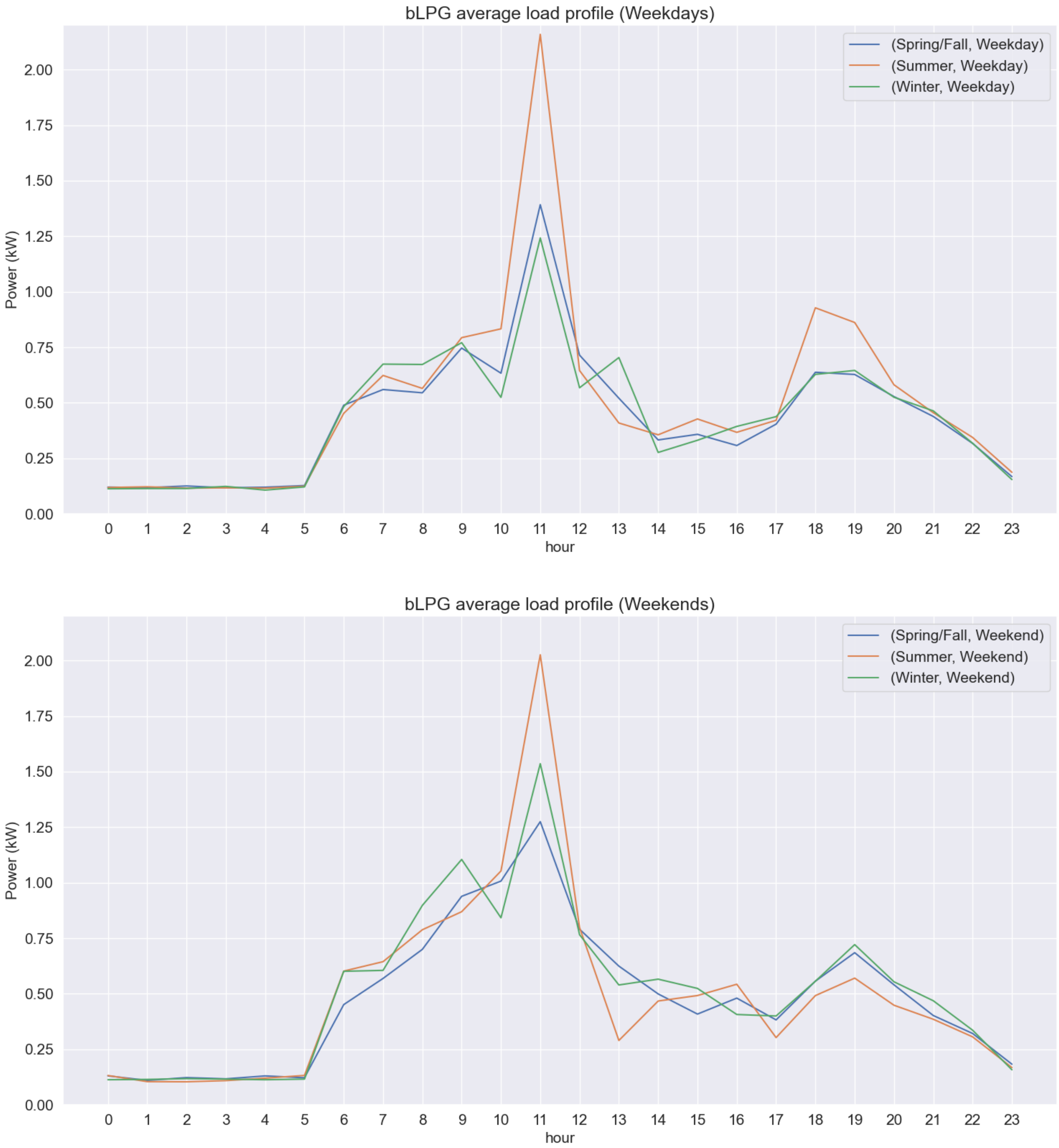

Figure 4.4 LoadProfileGenerator clustered load profiles

SynPRO

The synPRO simulation exhibits great variety across the different seasons and between weekdays and weekends. Nevertheless, the peaks are moderate.

The baseload of synPRO is the highest among all the other simulations especially in the winter with a value of up to 0.3 It is also the only baseload that showed variety in value across the seasons. The rise start time also starts early at 4 but rises gradually. The early peak at 8 on the weekday is also, similar to demandlib's 7 peak, very shallow. Moreover, no early peak exists on the weekends and the rise is more gradual and starts at around 7.

A variety in the average midday peaks means that during the day peaks usually exist at different times. In the case of synPRO, the peaks can happen between 11 and 14. The peaks are more intense in the winter and less so in the summer. The high load sustains in the winter and spring but not in the summer.

The evening peak happens early on 18 or 19 on the weekdays and later in the winter. Again, the winter peaks are more intense, and the summer peaks are weak. The higher peaks in the winter can partially be attributed to a higher baseload as this increases the overall power consumption through the day.

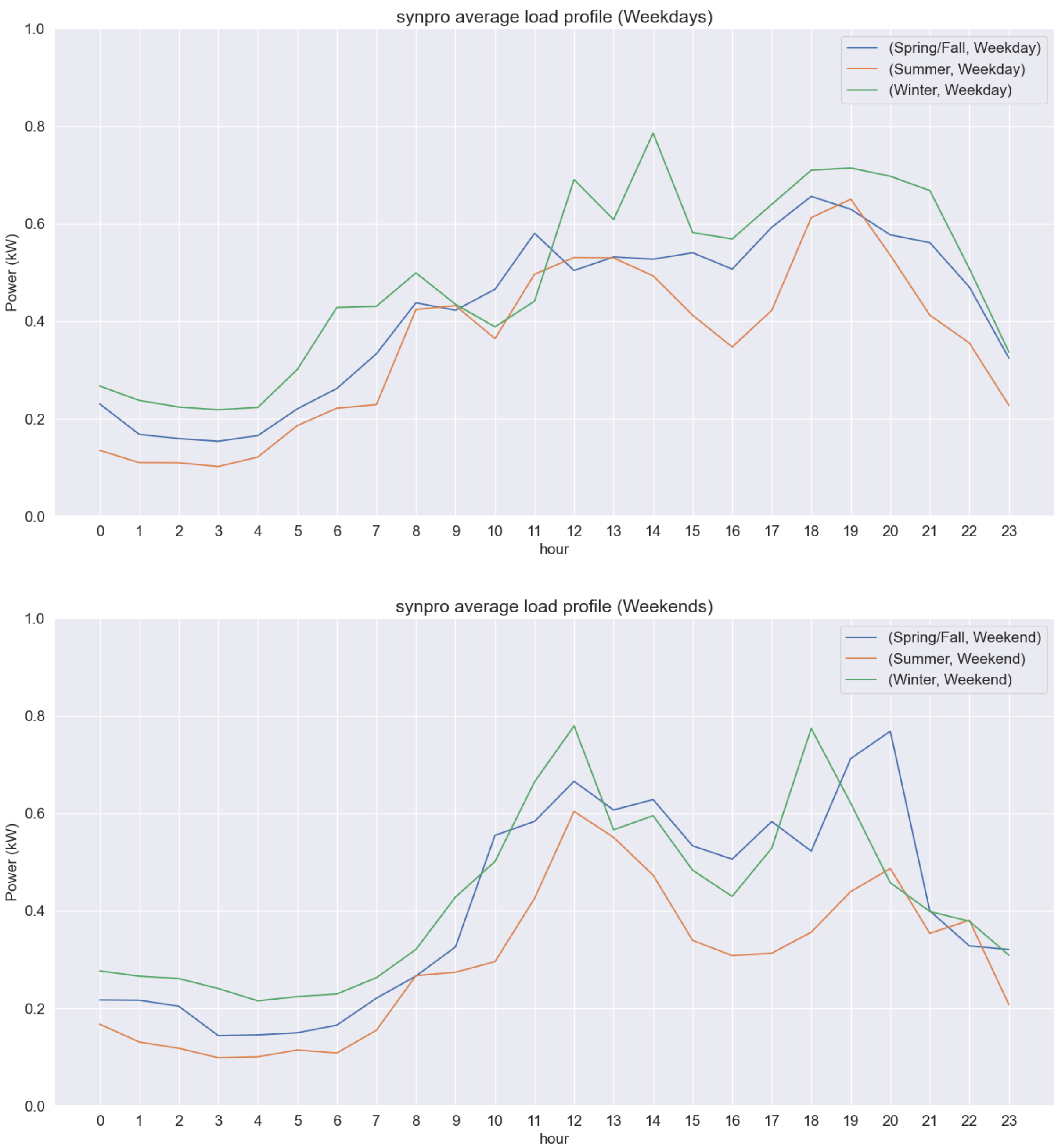

Figure 4.5 synPRO clustered load profiles

Comparative Analysis

The UKdale dataset, which served as the benchmark in this analysis, presented a consistent base load during nighttime, particularly noticeable between 0.2 kW to 0.3 kW. This consistency contrasted with the simulations in various ways. For instance, the Demandlib simulation delivered a smoother base load throughout the year, but it was lower in magnitude, hovering between 0.14 kW to 0.17 kW. The Districtgenerator's base load profile was markedly different, characterized by its short duration and instability. LoadProfileGenerator maintained a steady base load of around 0.11 kW. SynPRO's profile emerged as distinct among the simulations, particularly during the winter when it reached a base load as high as 0.3 kW, even surpassing UKdale's own winter base load which sustained base load of over 0.25 rarely and for a short period of time.

In terms of peak consumption patterns, UKdale generally witnessed its morning peaks around 8, with slight variations across seasons. Evening peaks were predominantly observed around 18 or 19. In contrast, Demandlib, while maintaining consistency, did not mirror UKdale's pronounced morning peak, instead showing an early semi-peak leading up to a significant midday rise. The Districtgenerator's approach was even more distinct, with its load profile punctuated by small midnight peaks and inconsistent early morning peaks across seasons. LoadProfileGenerator's sharp midday peak, especially evident during the summer months, starkly contrasted with UKdale's moderated consumption pattern. SynPRO offered a diverse pattern, with multiple peaks throughout the day, and a particularly pronounced evening peak in winter that reflected some characteristics of the UKdale winter behavior.

The seasonal variations in the load profiles further highlighted the differences between UKdale and the simulations. UKdale's winter consumption was characteristically higher, especially during its peak hours. Demandlib's profile, on the other hand, exhibited a trend of reduced power during the winter, while maintaining more uniformity during the summer. This differed from UKdale's clear winter elevation in consumption. Districtgenerator's unique approach resulted in two distinct winter peaks during the afternoon and evening. LoadProfileGenerator's summer profile stood out due to its pronounced midday peak, deviating considerably from UKdale's moderated summer consumption. SynPRO's winter profile, although exhibiting higher overall consumption, also stood out due to its elevated base load.

Among the simulations themselves, distinctions were apparent. Demandlib's consistent load profile across seasons set it apart from the more varied profiles of the other tools. Districtgenerator, with its erratic load variations, starkly contrasted with the other simulations. LoadProfileGenerator, while maintaining shape consistency across seasons, was characterized by intense midday peaks, especially in summer, making it divergent from the others. SynPRO offered a mix, with its winter profile indicating higher consumption, providing a unique perspective among the simulations but being the only one to match with UKdale’s seasonalities.

While UKdale offered a grounded and consistent benchmark, each simulation reflected varied patterns and behaviors. Some simulations, like Demandlib, exhibited overall consistency, while others, such as Districtgenerator, presented more fluctuating patterns.

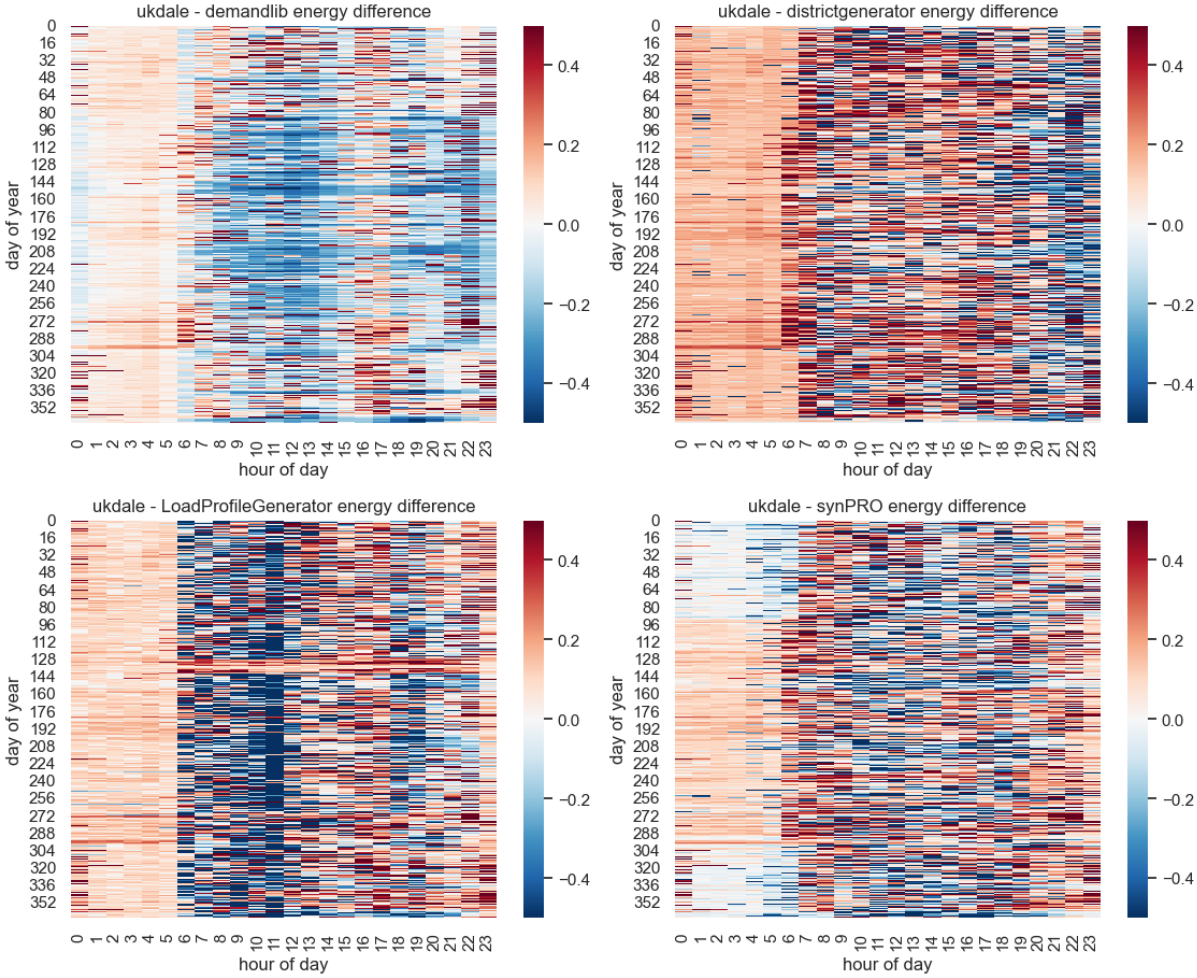

Figure 4.6 relative heatmaps comparing UKdale and the simulations

The relative heatmaps method visually represented differences between the actual and simulated load profiles. By subtracting corresponding timesteps, the heatmap was generated. Although a direct statistical comparison of these timesteps wasn't advantageous, the visualization highlighted periods when the simulated profile either over- or underestimated demands.

Comparing the baseload in the early hours between around 0 to 6 shows that all the simulations underestimated the baseload except for synPRO in the winter. This is consistent with the findings from the load profile characterisation.

The early peaks between 6 and 12 are slightly overestimated by demandlib in the summer and heavily overestimated by LoadProfileGenerator all year round. Again, this was evident from the sharp peaks in the load profiles. A red patch however can be seen after day 130 during midday. This corresponds to a drop in demand in the 5th month which will be seen in the analysis in the next section. The very messy midday map of synPRO and Districtgenerator indicates irregularities in the peaks during these hours throughout the year.

4.1.2 Statistical Analysis

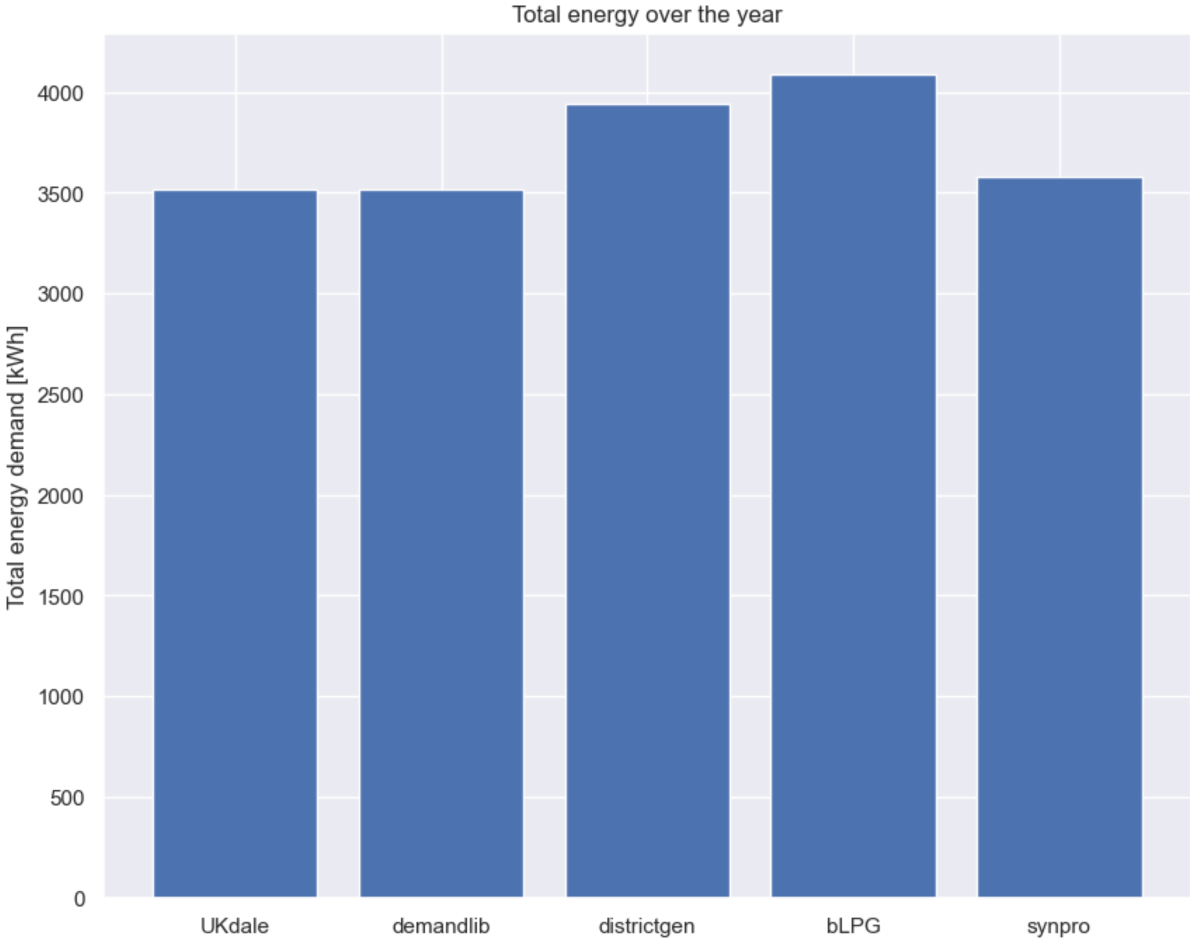

Figure 4.7 Total energy expenditure of real and simulated load profiles over a year.

Figure 4.7 depicts the total annual energy expenditure of the real and simulated load profiles over a year. The UKdale dataset serves as the benchmark, representing the actual energy demand, with a value of 3517.8 kWh. The remaining bars represent the output of various tools employed to simulate load profiles that aim to resemble the UKdale dataset.

Both the "demandlib" tool and the UKdale dataset exhibit identical total energy demand values at 3517.8 kWh. This is due to the input requirements of demandlib. When simulated the total annual energy expenditure was inputted into the simulation.

The synPRO tool's simulation results in an energy demand of 3581.5 kWh, which is slightly higher than the UKdale. The precent error is 8%.

On the other hand, the districtgenerator and LoadProfileGenerator tools yield significantly higher energy demand values. Specifically, districtgenerator results in a value of 3938.6 kWh, an error of 12%, and LoadProfileGenerator surpasses all with a value of 4087.2 kWh and an error of 16%.

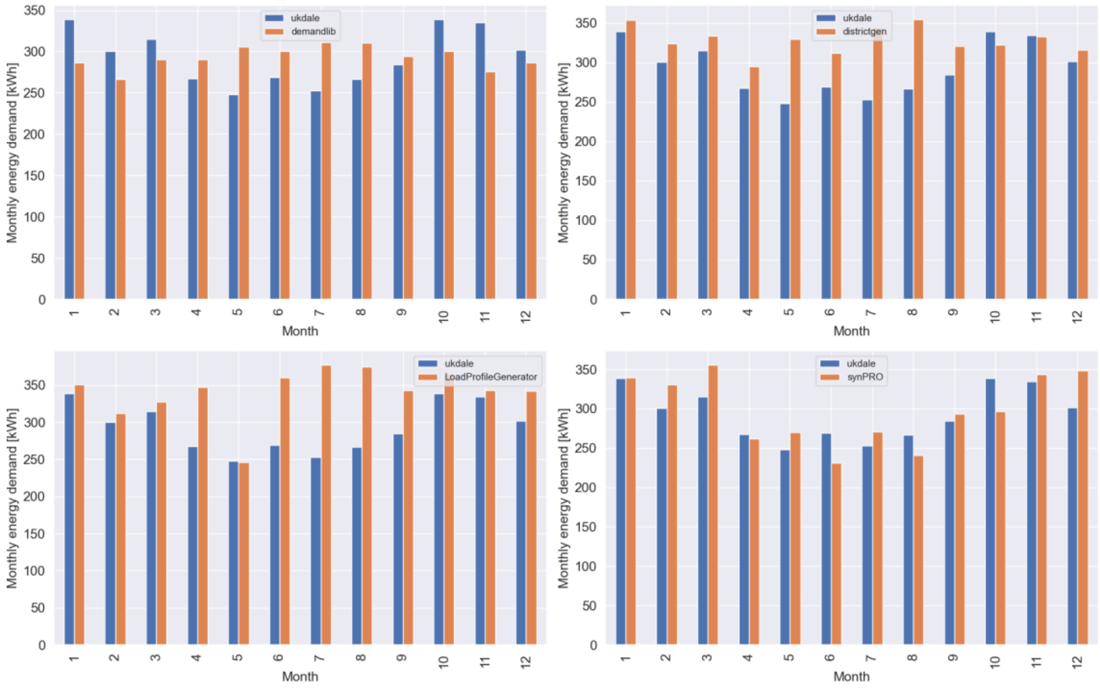

Figure 4.8 Monthly energy consumption of the UKdale dataset and the simulations

The series of graphs present the monthly electricity demand (in kWh) for a year, comparing the actual dataset, UKdale, with simulated results from different load profile generation tools. The UKdale monthly electricity demand displays a seasonal trend with lower demand in the summer months.

The energy demand from the demandlib simulation exhibited a weak seasonal trend. Contrary to UKdale, demandlib displayed higher electricity demand in the summer months. Hence, the summer demand is overestimated, and the winter demands are underestimated. The calculated MAPE of the demandlib simulation is 12.7%

Districtgenerator's simulation does not display a seasonal trend. The simulation provided similar values during the winter months but overestimated them in the summer. The calculated MAPE of the Districtgenerator simulation is 13.0%

LoadProfileGenerator also shows weak seasonal trend that increases in the summer except for a big dip in May. The simulation closely follows the UKdale dataset in the winter but shows a large margin of error in the summer months. The calculated MAPE is 16.3%

The synPRO simulation exhibits the lowest margin of error among all simulations. The calculated MAPE is 8.2%. This tool is also the only tool that matched the seasonal trend and showed a dip electricity usage during the summer months.

The same analysis was also done using a different dataset and simulations but with the same LPGs. This time the the analysis was benchmarked against the 21 selected houses in the WPuQ dataset. And because this dataset is an average of 21 houses, the seasonalities were slightly averaged out. Nevertheless winter in WPuQ datasets also showed higher energy usage when compared to summer.

Similar outcomes were observed in this anaylsis. The similarities are as follows. Demandlib underestimated enery usage in the winter and overestimated in the summer. It also exhibited weak seasonalities albeit slightly more energy usage in the summer. bLPG again underestimated energy usage in month 5. This is due to simulating vacation and time away from the house. It also overestimated in the summer months. In general the seasonalities among the tools did not change from one simulation to the other, but the intensities did change.

Some contrasts include districtgenerator underestimating instead of overestimating in energy usage. One notable difference is the very low forecasted monthly usage by synPRO. The main parameter that lead to this drop in occupants from four in the UKdale simulation to two in the WPuQ simulation.

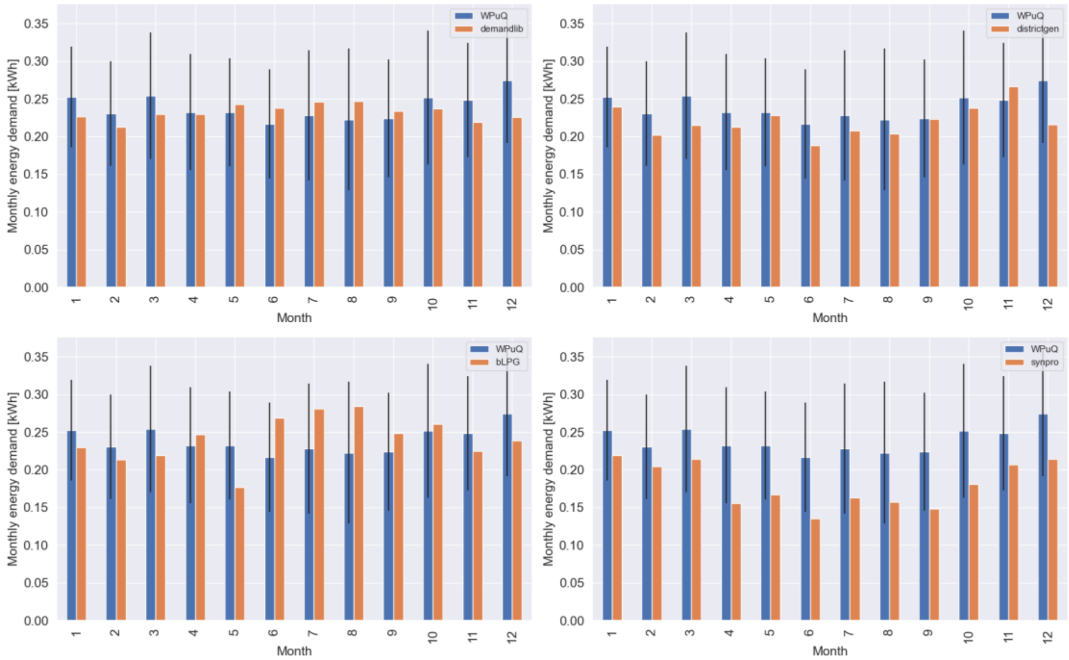

Figure 4.9 Monthly Energy consumption of the WPuQ datasets and the simulations

The frequency analysis undertaken delved into the occurrence rates of specific power levels over the year's hours. While average values can sometimes be deceptive, examining the frequency of power levels paints a clearer picture of the alignment between synthetic and historical datasets.

The load duration curve representing the UKdale shows that power throughout the year rarely crosses the 1kW mark and that only 500 total hours (6%) of electricity consumption happens while the power is above 2kW. On the contrary, progressive load profile generators like LoadProfileGenerator and district generator produce values above 1kW 14% and 11% of the time respectively.

On the other hand, demandlib is too conservative and never produces any values above 1kW. Its max power value is 0.74kW. In contrast, the max value of UKdale dataset is 2.9 kW. Despite these low values, demandlib’s power values over the hours does decay slower than the other simulations.

Looking at the load duration curve of the WPuQ dataset simulation a similar conclusion can be done. In this case however, since the WPuQ curve is an average of many households, its more conservative.

Nevertheless, standard deviation lines show that it is unlikely for households to be producing some of the high-power values that the load profile generators are producing.

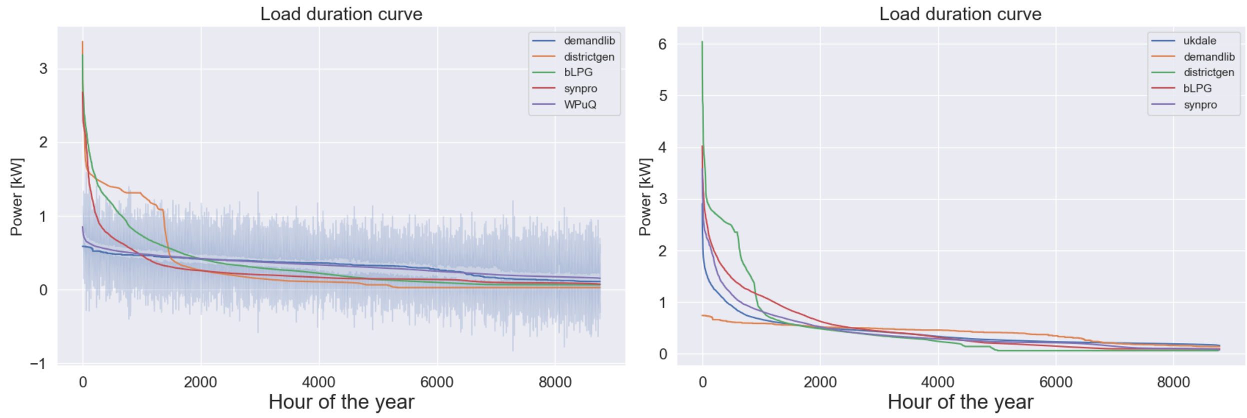

Figure 4.10 Load duration curve of UKdale simulations on the right and WPuQ simulations on the left

4.2 Analysis of Optimization Results

Section 4.2 addresses the third objective, focusing on the impact of simulated load profiles on optimization outcomes. Using a combined black box and inductive methodology, an energy model was developed that considers both design and operation optimization. The design optimization pertains to the configuration and sizing of energy components, with the total system cost as a primary metric. Operation optimization, on the other hand, concerns daily energy system management, determining optimal strategies for energy storage and release. By analysing strategies derived from both actual and simulated profiles, the study highlights the nuances introduced by LPG assumptions. In essence, the research offers insights into how synthetic load profiles, and the assumptions behind them, influence design and operational decisions in energy system models.

4.2.1 Design Optimization

In the design optimization results, a tenuous relationship was observed between demand and storage investment. Notably, even with comparable energy demands, synPRO and demandlib displayed distinct optimized energy storage capacities. Demandlib necessitated just 60% of the storage investment of synPRO and a marginally smaller PV system. This implied that the demandlib model sourced less overall energy from PV production and battery discharge. However, this was offset by a reduced energy excess or curtailment.

The correlation between grid imports and demand was also subtle, as highlighted by the results from districtgenerator and LoadProfileGenerator. While LoadProfileGenerator's demand was approximately 77% of districtgenerator’s, its grid imports slightly exceeded those of districtgenerator. Moreover, LoadProfileGenerator called for more significant PV investments, but this was accompanied by a considerably higher curtailment rate.

In comparison to UKdale, synPRO's design optimization results were closely aligned across most parameters. Interestingly, even with identical demand, Demandlib necessitated a more modest investment than UKdale and executed fewer energy curtailments. The most pronounced curtailment was observed with LoadProfileGenerator. Yet, it also possessed the largest PV system. In terms of curtailment as a fraction of PV production, LoadProfileGenerator ranked second with 23%, slightly behind synPRO's 23.5%. UKdale exhibited a median curtailment ratio of 21%, with demandlib presenting the lowest curtailment.

Regarding storage capacity investments, districtgenerator's model demanded the most significant commitment. In stark contrast, the investment needed by demandlib's model was markedly less than its counterparts, standing at merely 41% of what districtgenerator's model required.

Table 4.1 results of the design optimization of the different energy models

Parameter	UKdale	Districtgenerator	SynPro	Demandlib	bLPG
Grid Imports (kWh)	1929.1	2225.0	1910.0	1824.5	1714.8
PV Production (kWh)	2151.7	2432.7	2342.4	2143.8	3347.1
Storage Charge (kWh)	499.1	954.5	526.6	281.1	894.2
Storage Discharge (kWh)	382.6	736.5	406.5	217.9	692.2
Demand (kWh)	3517.8	3938.6	3581.3	3517.8	4087.2
Excess (kWh)	446.5	501.1	551.0	387.4	772.7
PV capacity (kWp)	2.3	2.6	2.5	2.3	3.5
Storage capacity (kWh)	1.9	2.9	1.8	1.2	2.6
Storage power (kWp)	0.9	1.4	0.9	0.6	1.3

4.2.2 Operation Optimization

The battery's behavior in different models was closely studied. To better understand how the battery operated over time, a charge duration curve was used. This curve showed how long the battery stayed at each charge level.

The result of almost all the charge duration curves is very similar. A notable outlier is the charge duration curve of demandlib's model. In this model the battery spent relatively more time at lower charge levels when compared to the other models.

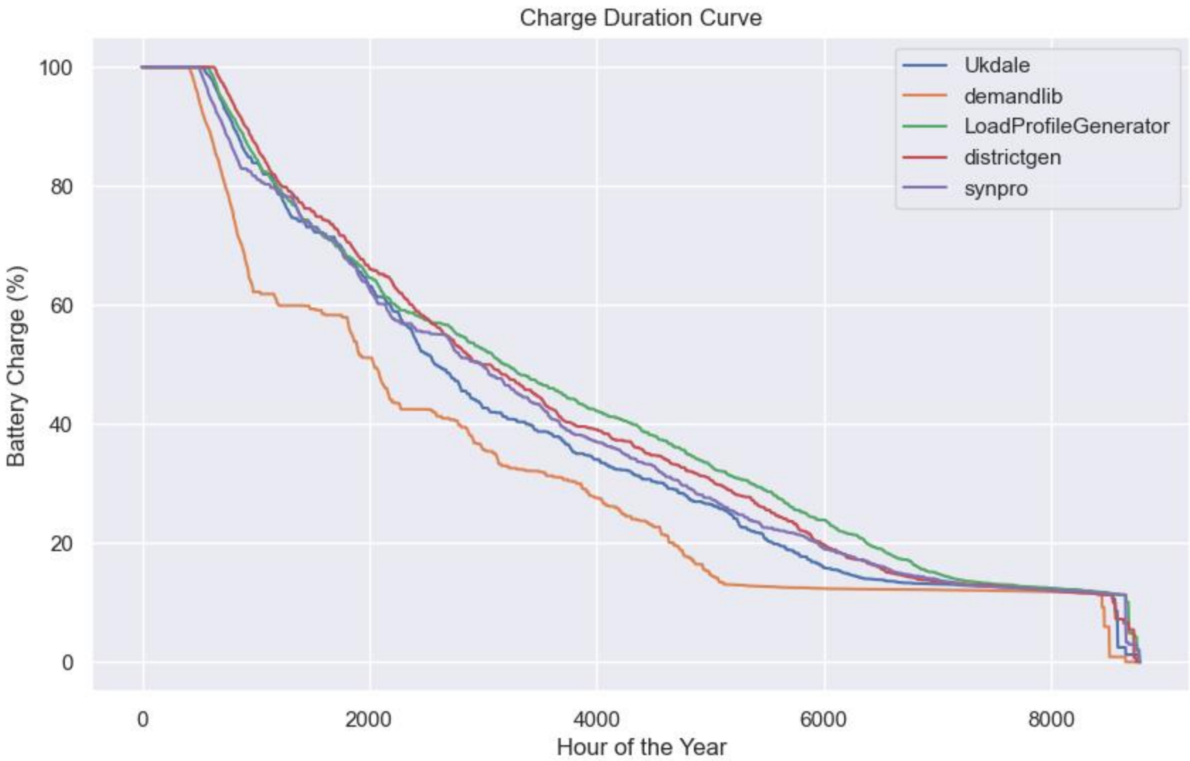

Figure 4.11 Charge duration curve of UKdale simulation

Chapter 5

Conclusion

5.1 Problem reformulation

Modelling and optimization of energy systems are essential for sustainable energy planning. A challenge within the residential household sector is the limited availability of detailed load profile data for research [6]. Load Profile Generators (LPGs) are often used to generate synthetic demand profiles in the absence of real data. However, the implications of the assumptions used in LPGs on the results of energy models are not fully understood. Although Hernandez et al. 2020 [7] analysed the effects of data sampling frequency on load profiles, further research is required to understand the broader impacts. Focusing on models like residential systems offers detailed insights. The primary concern is ensuring the accuracy of energy models, as inaccurate assumptions can lead to misaligned results, affecting system design and operation.

In response to this challenge, the research examined the characteristics of various LPG tools, compared synthetic demand profiles to historical data, analysed the impact of LPG input assumptions on energy system model outcomes, and sought to determine optimal practices for using LPGs effectively.

5.2 LPG Biases

In the analysis, demandlib exhibited very conservative simulation. Compared to the actual data, as well as the other simulations, the clustered load profiles had little variation across weekend/weekdays and across the seasons. These characteristics are mainly due to the fact that demandlib is a tool that scales the H0 load profile which is a standard set of load profiles that is scaled to fit the model requirements. In the context of EU households with weather patterns similar to German weather patterns, and behaviours similar to German household behaviours. Demandlib provides relatively accurate results as can be seen from the optimization results.This however could not hold true for other regions were behaviour is different as the H0 load profile might not be an accurate representative of other residential household load profiles.

Ukdale and Demandlib also both had a recommended PV system size of 2.3 kW. Furthermore, the grid import error of demandlib was only 5%. This could be attributed to the accuracy in annual energy expenditure as well as the accurate shape characteristics. Nevertheless since demandlib's load profile is relatively flat and even, it highly underestimates the required energy storage sizing since there are no peaks to require the flexibility provided by energy storage. Moreover, the stability of the load profile across the seasons also means that the batteries can be used more efficiently as demonstrated by the charge duration results.

Districtgen on the other hand exhibited inconsistencies in the shape and peaks. This can be mainly attributed to its stochastic nature. This also means that every new simulation will yield different results, albeit with the same inconsistencies and wild peaks. This means that using the tool will lead to overestimates in investments and costs. In the design optimization analysis, district generator required the biggest storage capacity. This is required to cover the big evening peaks. Furthermore, Districtgenerator had the highest requirement of grid imports to cover the excess peaks that solar could not cover especially the ones that happened at midnight. This also means that for the same energy demand, Districtgenerator will underestimate the required PV system size as it overestimates the demand in the summer when sun is plenty, underestimates the winter requirements, and creates peaks and demand in hours of the day when there is no sun hence also leading to an overestimate in storage investments.

Load Profile Generator	**Characteristics of Load Profile**	**Characteristics of Energy System Optimization**
Demandlib	- Conservative values - Mild peaks - Consistent shape - Exact total annual energy demand - Highest demand in summer	- Underestimation of storage investment - Efficient use of storage - Accurate PV investment
Districtgen	- Inconsistent shape - Wild Peaks - Generated stochastically - High frequency of high power loads	- Overestimation of storage investment - Overestimation of grid imports
LoadProfileGenerator	- Overestimation of midday peaks - large number of parameters provides flexibility in load profile design	- Overestimation of total investment
SynPro	- Mild peaks - High seasonal variations - Occupant parameter has	- Optimization results depend on choice of occupant

		significant impact on total annual energy demand	

Table 5.1 A summary of the biases found in each tool and their effect on the result of the energy system optimization.

In both instances of the simulation, LoadProfileGenerator's overestimated the midday peaks which led to a huge PV capacity requirement in the optimization results since a lot of PV is needed to cover the midday peak. The high midday peak and large PV system also led to the highest energy curtailment. This assumption is true for any load profile with big morning or midday peaks. It is important to note however that LoadProfileGenerator's simulation model has many parameters, and the load profile can be engineered by introducing/removing appliances, usage limits, and occupant behaviour. Nevertheless, in a real situation where the researcher usually does not have all the details of a household, or if modelling multiple households is required, then creating the simulation model is either impossible or at the very least time consuming. It is also important to note that any mistakes in modelling will introduce different biases in the load profile and despite the time investment required, stay-at-home be very inaccurate as was the case in the models simulated for UKdale and WPuQ where random garden appliances that were included in the simulation and a stay-at-home occupant usage of these appliances led to huge midday peaks and huge PV investments in the optimization results.

SynPRO's results depended very heavily on the number of occupants. The tool itself has limited parameters for electric profile measurements. So, each of these parameters can have a huge effect on the simulated load profile. In the results, modelling a house with two occupants produced a total annual energy demand of 2165.5 kWh. However, modelling a house with 4 occupants produced an energy of 3581 kWh. In the first case, synPRO was very inaccurate and underestimated the energy demand. In the second case, it was accurate and even in the monthly energy demand analysis, it had the lowest MAPE of 8.2%. However, it is important to mention that synPRO was the only tool to assume higher electricity expenditure in the winter compared to the summer. And although in the UKdale case that was study, this was the correct assumption, some of the houses in the WPuQ dataset exhibited the opposite seasonality. Nevertheless, synPRO's design optimization results were very accurate and close to UKdale's results. The load duration curve produced was also accurate and closely related to the actual data. The details of the biases are summarised in table 5.1.

5.3 Best Practices

The following section elucidates best practices derived from a comprehensive analysis of various load profile generation tools. These guidelines aim to aid researchers in achieving unbiased energy model optimization results and ensuring that simulated load profiles closely mirror real-world scenarios.

Parameter Selection: Choose tools that offer flexibility in parameter selection, enabling accurate representation of the specific scenario under study. The time

Regional Compatibility: Ensure that the foundational dataset of the tool (such as H0 for Demandlib) is in alignment with the behavioural and weather patterns of the region in question. Other tools, like RAMP, assert their capability to simulate load profiles in diverse contexts.

Stochasticity Awareness: For tools like Districtgen, be cognizant of the inherent unpredictability and its implications on results. Some tool rely on a stochastic methodology to produce load profiles. While this might provide non-representative load profiles, they might be more appropriate to use for aggregate households and communities as they will avoid being too specific.

Avoid Overfitting: While it's essential to align simulations closely with real-world data, be wary of overfitting. A model that is too tightly fit to a specific dataset might not generalize well to other similar scenarios.

Seasonal Analysis: Account for seasonal variations in energy consumption, as they can significantly influence PV and storage investment estimations.

Sensitivity Analysis: When using a new tool or creating a new model, it might be useful to perform a quick sensitivity analysis to understand how each parameter could affect the resultant load profile or the overall result of an energy model optimization.

5.4 Next Steps

The field of load profile generation, particularly concerning the tools examined in this research, is still emerging. This highlights a significant gap in current academic literature and research.

A logical progression for further research would be to extend the application of load profile generators beyond the residential sector. There is potential in exploring the intricacies of energy consumption within the agricultural, industrial, and electric vehicle sectors. Moreover, the study of load profiles for heat, as opposed to electricity, presents another avenue for investigation, especially in regions where heating demands are predominant.

Additionally, the outcomes of energy model optimization are influenced by a myriad of parameters and inputs. While this study addressed certain aspects, there is a vast expanse of variables yet to be studied in relation to the latest load profile generation tools. Subsequent research could delve deeper into how different parameters and inputs, when integrated with these tools, affect the results of energy model optimization. This would further enhance the accuracy and reliability of energy models in diverse contexts.

References

[1] S. Pfenninger, A. Hawkes, and J. Keirstead, “Energy systems modeling for twenty-first century energy challenges,” *Renew. Sustain. Energy Rev.*, vol. 33, pp. 74–86, May 2014, doi: 10.1016/j.rser.2014.02.003.

[2] D. Huppmann *et al.*, “The MESSAGE\(_\mbox\emphix\) Integrated Assessment Model and the \emphix modeling platform (ixmp): An open framework for integrated and cross-cutting analysis of energy, climate, the environment, and sustainable development,” *Env. Model Softw*, vol. 112, pp. 143–156, 2019, doi: 10.1016/J.ENVSOFT.2018.11.012.

[3] E. A. Martínez Ceseña *et al.*, “Quantifying the Impacts of Modelling Assumptions on Accuracy and Computational Efficiency for Integrated Water-Energy System Simulations Under Uncertain Climate,” *IEEE Trans. Sustain. Energy*, vol. 13, no. 3, pp. 1370–1382, Jul. 2022, doi: 10.1109/TSTE.2022.3155073.

[4] A. Almuhtady, A. Alshwawra, M. Alfaouri, W. Al-Kouz, and I. Al-Hinti, “Investigation of the trends of electricity demands in Jordan and its susceptibility to the ambient air temperature towards sustainable electricity generation,” *Energy Sustain. Soc.*, vol. 9, no. 1, p. 39, Dec. 2019, doi: 10.1186/s13705-019-0224-1.

[5] H. Kazmi, Í. Munné-Collado, F. Mehmood, T. A. Syed, and J. Driesen, “Towards data-driven energy communities: A review of open-source datasets, models and tools,” *Renew. Sustain. Energy Rev.*, vol. 148, p. 111290, Sep. 2021, doi: 10.1016/j.rser.2021.111290.

[6] E. Proedrou, “A Comprehensive Review of Residential Electricity Load Profile Models,” *IEEE Access*, vol. 9, pp. 12114–12133, 2021, doi: 10.1109/ACCESS.2021.3050074.

[7] J. C. Hernandez, F. Sanchez-Sutil, A. Cano-Ortega, and C. R. Baier, “Influence of Data Sampling Frequency on Household Consumption Load Profile Features: A Case Study in Spain,” *Sensors*, vol. 20, no. 21, p. 6034, Oct. 2020, doi: 10.3390/s20216034.

[8] C. Büttner, J. Amme, J. Endres, A. Malla, B. Schachler, and I. Cußmann, “Open modeling of electricity and heat demand curves for all residential buildings in Germany,” *Energy Inform.*, vol. 5, no. S1, p. 21, Sep. 2022, doi: 10.1186/s42162-022-00201-y.

[9] “Overview — demandlib 0.1.10 documentation.” Accessed: Oct. 24, 2023. [Online]. Available: https://demandlib.readthedocs.io/en/latest/readme.html

[10] M. Hellwig, “Enwicklung und Anwendung parametrisierter Standard-Lastprofile,” Jan. 2003.

[11] “1999_Repraesentative-VDEW-Lastprofile.pdf.” Accessed: Oct. 24, 2023. [Online]. Available: https://www.bdew.de/media/documents/1999_Repraesentative-VDEW-Lastprofile.pdf

[12] P. Schönfeldt, “demandlib.” oemof community, Sep. 18, 2023. Accessed: Oct. 24, 2023. [Online]. Available: https://github.com/oemof/demandlib

[13] “RWTH-EBC/districtgenerator: Tool for demand profile generation in districts.” Accessed: Oct. 24, 2023. [Online]. Available: https://github.com/RWTH-EBC/districtgenerator

[14] “TEASER - Tool for Energy Analysis and Simulation for Efficient Retrofit.” RWTH Aachen University - E.ON Energy Research Center - Institute for Energy Efficient Buildings and Indoor Climate, Oct. 12, 2023. Accessed: Oct. 24, 2023. [Online]. Available: https://github.com/RWTH-EBC/TEASER

[15] “richardsonpy.” RWTH Aachen University - E.ON Energy Research Center - Institute for Energy Efficient Buildings and Indoor Climate, Jul. 07, 2023. Accessed: Oct. 24, 2023. [Online]. Available:

https://github.com/RWTH-EBC/richardsonpy

[16] T. Loga, B. Stein, and N. Diefenbach, “TABULA Building Typologies in 20 European countries − making energy-related features of residential building stocks comparable,” *Energy Build.*, vol. 132, Jul. 2016, doi: 10.1016/j.enbuild.2016.06.094.

[17] A. Malhotra, J. Frisch, and C. Treeck, *Parametric Study of TEASER*. 2018. doi: 10.18154/RWTH-2018-230338.

[18] I. Richardson, M. Thomson, D. Infield, and C. Clifford, “Domestic electricity use: A high-resolution energy demand model,” *Energy Build.*, vol. 42, no. 10, pp. 1878–1887, Oct. 2010, doi: 10.1016/j.enbuild.2010.05.023.

[19] “LoadProfileGenerator.” Accessed: Oct. 24, 2023. [Online]. Available: https://www.loadprofilegenerator.de//

[20] N. Pflugradt, “Modellierung von Wasser und Energieverbräuchen in Haushalten,” 2016.

[21] C. Pinl, “Wo bleibt die Zeit? Die Zeitbudgeterhebung 2001/02 des Statistischen Bundesamts.,” 2004, Accessed: Oct. 24, 2023. [Online]. Available: https://orlis.difu.de/handle/difu/265376

[22] “Erhebung Wo im Haushalt bleibt der Strom? Stromverbrauchsanteile verschiedener Anwendungsbereiche in Ein- bis Fünf-Personen-Haushalten 2015 und PDF Free Download.” Accessed: Oct. 24, 2023. [Online]. Available: https://docplayer.org/10340556-Erhebung-wo-im-haushalt-bleibt-der-strom-stromverbrauchsanteile-verschiedener-anwendungsbereiche-in-ein-bis-fuenf-personen-haushalten-2015-und-2011.html

[23] F. ISE, “SynPRO Home.” Accessed: Oct. 25, 2023. [Online]. Available: https://synpro-lastprofile.de/

[24] D. Fischer, A. Härtl, and B. Wille-Haussmann, “Model for electric load profiles with high time resolution for German households,” *Energy Build.*, vol. 92, pp. 170–179, Apr. 2015, doi: 10.1016/j.enbuild.2015.01.058.

[25] H. Hersbach *et al.*, “ERA5 hourly data on single levels from 1979 to present,” *Copernic. Clim. Change Serv. C3s Clim. Data Store Cds*, vol. 10, no. 10.24381, 2018.

[26] D. Krähenmann, “Ortsgenaue Testreferenzjahre von Deutschland für mittlere, extreme und zukünftige Witterungsverhältnisse.” 2017. [Online]. Available: https://www.bbsr.bund.de/BBSR/DE/forschung/programme/zb/Auftragsforschung/5EnergieKlimaBauen/2013/testreferenzjahre/try-handbuch.pdf;jsessionid=32F96E889EBF9043C884BC440CF32493.live21301?__blob=publicationFile&v=1

[27] S. I. A. Merkblatt, “Raumnutzungsdaten für die Energie-und Gebäudetechnik,” *SIA Zür. Switz.*, 2015.

[28] B. Bagherian, “Erweiterung und Verbesserung der Bedienbarkeit einer Berechnungsmethode (Teilenergiekennwert-Methode) für die energetische Bilanzierung von Nicht-wohngebäuden im Bestand.” Accessed: Oct. 25, 2023. [Online]. Available: https://www.iwu.de/forschung/energie/tek2go/

[29] European Commission. Statistical Office of the European Union., *Harmonised European Time Use Surveys: 2018 guidelines : re edition, 2020 edition.* LU: Publications Office, 2020. Accessed: Oct. 25, 2023. [Online]. Available: https://data.europa.eu/doi/10.2785/160444

[30] “[PDF] Intelliekon. INTELLIEKON Projektverbund: Ergebnisse, Fazits, Interpretationen und Folgerungen aus dem INTELLIEKON Feldversuch - Free Download PDF.” Accessed: Oct. 25, 2023. [Online]. Available: https://silo.tips/download/intelliekon-intelliekon-projektverbund-ergebnisse-fazits-interpretationen-und-fo

[31] “A modular open source framework to model energy supply systems.” Accessed: Oct. 24, 2023. [Online]. Available: https://oemof.org/

[32] “oemof community.” Accessed: Oct. 24, 2023. [Online]. Available: https://github.com/oemof

[33] J. Nagel, *Optimization of Energy Supply Systems: Modelling, Programming and Analysis*, vol. 69. in Lecture Notes in Energy, vol. 69. Cham: Springer International Publishing, 2019. doi: 10.1007/978-3-319-96355-6.

[34] "design optimization — ESyOpT 30.01.2020 documentation." Accessed: Oct. 24, 2023. [Online]. Available: https://ob220.pages.fraunhofer.de/esyopt/Modell/theoretical_background/design_opt.html

[35] P. Price, "Methods for analyzing electric load shape and its variability," Lawrence Berkeley National Lab.(LBNL), Berkeley, CA (United States), 2010. Accessed: Oct. 27, 2023. [Online]. Available: https://www.osti.gov/biblio/985909

[36] H. Li, Z. Wang, T. Hong, A. Parker, and M. Neukomm, "Characterizing patterns and variability of building electric load profiles in time and frequency domains," *Appl. Energy*, vol. 291, p. 116721, 2021.

[37] L. Chuan and A. Ukil, "Modeling and Validation of Electrical Load Profiling in Residential Buildings in Singapore".

[38] G. Hoogsteen, "A cyber-physical systems perspective on decentralized energy management," 2017, Accessed: Oct. 27, 2023. [Online]. Available: https://research.utwente.nl/files/18822924/Hoogsteen_A_Cyber_Physical_Systems_Perspective_on_Decentralized_Energy_Management.pdf

[39] "RAMP," RAMP. Accessed: Oct. 27, 2023. [Online]. Available: http://localhost:4000/

[40] "World Bank Open Data," World Bank Open Data. Accessed: Oct. 28, 2023. [Online]. Available: https://data.worldbank.org

[41] "Residential buildings: Energy Efficiency & Consumption evolution in Europe." Accessed: Oct. 28, 2023. [Online]. Available: https://www.enerdata.net/publications/executive-briefing/households-energy-efficiency.html

[42] J. Kelly and W. Knottenbelt, "The UK-DALE dataset, domestic appliance-level electricity demand and whole-house demand from five UK homes," *Sci. Data*, vol. 2, Mar. 2015, doi: 10.1038/sdata.2015.7.

[43] "TABULA WebTool." Accessed: Oct. 28, 2023. [Online]. Available: https://webtool.building-typology.eu/#bd

[44] M. Schlemminger, T. Ohrdes, and E. Schneider, "Dataset on electrical single-family house and heat pump load profiles in Germany.," *URL Www Nat. Comarticless41597-022-01156-1 Cit P 27*, 2022.

[45] T. Kraschewski, T. Brauner, M. Heumann, and M. H. Breitner, "Disentangle the price dispersion of residential solar photovoltaic systems: Evidence from Germany," *Energy Econ.*, vol. 121, p. 106649, May 2023, doi: 10.1016/j.eneco.2023.106649.

[46] C. Kost, S. Shammugam, V. Fluri, D. Peper, A. D. Memar, and T. Schlegl, "Stromgestehungskosten Erneuerbare Energien," 2021.

[47] NREL, "2021 CMTS," Tableau Software. Accessed: Oct. 31, 2023. [Online]. Available: https://public.tableau.com/views/2021CMTS/TechSummary?:embed=y&Technology=Residential%20Battery%20Storage&:embed=y&:showVizHome=n&:bootstrapWhenNotified=y&:apiID=handler0

[48] B. U. Schyska and A. Kies, "How regional differences in cost of capital influence the optimal design of power systems," *Appl. Energy*, vol. 262, p. 114523, Mar. 2020, doi: 10.1016/j.apenergy.2020.114523.

Annexe A

Code Snippets

A.1 LPG simulations

A.1.1 Demand Lib model

Table 0.1 Simulation of house 1 load profile using demandlib

```
# The following dictionary is create by “workalendar”
cal = UnitedKingdom()
holidays = dict(cal.holidays(2016))
print(holidays)

# my note: ann_el_demand is the annual electricity demand in kWh
ann_el_demand_per_sector = {
    “h0”: 3517.79,
}

year = 2016

# read standard load profiles
e_slp = bdew.ElecSlp(year, holidays=holidays)

# multiply given annual demand with timeseries
elec_demand = e_slp.get_profile(ann_el_demand_per_sector)

# Add the slp for the industrial group
ilp = profiles.IndustrialLoadProfile(e_slp.date_time_index, holidays=holidays)


print(
    “Be aware that the values in the DataFrame are 15 minute values”
    + “with a power unit. If you sum up a table with 15min values”
    + “the result will be of the unit ‘kW15minutes’.”
)
print(elec_demand.sum())

print(“You will have to divide the result by 4 to get kWh.”)
print(elec_demand.sum() / 4)

print(“Or resample the DataFrame to hourly values using the mean() “ “method.”)

# elec demand in W
elec_demand = elec_demand*1000

# Resample 15-minute values to hourly values.
Elec_demand_resampled = elec_demand.resample(“H”).mean()
print(elec_demand_resampled.sum())

# Plot demand
ax = elec_demand_resampled.plot()
ax.set_xlabel(“Date”)
ax.set_ylabel(“Power demand”)
# fig size (20, 10)
```

```
plt.rcParams["figure.figsize"] = (20, 10)
plt.show()

print(elec_demand)
```

Table 0.2 Simulation of WpuQ load profile using demandlib

```
# The following dictionary is create by "workalendar"
cal = Germany()
holidays = dict(cal.holidays(2020))
print(holidays)

# my note: ann_el_demand is the annual electricity demand in kWh
ann_el_demand_per_sector = {
    "h0": 2786.66,
}

year = 2020

# read standard load profiles
e_slp = bdew.ElecSlp(year, holidays=holidays)

# multiply given annual demand with timeseries
elec_demand = e_slp.get_profile(ann_el_demand_per_sector)

# Add the slp for the industrial group
ilp = profiles.IndustrialLoadProfile(e_slp.date_time_index, holidays=holidays)


print(
    "Be aware that the values in the DataFrame are 15 minute values"
    + "with a power unit. If you sum up a table with 15min values"
    + "the result will be of the unit 'kW15minutes'."
)
print(elec_demand.sum())

print("You will have to divide the result by 4 to get kWh.")
print(elec_demand.sum() / 4)

print("Or resample the DataFrame to hourly values using the mean() " "method.")

# elec demand in W
elec_demand = elec_demand * 1000

# Resample 15-minute values to hourly values.
Elec_demand_resampled = elec_demand.resample("H").mean()
print(elec_demand_resampled.sum())

# Plot demand
ax = elec_demand_resampled.plot()
ax.set_xlabel("Date")
ax.set_ylabel("Power demand")
# fig size (20, 10)
plt.rcParams["figure.figsize"] = (20, 10)
plt.show()
```

```
print(elec_demand)
```

A.2 Functions

Figure 0.1 Clustering function used for the load profile characterization method

```
def clustering(data):
    # add the hourly cycle to the dataframe
    feature = data.copy()
    feature['hour'] = feature.index.hour

    # add the season cycle to the dataframe
    feature['season'] = pd.cut(feature.index.month, [0, 2, 5, 8, 11, 12],
labels=['Winter', 'Spring/Fall', 'Summer', 'Spring/Fall', 'Winter'],
ordered=False)

    # add the daily cycle to the dataframe
    feature['day'] = feature.index.dayofweek

    # group the houses of the week into weekdays and weekends
    feature['day'] = feature['day'].replace([0, 1, 2, 3, 4], 'Weekday')
    feature['day'] = feature['day'].replace([5, 6], 'Weekend')

    # pivot the dataframe to create a column for each season and
weekend/weekday combination
    cluster = feature.pivot_table(index='hour', columns=['season', 'day'],
values='demand_el', aggfunc=np.mean)

    return cluster
```

A.3 Optimization Model

```
# initialize the energy system
date_time_index = solph.create_time_index(2016, number=number_timesteps)
energysystem = solph.EnergySystem(timeindex=date_time_index,
infer_last_interval=False)

# create the OEMOF objects and add them to the energy system

# create electricity bus
bel = solph.Bus(label="electricity")

energysystem.add(bel)
```

```
# create fixed source object representing electricity grid feed-in
grid = solph.components.Source(
    label="grid",
    outputs={bel: solph.Flow(variable_costs=0.3)},
)

# create simple sink object representing the electrical demand
demand = solph.components.Sink(
    label="demand",
    inputs={bel: solph.Flow(fix=data["demand_el"], nominal_value=1)},
)

# create excess component for the electricity bus to allow overproduction
excess = solph.components.Sink(label="excess_bel", inputs={bel:
solph.Flow()})

# create pv esyopt component
pv_esyopt = PVPlant(
label = 'pv_esyopt',
el_output = {bel: solph.Flow()},
c_FIX_ANNU=26,
normalized_output= data['pv'],
design_opt=True,
CAPEX_PV=1392.13,
LIFETIME_PV=30,
CAPEX_INR=0,
LIFETIME_INR=10,
WACC=0.034,
t_OPT=1,
c_INV_FIX=0,
P_MAX_UPPER=100,
P_MAX_LOWER=0,
A_kWp=10,
A_MAX=300)

# create storage esyopt component
storage_esyopt= BatteryStorage(
label='storage_esyopt',
el_inputs={bel: solph.Flow()},
el_outputs={bel: solph.Flow()},
soc_0=0,
soc_MAX=1,
eta_STO=0.99,
eta_IN=0.90,
eta_OUT=0.90,
kappa_P_E=0.5,
LT_CYC=8000,
t_OPT=1,
```

```
c_INR_FIX_ANNU=0,
c_BATT_FIX_ANNU=0,
design_opt=True,
CAPEX_BATT=500,
CAPEX_INR=0,
LIFETIME=15,
WACC=0.034,
E_MAX_UPPER=500,
P_MAX_UPPER=500,
E_MAX_LOWER=0,
P_MAX_LOWER=0,
c_INR_INV_FIX=0,
c_BATT_INV_FIX=0)

# don't forget to add all the components to the energy system.
energysystem.add(excess, pv_esyopt, demand, storage_esyopt, grid)

# optimize the energy system

om = solph.Model(energysystem)

om.solve(solver="cbc")
```

Annexe B

Visual Analysis Power Tables

season	*Spring/Fall*		*Summer*		*Winter*	
day	Weekday	Weekend	Weekday	Weekend	Weekday	Weekend
hour						
0	0.251	0.341	0.223	0.263	0.312	0.323
1	0.219	0.252	0.218	0.234	0.211	0.267
2	0.211	0.239	0.218	0.211	0.202	0.220
3	0.214	0.212	0.214	0.212	0.205	0.199
4	0.228	0.231	0.239	0.232	0.215	0.214
5	0.226	0.218	0.221	0.213	0.213	0.214
6	0.390	0.256	0.342	0.285	0.224	0.199
7	0.501	0.390	0.444	0.350	0.475	0.259
8	0.476	0.441	0.480	0.436	0.500	0.479
9	0.499	0.419	0.376	0.415	0.471	0.576
10	0.448	0.462	0.332	0.318	0.455	0.598
11	0.393	0.475	0.371	0.387	0.544	0.502
12	0.420	0.439	0.347	0.432	0.565	0.518
13	0.413	0.380	0.346	0.318	0.575	0.417
14	0.416	0.385	0.371	0.344	0.478	0.375
15	0.407	0.381	0.365	0.379	0.358	0.482
16	0.490	0.465	0.424	0.455	0.467	0.543
17	0.527	0.578	0.447	0.460	0.582	0.557
18	0.552	0.579	0.455	0.520	0.610	0.501
19	0.600	0.524	0.426	0.467	0.641	0.559
20	0.512	0.475	0.488	0.471	0.608	0.498
21	0.522	0.474	0.447	0.481	0.518	0.443
22	0.549	0.518	0.451	0.515	0.522	0.464
23	0.394	0.422	0.273	0.307	0.560	0.526

Table 0.3 UKdale cluster load profile of house 1

season	*Spring/Fall*		*Summer*		*Winter*	
day	Weekday	Weekend	Weekday	Weekend	Weekday	Weekend
hour						
0	0.121	0.131	0.119	0.132	0.114	0.114
1	0.119	0.110	0.122	0.105	0.115	0.115
2	0.126	0.123	0.116	0.104	0.115	0.118
3	0.118	0.117	0.118	0.109	0.124	0.116
4	0.120	0.131	0.116	0.120	0.108	0.113
5	0.128	0.122	0.125	0.133	0.122	0.116
6	0.490	0.452	0.453	0.602	0.484	0.601
7	0.560	0.569	0.624	0.645	0.675	0.606
8	0.545	0.700	0.565	0.788	0.673	0.898
9	0.747	0.939	0.794	0.870	0.771	1.105
10	0.634	1.007	0.833	1.053	0.525	0.842
11	1.392	1.275	2.159	2.026	1.242	1.536
12	0.715	0.791	0.645	0.798	0.568	0.766
13	0.521	0.625	0.410	0.290	0.704	0.540

14	0.333	0.500	0.356	0.468	0.277	0.566
15	0.359	0.409	0.428	0.492	0.332	0.524
16	0.308	0.481	0.368	0.543	0.394	0.407
17	0.404	0.383	0.422	0.303	0.438	0.400
18	0.638	0.557	0.928	0.492	0.628	0.557
19	0.628	0.685	0.862	0.571	0.646	0.722
20	0.528	0.541	0.581	0.449	0.525	0.554
21	0.439	0.402	0.456	0.386	0.464	0.469
22	0.318	0.322	0.345	0.306	0.318	0.336
23	0.168	0.183	0.188	0.168	0.155	0.158

Table 0.4 LoadProfileGenerator cluster load profile of house 1 simulation

season	*Spring/Fall*		*Summer*		*Winter*	
day	Weekday	Weekend	Weekday	Weekend	Weekday	Weekend
hour						
0	0.232	0.275	0.259	0.303	0.208	0.254
1	0.172	0.207	0.193	0.230	0.154	0.191
2	0.155	0.167	0.175	0.187	0.140	0.153
3	0.149	0.155	0.165	0.177	0.136	0.142
4	0.153	0.151	0.171	0.176	0.137	0.135
5	0.182	0.154	0.207	0.176	0.163	0.138
6	0.321	0.182	0.341	0.202	0.308	0.166
7	0.448	0.267	0.452	0.285	0.441	0.234
8	0.478	0.408	0.496	0.423	0.461	0.368
9	0.472	0.532	0.507	0.555	0.435	0.492
10	0.458	0.600	0.491	0.617	0.418	0.564
11	0.466	0.653	0.502	0.662	0.429	0.625
12	0.518	0.676	0.557	0.687	0.471	0.653
13	0.502	0.606	0.537	0.617	0.462	0.594
14	0.438	0.520	0.463	0.529	0.408	0.509
15	0.394	0.470	0.417	0.479	0.370	0.459
16	0.377	0.444	0.403	0.445	0.370	0.444
17	0.423	0.489	0.429	0.454	0.451	0.541
18	0.527	0.576	0.498	0.515	0.582	0.646
19	0.613	0.635	0.579	0.584	0.655	0.681
20	0.588	0.586	0.584	0.583	0.589	0.584
21	0.533	0.505	0.553	0.534	0.494	0.473
22	0.462	0.461	0.507	0.507	0.409	0.417
23	0.345	0.366	0.393	0.409	0.303	0.332

Table 0.5 Demandlib cluster load profile of house 1 simulation

season	*Spring/Fall*		*Summer*		*Winter*	

day	Weekday	Weekend	Weekday	Weekend	Weekday	Weekend
hour						
0	0.060	0.066	0.060	0.058	0.059	0.056
1	0.138	0.174	0.114	0.245	0.149	0.296
2	0.062	0.060	0.064	0.062	0.082	0.065
3	0.071	0.056	0.103	0.059	0.072	0.070
4	0.075	0.056	0.060	0.056	0.060	0.056
5	0.085	0.060	0.056	0.070	0.056	0.248
6	0.137	0.251	0.206	0.584	0.180	0.084
7	0.331	0.441	0.322	0.430	0.340	0.521
8	0.651	0.895	0.639	0.524	0.616	0.521
9	0.456	0.450	0.457	0.547	0.680	0.536
10	0.547	0.422	0.537	0.793	0.654	0.392
11	0.480	0.483	0.636	0.269	0.472	0.328
12	0.538	0.367	0.626	0.301	0.500	0.278
13	0.477	0.320	0.365	0.559	0.501	0.411
14	0.406	0.373	0.430	0.485	0.361	0.306
15	0.407	0.530	0.435	0.385	0.520	0.373
16	0.536	0.587	0.549	0.441	0.889	0.794
17	0.639	0.543	0.503	0.644	0.687	0.624
18	0.585	0.674	0.634	0.662	0.515	0.579
19	0.795	0.615	0.850	0.919	0.706	0.784
20	0.774	0.945	0.657	0.889	0.795	0.994
21	0.747	0.931	0.910	0.851	0.981	1.016
22	0.873	0.876	1.002	0.676	0.658	0.908
23	0.570	0.690	0.735	0.589	0.527	0.774

Table 0.6 Districtgenerator cluster load profile of house 1 simulation