UCLouvain

epl

École polytechnique de Louvain

Integration of LCA into Renewable Energy Hub Optimizer (REHO)

Multi-crilteria analysis based on LCA of a positive energy district accounting for a country context - application to real case study

Author: **Zhichuan MA**
Supervisors: **Hervé JEANMART, Gauthier LIMPENS**
Academic year 2023–2024
Master of Science and Technology, STEEM program, École polytechnique

Abstract

Achieving net-zero emissions requires a significant expansion in the deployment of clean technologies compared to current levels. These renewable technologies play a critical role in modern energy systems, increasing flexibility and better meeting end-use demands. Accurately assessing the potential of these energy systems necessitates a comprehensive understanding of their costs and environmental impacts. This research addresses this need by integrating Life Cycle Assessment (LCA) methodologies with energy system optimization models, creating a robust framework that evaluates both environmental impact and cost. A more generalized LCA indicator set was developed, and a refined database, free of double-counting, was implemented. The new methodology and database were validated through a comparative case study and further applied to a real-world scenario in Spain.

The analysis highlights the improvements in the LCA framework, where correlation analysis and multi-objective optimization using K-means clustering enabled the selection of representative indicators, simplifying normalization in optimization tasks. This integrated framework, implemented through Python scripts and AMPL models, provides key insights into the environmental impact of energy systems. Its application in the case study for a community in Spain offers valuable guidance for decarbonization strategies, contributing to more sustainable energy solutions.

Table of Contents

List of Figures

List of Tables

1 Introduction

1.1 Background

1.1.1 Energy system

In response to climate change and its impact to all kinds of industries, especially through increasingly frequent extreme weather events[1][2], energy systems are evolving to become more adaptive and flexible[3]. With international frameworks like the Paris Agreement pushing for significant reductions in global greenhouse gas emissions[4], there is a growing demand for energy systems that align with these sustainability goals. Traditional energy systems, which were primarily evaluated based on economic efficiency and profit, are now being reimagined to balance environmental sustainability, resilience, and cost-effectiveness[5][6]. Modern energy systems emphasize not only producing energy but also managing it in a way that minimizes environmental impacts, ensures resilience against unpredictable conditions, and incorporates renewable energy sources. This shift underscores the need for a comprehensive approach that integrates advanced technologies, flexible energy distribution, and sustainable practices to meet the demands of a changing world and the commitments outlined in the Paris Agreement.

1.1.2 LCA

Life Cycle Assessment (LCA) is a widely used methodology for assessing the environmental impacts associated with various clean technologies. It encompasses a broad range of factors, including global warming potential (GWP), ozone depletion, particulate matter emissions, and more. LCA methods typically involve the detailed tracking of material and energy flows entering and leaving the environment throughout a product's life cycle. By quantifying these flows using comprehensive data, the overall environmental impact of a technology or process can be accurately assessed. This holistic approach enables more informed decision-making, promoting the adoption of technologies that minimize environmental footprints.

The process of conducting an LCA generally follows four key stages: (1) defining the goal and scope of the assessment, (2) compiling a life cycle inventory of all relevant material and energy inputs and outputs, (3) performing a life cycle impact assessment by evaluating potential environmental impacts, and (4) interpreting the results to inform improvements or decision-making. This systematic approach ensures that environmental impacts are considered across all phases of a product's life, from raw material extraction to production, use, and disposal[7].

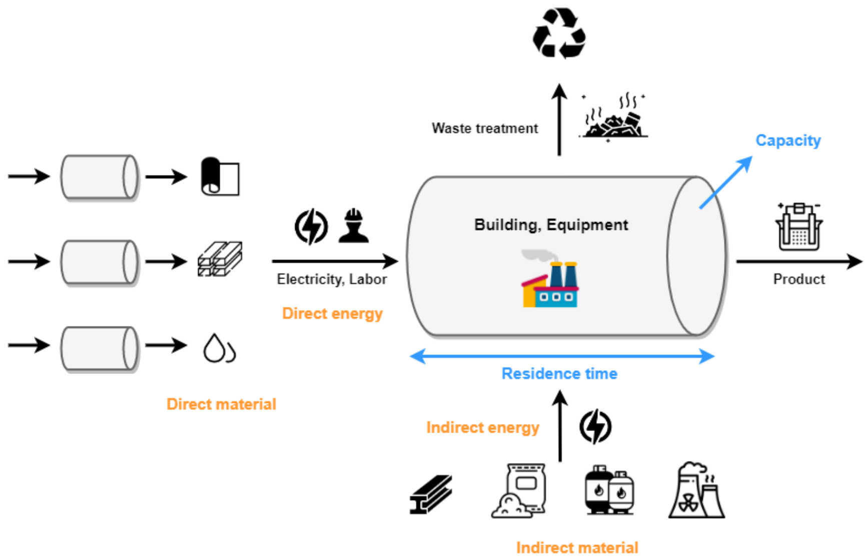

Figure 1: Overview of the Life Cycle Assessment (LCA) Mechanism

The production process serves as a conduit for calculating the environmental impacts of a product. The incoming flows include both direct materials and direct energy, along with indirect energy and material inputs. These flows, and their associated environmental impacts, contribute to the overall environmental footprint of the product. In essence, the LCA of a product is the cumulative sum of the environmental impacts from all relevant exchange flows associated with its production.

The popularity of LCA has grown since 1990, with increased environmental awareness [8]. However, LCA faced criticism during its early development due to high expectations. Over the years, there has been substantial progress and refinement in LCA standards and practices. In 2008, Jørgensen et al. [9] conducted a comprehensive LCA technology review. Since then, LCA's applications have expanded into various areas, including waste management, technology assessment, energy sector decision-making, product system improvement and energy system optimization[10][11].

1.2 REHO - Renewable energy hub optimizer

REHO, the abbreviation for the Renewable Energy Hub Optimizer, is an open-source tool developed by EPFL. This model integrates multiple technologies to address district-level energy optimization challenges. The problem is decomposed into several sub-problems (individual house optimization) and a master problem (district optimization), as illustrated in Figure 2. The optimization starts at the individual house level, generating several potential configurations for each house. These configurations are then passed to the district-level optimization, where all possible combinations of house configurations, along with district-specific units, are iterated to identify the global optimal configuration for the entire district.

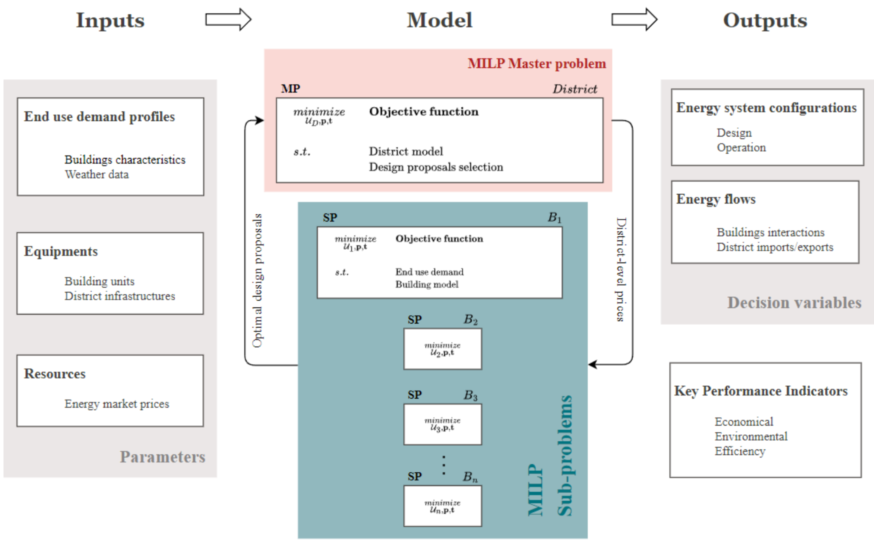

Figure 2: Overview of the Renewable Energy Hub Optimizer (REHO) Mechanism [12]

In this model, given the end-use demand (EUD) profiles derived from meteorological data and building characteristics, and the available resources (such as grids and equipment for converting resources into the required energy services), REHO performs multi-objective optimization across different scenarios. This optimization is powered by AMPL. REHO provides a flexible platform for energy system planning.

1.3 Research gap

The current REHO model incorporates certain aspects of Life-Cycle Assessment (LCA), but significant limitations reveal a clear research gap. One major issue is the quality of the LCA database, which is insufficiently robust, leading to inaccuracies in assessment results. This deficiency is evident in two key areas: firstly, the existing database lacks units for its values, creating ambiguity in the data; secondly, the absence of declared data sources means it is unclear what methods have been employed for the LCA, reducing transparency and reliability. Moreover, the methodological framework lacks the rigor and detail required to produce dependable LCA outcomes, which limits the overall effectiveness of the model.

In addition, the REHO model is expected to be used in conjunction with or compared against another energy model, Energyscope[13], as noted by REHO's developers. This comparison underscores the need for a consistent and standardized LCA methodology to ensure valid comparisons. Addressing these issues calls for the development of a more comprehensive and transparent LCA database, coupled with a more rigorous methodological framework. These improvements are essential for enhancing both the accuracy and credibility of LCA within the model. Furthermore, integrating advanced LCA techniques alongside multi-objective optimization and real case studies for pilot sites will ensure a more holistic and reliable approach to sustainable energy system planning.

2 Methodology

The methodology developed is primarily designed to complete the integration of Life Cycle Assessment (LCA) while also paving the way for future multi-objective optimization. The overall process is divided into three key phases: LCA Database Creation, LCA Methodology Integration, and Normalization. These steps are carried out using a series of toolchains, including Mescal, Brightway2, and Energyscope. In some cases, new concepts were learned, introduced and applied in my own energy system model, while in others, specific tasks were accomplished through the capabilities of these tools. The general framework for this methodology is illustrated in the figure below.

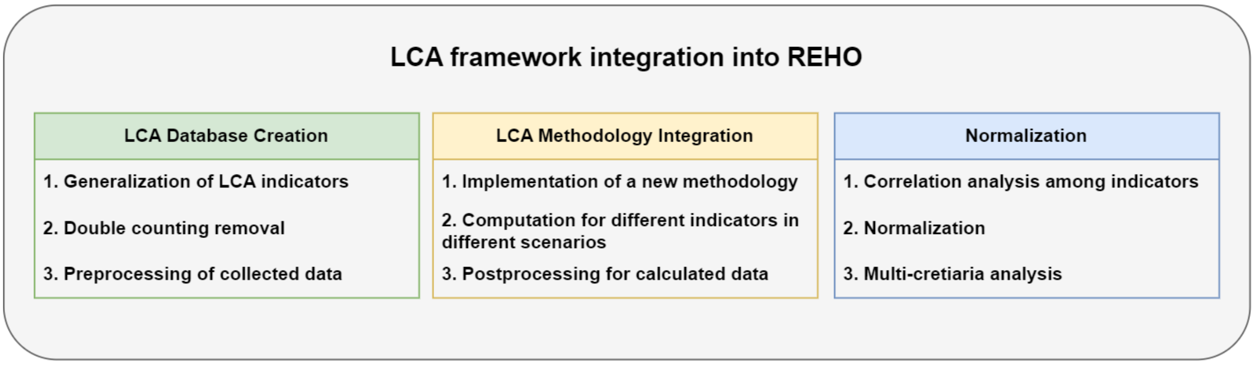

Figure 3: General Methodology for Integrating the Life Cycle Assessment (LCA) Framework into the Renewable Energy Hub Optimizer (REHO)

2.1 Original LCA Methodology in REHO

In the original REHO methodology, lifecycle analysis (LCA) for technologies is divided into two stages: Construction and Operation. The construction phase accounts for the environmental impacts over the entire lifecycle of each technology. Instead of analyzing the operation phase seperately, the total exchange between resources/grids and the technologies/units represents the operation phase. As illustrated in the figure below, the resources and grids are positioned on the left side of the dotted line, while the technologies and units are on the right. The total exchange between these two sides reflects the amount of energy consumed from or returned to the external environment.

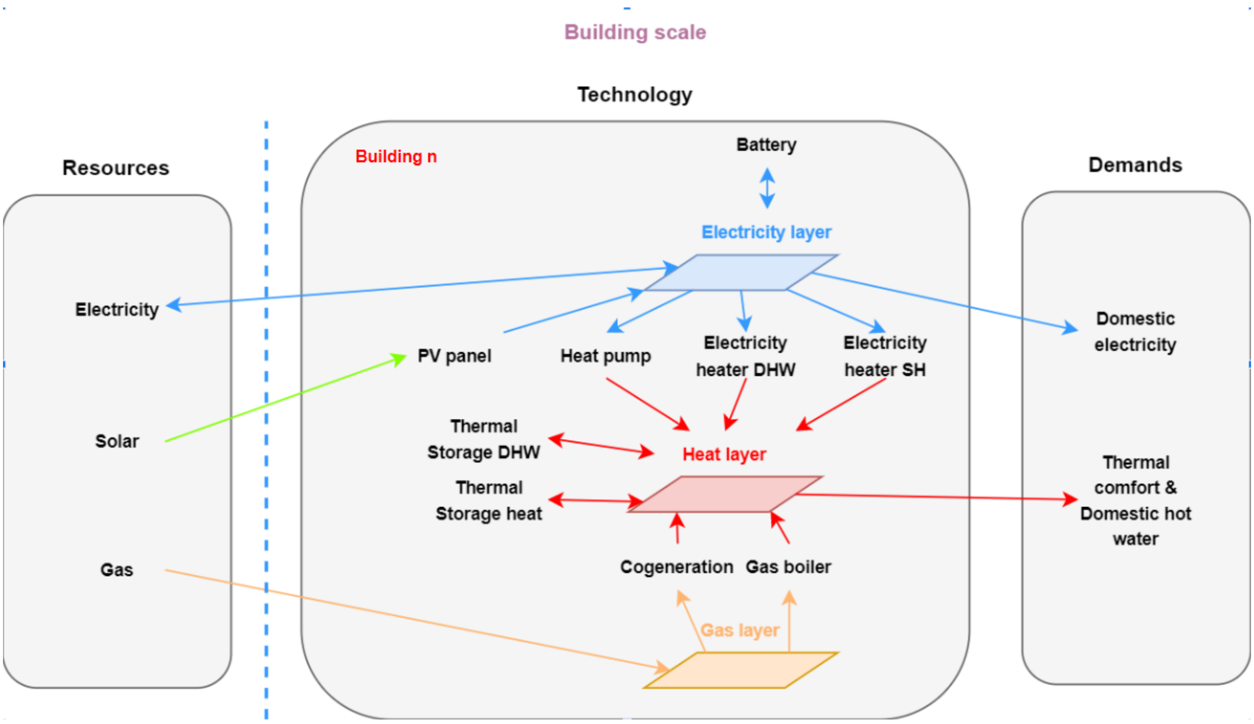

Figure 4: Predefined Energy System Model at Building Scale in the Renewable Energy Hub Optimizer (REHO)

3 layers are represented here(Electricity, Heat and Gas) and 9 technologies allow conversion of quantities between layers.

However, the LCA in the original REHO is not sufficiently comprehensive or convincing for several reasons. In the following, we will examine how the LCA was conducted to identify areas for potential improvement. For the district level, please refer to the Appendix Fig 34

2.1.1 REHO Construction Model

As we all know, environmental impacts can be viewed as another kind of cost, as they share many similar properties with economic costs, particularly in the context of energy systems [14]. The consideration of Renewable Energy Hub Optimization (REHO) for construction life cycle assessment (LCA) is similarly based on the distinction between fixed and variable costs. In cost analysis, fixed costs typically refer to indirect contributions, such as rent, technology repayments, and other overhead expenses. In contrast, variable costs are tied to direct inputs and fluctuate with production volumes, including raw materials [15]. Consequently, the LCA within REHO for construction can be represented as the following calculation:

$$LCA_{constr} = \sum_{u \in Units} \frac{UseofUnits(u) * ei_{fixed}(u) + SizeofUnits(u) * ei_{variable}(u)}{lifetime(u)} \quad (1)$$

where

- LCA_{constr} represents the total construction environmental impacts in the energy system.

- $UseofUnits(u)$ is a binary variable for denoting the use of units; 1 means using while 0 means not using.
- $ei_{fixed}(u)$ is the fixed environmental impact of unit u.
- $ei_{variable}$ is the variable environmental impact of unit u.
- $SizeofUnits(u)$ is the installed capacity for the unit u.
- $lifetime(u)$ is lifetime of unit u.

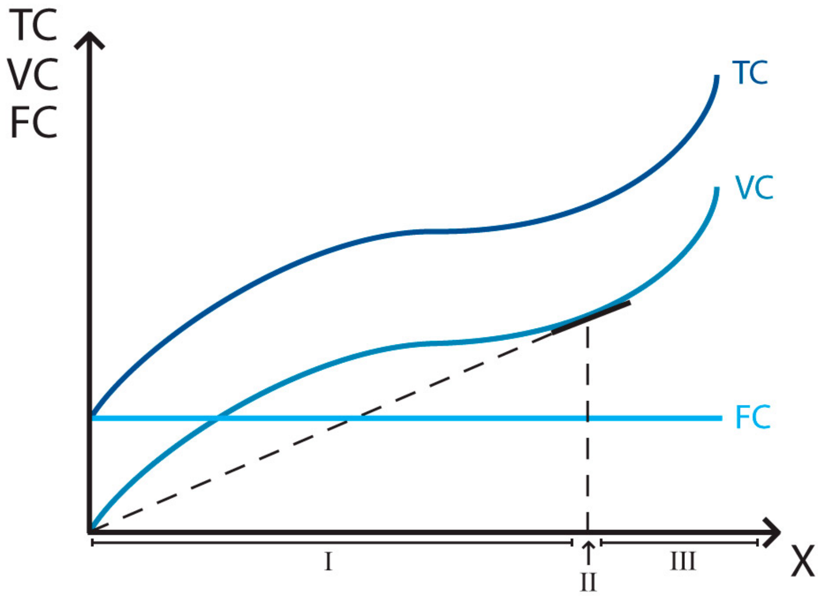

Figure 5: Comparison of Fixed and Variable Costs in Business Operations

In cost analysis, fixed and variable costs can be determined by selecting two points on the cost evolution curve and drawing a secant line in the Cartesian coordinate system. The slope of this secant represents the variable cost, while its intersection with the Y-axis represents the fixed cost, as illustrated in the figure above.

However, in the context of LCA, there is limited literature addressing the scale-up effect, i.e., the relationship between environmental impacts and production volume. Moreover, unlike cost analysis, it is challenging to clearly define fixed and variable environmental impacts. Therefore, the methodology for assessing the environmental impacts of construction requires refinement.

2.1.2 REHO Operation Model

As mentioned before, the LCA for operation in REHO considers the total exchange between the unit side and the grid side. The calculation could be formulated as:

$$EI_{sup}(l,u,t) = ei_{supply}(l,u,t) \cdot Network_{supply}(l,u,t) \tag{2}$$

$$EI_{dem}(l,u,t) = ei_{demand}(l,u,t) \cdot Network_{demand}(l,u,t). \tag{3}$$

$$LCA_{op} = \sum_{l \in \text{ResourceBalances}, u \in \text{units}, t \in \text{Time}} (EI_{sup}(l,u,t) - EI_{dem}(l,u,t)) \tag{4}$$

Where,

- LCA_{op} represents the total operation environmental impacts in the energy system.
- $EI_{sup}(l,u,t)$ is the total environmental impacts of unit u supply to the energy layer$l$ at time t.
- $EI_{dem}(l,u,t)$ is the total environmental impacts of unit u demand from the energy layer$l$ at time t.
- $ei_{supply/demand}$ is the unit environmental impact.
- $Network_{supply}$ is the quantity of energy supply from the network to the system.
- $Network_{demand}$ is the quantity of energy supply to the network from the system.

The Life Cycle Assessment (LCA) for operational performance typically assumes that the system functions like a conduit, where energy flows in, is converted into another form, and flows out without any significant consumption of energy. However, this assumption is unrealistic due to the Second Law of Thermodynamics. In reality, during such processes, entropy will inevitably increase, meaning that energy cannot be fully transformed from one ordered state to another without losses.

For instance, consider the example of a gas boiler. The figure below illustrates the material and energy flow exchanges within the gas boiler operation phase.

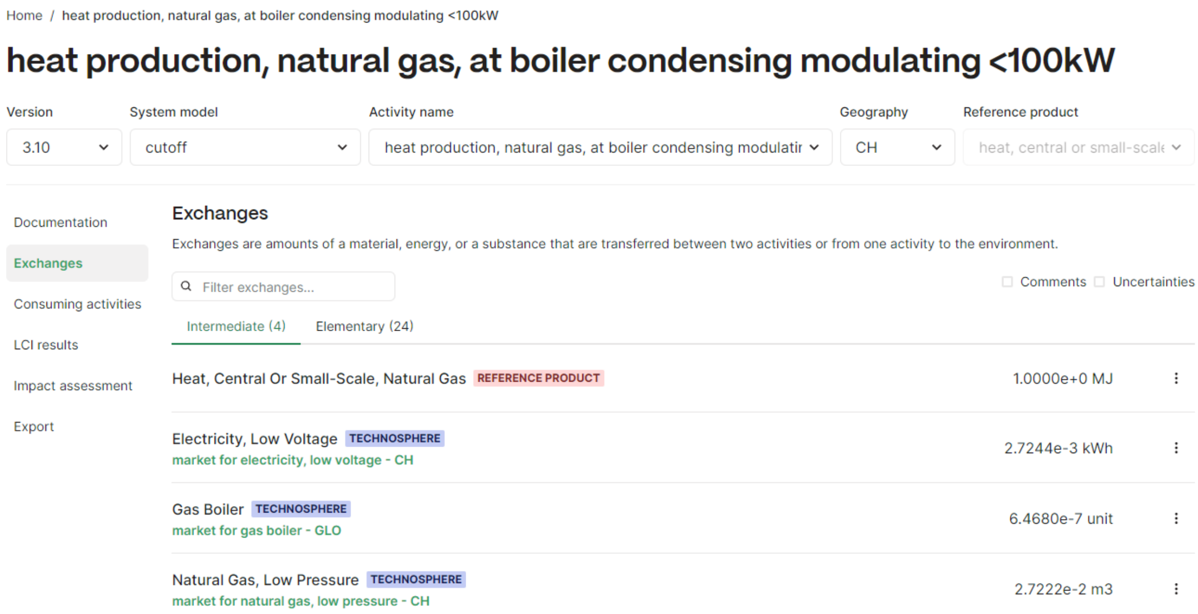

Figure 6: Detailed Exchange Flows for a Natural Gas Boiler in Ecoinvent

In addition to natural gas, which is considered in the operation phase of REHO, and the construction of natural gas boilers, which is accounted for in the construction phase, electricity usage has been neglected in the Life Cycle Assessment (LCA) of REHO (see Figure 3). This emphasizes the need for a new methodology that incorporates electricity consumption into the operation phase, ensuring a more comprehensive assessment.

2.1.3 LCA indicators

There are two main issues with the current LCA datasets in REHO. First, the LCA data being used in REHO lacks credibility, as it is not accompanied by units or standards. Typically, it is essential to declare the specific LCA methodology being employed, such as IPCC 2013 or ReCiPe Midpoint (I), since the evaluation approaches vary significantly between methods. In comparative studies, clearly stating the method used is critical to ensure consistency and reliability.

The second issue is that the indicators used in REHO are limited. As shown in the table below, there are only 13 indicators, which is insufficient to capture the full scope of environmental impacts. Therefore, expanding the range of indicators and improving the generalization of the dataset are key areas that need attention for improvements.

Table 1: Original LCA Indicators and Corresponding Environmental Impacts

Abbreviation	Environmental Impact
GWP	Global Warming Potential
land_use	Land Use
mine_res	Mineral Resource Utilization
water_res	Water Resource Utilization
energy_res	Energy Resource Utilization
human_toxicity	Human Toxicity
water_pollutants	Water Pollutants
metals_water	Metals in Water
pop_water	Persistent Organic Pollutants in Water
metals_soil	Metals in Soil
pollutants_pm	Particulate Matter Pollutants
metal_air	Metals in Air
ozone_depletion	Ozone Layer Depletion

2.2 New Methodology

Due to several shortcomings in REHO's Life Cycle Assessment (LCA), a new LCA methodology has been proposed and integrated into REHO. While Energyscope [13] represents another advanced energy system optimization model, it differs from REHO in scale and application. Energyscope is a novel open-source model designed for strategic energy planning at a national level, whereas REHO focuses on a district scale. Despite its advantages, Energyscope also has limitations, such as a lower level of technico-economic resolution and lack of market equilibrium.

Nevertheless, Energyscope employs a comprehensive LCA methodology for energy systems, which offers valuable insights that can be adapted for use in REHO. Consequently, the LCA methodology from Energyscope has been adapted and implemented into REHO [16].

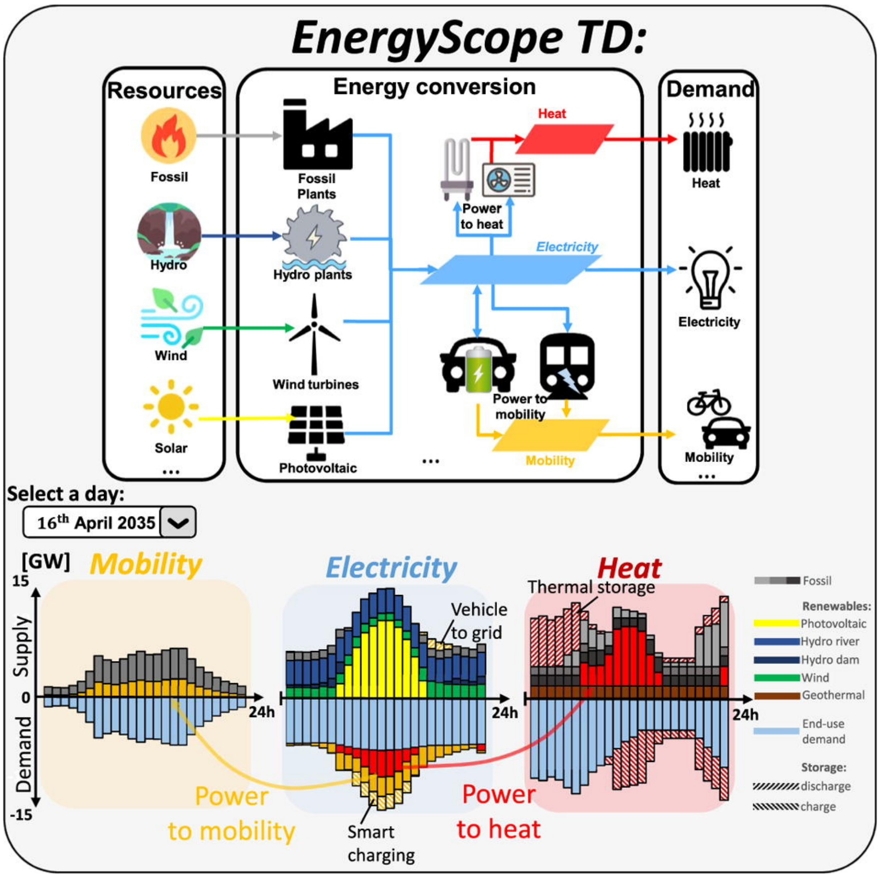

Figure 7: Overview of an energy system in Energyscope.

4 layers are represented here(Electricity, Heat, Passenger Mobility and Gas) and 7 technologies allow conversion of quantities between layers.

2.2.1 LCA Methodology integration

From the perspective of Energyscope, the lifecycle of technologies is categorized into three distinct phases: resource extraction, operation, and construction. This approach extends beyond merely considering the operation and construction of units. The general formulation of this lifecycle assessment can be expressed as follows:

$$LCA_{tot} = \sum (LCA_{res} + LCA_{op} + LCA_{constr}) \tag{5}$$

2.2.2 Resource

Considering the resource extraction phase is crucial because, for example, a gas boiler can use either renewable gas or fossil gas as input. In the original REHO model, however, the environmental impacts of using both types of gas in the boiler are considered identical since it only accounts for the construction and operation of the natural gas boiler. This approach is unrealistic.

Including the resource extraction stage allows us to capture the environmental impacts associated with the extraction of resources. In REHO, this extraction process is indeed the energy exchange between districts and networks, the formulation for this stage can be expressed as follows:

$$EI_{sup}(l,u,t) = ei_{supply}(l,u,t) \cdot Network_{supply}(l,u,t) \tag{6}$$

$$EI_{dem}(l,u,t) = ei_{demand}(l,u,t) \cdot Network_{demand}(l,u,t). \tag{7}$$

$$LCA_{res} = \sum_{l \in \text{ResourceBalances}, u \in \text{units}, t \in \text{Time}} (EI_{sup}(l,u,t) - EI_{dem}(l,u,t)) \tag{8}$$

Where,

- LCA_{res} represents the total resource environmental impacts in the energy system.
- $EI_{sup}(l,u,t)$ is the total environmental impacts of unit u supply to the energy layer$l$ at time t.
- $EI_{dem}(l,u,t)$ is the total environmental impacts of unit u demand from the energy layer$l$ at time t.
- $ei_{supply/demand}$ is the unit environmental impact.
- $Network_{supply}$ is the quantity of energy supply from the network to the system.
- $Network_{demand}$ is the quantity of energy supply to the network from the system.

The original operation stage is the same with the resource stage.

2.2.3 Construction

As previously mentioned, determining fixed environmental impacts is unnecessary due to the lack of literature on scale-up effects in LCA and the absence of intuitive physical meaning for these fixed impacts. Consequently, fixed environmental impacts are excluded from this research. Instead, we assume that the environmental impacts of construction are proportional to the equipment's capacity. The formulation can thus be summarized as follows:

$$LCA_{constr} = \sum_{u \in Units} \frac{SizeofUnits(u) * ei_{variable}(u)}{lifetime(u)} \tag{9}$$

where

- LCA_{constr} represents the total construction environmental impacts in the energy system.
- $ei_{variable}$ is the variable environmental impact of unit u.
- $SizeofUnits(u)$ is the installed capacity for the unit u.
- $lifetime(u)$ is lifetime of unit u.

2.2.4 Operation

The operation phase is the truly neglected aspect in REHO. To accurately assess the environmental impacts of operation, we exclude energy and material flows originating from the resource extraction and construction phases in our database. The operational LCA is calculated by multiplying the working hours of the units by the operational lifecycle assessment inventory and the unit's output capacity. The calculation of operational environmental impacts is detailed below, and the process for eliminating double counting will be discussed subsequently in LCA database creation part.

$$LCA_{op} = \sum_{l \in ResourceBalances, u \in Units, t \in Time} ei_{op}(u) * Units\_Supply(l, u) * t \tag{10}$$

where

- LCA_{op} represents the total operation environmental impacts in the energy system.
- ei_{op} is the operation environmental impact of unit u.
- $Units\_Supply(l, u)$ is the unit u's output in energy layer l.
- t is the unit working time.

2.3 LCA database creation

The creation of new LCA database is based on an open-source framework for Life Cycle Assessment, brightway2[17]. Its combination of a modular structure, the expressiveness and interactivity of Python and in particular Jupyter notebooks, and tuned calculation pathways allows for new research directions in Life Cycle Assessment. The database we are going to use is ecoinvent[18], which is widely used in LCA especially in Europe.

2.3.1 Generalization of LCA indicators

Due to insufficient and unconvincing LCA indicators in REHO, and to facilitate a better comparison between REHO and other energy system optimization models, particularly Energyscope, a generalization of the LCA indicators is necessary. The new LCA method utilized in REHO is World IMPACT + [19]. World IMPACT + offers a comprehensive and detailed assessment of environmental impacts across various categories, ensuring a thorough evaluation of sustainability. It integrates multiple impact categories, including resource use, emissions, and environmental degradation, providing a holistic view of environmental performance. Additionally, since Energyscope also employs World IMPACT +, using the same method will simplify comparisons. Furthermore, World IMPACT + includes internal packages specifically for Brightway, making it more manageable and user-friendly.

After generalization, 37 different environmental impact indicators can be used for evaluation as the table listed below:

Table 2: Environmental Impact Categories

Impact category	Unit	Abbrev
Climate change, ecosystem quality, long term	PDF.m2.yr	CCEQL
Climate change, ecosystem quality, short term	PDF.m2.yr	CCEQS
Climate change, ecosystem quality, long term, fossil and biogenic	PDF.m2.yr	CCEQLB
Climate change, ecosystem quality, short term, fossil and biogenic	PDF.m2.yr	CCEQSB
Climate change, human health, long term	DALY	CCHHL
Climate change, human health, short term	DALY	CCHHS
Climate change, human health, long term, fossil and biogenic	DALY	CCHHLB
Climate change, human health, short term, fossil and biogenic	DALY	CCHHSB
Freshwater acidification	PDF.m2.yr	FWA
Freshwater ecotoxicity, long term	PDF.m2.yr	FWEXL
Freshwater ecotoxicity, short term	PDF.m2.yr	FWEXS
Freshwater eutrophication	PDF.m2.yr	FWEU
Human toxicity cancer, long term	DALY	HTXCL
Human toxicity cancer, short term	DALY	HTXCS
Human toxicity non-cancer, long term	DALY	HTXNCL
Human toxicity non-cancer, short term	DALY	HTXNCS
Ionizing radiation, ecosystem quality	PDF.m2.yr	IREQ
Ionizing radiation, human health	DALY	IRHH
Land occupation, biodiversity	PDF.m2.yr	LOBDV
Land transformation, biodiversity	PDF.m2.yr	LTBDV
Marine acidification, long term	PDF.m2.yr	MAL
Marine acidification, short term	PDF.m2.yr	MAS
Marine acidification, long term, fossil and biogenic	PDF.m2.yr	MALB
Marine acidification, short term, fossil and biogenic	PDF.m2.yr	MASB
Marine eutrophication	PDF.m2.yr	MEU
Ozone layer depletion	DALY	OLD
Particulate matter formation	DALY	PMF
Photochemical oxidant formation	DALY	PCOX
Terrestrial acidification	PDF.m2.yr	TRA
Thermally polluted water	PDF.m2.yr	TPW
Water availability, freshwater ecosystem	PDF.m2.yr	WAVFWES
Water availability, human health	DALY	WAVHH
Water availability, terrestrial ecosystem	PDF.m2.yr	WAVTES
Total ecosystem quality	PDF.m2.yr	TTEQ
Total human health	DALY	TTHH
Total ecosystem quality, fossil and biogenic	PDF.m2.yr	TTEQB
Total human health, fossil and biogenic	DALY	TTHHB

Meanwhile, the GWP (Global Warming Potential) indicator continues to rely on the IPCC 2013 methodology primarily due to its widespread recognition for consistency and stability. Many LCA databases and software tools, such as Ecoinvent and SimaPro, still frequently default to IPCC 2013 values. This means that IPCC 2013, GWP 100 remains a common choice for calculating GWP, ensuring compatibility with previous studies and established datasets.

2.3.2 Double counting removal

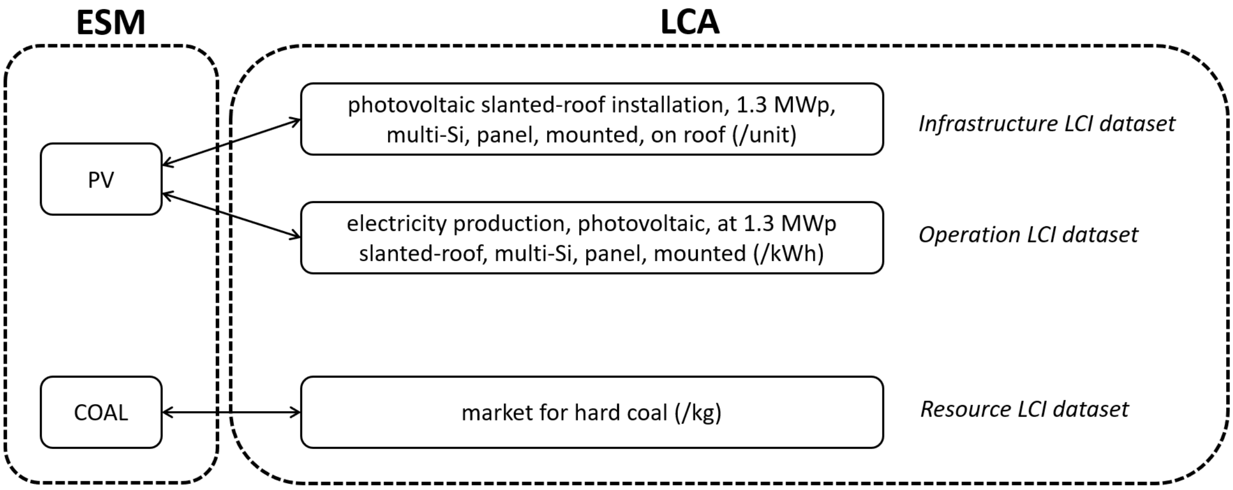

Figure 8: Mapping Examples Between Energy System-Specific Technologies and the LCA Database

In LCA, we use the lifecycle inventory like ecoinvent. While in energy system modeling, units are manually defined by their inputs and outputs. To coordinate, the units in energy system model will be corresponded to exsiting items in the LCA database. And double counting occurs during this process.

For instance, a natural gas boiler is modeled with natural gas as its input and heat as its output. Both the construction of the natural gas boiler and the provision of natural gas are accounted for in the construction and resource phases, respectively. However, additional energy and material flows occur during the boiler's operation. These flows are addressed in the operation life cycle assessment (LCA). In the Ecoinvent database, the operation of technologies considers both construction and resource flows. Therefore, it is necessary to eliminate the effects from these flows to align with our methodology and avoid double counting.

The tool used for double counting removal is Mescal [20], developed specifically for use with the same LCA methodology as another energy system optimization model, EnergyScope. The basic logic for double counting removal employed by Mescal is shown below:

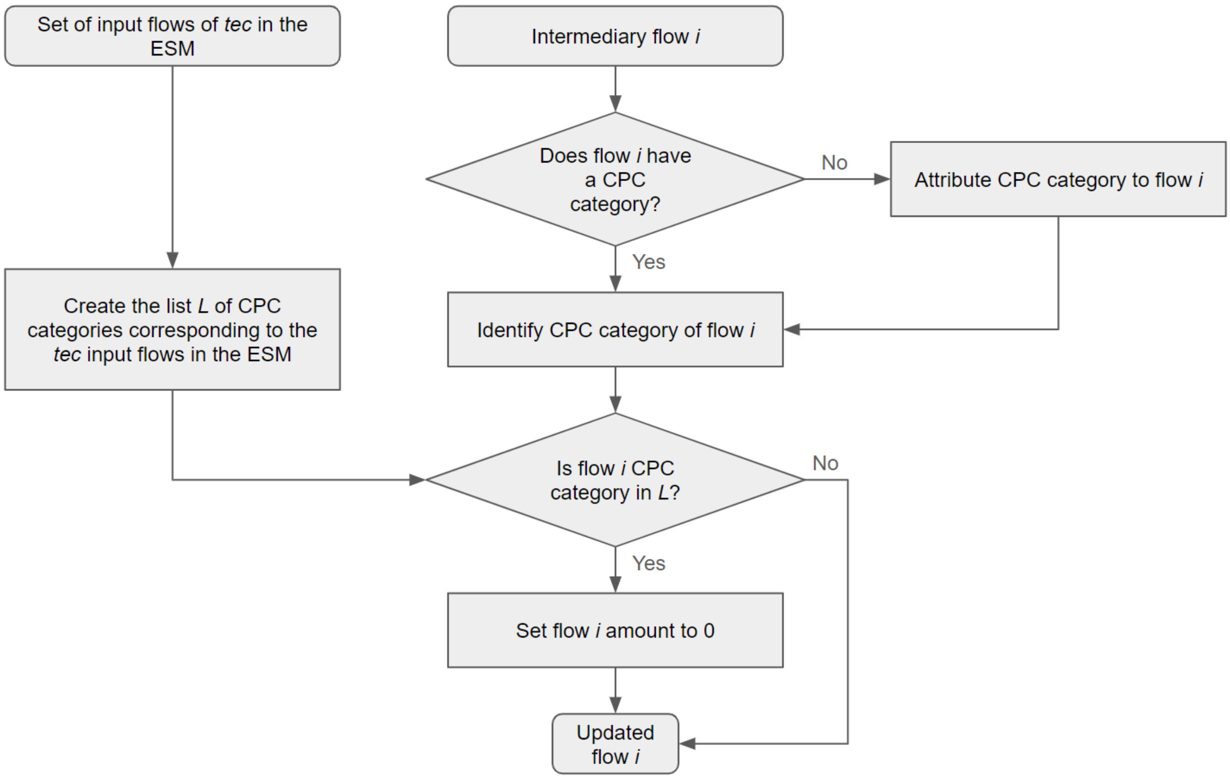

Figure 9: Flowchart of the Double Counting Removal Mechanism

The Central Product Classification (CPC) system is crucial for this task. It provides a coherent and consistent framework for classifying products (both goods and services) based on a set of internationally agreed concepts, definitions, principles, and classification rules [21]. Initially, each input flow, predefined in our energy system model, is assigned a CPC category. This process results in a comprehensive list L of CPC categories encompassing the technology inputs. Subsequently, each technology is analyzed to assign CPC categories to its intermediary flows. These are then compared with the list L. If an intermediary flow is found in list L, indicating it is already defined as an input flow in our energy system, its flow amount is set to zero to avoid double counting. This method effectively ensures the removal of double counting.

2.3.3 Refactor

In practice, after the removal of double counting, the quantity levels of several LCA indicators can drop to extremely low values, ranging from 10^{-14} to 10^{-16}. These values are too small for AMPL to handle effectively. To avoid computational errors, all indicators are scaled by a refactor, with a default value of 1000.

2.3.4 Files mapping

In the process of removing double counting, we assign numerous flows with their respective CPC classifications using generative methods to boost mapping efficiency. The Industrial Ecology Machine Learning Mapping [22] is a Python module that employs machine learning to align two classifications based on word similarity. However, despite the utilization of this module, a manual review of the mappings is still necessary to guarantee their accuracy.

```
[1]
import ML_mapping
Executed at 2024.08.20 10:51:50 in 15s 359ms

[2]
self = ML_mapping.Mapping(reference_classification='CPC 2.1',
                          transformer_model='all-MiniLM-L6-v2',
                          number_of_guesses=5)
Executed at 2024.08.19 18:09:56 in 2m 9s 981ms

[3]
self.match_inputs(['OIL_Boiler'])
self.calculate_scores()
self.format_results()
Executed at 2024.08.19 18:09:56 in 77ms
```

5 rows 5 rows × 3 columns

product	order	code sector	sector	similarity
OIL_Boiler	1	33360	Gas oil	0.572717
	2	3336	Gas oil	0.572717
	3	33370	Fuel oils n.e.c.	0.522871
	4	3337	Fuel oils n.e.c.	0.522871
	5	21671	Olive oil, crude	0.492022

Figure 10: Example Codes for Mapping Files Using Machine Learning Methods

Additionally, files are required to characterize the model by defining the output and input flows for each technology. These flows are manually calculated based on both the technology's efficiency and power, establishing the relationship between resource inputs and technology operations. It is crucial to determine how many resource inputs are required to produce one unit of output. A portion of the model file is shown in the figure below:

Name	Flow	Amount
NG_Boiler	Heat	1
NG_Boiler	NaturalGas	-1.052631579
OIL_Boiler	Heat	1
OIL_Boiler	Oil	-1.111111111
WOOD_Stove	Heat	1
WOOD_Stove	Wood	-1.086956522

Figure 11: Example Model Characteristic Files for Defining Inputs and Outputs for Each Technology

Some technologies can be quite complex. To accurately characterize these, they need to be broken down into several simpler components. Therefore, a 'technology_specific' file is required. For each technology, the lifetime must be specified for calculating its construction LCA. Additionally, when using our custom energy system, we may have unique units, necessitating a conversion file to bridge our energy system model with Ecoinvent. For a more comprehensive understanding for these files, please refer to my file folder.

2.4 Normalization

The purpose of normalization is to facilitate multi-objective optimization. The multi-objective function is defined as:

$$Objective = \sum w_i X_i \tag{11}$$

where,

- w_i is the weights for different KPIs.
- X_i is the i th indicator's value.

The weights w_i will be determined using generative algorithms, which requires all indicators to be normalized to the same quantitative scale to ensure fair comparison and balanced contribution to the overall objective function. However, in this project, only the preparatory steps for normalization have been completed. The methodology for multi-objective optimization still follows the original REHO approach, which will be introduced in a later section.

2.4.1 Normalization

The normalization utilized the formulation below:

$$X_{normalized} = \frac{X - X_{min}}{X_{max} - X_{min}} \tag{12}$$

Where X means the LCA indicators. To obtain the minimum value of X, we can simply set the objective to the specific LCA indicators we wish to normalize. However, because all constraints are aimed at minimizing these indicators, determining the maximum value of X poses a challenge, as the system will not converge under such conditions. Therefore, our approach is to minimize the other LCA indicators. Among these results, we can then identify the maximum value for the LCA indicators we wish to normalize.

2.4.2 Correlation analysis among indicators

Since there are as many as 29 LCA indicators, optimizing them individually for complex scenarios—such as 4,000 buildings with varying features—would be computationally intensive, increasing the time complexity 29-fold. To simplify this process, a correlation analysis is employed. By calculating the Pearson correlation coefficient between each indicator [23], we can identify the most influential indicators. This allows us to focus optimization efforts on these key indicators rather than optimizing all 22, significantly reducing computational complexity while maintaining accuracy in the results.

The Pearson correlation coefficient, denoted as r, between two LCA indicators X and Y is calculated using the formula:

$$r_{XY} = \frac{\sum(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum(X_i - \bar{X})^2 \sum(Y_i - \bar{Y})^2}} \tag{13}$$

where:

- X_i and Y_i are the individual data points of variables X and Y,
- $\bar{X}$ and $\bar{Y}$ are the mean values of X and Y, respectively,
- $\sum$ represents the summation over all data points.

2.4.3 Multi-objective optimization

In multi-objective optimization, various methodologies exist, including scalarization methods [24], Pareto-based approaches [25], and epsilon-constraint methods [26], among others. In REHO, the epsilon-constraint method is employed for optimization.

The epsilon-constraint method is a widely used approach in multi-objective optimization where one objective is optimized, while the remaining objectives are transformed into constraints with upper bound limits, called epsilon (ϵ) values. In REHO, these constraints are defined as below:

$$Objective + Slack\_Variable = \epsilon Area\_tot \tag{14}$$

This constraint ensures that the indicators remain within a limit defined by ϵ (represented as EMOO_CAPEX, EMOO_GWP, etc.), adjusted by the total area. In the REHO objective functions, penalty terms are incorporated to address potential constraint violations. These penalties act as a corrective measure: if any objective exceeds its specified limit, the penalty term increases, driving the solution toward feasible regions that comply with the epsilon constraints. This mechanism effectively guides the optimization process to respect the imposed limits while balancing trade-offs between objectives.

```
# Multi objective optimization
subject to EMOO_CAPEX_constraint: # beta_cap
Costs_inv + EMOO_slack = EMOO_CAPEX * Area_tot;

subject to EMOO_OPEX_constraint: # beta_op
Costs_op + EMOO_slack_opex = EMOO_OPEX * Area_tot;

subject to EMOO_GWP_constraint: # beta_gwp
GWP_tot + EMOO_slack_gwp = EMOO_GWP * Area_tot;

subject to EMOO_TOTEX_constraint: # beta_tot
Costs_tot + EMOO_slack_totex = EMOO_TOTEX * Area_tot;

subject to EMOO_lca_constraint{k in Lca_kpi} :
lca_tot[k] <= EMOO_lca[k] * Area_tot;
```

Figure 12: Multi-Objective Optimization (MOO) Using Epsilon Constraints in the Renewable Energy Hub Optimizer (REHO)

```
param penalty_ratio default 1e-6;
var penalties default 0;

subject to penalties_contraints:
penalties = penalty_ratio * (Costs_inv + Costs_op + sum{k in Lca_kpi} lca_tot[k] +
            sum{l in ResourceBalances,p in PeriodExtreme,t in Time[p]} (Network_supply[l,p,t] + Network_demand[l,p,t])) + Costs_cft;


# objective functions
minimize TOTEX:
Costs_tot + Costs_grid_connection + penalties;

minimize OPEX:
Costs_op + Costs_grid_connection + penalties;

minimize CAPEX: # the second term is added to correspond to objective set in design_cst
Costs_inv + penalties;

minimize GWP:
GWP_tot + penalties;
```

Figure 13: Components of the Objective Functions in the Renewable Energy Hub Optimizer (REHO)

3 Case Study

3.1 Case Study for Lugaritz-Matía Community in Spain[27]

3.1.1 Introduction

Located in the vibrant city of San Sebastián, in the heart of Basque Country, the Lugaritz-Matía Community pilot is a pioneering initiative integrating modern healthcare facilities with a cutting-edge, sustainable energy network. This project encompasses three key institutions: the Birmingham Hospital and two nursing homes, Rezola and Lugaritz, which collectively serve around 700 residents and span a total heated area of 40,000 m^2. The pilot is a model for how healthcare institutions can adopt innovative energy solutions to improve efficiency and sustainability.

3.1.2 Facility Overview

The Lugaritz-Matía Community pilot comprises:

- **Birmingham Hospital**: The primary medical institution in the network, composed of hospital and swimming pool. Several technologies are utilized, including a condensing boiler using biomass (District Heating), a conventional boiler, an air-source heat pump (ASHP), and solar thermal systems.
- **Rezola Nursing Home**: A facility dedicated to the long-term care of elderly residents, composed of only residential nursing home. The system utilizes district heating powered by a biomass-fueled condensing boiler, along with additional conventional and condensing boilers.
- **Lugaritz Nursing Home**: Another key institution composed of residential nursing home and administrative office. The setup includes district heating with a biomass-fueled condensing boiler, a ground-source heat pump (GSHP), and a chiller recovery system.

For more details on the three buildings, refer to the appendix (Fig 37, Fig 38). However, before diving into the case study, several clarifications are needed. First, the database used lacks credibility and is missing some crucial information. For instance, as shown in the appendix figures, the average thermal transmittance for facades and roofs is extremely high, with the value for the building Rezola exceeding $1kW/(m^2 \cdot K)$. In contrast, single glazing in typical building structures accounts for only 5.7 $W/(m^2 \cdot K)$. Even when the unit is converted from $kW/(m^2 \cdot K)$ to $W/(m^2 \cdot K)$, these values remain unrealistic.

To address this, we replaced the unrealistic values by reviewing REHO's internal building files and selecting data from buildings constructed in recent decades. Additionally, instead of using latitude and longitude directly—information which is also missing from the datafile—REHO requires converting these coordinates into an X and Y coordinate system using the EPSG:2054 standard [28].

In cases where data, such as DHN and DataHeat, could not be found in the Ecoinvent database, I replaced missing values with 0.0. However, this caused the problem to fail to converge. As a result, we can only set the objectives using the original indicators (GWP, TOTEX, OPEX).

Figure 14: Pilot View from Satellite[27]

3.1.3 Information to input

In the REHO model, three primary types of inputs are utilized:

- **End Use Demands (EUDs):** Calculated based on weather data and building characteristics. Weather data is automatically fetched from the PVGIS database, providing yearly data which is then clustered into typical days for use in REHO.
- **Resources:** Includes grids to which the system has access for providing the EUDs.
- **Equipment:** Consists of the technologies used to convert resources into required services. Detailed in the Life Cycle Assessment (LCA) for building units, district units, and grids.

In the methodology section, I mentioned the integration of equipment and resource data within the REHO framework, emphasizing our strategy for developing a comprehensive Life Cycle Assessment (LCA) database. This database was constructed based on initial files from original technologies, which included detailed information on each technology's capacity, cost, lifetime, and other pertinent metrics. These LCA features were then seamlessly merged with the original data files, creating a unified resource that integrates environmental impact assessments with functional characteristics. This modification was crucial because REHO has a unique data processing approach; it reads CSV files and converts the dataframes into parameters that the AMPL model requires, ensuring efficient data flow and model integration.

3.1.4 Building Characteristics Datafile Format

The building characteristics require manual generation of a datafile with the necessary columns as follows:

area_era_m2 Total floor area in square meters.

id_building Identifier for the building.

temperature_cooling_supply_C Supply temperature for cooling in degrees Celsius.

temperature_cooling_return_C Return temperature for cooling in degrees Celsius.

temperature__heating__supply__C Supply temperature for heating in degrees Celsius.

temperature__heating__return__C Return temperature for heating in degrees Celsius.

area__facade__m2 Area of the facade in square meters.

area__roof__solar__m2 Area of solar panels on the roof in square meters.

temperature__interior__C Interior temperature in degrees Celsius.

ratio Usage ratio or occupancy rate.

status Operational status of the building.

id__class Classification ID for the building.

thermal__specific__capacity__Wh__m2__K Thermal capacity per unit area.

thermal__transmittance__signature__kW__m2__K Thermal transmittance signature.

height__m Building height in meters.

x X-coordinate for geographic location.

y Y-coordinate for geographic location.

geometry Geometric data or shape of the building.

3.1.5 2 different scenarios

Two distinct scenario case studies were performed for the pilot site:

- **Scenario 1**: This scenario involved simulating the existing energy system of the pilot site to understand its current performance and identify potential areas for improvement.
- **Scenario 2**: This scenario focused on optimization without constraints, meaning no consideration was given to prior investments. Additionally, it aimed for a fossil-free operation by excluding the use of non-renewable energy sources such as natural gas (NG) boilers and oil boilers.

3.2 Case study for verification

Meanwhile, an existing error in REHO arises when applying the district heating network feature to Spain's case study. End-use demands (EUDs), which should be determined by building characteristics and weather data, are expected to remain consistent for a specific location across different scenarios. However, due to an internal error in REHO, the EUDs for Spain differ between scenarios, making comparisons invalid as the EUDs should remain constant.

To address this issue, a comparative case study was conducted using REHO's example scripts to validate the integration of the LCA framework and assess the generalization of LCA indicators across different scenarios. This study ensures that the LCA methodology is implemented robustly, accurately capturing and generalizing environmental impacts. The case study was performed for Sion, Switzerland, under a basic scenario setup—without any enforced or excluded units. For a comprehensive comparison, both the old and new datafiles were used separately to analyze variations in outcomes and identify any discrepancies.

For detailed information regarding the default buildings, please refer to the building.csv file. The figure below presents part of the buildings' characteristics.

egid	period	class	area_era_m2	is_hotwater_system	x	y	z	transformer	gas_grid	h2_grid
954135	1946-1960	Residential	1095.0	True	2592820.444606846	1120168.6329999994	573.56	3658	True	False
9038773	1991-2000	Residential	100.0	True	2592379.023002238	1120317.4309999999	653.44	3658	True	False
954108	1961-1970	Residential	100.0	True	2592459.3597605387	1120373.5124999993	653.64	3658	True	False
3113265	1991-2000	Residential	961.0	True	2592564.393124838	1119898.0670000017	565.17	3658	True	False
954118/954120/954119	1971-1980	Residential	508.0	True	2592659.8565228325	1120015.1919999998	569.7152961123926	3658	True	False
954122/954121	1981-1990/1971-1980	Residential	310.0	True	2592689.866545876	1120027.3990000002	562.4250619209357	3658	True	False
954130	1946-1960	Residential	894.0	True	2592713.3476149514	1119984.613500001	559.3	3658	True	False
954124/954123	1971-1980	Residential	330.0	True	2592708.284098482	1120033.3184999991	559.57	3658	True	False
954129/954128	1981-1990	Residential	1260.0	True	2592752.9730697055	1120013.6225500014	559.0400000000001	3658	True	False
954127	1981-1990	Residential	720.0	True	2592777.2309807995	1120046.8227499984	559.4769412831165	3658	True	False
954484	1961-1970	Residential	776.0	True	2592762.2514200956	1119976.0559999999	556.62	3658	True	False

Figure 15: Examples of Partial Building Characteristics for Sion

4 Results

Using the parameters, formulas, and datasets outlined earlier, we can derive a comprehensive set of results. In this section, various results will be presented and analyzed, including figures, data visualizations, and Sankey diagrams. Each of these elements will provide insight into the system's performance, highlighting key trends and outcomes. The following analysis will delve into the implications of the results, examining correlations, discrepancies, and any unexpected behavior, allowing for a thorough understanding of the system dynamics across different scenarios.

This section features multiple Sankey diagrams, where colors are used to represent different energy sources and systems:

- **Orange**: District Heating Network (DHN)
- **Black**: Gas Boiler
- **Green**: Electricity
- **Pink**: Heat Pump
- **Yellow**: Photovoltaic (PV)
- **Red**: Domestic Hot Water (DHW)

4.1 Results for Sion's case study

4.1.1 Comparative Analysis Between the Old and New Methodologies

To validate the successful integration of the new LCA methodology into REHO, a comparative study between the old and new methodologies was conducted. The internal AMPL code and the associated Python data processing scripts were updated to reflect the new methodology. The scenarios analyzed are based on the 40 buildings included in the REHO example code, without any additional features, and the objective function for optimization remains minimizing Global Warming Potential (GWP). For consistency in comparison, the lifecycle inventory has been kept unchanged. The End-Use Demands (EUDs) are displayed below:

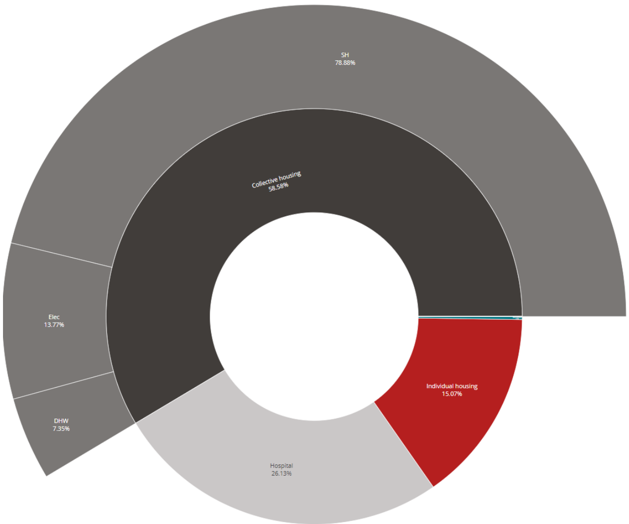

Figure 16: End-Use Demands (EUDs) Distribution for Sion's Case

The primary contributor to the EUDs is collective housing, which accounts for 78.88% space heating (SH), 13.77% electricity (elec), and 7.00% domestic hot water (DHW). As expected, both the old and new methodologies resulted in the same energy and material flows, as illustrated by the following Sankey diagram:

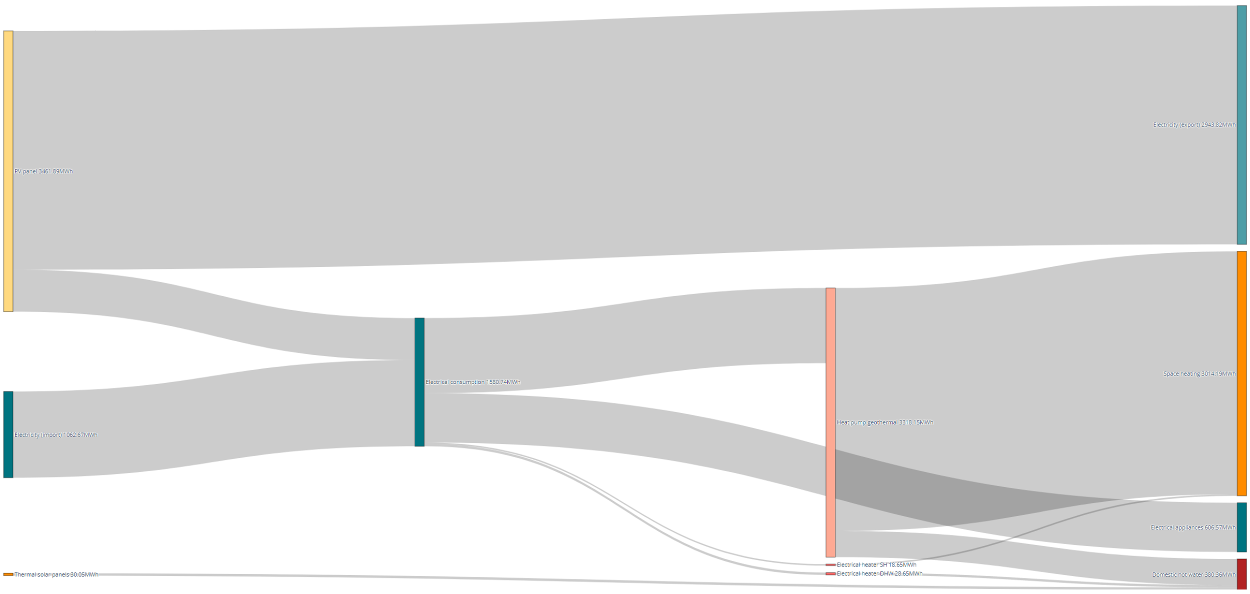

Figure 17: Sankey Diagram of Energy and Material Flows to Fulfill End-Use Demands (EUDs) in Sion's case

This outcome implies that both methodologies yield the same optimal configuration of the energy system. However, the key difference between the two methodologies lies in the operation phase of the energy system, which ultimately affects the final GWP results. The comparison between the final values for the new and original methodologies is shown below:

```
0 simplex iterations
               GWP
0 -121598.471405
```

Figure 18: Global Warming Potential (GWP) Optimization Results from the New Methodology for Sion's Case Study

```
0 simplex iterations
               GWP
0 -122008.748889
```

Figure 19: Global Warming Potential (GWP) Optimization Results from the Original Methodology for Sion's Case Study

The results indicate that both energy systems could reduce GWP by over 120,000 units (no specific units are available, as the old database lacks this information, which also contributes to its lower credibility). The differences observed between the two methodologies can be attributed to their treatment of the operation and construction phases. Specifically, the new methodology accounts for both operational and construction GWP in a different manner, resulting in variations in the final GWP values.

4.1.2 Generalization for LCA indicators

To explore the generalization of LCA indicators, the new database, after the removal of double counting, is utilized. The scenarios remain the same for consistency. To better interpret the results, various Sankey diagrams are compared. Since the purpose of this section is to showcase the new methodology and database, the scenarios are selected from REHO's internal examples. Therefore, the specific numerical values generated by REHO are not the primary focus. Instead, the emphasis is on comparing the results with the original REHO outputs to highlight any differences.

If we continue to set GWP as the optimization objective, the resulting Sankey diagram closely resembles the original:

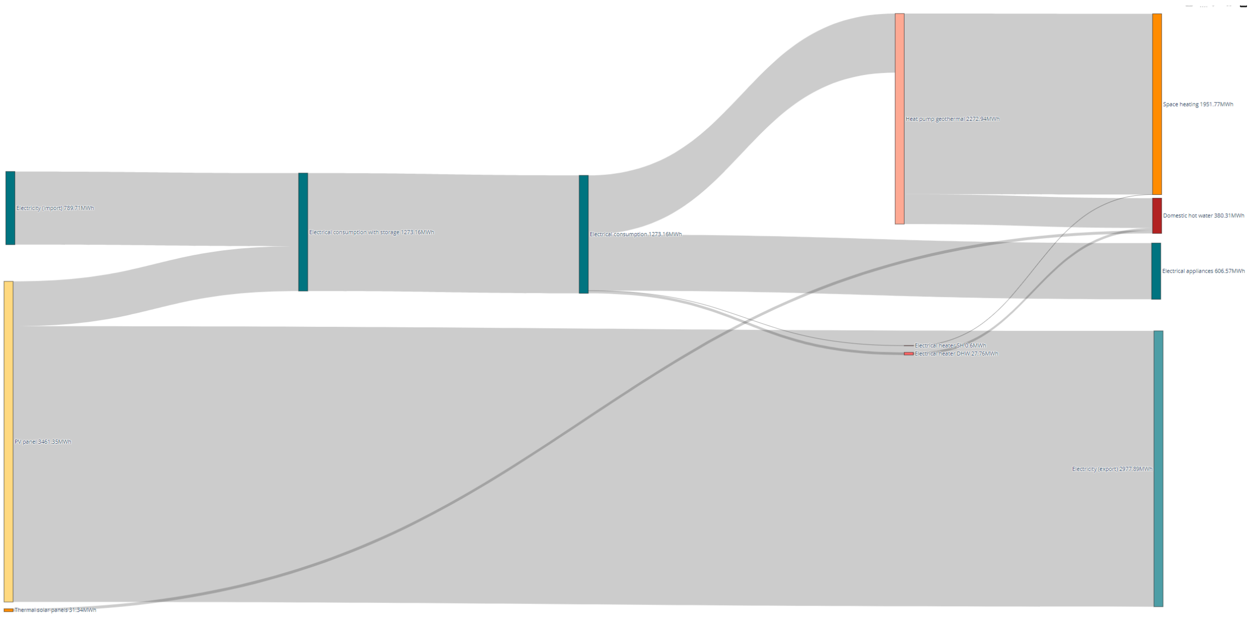

Figure 20: Sankey Diagram of Energy and Material Flows to Optimize GWP in Sion's Case

This demonstrates that when optimizing for GWP, photovoltaic (PV) systems are extensively utilized, while fossil-based technologies are excluded in order to minimize global warming potential. The corresponding optimized results are shown below:

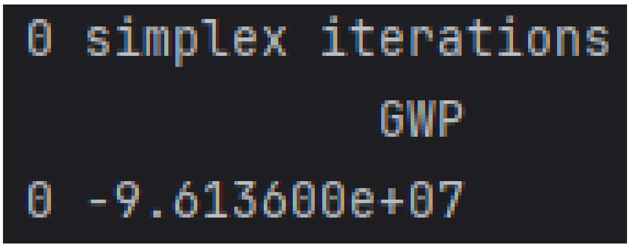

Figure 21: Optimized Results for Global Warming Potential (GWP)

Although the result appears significantly smaller than those previously shown, this is due to the scaling factor of 1000 applied to all indicators, as mentioned earlier. Using the IPCC 2013 method, the actual value is approximately $-96,136$ (kg CO_2 eq), which is much more convincing.

If we change the optimization objective from GWP to LOBDV (Land Occupation, Biodiversity), the system behaves differently:

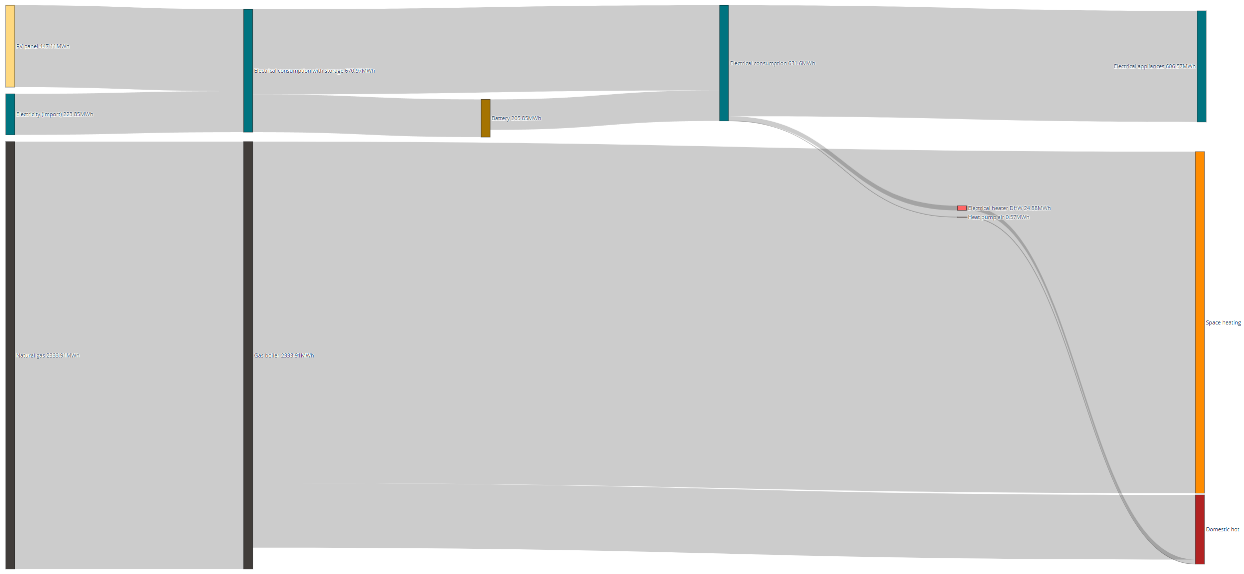

Figure 22: Sankey Diagram of Energy and Material Flows for Optimizing Land Occupation and Biodiversity (LOBDV) in Sion's Case

In this case, the energy system limits the use of PV panels due to their large land occupation. Instead, gas boilers are widely used to provide heat, reducing the reliance of electricity on heat pumps that would otherwise require significant land area to operate.

There are 27 additional LCA indicators, each with its own optimal configuration. By considering these indicators, the generalization of the LCA indicators is successfully achieved.

4.2 Results for Spain's case study

The scenarios for Spain utilized multiple energy technologies, with the data sourced from the InterPED project. Two optimization scenarios were conducted. In the first scenario, the technologies remained unchanged, and a comprehensive simulation of the existing energy system was performed to identify potential areas for improvement. The second scenario focused on achieving a carbon-free energy system by excluding carbon-emitting technologies, with the aim of maximizing the system's performance based on various objectives. These optimizations provide valuable insights into the trade-offs and challenges involved in transitioning to a more sustainable energy system.

4.2.1 Scenario 1: current scenario simulation

In the current scenario, technologies such as 'HeatPump_Air', 'HeatPump_Geothermal', 'NG_Boiler', 'OIL_Boiler', 'ThermalSolar', and 'PV' are fixed based on the pilot datasets, as shown in Figure 37. It means the composition of the energy system is fixed. Thus it's not optimization and it's just simulation. Additionally, a water district heating network (DHN) is utilized in this case. To integrate these technologies into REHO, some minor adjustments are necessary. For instance, in the CSV file, the thermal solar system for the Bermingham building is specified as 40×0.4. However, in REHO, the required unit for thermal solar is 2.32. Therefore, an integer multiple of 16.24 is used to substitute for 16.

After all the necessities prepared, the simulation based on the old datasets are performed:

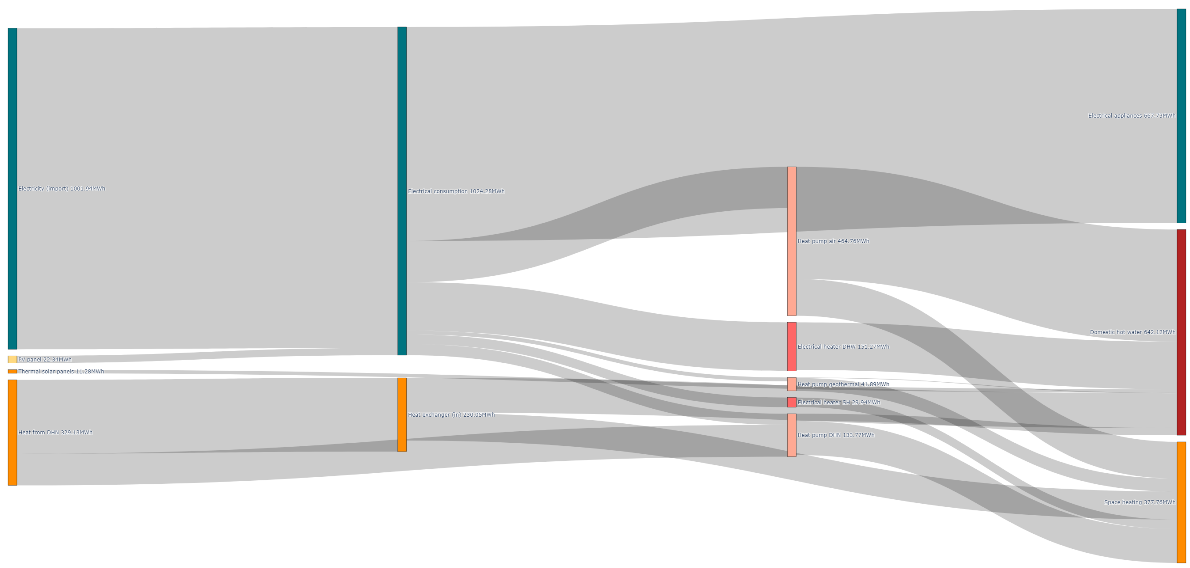

Figure 23: Sankey Diagram of Energy and Material Flows Using the Old Dataset in REHO for GWP Optimization

Nevertheless, for the new double counting removal database, the simulation sankey performs like this:

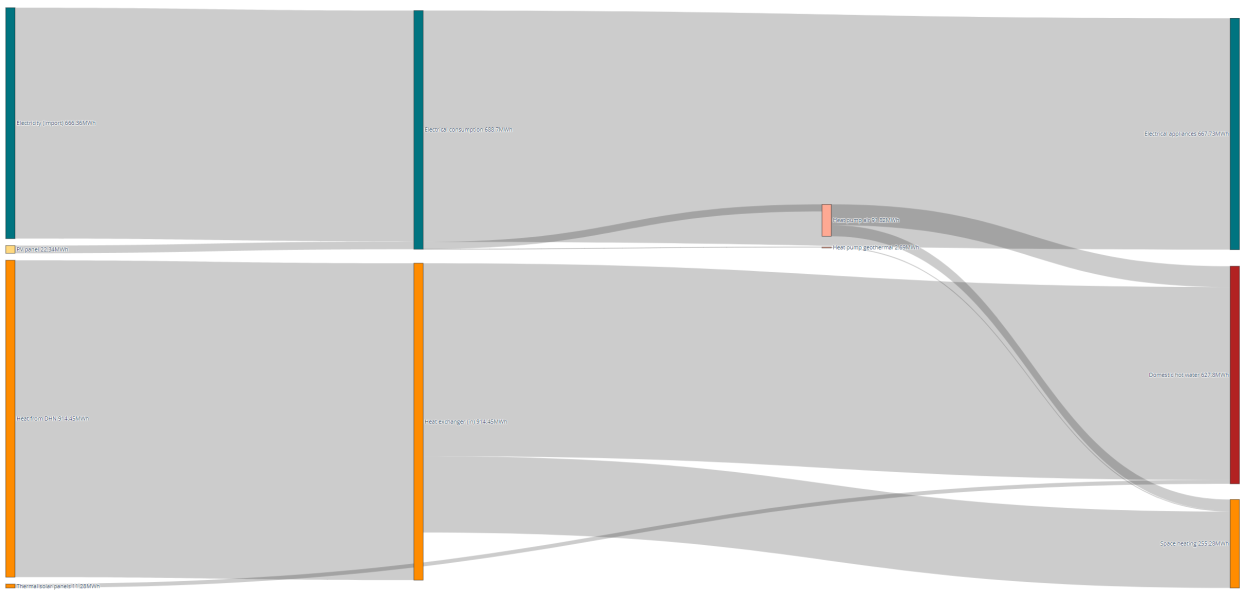

Figure 24: Sankey Diagram of Energy and Material Flows Using the Double Counting Removed Dataset in REHO for GWP Optimization

The most significant difference between these two results is that the former primarily uses electricity to power heat pumps in order to produce heat and meet the demand, while the latter relies mainly on direct heat from the district heating network. The root cause of this behavioral difference is that, in the original dataset, heat is more expensive than electricity in terms of GWP, whereas the new database does not reflect this. However, due to the direct and more efficient use of heat, the GWP of heat is indeed lower than that of electricity, as is also evident from the Ecoinvent database:

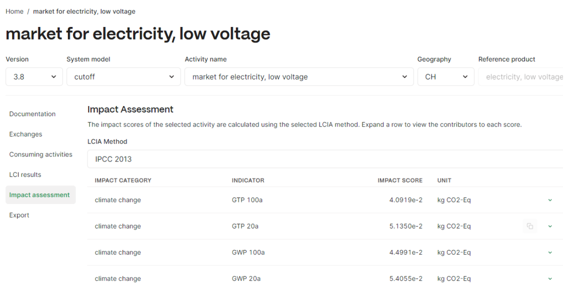

Figure 25: Electricity Market Activity in Ecoinvent

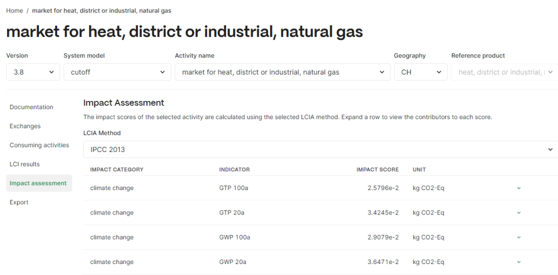

Figure 26: Heat Market Activity in Ecoinvent

The final optimal results for GWP are presented below:

```
           GWP
0  42472.242005
Results are saved in results/fix.pickle
Results are saved in results/fix_GWP.xlsx
```

Figure 27: Final Simulation Results for GWP Using the New Methodology for Spain's Case Study in the Current Scenario

This indicates that 42,472.24 kg CO_2-eq is emitted over one year in the current energy system. The corresponding energy profiles, with a weekly moving average, are shown below:

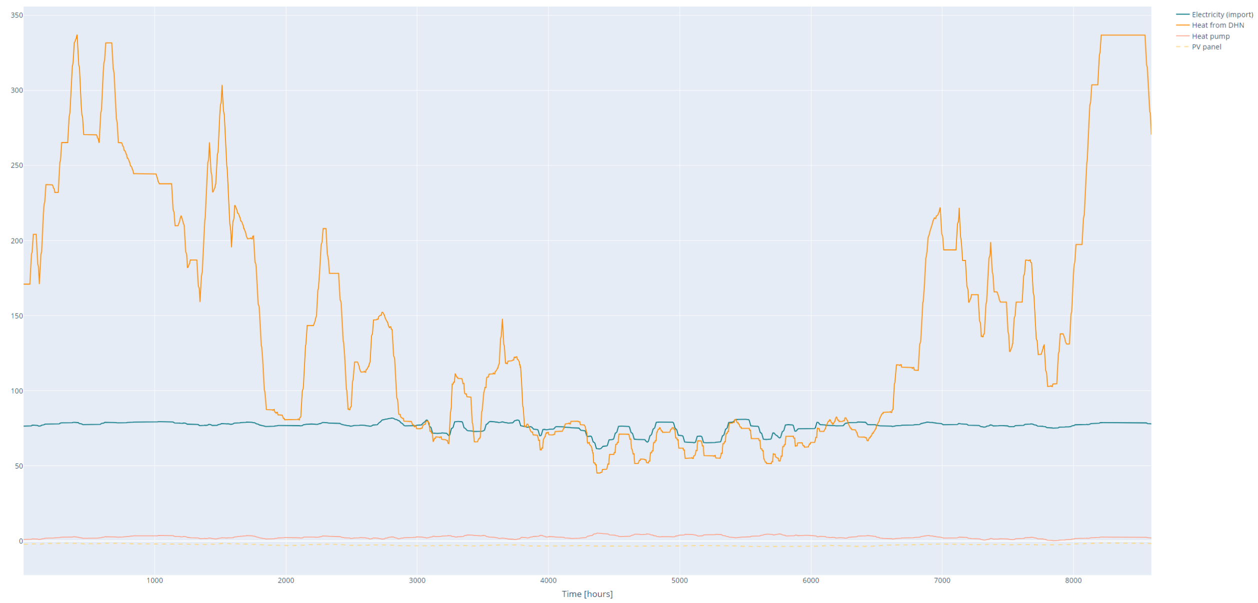

Figure 28: Energy Profiles with a Weekly Moving Average for Spain's Case Study

4.2.2 Scenario 2: Fossil-Free Scenario

A fossil-free optimization scenario was conducted, excluding fossil-fuel-based technologies, while maintaining the same EUDs (End-Use Demands) through consolidated building features. The optimized GWP results are shown below:

```
             GWP
0  26827.551036
Results are saved in results/gwp.pickle
Results are saved in results/gwp_gwp.xlsx
```

Figure 29: Fossil-Free Scenario GWP Optimization for Spain's Case Study

This scenario achieves a reduction of approximately 36.8% in global warming potential. And the corresponding Sankey is shown below:

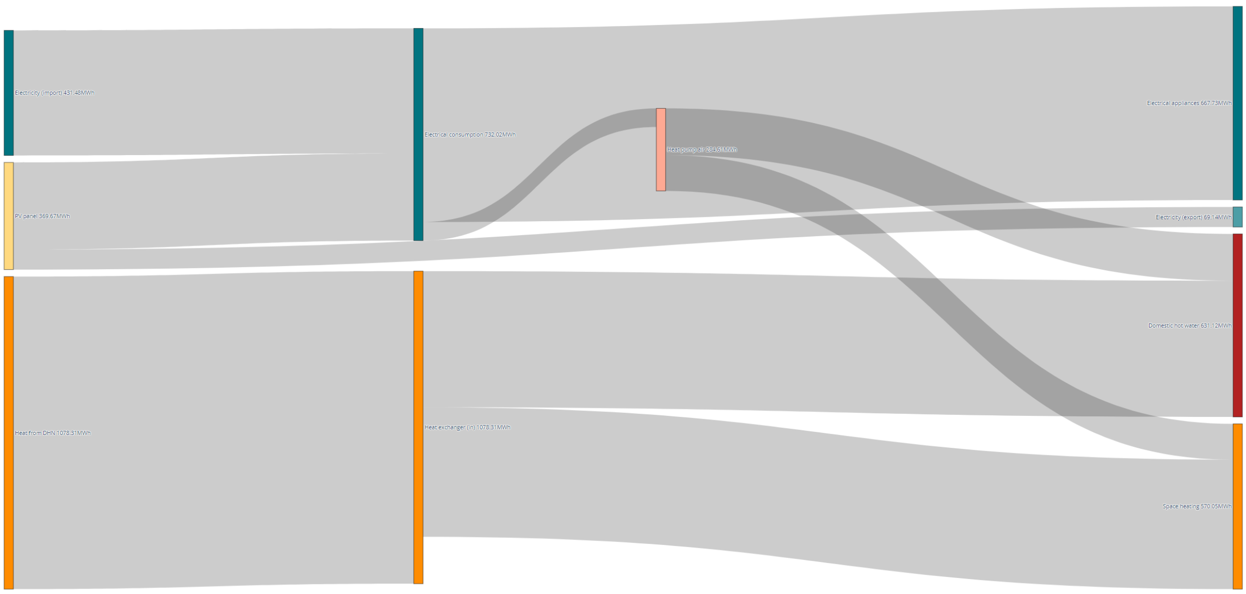

Figure 30: Sankey Diagram of Energy and Material Flows for Fossil-Free Scenario GWP Optimization in Spain's Case Study

As shown in the figure, fossil fuel-based technologies have been completely replaced by photovoltaics (PV). The energy system, utilizing PV panels, not only meets the district's electricity demand but also exports excess electricity to the grid, thus offsetting some of the global warming potential. Over the course of one year, the PV system produces 369.67 MWh of electricity. A portion of this energy satisfies the district's electricity demand, while 284.61 MWh is used for domestic hot water and space heating. Despite this, heat from the district heating network still plays a major role in meeting these demands. The energy profile evolution over time is illustrated below:

The energy demand is still the same as the fig 28. At the beginning and end of the year, the demand for heat is higher, resulting in an increased contribution from the district heating network (DHN). For other energy demands, the trend remains relatively stable. Future improvements to this energy system, from a GWP perspective, should focus on increasing PV panel production capacity to further reduce emissions.

4.3 Results for normalization

For the normalization process, we chose the simplest scenario, meaning there are no enforced units, no additional units to include, and no extra features. This allows us to optimize 29 times, resulting in a 29×29 matrix where each index represents an optimized objective, and each column contains the values of other indicators when optimizing for the corresponding indicator.

Normalization is performed by extracting the maximum and minimum values for each indicator and applying the following equation:

$$X_{\text{normalized}} = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \tag{15}$$

To further simplify the problem, we calculate a 29×29 Pearson correlation coefficient matrix.

$$r_{XY} = \frac{\sum(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum(X_i - \bar{X})^2 \sum(Y_i - \bar{Y})^2}} \tag{16}$$

This approach allows us to select a few representative indicators that can capture the essential characteristics of the entire dataset. Instead of optimizing all the indicators, we only need to optimize the selected representative indicators. From these, we can then derive the maximum and minimum values for all the indicators, significantly reducing the computational complexity. Normalization is done by changing the objective for the scenario. The next question is that how to reduce the complexity for normalization.

By visualizing the Pearson correlation matrix with a heatmap, as shown in Figure 31, we can identify three distinct clusters of indicators. For the detailed code scripts, newly generated figures, and CSV data, please refer to my GitHub repository.

The heatmap reveals three distinct areas of correlation. Indicators from CCEQL to PWEQS, and from MEU to PWEU, including LTBDV and TTEQ, exhibit strong positive correlations with each other and strong negative correlations with indicators such as IREQ, LOBDV, TPW, WAVHH, and THH, forming a second cluster. The remaining indicators belong to a third group, as they display slight positive correlations with the first group and slight negative correlations with the second group. To further refine the selection process, the K-means clustering method is applied to identify three representative indicators from these clusters.

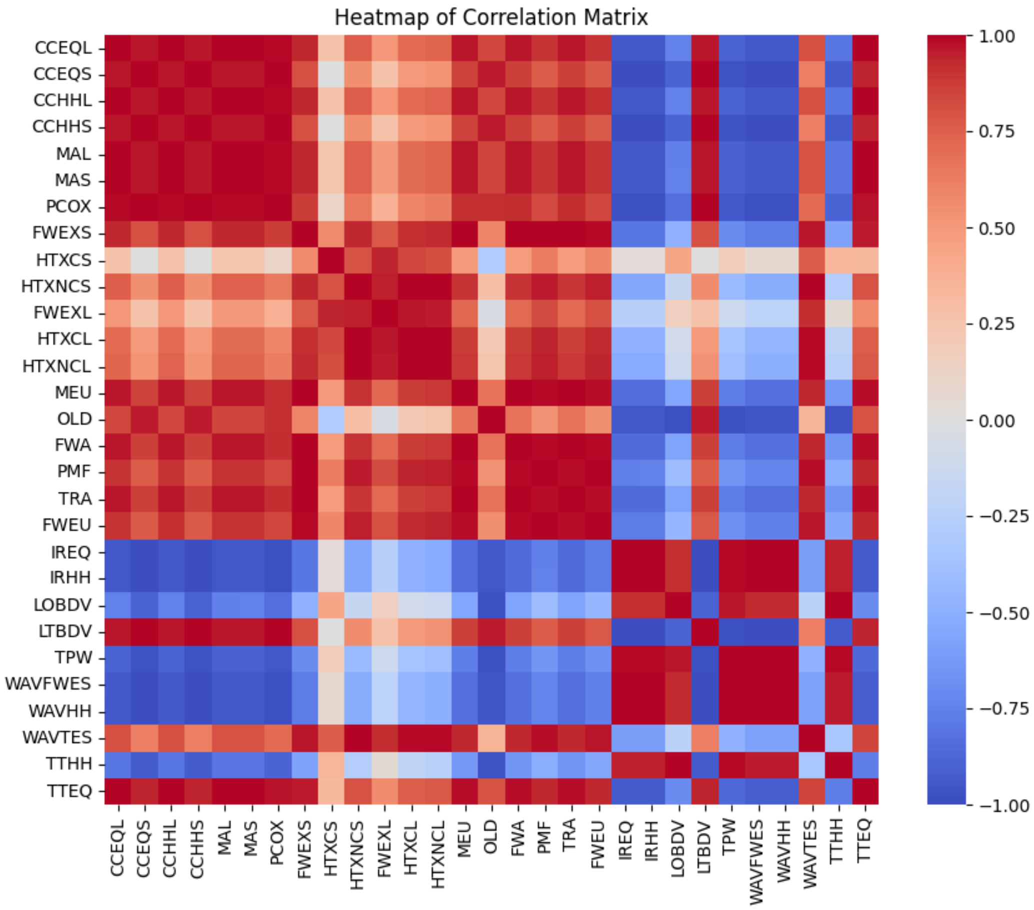

Figure 31: Pearson Correlation Matrix Heatmap

Thus, by means of k-means clustering method, we could get 3 representative indicators:

```
Selected representative indicators from each cluster:
Cluster
0    CCEQL
1     IREQ
2    HTXCS
```

Figure 32: K-means Selected Indicators

Instead of optimizing all indicators, we focus on optimizing three representative categories: Climate Change, Ecosystem Quality, Long Term (CCEQL); Ionizing Radiation, Ecosystem Quality (IREQ); and Human Toxicity, Cancer, Short Term (HTXCS). This allows us to derive a comprehensive 3x29 environmental impact matrix for all indicators. From this matrix, we extract the maximum value for each indicator, significantly reducing time complexity. PCA is used for clusters visualization as the Appendix Fig 35 shows.

4.4 Results for double counting removal

In fact, all the results above are on the basis of double counting removed database. But to make results more comprehensive, the intermediate results of this process will be analyzed in this section.

As previously mentioned, to create a more reliable and convincing database, we addressed the issue of double counting by eliminating repeated energy and material flows within the energy system. This process ensures that no flow is accounted for more than once, leading to more accurate results. Several intermediate results were generated to demonstrate the impact of double counting removal, highlighting how this step refined the overall data and improved the consistency of energy and material flow tracking throughout the system. These results provide a clearer picture of the system's actual performance and efficiency after the correction.

Table 3: Energy Technologies Double Counting Removal (Compact)

Operation	**Value**	**Unit**	**Removed Activity**
NG_Boiler, Operation	6.86e-07	unit	gas boiler (GLO)
NG_Boiler, Operation	0.02865	m^3	natural gas (CH)
OIL_Boiler, Operation	0.0249	kg	light fuel oil (CH)
OIL_Boiler, Operation	7.04e-07	unit	oil boiler, 10kW (GLO)
WOOD_Stove, Operation	1.39e-07	unit	wood furnace, 50kW (CH)
WOOD_Stove, Operation	0.0705	kg	wood chips (RER)
HP_Air, Operation	0.0992	kWh	electricity (CH)
Elec. transf. med to low	1.0276	kWh	elec. (med voltage, CH)
PV, Operation	1.2051e-05	unit	PV, slanted roof (CH)
NG_Cogeneration, Operation	0.2377	m^3	natural gas (CH)

The column titled "Removed Activity" lists the flows that were removed from the corresponding entries in the "Operation" column. For instance, 6.86e-07 units of gas boiler (GLO) construction were removed from "*NG_Boiler, Operation.*" This method allows us to easily validate the double counting removal process by comparing these intermediate results with the activity exchanges in the Ecoinvent database. Through this comparison, we can ensure that all duplicated flows have been accurately identified and removed, thereby enhancing the reliability and integrity of the dataset.

There are additional intermediate files, such as double_counting_removal.csv and double_counting_removal_co which provide a more direct representation of the removed energy and material flows. These files contain columns for energy and material flows, with indexes corresponding to different technologies. They show both the amount of flows and the number of flows removed for each technology. For instance, the figure below illustrates the count of double counting flows removed. In the case of the Electrical Heater, the count reaches 31. This is because the operation for the Electrical Heater in Ecoinvent relies on the "market for electricity (CH)," which encompasses multiple electricity generation methods. Since each method involves different locations, the number of removed flows is relatively high for this technology.

These files offer a clear and structured way to trace and quantify the double counting removal process, ensuring transparency and accuracy in the adjustment of the dataset.

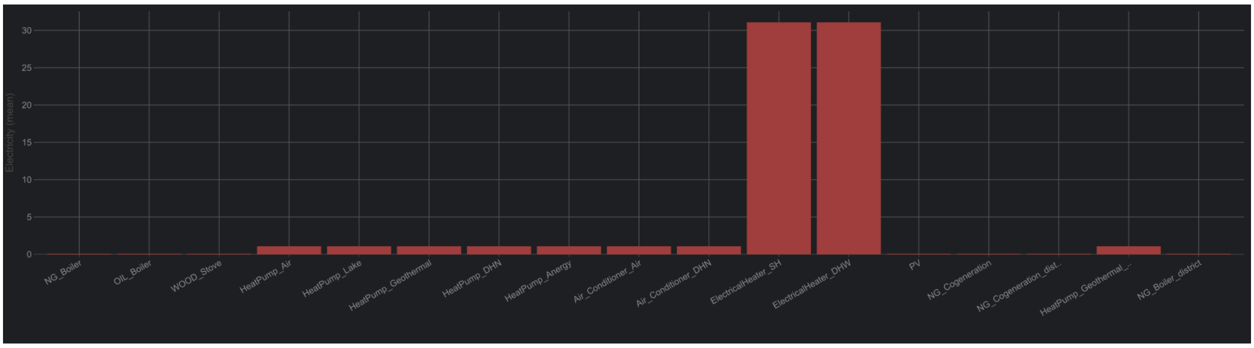

Figure 33: Double Counting Removal Count for Each Technology

For more detailed results on the newly created database, and due to the size and complexity of the datasets, please refer to my GitHub repository. The relevant outputs can be found in the directory `REHO/REHO_database/results`.

5 Conclusion

Implementation of a New LCA Methodology

In this research, I successfully implemented a new Life Cycle Assessment (LCA) methodology, integrating the LCA framework from the Energyscope energy system tool into the existing model. The generalization of LCA indicators through Brightway2, combined with the incorporation of World IMPACT+ methods, allowed for a more comprehensive environmental impact assessment. This integration ensures that a wider array of environmental factors is considered, contributing to more robust analyses.

Improvement in Data Accuracy

To address the issue of double-counting within the ecoinvent database, I applied machine learning techniques to create mapping files. These files enabled the accurate mapping of REHO technologies to corresponding activities in the ecoinvent database. The development of model characteristic files, unit conversion profiles, and other data files facilitated a seamless integration between the energy system model and ecoinvent. This approach significantly improved the accuracy of the data used for the LCA.

Integration and Validation of the New Framework

The enhanced LCA framework was successfully incorporated into both the AMPL model and Python scripts, enabling more efficient optimization. A comparative case study for Sion, utilizing both the old and new databases, highlighted the advancements made through this new methodology. The case study results provided strong validation of the methodology's effectiveness and its potential to enhance environmental impact assessments.

Real-World Application and Insights

I applied the new REHO version to a real-world case study in Spain. Through correlation analysis and the use of K-means clustering, representative LCA indicators were selected and employed in multi-objective optimization. This provided valuable insights into the environmental impact of different energy system configurations, illustrating the practical benefits of the new LCA methodology.

Overall Impact and Future Potential

The integration of the new LCA framework demonstrated significant improvements in both the accuracy and comprehensiveness of environmental assessments. The case studies conducted offer strong validation of the approach, and the results from the multi-objective optimization emphasize its applicability to real-world scenarios. This research lays a solid foundation for future applications, with the potential to further improve decision-making in energy system planning.

6 Discussion

Data Availability Gaps

Despite efforts to ensure comprehensive data coverage, certain technologies and resources were not adequately represented in the ecoinvent database. Notably, information on Information and Communication Technology (ICT) and data heat was unavailable, as these components are not included in the database. This limitation in data availability hindered a full assessment of these aspects within the study.

District Heating Network Generalization Challenges

The generalization of certain features, particularly the District Heating Network (DHN) scenario, posed challenges. The specific modeling approach used for DHN in this project limited its ability to represent the broader variability of such systems. Further refinement in the modeling of DHN is necessary to ensure more accurate and generalizable LCA results.

Partial Completion of Multi-Objective Optimization

Due to time constraints, the multi-objective optimization process was only partially completed. One of the major tasks left unfinished was the allocation of weights for each Life Cycle Assessment (LCA) indicator, which was intended to be handled by a generative algorithm. The incomplete optimization limits the comprehensiveness of the results and points to an area that requires further attention in future studies.

Double Counting Removal Database Globalization

The generation of the double counting removal database was another area with incomplete coverage. Only the database for Switzerland was fully developed, although ecoinvent primarily includes data from Switzerland and Canada. This gap limits the robustness of the current study's findings, as a more expansive dataset is needed to avoid double counting across other regions. A sensitivity analysis should be conducted to assess the implications of this limitation.

References

[1] P. Stott, "How climate change affects extreme weather events," *Science*, vol. 352, no. 6293, pp. 1517–1518, Jun. 2016, Publisher: American Association for the Advancement of Science. [Online]. Available: https://www.science.org/doi/full/10.1126/science.aaf7271 (visited on 09/17/2024).

[2] A. Cogato, F. Meggio, M. De Antoni Migliorati, and F. Marinello, "Extreme Weather Events in Agriculture: A Systematic Review," en, *Sustainability*, vol. 11, no. 9, p. 2547, Jan. 2019, Number: 9 Publisher: Multidisciplinary Digital Publishing Institute. [Online]. Available: https://www.mdpi.com/2071-1050/11/9/2547 (visited on 09/17/2024).

[3] P. D. Lund, J. Lindgren, J. Mikkola, and J. Salpakari, "Review of energy system flexibility measures to enable high levels of variable renewable electricity," *Renewable and Sustainable Energy Reviews*, vol. 45, pp. 785–807, May 2015. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1364032115000672 (visited on 09/17/2024).

[4] A. Savaresi, "The Paris Agreement: A new beginning?" *Journal of Energy & Natural Resources Law*, vol. 34, no. 1, pp. 16–26, Jan. 2016, Publisher: Routledge _eprint: https://doi.org/10.1080/02 [Online]. Available: https://doi.org/10.1080/02646811.2016.1133983 (visited on 09/17/2024).

[5] S. C. Bhattacharyya and G. R. Timilsina, "A review of energy system models," *International Journal of Energy Sector Management*, vol. 4, no. 4, pp. 494–518, Jan. 2010, Publisher: Emerald Group Publishing Limited. [Online]. Available: https://doi.org/10.1108/17506221011092742 (visited on 09/17/2024).

[6] C. Milan, C. Bojesen, and M. P. Nielsen, "A cost optimization model for 100% renewable residential energy supply systems," *Energy*, 6th Dubrovnik Conference on Sustainable Development of Energy Water and Environmental Systems, SDEWES 2011, vol. 48, no. 1, pp. 118–127, Dec. 2012. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0360544212004148 (visited on 09/17/2024).

[7] M. Z. Hauschild, "Introduction to LCA Methodology," en, in *Life Cycle Assessment: Theory and Practice*, M. Z. Hauschild, R. K. Rosenbaum, and S. I. Olsen, Eds., Cham: Springer International Publishing, 2018, pp. 59–66. [Online]. Available: https://doi.org/10.1007/978-3-319-56475-3_6 (visited on 09/17/2024).

[8] J. B. Guinée, H. A. Udo de Haes, and G. Huppes, "Quantitative life cycle assessment of products: 1:Goal definition and inventory," *Journal of Cleaner Production*, vol. 1, no. 1, pp. 3–13, Jan. 1993. [Online]. Available: https://www.sciencedirect.com/science/article/pii/0959652693900279 (visited on 09/17/2024).

[9] A. Jørgensen, A. Le Bocq, L. Nazarkina, and M. Hauschild, "Methodologies for social life cycle assessment," en, *The International Journal of Life Cycle Assessment*, vol. 13, no. 2, pp. 96–103, Mar. 2008. [Online]. Available: https://doi.org/10.1065/lca2007.11.367 (visited on 09/17/2024).

[10] A. Bjørn, M. Owsianiak, C. Molin, and M. Z. Hauschild, "LCA History," en, in *Life Cycle Assessment: Theory and Practice*, M. Z. Hauschild, R. K. Rosenbaum, and S. I. Olsen, Eds., Cham: Springer International Publishing, 2018, pp. 17–30. [Online]. Available: https://doi.org/10.1007/978-3-319-56475-3_3 (visited on 09/17/2024).

[11] A. Laurent, N. Espinosa, and M. Z. Hauschild, "LCA of Energy Systems," en, in *Life Cycle Assessment: Theory and Practice*, M. Z. Hauschild, R. K. Rosenbaum, and S. I. Olsen, Eds., Cham: Springer International Publishing, 2018, pp. 633–668. [Online]. Available: https://doi.org/10.1007/978-3-319-56475-3_26 (visited on 09/17/2024).

[12] D. Lepour, C. Terrier, and J. Loustau, *Renewable Energy Hub Optimizer (REHO) - A Comprehensive Decision Support Tool for Sustainable Energy System Planning.* [Online]. Available: https://github.com/IPESE/REHO (visited on 09/18/2024).

[13] G. Limpens, S. Moret, H. Jeanmart, and F. Maréchal, "EnergyScope TD: A novel open-source model for regional energy systems," *Applied Energy*, vol. 255, p. 113 729, Dec. 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0306261919314163 (visited on 09/11/2024).

[14] Z. Ma, "Estimating Future Costs and Carbon Footprints of PEMEC and SOEC Manufacturing," en,

[15] *Fixed and variable costs - Policonomics*, en-US, Nov. 2012. [Online]. Available: https://policonomics.com/typology-of-costs/ (visited on 09/10/2024).

[16] J. Schnidrig, J. Brun, F. Maréchal, and M. Margni, "Integration of Life Cycle Impact Assessment in Energy System Modelling," en, in *36th International Conference on Efficiency, Cost, Optimization, Simulation and Environmental Impact of Energy Systems (ECOS 2023)*, Las Palmas De Gran Canaria, Spain: ECOS 2023, 2023, pp. 2810–2822. [Online]. Available: http://www.proceedings.com/069564-0252.html (visited on 06/21/2024).

[17] C. Mutel, "Brightway: An open source framework for Life Cycle Assessment," en, *The Journal of Open Source Software*, vol. 2, no. 12, p. 236, Apr. 2017. [Online]. Available: http://joss.theoj.org/papers/10.21105/joss.00236 (visited on 09/11/2024).

[18] G. Wernet, C. Bauer, B. Steubing, J. Reinhard, E. Moreno-Ruiz, and B. Weidema, "The ecoinvent database version 3 (part I): Overview and methodology," en, *The International Journal of Life Cycle Assessment*, vol. 21, no. 9, pp. 1218–1230, Sep. 2016. [Online]. Available: https://doi.org/10.1007/s11367-016-1087-8 (visited on 09/11/2024).

[19] C. Bulle, M. Margni, L. Patouillard, *et al.*, "IMPACT World+: A globally regionalized life cycle impact assessment method," en, *The International Journal of Life Cycle Assessment*, vol. 24, no. 9, pp. 1653–1674, Sep. 2019. [Online]. Available: https://doi.org/10.1007/s11367-019-01583-0 (visited on 09/11/2024).

[20] M. Souttre, *Mescal — mescal.* [Online]. Available: https://mescal.readthedocs.io/en/latest/index.html (visited on 09/11/2024).

[21] M. F. Mohr and A. S. Russell, "Concepts and Process of Identifying Service Products," en,

[22] *CIRAIG/IE_ml_mapping: Module to automate mapping of classifications based on machine learning word association.* [Online]. Available: https://github.com/CIRAIG/IE_ML_mapping?tab=readme-ov-file (visited on 09/12/2024).

[23] P. Sedgwick, "Pearson's correlation coefficient," en, *BMJ*, vol. 345, e4483, Jul. 2012, Publisher: British Medical Journal Publishing Group Section: Endgames. [Online]. Available: https://www.bmj.com/content/345/bmj.e4483 (visited on 09/16/2024).

[24] N. Gunantara, "A review of multi-objective optimization: Methods and its applications," *Cogent Engineering*, vol. 5, no. 1, Q. Ai, Ed., p. 1 502 242, Jan. 2018, Publisher: Cogent OA _eprint: https://doi.org/10.1080/23311916.2018.1502242. [Online]. Available: https://doi.org/10.1080/23311916.2018.1502242 (visited on 09/14/2024).

[25] I. Giagkiozis and P. J. Fleming, "Methods for multi-objective optimization: An analysis," *Information Sciences*, vol. 293, pp. 338–350, Feb. 2015. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0020025514009074 (visited on 09/14/2024).

[26] Y. Du, L. Xie, J. Liu, Y. Wang, Y. Xu, and S. Wang, "Multi-objective optimization of reverse osmosis networks by lexicographic optimization and augmented epsilon constraint method," *Desalination*, vol. 333, no. 1, pp. 66–81, Jan. 2014. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0011916413005109 (visited on 09/14/2024).

[27] *Pilot 1: Lugaritz-Matía Community - InterPED*, en-US, May 2024. [Online]. Available: `https://interped.eu/pilot-1-lugaritz-matia-community/` (visited on 09/14/2024).

[28] K. T. GmbH (https://www.klokantech.com/), *Hartebeesthoek94 / Lo31 - EPSG:2054*, en. [Online]. Available: `https://epsg.io` (visited on 09/16/2024).

A Appendix

For all the codes, please refer to [my github].

More specifically:

- **REHO Methodology Integration**:
  `./REHO/reho`
- **REHO Database Generation**:
  - For database generation results:
    `./REHO_databases/results`
  - For double counting removal:
    `./REHO_databases/datasets`
  - For mapping files:
    `./REHO_databases/REHO_data/CH`
- **Normalization**:
  `./reho/normalization`
- **Case Study**:
  - Scenario 1:
    `./VEO_Spain/scenario1`
  - Scenario 2:
    `./VEO_Spain/scenario2`
- **Scripts for Case Study Results**:
  `./scripts/example/3a_read.csv`

District level for REHO

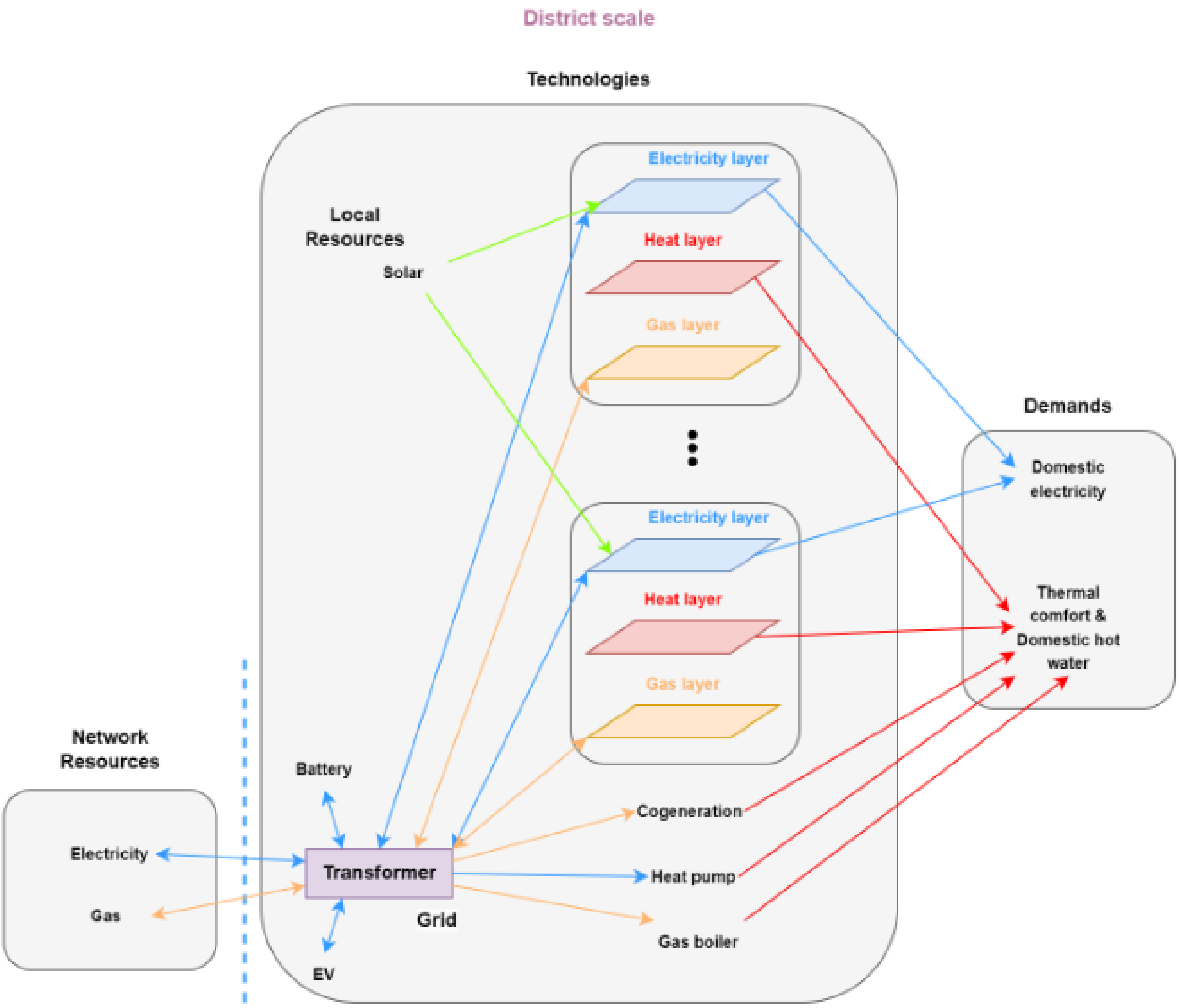

Figure 34: Predefined Energy System Model at District Scale in the Renewable Energy Hub Optimizer (REHO)

The primary difference between district-scale and building-scale systems lies in whether the network is treated as equivalent to the grid. At the district level, the network encompasses not only the grid but also additional connections between the transformer and the district units.

Mapping files

All the required files generated manually can be found in [my github]

Visulization of clusters by means of PCA

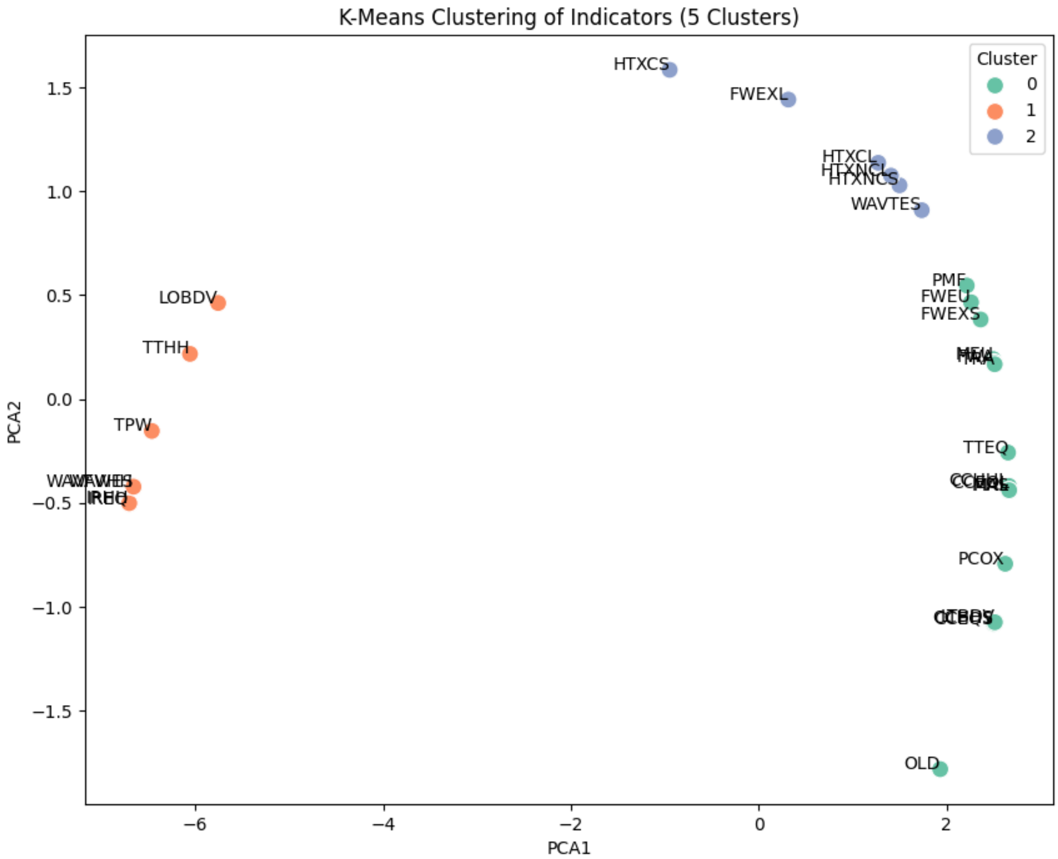

Figure 35: Three Clusters Classified by Principal Component Analysis (PCA)

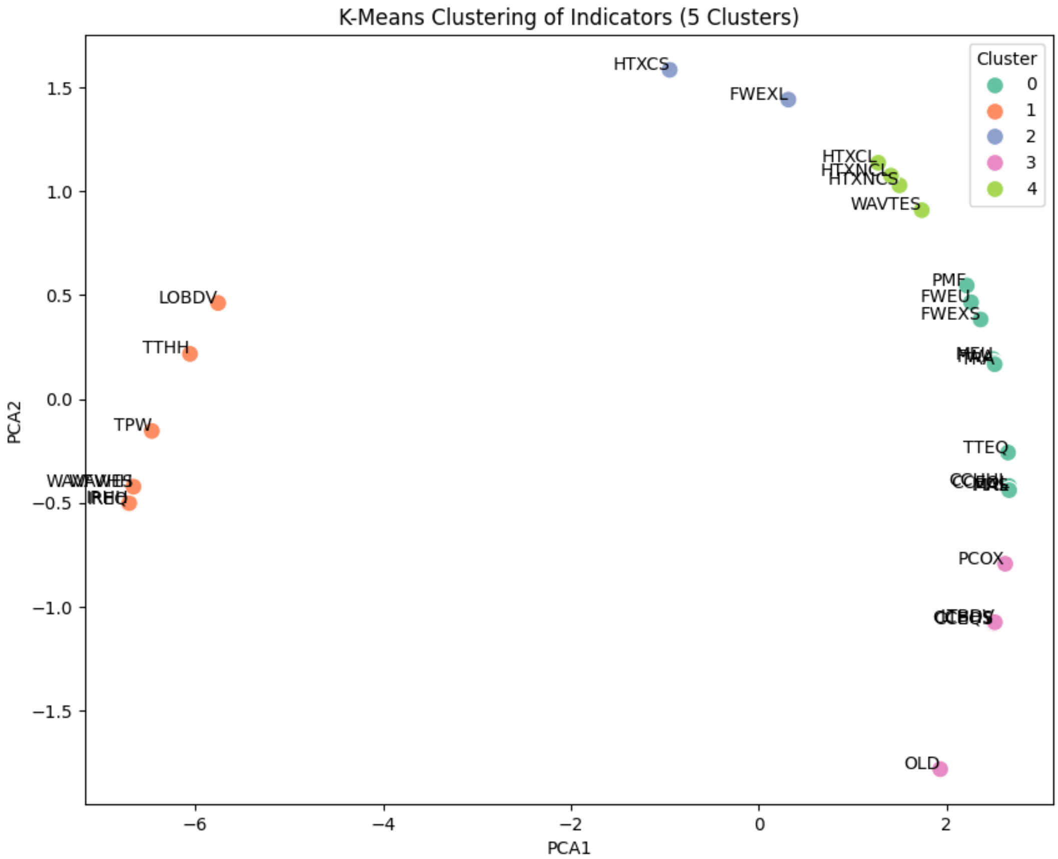

Figure 36: Five Clusters Classified by Principal Component Analysis (PCA)

Detailed description of building characters in Spain

	PARAMETERS	HISTORICAL	Berminghan	Rezola	Lugaritz
NAME	Building identifier	NO	BER	REZ	LUG
TYPE	Building type	NO	Hospital	Nursing home	Nursing home (54 apartments + 40 day beds) + Office building
ADDRESS	Town or address	NO	San Sebastián, Basque Country (Spain)Matia - Bermingham Ospitalea	an Sebastián, Basque Country (Spain)Centro Gerontológico Julián Rezol	ebastián, Basque Country (Spain)Matia - Lugaritz - Babesdun Etxebi
COORDINATES	Latitude	NO	43,3038	43,3034	43,3045
	Longitude	NO	-2,0014	-2,0031	-2,0024
NOTE	If any additional information is needed	NO	The DH is located in Lugaritz building	The DH is located in Lugaritz building	The DH is located in Lugaritz building
GEOMETRY	Floor number	NO	9	4	7
	Height up to the last ceiling [m]	NO	24,65	11	19,5
	Floor area [m2]	NO	18 000	Floor 0: 1,459.42; Floor 1: 1,589.35;Floor 2: 1,517.53; Floor 3: 742.92	,269.87; Floor 2: 3,214.9; Floor 3: 3,335.5; Floor 4: 3,244.18; Floor
	Vertical facades area [m2]	NO	25 932	1 235	6 424
	Vertical windows area [m2]	NO	300	68	163
	Solar vertical facades area [m2]	NO	12 966	600	3 200
	Solar roof area [m2]	NO	330	800	220
EPB	Construction period [y]	NO	2011	1989	2023
	Yearly electricity demand [kWh/y]	NO	368,000 for heating	N/A	30 for cooling
	Yearly heating demand [kWh/y]	NO	1 352 000	979 000	910 000
	Yearly DHW demand [kWh/y]	NO	425 000	246 000	56 940
	Averaged thermal transmittance [kW/m2·K]	NO	Facade: k=0.66 Roof: k=0.38	Facade: k=1.8 Roof: k=1.4	Facade: k=0.27 Roof: k=0.22
	Thermal specific capacity [Wh/m2·K]	NO			
USAGE		NO	[Hospital, Swimming pool]	Residential (Collective housing)	[Residential (Collective housing), Administrative]
	Percentage of each utilisation [%]	NO	[87.5, 12.5]	100	[70, 30]
	Average users of the building	NO	336	143	148
	Electrical appliances consumption (Low, Medium, High)	NO	High	High	High
THERMAL	Heating technology (indicate all available)	NO	[DH(Biomass, Condensing boiler), Conv. boiler, ASHP, Solar thermal]	[DH(Biomass, Condensing boiler), Conv. boiler, Condensing boiler]	[DH(Biomass, Condensing boiler), GSHP, Chiller recov.]
	Heating technology power [kW]	NO	[DH(500+300, 620+620), 270+270+270, 16, 40x0.4]	[DH(500+300, 620+620), 270, 115]	[DH(500+300, 620+620), 160, 14.5]
	Heating technology energy [kWh] (include link to time seri	YES	Datos Berminghan desde SCadaDatos Lugaritz desde SCada	Datos Rezola desde SCada Datos Lugaritz desde SCada	Datos Lugaritz desde SCada
	Heating storage volume [m3]	NO	2x2[DH(4, N/A), N/A, N/A, 2.5]	[DH(4, N/A, N/A, N/A), N/A, N/A]	3.1[DH(4, N/A, N/A, N/A)]
	Heating storage maximum power [kW]	NO			
	Heating storage capacity [kWh]	NO			
	DHW technology (indicate all available)	NO	[DH(Biomass, Condensing boiler), Conv. boiler, ASHP, Solar thermal]	[DH(Biomass, Condensing boiler), Conv. boiler, Condensing boiler]	[DH(Biomass, Condensing boiler), GSHP, Chiller recov.]
	DHW technology power [kW]	NO	[DH(500+300, 620+620), 270+270+270, 16, 40x0.4]	[DH(500+300, 620+620), 270, 115]	[DH(500+300, 620+620), 160, 14.5]
	DHW technology energy [kWh] (include link to time series)	YES	Datos Berminghan desde SCadaDatos Lugaritz desde SCada	Datos Rezola desde SCada Datos Lugaritz desde SCada	Datos Lugaritz desde SCada
	DHW storage volume [m3]	NO	2x2[DH(4, N/A), N/A, N/A, 2.5]	2x0.8[DH(4, N/A), N/A, N/A]	3.1[DH(4, N/A), N/A, N/A]
	DHW storage maximum power [kW]	NO			
	DHW storage capacity [kWh]	NO			
	Cooling technology (indicate all available)	NO	N/A	N/A	[Chiller, GSHP]
	Cooling technology power [kW]	NO	N/A	N/A	[218, 141]
	Cooling technology energy [kWh] (include link to time seri	YES	N/A	N/A	Datos Lugaritz desde SCada
	Cooling storage volume [m3]	NO	N/A	N/A	1.5[N/A, N/A]
	Cooling storage maximum power [kW]	NO	N/A	N/A	N/A
	Cooling storage capacity [kWh]	NO	N/A	N/A	N/A
	Target or ideal internal temperature [°C]	NO	21 (Winter) / 25 (Summer)	21 (Winter) / 25 (Summer)	21 (Winter) / 25 (Summer)
	Heating supply temperature [°C]	NO	80 (Winter) / 70 (Summer)	80 (Winter) / 70 (Summer)	80 (Winter) / 70 (Summer)
	Heating return temperature [°C]	NO	60 (Winter) / 50 (Summer)	60 (Winter) / 50 (Summer)	60 (Winter) / 50 (Summer)

Figure 37: Raw File of Building Characteristics for Spain's Case Study - 1

	Cooling supply temperature [°C]	NO	7	7	7
	Cooling return temperature [°C]	NO	12	12	12
HEAT/RECOVERYTECHN	Waste heat sources (indicate all available)	NO	N/A	N/A	Heat recovery from chiller / Heat recovery from air extraction
	Heat transfer fluid to be recovered (Liquid water, Steam, A	NO	N/A	N/A	Water / Air
	Heat transfer fluid temperature [°C] (include link to time se	YES	N/A	N/A	35-40 Datos Lugaritz desde SCada
	Heat transfer fluid pressure [bar]	NO	N/A	N/A	1
	Heat transfer fluid flow (include link to time series)	YES	N/A	N/A	Datos Lugaritz desde SCada
	Distance between heat source and final use	NO	N/A	N/A	
	Space available next to the heat source	NO	N/A	N/A	
	Recovery technologies already installed (indicate all availat	NO	N/A	N/A	DH(Chiller heat recovery)
	Recovery technology power [kW]	NO	N/A	N/A	10
	Recovery technology energy [kWh] (include link to time se	YES	Datos Berminghan desde SCadaDatos Lugaritz desde SCada	Datos Rezola desde SCada Datos Lugaritz desde SCada	Datos Lugaritz desde SCada
ELECTRIC	Electricity consumption profile [kWh] (include link to time	YES	N/A	N/A	N/A
	PV capacity [kW]	NO	N/A	N/A	16 335
	PV production [kWh] (include link to time series)	YES	N/A	N/A	17,839 Datos Lugaritz desde SCada
	Wind turbines capacity [kW]	NO	N/A	N/A	N/A
	Wind turbines production [kWh] (include link to time serie	YES	N/A	N/A	N/A
	Battery maximum discharge power [kW]	NO	N/A	N/A	N/A
	Battery capacity [kWh]	NO	N/A	N/A	N/A
	Battery minimum SoC fixed [%] / [kWh]	NO	N/A	N/A	N/A
	Battery maximum SoC fixed [%] / [kWh]	NO	N/A	N/A	N/A
	Battery SoC [%] / [kWh] (include link to time series)	YES	N/A	N/A	N/A
	EV charger power [kW]	NO	N/A	N/A	N/A
	EV battery capacity [kWh]	NO	N/A	N/A	N/A
	EV minimum SoC fixed [%] / [kWh]	NO	N/A	N/A	N/A
	EV maximum SoC fixed [%] / [kWh]	NO	N/A	N/A	N/A
	EV SoC [%] / [kWh] (include link to time series)	YES	N/A	N/A	N/A
	Other [?]	NO	N/A	N/A	N/A
GRID	Feed-in tariff rate [€/kWh]	NO	N/A	N/A	N/A
	Energy price of the community member [€/kWh]	NO	N/A	N/A	N/A
	Grid tariffs / network costs [€/kWh]	NO	N/A	N/A	N/A
	Type of trading allowed if any	NO	N/A	N/A	N/A

Figure 38: Raw File of Building Characteristics for Spain's Case Study - 2

UNIVERSITÉ CATHOLIQUE DE LOUVAIN
École polytechnique de Louvain
Rue Archimède, 1 bte L6.11.01, 1348 Louvain-la-Neuve, Belgique | **www.uclouvain.be/epl**