Model formulation

Model overview:

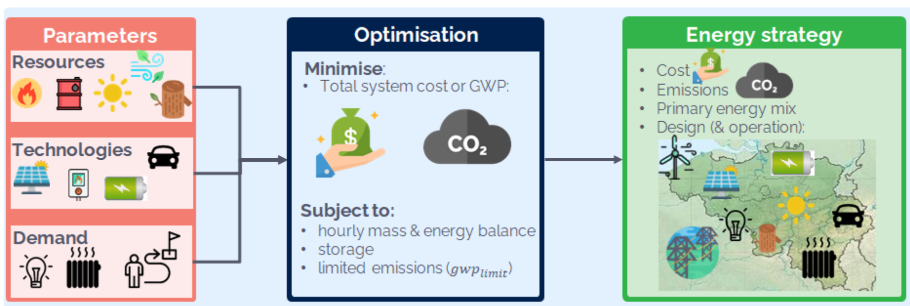

Figure 1: Overview of the LP modelling framework.

Overview

Due to computational restrictions, energy system models rarely optimise over the 8760 hours of the year. For example, running our model with 8760h time series takes more than 19 hours, while it takes only around 1 minute with the methodology presented hereafter.

A typical solution is to use a subset of Typical (i.e. representative) Days called TD; this is a trade-off between introducing an approximation error in the representation of the energy system (especially for short-term dynamics) and computational time.

Running the EnergyScope Typical Day (ESTD) model consists in two steps:

- the first step consists in pre-processing the time series and solving a MILP problem to determine the adequate set of typical days (`sec_td_selection`).
- the second step is the main energy model: the optimal design and operation of the energy system over the selected typical days is computed i.e. technology selection, sizing and operation for a target future year (`sec_estd`)

These two steps can be performed independently. Usually, the first step is computed only once for an energy system with given weather data whereas the second step is computed several times (once for each different scenario). `Figure %s <fig:ProcessStructure>` illustrates the overall structure of the code.

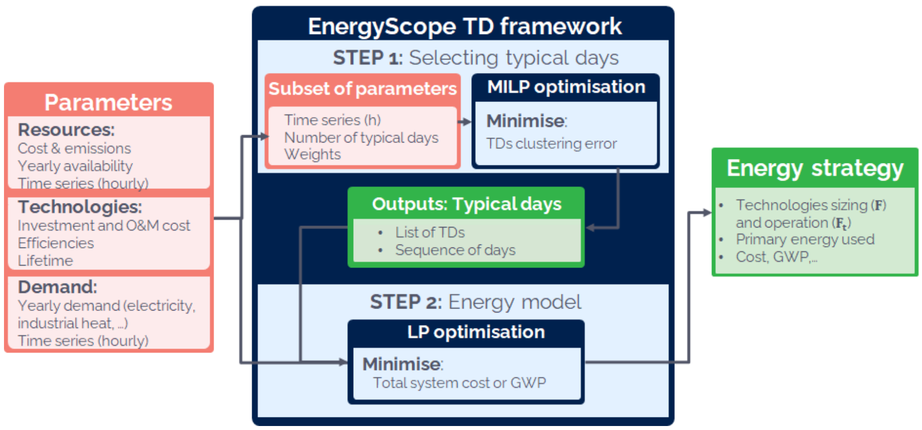

Figure 2: Overview of the EnergyScope TD framework in two-steps. **STEP 1**: optimal selection of typical days (`sec_td_selection`). **STEP 2**: Energy system model (`sec_estd`). The first step processes only a subset of parameters, which account for the 8760h time series. Abbreviations: Typical Day (TD), Linear Programming (LP), Mixed Integer Linear Programming (MILP), Global Warming Potential (GWP).

This documentation is built based on previous works Moret2017PhDThesis,Limpens2019,Limpens2021thesis. For more details about the research approach, the choice of clustering method or the reconstruction method, see Limpens2021thesis.

Typical days selection

Resorting to TDs has the main advantage of reducing the computational time by several orders of magnitude. Usually, studies use between 6 and 20 TDs Gabrielli2018,Despres2017,Nahmmacher2014,Pina2013 and sometimes even less Poncelet2017,Dominguez-Munoz2011.

Clustering methods

In a previous work Limpens2019, it has been estimated that 12 typical days were appropriate for this model. Moreover, a comparison between different clustering algorithms showed that the method of Dominguez-Munoz2011 had the best performances.

Implementing seasonality with typical days

Using TDs can introduce some limitations. For example, model based on TDs are traditionally not able to include inter-days or seasonal storage due to the discontinuity between the selected days. Thus, they assess only the capacity of production without accounting for storage capacities. Carbon-neutral energy system will require long term storage and thus, this limitation had to be overcome. Therefore, we implemented a method proposed by Gabrielli2018 to rebuild a year by defining it as a sequence of typical days. This allows to optimise the storage level of charge over the 8760 hours of the year. Gabrielli2018 assigned a TD to each day of the year; all decision variables are optimised over the TDs, apart from the amount of energy stored, which is optimised over 8760 hours. This methodology is illustrated in `Figure %s <fig:SeasonalityImplementation>`.

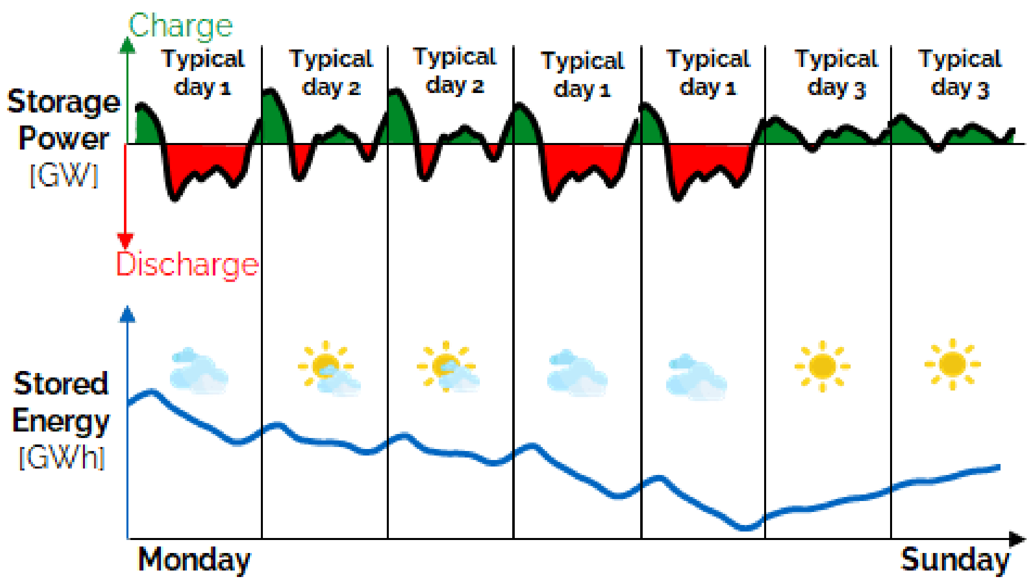

Figure 3: Illustration of the typical days reconstruction method proposed by Gabrielli2018 over a week. The example is based on 3 TDs: TD 1 represents a cloudy weekday, applied to Monday, Thursday and Friday; TD 2 is a sunny weekday, applied to Tuesday and Wednesday; finally, TD 3 represents sunny weekend days. The power profile (above) depends solely on the typical day but the energy stored (below) is optimised over the 8760 hours of the year (blue curve). Note that the level of charge is not the same at the beginning (Monday 1 a.m.) and at the end of the week (Sunday 12 p.m.).

The performances of this method have been quantified in a previous work Limpens2019. With 12 typical days, the key performance indicators (cost, emissions, installed capacity and primary energy used) are well captured. The only exception are the long term storage capacities, which are slightly underestimated (by a factor of 2 at most).

Energy system model

In this section, we present the core of the energy model. First, we introduce the conceptual modelling framework with an illustrative example. This helps to clarify the nomenclature as well. Second, we introduce the constraints of the optimization problem. The data used in the model, on the other hand, are detailed in the sections `/sections/Input Data - Colombia` and `/sections/Input Data - Turkey`.

Linear programming formulation

The model is mathematically formulated as a LP problem fourer1990modeling. `Figure %s <fig:linear_programming_example>` represents - in a simple manner - what is a LP problem and the related nomenclature. In italic capital letters, *SETS* are collections of distinct items. For example, the *RESOURCES* set regroups all the available resources (NG, WOOD, etc.). In italic lowercase letters, *parameters* are known values (inputs) of the model, such as specific end-use demands and resource availabilities. In bold with first letter in uppercase, **Variables** are unknown values of the model, such as the installed capacity of PV. These values are determined (i.e. optimised) by the solver within an upper and a lower bound (both being parameters). For example, the installed capacity of wind turbines is a *decision variable*. This quantity is bounded between zero and the maximum available wind potential. *Decision variables* can be split in two categories: independent decision variables, which can be freely fixed, and dependent decision variables, which are linked via equality constraints to the previous ones. For example, the investment cost for wind turbines is a variable but it directly depends on the installed capacity of wind turbines, which is an independent decision variable. *Constraints* are inequality or equality that must be satisfied between variables parameters. The problem is subject to (*s.t.*) constraints that can enforce, for example, an energy balance or an upper limit for the availability of resources. Finally, an *objective function* is a particular variable whose value is to be minimised (or maximised).

Conceptual modelling framework

The proposed modelling framework is a simplified representation of an energy system, accounting for the energy flows within its boundaries. Its primary objective is to satisfy the energy balance constraints, meaning that the demand is given and the supply has to meet it. In energy modelling practice, the energy demand is often expressed in terms of Final Energy Consumption (FEC). According to the definition of the European commission, FEC is defined as "*the energy which reaches the final consumer's door*" EU_FEC. In other words, the

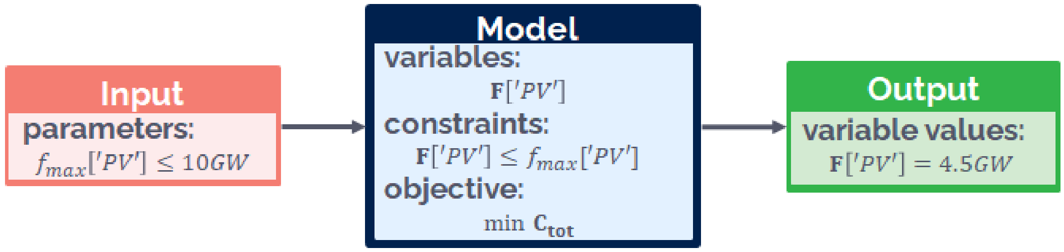

Figure 4: Conceptual illustration of a LP problem, with the related nomenclature. Description of symbols: maximum installable capacity for a technology (f_{max}), installed capacity of a technology (**F**) and total system cost (**C_tot**). In this example, a specific technology (**F** $['PV']$) has been chosen from the set TECHNOLOGY.

FEC is the amount of input energy needed to satisfy the end-use demand (EUD). For example, in the case of decentralised heat production with a natural gas boiler, the FEC is the amount of natural gas consumed by the boiler, while the EUD is the amount of heat produced by the boiler i.e. the heat delivered to the user.

The input of the proposed modelling framework here is not the FEC but the EUD, related to five energy sectors: electricity, heating, cooling, mobility and non-energy. This replaces the classical economic-sector based representation of energy demand. Heat is divided in three end-use types (EUT): high temperature heat for industry, low temperature heat for space heating and low temperature heat for hot water. Cooling is divided in two EUTs: process cooling (for industry) and space cooling. Mobility is divided in two EUTs: passenger mobility and freight[1]. Non-energy demand is, based on the IEA definition, "*fuels that are used as raw materials in the different sectors and are not consumed as a fuel or transformed into another fuel.*" IEA_websiteDefinition. For example, the European Commission classifies as non-energy the following materials: "*chemical feed-stocks, lubricants and asphalt for road construction.*" EuropeanCommission2016.

A simplified conceptual example of the energy system's structure is proposed in Figure %s <fig:conceptual_example>. The system is split in three parts: resources, energy conversion and demand. In this illustrative example, resources are solar energy, electricity and natural gas (NG). The EUDs are demands for electricity, space heating and passenger mobility. The energy system encompasses all the energy conversion technologies needed to transform resources in order to fulfill the EUD. In this example, solar and NG resources cannot be directly used to supply heat. Thus, technologies are used, such as boilers or cogenerations of heat and power (CHP) using NG, to supply the EUT layers (in this case, the *high-temperature heat for industry* layer). *Layers* are defined as all the elements

[1]Air passenger transport is accounted for in passenger mobility (excluding international flights).

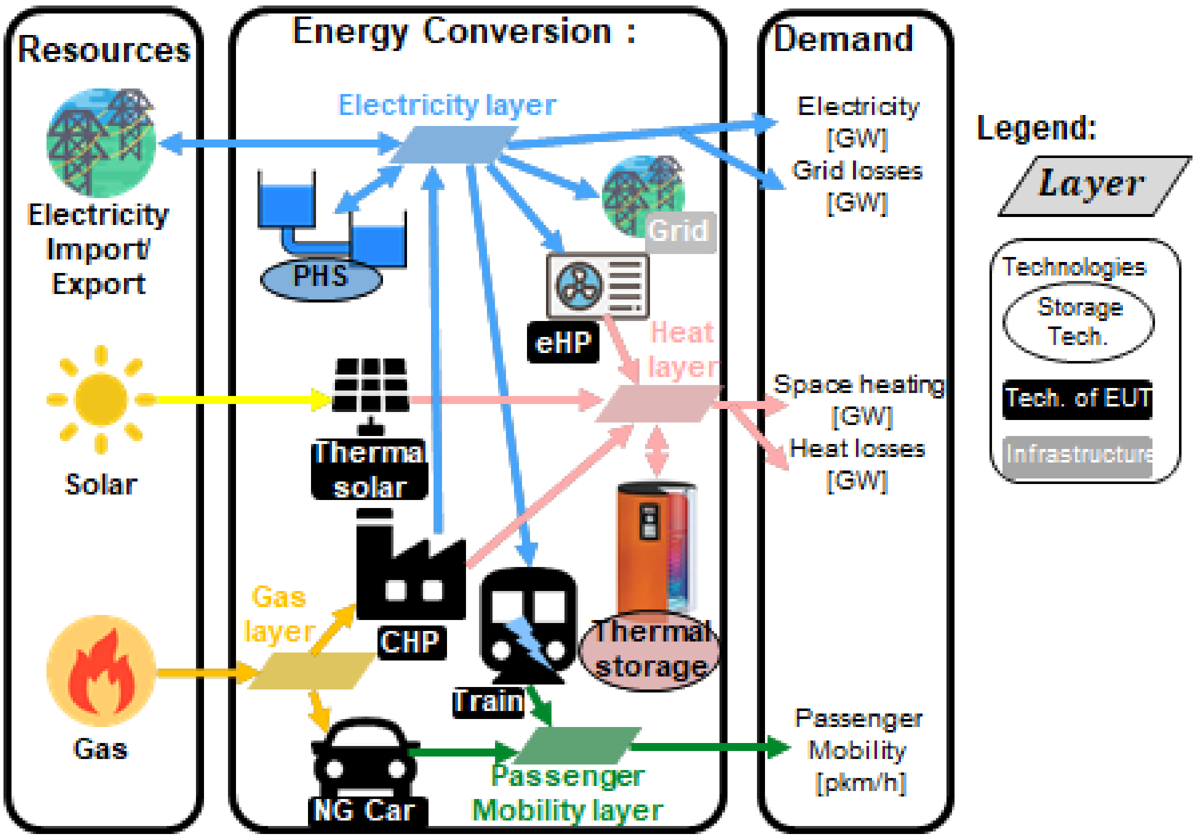

Figure 5: Conceptual example of an energy system with 3 resources, 3 EUDs and 8 technologies, among which 2 of storage type (coloured oval) and 1 of infrastructure type (grey rectangle). Abbreviations: pumped hydro storage (PHS), electrical heat pump (eHP), cogenerations of heat and power (CHP), natural gas (NG). Some icons from FlatIcon.

in the system that need to be balanced in each time period; they include resources and EUTs. For example, the electricity layer must be balanced at any time, meaning that the production and storage must equal the consumption and losses. These layers are connected to one another by *technologies*. We define three types of technologies: *technologies of end-use type*, *storage technologies* and *infrastructure technologies*. A technology of end-use type can convert the energy (e.g. a fuel resource) from one layer to an EUT layer, such as a CHP unit that converts NG into heat and electricity. A storage technology converts energy from a layer to the same one, such as thermal storage (TS) that stores heat to provide heat. In our example (`Figure %s <fig:conceptual_example>`), there are two storage technologies: TS for heat and pumped hydro storage (PHS) for electricity. Infrastructure technologies include all the remaining technologies, including the networks, such as the power grid and district heating networks (DHN), but also technologies linking non end-use layers, such as methane production from wood gasification or hydrogen production from methane reforming.

As an illustrative example of the concept of *layer*, `Figure %s <fig:LayerElec>` presents a sketch of the electricity layer, which is the most complex one since the electrification of other sectors is foreseen as a key element of the energy transition Sugiyama2012. In the version of EnergyScope described in this document, 54 technologies are related to the electricity layer. 16 technologies produce electricity exclusively, such as combined cycle gas turbine (CCGT), rooftop PV or onshore wind. 14 cogenerations of heat and power (CHPs) produce both heat and electricity, such as industrial waste CHP. 10 technologies are related to the production of synthetic fuels and to carbon capture and storage (CCS). 1 infrastructure technology represents the electrical grid. 6 storage technologies are implemented, such as PHS, batteries or vehicle-to-grid (V2G). The rest relates to the electrification of heat and mobility. Electrification of the heating sector is supported by direct electric heating and, most importantly, by electrical heat pumps. Electrification of mobility is achieved via electric public transportation (train, trolley, metro and electrical/hybrid bus), electric private transportation including battery and hydrogen cars[2] and electric freight with trains.

The model is formulated as a LP problem. It optimises the design of the energy system by computing the installed capacity of each technology, as well as the operation in each period, that minimizes the total annual cost of the system while meeting the energy demand. In the following sections, we present the complete formulation of the model in two steps. First, all the terms used are summarised in a figure and a set of tables (`Figure %s <fig:sets>` for sets, Tables `%s <tab:paramsDistributions>` and `%s <tab:params>` for parameters, Tables `%s <tab:variablesIndependent>` and `%s <tab:variablesdependent>` for variables). Second, on this basis, the equations representing the constraints and the objective function are presented in `Figure %s <fig:EndUseDemand>`

[2] Hydrogen can be produced based on many feedstocks, among which electricity via the use of electrolysers.

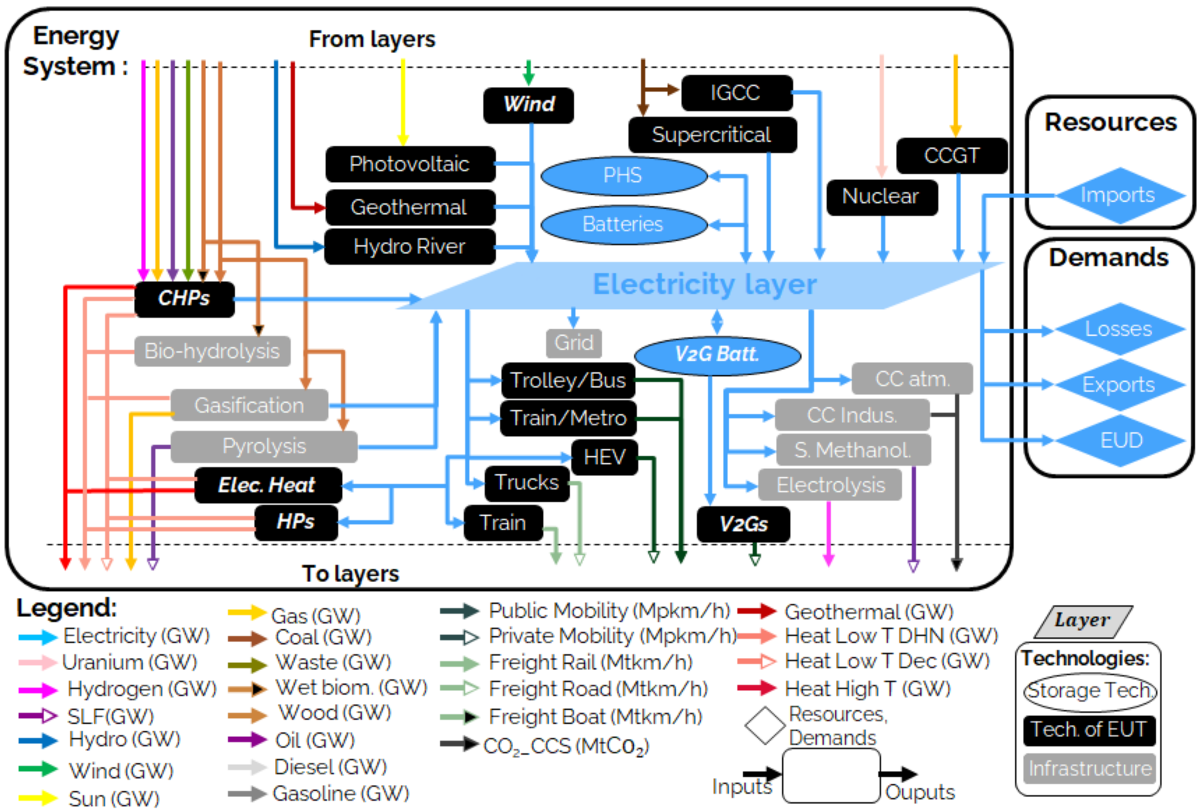

Figure 6: Representation of the *electricity* layer with all the technologies implemented in ESTD v2.1. (The version descripted in this document is v2.2.) Bold italic technologies represent a group of different technologies. Abbreviations: electricity (elec.), industrial (ind.), combined cycle gas turbine (CCGT), integrated gasification combined cycle with coal (IGCC), cogeneration of heat and power (CHP), heat pump (HP), pumped hydro storage (PHS), vehicle-to-grid (V2G), synthetic methanolation (S. Methanol.), atmospheric (atm.), carbon capture (CC), end-use demand (EUD).

and detailed in Eqs. `eq:obj_func` - `eq:efficiency`.

Sets, parameters and variables

`Figure %s <fig:sets>` gives a visual representation of the sets used in EnergyScope, together with their respective indices. Tables `%s <tab:paramsDistributions>` and `%s <tab:params>` list and describe the model parameters. Tables `%s <tab:variablesIndependent>` and `%s <tab:variablesdependent>` list and describe the independent and dependent variables, respectively.

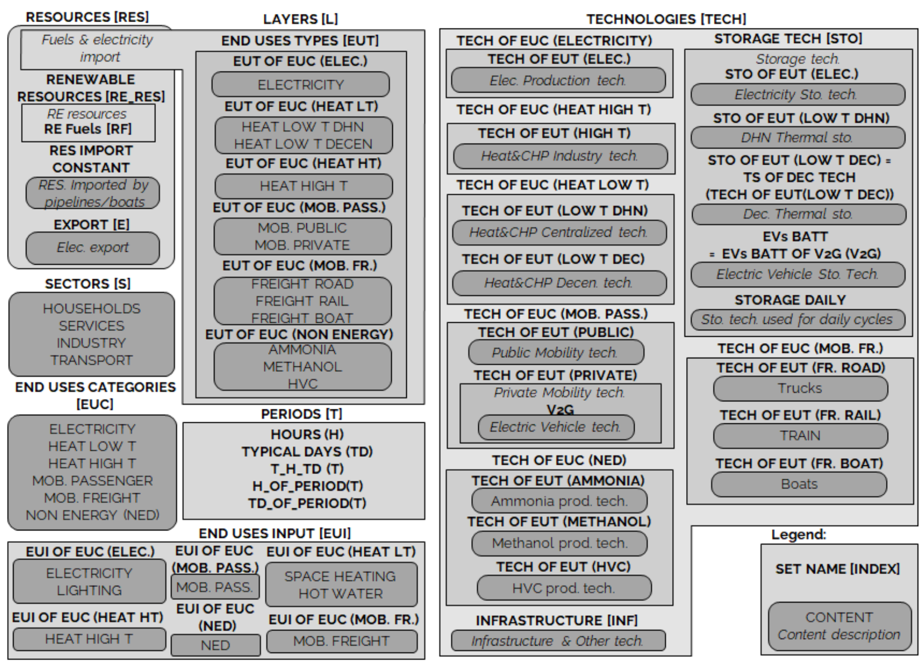

Figure 7: Visual representation of the sets and indices used in the LP framework. This figure was produced for ESTD v2.1 and does not include the latest technologies and EUDs included in v2.2, such as the ones related to cooling. Abbreviations: space heating (SH), heating water (HW), temperature (T), mobility (MOB), passenger (*Pass.*), vehicle-to-grid (V2G), thermal storage (TS).

Table 1: Exhaustive list of time series parameters used in ESTD v2.2.

Parameter	Units	Description
$\%_{elec}(h,td)$	[-]	Yearly time series (adding up to 1) of electricity EUD

Parameter	**Units**	**Description**
$\%_{sh}(h,td)$	[-]	Yearly time series (adding up to 1) of space heating EUD
$\%_{sc}(h,td)$	[-]	Yearly time series (adding up to 1) of space cooling EUD
$\%_{pass}(h,td)$	[-]	Yearly time series (adding up to 1) of passenger mobility EUD
$\%_{fr}(h,td)$	[-]	Yearly time series (adding up to 1) of freight mobility EUD
$c_{p,t}(tech,h,td)$	[-]	Hourly maximum capacity factor for each technology (default 1)

Table 2: Exhaustive list of parameters (except time series) used in ESTD v2.2.

Parameter	Units	Description
τ(tech)	[-]	Investment cost annualization factor
i_{rate}	[-]	Real discount rate
$endUses_{year}(eui,s)$	[GWh/y] [a]	Annual EUD values divided per sector
$endUsesInput(eui)$	[GWh/y] [a]	Total annual EUD values
re_{share}	[-]	Minimum share [0;1] of primary energy coming from renewables
gwp_{limit}	[$ktCO_2$-eq/y]	Upper CO_2-eq emissions limit
$\%_{public,min}, \%_{public,max}$	[-]	Lower and upper limit to $\%_{\mathbf{Public}}$
$\%_{fr,rail,min}, \%_{fr,rail,max}$	[-]	Lower and upper limit to $\%_{\mathbf{Fr,Rail}}$
$\%_{fr,boat,min}, \%_{fr,boat,max}$	[-]	Lower and upper limit to $\%_{\mathbf{Fr,Boat}}$
$\%_{fr,truck,min}, \%_{fr,truck,max}$	[-]	Lower and upper limit to $\%_{\mathbf{Fr,Truck}}$

Parameter	Units	Description
$\%_{private,motorc,max}$	[-]	Max. share of private mobility supplied by motorcycles
$\%_{dhn,min}, \%_{dhn,max}$	[-]	Lower and upper limit to $\mathbf{\%_{Dhn}}$
$\%_{ned}(EUT\_OF\_EUC(NON\_ENERGY))$	[-]	Share of non-energy demand per type of feedstock
$t_{op}(h, td)$	[h]	Duration of each time period (default 1h)
$f_{min}, f_{max}(tech)$	[GW] [a] [b]	Min./max. installed size of each technology
$f_{min,\%}, f_{max,\%}(tech)$	[-]	Min./max. relative share of each technology in a layer
$avail(res)$	[GWh/y]	Yearly total availability of each resource
$c_{op}(res)$	[M USD/GWh]	Specific cost of resource
veh_{capa}	[km-pass/h/veh.] [a]	Mobility capacity per vehicle (veh.)
$\%_{Peak_{sh}}$	[-]	Ratio between highest yearly demand and highest demand in TDs for space heating
$\%_{Peak_{sc}}$	[-]	Ratio between highest yearly demand and highest demand in TDs for space cooling
$f(res \cup tech \setminus sto, l)$	[GW] [c]	Input from (<0) or output to (>0) layers . f(i,j) = 1 if j is main output layer for technology/resource i.
$c_{inv}(tech)$	[M USD/GW] [c] [b]	Technology specific investment cost
$c_{maint}(tech)$	[M USD /GW/y] [c] [b]	Technology specific yearly maintenance cost
$lifetime(tech)$	[y]	Technology lifetime
$gwp_{constr}(tech)$	[ktCO$_2$-eq./GW] [a] [b]	Technology construction specific GHG emissions
$gwp_{op}(res)$	[ktCO$_2$-eq./GWh]	Specific GHG emissions of resources

Parameter	Units	Description
$c_p(tech)$	[-]	Yearly mean capacity factor
$\eta_{sto,in}, \eta_{sto,out}(sto, l)$	[-]	Efficiency [0;1] of storage input from/ output to layer. Set to 0 if storage not related to layer
$\%_{sto_{loss}}(sto)$	[1/h]	Losses in storage (self-discharge)
$t_{sto_{in}}(sto)$	[-]	Time to fully charge storage (energy to power ratio)
$t_{sto_{out}}(sto)$	[-]	Time to fully discharge storage (energy to power ratio)
$\%_{sto_{avail}}(sto)$	[-]	Storage technology availability for charge/discharge
$\%_{net_{loss}}(eut)$	[-]	Losses coefficient $[0; 1]$ in the networks (grid and DHN)
$ev_{batt,size}(v2g)$	[GWh]	Battery size of each V2G car technology
$soc_{min,ev}(v2g, h)$	[GWh]	Minimum state of charge for electric vehicles
$\%_{max,motorcycle}$	[GWh]	Maximum share of motorcycles in private mobility
$c_{grid,extra}$	[M USD/GW	Cost to reinforce the grid per GW of installed intermittent renewable
$elec_{import,max}$	[GW]	Maximum import capacity for electricity
$elec_{export,max}$	[GW]	Maximum export capacity for electricity
$solar_{area,rooftop}$	[km^2]	Available area for solar technologies on rooftops
$solar_{area,ground}$	[km^2]	Available land area for solar technologies on the ground

Parameter	Units	Description
$solar_{area,ground,highirr}$	[km^2]	Available land area with high irradiation for solar technologies on the ground
sm_{max}	[-]	Maximum solar multiple for CSP plants
$power\_density_{pv}$	[GW/km^2]	Maximum power irradiance for PV
$power\_density_{csp}$	[GW/km^2]	Maximum power irradiance for CSP
$power\_density_{solar\ thermal}$	[GW/km^2]	Maximum power irradiance for solar thermal

Table 3: Exhaustive list of dependent variables used in ESTD v2.2. All variables are continuous and non-negative, unless otherwise indicated.

Variable	**Units**	**Description**
$\mathbf{EndUses}(l, h, td)$	[GW] [e]	EUD. Set to 0 if $l \notin EUT$
$\mathbf{C_{tot}}$	[M USD/y]	Total annual cost of the energy system
$\mathbf{C_{inv}}(tech)$	[M USD]	Total investment cost of technology
$\mathbf{C_{maint}}(tech)$	[M USD/y]	Yearly maintenance cost of technology
$\mathbf{C_{op}}(res)$	[M USD/y]	Total cost of resource
$\mathbf{GWP_{tot}}$	[ktCO_2-eq./y]	Total yearly GHG emissions of the energy system
$\mathbf{GWP_{constr}}(tech)$	[ktCO_2-eq.]	GHG emissions during construction of technology
$\mathbf{GWP_{po}}(res)$	[ktCO_2-eq./y]	Total GHG emissions of resource
$\mathbf{Net_{losses}}(eut, h, td)$	[GW]	Losses in the networks (grid and DHN)
$\mathbf{Sto_{level}}(sto, t)$	[GWh]	Energy stored over the year
$\mathbf{ImportConstant}(RES\ IMPORT\ CONSTANT)$	[GWh]	Constant value of import over the year

Variable	Units	Description
$\mathbf{Export}_{\mathbf{constant}}(EXPORT\_EFUEL)$	[GWh]	Constant value of export of each e-fuel over the year

Energy model formulation

In the following, sub-sections, the overall LP formulation is proposed through Figure %s <fig:EndUseDemand> and equations eq:obj_func - eq:solarAreaLandLimited. The first constraints presented relate to the computation of the EUDs. Then, the cost, the global warming potential (GWP) and the objective functions are introduced. The sub-sections coming after are more specific, describing for example the implementations of *storage* or *vehicle-to-grid*.

End-use demand

Giving as input to the model the EUD instead of the FEC has two advantages. First, it introduces a clear distinction between demand and supply. On the one hand, the demand concerns end-uses (e.g. mobility needs). On the other hand, the supply concerns the choice of the energy conversion technologies to supply these services (e.g. the types of vehicles used to satisfy the mobility needs). Based on technology choice, the same EUD can be satisfied with different FECs. Second, using the EUD facilitates the inclusion in the model of electric technologies for heating and transportation.

The hourly EUDs (**EndUses**) are computed based on the yearly EUDs (*endUsesInput*), distributed according to the time series listed in Table %s <tab:paramsDistributions>. Figure %s <fig:EndUseDemand> graphically presents the constraints associated to the hourly EUDs (**EndUses**). For example, the public mobility demand at time t is equal to the hourly passenger mobility demand multiplied by the public mobility share ($\mathbf{\%_{Public}}$).

Electricity EUD results from the sum of the electricity-only demand, assumed constant throughout the year, and the variable demand for electricity, distributed across the periods according to $\%_{elec}$. Low-temperature heat demand results from the sum of the demand for hot water (HW), evenly shared across the year, and the demand for space heating (SH), distributed across the periods according to $\%_{sh}$. The percentage repartition between centralized (DHN) and decentralized heat demand is defined by the variable $\mathbf{\%_{Dhn}}$. High temperature heat for industrial processes is evenly distributed across the periods. Passenger mobility demand is expressed in passenger-kilometers (pkms), while freight demand is in ton-kilometers (tkms). The variable $\mathbf{\%_{Public}}$ defines the penetration of public transportation in passenger mobility. Similarly, $\mathbf{\%_{Rail}}$, $\mathbf{\%_{Boat}}$ and $\mathbf{\%_{Truck}}$ define the penetration of train, boat and trucks for freight mobility, respectively.

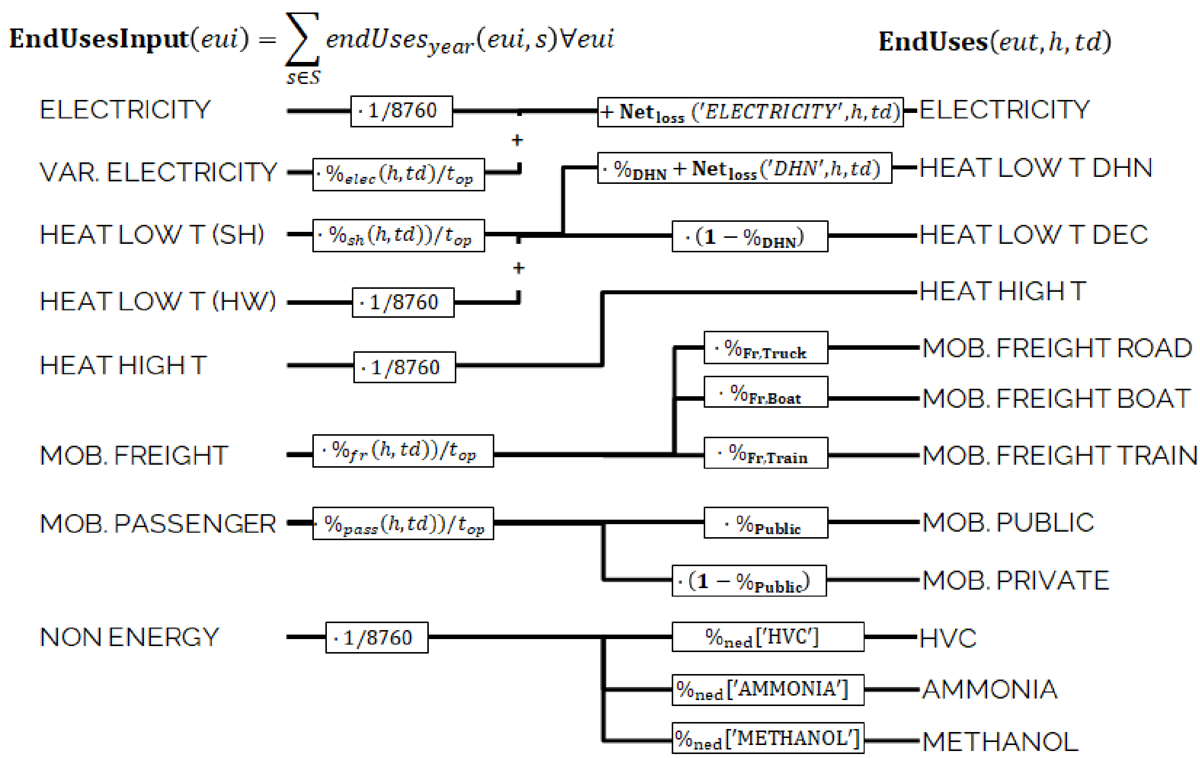

Figure 8: Hourly **EndUses** demands calculation, starting from yearly demand inputs (*endUsesInput*). Adapted from Moret2017PhDThesis. This figure was produced for ESTD v2.1. It does not show the latest EUDs related to cooling and included in v2.2. Abbreviations: space heating (sh), district heating network (DHN), high value chemicals (HVC), hot water (HW), passenger (pass), freight (fr) and non-energy demand (NED).

Space cooling demands were added in ESTD v2.2 and are not represented on `Figure %s <fig:EndUseDemand>`. The demand for space cooling (SC) is distributed across the periods according to $\%_{sc}$, while the cooling demand for industrial processes is uniform.

Cost, emissions and objective function

$$\min \mathbf{C_{tot}} = \sum_{j \in \text{TECH}} \Big(\tau(j)\mathbf{C_{inv}}(j) + \mathbf{C_{maint}}(j)\Big) + \sum_{i \in \text{RES}} \mathbf{C_{op}}(i)$$

$$\text{s.t. } \tau(j) = \frac{i_{\text{rate}}(i_{\text{rate}}+1)^{lifetime(j)}}{(i_{\text{rate}}+1)^{lifetime(j)} - 1} \qquad \forall j \in \text{TECH}$$

$$\mathbf{C_{inv}}(j) = c_{\text{inv}}(j)\mathbf{F}(j) \qquad \forall j \in \text{TECH}$$

$$\mathbf{C_{maint}}(j) = c_{\text{maint}}(j)\mathbf{F}(j) \qquad \forall j \in \text{TECH}$$

$$\mathbf{C_{op}}(i) = \sum_{t \in T | \{h,td\} \in T\_H\_TD(t)} c_{\text{op}}(i)\mathbf{F_t}(i,h,td)t_{op}(h,td) \qquad \forall i \in \text{RES}$$

The objective function to minimize is given in Eq. `eq:obj_func`. It is the the total annual cost of the energy system ($\mathbf{C_{tot}}$), defined as the sum of the annualized investment cost of the technologies ($\tau\mathbf{C_{inv}}$), the operating and maintenance costs of the technologies ($\mathbf{C_{maint}}$) and the operating cost of the resources ($\mathbf{C_{op}}$). The total investment cost ($\mathbf{C_{inv}}$) of each technology results from the multiplication of its specific investment cost (c_{inv}) by its installed capacity ($\mathbf{F}$) (Eq. `eq:c_inv`), the latter being defined with respect to the main end-use output type[3]. $\mathbf{C_{inv}}$ is annualised using the factor τ, calculated based on the interest rate (t_{op}) and the technology lifetime (*lifetime*) in Eq. `eq:tau`. The total operation and maintenance cost is calculated in the same way in Eq. `eq:c_maint`. In Eq. `eq:c_op`, the total cost of the resources is calculated as the sum of the end-use over different periods multiplied by the periods' duration (t_{op}) and the specific cost of the resource (c_{op}). Note that in Eq. `eq:c_op`, summing over the typical days using the set TH_TD[4] is equivalent to summing over the 8760h of the year.

[3] Indeed, some technologies have several outputs, such as a CHP. Thus, the installed size must be defined with respect to one of these outputs. For example, CHP are defined based on the thermal output rather than the electrical one.

[4] To simplify the reading, the formulation $t \in T|\{h,td\} \in T\_H\_TD(t)$ is used. However, this cannot be directly implemented in the code and it requires two additional sets : $HOUR\_OF\_PERIOD(t)$ and $TYPICAL\_DAY\_OF\_PERIOD(t)$. Hence, we have: $t \in T|\{h,td\} \in T\_H\_TD(t)$, which is equivalent in the code to $t \in T|h \in HOUR\_OF\_PERIOD(t), td \in TYPICAL\_DAY\_OF\_PERIOD(t)$.

$$\mathbf{GWP_{tot}} = \sum_{j \in \text{TECH}} \frac{\mathbf{GWP_{constr}}(j)}{lifetime(j)} + \sum_{i \in \text{RES}} \mathbf{GWP_{op}}(i)$$

$$\left(\text{in this version of the model : } \mathbf{GWP_{tot}} = \sum_{i \in \text{RES}} \mathbf{GWP_{op}}(i)\right)$$

$$\mathbf{GWP_{constr}}(j) = gwp_{\text{constr}}(j)\mathbf{F}(j) \qquad \forall j \in \text{TECH}$$

$$\mathbf{GWP_{op}}(i) = \sum_{t \in T | \{h,td\} \in T\_H\_TD(t)} gwp_{\text{op}}(i)\mathbf{F_t}(i,h,td)t_{op}(h,td) \qquad \forall i \in \text{RES}$$

The global annual GHG emissions are calculated using a life-cycle assessment (LCA) approach, i.e. taking into account emissions of the technologies and resources '*from cradle to grave*'. For climate change, the natural choice as indicator is the global warming potential (GWP), expressed in ktCO$_2$-eq./year. In Eq. `eq:GWP_tot`, the total yearly emissions of the system ($\mathbf{GWP_{tot}}$) are defined as the sum of the emissions related to the construction and end-of-life of the energy conversion technologies ($\mathbf{GWP_{constr}}$), annualized based on the technology lifetime ($lifetime$), and the emissions related to resources ($\mathbf{GWP_{op}}$). Similarly to the costs, the total emissions related to the construction of technologies are computed in Eq. `eq:GWP_constr` as the product of the specific emissions (gwp_{constr}) by the installed capacity ($\mathbf{F}$). In Eq. `eq:GWP_op`, the total emissions of the resources are computed as the emissions associated to fuels from cradle to combustion and imports of electricity (gwp_{op}), multiplied by the period duration (t_{op}). GWP accounting can be conducted in different manners depending on the choice of scope. The European Commission and the IEA mainly use resource-related emissions ($\mathbf{GWP_{op}}$) while neglecting indirect emissions related to the construction of technologies ($\mathbf{GWP_{constr}}$). To facilitate the comparison with their results, a similar implementation is proposed in Eq. `eq:GWP_tot`.

System design and operation

$$f_{\min}(j) \leq \mathbf{F}(j) \leq f_{\max}(j) \qquad \forall j \in \text{TECH}$$

Eq. `eq:fmin_fmax` imposes that the installed capacity of a technology ($\mathbf{F}$) is constrained by upper and lower bounds (f_{max} and f_{min}). This formulation allows for accounting for old technologies still existing in the target year (lower bound), but also for the maximum deployment potential of a technology. For example, regarding offshore wind turbines, (f_{min}) represents the existing installed capacity (which will still be available in the near future), while (f_{max}) represents the maximum potential.

$$\mathbf{F_t}(i,h,td) \leq \mathbf{F_t}(i) \cdot c_{p,t}(i,h,td) \qquad \forall i \in \text{TECH}, h \in H, td \in TD$$

$$\sum_{t \in T | \{h,td\} \in T\_H\_TD(t)} \mathbf{F_t}(j,h,td) t_{op}(h,td) \leq \mathbf{F}(j) c_p(j) \sum_{t \in T | \{h,td\} \in T\_H\_TD(t)} t_{op}(h,td)$$
$$\forall j \in \text{TECH}$$

$$\sum_{t \in T | \{h,td\} \in T\_H\_TD(t)} \mathbf{F_t}(i,h,td) t_{op}(h,td) \leq \text{avail}(i) \qquad \forall i \in \text{RES}$$

The operation of resources and technologies in each period is determined by the decision variable $\mathbf{F_t}$. The capacity factor of technologies is conceptually divided into two components: a capacity factor for each period ($c_{p,t}$) depending on resource availability (e.g. renewables) and a yearly capacity factor (c_p) accounting for technology downtime and maintenance. For a given technology, the definition of only one of these two is needed, the other one being fixed to the default value of 1. For example, intermittent renewables are constrained by an hourly load factor ($c_{p,t} \in [0;1]$) while CCGTs are constrained by an annual load factor (c_p) (with a value in that case of 96% in 2035). Eqs. `eq:cp_t` and `eq:c_p` link the installed size of a technology to its actual use in each period ($\mathbf{F_t}$) via the two capacity factors. The total use of resources is limited by the yearly availability (*avail*) in Eq. `eq:res_avail`.

$$\sum_{i \in \text{RES } \cup \text{TECH} \setminus \text{STO}} f(i,l) \mathbf{F_t}(i,h,td) + \sum_{j \in \text{STO}} \Big(\mathbf{Sto_{out}}(j,l,h,td) - \mathbf{Sto_{in}}(j,l,h,td) \Big)$$
$$- \mathbf{EndUses}(l,h,td) = 0$$
$$\forall l \in L, \forall h \in H, \forall td \in TD$$

The matrix f defines, for all technologies and resources, their output layers (positive) and input layers (negative). Eq. `eq:layer_balance` expresses the balance for each layer: all outputs from resources and technologies (including storage) are used to satisfy the EUDs or as inputs to other resources and technologies.

Storage

$$\mathbf{Sto_{level}}(j,t) = \mathbf{Sto_{level}}(j,t-1) \cdot (1 - \%_{sto_{loss}}(j))$$
$$+ t_{op}(h,td) \cdot \Big(\sum_{l \in L | \eta_{\text{sto,in}(j,l)} > 0} \mathbf{Sto_{in}}(j,l,h,td) \eta_{\text{sto,in}}(j,l)$$

$$- \sum_{l \in L | \eta_{\text{sto,out}(j,l)>0}} \mathbf{Sto_{out}}(j,l,h,td) / \eta_{\text{sto,out}}(j,l) \Big)$$

$$\forall j \in \text{STO}, \forall t \in \text{T} | \{h, td\} \in T\_H\_TD(t)$$

$$\mathbf{Sto_{level}}(j,t) = \mathbf{F_t}(j,h,td) \qquad \forall j \in \text{STO DAILY}, \forall t \in \text{T} | \{h, td\} \in T\_H\_TD(t)$$

$$\mathbf{Sto_{level}}(j,t) \leq \mathbf{F}(j) \qquad \forall j \in \text{STO} \setminus \text{STO DAILY}, \forall t \in \text{T}$$

In Eq. `eq:sto_level`, the storage level ($\mathbf{Sto_{level}}$) at time step t is defined as the storage level at $t-1$ (accounting for the losses in $t-1$), plus the inputs to the storage, minus the output from the storage (accounting for input/output efficiencies). The storage systems which can only be used for short-term (daily) applications are included in the daily storage set (STO DAILY). For these units, Eq. `eq:Sto_level_bound_DAILY`: imposes that the storage level be the same at the end of each typical day[5]. Adding this constraint drastically reduces the computational time. For the other storage technologies, which can also be used for seasonal storage, the capacity is bounded by Eq. `eq:Sto_level_bound`. For these units, the storage behaviour is thus optimized over 8760 hours.

$$\mathbf{Sto_{in}}(j,l,h,td) \cdot \Big(\lceil \eta_{sto,in}(j,l) \rceil - 1 \Big) = 0 \qquad \forall j \in \text{STO}, \forall l \in \text{L}, \forall h \in \text{H}, \forall td \in \text{TD}$$

$$\mathbf{Sto_{out}}(j,l,h,td) \cdot \Big(\lceil \eta_{sto,out}(j,l) \rceil - 1 \Big) = 0 \qquad \forall j \in \text{STO}, \forall l \in \text{L}, \forall h \in \text{H}, \forall td \in \text{TD}$$

$$\Big(\mathbf{Sto_{in}}(j,l,h,td) t_{sto_{in}}(\text{j}) + \mathbf{Sto_{out}}(j,l,h,td) t_{sto_{out}}(\text{j}) \Big) \leq \mathbf{F}(j) \%_{sto_{avail}}(j)$$

$$\forall j \in STO \setminus V2G, \forall l \in L, \forall h \in H, \forall td \in TD$$

Eqs. `eq:StoInCeil` - `eq:StoOutCeil` force the power input and output to zero if the layer is incompatible[6]. For example, a PHS will only be linked to the electricity layer (input/output efficiencies > 0). All other efficiencies will be equal to 0, to impede that the PHS exchanges with incompatible layers (e.g. mobility,

[5] In most cases, the activation of the constraint stated in Eq. `eq:sto_level` will have as a consequence that the level of storage be the same at the beginning and at the end of each day — hence the use of the terminology '*daily storage*'. Note, however, that such daily storage behaviour is not always guaranteed by this constraint and thus, depending on the typical days sequence, a daily storage behaviour might need to be explicitly enforced.

[6] In the code, these equations are implemented with a *if-then* statement.

heat, etc). Eq. `eq:LimitChargeAndDischarge` limits the power input/output of a storage technology based on its installed capacity ($\mathbf{F}$) and three specific characteristics. First, storage availability ($\%_{sto_{avail}}$) is defined as the ratio between the available storage capacity and the total installed capacity (default value is 100%). This parameter is only used to realistically represent V2G, for which we assume that only a fraction of the fleet (i.e. 20% in these cases) can charge/discharge at the same time. Second and third, the charging/discharging time ($t_{sto_{in}}$, $t_{sto_{out}}$), which are the time to complete a full charge/discharge from empty/full storage[7]. For example, a daily thermal storage needs at least 4 hours to discharge ($t_{sto_{out}} = 4$[h]), and another 4 hours to charge ($t_{sto_{in}} = 4$[h]). Eq. `eq:LimitChargeAndDischarge` applies for all storage except electric vehicles which are limited by another constraint Eq. `eq:LimitChargeAndDischarge_ev`, presented later.

Networks

$$\mathbf{Net_{loss}}(eut,h,td) = \Big(\sum_{i \in \text{RES} \cup \text{TECH} \setminus \text{STO} | f(i,eut)>0} f(i,eut)\mathbf{F_t}(i,h,td)\Big)\%_{\text{net}_{loss}}(eut)$$
$$\forall eut = \text{EUT}, \forall h \in H, \forall td \in TD$$

$$\mathbf{F}(Grid) = 1 + \frac{c_{grid,extra}}{c_{inv}(Grid)}\Big(\mathbf{F}(Wind_{onshore}) + \mathbf{F}(Wind_{offshore}) + \mathbf{F}(PV)$$
$$-\big(f_{min}(Wind_{onshore}) + f_{min}(Wind_{offshore}) + f_{min}(PV)\big)\Big)$$

$$\mathbf{F}(DHN) = \sum_{j \in \text{TECH} \setminus STO | f(j,\text{HeatLowTDHN})>0} f(j,\text{HeatLowTDHN}) \cdot \mathbf{F}(j)$$

$$\mathbf{F}(H_{2,infrastructure}) = \mathbf{F}(H_{2,electrolysis})$$

$$\mathbf{F}(ChargingStations) = \mathbf{F}(CAR_{BEV})$$

Eq. `eq:loss` calculates network losses as a share ($\%_{\text{net}_{loss}}$) of the total energy transferred through the network. As an example, losses in the electricity grid are

[7] In this linear formulation, storage technologies can charge and discharge at the same time. On the one hand, this avoids the need of integer variables; on the other hand, it has no physical meaning. However, in a cost minimization problem, the cheapest solution identified by the solver will always choose to either charge or discharge at any given t, as long as cost and efficiencies are defined. Hence, we recommend to always verify numerically the fact that only storage inputs or outputs are activated at each t, as we do in all our implementations.

estimated to be 4.5% of the energy transferred in 2015[8]. Eqs. `eq:mult_grid` - `eq:DHNCost` define the extra investment for networks. Integration of intermittent RE implies additional investment costs for the electricity grid ($c_{grid,ewtra}$). For example, the reinforcement of the electricity grid is estimated to be 518.5 million USD_{2021} per Gigawatt of intermittent renewable capacity installed (see Data for the grid for details). Eq. `eq:DHNCost` links the size of DHN to the total size of the installed centralized energy conversion technologies. Finally, Eq. `eq:h2_network` links the size of the hydrogen network to the installed capacity for hydrogen production and `eq:charging_stations` does the same for the infrastructure of charging stations and electrical vehicles.

Additional Constraints

$$\mathbf{F_t}(Nuclear, h, td) = \mathbf{P_{Nuclear}} \qquad \forall h \in H, \forall td \in TD$$

Nuclear power plants are assumed to have no power variation over the year, Eq. `eq:CstNuke`. If needed, this equation can be replicated for all other technologies for which a constant operation over the year is desired.

$$\mathbf{F_t}(j, h, td) = \mathbf{\%_{PassMob}}(j) \sum_{l \in EUT\_of\_EUC(PassMob)} \mathbf{EndUses}(l, h, td)$$
$$\forall j \in TECH\_OF\_EUC(PassMob), \forall h \in H, \forall td \in TD$$

$$\mathbf{F_t}(j, h, td) = \mathbf{\%_{FreightMob}}(j) \sum_{l \in EUT\_of\_EUC(FreightMob)} \mathbf{EndUses}(l, h, td)$$
$$\forall j \in TECH\_OF\_EUC(FreightMob), \forall h \in H, \forall td \in TD$$

$$\mathbf{\%_{Fr,Rail}} + \mathbf{\%_{Fr,Train}} + \mathbf{\%_{Fr,Boat}} = 1$$

Eqs. `eq:mob_share_fix - eq:freight_share_fix` impose that the share of the different technologies for mobility ($\mathbf{\%_{PassMob}}$) and ($\mathbf{\%_{Freight}}$) be the same at each time step[9]. In other words, if 20% of the mobility is supplied by train, this share remains constant in the morning or the afternoon. Eq. `eq:freight_share_constant` verifies that the freight technologies supply the overall freight demand (this constraint is related to `Figure %s <fig:EndUseDemand>`).

[8] This is the ratio between the losses in the grid and the total annual electricity production in Belgium in 2015 Eurostat2017.

[9] [foot:nonLinear]All equations expressed in a compact non-linear form in this section Eqs. `eq:mob_share_fix`, `eq:freight_share_fix`, `eq:heat_decen_share` and `eq:dhn_peak` can be linearised. For these cases, the **EndUses** is defined with parameters and a variable representing a constant share over the year (e.g. $\mathbf{\%_{public}}$). As an example, **EndUses** in Eq. `eq:mob_share_fix` is equal to $\mathbf{EndUsesInput}(PassMb) \cdot \%pass(h, td)/t_op(h, td)$. The term $\mathbf{\%_{public}}$, is missing in the equation, but is implicitly implemented in $\mathbf{\%_{PassMob}}$.

Decentralised heat production

$$\mathbf{F}(Dec_{Solar}) = \sum_{j \in \text{TECH OF EUT(HeatLowTDec)} \setminus \{Dec_{Solar}\}} \mathbf{F_{sol}}(j)$$

$$\mathbf{F_{t_{sol}}}(j,h,td) \leq \mathbf{F_{sol}}(j) c_{p,t}(Dec_{Solar},h,td)$$
$$\forall j \in \text{TECH OF EUT(HeatLowTDec)} \setminus \{Dec_{Solar}\}, \forall h \in H, \forall td \in TD$$

Thermal solar is implemented as a decentralized technology. It is always installed together with another decentralized technology, which serves as backup to compensate for the intermittency of solar thermal. Thus, we define the total installed capacity of solar thermal $\mathbf{F}(Dec_{solar})$ as the sum of $\mathbf{F_{sol}}(j)$, Eq. `eq:de_strategy_dec_total_ST`, where $\mathbf{F_{sol}}(j)$ is the solar thermal capacity associated to the backup technology j. Eq. `eq:op_strategy_dec_total_ST` links the installed size of each solar thermal capacity $\mathbf{F_{sol}}(j)$ to its actual production `\textbf{F}_{\textbf{t}_\textbf{sol}}(j,h,td)` via the solar capacity factor $(c_{solar_{area,rooftop}p,t}(Dec_{solar}))$.

$$\begin{aligned}
&\mathbf{F_t}(j,h,td) + \mathbf{F_{t_{sol}}}(j,h,td) \\
&+ \sum_{l \in \mathrm{L}} \Big(\mathbf{Sto_{out}}(i,l,h,td) - \mathbf{Sto_{in}}(i,l,h,td) \Big) \\
&= \mathbf{\%_{HeatDec}}(\mathrm{j}) \mathbf{EndUses}(HeatLowT,h,td) \\
&\forall j \in \text{TECH OF EUT(HeatLowTDec)} \setminus \{Dec_{Solar}\}, \\
&i \in \text{TS OF DEC TECH}(j), \forall h \in H, \forall td \in TD
\end{aligned}$$

A thermal storage i is defined for each decentralised heating technology j, to which it is related via the set *TS OF DEC TECH*, i.e. i=*TS OF DEC TECH(j)*. Each thermal storage i can store heat from its technology j and the associated thermal solar $\mathbf{F_{sol}}$ (j). Similarly to the passenger mobility, Eq. `eq:heat_decen_share` makes the model more realistic by defining the operating strategy for decentralized heating. In fact, in the model we represent decentralized heat in an aggregated form; however, in a real case, residential heat cannot be aggregated. A house heated by a decentralised gas boiler and solar thermal panels should not be able to be heated by the electrical heat pump and thermal storage of the neighbours, and vice-versa. Hence, Eq. `eq:heat_decen_share` imposes that the use of each technology $(\mathbf{F_t}(j,h,td))$, plus its associated thermal solar $(\mathbf{F_{t_{sol}}}(j,h,td))$ plus its associated storage outputs $(\mathbf{Sto_{out}}(i,l,h,td))$ minus its associated storage inputs $(\mathbf{Sto_{in}}(i,l,h,td))$ should be a constant share $(\mathbf{\%_{HeatDec}}(j))$ of the decentralised heat demand $(\mathbf{EndUses}(HeatLowT,h,td))$. `Figure %s <fig:FsolAndTSImplementation>` shows, through an example with two technologies (a gas boiler and a HP), how decentralised thermal storage and thermal solar are implemented.

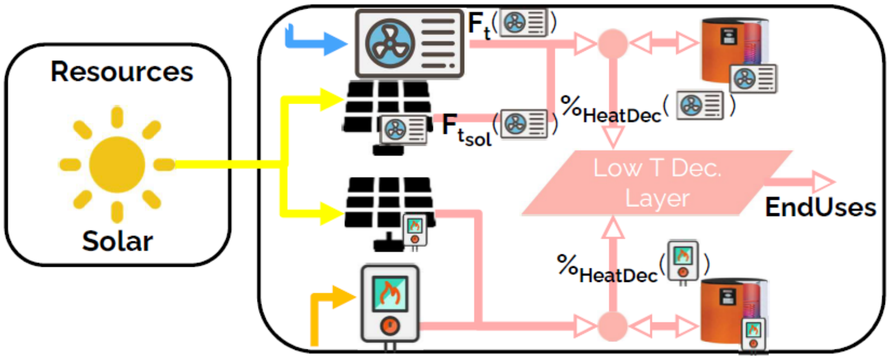

Figure 9: Illustrative example of a decentralised heating layer with thermal storage, solar thermal and two conventional production technologies, gas boilers and electrical HP. In this case, Eq. `eq:heat_decen_share` applied to the electrical HPs becomes the equality between the two following terms: left term is the heat produced by: the eHPs ($\mathbf{F_t}(eHPs, h, td)$), the solar panel associated to the eHPs ($\mathbf{F_{t_{sol}}}(eHPs, h, td)$) and the storage associated to the eHPs; right term is the product between the share of decentralised heat supplied by eHPs ($\boldsymbol{\%_{HeatDec}}(eHPs)$) and heat low temperature decentralised demand ($\mathbf{EndUses}(HeatLowT, h, td)$).

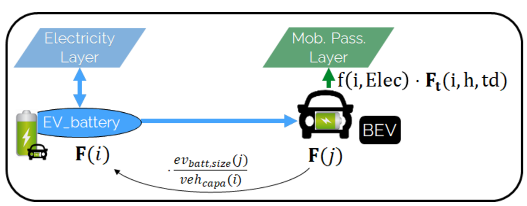

Figure 10: Illustrative example of a V2G implementation. The battery can interact with the electricity layer. The size of the battery is directly related to the number of cars (see Eq. `eq:SizeOfBEV`). The V2G takes the electricity from the battery to provide a constant share ($\boldsymbol{\%_{PassMob}}$) of the passenger mobility layer (*Mob. Pass.*). Thus, it imposes the amount of electricity that electric car must deserve (see Eq. `eq:BtoBEV`). The remaining capacity of battery available can be used to provide V2G services (see `eq:LimitChargeAndDischarge_ev`).

Vehicle-to-grid

$$\mathbf{F}(i) = \frac{\mathbf{F}(j)}{veh_{capa}(j)} ev_{batt,size}(j) \quad \forall j \in V2G, i \in \text{EVs}_B ATTOFV2G\text{EVs}_B ATTOFV2G\text{EVs}_B ATTOFV2G\text{EVs}_B ATTO$$

Vehicle-to-grid dynamics are included in the model via the *V2G* set. For each vehicle $j \in V2G$, a battery i ($i \in EVs\_BATT$) is associated using the set EVsBATT_OF_V2G ($i \in$

$$\mathbf{F}(j) \geq \%_{Peak_{sh}} \max_{h \in H, td \in TD} \{\mathbf{F_t}(j,h,td)\}$$

$$\forall j \in \text{TECH OF EUT}(HeatLowTDEC) \setminus \{Dec_{Solar}\}$$

$$\sum_{j \in \text{TECH OF EUT}(HeatLowTDHN), i \in \text{STO OF EUT}(HeatLowTDHN)}$$

$$\Big(\mathbf{F}(j) + \mathbf{F}(i)/t_{sto_{out}}(i, HeatLowTDHN)\Big)$$

$$\geq \%_{Peak_{sh}} \max_{h \in H, td \in TD} \big\{\mathbf{EndUses}(HeatLowTDHN, h, td)\big\}$$

$$\mathbf{F}(j) \geq \%_{Peak_{sc}} \max_{h \in H, td \in TD} \{\mathbf{F_t}(j,h,td)\}$$

$$\forall j \in \text{TECH OF EUT}(SpaceCooling)$$

Finally, Eqs. `eq:dec_peak` - `eq:dhn_peak` constrain the installed capacity of low temperature heat supply. Based on the selected TDs, the ratio between the yearly peak demand and the TDs peak demand is defined for space heating ($\%_{Peak_{sh}}$). Eq. `eq:dec_peak` imposes that the installed capacity for decentralised technologies covers the real peak over the year. Similarly, Eq. `eq:dhn_peak` forces the centralised heating system to have a supply capacity (production plus storage) higher than the peak demand. These equations force the installed capacity to meet the peak heating demand, i.e. which represents, somehow, the network adequacy[10]. Similarly to `eq:dec_peak`, `eq:sc_peak` imposes that the installed capacity for space cooling technologies covers the real peak cooling demand over the year.

[10]The model resolution of the dispatch is not accurate enough to verify the adequacy. As one model cannot address all the issues, another approach has been preferred: couple the model to a dispatch one, and iterate between them. Percy and Coates percy_coates_coupling_2020 demonstrated the feasibility of coupling a design model (ESTD) with a dispatch one (Dispa-SET Quoilin2017). Based on a feedback loop, they iterated on the design to verify the power grid adequacy and the strategic reserves. Results show that the backup capacities and storage needed to be slightly increased compared to the results of the design model alone.

Adaptations for the case study

Additional constraints are required to implement scenarios. Scenarios require six additional constraints (Eqs. `eq:LimitGWP` - `eq:solarAreaLandLimited`) to impose a limit on the GWP emissions, the minimum share of RE primary energy, the relative shares of technologies, such as gasoline cars in the private mobility, the cost of energy efficiency measures, the electricity import power capacity and the available surface area for solar technologies.

$$\mathbf{GWP_{tot}} \leq gwp_{limit}$$

$$\begin{aligned}
&\sum_{j \in \text{RES}_{\text{re}}, t \in T | \{h,td\} \in T_H_TD(t)} \mathbf{F_t}(j,h,td) \cdot t_{op}(h,td) \\
&\geq re_{share} \sum_{j \in \text{RES}, t \in T | \{h,td\} \in T_H_TD(t)} \mathbf{F_t}(j,h,td) \cdot t_{op}(h,td)
\end{aligned}$$

To force the energy system to decrease its emissions, two lever can constraint the annual emissions: Eq. `eq:LimitGWP` imposes a maximum yearly emissions threshold on the GWP (gwp_{limit}); and Eq. `eq:LimitRE` fixes the minimum renewable primary energy share.

$$\begin{aligned}
&f_{\min,\%}(j) \sum_{j' \in \text{TECH OF EUT}(eut), t \in T | \{h,td\} \in T_H_TD(t)} \mathbf{F_t}(j',h,td) \cdot t_{op}(h,td) \\
&\leq \sum_{t \in T | \{h,td\} \in T_H_TD(t)} \mathbf{F_t}(j,h,td) \cdot t_{op}(h,td) \\
&\leq f_{\max,\%}(j) \sum_{j'' \in \text{TECH OF EUT}(eut), t \in T | \{h,td\} \in T_H_TD(t)} \mathbf{F_t}(j'',h,td) \cdot t_{op}(h,td) \\
&\forall eut \in EUT, \forall j \in \text{TECH OF EUT}(eut)
\end{aligned}$$

To represent the national energy system under study, Eq. `eq:fmin_max_perc` imposes the relative share of a technology in its sector. Eq. `eq:fmin_max_perc` is complementary to Eq. `eq:fmin_fmax`, as it expresses the minimum ($f_{min,\%}$) and maximum ($f_{max,\%}$) yearly output shares of each technology for each type of EUD. In fact, for a given technology, assigning a relative share (e.g. boilers providing at least a given percentage of the total heat demand) is more intuitive and close to the energy planning practice than limiting its installed size. $f_{min,\%}$ and $f_{max,\%}$ are fixed to 0 and 1, respectively, unless otherwise indicated.

$$\sum_{t \in T | \{h,td\} \in T\_H\_TD(t)} \big(\mathbf{F_t}(Motorcycle, h, td) + \mathbf{F_t}(Motorcycle\ Electric, h, td)\big) \cdot t_{op}(h, td)$$
$$\leq \%_{max,motorcycle} \sum_{j \in \text{TECH OF EUT}(Mob\ Private), t \in T | \{h,td\} \in T\_H\_TD(t)} \mathbf{F_t}(j, h, td) \cdot t_{op}(h, td)$$

Similarly to eq. `eq:fmin_max_perc`, eq. `eq:f_max_perc_motorcycle` imposes the maximum share of private passenger mobility that can be supplied by motorcycles.

$$\mathbf{F}(Efficiency) = \frac{1}{1 + i_{rate}}$$

Eq. `eq:efficiency` is supposed to compute the cost of efficiency measures. This equation was used to put a cost on the energy efficiency measures envisaged by the EU Commission for European countries. It is not used for the case of Colombia and Turkey (i.e. F(Efficiency) is multiplied by a zero-cost later on).

Additional contraints on imports, exports and renewables

$$\mathbf{F_t}(Electricity, h, td) \leq elec_{import,max} + \mathbf{F}(HVAC\ Line) \qquad \forall h \in H, \forall td \in TD$$

$$\mathbf{F_t}(Elec\ Export, h, td) \leq elec_{export,max} + \mathbf{F}(HVAC\ Line) \qquad \forall h \in H, \forall td \in TD$$

$$\mathbf{F_t}(i, h, td) \cdot t_{op}(h, td) = \mathbf{Import_{constant}}(i) \quad \forall i \in \text{RES IMPORT CONSTANT}, h \in H, td \in TD$$

$$\mathbf{F_t}(i, h, td) \cdot t_{op}(h, td) = \mathbf{Export_{constant}}(i) \quad \forall i \in \text{EXPORT E FUEL}, h \in H, td \in TD$$

Eqs. `eq:elecImpLimited` and `eq:elecExpLimited` limit the power grid import and export capacity from/to neighbouring countries, based on the 2021 import/export capacity plus the construction of new High-Voltage transfer capacity (HVAC Line). Eq. `eq:import_resources_constant` imposes that some resources are imported at a constant power. For example, gas and hydrogen are supposed to be imported at a constant flow during the year. In addition to offering a more realistic representation, this implementation makes it possible to visualise the level of storage within the region. Eq. `eq:export_efuels_constant` imposes the same constraint as `eq:import_resources_constant`, but for the *export* of e-fuels.

Caution

Adding too many ressource to Eq. `eq:import_resources_constant` increases drastically the computational time. In this implementation, only resources expensive to store have been accounted for i.e. hydrogen and gas. Other resources, such as diesel or ammonia, can be stored at a cheap price with small losses. By limiting EXPORT-E-FUEL to two types of resources (hydrogen and gas), the computation time is below a minute. When adding all imported resources to EXPORT-E-FUEL, the computational time becomes above 6 minutes.

$$\frac{\mathbf{F}(Dam\ Storage) - f_{min}(Dam\ Storage)}{f_{max}(Dam\ Storage) - f_{min}(Dam\ Storage)} \leq \frac{\mathbf{F}(Hydro\ Dam) - f_{min}(Hydro\ Dam)}{f_{max}(Hydro\ Dam) - f_{min}(Hydro\ Dam)}$$

$$\mathbf{Sto_{in}}(Dam\ Storage, Electricity, h, td) = \mathbf{F_t}(Hydro\ Dam, h, td) \quad \forall h \in H, \forall td \in TD$$

$$\mathbf{Sto_{out}}(Dam\ Storage, Electricity, h, td) \leq \mathbf{F}(Hydro\ Dam, h, td) \quad \forall h \in H, \forall td \in TD$$

In EnergyScope, there are two technologies related to hydro-electric dams: *Hydro Dam* and *Dam Storage*. The former relates to the electricity production function of hydro-electric dams, while the second relates to their storage function. These two functions are defined as separate technologies in EnergyScope, but of course they relate to the same physical asset. The constraints defined in eqs `eq:link_dam_storage_to_hydro_dam` - `eq:dam_storage_out` hence bind these two technologies to each other. First, eq. `eq:link_dam_storage_to_hydro_dam` imposes that the installed capacity of both technnologies must remain proportional to each other, in the proportions defined by their respective input parameters f_{min} and f_{max}. (Note that to improve readability, eq. `eq:link_dam_storage_to_hydro_dam` has been written in a non-linear fashion in this documentation. The equation is linear in the actual model's code). Second, eq. `eq:dam_storage_in` imposes that all electricity produced by *Hydro Dam* is immediately absorbed by *Dam Storage*. This electricity is then released by *Dam Storage*, under the constraint that the maximum electricity production of *Dam Storage* is the same one as for *Hydro Dam*.

$$-\frac{\mathbf{F}(ST\ Collector)}{f(ST\ Power\ Block, ST\ Heat)} \leq sm_{max} \cdot \mathbf{F}(ST\ Power\ Block)$$

$$-\frac{\mathbf{F}(PT\ Collector)}{f(PT\ Power\ Block, PT\ Heat)} \leq sm_{max} \cdot \mathbf{F}(PT\ Power\ Block)$$

Concentrated solar power (CSP) technologies are modelled with 3 elements: *Collectors*, *Storage* and *Power Block*. The link between the 3 elements is kept into a "realistic" range thanks to Equations `eq:limit_solar_mulitple_ST` and `eq:limit_solar_mulitple_PT`.

$$\frac{\mathbf{F}(PV\ Rooftop)}{power\_density_{pv}} + \frac{\mathbf{F}(Dec_{Solar}) + \mathbf{F}(DHN_{Solar})}{power\_density_{solar\ thermal}} \leq solar_{area,rooftop}$$

$$\frac{\mathbf{F}(PV\ Utility)}{power\_density_{pv}} - \frac{\mathbf{F}(ST\ Collector)}{f(ST\ Power\ Block, ST\ Heat) \cdot power\_density_{csp}}$$
$$- \frac{\mathbf{F}(PT\ Collector)}{f(PT\ Power\ Block, PT\ Heat) \cdot power\_density_{csp}} \leq solar_{area,ground}$$

$$- \frac{\mathbf{F}(ST\ Collector)}{f(ST\ Power\ Block, ST\ Heat) \cdot power\_density_{csp}}$$
$$- \frac{\mathbf{F}(PT\ Collector)}{f(PT\ Power\ Block, PT\ Heat) \cdot power\_density_{csp}} \leq solar_{area,ground,high\ irr}$$

In this version of EnergyScope, the upper limit for the deployment of solar technologies is calculated based on the available areas ($solar_{area}$) and power densities ($power\_density$) of solar technologies. The conversion factor between an installed capacity (in watt peak (Wp)) and the surface used (in km^2) is calculated based on the peak power density (in [Wp/m^2]). Put simply, the peak power density represents the peak power of one square meter of solar panel. Thus, the land use of a solar technology is its installed capacity ($\mathbf{F}(\cdot)$, in [GW]) divided by its power peak density (in [GW/km^2]). Eq. `eq:solarAreaRooftopLimited` imposes a constraint on the available rooftop area for solar energy. Eq. `eq:solarAreaLandLimited`, does the same for ground area. Finally, eq. `eq:solarAreaGroundHighIrrLimited` proceeds similarly for ground area with high irradiation, suitable for the installation of CSP plants (i.e. with a Direct Normal Irradiation (DNI) superior to 1800 [kWh/m^2/year]). The area limitation is applied on the *Collector* element of the CSP. The capacity of the *Collector* (which produces heat) is then transformed into the equivalent electrical power capacity in eqs `eq:solarAreaLandLimited - eq:solarAreaGroundHighIrrLimited`, using the heat-electricity conversion efficiency of the *Power block*. Indeed, the $power\_density_{csp}$ is expressed in electrical GW, not in thermal GW. Note that in eqs. `eq:solarAreaLandLimited - eq:solarAreaGroundHighIrrLimited`, the terms associated with CSP are counted as positive (the minus signs are present to compensate for the negative signs of $f(\cdot)$).

Implementation

The implementation into code of the MILP and LP problems has been carried out using an algebraic modelling language. Such modelling language allows for the representation of large LP and MILP problems. Its syntax is similar to AMPL, which is -according to the NEOS-statistics[11] - the most popular format for representing mathematical programming problems. The chosen formulation enables the use of different solvers, be it open source ones (e.g. GLPK) or commercial ones (e.g. CPLEX, Gurobi). Each of the equations defined in this documentation is found in identical form in the code, together with the corresponding numbering. SETS, Variables and parameters have the same names (unless explicitly stated in the definition of the term). Figure %s <fig:ch2_LP_formulation_implementation_colored> illustrates, for the balance constraint `eq:layer_balance`, the mathematical formulation presented in this work and its implementation in the code. Colors highlight corresponding elements. In the code, each constraint has a comment (starting with #) and a name (colored in black), in this case *layer_balance*. In addition, most of the SETS, Variables and parameters are more explicitly named in the code. For example, the set "layers" is named L in the documentation and $LAYERS$ in the code; similarly, the input efficiency is named f in the documentation and *layers_in_out* in the code.

$$\sum_{i \in RES \cup TECH \setminus STO} f(i,l)\mathbf{F_t}(i,h,td) + \sum_{j \in STO} \Big(\mathbf{Sto_{out}}(j,l,h,td) - \mathbf{Sto_{in}}(j,l,h,td)\Big) - \mathbf{EndUses}(l,h,td) = 0$$

$$\forall l \in L, \forall h \in H, \forall td \in TD$$

Figure 11:

```
# [Eq. 2.13] [...]
subject to layer_balance {l in LAYERS, h in HOURS, td in TYPICAL_DAYS}:
	sum {i in RESOURCES union TECHNOLOGIES diff STORAGE_TECH }
	(layers_in_out[i, l] * F_t [i, h, td])
	+ sum {j in STORAGE_TECH} ( Storage_out [j, l, h, td] - Storage_in [j, l, h, td] )
	- End_uses [l, h, td]
	= 0;
```

Figure 12: Comparison of equations' formulation in this documentation (upper) and in the code (lower). Example based on Eq. `eq:layer_balance`.

The entire implementation is available on the directory ESTD_v2_1_repo (standard model for Belgium) and ESTD_v2_1_repo2 (model with python wrapper applied to Colombia and Turkey). In the case of the model applied to Colombia and Turkey, the directory contains 4 main folders:

- 'energyscope' contains the code describing the model formulation. In particular, all equations described above are written as is in the sub-folder 'energymodel', in the file 'esmodel.mod'.

[11] NEOS Server is an Internet-based client-server application that provides free access to a library of optimization solvers. Statistics are available at: https://neos-server.org/neos/report.html, consulted the 27/01/2021.

- 'Data' contains several data sets in different sub-folders, each corresponding to a different case study. For example, 'CO 2035' contains the values of all parameters for applying the model to Colombia in the year 2035.
- 'casestudies' contains the ready-to-run AMPL formulation of the model as well as the ouptuts of the runs for each case study, in different sub-folders.
- 'scripts' contains two important files:
  - configref.yaml indicates which case study is chosen, by pointing towards the right subfolders in 'Data' and 'casestudies' and by indicating the chosen limit for GHG emissions.
  - runenergyscope.py is the python file which runs everything.

a Generally [GWh/y], but [Mpkm] (millions of passenger-km) for passenger mobility EUD and [Mtkm] (millions of ton-km) for freight mobility EUD

b GWh instead of GW if $tech \in STO$

c [Mpkm/h] for passenger mobility EUD, [Mtkm/h] for freight mobility EUD

d [Mpkm] (millions of passenger-km) for passenger mobility EUD, [Mtkm] (millions of ton-km) for freight EUD

e [Mpkm] (millions of passenger-km) for passenger mobility EUD, [Mtkm] (millions of ton-km) for freight EUD