D3.1

Technical Documentation of the WHY-toolkit

LEGAL DISCLAIMER

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 891943.

DOCUMENT INFORMATION

Deliverable title	Technical Documentation of the WHY-tool
Dissemination level	Public
Submission deadline	28/02/2022
Version number	1
Authors	Thomas Nacht (4ER) Johanna Ganglbauer (4ER) Cruz Enrique Borges Hernandez (UD) Armando Aguayo Mendoza (UD) Ruth Carbajo Garcia (UD) Noah Pflugradt (FZJ) Panagiotis Fragkos (E3M)
Internal reviewers	Thomas Nacht (4ER) Panagiotis Fragkos (E3M) Cruz Enrique Borges Hernandez (UD)
External peer reviewers	Not needed to be reviewed.
Document approval	Not needed to be approved.
Scope of the document according to the DoA	The report will include the description of the different models and components developed. The report describes the architecture of the WHY Toolkit and its components as well as the public APIs.

EXECUTIVE SUMMARY

This deliverable contains the definition of the WHY Model Architecture. The architecture is described by **Code-Elements**, a collection of models and modules designated to perform a certain task within the WHY Model Architecture. Code-elements are linked by **Interfaces**. Each Interface is described by their corresponding data format and the data they will contain. Lastly the architecture contains **Components**, which are the key parts that make up the Code-Elements and are either Models, Data or Algorithms performing different tasks in the project.

The WHY Model Architecture presented in this Deliverable is a preliminary version that will be developed further in the course of the project when the different elements will be developed. This preliminary version was developed to provide first insights into the inner workings of the WHY Toolkit. Fig. 1 shows a simplified version of the architecture with only the Code-Elements considered. Between each of the Code-Elements is an Interface which provides the required data in the necessary format. The Components setting up the different Code-Elements are mentioned in the list below. A detailed description can be found in the corresponding section of the Deliverable.

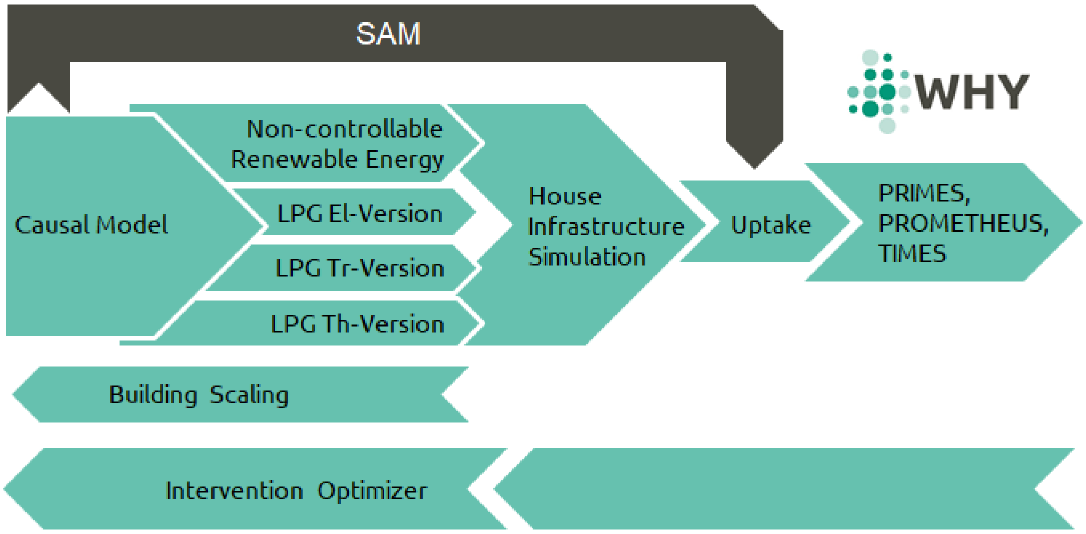

Figure 1: Architecture, only Code-Elements considered

Causal Model:

- Environment Data:
  - Climate Data projections
  - Exogenous Data projections
  - Legal and interventions
  - Grid Data
  - Geographic Data
- Behavioural Models:
  - Data on energy use classes
  - Energy behaviour classifier Model
  - Investment decision Model
  - Energy needs model
  - Thermal needs model
  - Transport needs model

Non-controllable Renewable Energy:

- Energy Generation Model

Building Scaling:

- Investment Effects Model
- Technology Data

Upscale:

- Upscale Algorithm

Others:

- Sustainability Assessment Model
- Intervention Optimizer

LPG Tr-Version:

- Transport Data
- Transport needs model

LPG El-Version:

- Energy Generator Models
- Non-Controllable Appliance Models
- Multi Agent System Model

LPG Th-Version

- Building Data and Model
  - Thermal Model
  - Electrical Data

House Infrastructure Simulation:

- Stationary Energy Storage Models: Batteries
- Movable Energy Storage Models: EV
- Power2Gas Models
- Transport Energy Models
- Controllable Appliances Models
- Energy Management System Model

PRIMES, PROMETHEUS, TIMES:

- Energy System Models

TABLE OF CONTENTS

LIST OF ACRONYMS AND ABBREVIATIONS

Acronym	Long text
API	Application Programming Interface
ATE	Average Treatment Effect
BMS	Building Scaling Model
BuiMo	Building Module
CAPEX	Capital Expenditures
CATE	Conditional Average Treatment Effect
CCS	Carbon Capture and Storage
CSV	Comma Separated Values
DAC	Directly Acyclic Graphs
ESM	Energy System Model
EV	Electric Vehicles
IEEE	Institute of Electrical and Electronics Engineers
ITE	Individual Treatment Effect
JSON	JavaScript Object Notation
KPI	Key Performance Indicator
LPG	Load Profile Generator
LPG-Tr	Load Profile Generator Transport
LPG-El	Load Profile Generator Electricity
LPG-Th	Load Profile Generator Thermal
MAS	Multi-Agents-System
MILP	Mixed Integer Linear Programming
OPEX	Operational Expenditures
PRIMES	Price-Induced Market Equilibrium System
PV	Photovoltaics
SCM	Structured Causal Model

1 Introduction

The project WHY aims at the development of the WHY Toolkit which will provide the means to better understand the consumption behaviour of European households, considering both technical models and behavioural models in the form of a causal model. The software will thus enable the users to:

- make better forecasts for electricity consumption in households,
- analyse, evaluate, and validate policy decisions or other interventions such as changes in regulation, policy measures, funding, etc.,
- examine how today's world would look like if certain energy policy decisions had or had not been taken.

The WHY Toolkit will consist of many different models which interact with one another to provide the results mentioned above. These models make up different core components of the WHY Toolkit:

- The Causal Model responsible for setting a relation between a cause and a decision made by inhabitants of residential buildings
- Building Scaling and Simulation models considering electrical and thermal consumption as well as generation capacities
- Large Scale Energy System Models which will allow the upscaling of individual households to analyse large scale scenarios at national, EU and global levels

These different groups of models need to be put together in a common architecture and the methods applied to these models need to be defined. Deliverable D3.1 will provide an overview of the methods and approaches used in the models and will highlight how they will interact. The Deliverable will contain the following chapters:

1. **General WHY Model Architecture:**
   This chapter will introduce the WHY Model Architecture and show how the different components of the WHY Toolkit will interact with one another. Within the chapter the different components will be described in a superficial manner to provide a general overview and understanding of the components and how they are intended to work.
2. **Models of the WHY Model Architecture**
   While chapter 1 provides a general overview of the Architecture and the models, chapter 2 will provide a more detailed perspective on the general methodology applied to each of the models. This description will not be an in depth description of the coding of these models but rather an explanation of what the models should do and how they plan to do it.

2 General WHY Model Architecture

The WHY model architecture is the description of the main components of the different models developed during the WHY project and their corresponding interactions. Furthermore, it links the components defined in the Volere Requirements Specification Template (see Deliverable D1.1) with the general structure of the models. To convey the components and their interactions with the model architecture, as well as the required APIs between the different modelling components, a representation of the WHY-Model-Architecture is presented in Fig. 2.

The components in Fig. 2 are colour-coded according to their functionality. The colour codes can be found directly in the figure. It needs to be stated at this point that the architecture in Fig. 2 is only a preliminary architecture which is developed further during the course of the project as the different model componentes are defined.

Following the figure a very brief description of the main code blocks, interfaces and components is provided, giving a general overview and understanding of the WHY model architecture. In the subchapters of chapter 3 the methods of the different main code blocks are described in further detail to provide a more in depth insight into the realised approach and functionality.

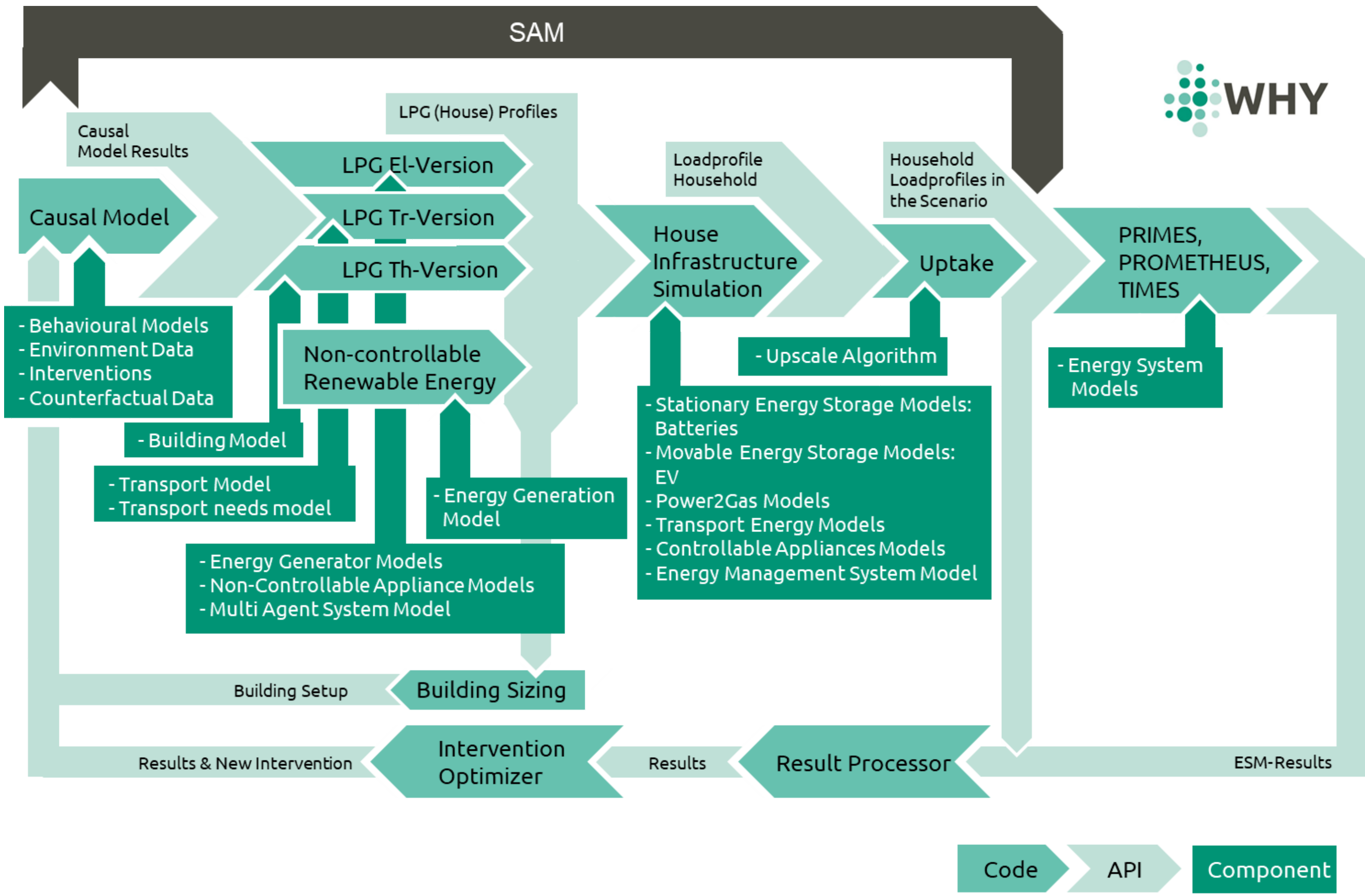

Figure 2: WHY Model Architecture

2.1 Components

Following the Volere Requirements Methodology, it is needed to define the *Components* of the project. In the case of the WHY project, the *Components* are either the (simulation/optimisation) models or sets of data that are needed to acquire the results needed for the different *Code*-Elements as well as the ESM. The *Components* are grouped by different categories according to their overall function.

The description of the *Components* presented in this chapter is only preliminary. The detailed description and the methodology behind each of the models will be described once the modelling approaches have been researched and defined. The *Components* described in this chapter facilitate the research results provided in Deliverable D1.2.

- **Environment Data:**
    - **Climate Data:**
      This set of data contains the climate and climate-related data (wind speed, irradiation, etc.) and its future projections due to climate change, needed for forecasting the thermal load of buildings as well as supply dependent generation.
    - **Exogenous Data:**
      Input data to the different models (like population, socio-economic distribution of the population, household income distribution, future projection of these, etc.)
    - **Grid Data:**
      Data on the buses, lines and the technical parameters of an actual or representative grid (IEEE-grids) needed for load flow calculations and grid analysis.
    - **Geographic Data:**
      Data containing information on the road or transport network (including the public transport network).

- **Building Data and Model**
    - **Thermal Model:**
      The thermal model is a mathematical simulation model of the building, whose purpose is the calculation of the heat losses through the hull of the building and thermal yield through internal loads and solar yield. Thus the model can calculate the thermal demand of the building.
    - **Electrical Data:**
      List of controllable and non-controllable devices that are installed at the building (including generation, power2x, etc.).
    - **Transport Data:**
      List of transport modes that can be used and their "costs" (in terms of economic, time, emissions and other KPIs)

- **Behavioural Models**
    - **Data on energy use classes:**
      Classes of households by their energy use. This data will be a result of Task T2.1 and will be provided in Deliverable D2.1. Each class represents a certain type of household, its statistical description of the energy

consumption and the features (technical and non-technical) that this type of household has.

- **Energy behaviour classifier Model:**
  Instrument to classify each household in the different energy use classes. The instrument will be a fuzzy logic decision tree (or similar instrument) that given a load profile plus a set of socio-economic information will assign a given household to one of the energy use classes. This instrument will be a result of Task T2.1 and will be provided in Deliverable D2.1.
- **Investment decision Model:**
  For each of the energy use classes, this model will decide when and why the household represented by its energy use class will invest in any of the appliances, energy generation possibilities, the energy management system and the electrification of services (for instance mobility) studied in the project. The model will consider a large pool of socioeconomic and psychological variables and, by following a causal model, will try to mimic the mental process to decide when and why investments in any of the components happen. This model will be a result of Tasks T2.2 and T2.3 and will be provided in Deliverable D2.2 and D2.3.
- **Energy needs model**:
  A mathematical model that will estimate the user's preference and desires to use the different appliances within the household. This model will, similar to the investment decision model, try to mimic the mental process behind the actions of the inhabitants that lead to the use of an appliance. This model will be created for each of the energy use classes. This model will be a result of Tasks T2.2 and T2.3 and will be provided in Deliverable D2.2 and D2.3.
- **Thermal needs model:**
  A mathematical model that will estimate the user's preferences and desires when it comes to temperature comfort levels within the house and their preferences and desires concerning the use of warm water. This model will also apply the causal model to mimic the decisions of the inhabitants. This model will be a result of Tasks T2.2 and T2.3 and will be provided in Deliverable D2.2 and D2.3.
- **Transport needs model:**
  A mathematical model that will estimate the user's preferences and desires when it comes to transport activities. This model will also apply the causal model to mimic the decisions of the inhabitants. This model will be a result of Tasks T2.2 and T2.3 and will be provided in Deliverable D2.2 and D2.3.

- **Appliance Models**
  - **Stationary Energy Storage Models: Batteries**
    The stationary energy storages (batteries) will be represented by these mathematical simulation models. This set of models will simulate different types of battery storages (Lithium-Ion, Redox-Flow, Pb-acid, etc.) represented by their respective parameters. These models will consider constant charging and discharging efficiencies as well as the capacity and the maximal charging and discharging power. A basic charging and discharging control can also be implemented, which has the focus on balancing the exchange with the public grid (but will most probably be

implemented in the Energy Management Model). The models will be fed by an input control signal, indicating whether charging or discharging should take place, for what duration and with what power.

- **Movable Energy Storage Models: EV**
  Electric vehicles (movable batteries) will be represented by these mathematical simulation models, which are similar to the above mentioned models of the batteries: electric vehicles will be modelled as batteries, which are underlaid with the information originated from the LPG-transport version. This information includes (a) if the electric vehicle (battery) is plugged in and available for charge or not and (b) how much energy of the battery is used off while driving. Different types of electric vehicles can be simulated by implementing different parameters. For the batteries in electric vehicles, charging will be possible for each type of vehicle whereas discharging will only be possible if the type of electric vehicle will allow it. Charging does not necessarily need to happen at the "home-site" of the EV, but discharging will only happen at the "home-site". The model will consider the time when the charging occurs, as well as the duration of the charging (or and discharging) process. For discharging, two modes are possible: Active Discharge: This mode is only available for EVs which are capable of discharging at a charging station. This will only happen at the "home-site". Passive Discharge: This discharging happens due to the EV driving around. The data for that will be provided by the *Transport Energy Model.*
- **Energy Generator Models:**
  Mathematical models for different types of controllable generation capacities (such as PV, wind, heat pumps, boilers, gas heaters and oil heaters, excluding non-controllable and battery generation). These models simulate the behaviour of the energy generators that will be operated to satisfy the demand of the household.
- **Power2gas Models:**
  The Power2Gas appliance model will be a mathematical representation of an electrolyser. Generating electricity from H_2 will be tackled by the corresponding *Energy Generator Models*. The model will allow the calculation of the H_2 output of the electrolyser depending on the electrical power fed into the device. The model will also consider the internal H_2 storage as part of the appliance.
- **Transport Energy Models:**
  Mathematical models that will calculate the energy demand from moving a vehicle based on the transport needs of the inhabitants of the building. The models will consider different transport means like, regular cars, electric cars, public transport, personal travel devices, etc. Furthermore, it is planned that the models will also consider different driving styles to better reflect the behaviour of the inhabitants.
- **Controllable Appliance Models:**
  Models for all appliances that can be controlled by an external energy management system or by the inhabitants directly. Consumption appliances of this category must have some sort of smartness or controllability or be able to receive control signals from another source. These appliances can either be switched on and off or be adjusted in the power consumption (for instance, dimmable lights).

Additionally devices, which are started by the users themselves but are not directly linked to the satisfaction of a desire and do have some sort of internal storage (dish-washer, washing machine, ovens, dryers etc.), are within this category, although they would require a change in the user behaviour and a tool (preferably the energy management system), which provides the information for users to change their behaviour

 - **Non-Controllable Appliance Models:**
   These models include all appliances whose operation is directly linked to the fulfilment of a desire of the inhabitants of the building. Additionally, these devices do not have any type of storage, so the time at which the desire of the inhabitant of the building occurs is also the time at which this particular desire needs to be satisfied and the device activated. Furthermore, the power of the appliances can not be regulated nor are the devices capable of receiving external signals. Members of this category of appliances are TV, Stereos, Desktop-PCs, and other similar devices.

- **Energy Management System Model:**
  The Energy Management System Model is the mathematical representation of an actual energy management device. These devices are able to control other appliances based on internal or external signals (price signals or exchange of power with the electricity grid). Additionally, if predictive control is an option, the energy management model will also be able to operate on forecasts for load and generation (as an example) to optimise the use of appliances. Energy Management Models work on decentralised (single home application) or centralised (neighbourhood) level, resulting in an increasing degree of complexity and issues concerning competition of different devices.
  The Energy Management System Model will be developed as optimisation models as it will need to be able to evaluate and determine the optimal operation of different appliances. There can be multiple different goals and approaches for the optimisation, each of which is represented by an individual target function within the optimisation model.
- **Investment Effects Model:**
  The Investment Effects Model is a mathematical model which can be used to simulate how an investment decision translates to the technical parameters of the investment. As an example for an investment into PV: Once the decision is made, the technical parameters of the new PV-generator need to be defined. This will depend on the existing load profile, the requirements of the investor, the available money and the technical parameters of the building. As an example for an investment into the building itself: Once the decision is made, the technical parameters of the building will change, for example due to the investment into new windows. But this strongly depends on the type of windows, the number of windows, etc. Finally, this model will also include the description of time delays in the implementation due to the construction and set up of the investment.
- **Technology Data:**
  The Technology Data represents the relevant data of potential investments into the building itself or into generation capacities. This data contains, for instance, information on the space requirements of new technologies or the insulation factors. A multitude of different technologies and parameters will be considered.
- **Multi Agent System Model:**
  The Multi-Agent-System (MAS) is the overall simulation system in which the

different *Components* are nested. A MAS simulates different agents which are described by a set of parameters and rules. These agents can either be completely sentient, meaning that they can act on their own will (represented by a certain set of rules) or be passive, meaning that they will be only active when a sentient agent interacts with them.
In the case of the WHY Toolkit, the inhabitants are represented in the MAS as sentient agents, which will act independently under certain constraints and follow their desires and will interact with the different appliances and thus generate an electrical load / consumption.

- **Upscale Algorithm:**
  Algorithm that uses time series bootstrapping techniques to create as many load profiles from the single household load profiles as are needed for the scenario that will be analysed. The amount of profiles and their distribution is taken from the scenario description.
- **Energy System Model:**
  An Energy System Model is a large-scale mathematical representation of our current and future energy system. These Models are used to analyse future trends and scenarios and calculate the effects of policy measures on, for instance, energy consumption, CO_2 emissions, technology deployment, power generation mix and so on. These models represent multiple different aspects of the energy system from consumption, generation, the electricity markets, grids etc. There are multiple different Energy System Models currently being used to analyse national, European and global policy scenarios, with TIMES and PRIMES being two of the most prominent ones, extensively used for energy policy impact assessment

2.2 Code

The term *Code* in this particular case refers to a collection of models and modules designated to perform a certain task within the WHY-Architecture. Different *Code*-Elements are interconnected via *Interfaces*. These *Code*-Elements can operate on their own and return the requested output if the required input is provided. Within the WHY-Architecture the *Code*-Elements interact with one another.

The following sections provide a summarised description of all *Code*-Elements considered in the WHY-Architecture.

2.2.1 Causal Model

This *Code*-Element contains the implementation of the Causal Model. The Causal Model will be responsible for the generation of the configuration of the infrastructure that households to be simulated has, including the behaviour of its inhabitants and their desires to participate in all the aspects of the energy transition.

- Inputs:
  - Technical composition of the households
  - Socio-economic description of the individuals at the household
- Outputs:
  - Behavioural information regarding desires and lifestyles

  - Type of households
  - Investment decisions related to energy

2.2.2 Load Profile Generator El-Version

This *Code*-Element contains all the *Components* required to calculate the electric load profiles of a household, stemming from the user behaviour and their desires. This element is the WHY-implementation of the Load Profile Generator[1], which is going to be adapted to fit into the WHY-Architecture. The *Code*-Element will facilitate a "***Multi Agent System"*** to represent and calculate the behaviour of the building's inhabitants and their desires, resulting in electricity consumption. For that purpose, this *Component* will be responsible for calculating the:

- ***Non-Controllable Appliances:*** a mathematical representation of the technical devices that will be directly controlled by an actual person and not by a control algorithm will be included in this *Component*.
- ***Controllable Appliances:*** a mathematical representation of the desires for a comfort living at the households for devices that will be directly controlled by a control algorithm will be also included in this Component.
- ***Energy Generator:*** a mathematical representation of the controlled energy generation including the possibility of in situ generation of fuels (***"Power 2 X")***.

As a result, this *Code*-Element will provide insight and results for the human-driven electrical energy consumption within a building.

- Inputs:
  - Technical composition of the households
  - Behavioural information regarding desires and lifestyles
  - Type of households
  - Investment decisions related to energy
- Outputs:
  - Electrical load profiles for non controllable loads
  - Use desires for controllable loads

2.2.3 Load Profile Generator Tr-Version

While the LPG El-Version focuses on electrical energy consumption, the LPG Tr-Version will generate data on the transport behaviour or transport actions resulting from the desires of the inhabitants of the building. The data is used to simulate the mobility behaviour of the inhabitants and, therefore, provide results on the energy consumption due to mobility actions. Within the WHY-Architecture, there is a strong focus on electricity demand, thus this *Code*-Element primarily focuses on electric vehicles.

As a result, the information on the locations of the electric vehicles, their routes and their arrivals at charging points is available. This can be achieved by considering the ***"Transport***

[1] https://www.loadprofilegenerator.de/moreinfo/

Needs" of the inhabitants of the building as well as the available modes of transport considered in the ***"Transport model"***.

- Inputs:
    - Mobility devices of the households (e.g. electric cars, electric scooters)
    - Behavioural information regarding mobility desires
    - Routing information
    - Investment decisions
- Outputs:
    - Mobility behaviour (trip details[2]) of each vehicle

2.2.4 Load Profile Generator Th-Version

The load profile generator has a focus on the thermal behaviour of the buildings considered in each simulation. This *Code*-Element contains the collection *Components* and data needed to calculate the thermal energy demand of the buildings as well as the thermal energy demand resulting from the use of warm water of the inhabitants of the building. Thermal demand is one of the main contributors to CO_2 emissions in the residential sector and its electrification one of the main drivers in the decarbonization of the latter. Furthermore, it has a high relevance due to its storage capacity and thus the possibility to provide flexibility to the system, which will later on be used. The LPG Th-Version will be part of the HiSim[3] Python package for simulation of household energy consumption.

Most crucially, the LPG Th-Version will facilitate the *Component* ***"Building Data and Model"***, as well as data and information on user behaviour, and preferences for heating/cooling and warm water.

- Inputs:
    - Technical parameters of the building (e.g. heating requirements)
    - Comfort requirements of the users
    - Behavioural information regarding heating/cooling and warm water usage desires
- Outputs:
    - Thermal load profiles for the household

2.2.5 Non-controllable Renewable Energy

In parallel to three Load Profile Generators also energy generation profiles of sources that cannot be controlled are relevant. Therefore, the purpose of this *Code*-Element is to provide energy generation data for generation capacities that depend on the availability of their corresponding primary energy source, like PV-generation or wind-generation.

[2] Including at least the distance, energy and time spent in each trip.
[3] https://pythonrepo.com/repo/FZJ-IEK3-VSA-HiSim-python-miscellaneous

The Non-controllable Renewable Energy Code section will be part of the HiSim[4] Python package for simulation of household energy management.

- Inputs:
  - Technical parameters of the generating capacities
  - Data on the availability of the primary energy source
- Outputs:
  - Electrical generation profile of the non-controllable generation capacity

2.2.6 Building Scaling

This *Code*-Element is needed to scale the building itself as well as the investments into generation capacities considering the parameters of the building. For instance, if the behavioural model identifies the decision of the inhabitants to invest into a PV-generator, it is necessary to define the installed capacity / amount of panels for further calculations. Also for the calculation of the thermal demand, the technical parameters of the building itself (for instance insulation) as a result of an investment decision are necessary.

The Building Scaling *Code*-Element uses a similar approach as the House Infrastructure Simulation, which will be described later on. Both of these *Code*-Elements facilitate HiSim[5] where the different LPG-*Code*-Elements as well as the non-controllable renewables are combined or called upon within one simulation environment.

To generate the results, this *Code*-Element facilitates the *Component* ***"Technology Data"***, which contains all relevant data relevant to the building scaling process as well as the ***"Investment Effects Model"***, which performs the actual scaling.

- Inputs:
  - Electrical load profiles
  - Thermal load profiles
  - Behavioural information regarding mobility desires
  - Investment decisions
- Outputs:
  - Technical parameters for building related and generation related investments.

2.2.7 House Infrastructure Simulation

Considering the non-controllable electrical loads, this *Code*-Element combines the capacities of a household to provide flexibility with the system that has the necessary intelligence to control flexible devices and the signals needed to make a decision on when to use the devices. Provided that the required data on user preferences, devices, etc. are available, this *Code*-Element will optimise the use of flexible loads (***"Stationary Energy Storage: Batteries"***, **"Seasonal Storages: Power 2 Hydrogen combined to Hydrogen to Power"**, ***"Movable Energy Storage: EV"***, ***"Power 2 Gas", "Electric Heating devices"*** and

[4] https://pythonrepo.com/repo/FZJ-IEK3-VSA-HiSim-python-miscellaneous
[5] https://pythonrepo.com/repo/FZJ-IEK3-VSA-HiSim-python-miscellaneous

"Controllable Appliances"). Furthermore it will build the link between transport needs and energy consumption for charging electric vehicles, thus facilitating the ***"Transport Energy Models"***.

The optimisation of the time of use of the devices will be achieved through an ***"Energy Management System"***, which will react to a (price) signal.

Therefore, both loads whose time of use can be shifted according to a price signal or the availability of renewable energy are considered as well as an energy management system capable of activating the flexible loads and shifting their time of use.

The House Infrastructure Simulation will be implemented in the HiSim[6] simulation environment and advance the functionality of said Python package.

- Inputs:
  - Load profiles
  - Generation profiles
  - Thermal profiles
  - Mobility use profiles
  - Technical parameters for the flexible loads
  - Household equipment (flexible loads)
  - Price signal
  - User behaviour

- Outputs:
  - Load profile of the household with consideration of flexible loads

2.2.8 Sustainability Assessment Module

This *Code*-Elements will be responsible for the estimation of a panel of KPIs to evaluate the impacts produced by the energy consumption[7]. The focus will be on the environmental and social impacts of the energy demand but the traditional technical and economic KPIs will also be calculated here. The data for the estimation of these KPIs will be collected during all the previous steps and the final calculations will be made in this component.

- Inputs:
  - Load profiles (by energy service)
  - Generation profiles (by primary energy source)
  - Thermal profiles (by primary energy source)
  - Mobility use profiles (by transport mean used)
  - User behaviour (including ethical and social drivers)
  - Others depend on the final set of KPIs defined by the stakeholders of each use case.

- Outputs:
  - Estimation of the impacts on all KPIs defined

[6] https://pythonrepo.com/repo/FZJ-IEK3-VSA-HiSim-python-miscellaneous
[7] Please note that the KPIs will be defined in Task T1.3 and included in Deliverable D1.3.

2.2.9 Upscale

Up to this point, the *Code*-Elements have provided information for single households or inhabitants of single households. Nevertheless for *Energy System Models* the information on single households is not overly relevant as they have a much broader scope. Thus, it is required to scale up the results to the point where a representation of a large proportion of the population or households within the scope of the Energy System Model analysis is reached. The purpose of this *Code*-Element is to perform such Upscale for which the Component ***"Upscale Algorithm"*** is facilitated.

- Inputs:
  - Scenario Data on the distribution and number of representative households
  - Probability distribution of Load Profiles of single households

- Outputs:
  - Load profiles of residential customers within the scope of the scenario analysed

2.2.10 ESM (PRIMES, PROMETHEUS, TIMES)

This *Code*-Element represents the existing and well established Energy System Models which provide information on Energy Systems at national, European and global levels within a certain given scenario.

- Inputs:
  - Energy data of households in the format requested by the corresponding Energy System Model

- Outputs:
  - Information on the Energy System within the scope of the scenario defined by the modellers, including e.g. energy consumption, fuel mix by sector, power production mix, CO_2 emissions, energy system costs, investment etc.

2.2.11 Interventions Optimizer

Finally, this *Code*-Element will take the results of the simulation and modify the parameters of the different interventions in order to look for an optimal combination of them.

- Inputs:
  - Results of a simulation.

- Outputs:
  - Parameters of the different interventions.

2.3 Interfaces

Interfaces are the coupling points between the different *Code*-Elements. Each interface is described by the data it contains and of the format the data provided. The *Interfaces* presented in this deliverable only represent a preliminary description of the actual *Interfaces* developed later on in the project. The purpose of this preliminary description is to provide an overview and get a first idea of the data required for the different *Code*-Elements and the *Components* they contain.

2.3.1 Data provided to the Causal Model

The data for the Causal Model will be provided as a **JavaScript Object Notation** (JSON) File containing the following information:

- **Location Specific Information:**
  Geographical and statistical data on the location of the household to be analysed.
- **Climate data:**
  Climate information of the localization.
- **Household / People Specification for each household separately:**
  Information on the inhabitants of the building, their agendas, demographic information, etc.
- **House / Building Specification:**
  Technical Data of the building and the devices installed there (including insulation, generation, storage, etc.).
- **Interventions and other legal instruments:**
  Parametrization of the legal or of any other intervention type to be implemented in the scenario.
- **Counterfactual data:**
  Actual measurements on the scenario used for counterfactual simulation.

2.3.2 Causal Model Results

This interface will be implemented as a JSON file (or similar), which will provide all the data generated by the causal model:

- Investment Decision of Inhabitants
- Reaction of Inhabitants to external interventions
- "Desire Profiles" of the Inhabitants concerning use of devices, travel and heating / warm water

Furthermore, all the general Data provided to the Causal Model in the JSON file will also be provided at this *Interface*. This interface will be integrated in the house infrastructure simulator HiSIM, which will pass the information to the noncontrollable renewable generation module and the electric and transport version of the Loadprofile Generator.

2.3.3 Load Profile Generator (House) Profiles

This interface, which will be part of the house infrastructure simulation HiSIM, connects the outputs of the non controllable renewable energy generation module, the Loadprofile Generator (transport + electric) and the results from the causal model to the house infrastructure simulator, which will incorporate the thermal Loadprofile Generator.

The results of the electric and transport version of the Load Profile Generator will be passed either via a JSON file or as csv files.

2.3.4 Building Setup

The Data describing the building setup of generation, the building parameters required to calculate the thermal demand and the non-controllable generation will also be available in a JSON file. Furthermore, the JSON File will contain the general data provided to the Causal Model as it will be used in the House Infrastructure Simulation.

2.3.5 Load Profile Household

After the optimization of the flexible loads, the resulting data of the building will be generated and provided as a .CSV file. The file will contain the load profile of all devices within the building (controllable and non-controllable) and also the generation profile of the considered generation capacities.

2.3.6 Household Load Profiles in the Scenario

After the upscaling process, the resulting data will be saved into a .CSV file for use in the ESMs. The data format will strongly depend on the requirements of the different ESMs. Currently, the following data is considered to be provided:

- Power values: Load profile (different temporal resolutions)
- Power values: Generation profile (different temporal resolutions)
- Power values: Peak Generation (daily, weekly, monthly, annually)
- Power values: Peak Load (daily, weekly, monthly, annually)
- Energy values: Grid demand (different temporal resolutions)
- Energy values: Surplus (different temporal resolutions)
- Energy values: Self-Consumption (different temporal resolutions)
- Energy values: Use of flexible loads

Furthermore data and information on the scenario, such as distribution of households considered, etc. will be provided in the .CSV File.

2.3.7 ESM results

The results of the ESMs will be provided in the .CSV (or .xlsx) file format. The content will strongly depend on what needs to be analysed in relation to the input parameters. At this point, it is very likely that energy related results (e.g. energy consumption by sector, fuel mix, power mix), emission, technology uptake, and economic results will be relevant.

2.3.8 Results

In case of applying an iterative approach where the results of a simulation will be used for another iteration of the same scenario, the results of each iteration will be added to the JSON file for the *Causal Model.*

2.3.9 New interventions

As in the previous case, if an iterative approach is followed to optimise the interventions carried out, the results of an iteration will be used to derive a new set of interventions (or a reparametrization of the used ones). The new set of interventions will be parametrized and a new JSON file with this data will be created for being used in the *Causal Model.*

3 Models of the WHY Model Architecture

To provide a more in-depth insight into the workings of the different *Code*-Blocks within the WHY Model Architecture a description of the main methods applied to each of the *Code*-Blocks is provided in this chapter. The aim of this chapter is to provide the reader with general knowledge on the functionalities of the different *Code*-Blocks, there will be no source codes or similar provided in this chapter, but where possible reference to repositories etc. will be provided.

3.1 Causal Model

The causal model module aims to generate structured causal models (SCM) representing the causal existing relationships in a variable set. The SCM can measure how one variable affects another one, and they can intervene in a causal relationship to predict the effects of an intervention.

A SCM is also a predictive model. Predictive models uncover patterns that connect the inputs and outcomes in observed data. To intervene, however, we need to estimate the effect of changing an input from its current value, for which no data exists. Such questions, involving estimating a counterfactual, are common in decision-making scenarios.

- Will it work? Does a proposed change to a system improve people's outcomes?
- Why did it work? What led to a change in a system's outcome?
- What should we do? What changes to a system are likely to improve outcomes for people?
- What are the overall effects? How does the system interact with human behaviour? What is the effect of a system's recommendations on people's activity?

Answering these questions requires causal reasoning.[8]

Judea Pearl[9] drafted a "causal inference engine" that might handle causal reasoning for future artificial intelligence. It's important to realise that the draft in Fig. 3, is not only a blueprint for the future but also a guide to how causal models work in scientific applications today and how they interact with data.

[8] Amit Sharma, Emre Kiciman. DoWhy: An End-to-End Library for Causal Inference, https://github.com/microsoft/dowhy
[9] Pearl, J., & Mackenzie, D. (2018). *The book of why : The new science of cause and effect* (1st ed.). New York: Basic Books.

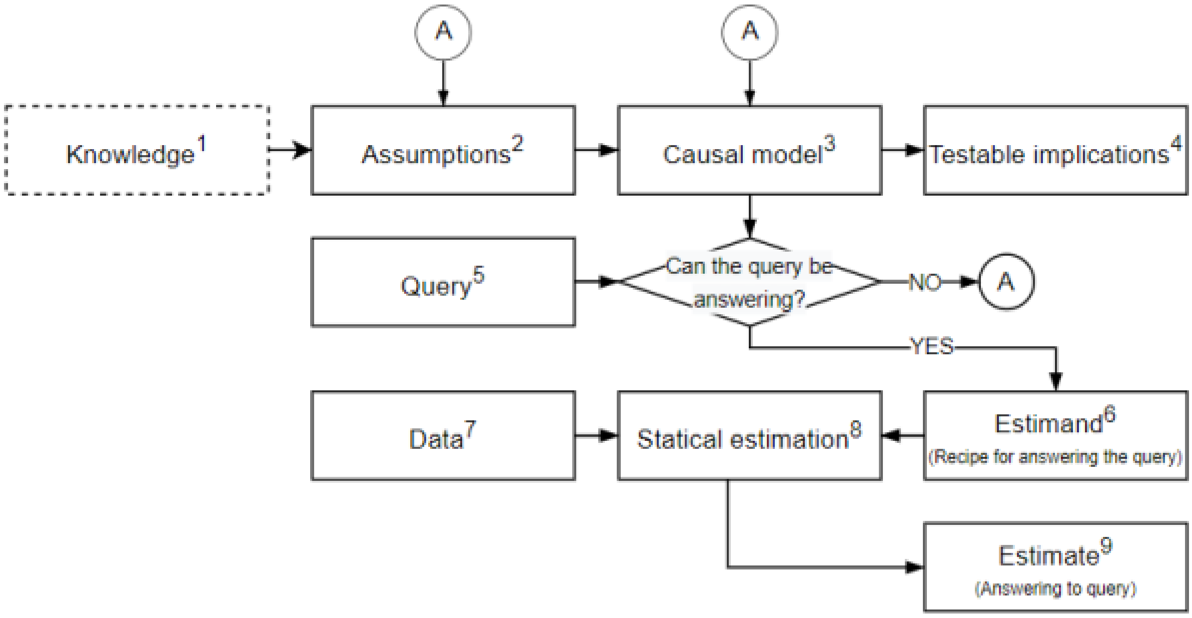

Figure 3: Causal inference engine[10]

Fig. 4 shows the architecture of the structured causal modelling. The core of the module is the causal inference engine that will generate structured causal models. The application programming interface (API) enables the integration of the module with external Applications, such as Load profile generator (LPG). The diagram also shows the SCM engine, an optional component, that enables the module to perform SCM on consuming live streaming data from external applications.

[10] Pearl, J., & Mackenzie, D. (2018). *The book of why : The new science of cause and effect* (1st ed.). New York: Basic Books.

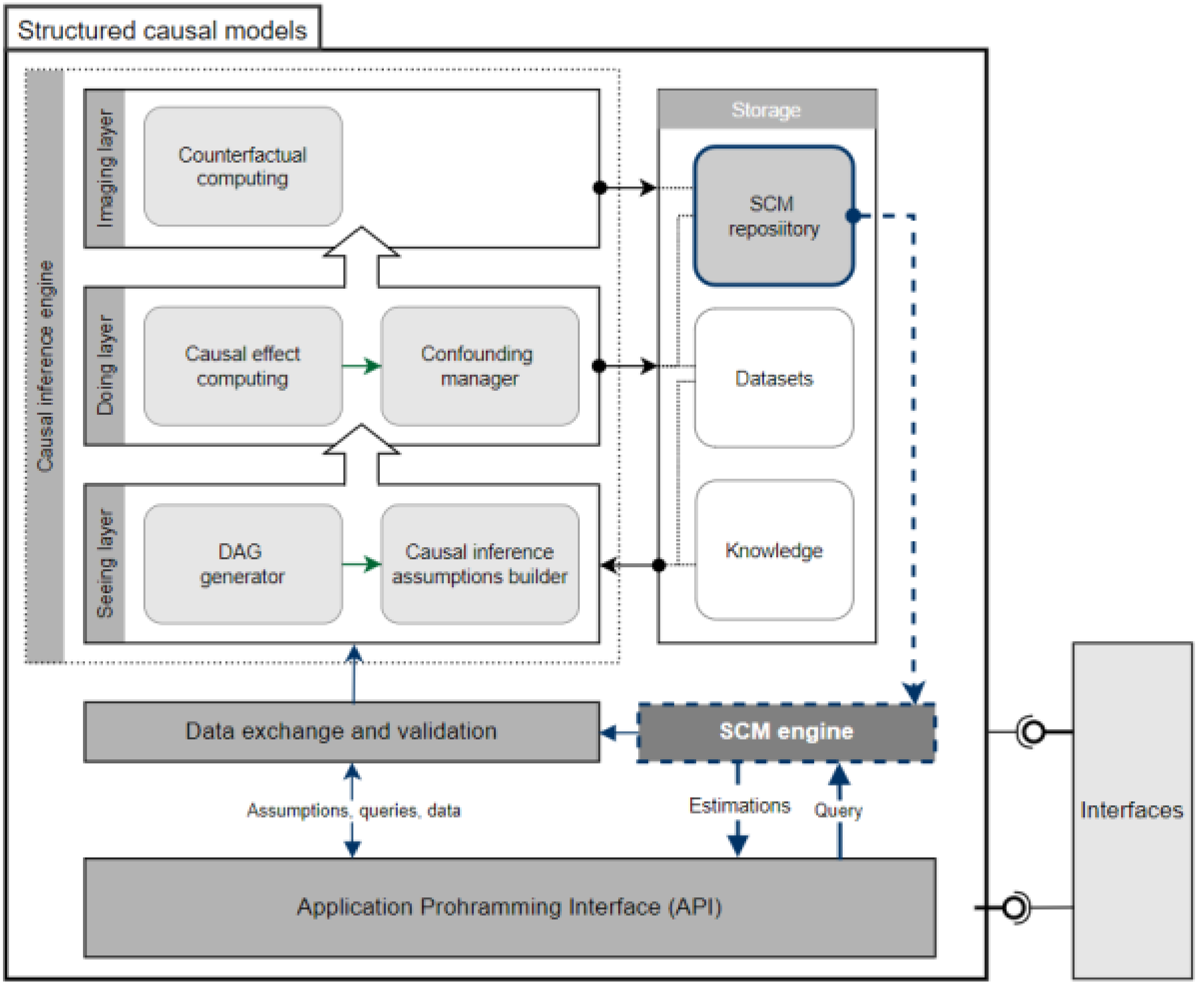

Figure 4: Causal model architecture

I. Application Programming Interface (API): It enables the Causal model to communicate with other products and services, but it also allows sharing data with either other external Applications or users.

II. Data exchange validator: Data Exchange Validator: it will validate that the input data format is well-formed JSON files, and it will perform data transformation routines to guarantee values match to the required data type.

III. Causal inference engine. This component consists of three layers that turn assumptions into SCM.

a. ***Seeing layer:*** It calls for predictions based on passive observations and it is characterised by the question "What if I see …?". The seeing layer processes the input data, which means previous knowledge of the scenario, and involves the detection of regularities in the input data as well as the identification of the causal assumptions that they imply. The generated assumption will be the input of the next layer. The seeing layer has two components described below.

- ***DAG generator:*** It uses graph-based criteria and do-calculus for modelling assumptions and identifies a non-parametric causal effect. Each analysis starts with building a causal model. The Causal model creates an underlying causal graphical model that serves to make each causal assumption explicit. This graph does not need to be complete—one can provide a partial graph, representing prior knowledge about some of the variables. The assumptions can be viewed graphically via directly acyclic graphs (DAGs) or in terms of conditional independence statements.

- ***Causal inference assumption builder***. It will model the current causal inference scenario using assumptions. Some associations might have obvious causal interpretations; others may not. Wherever possible, this competent also automatically tests for stated assumptions using observed data.

  While not recommended, one can also specify common causes and/or instruments directly instead of providing a graph. The supported formats for specifying causal assumptions are the following.

  - **Graph**: Provide a causal graph in either gml or dot format. Can be a text file or a string.
  - **Named variable sets**: Instead of the graph, provide variable names that correspond to relevant categories, such as common causes, instrumental variables, effect modifiers, frontdoor variables, etc.

b. ***Doing layer***: It can try to change the current fact based on data. The questions that define this layer are "What if we do…?". Intervention involves not just seeing but changing what is. A very direct way to predict the result of an intervention is to experiment with it under carefully controlled conditions.

- ***Causal effect computing:*** Based on the causal graph, the Causal model will implement a set of causal inference libraries, to find all possible ways of identifying a desired causal effect based on the graphical model. It uses graph-based criteria and do-calculus to find potential ways to find expressions that can identify the causal effect. It provides a non-parametric confidence interval and a permutation test for testing the statistical significance of obtained estimates. Causal computing can measure the individual (ITE), the Average Treatment Effects in a population (ATE) and the Conditional Average Treatment Effects (CATE).

- ***Confounding manager***: works for scenarios that account for the observed data. The data can be determined by many factors which can affect the effect of interest, then the effects of those factors become entangled with the effect of treatment. The confounding manager implements identification criteria like Back-door, Front-door and Instrumental Variables to manage confounding[11].

c. ***Imaging layer***: It will answer the question "What if I have done…?" and "Why?" Both involve comparing the outcome based on observed data (factual) to a counterfactual, which stands for data that either has not or cannot be observed. The aim of this layer is also to make better predictions.

[11] Hernán, M. A., & Robins, J. M. (2020). *Causal Inference: What If.*

- ***Counterfactual computing***: It will estimate counterfactual outcomes that are linked to the factual one. Estimating a counterfactual is a common step within decision-making scenarios. This component also will estimate the counterfactual outcome under different intervention values.

IV. Storage: It will enable the data persistence. It will store the prio and generated knowledge, input datasets, data for trial experiments, and the generated structured causal models.

V. SCM Engine: This optional component enables the Causal Model to consume live data streamed from external applications. By running an online causal model it can evaluate, in real time, the effect of interventions on the outcome of interest (factual or counterfactual) in order to reply to incoming queries.

VI. Interfaces: they will provide the causal model with JavaScript Object Notation (JSON) data files. The data contains information such as location, weather, and household/person specifications (see section 2.3.1).

3.2 Load Profile Generator

The Load Profile Generator (LPG)[12] is a behaviour-based multi-agent model that models every person as an individual, desire-driven agent. That means that at every step, every agent is looking at the available options and decides, based on their internal state, what to do next. The model is based on studies by the German psychologist D. Dörner, who first developed a similar behaviour model to study behaviour and learning[13].

The basic principle of the model is shown in Fig. 5. The devices in the household offer different activities to the resident and the residents then choose what they would like to do. Once they trigger an activity, then the corresponding devices are activated and generate an electricity or water demand. The profiles are added up to generate the summed-up profile for the household. The model is described in detail in (Pflugradt, 2016)[14].

[12] https://loadprofilegenerator.de/
[13] Dörner D. Bauplan für eine Seele. 2nd ed. Reinbek bei Hamburg: Rowohlt Taschenbuch Verlag; 2001
[14] Pflugradt, N. D. (2016). Modellierung von wasser und energieverbräuchen in haushalten.

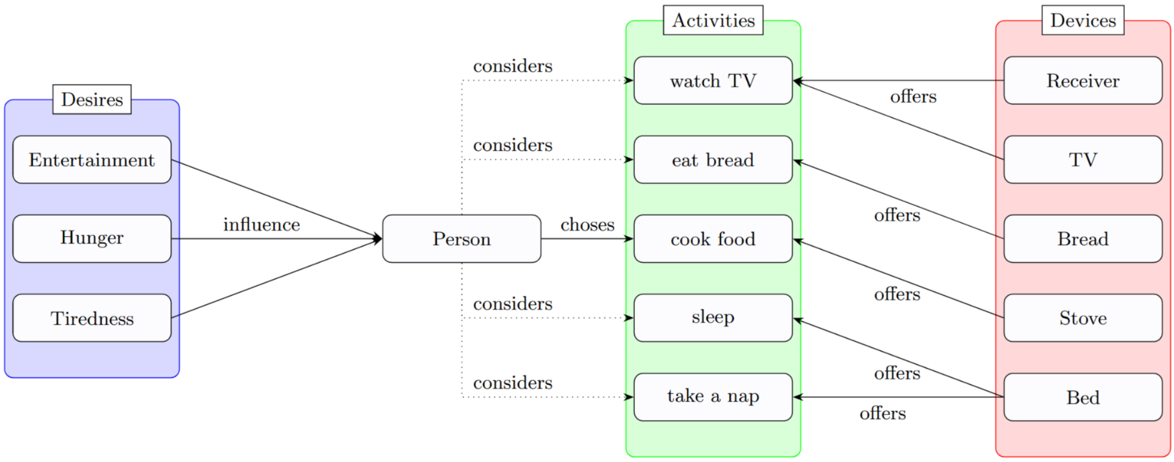

Figure 5: Basic principle of the decision model[15]

This approach has a number of advantages over alternative, probability-based approaches. In particular, it requires far less data to generate usable profiles. This is especially relevant in privacy-heavy areas such as residential behaviour (or studies about future behaviour) where there is a major shortage of available data.

The model has been extended beyond the simple decision model with a large number of additional features:

- **Vacations**: People can leave their homes behind and go travelling. Depending on the settings they turn off their standby devices or leave them running.
- **Holidays**: People will not go to work on local holidays.
- **Illness**: People can get sick during the year for one or more days and then change their behaviour in configurable ways, for example taking naps, not going to work and watching lots of TV.
- **Triggered behaviour**: To make sure that the frequency of dishwashers and washing machines are reasonable, the LPG contains functionality to track for dirty laundry, dirty dishes and other things. Devices can be configured to be only available if a certain amount of dirty dishes or dirty laundry is available for washing. The same functionality can also be used to make sure that the dryer is only turned on if there is wet laundry available.
- **Accurate lighting model**: People turn on the lights in rooms that are too dark based on global radiation, but only if they execute activities that need light.
- **Complex activity model with independent device profiles and person profiles**: While cooking, people will first prepare food, maybe use a mixer, then turn on the stove for a bit and then maybe eat. Each device can have their own, independent profiles.
- **Multiple load types**: The LPG models effective power, reactive and apparent power, warm and cold water consumption and other things
- **Occupancy tracking**: The LPG provides data on how many people are home and how active they are to provide accurate data on the heat generated by the residents.

[15] Pflugradt, N. D. (2016). Modellierung von wasser und energieverbräuchen in haushalten.

- **Autonomous devices**: To model devices that run independently of human interaction such as fridges, the LPG contains a sophisticated device model that can handle things such as a TV that is only in standby until it is turned on, a fridge that runs for a few minutes every hour and anything in between.
- **Joint actions**: to accurately model behaviour such as the mother calling the children to dinner that she cooked, the LPG offers sub-actions that get enabled by someone executing the main activity (e.g. “make dinner and eat”) and that can interrupt other activities (e.g. “watching TV”).
- **Templating**: One of the largest challenges when modelling a large number of households is to make sure that they vary in a realistic way so that in a neighbourhood a realistic smoothness of the aggregated profile develops. For this the LPG contains a sophisticated templating functionality that enables the user to describe a household in rather abstract terms (basically “office worker with 3 hobbies” and the templating function then selected the hobbies from a list of predefined activities, if the hobbies require for example a computer, the LPG will select on specific model from the list of predefined computers and so on. By utilising this functionality it is possible to automatically generate hundreds or thousands of nearly identical but slightly different households to simulate for example districts.

In conclusion, the LPG is a sophisticated open-source model that has been in development for over 10 years and that covers nearly all aspects of daily life. The challenges that will be solved in the WHY project are primarily focused on the data side: Previously the LPG contained only a limited number of household templates calibrated for Germany. In the project it is planned to create additional templates, thoroughly validate them with Time-Of-Use survey data, add additional devices and measurements and provide calibrated templates for other European countries.

3.3 HiSim

HiSim[16] is an open-source Python-package for simulation in building energy systems. It is a component-based time-step simulation with a substitution solver. The goal of this package is to enable the research of different building energy system strategies considering the many load profiles, physical conditions and alternative components from fossil fuel sources. The user can also design their own components to be implemented in its own building energy system. Extendability and modularity are the two concepts that HiSim is designed around. In the context of the WHY-toolkit architecture (sketched in Fig. 2) HiSim represents the thermal version of the load profile generator, the non-controllable renewable energy module as well as the house infrastructure simulator, containing interfaces to the electrical - and transport version of the load profile generator.

It is written in Python and has been in development for about 2 years now. The goal of the simulation package is to provide a free, fast and extendable building simulation on a high level.

HiSim can model arbitrary building energy systems by combining and connecting individual, reusable components. Examples for the components are:

[16] https://github.com/FZJ-IEK3-VSA/HiSim

- Oil heater
- Gas heater
- Warm water storage
- Heat pump
- Different controllers
- Photovoltaic
- Residential energy demand
- Batteries
- 5R1C building model[17]

Every component can have multiple inputs or outputs that can be connected as needed. For example, a controller to optimise self consumption of a PV system needs to be connected to the residential electricity demand, the PV system, and one or more controllable loads, such as a heat pump or a battery and can then turn on the load whenever there is excess energy available. HiSim has a default time resolution of 1 minute with a fixed time step width, though the step size is customizable. Fig. 6 shows an example system. It is noteworthy that HiSim does not aim to model every pipe, every valve and every pump at a physical level, but instead aims to use a slightly higher abstraction level to enable a quick sizing of components and analysis of households.

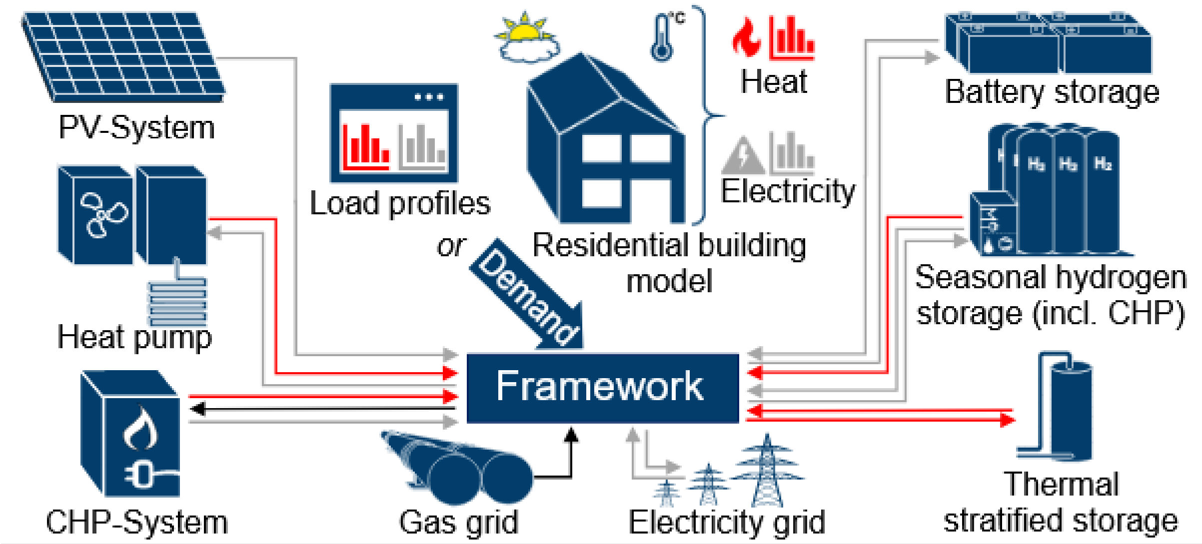

Figure 6: Example for a building energy system which can be modelled with HiSim[18]

The substitution solver is able to deal with circular dependencies and resolve issues within the same time step. A circular dependency is for example shown in Fig. 7. Here the controller acts based on the building temperature and controls a heat pump. The heat pump charges a thermal energy storage, which is used to heat the building. Most reasonably complex building energy systems have such circular dependencies built in.

[17] https://www.iso.org/standard/41974.html
[18] https://household-infrastructure-simulator.readthedocs.io/en/latest/index.html

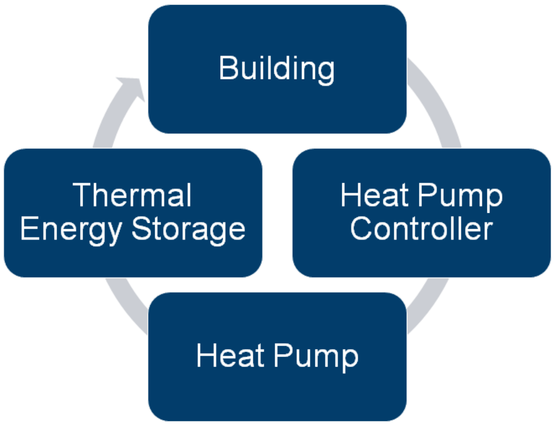

Figure 7: Example of a typical circular dependency when modelling a building.

The WHY-Project will add new components to the HiSim package, such as oil heaters, gas heaters and advanced controllers. Additionally it is planned to add a better interface to make it easier and faster to perform thousands or tens of thousands of simulations.

3.4 LPG Transport

The LoadProfileGenerator (LPG) referenced above contains a full mobility model inside the behaviour simulation. If people decide to go to work, they can use different mobility options, for example, bicycle, bus or car based on their preferences. The LPG keeps track of the behaviour simulation of where people are and where they need to go for different activities. The LPG contains a decision model that takes into consideration the activity to be carried out, the vehicles availables and the preferences to select the transport to be used. The decision model will be further improved during the WHY-project by considering time, cost and other qualitative metrics (like number of hops, time looking for parking, stress due to high traffic, etc.). In order to avoid the need to include geographical components to the model (that are not considered at WHY) the LPG contains a simple distance-model that describes the typical distances for different types of activities. The model is described in more detail for example in[19]. The rest of the variables are obtained directly (time, costs, for example) or using indirect methods from the rest of variables.

The output of the model is the energy demand of the chosen mode of transportation and the time that each vehicle remains at certain locations and is available for charging/discharge. In other words, the LPG will yield result files that for example list that a vehicle was home from 19:53 to 6:34, the state of charge of the battery of the vehicle and the energy that it requires for the next day.. The model can be fully configured in respect to modal preferences, travel distances, available transport choices and more.

3.5 Energy System Models

The section provides a brief description of the large-scale Energy System Models used and further enhanced in the WHY project. The PRIMES model focuses on the EU and national level, while PROMETHEUS analyses the global energy system and emission developments.

[19] https://www.cired-repository.org/handle/20.500.12455/595

3.5.1 The PRIMES model

The PRIMES (Price-Induced Market Equilibrium System) is a large-scale applied energy system model that provides detailed projections of energy demand, supply, prices and investment to the future, covering the entire energy system including emissions (Fig. 8). The distinctive feature of PRIMES is the combination of behavioural modelling (following a micro-economic foundation) with engineering aspects, covering all energy sectors and markets. PRIMES has a detailed representation of instruments policy impact assessment related to energy markets and climate, including market drivers, standards, and targets by sector or overall. It handles multiple policy objectives, such as GHG emissions reductions, energy efficiency, renewable energy targets and provides pan-European simulation of internal markets for electricity and gas. PRIMES offers the possibility of handling market distortions, barriers to rational decisions, behaviours and market coordination issues and it has full accounting of costs (CAPEX and OPEX) and investment on infrastructure needs. The model covers the horizon up to 2070 in 5-year interval periods and includes all Member States of the EU28 individually. PRIMES is well placed to simulate long term transformations (rather than short term) and includes non-linear formulation of potentials by type (renewable or fossil resources, sites, acceptability etc.) and technology learning. PRIMES has a modular structure, with modules focusing on different aspects of the energy system, both on the demand and the supply side, as well as energy networks.

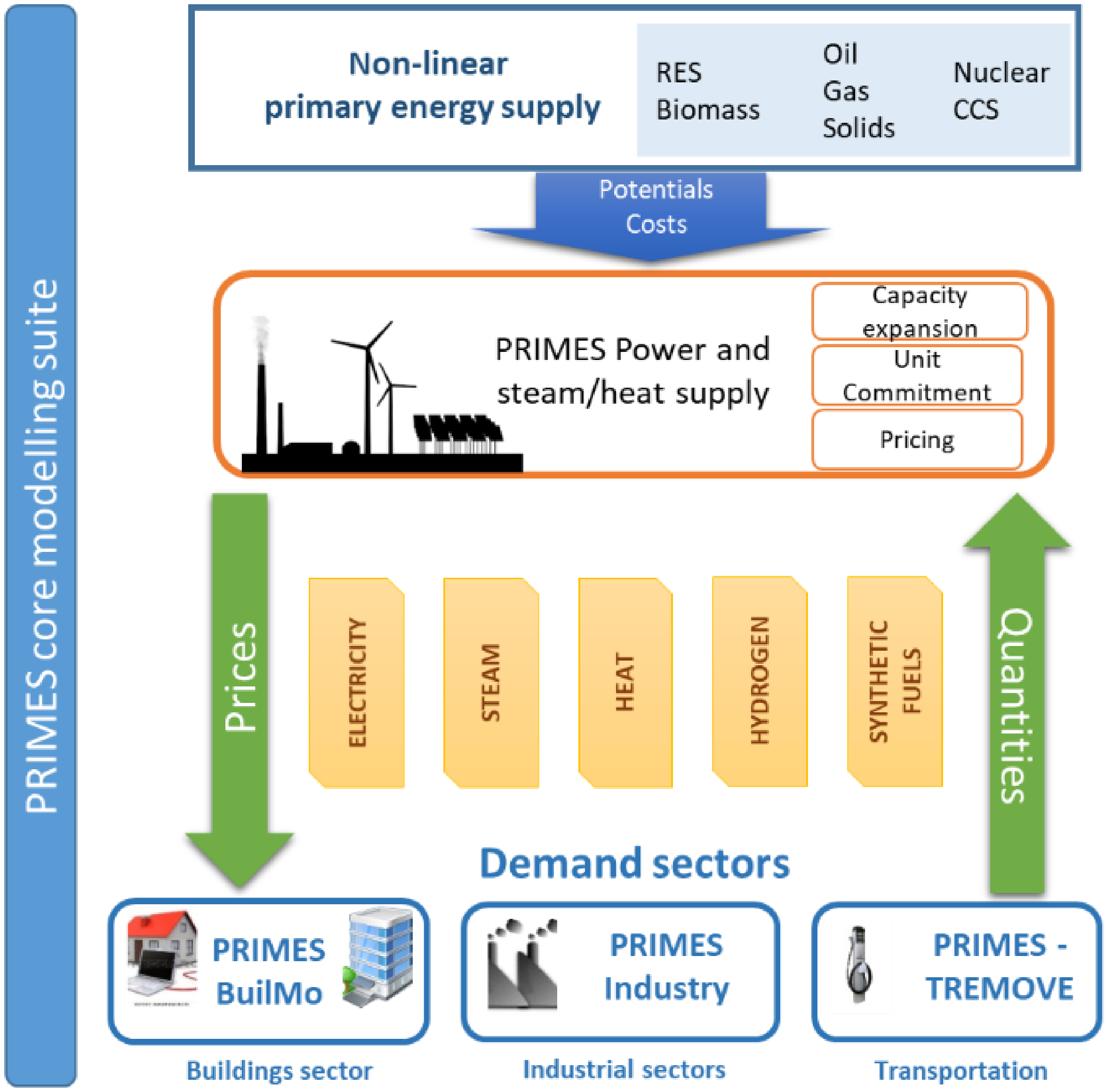

Figure 8: Schematic representation of the PRIMES energy system model

In the context of the WHY project, the PRIMES BuiMo is the most relevant module. PRIMES BuiMo integrates a high resolution representation of the European housing and office building stock embedded in an economic-engineering model of multi-agent choice of building renovation, heating system and equipment/appliances by energy use. The model projects the final energy consumption, fuel mix, CO_2 emissions, renovation rates and depth, equipment choice and replacement rates in the residential and services sectors, under alternative policy and regulatory measures (European Commission 2021[20]; Fotiou, T. et al., 2019[21]). It covers market and non-market barriers; hidden costs and perceptions

[20] European Commision, Impact Assessment Report Accompanying the Proposal for a Directive of the European Parliament and of the Council on energy efficiency (recast), https://eur-lex.europa.eu/resource.html?uri=cellar:c20a8b93-e574-11eb-a1a5-01aa75ed71a1.0001.02/DOC_2&format=PDF

[21] Fotiou, T.; de Vita, A.; Capros, P. Economic-Engineering Modelling of the Buildings Sector to Study the Transition towards Deep Decarbonisation in the EU. Energies 2019, 12, 2745. https://doi.org/10.3390/en12142745

affecting consumer behaviour and models a variety of policy instruments influencing decisions and possibly removing barriers. The model accounts for behavioural aspects of energy consumers while also respecting engineering constraints and specificities and tapping possibilities for building transformation. PRIMES-BuiMo splits the stock of buildings in many categories, by:

- geographic locations,
- age of construction,
- income classes and
- service sector sub-sectors.

Income classes help simulate the heterogeneity of energy consumers and their idiosyncratic behaviour, which depends on a multitude of factors, including income, preferences, weather patterns, access to loans, location, and household composition. Instead of a single representative actor, the model includes a variety of actors with distinct behavioural patterns. Discount rates differ by income class with low-income classes typically facing higher discount rates, representing their difficulty to access low-cost loans. Through the differentiation of the discount rates based on real-world estimates, the model can reproduce consumer decisions (e.g. to purchase a heat pump or improve thermal insulation) capturing the heterogeneity of consumers in each class, thus addressing the drawbacks of the representative consumer assumption.

To further enhance the representation of energy consumption, modelling improvements are needed, especially on how consumer behaviour is represented, which factors influence the decisions of energy consumers, how new markets and business models are integrated (e.g. prosumaging, distributed generation, smart appliances) and how various instruments influence consumer decisions for energy-related equipment (including electric appliances and heating and cooling equipment). The soft-link of PRIMES-BuiMo with the WHY Toolkit will enable an improved assessment of households' energy demand and a better representation of the above factors while considering the complex, systemic interactions of energy demand, supply, fuel prices and investment dynamics as captured by PRIMES.

The modelling of renovation is based on the concept of dynamic discrete choice, where heterogeneous agents choose the most cost-efficient ones. A dynamic strategy may involve renovation of building envelope, technical building system equipment selection, including self-production equipment for electricity, premature replacement of equipment, and fuel switching. Based on the chosen renovation strategy by building class, the thermal demand of households will change, influencing also consumer decisions to purchase energy products. This, in turn, is translated to the final energy consumed (through dynamic evolution and keeping track of equipment vintages) by the space heating system. The choice of the space heating strategy depends on the timing and depth of the envelope renovation. Likewise, the dynamic strategy of hot tap water and cooking equipment depends on the space heating system. Keeping track of capital turnover as technology vintages, PRIMES-BuiMo determines the fuel mix for the technical building equipment. As a result, it derives energy consumption by fuel, CO_2 emissions, operating costs, and investment expenditures. In addition, PRIMES-BuiMo includes a sub-module projecting electricity use in households, which first determines the energy service to be provided and then chooses the type of technology to purchase to meet the desired level of energy use. The turnover of the stock of appliances is dynamic and endogenous to the model. Policy instruments including eco-design regulations influence the technology types that are offered to consumers.

Energy labelling and other policies are represented in the model and facilitate the uptake of highly efficient, yet more expensive, technology types through reducing the uncertainty and lack of information. PRIMES-BuiMo can represent various policy instruments including: Taxes for energy products, financial facilitation for renovation and purchase of low-carbon technologies, Information campaigns, eco-labelling of technical equipment, eco-design standards, Building code standards and levels of compliance, Energy efficiency standards, carbon pricing, white certificates and targets for RES (Renewable Energy Sources)- Heating and cooling, see Fig. 9

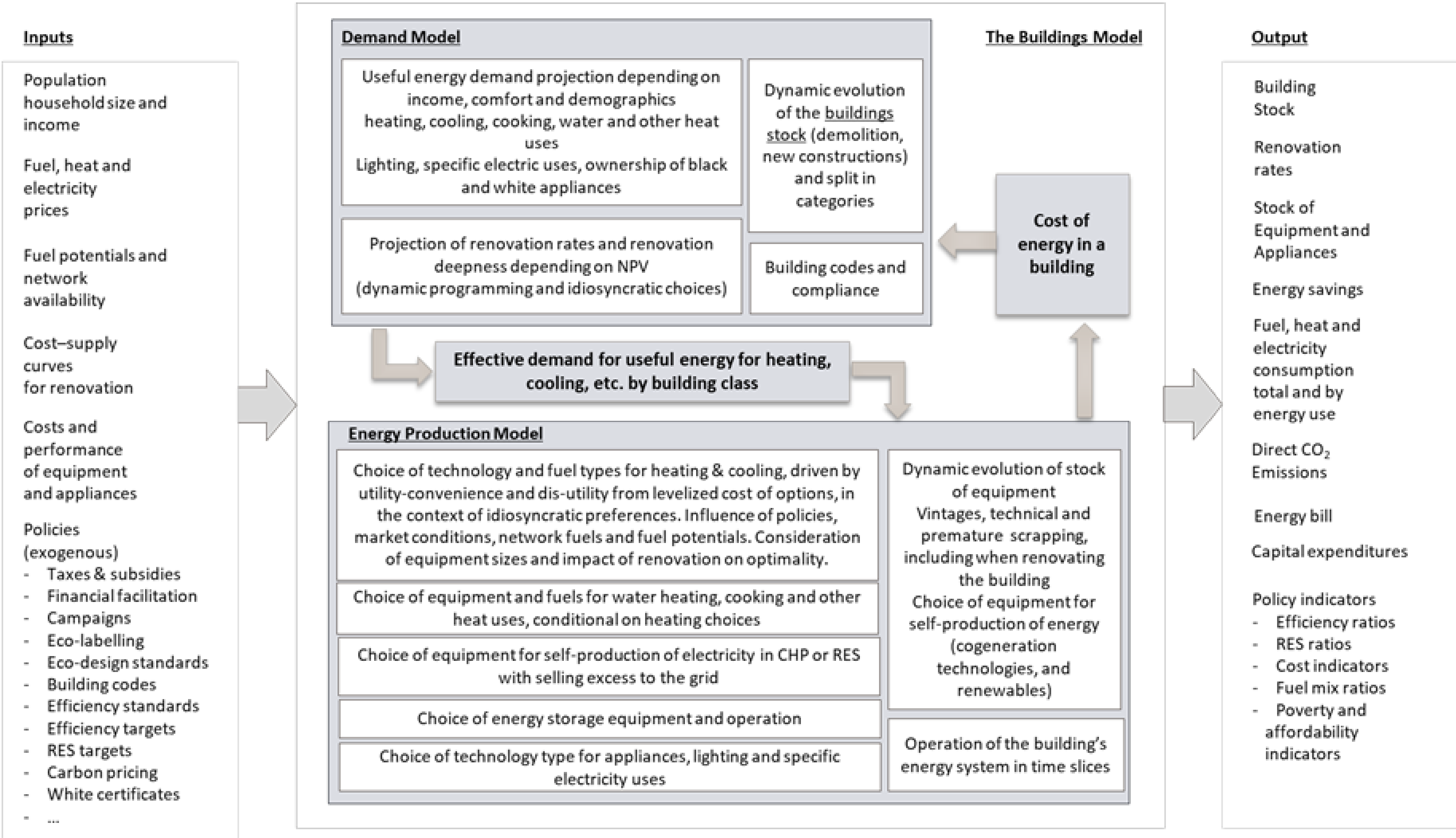

Figure 9: Flowchart of the PRIMES-BuiMo model

3.5.2 The PROMETHEUS model

PROMETHEUS is a fully fledged global energy demand and supply simulation model aiming at addressing energy system analysis, energy price projections, power generation planning and climate change mitigation policies. The PROMETHEUS model provides detailed projections of energy demand, supply, power generation mix, energy-related carbon emissions, energy prices, and investment to the future covering the global energy system. PROMETHEUS is designed to provide medium and long term energy system projections and system restructuring up to 2050, both in the demand and the supply sides. The model produces analytical quantitative results in the form of detailed energy balances in the period 2015 to 2050 annually (Fig. 10). The model can support impact assessment of specific energy and environment policies and measures, applied at regional and global level, including price signals, such as energy or carbon taxation, subsidies to technologies, renewable energy and energy efficiency promoting policies, efficiency standards, environmental policies and technology standards.

PROMETHEUS incorporates a recursive dynamic (partial equilibrium energy system) model with annual resolution currently served to run up to the year 2050 (the process to extend model horizon to 2100 is ongoing). The model has a triangular structure in order to avoid contemporaneous simultaneity. On the other hand, simultaneity is modelled through lagged instances of endogenous variables. Most of the model equations are specified in different terms in order to avoid excessive early variability and adequately represent accumulation of uncertainty in the longer term. The model simulates both demand and supply of energy, interacting with each other to form market equilibrium at different regional scales: detailed regional balances are aggregated in order to simulate world energy markets. The current model version identifies 10 world regions: EU-27 & UK (split into western and eastern EU countries), North America (USA and Canada), China, India, Western Pacific (Japan, Korea, Australia, New Zealand), Russia and other CIS, Middle East and North Africa, Emerging Economies and Rest of world. Apart from international fuel prices, regional energy systems influence each other particularly through trade, technical progress and network effects including changing patterns of consumption and spillover effects with regard to technology diffusion.

PROMETHEUS contains relations and/or exogenous variables for all the main quantities, which are of interest in the context of general energy systems analysis. These include demographic and economic activity indicators, primary and final energy consumption by fuel/sector, fuel resources and prices, CO_2 emissions, and technology dynamics (for power generation, road transport, hydrogen production and residential end-use technologies). PROMETHEUS quantifies CO_2 emissions and includes environmentally oriented emission abatement technologies (like renewable energy, electric vehicles, Carbon Capture and Storage (CCS), energy efficiency) and policy instruments. The latter include both market based instruments such as cap and trade systems with differential application per region and sector specific policies and measures focusing on specific carbon emitting activities. Key characteristics of the model, that are particularly pertinent for analysing the impacts of climate mitigation scenarios, include world supply/demand resolution for determining the prices of internationally traded fuels and technology dynamics mechanisms for simulating spill-over effects for technological improvements (increased uptake of a new technology in one region leads to improvements through learning by experience which eventually benefits other parts of the World). PROMETHEUS is a technology-rich energy system model that represents most fossil fuel and low-carbon technologies that are envisaged to be available for at least the first half of the 21st century, and also including disruptive technological options, like negative-emission technologies (including Biomass with CCS and Direct Air Capture). By simulating the substitution of low or zero-carbon for high-carbon technologies in response to changes in relative costs, as well as emissions constraints and/or carbon prices, the PROMETHEUS energy system model simulates mitigation through a large set of different measures and technological configurations differentiated by region and sector.

The economic decisions regarding the investment and operation of the energy system are based on the current state of knowledge of parameters (costs and performance of technologies, prices, ...) or with a myopic anticipation of future costs and constraints. The model does not use foresight but myopic anticipation. Some foresight can be forced in the electricity production sector. The core operating principle of the PROMETHEUS model is that of market equilibrium. The representative agents in the modules use information on prices and make decisions about the allocation of resources. Markets are the means by which these representative agents interact with one another. The model solves for a set of market prices so that supplies and demands are balanced in all energy markets across the model (partial equilibrium). The solution process is the process of iterating on market prices until this equilibrium is reached. The regional fuel markets are also integrated to form an international market equilibrium for crude oil, natural gas and coal.

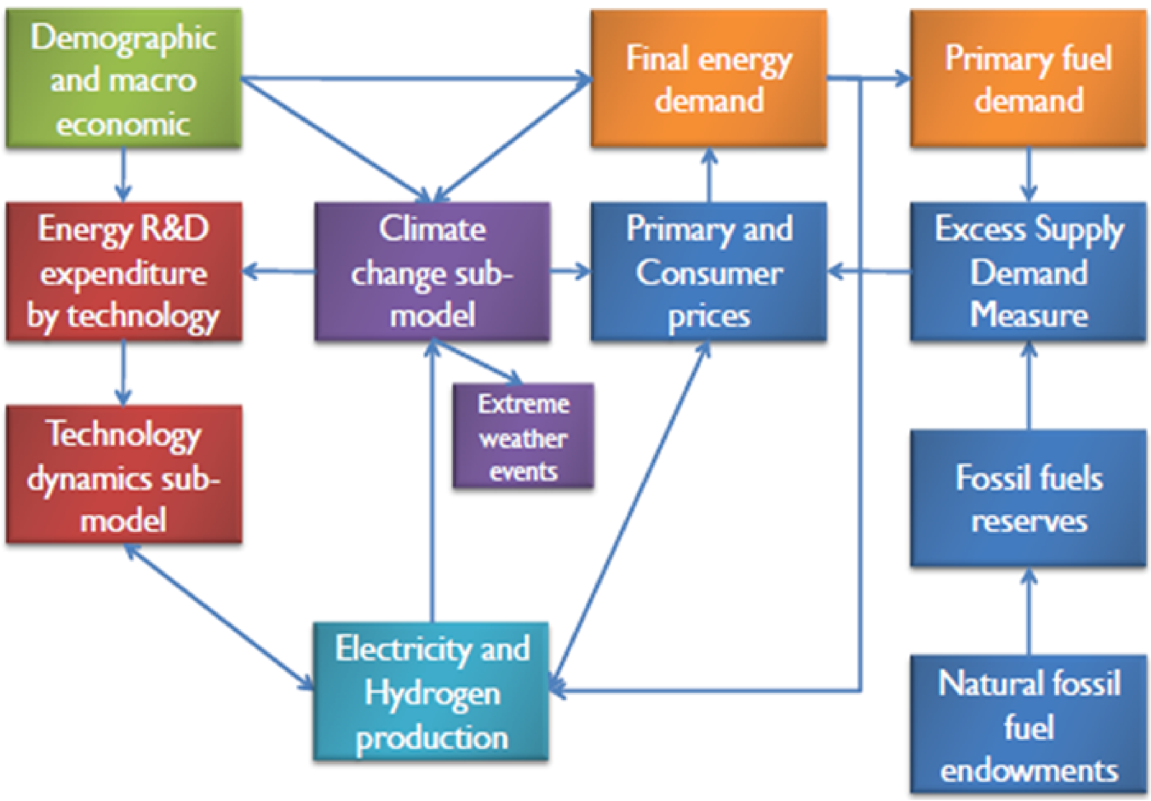

Figure 10: Schematic representation of the PROMETHEUS model

In the residential sector of PROMETHEUS, energy is consumed as input in processes that provide services to the households, such as space heating, water heating, cooking, cooling, specific electricity uses, lighting and other needs. The model distinguishes between households' demand for specific electric uses (e.g. electric appliances for non-heating purposes, air-conditioning, lighting, electronic equipment etc.) and useful energy demand for space and water heating. Demand for non-substitutable electricity is driven by growth in disposable income of households and residential electricity price, while useful energy demand for heating purposes is related to income growth and the evolution of fuel prices. Consumers decide about the level of energy consumption taking into account their need for heating, which is related to changes in income and fuel prices. Different iso-elastic demand equations are estimated for each type of residential sector's demand and for each region. As the pattern of energy consumption is not usually controlled directly by the consumer, but is determined by the installed technology and is largely embodied in the characteristics of the durable equipment, responses to price shifts and policies usually involve lags. Changes in consumption patterns for developing regions are also modelled through a gradual convergence procedure to developed countries' consumption patterns.

The competition between technologies to cover energy demand for heating is modelled using the substitution specification as described in[22], based on the Weibull function depending on the relative costs and commercial maturities of competing technological options. The model differentiates between "cold" and "warm" regions based on their climatic conditions (like India, Emerging economies, the Middle East and North Africa), as in the latter energy demand for space heating is relatively insignificant, i.e. energy demand for water heating dominates. The evolution of useful energy demand is also assumed to depend on regional climatic characteristics.

Energy demand for heating purposes is covered by gas, oil, coal, electric resistances, fuel cells (using hydrogen or natural gas as a fuel) and heat-pumps. Substitution between fuels and technologies is triggered by their total production cost, which includes capital, fixed O&M, variable O&M and fuel cost, their transformation efficiency, the scrapping rates of their equipment and their relative "technology maturity" factors. Technological trends, infrastructure and social network effects influence technologies' maturities, especially for fuel cells and heat-pumps, are incorporated in the decision mechanism, in order to represent in a realistic way the consumption patterns, the technology evolution, and the rigidities involved in the decision mechanism. Energy performance largely depends on the characteristics of the dwelling (thermal integrity) and the technology of the equipment which uses energy. Individual energy consumers can spend money to improve energy efficiency and select solutions with upfront costs and utilisation performance leading to reasonable pay-back periods. Energy efficiency progress implies high upfront cost but saves on variable and operating costs during the lifetime of the energy equipment.

In the context of the WHY project, the buildings module of PROMETHEUS is significantly improved with an enhanced representation of different technology types (e.g. based on their efficiencies), electricity load of households, the impacts of energy efficiency policies and the distributional impacts of decarbonisation of buildings across regions.

[22] Fragkos, P., Kouvaritakis, N. & Capros, P. Incorporating Uncertainty into World Energy Modelling: the PROMETHEUS Model. Environ Model Assess 20, 549–569 (2015). https://doi.org/10.1007/s10666-015-9442-x

3.6 Building Scaling Model and House Infrastructure Model

The *Code*-Elements “Building Scaling” and the “House Infrastructure” are described in one chapter as they feature similar methods, characteristics and approaches. The models and methods developed for and applied in the “House Infrastructure” will also be used in the “Building Scaling”. After the Causal Model has decided that an investment in new technologies is due within a certain residential building, the “Building Scaling” *Code*-Element is used to define which technology will be applied, what the technical parameters of the technology are and how it will be applied. As an example:

> “*The Causal Model decided that the inhabitants of a residential building want, due to external stimuli, invest into renewable generation capacities and maybe even install a battery storage system. The Causal Model will not define the technology itself nor the technical parameters to be applied. This is where the Building Scaling Model will come in”*

The House Infrastructure Model on the other hand considers the already existing and the newly installed components and performs an optimisation of their behaviour towards a certain goal (minimisation of operational costs, maximisation of own consumption, etc.). As an example:

> *“A residential building has a PV generation capacity installed as well as a heat pump that can be controlled via a smart home system. The House Infrastructure Model will then optimise the time at which the heat pump will be used considering the comfort parameters of the inhabitants and the available renewable energy.”*

The main components and approaches featured in those two models are shown in Fig. 11.

Both models will facilitate the HiSim simulation environment and certain versions will be implemented within HiSim.

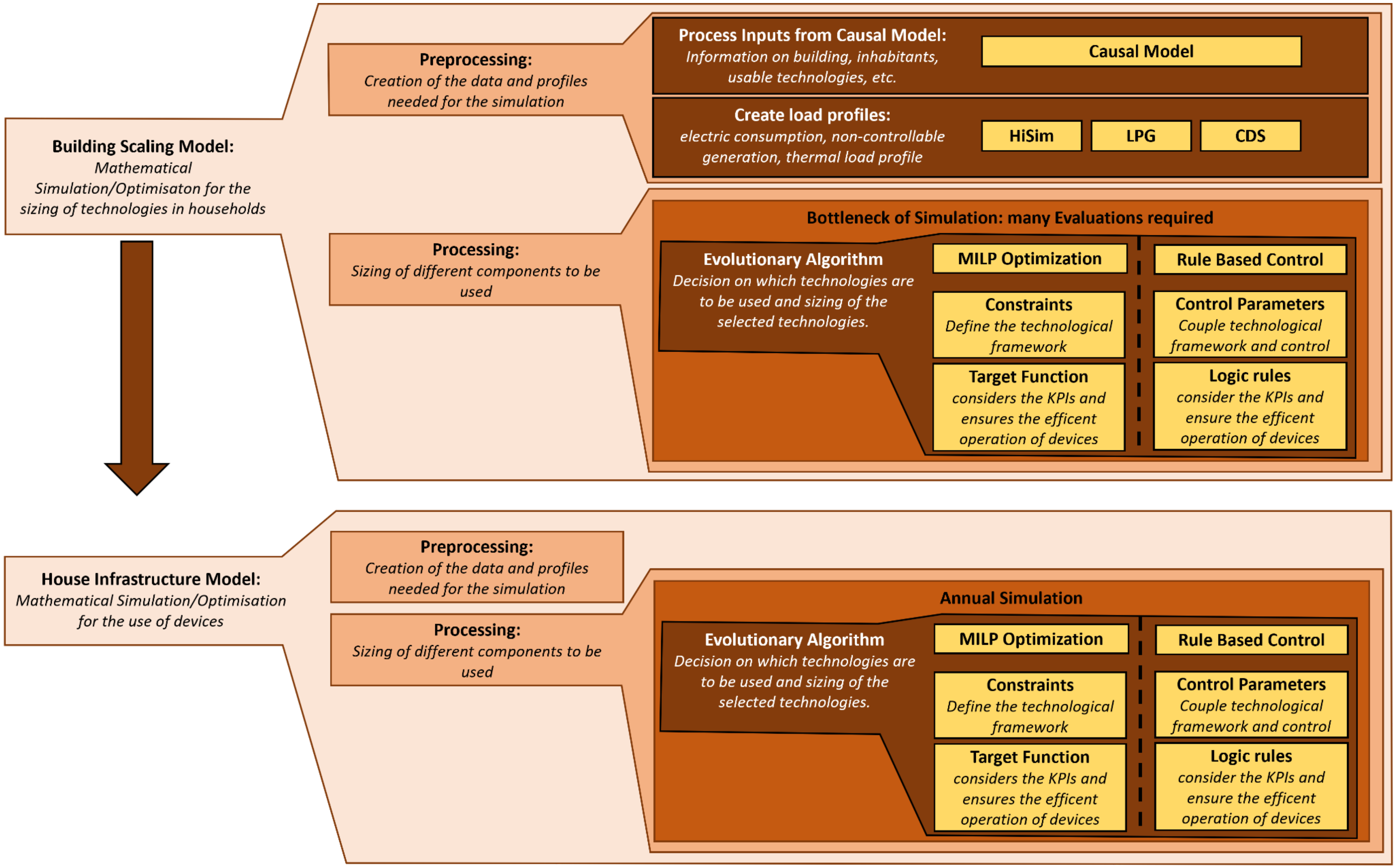

Figure 11: Architecture of the Building Scaling Model and the House Infrastructure Model

3.6.1 Building Scaling Model

As mentioned the purpose of the Building Scaling Model (BMS) is the optimisation and sizing of new technologies within a residential building. The model is separated into two distinct parts:

- **Preprocessing:**
  During Preprocessing all steps necessary to provide or generate the data required for the sizing algorithm will be performed. First off the data from the causal model will be processed. This includes information on the building and the inhabitants but more importantly it contains information on the decision of the inhabitants of the residential building on what they want to invest in (if different technologies are an option) and what the investment parameters are. Furthermore the individual goals for the investment are defined. As an example:

  > *"The inhabitants of the residential building want to invest into a renewable energy source to maximise their independence from grid supply."*

  Building on the general data on inhabitants, the building itself, etc. the second step of the preprocessing phase will be initiated, where the models for HiSim, which will incorporate the models for the LPG-El and the models for the LPG-Tr will be used to generate the corresponding (load) profiles for the processing step. During the preprocessing step all non-controllable load and generation profiles will be created to be used during the processing phase. As an example:

  > *"The residential building is a 3 person household with regular use of devices and no electric mobility. The building is of low energy standard, warm water and heating are both provided by a heat pump. This will create electric load profiles for all devices except the heat pump (since it is controllable) and the thermal load profile for the building, which will be used to simulate the behaviour of the heat pump later on."*

  Thus the preprocessing step generates all the (load profile) data that is fixed for a simulation, for instance non flexible generation capacities, or non-shiftable loads.

- **Processing:**
  During the processing phase two things are being considered. Firstly the sizing of the components that need to be sized and secondly the simulation of the households considering the flexible loads and generation capacities. The sizing phase involves a Evolutionary algorithm that calls the simulation of the household with various configurations. Evolutionary algorithms find high-quality solutions to optimization problems, without guaranteeing that the given solutions are the best available (metaheuristics). The advantage of Evolutionary algorithms is that they can be parallelized and perform well. For that reason, they have been applied to industrial problems like the design of water distribution systems, the operation of

reservoir systems, or the allocation of waste load, as well as to optimise decision trees or hyperparameters in advanced modelling applications[23].

Evolutionary algorithms are inspired by the process of biological evolution. The core idea of Evolutionary algorithms is to let the fittest individuals survive and procreate. An individual is a possible solution to a given optimization problem and its fitness can be evaluated by the corresponding objective function. Starting from a randomly generated set of individuals (population) the fittest are stochastically selected and their encoding is modified by operations such as mutation or cross over to create a new population. The fitness of each individual of the new generation is evaluated and the fittest ones are again stochastically selected to be parents of a new population. Evolutionary algorithms are iterative processes, which population by population evolute towards better solutions of the optimization problem.[24]

A further advantage of Evolutionary algorithms is that they are robust to different kinds of fitness functions. More concretely, within the WHY toolkit the application of Evolutionary algorithms makes the decoupling of the sizing problem and the control/operation of the house infrastructure possible. As the coupled problem is quite complex and yields high computational efforts the application of Evolutionary algorithms is considered as a significant improvement in terms of computational performance when compared to pure mixed integer linear programming formulations.

The Evolutionary algorithm will be used for the sizing part of the Building Scaling Model, defining for instance the size of a PV-Generator within the technical parameters of the building (roof size, etc.), or the sizing of a heat pump system. As an example:

> *"The Evolutionary Algorithm changes the technology of the renewable generation capacity from wind power to PV from one evolution to the other and then sets the number of PV-modules, the inclination and the general direction with each new evolution."*

To compute how fit the solution (technology and parameters) is, the components defined by the Evolutionary algorithm need to be applied to the building simulation and their performance simulated. To continue the example above:

> *"In the current evolution step, a 5 kWp PV generator is installed. The smart home system is capable of shifting the time of use of the already installed heat pump to maximise the own-consumption rating which is simulated. In the next evolution step, the installed capacity is increased to 10 kWp and the simulation is rerun. The degree of fitness is the reached own-consumption values which are compared before the next evolution step."*

[23] Maier, H. R., Razavi, S., Kapelan, Z., Matott, L. S., Kasprzyk, J., & Tolson, B. A. (2019). Introductory overview: Optimization using evolutionary algorithms and other metaheuristics. Environmental modelling & software, 114, 195-213.
[24] Eiben, A. E., & Smith, J. E. (2003). Introduction to evolutionary computing. Berlin: springer.

For this simulation to obtain the fitness of the current evolutionary step (namely, the actual result in terms of technical, economic and environmental impacts of the solution), two different approaches will be implemented and thoroughly tested. The objective is to evaluate which one performs better in terms of computation time and the results obtained:

- **MILP Optimisation:**
  The household simulation will be implemented in a mixed integer linear programming (MILP) problem, consisting of equations and inequations describing the technical components and their interdependencies. Fixed values, such as non-flexible generation or load profiles, will be considered but not changed during the optimisation. The MILP optimisation will be implemented in Python using the [Pyomo optimisation library](http://www.pyomo.org/)[25], or the [PyPSA linopy optimization interface](https://github.com/PyPSA/linopy)[26], which is currently under development. The MILP algorithm will ensure that the set constraints concerning the operation of the devices are met. The target function of the optimisation will be defined through the parameters provided by the Causal Model, which basically is the key performance indicator or fitness of the Evolutionary algorithm. Thus the MILP Optimisation will be used to simulate flexible loads or other flexibilities and define their time and duration of use. In case that the Evolutionary algorithm sizes a non-controllable load or a non-controllable generation capacity and there are no flexible loads or nothing to be optimised, the MILP algorithm can still be executed but only a non optimised solution under the given circumstances will be generated. To continue the example above:

  > *"The MILP optimisation will optimise the operation time and operation parameters of the heat pump considering the temperature requirements of the inhabitants (constraints)"*

- **Rule Based Control Scheme:**
  The second approach features a ruled based control scheme for controllable loads and controllable generation devices. For that purpose a set of fixed rules provide the foundation for the decision of when and how to use a controllable load or a controllable generation capacity. This set of rules will furthermore decide upon a hierarchy of the different options of controllable assets, to ensure if two options are available at a certain time, the more relevant one will be chosen. The Rule Based Control Scheme will be implemented directly in HiSim and will be a better representation of smart home control units than the MILP approach, as today's technology for home automation is rather rule based than using a model predictive control algorithm. To continue the example above:

  > *"For the Rule Based Control Scheme the following rules have been defined for the use of the heat pump:*

[25] http://www.pyomo.org/
[26] https://github.com/PyPSA/linopy

1. *Activate the heat pump when the temperature in the household gets too low, regardless of the current availability of PV generation.*
2. *Activate the heat pump whenever the surplus of the PV generation is above a certain threshold and the room temperature is below a certain temperature.*
3. *Active the heat pump only when forecasts predict that a surplus will persist over the next quarter hour to reduce too many on/off-switches.*

The Algorithm then sets the time of use of the heat pump accordingly and calculates the KPI or fitness values.

The Evolutionary Algorithm will define the parameters of the technology to be installed and call either the Python Script featuring the MILP or HiSim featuring the Rule Based Control Scheme which will then simulate the household and return the relevant KPIs to define the fitness of the solution. After multiple rounds of "evolutions", the final result will provide the best solution for the technical parameters of the technology to be installed under the given circumstances.

Since the Evolutionary Algorithm used in the Building Scaling Model will use quite some computation time, only a single year will be considered for each run at the point where an investment decision is made.

3.6.2 Household Infrastructure Model

The Household Infrastructure Model works in the same manner and uses the same methods as the Building Scaling Model, aside from the Evolutionary Algorithm. For the Household Infrastructure Model the technical parameters of all technologies are already set and the simulations are run with these given parameters. Furthermore there is a difference between the Household Infrastructure Model and the Building Scaling Model as the Household Infrastructure Model will simulate multiple years of data, whereas the Building Scaling Model will be run just for a single year in order to define the technical parameters.

3.7 Uptake

This Code-Element is responsible for translating the results of the previous Code-Elements to useful inputs for large scale ESMs. Two different approaches will be followed: bootstrapping and AI modelling. In both cases, the two methods are built over a potentially large database of precomputed results calculated with the previous Code-Elements. As most of the algorithms used in the previous Code-Elements are stochastic, this database will contain several repetitions of the same scenario in order to assess the probability distribution of the actual intervention. Table 1 shows the layout of the database constructed. Please note that each cell does not denote a single value but several copies of the yearly distribution of the energy consumption for this particular intervention at this particular Cluster Type (Cluster of Electric Behaviour are defined in

Deliverable D2.1 and consist of the different electric behaviour found in the load profiles datasets assess in the project).

Table 1: Example of individual for the optimization of interventions

	Cluster Type$_1$	**Cluster Type$_2$**	$\cdots$	**Cluster Type$_m$**
Intervention$_1$	$Load\ profile_1^1$	$Load\ profile_1^2$	$\cdots$	$Load\ profile_1^m$
Intervention$_2$	$Load\ profile_2^1$	$Load\ profile_2^2$	$\cdots$	$Load\ profile_2^m$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
Intervention$_n$	$Load\ profile_n^1$	$Load\ profile_n^2$	$\cdots$	$Load\ profile_n^m$

From this table, the two different approaches diverge.

- **Bootstrapping**: the objective of this task is to build a realistic aggregated load profile for a particular population. To this end, first it is needed a description of the population in terms of the Cluster Types, namely its composition. Then, following the bootstrapping resampling method[27], a new sample of the right size is built. The method just takes an element from the database at random following the distribution of the Cluster Types and adds it to the total load profile. As generating random elements and selecting a cell in the table is a fast operation, it is expected to be able to construct realistic load profiles for large populations quickly. Moreover, as we are using a bootstrap approach, it is possible to easily get relevant parameters of the load profile distribution used (like confidence intervals, etc.).
- **AI modelling**: Table 1 will be enriched with the socio economic description of each Cluster Type (see Deliverable D2.1 and D2.3 for a detailed description of this). Then, an AI method (e.g. Deep Neural Networks[28]) will be fitted to this data. Standard validation methods[29] will be used to assess the quality of the fit. If a suitable model is found, then this method can be used as a surrogate model[30] of the WHY toolkit whenever it is needed (for example, to generate new bootstrapping samples).

3.8 Result Processor and Intervention Optimizer

This Code-Element is responsible for the final processing of the information created by the WHY toolkit. In particular, it will post process the results to produce:

- **The social impact assessment:** to complete this task, this Code-Element will follow the methodologies presented in Deliverable D1.3 to estimate each KPIs defined for each use case.

[27] https://en.wikipedia.org/wiki/Bootstrapping_(statistics)
[28] Schmidhuber, J. (2015). "Deep Learning in Neural Networks: An Overview". *Neural Networks*. **61**: 85–117. arXiv:1404.7828. doi:10.1016/j.neunet.2014.09.003.
[29] Lalli Myllyaho et al. Systematic literature review of validation methods for AI systems, Journal of Systems and Software, Volume 181, 2021, 111050, https://doi.org/10.1016/j.jss.2021.111050.
[30] https://en.wikipedia.org/wiki/Surrogate_model

- **The ethics evaluation of the end results** including an assessment of all the algorithms provided: The ethics evaluation is going to be based in the development of an constructive ethics-regulation interface (ERI) **capable** to support the process of cross- fertilisation between the ethics, regulation approaches and ethical perspectives (Utilitarianism, Deontologist, Process-governance, Virtue theory and Care-vulnerability ethics) and **ensure** the engagement (constructive capacity), necessaries to achieve trustworthy results in terms of lawful, ethical and robustness[31] following the EU guidelines for data intensive technologies.

  Moreover, and despite the normative consideration, the ERI, will advocate for a strong shift towards 'ethical governance approach' instead of normative utility, which can be achieved integrating an adaptive and flexible frameworks, based on the insights from: *ethical design manifestos and transformative change innovation paradigms* focused on human rights integration[32],[33] *adaptive organisational roles and practices* proposals, *alternative technology design principles and ethical systems designs assessments*[34],[35]*FATE in algorithms frames*[36] and transparency, autonomy, data privacy principles of Digital integration charts[37].

  At methodological level, the development of the **constructive ERI** will be based in the consideration of three operational elements:

  (1) Professional ethics frameworks (since have resources to problematize some overly simplistic assumptions and transcend to 'moral sums' and provide systematised tools committed to real action);
  (2) Regulation ethics frameworks (since have well-established and evolving tradition of cross-fertilisation between the political, ethical and legal considerations) and
  (3) Evolving ethical perspectives such as Virtue ethics[38] (applicable for example in the case of algorithmic opacity and datafication) or Process ethics regarding decision making process (for example promoting citizen-centred data governance).

  In addition, constructive capacity will be based in the acknowledgement of human agency relevance (since plays a crucial role in decision-making and demands a reflection on the psychological and environmental factors that impact such decision-making) and the commitment regarding the decisions around the

[31] EU Ethics Guidelines for Trustworthy AI (8 April 2019) and Draft Ethics Guidelines for Trustworthy AI (18 December 2018) of the European Commission's High-Level Expert Group on Artificial Intelligence https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai15
[32] Balkan A. (2017) Ethical design manifesto. Available at: https://2017.ind.ie/ethical-design/
[33] University of Deusto Declaration Human Rights in Digital Environments, Available at: https://www.deusto.es/cs/Satellite/deusto/en/university-deusto/information-about-deusto/human-rights-in-digital-environments?cambioidioma=si
[34] European Professional Ethics Framework for the ICT Profession, 2021 Available at:https://www.ict-ethics.eu/
[35] EEE Standard Model Process for Addressing Ethical Concerns during System Design, 2021, Available at: https://standards.ieee.org/ieee/7000/6781/
[36] Fairness, accountability, transparency, and ethics (FATE) in algorithmic systems.
[37]Spanish charter of digital rights, 2021, Available at: https://www.lamoncloa.gob.es/lang/en/presidente/news/Paginas/2021/20210714_digital-rights.aspx
[38] Gal, U., Jensen, T. B., & Stein, M. K. (2020). Breaking the vicious cycle of algorithmic management: A virtue ethics approach to people analytics. Information and Organisation, 30(2), 100301. https://doi.org/10.1016/j.infoandorg.2020.100301

management of personal data, to safeguard every person's 'self-vulnerability' and 'rights vulnerability'[39].

- **Any contrafactual assessment**: simulations with or without the interventions will be carried out and the relevant statistical assessment will be carried out in this action in order to assess the potential impact that the interventions could have achieved. Please note that this activity requires the input of the actual results, so for this end, this will be constrained to just one year or retrospective assessments.

Finally, this Code-Element will be used to optimise the interventions to be carried out in each of the use cases. To this end, a similar approach as the one followed on the Building Scaling Model will be used. In this regard, the interventions considered in each use case will be parametrized (for a detailed description of the interventions, please consult Deliverable D1.3). In most of the cases, the parametrization will be trivial (namely, the intervention is set or not) but in others one or several parameters will be present. Table 2 shows an example of the parametrization.

Table 2: Example of individual for the optimization of interventions

	Parameters
Intervention$_1$	$\varnothing$
Intervention$_2$	μ, σ
$\vdots$	$\vdots$
Intervention$_n$	κ, θ

Then, an Evolutionary Algorithm will be used to optimise the set of intervention used. This algorithm will create populations of potential interventions. Evolutionary operators able to evolve a set of interventions will be created for example by creating a small random mutation or crossing two sets of interventions. Finally, to assess the quality of the solution (fitness function), a particular KPI of each use case will be taken and optimised (maximised or minimised).

Through the use of this algorithm, an optimised set of interventions will be constructed for each use case. Single objective evolutionary algorithms will be tested in general. Nevertheless, it is not clear that it would be feasible to run this method on the large scale ESM use case (namely, the EU and Global use cases) as the run-time of each simulation will be prohibitive. Simplification and the use of cloud computing will be examined before dropping the use of this Code-element on these use cases. Moreover, if enough time is provided, it would be considered to test also multi criteria optimization algorithms such as Non-dominated Sorting Genetic Algorithm-II[40]. In this case, several KPIs will be selected and the Pareto Frontier[41] will be constructed. This activity is even more computationally expensive than the previous ones, so it is only expected to work on the energy community use case (if at all).

[39] Delacroix, S., & Wagner, B. (2021). Constructing a mutually supportive interface between ethics and regulation. Computer Law and Security Review, 40, 105520. https://doi.org/10.1016/j.clsr.2020.105520

[40] Deb, K.; Pratap, A.; Agarwal, S.; Meyarivan, T. (2002). "A fast and elitist multiobjective genetic algorithm: NSGA-II". IEEE Transactions on Evolutionary Computation. 6 (2): 182. CiteSeerX 10.1.1.17.7771. doi:10.1109/4235.996017.

[41] https://en.wikipedia.org/wiki/Pareto_front