Synergistic Approach of Multi-Energy Models for an European Optimal Energy System Management Tool

Deliverable D 3.1
Description of model interconnections

Deliverable No.	D 3.1	Work Package No.	WP3	Task No.	TASK 3.1
Work Package Title	Modelling the system and the decision process				
Linked Task/s Title	Description of model interconnections				
Status	Final	Draft / Draft Final / Final			
Dissemination level	PU	PU-Public / CO-Confidential			
Due date deliverable	2018-04-30	Submission date		2018-04-30	
Deliverable version	`plan4res_D3.1_Description_of_model_interconnections.pdf`				

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 773897

Deliverable Contributors:	Name	Organisation	Date
Deliverable Leader	Daniel Beulertz	RWTH	26.04.2018
Work Package Leader	Daniel Beulertz	RWTH	26.04.2018
Contributing Author(s)	Daniel Beulertz	RWTH	26.04.2018
	Marco Franken	RWTH	29.03.2018
	Nadia Oudjane	EDF	04.04.2018
	Wim van Ackooij	EDF	04.04.2018
	Jonas Schweiger	ZIB	06.04.2018
	Ioannis Konstantelos	Imperial	05.04.2018
	Predrag Djapic	Imperial	29.03.2018
	Danny Pudjianto	Imperial	29.03.2018
Reviewer(s)	Felix Trieu	EDF	18.04.2018
	Antonio Frangioni	ICOOR	16.04.2018
Final Review and Approval	Sandrine Charousset	EDF	30.04.2018

History of Changes:

Release	Date	Reason for Change	Status
0.1	08.12.2017	Draft	in progress
0.2	28.02.2018	Draft	in progress
1.0	29.03.2018	First Draft	Almost finished
2.0	09.04.2018	Second Draft	Internal review accomplished
3.0	26.04.2018	Final draft	External review accomplished
4.0	30.04.2018	Final version	Final review accomplished

List of Figures

List of Tables

List of acronyms used in this document

CEM Capacity expansion model

CHP Combined heat and power

CP Cutting plane

CTS Commercial/trade/service

CWE Central western europe

DER Distributed energy ressources

DG Distribution grid

DSR Demand side response

EUC European unit commitment

LODF Line outage distribution factor

LV Low voltage

NUTS Nomenclature des unités territoriales statistiques

PTDF Power transfer distribution factor

PtX Power-to-X

RES Renewable energy source

SDDP Stochastic dual dynamic programming

SSV Seasonal storage valuation

TGEM Transmission grid expansion model

UC Unit commitment

VC Voltage control

WACC Weighted average cost of capital

DISCLAIMER / ACKNOWLEDGMENT

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 773897

Contents

Executive Summary

The goal of plan4res is to develop a modeling framework that allows to obtain a holistic assessment of the energy system. Having such an ambitious goal, it is required to divide the energy system in models that cover the different aspects of the energy system. This modular framework allows to make use of the most promising solving techniques and the most efficient optimization solvers, each tailored towards the needs of every single sub-model. In order to guarantee a flawless workflow, it is vital to have a detailled description of the interconnections between these models. The goal of this deliverable is to give an overview of the plan4res modeling framework and describe these model interconnections.

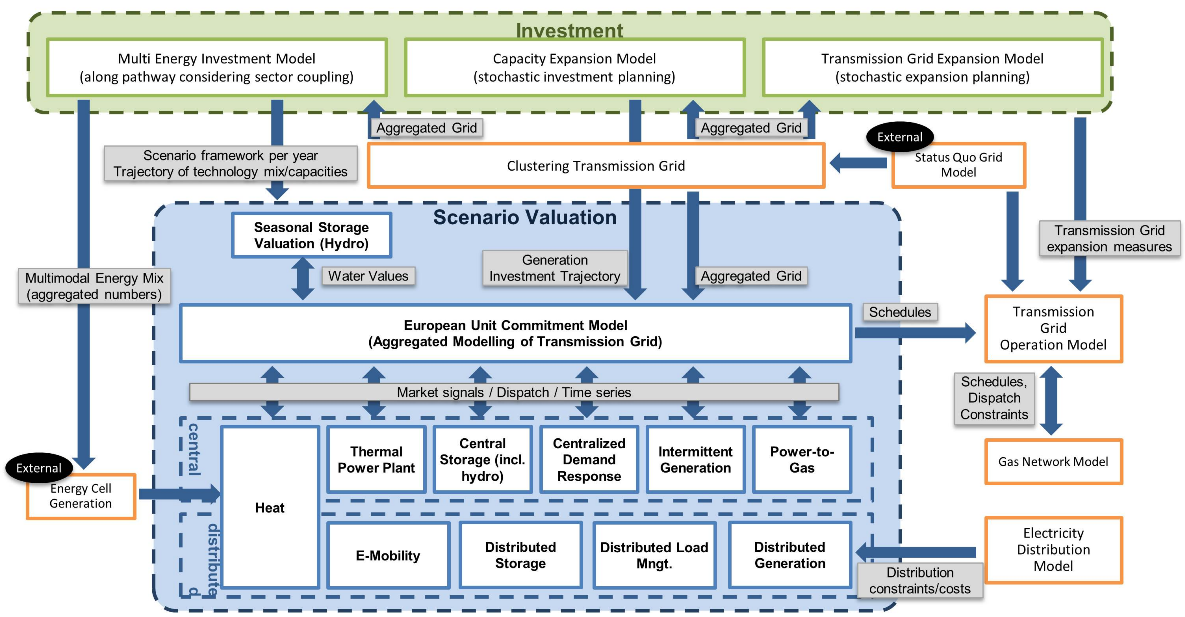


Figure 1: The plan4res model framework

Figure 1 gives an overview of the modeling framework, that is divided into

- Expansion models
- Valuation/operation models
- Supplemental models

The goal of the *expansion models* is to determine the optimal investment decisions for the future energy system. Since the case studies of plan4res have different key aspects, three investment models are defined that are tailored towards the needs of each case study.
The core of the *scenario valuation* is the European unit commitment (EUC) model, that optimizes the operation of the generation units determined by the investment models. A Lagrangian relaxation approach enables to decouple the generation units and define submodels for the different assets in the energy system. This modular approach also allows to only take the submodels into consideration, that are important for the respective case study.
Supplemental models are needed to either make input data available that are needed within the investment or valuation models (e.g. clustered version of the transmission grid, distribution reinforcement cost curves) or to do grid operation calculations (transmission grid as well as gas grid). The latter allow to also analyse the energy system regarding grid congestions, the amount of redispatch to clear these congestions and the capability of the gas grid to include gas provided by power-to-gas units.

1 Introduction

In the past, energy systems (electricity and gas) used to be mostly centrally planned and operated at the scale of each member state. Since the liberalisation of the energy system at the end of the 1990s, generation and supply of energy are planned and operated in a market framework, while transmission is still centrally managed at the scale of each member state. Another aspect having a high impact on the energy system transformation is the goal of reducing the CO_2 emissions that has been stated by the European Union. Incentive mechanisms, introduced to reach this goal, were the reasons for the massive expansion of renewable energies within the last decades. Thus the energy system is undergoing major changes and will have different facets in the future.

These changes, especially the emergence of a high share of intermittent renewable energy sources (RES) in the energy system, create completely new challenges. The volatile character of generation from renewable energy sources and the dependency on weather conditions increase the need for flexibility for the energy system many times over. These flexibilities can be provided by:

- Traditional centralized generation (including hydro)
- New grid equipment and improved operation techniques
- Storage (either distributed or centralized)
- Load management tools, i.e. flexible loads (either centralized or distributed)
- Flexibility provided by other energy sources (heat, gas, mobility), e.g. Power-to-gas, Power-to-heat or electro-mobility which can also be seen as storage capacities, thus introducing coupling between electricity and gas/heat
- The European electricity exchange

The idea of plan4res is to tackle all these aspects in an end-to-end energy system planning tool.

1.1 The plan4res model framework

In order to obtain a holistic assessment of the energy system all relevant aspects have to be taken into account (investment/operation, grid/market, central/distributed, different energy carriers). Since the modeling of theses aspects requires customized approaches, the idea of the framework is to seperate the individual components of the energy system into seperate model blocks. Figure 2 gives on overview of the model framework to be build within plan4res. Having separate models, allows the use of the most promising techniques regarding the mathematical formulation and solving methods for the specific models, thus increasing the computational efficiency of every single model within the framework. However to secure the functionality of the overall framework, the interconnections between the models have to be well defined. The task of this deliverable is to describe these interconnections.

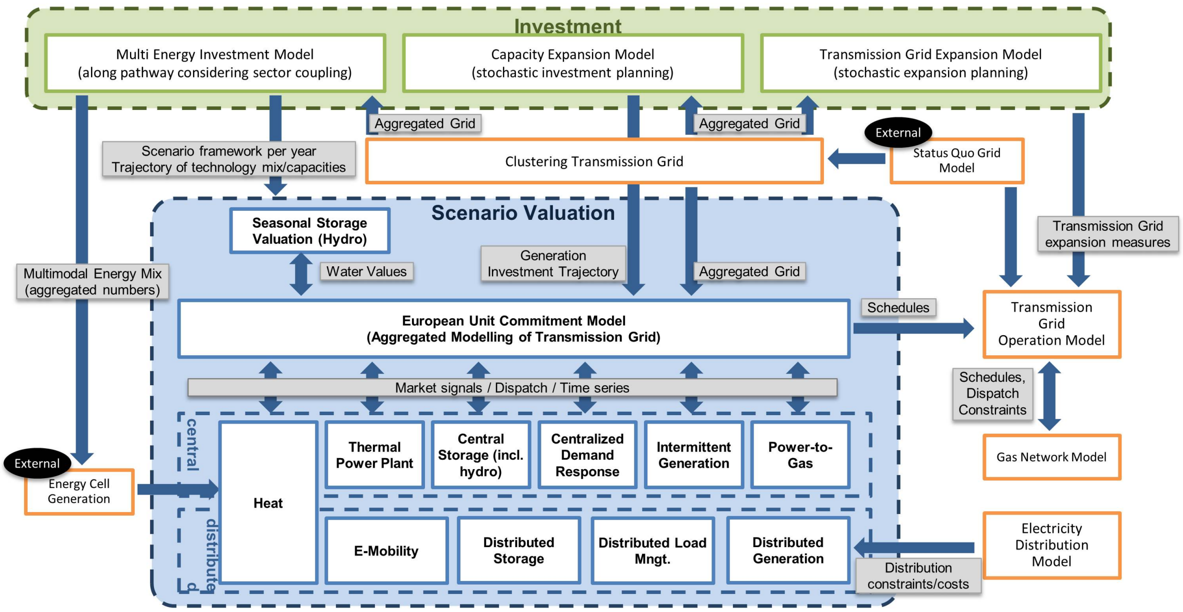


Figure 2: The plan4res model framework

1.2 On the representation of data involving time

Throughout the plan4res project, the code/models/equations will have to handle various types of data. Such data can either come as fixed numbers or as a "time-series". Such data may be available at a natural granularity quite different from the granularity of resolution of whatever model. For obvious reasons it is not desirable to have to smear out not so granular data over a "finely" discretized time horizon or to "aggregate" it whenever "not so granular" time horizons are considered. To this end, we will define a "time-series" (not to be confounded with "time-series model", popular in statistics) as follows. Time series data d is considered a function of continuous time $d : [t_0, \infty) \to \mathbb{R}$, given in the form of a "step-function". It is considered undefined prior to t_0.

Example 1.1 *The price series with provided data:*

- *2013-Sep-23T23:00 77.85*
- *2013-Sep-24T01:15 78.19*

is assumed to be "unspecified" prior to 2013-Sep-23T23:00 ; take on the value 77.85 for any time instant between [2013-Sep-23T23:00, 2013-Sep-24T01:15) and to be equal to 78.19 for all t >= 2013-Sep-24T01:15 ;

Under such a convention, it becomes easy to play with different discretizations of time, all while keeping the original data fixed. For some "time step" encompassing several values, it suffices to integrate over the interval and divide by total time (time-weighted average).

1.3 Cutting-plane models

Some of the model interconnections will physically take the form of transmitting a *cutting-plane* model. The reason for this is that the model interconnection must, to give an example, transfer some vision of future cost. The cutting plane model is, at the moment of transfer, the best possible vision of such cost. In order to precisely explain what is actually transferred, it is important to precisely describe what a cutting plane model is. To this end, let $f : \mathbb{R}^n \to \mathbb{R}$ be a convex function (see figure 3). With this convex function f, one can associate a cutting plane (CP) model consisting of a certain number of pieces, for instance k. This model will be denoted $\check{f}_k(x)$ and is given as the following maximum function:

$$\check{f}_k(x) = \max_{i=1,\ldots,k} \left\{ f(x^i) + \left\langle s^i, x - x^i \right\rangle \right\}, \tag{1}$$

where s^i can be understood as the "gradient / derivative" of f at x^i. Such models will form vital blocks of the model interactions and are relatively easy to store. Indeed, one needs to store the set of triplets $\{(f(x^i), s^i, x^i)\}_{i=1,\ldots,k}$ thus requiring the storage of 1 scalar and 2 vectors of dimension n.

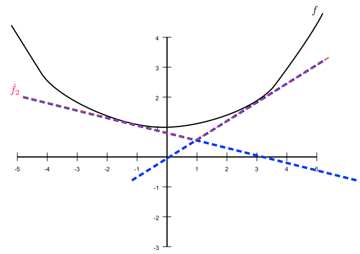


Figure 3: Example of a cutting plane model

1.4 Global picture

This preliminary section aims at providing a global view on the main optimization problems involved in the execution of the three case studies of plan4res. Since these case studies give a very detailled view on specific aspects of the energy system, the plan4res framework is tailored towards the needs of the respective case study. A detailled description of the models will be subject of a specific section in the sequel of the present document.

1.4.1 Overview case study 1

Case Study 1 aims at investigating a **Multi-modal European energy concept for achieving the COP 21 goal** with perfect foresight, considering sector coupling of electricity, gas, heat and transport demand.
Case study 1 will analyse the impact of sector coupling technologies, giving a detailled view on the overall energy system, including electricity, heat, cold and transport. It provides not only the optimal energy mix and operation schedules for one year, but for the entire pathway from today to 2050. The objective is to assess the plan4res framework to capture:

- The investment trajectory for a cluster of countries
- The impact of pan-European energy exchange
- The impact of sector coupling on the energy mix (Electrification of transport, heat, cooling; Flexibility provided by power2heat, heat storage, emobility, synthetic fuels; Coupling of electricity and gas sector by power2gas)

Considering sector coupling via heating/cooling technologies and transport demands adds many more technologies to be taken into account by the investment model. Optimizing the entire pathway and not just one year additionally increases the model complexity. Therefore the execution of case study 1 is divided into two modeling steps, with the results of step 1 providing the input data for step 2.

Step 1 - Determine investment decisions

Within step 1, the investment decisions along the whole pathway are determined, considering a simplified model for the operation of generation units. This step employs a linear optimization model with generic "input - conversion - output" processes to determine the investment decisions and operation schedules for an aggregated generation fleet. Hourly demand for "useful energies" (energy that can be assigned to a concrete benefit; e.g. space heating, industrial process heating, road/railroad/ship transport) are used as inputs for each of the considered sectors.

Step 2 - Determine operational schedules

Step 2 uses a more sophisticated operation optimization to give an even more detailled view on the electricity-, heat- and electric mobility sectors, using a single year approach. The results of step 1 are therefore used for building the input data for the step 2 optimization. Besides the installed generation capacities for the year under investigation, the energy mix data are used as input from step 1. Within the second step, the concept of energy cells is used to enable a detailled modeling of distributed heating technologies, which will be briefly described here.
The use of detailled socio economic data enables the construction of registers for households, commercial/trade/service (CTS) and industry within central western europe (CWE; Austria, Belgium, France, Germany, Netherlands, Switzerland). These registers model buildings and businesses on a spatial level in these countries. For every single entry (e.g. a specific household or a specific business location) the registers include a predefined heat demand (warm water, space heating, process heating) with an hourly temporal resolution. Furthermore the registers contain heat generation and storage technologies assigned to every single building/business based on the predefined energy-mix scenario generated by step 1. The heat generation units might also have an interaction to the electricity sector (e.g. heat-pumps as additional electricity loads or combined heat and power (CHP) units as electricity generators), thus providing flexibility to the energy system.
Within the process of building the input data these registers are aggregated to so called "energy cells". These energy cells define a regionally connected part of each country in

CWE and build an aggregation level between the registers and the heat submodel (Section 4.3). Each energy cell contains a matrix of aggregated heat demands (as hourly time series) per occurring technology combination (e.g. chp unit + gas boiler + heat storage unit) within this energy cell. Each heat demand profile is assigned an ID that clearly links this demand time series with the connected generation and storage units (e.g. generation units with ID 1,2,3 have to meet the heat demand with ID A). For each generator within the energy cells a variety of information is provided, e.g. the heat-ID that links this unit to the heat demand profile it should supply, the power-to-heat ratio, the efficiency, the energy cell ID and more (see the powerplant database described in section 4.1).
The heat submodel (4.3) will describe the hourly operation of those technologies to fulfill the heat demand, while maximizing its profits on the electricity market and minimizing the generation costs for electricity and heat. Not only distributed generation technologies are supplying the heat demands, but also conventional power plants are used to generate heat and supply buildings by district heating or industrial companies by process heating. That means that also conventional power plants can have a corresponding heat demand ID, and thus are not solely optimized regarding the energy price. The built energy cells are input for the EUC, that is used in step 2 for minimizing the operational costs of the generation units and flexibilities in the energy system for a single year.

1.4.2 Overview case study 2

Strategic development of pan-European network without perfect foresight and considering long-term uncertainties is the main object of examination in case study 2. It will analyze the optimal pan-European transmission grid investment strategy given the uncertainty that surrounds future system parameters. The objectives are:

- Demonstrate the ability of the plan4res frameworks to carry out system planning under uncertainty
- Identify optimal development pathways for the European transmission system under future uncertainty
- Assess the impact of long-term uncertainty on planning
- Assess the value of flexible non-network technologies

In particular, we will focus on the uncertainty around future demand due to electrification of heat and transport, future generation mix and location (e.g. north sea offshore developments), fuel cost (primarily coal vs. gas), future technology costs with a focus on energy storage as well as participation in demand-side schemes. The overall aim is to identify robust investment decisions that can be made in the near-term future and facilitate least-cost decarbonisation in the long term while minimizing the risk of stranded assets.

For simplicity we break down case study 2 in three levels.

Multi-stage scenario tree

The first task involves defining the scenario-tree that describes long-term evolution of the uncertain parameters. Note that depending on the sensitivity analysis we wish to carry out, the scenario tree may have a different shape or entail different uncertain parameters. The main idea is that each scenario tree node corresponds to a possible state of the 'world' and the model's aim is to identify the optimal investment and operational decisions that can be undertaken so as to minimize expected system cost throughout the study horizon. Due to the scenario tree's nested architecture it is possible to employ a nested decomposition approach to handle the huge computational complexity that emerges. In a nested benders decomposition scheme a master problem is first defined as the sum of the total cost (investment and operation) of the first stage plus the future cost function (i.e. an approximation of the expected cost of all emanating future scenarios). This approach can, in turn, be applied to all subsequent scenario tree nodes resulting in a nested problem structure where the cost of each node is expressed in terms of its children. In an iterative fashion, it is possible to apply trial investment solutions and then propagate backwards while computing the Lagrangian multipliers of the applied trial decision, thus obtaining local cost gradients. In the following iterations, this gradient information can be employed to iteratively drive our solution to the global optimum. Given the presence of binary variables in the subproblems, a relaxation is required - in cases of a larger-than-desired solution gap, convexification methods can be deployed as per the Disjunctive Branch and Bound (DBAB) algorithm. Focusing further into the problem corresponding to each individual scenario tree node, we can split it into an investment and an operational module. If the computational burden proves excessive, it is possible to use classical multi-cut Benders decomposition to split between investment and operation. However the right balance should be struck (across problem sizes) since the benefit obtained from stacked decomposition schemes can quickly become saturated due to model loading and input/output overheads.

Optimal investment

This sub-module enforces the constraints related to investing in network and non-network assets such as corridor capacity upgrades, building energy storage and implementing demand-side schemes. It entails aspects such as path-dependency across the scenario tree as well as mutual exclusiveness and other logical relationships that should be upheld.

Optimal operation

This sub-module determines the system operation along with all relevant constraints. A DC formulation will be adopted to reduce complexity. The aim is to combine the European unit commitment model with a transmission-constrained operation module in such a way as to achieve generation schedules that respect both unit commitment and locational constraints.

1.4.3 Overview case study 3

Identifying the **Cost of RES integration and impact of climate change for the European electricity system** in a future world with high shares of renewable energy sources will be the main focus of case study 3. The aim is to give a detailled view on the electricity system, combining generation investment, transmission grid investment and operations management. Considering uncertainties (e.g. demand, water inflows) allows giving an universal view on the electricity system in Europe. The objective is to assess the plan4res framework to capture:

- The cost of RES integration
- The value of different flexibility services
- The impacts of climate change

The generation investment problem involved in case study 3 includes three sub-problems related to three different time horizons.

1. At the long-term level, the objective is to design the optimal generation mix with the optimal transmission and distribution grid capacities for a given *long-term horizon* θ (say $\theta = 2050$). The problem then consists in minimizing the sum of two terms

$$\min_{\kappa} \left\{ C^{\mathtt{Inv}}(\kappa) + F(\kappa) \right\} ,$$

where

   (a) $C^{\mathtt{Inv}}(\kappa)$ denotes the investment cost, κ being a vector containing the investment capacities for each technology at each node;

(b) $F(\kappa)$ denotes the operation cost, associated with the given invested capacity vector κ.

2. The mid-term problem, also referred to as seasonal storage valuation (SSV), consists in evaluating the operation cost function, $F(\kappa)$, for a given vector of invested capacity, κ. In full generality, $F(\kappa)$ should represent the cost of optimally running the generation portfolio on the whole life of the portfolio, from the *long-term horizon*, θ to the end of the portfolio life. Fortunately, thanks of the seasonality of the problem, one can reduce the problem to optimally run the generation portfolio on a representative period $[\theta, \theta + T]$, typically corresponding to one single year. T will then be called the *mid-term horizon*. Notice that $C^{\mathtt{Inv}}(\kappa)$ will then represent an annualized investment cost. However, managing optimally the generation portfolio on the whole period $[\theta, \theta + T]$ (e.g. one year) cannot be treated as a deterministic optimization problem. Indeed, some random factors (such as reservoir inflows, or demand) are impacting the problem and the operation decisions are made dynamically while the random factors realizations are progressively revealed and the forecasts are accordingly updated. In fact, this consists of a multi-stage stochastic optimization problem. The mid-term period $[\theta, \theta + T]$ is divided into n sub-periods or stages $[t_k, t_{k+1})$ where $t_0 = \theta < \cdots t_k < t_{k+1} < \cdots < t_n = \theta + T$ (e.g. each stage corresponds to one week). We assume that for each stage, $k = 0, \cdots, n-1$, the whole information concerning the period $[t_k, t_{k+1})$ is simultaneously revealed, at the beginning of the stage, at time t_k. Hence, the mid-term problem can be stated as a stochastic dynamic problem and solved backwardly in time according to the dynamic programming principle. More precisely, on each stage, $[t_p, t_{p+1})$, the optimal operation decisions (related to seasonal storage and conventional power plants on sub-period p) can then be computed by solving a *transition problem* which consists in minimizing the production costs generated on sub-period p added to the value of a *cost to go function* evaluating the expected cost induced by optimally operating the system on the rest of the period $[t_{p+1}, t_n)$. These *cost to go functions* computed at each time step $t_0 < \cdots < t_{k+1} < \cdots < t_n$ depend mainly on the storage levels and constitute a way to decompose the optimization problem along the time by assigning a value to each storage level.

3. The short-term problem is related to the *transition problem*. This corresponds to the so-called Unit commitment (UC) problem where operation decisions are provided for one stage $[t_p, t_{p+1})$, in a deterministic horizon (random factors being fixed and known inside the sub-period p), taking into account the *value* that seasonal storages can bring to the system via the *cost to go functions*. This UC problem occurs in two ways.

(a) The *UC optimization mode* solves approximately the transition problem of the stochastic mid-term problem. In fact, it is intended to provide cutting plane approximations of *cost to go functions* and does not use any feasible recovery heuristic for the operation decisions. In this approach, the transition problem is, in general, not solved to optimality since operation decisions may be unfeasible and we rely on a cutting plane (lower bound) approximation of the *cost to go functions*. The advantage is that the *UC optimization mode* should run reasonably fast.

(b) The *UC simulation mode* computes a feasible generation dispatch, on a given sub-period. It uses the cutting plane approximations of the *cost to go functions* provided by the mid-term problem and is based on a feasible recovery heuristic ensuring the feasibility of operation decisions. This approach provides a sub-optimal solution to the original *transition problem* since the implemented strategy relies on a cutting plane (lower bound) approximation of the *cost to go functions*. The computing time required to run the *UC simulation mode* should be significantly greater than to run the *UC optimization mode*.

To compute the expected cost, $F(\kappa)$, it is more relevant to rely on feasible decisions and consequently to use the *UC simulation mode* implemented sequentially on each sub-period from $k = 0$ to $k = n - 1$. The expected optimal operation cost is then approximated as an average of the cumulative costs obtained by running *the UC simulation mode* successively on each sub-period, from 0 to $n - 1$, over N Monte Carlo simulations according to the uncertainties $(\xi^1, \cdots, \xi^N)$:

$$F(\kappa) \approx \frac{1}{N} \sum_{i=1}^{N} OptimalOperationalCost(\xi^i) \ .$$

2 Investment Layer

Having the framework divided into an *investment* and a *scenario valuation* layer, the investment layer will model investment decisions for the energy system. It will be composed of 3 models that are tailored to the specific needs of the three case studies to be analysed (compare global picture, section 1.4).

- A multimodal investment model which models investment along a pathway, taking into account the coupling of different energy sectors (Case study 1, section 1.4.1)
- A transmission grid expansion model focused on a detailed modelling of the transmission grid (Case Study 2, section 1.4.2)
- A capacity expansion model model with an aggregated modelling of the transmission grid (Case Study 3, section 1.4.3)

Taken together, these models provide a holistic view of the energy system investments. A detailled description of the interconnection with the *scenario valuation* layer and further input data will be given in the respective chapters.

2.1 Capacity expansion model

2.1.1 A tradeoff between operational costs and investment costs

The capacity expansion model is concerned with finding a (better) or ideally optimal set of assets including generation plants, interconnection capacities between clusters and distribution grid capacities, for the considered time horizon. Here optimal means, providing the least-cost set of assets, while accounting at best for the modelled constraints. In order to achieve this, the capacity expansion model can be operated in three different modes:

- Manual, "what-if" mode. In this elementary mode, we essentially will perform two runs: one with and one without a specific asset. The obtained differential of cost can then be compared to whatever investment cost the asset has.
- Sensibility-information. The seasonal-storage valuation tool will provide sub-gradient information with respect to the given set of possible assets to invest in. This information will provide a "direction" of investment.
- (Automatic mode). The previous sensibility information is automatically exploited to provide a more cost-effective set of assets. Then we obtain information regarding sensibility, which will allow us to update this set of choices and so on, until some convergence criterion is reached.

It is now clear, that we only need to describe the general structure of the last two items, since the first requires nothing special. To this end, and to fix thoughts, let κ denote the investment capacity in three types of assets

1. Generation plants
2. Interconnection between two clusters on the transmission grid
3. Distribution grid reinforcement

The objective of the Capacity Expansion Model (CEM) is then to:

$$\min_{\kappa} \left\{ C^{\mathtt{inv}}(\kappa) + \max_{\eta_1,\dots,\eta_S} F(\kappa, \eta_i) \right\},$$

where $\eta_1, ..., \eta_S$ are the distinct and finite set of "meta-scenarios" (e.g., some choice of climate-change trajectory). This corresponds to a robust optimization problem considering the "worst case" over the meta-scenarios. These "meta-scenarios" in turn impact the "distributions" of the regular set of scenarios. The "cost-function" $\kappa \mapsto C^{\texttt{inv}}(\kappa)$ are convex and "simple" to evaluate and refers to the cost of investment. For a fixed $i = 1, ..., S$, the map $\kappa \mapsto F(\kappa, \eta_i)$ refers to evaluating the expected cost given the investment κ, i.e., some run of the SSV. Within those runs, κ enters the constraint sets of certain assets. Thus making $F(\kappa, \eta_i)$ a parametric optimization problem. Under favorable structure, e.g., the resulting optimization problem is jointly convex in its regular optimization variables and κ, F will be convex in κ. In any situation, sensitivity, i.e., sub-gradient information, of F to κ can be associated with certain dual-multipliers.

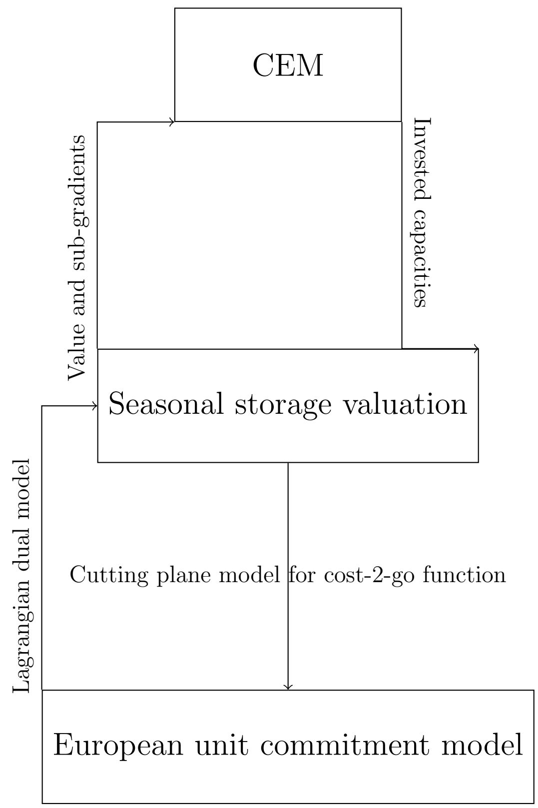


Figure 4: Capacity expansion model interconnections

2.1.2 Cost functions related to capacity investments

The cost function $\kappa \mapsto C^{\texttt{inv}}(\kappa)$ is obtained by concatenating the cost functions related to the three types of assets (generation plants, transmission capacities, distribution grid capacities). Distribution grid capacities limit the installation of distributed generation units and distributed storages. Thus distribution grid reinforcement is taken into account to increase this capacities. In particular, the cost function related to reinforcement costs of the distribution grid will be provided by the model described in Section 5.2.

The investment cost on the distribution grid is characterized by a curve, as illustrated on figure 5, providing the reinforcement cost of the distribution grid required to support an additional distributed generation capacity.

More precisely, for each region, the distribution network is modeled as an aggregation of three types of representative networks:

1. Aggregated rural reference network
2. Aggregated semi-urban reference network
3. Aggregated urban reference network

The distribution network in each region is characterized by a specific volume in each type of network and a specific reinforcement cost curve for each type of network.

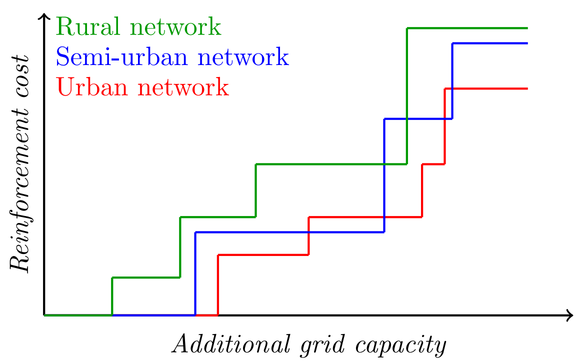


Figure 5: Distribution grid reinforcement cost as a function of the maximum capacity

Model requirements - Inputs

This model operates a tradeoff between the investment cost, $C^{\texttt{inv}}(\kappa)$ and the operation cost $F(\kappa)$ in order to select the optimal capacity κ. This requires to evaluate the operation cost, $F(\kappa)$ which requires to run successively SSV and EUC models.

Table 4: Required input data for the Capacity expansion model

Model	Input	Description	Format
External input	Long-term horizon	For instance 2050	scalar
External input	SSV and EUC data	All data required to run the SSV and EUC models	see section 3.1 and 3.2
External input	Meta-scenarios	Climate-change trajectories (Load, Inflows, Intermittent generation)	As a collection of time series (indexed by cluster and meta-scenario number)
External input	Investment costs in generation technologies	Costs functions related to each generation technology for each cluster	For each technology and each cluster, cost functions coefficients
External input	Investment costs in transmission grid	Costs functions related to the interconnection capacity for each couple of clusters	For each couple of clusters, cost functions coefficients
External input/Electricity distribution model	Investment costs in distribution grid	Cost reinforcement functions	For each cluster, cost reinforcement functions coefficients
External input	Constraints on investment capacities	Admissible set for the vector κ as a set of inequality constraints $A^t\kappa \leq B$	Matrix A and vector B
External input	Initial generation mix	Initial installed capacities in each generation technology for each cluster	For each cluster, a vector of installed capacity in each generation technology

External input/Clustering transmission grid	Initial transmission grid	Initial grid with a limited number of nodes (cluster nodes) and aggregated transmission lines (may be provided by the clustering model)	List of nodes and lines of the aggregated network
External input	Initial distribution grid	Distribution electrical nodes connected to each cluster of the transmission grid with a given initial maximal capacity	For each cluster, a list of distribution nodes with maximal capacity
European unit commitment	Operation costs	Minimal expected operation cost of the system required to satisfy the given demand constraints for a given installed capacity κ	one scalar
SSV	Sensibility information	Sub-gradients of the expected minimal operation cost with respect to each type of capacity (i.e. each coordinate of κ)	vector with the same size as the capacity vector κ

Model results - Outputs

The Capacity expansion model is intended to provide an approximation of the generation mix as well as reinforcements required to be operated on the transmission ans distribution grids.

Table 5: Results of the Capacity expansion model

Model	Output	Description	Format
Result	Investment capacity	Best obtained generation mix and transmission and distribution grid investments	vector of capacity
Result	Best obtained operation and investment costs	Best obtained tradeoff operated by the Capacity expansion model between operation and investment costs	scalar
Result	Lower bound for the optimal cost	Lower bound obtained by cutting-plane approximation	scalar
Result	Indicators of success	Provides an indication if solving the Capacity expansion model has been successful or failed for some reason	scalar

2.2 Multimodal investment model

General

The *multimodal investment model* considers different energy consuming sectors. It uses multiple energy carriers that satisfy the demand for useful energy e.g. for lighting, households, industry, heating, transport, etc. A graphical overview is given in Figure 6.

The modeling approach is multi-modal, i.e. it allows directly considering coupling between several energy sectors, like electricity, heat/cold, fuels/gas and chemicals. The system can be parameterized to determine optimal investment paths from today onwards for a chosen horizon of interest. The model can consider several distinct regions and the required interconnection capacities in between. A maximal amount of CO_2 emissions can be set as a constraint for the entire horizon which will result in a cost-optimal abatement across all sectors and years. Additionally the CO_2 emissions for every single planning step can be constrained.

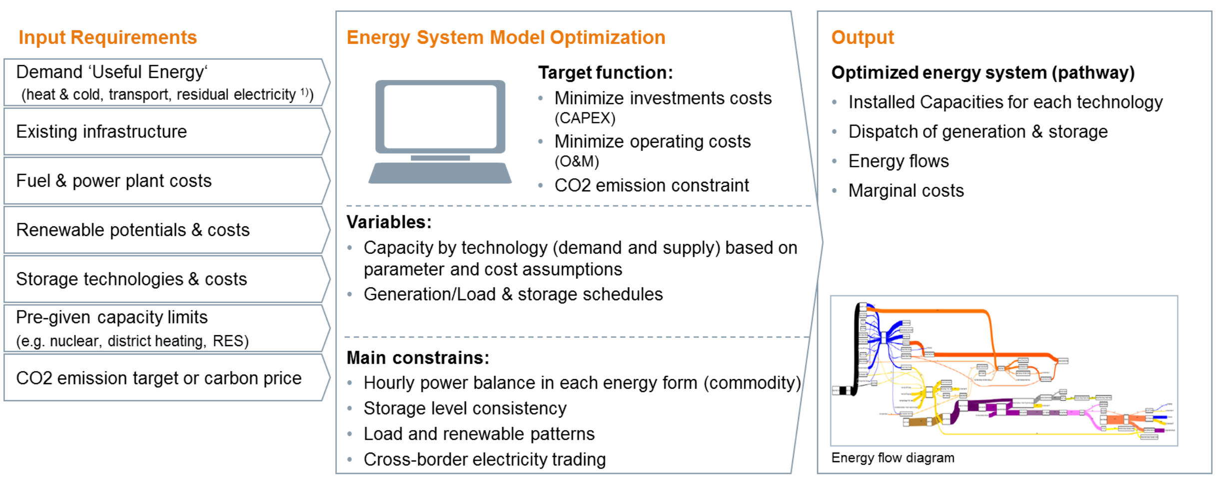


Figure 6: Overview of the multimodal energy investment model

A linear approximation of all operational details is considered to be sufficient for the design/investment decisions. For the investment decision, the entire costs including capital expenditure (CAPEX) as well as the required operating expenditure in system operation (OPEX) have to be considered. The investment and operation decisions for all time steps and all regions are determined in one single optimization run in order to achieve optimal intertemporal allocation. Several planning steps are considered (e.g. 5 years steps) to optimize the system development until the final year of the planning horizon (e.g. 2050).

In each of these planning steps the hourly operation for the entire year or representative weeks is considered including operational constraints of power plants and storages.

Investment and operational costs are aggregated over several simulated years and discounted to a net present value in today's terms that is effectively optimized to yield the system's optimal development pathways. Necessary constraints (e.g. energy conservations, power limitations, temporal consistency of storage levels) are also formulated using these abstractions.

Multimodality

The model considers all energy consuming sectors and all carriers relevant for energy system modeling, including electricity, heat/cold, liquid/gaseous fuels, and optionally chemicals. A possible approach to consider all of these commodities is to define processes that transform input commodities into output commodities in a one step approach. For each process, the technical details such as efficiency and costs are provided as inputs and the invested capacity and hourly operation is given as output. A manifold of these one-step approaches allows to flexibly model all kinds of processes, e.g. power plants, heat pumps, CHP and Power-to-X (PtX). In addition to the resulting technology mix for each year, the model also provides the electricity consumption for all sectors in hourly resolution.

Time and space

Depending on technical possibilities, ideally 8760 hours of each year are considered for each of the intervals in order to represent daily and seasonal patterns of load and renewable generation. In order to represent the regional characteristics as good as possible, the minimum approach is to model at least on a country level and consider the energy exchange between several countries.

Model requirements - Inputs

Technology data: The main inputs are available technologies and their technical and economic description (e.g. efficiency, investment costs per installed power, technical and financial lifetime, etc), scenario-dependent side conditions (e.g. fixed or maximum capacities for some technologies) and timeseries data such as the time-dependence structure of a given demand or renewable availability (if possible with regional resolution), usually as timeseries normalized between 0 to 1.

Table 6: Required input data for the multimodal investment model

Model	Input	Description	Format
External input	Discount Rate	Discount Rate	One value
External input	WACC	Weighted average cost of capital	One value
External input	Default financial lifetime	Default financial lifetime	One value
External input	Annual CO_2 emissions	Annual CO_2 emissions	One value per planning step
External input	Total CO_2 emissions	Total CO_2 emissions until 2050	One value
External input	(Fixed) Demand load curves	Demand load curves covering: • Residual electricity • transport demand • heat demand • cooling demand	One normalized timeseries per energy type (from characteristic reference year)
External input	Annual demand for usefuel energies	Usefuel energy demand covering: • Residual electricity • heat central (4 groups) • heat decentral • transport • cooling • Industrial (if applicable, e.g. H2)	Three values (TWh_{el}, TWh_{th}, Energy needed to match demand km*passenger) per planning step

External input	Generation curves RES	RES generation for PV, wind & solar thermal.	Normalized time-series from characteristic reference year
External input	Fuel price	Fuel price assumptions	One value per fuel type and planning step
External input	Price for other imports	Price assumptions for other imports (e.g. biomass)	One value per type and planning step
External input	Specific CO_2 Emissions	CO_2 emissions per conversion process	One value per conversion process
External input	Technical lifetime	Technical lifetime	One value per technology
External input	Financial lifetime	Financial lifetime	One value per technology
External input	Cost assumptions	Cost assumptions per conversion process consisting of • CAPEX (€ per kW) • O&M (€ per kWh-throughput) • O&M (€ per kW*a)	Three values per conversion process and planning step

External input	Cost assumptions storages	Cost assumptions per storage technology consisting of • CAPEX (€ per kW) • CAPEX (€ per kWh-throughput) • O&M (€ per kWh-throughput) • O&M (€ per kW*a)	Four values per storage technology and planning step
External input	C-rate min/max (storage only)	Maximum charge and discharge rate	Min & max value per storage technology
External input	Efficiency	Efficiency of conversion processes	One value per conversion process and planning step
External input	Efficiency (storage)	Efficiency of storages technologies	One value per storage technology and planning step
External input	Limiting min/max fraction of output commodity of this special conversion process	Commodity can be generated by several processes	As % per conversion process
External input	Limiting min/max fraction of input commodity of this special conversion process	Process can be provided by several input commodities	As % per conversion process

External input	(Limiting) Max ratio for new capacity installations	Maximum new capacity installations limited with respect to installations in neighbouring planning steps	One value per conversion process and planning step (as % of new capacity in beighbouring planning steps)
External input	Installed capacity	Installed generation capacities for the base year (e.g. 2020)	One value per conversion process
External input	Min/max capacity per conversion process	Lower/upper limitation for conversion process capacities	One value per conversion process and planning step
External input	Min/max energy generation per conversion process	Lower/upper limitation for conversion process generation	One value per conversion process and planning step
External input	Installed storage capacity and storage power	Installed storage capacities [kWh] and storage power [kW] for the base year (e.g. 2020)	Two values (capacity, power) per storage technology
External input	Min/max capacity per storage technology	Lower/upper limitation for storage capacities [kWh]	One value per storage technology and planning step
External input	Optional: Availability curves	Hourly availability of each conversion technology	One normalized timeseries from characteristic reference year per conversion technology

Model results - Outputs

The primary outputs are operation schedules for each conversion process, cost-optimal investment decisions for each conversion process and respective investment trajectories (including early retirements). Also, an indication for market prices (marginals) for each of the modeled energy forms for each modeled hour, can be given, assuming perfect market conditions.

From these primary outputs further secondary outputs can be derived, which include:

- Total system costs
- Total CO_2 emissions
- Technology specific CAPEX/OPEX/annual revenues per kW
- Statistics on the required flexibility of a technology's operation

Table 7: Results of the multimodal investment model

Model	Output	Description	Format
Result	Conversion process investment	Investment in conversion process per planning step	One value per conversion process and planning step (in MW)
Result	Storage capacities investment	Investment in storage capacities per planning step	One value per storage technology and planning step (in MWh)
Result	Electricity demand	Overall electricity demand (including heating, transport, etc.)	One timeseries per planning step
Result	Annual electricity demand	Overall electricity demand (including heating, transport, etc.)	One value (TWh) per planning step
EUC	CO_2 price	Mean CO_2 price for each planning step	One value per planning step (€ per ton CO_2)
EUC	CO_2 emissions per planning step	Overall CO_2 emissions per planning step	One value per planning step (in tons CO_2)

Successful modeling of complex energy systems requires a modular data model. Scenarios are based on the one side on large volume data like time series or spatiotemporal data sets and on the other side on complex interrelated techno-economical parameters of the energy conversion units. For detailed (post-)analysis of the scenarios, all results must be made avaible in an accessible format.

2.3 Transmission grid expansion model

The transmission grid expansion model (TGEM) makes use of two main sub-modules, namely the investment sub-module and the operation sub-module. The investment sub-module is responsible for identifying the optimal investment strategy across a multi-stage scenario tree. The operation sub-module is responsible for operating the given system in a cost-optimal way. These two sub-models can either be joined into a single optimization problem or kept separate and driven to global convergence through a decomposition technique. Due to the nature of the problem, multi-cut Bender decomposition (classical or hierarchical) is proposed as a suitable decomposition scheme.

The EUC is supposed to be used for the the transmission grid expansion model as operational model. Once the final solution of transmission grid expansions has been obtained, the transmission grid operation model can be used as an ex-post analysis tool for detailled AC transmission simulations to analyse congestions and the amount of redispatch to clear these congestions (see section 5.4).

Model Overview

The stochastic transmission planning problem will be formulated as a mixed integer-linear programming (MILP) problem. Uncertainty is modeled in the form of a multi-stage scenario tree consisting of $|\Omega_M|$ nodes spanning $|\Omega_E|$ stages (also referred to as epochs). The scenario tree portrays the possible states of the 'world' along with transition probabilities. The model will adopt a node-variable formulation where each scenario-dependent parameter will be expressed in terms of a particular scenario tree node. As such, if we choose to incorporate uncertainty on the evolution of future demand, we can use the parameter $D_{(m,t,n)}$ where indices m, t and n refer to the scenario tree node, time and system bus respectively. The objective of the transmission grid expansion model is the minimization of expected investment and operation cost. Given a node-variable formulation, this becomes the minimization of the probability-weighted cost corresponding to each scenario tree nodes as:

$$\min \sum_{m \in \Omega_M} \pi_m (\omega_m^I + \omega_m^O)$$

where π_m is the probability of occurrence of node m, ω_m^I and ω_m^O is the investment and operation cost corresponding to node m respectively. In general, investment cost ω_m^I is the sum of all investment decisions made at node m. We assume that investment decisions are irreversible and their capital cost is accrued at the time of decision (which may be different to the time of commissioning due to long building times):

$$\omega_m^I = r_m^I \sum_{l \in \Omega_L} \sum_{o \in \Omega_O^l} \kappa_{o,l}^B B_{m,l,o}$$

where Ω_L is the set of candidate line corridors, Ω_O^l the set of candidate investment options regarding line l (e.g. we may be able to choose between a small or a large reinforcement to line l, each adding a different amount of capacity and entailing a different cost), $\kappa_{o,l}^B$ is the cost of expansion (o, l). The parameter r_m^I denotes the discount factor corresponding to node m and depends on the undertaken assumption regarding macroeconomic situation and epoch length. The binary decision variable $B_{m,l,o}$ denotes the decision to invest in that option in the current scenario tree node m. In turn, operation cost ω_m^O is equal to

$$\omega_m^O = r_m^O \sum_{t \in \Omega_T} \tau_t \sum_{g \in \Omega_G} \kappa_{m,g}^G$$

where Ω_T is the set of operating points considered, each with duration τ_t, Ω_G is the set of generators and $\kappa_{m,g}^G$ is the generation cost of unit g at scenario tree node m - note that setting this parameter to be a function of m renders possible the modeling of uncertain future generation costs. The parameter r_m^O is the cumulative discount factor. In addition, there is a number of constraints that need to be respected. We indicatively show the state equation enforcing path-dependency related to investment in lines:

$$\tilde{B}_{m,l,o} = \sum_{m \in \Omega_M^{m-\gamma_{l,o}^B}} B_{m,l,o}$$

where $\tilde{B}_{m,l,o}$ is the state variable related to control variable $B_{m,l,o}$ (i.e. denotes whether the expansion option (l, o) has been commissioned at node m). Parameter $\gamma_{l,o}^B$ denotes the building time required for expansion option (l, o), expressed in terms of epochs. $\Omega_M^{m-\gamma_{l,o}^B}$ denotes the set of scenario tree nodes form the first stage up to stage $\varepsilon_m - \gamma_{l,o}^B$ (where ε_m is the stage corresponding to node m). This set is derived directly from the scenario tree's architecture and is required to impose path dependency including building delay.

We also show some basic constraints related to system operation, starting with the system balance equation:

$$\sum_{g \in \Omega_G} p_{m,t,g} + \sum_{l \in \Omega_L^{n+}} f_{m,t,l} - \sum_{l \in \Omega_L^{n-}} f_{m,t,l} + d_{m,t,n} = D_{m,t,n}$$

where $p_{m,t,g}$ is the power output of unit g, $f_{(m,t,l)}$ is the power flow over line l and $d_{m,t,n}$ is the demand curtailed at bus n. Note that sets Ω_L^{n+} and Ω_L^{n-} refer to lines defined as importing/exporting energy to bus n according to the system topology definition.

The main constraints coupling investment decisions with operation are the following:

$$|f_{m,t,l}| \leq F_l^0 + \sum_{o \in \Omega_O^l} \tilde{B}_{m,l,o} F_{l,o}$$

where F_l^0 is the initial capacity of corridor l.

Note that in the full model formulation, investment and operation of additional asset types will be available such as energy storage devices.

Model requirements - Inputs

Table 8: Required input data for the transmission grid expansion model

Model	Input	Description	Format
External input	Scenario tree stages	Description of scenario tree stages set Ω_M	List of $\|\Omega_E\|$ scenario stage objects. Each stage object has fields: unique identifier name (string), duration in years (integer)

External input	Scenario tree nodes	Description of scenario tree nodes set Ω_M	List of $\|\Omega_M\|$ scenario tree node objects. Each node object has fields: unique identifier (string), name of parent node (string - null for root node), probability of occurrence (real $\in [0,1]$), name of stage to which it belongs (string)
External input	Time periods set	Description of time periods set Ω_M	List of $\|\Omega_T\|$ time period objects. Each time period object has fields: unique identifier (string), duration (real number - hours), name of preceding node (string - required when modeling includes storages/DSR)
External input	Network - nodes	Set of all transmission network bus Ω_N	List of $\|\Omega_N\|$ bus objects. Each bus object has fields: unique identifier (string), country to which it belongs (string)

External input	Network - lines	Topology and capacity of existing and candidate future transmission corridors	List of $\|\Omega_L\|$ line objects. Each line object has fields: origin bus id (string), destination bus id (string), name (string), initial transfer capacity (MW)
External input	Network - generators	Topology and capacity of existing and future generators	List of $\|\Omega_G\|$ generator objects. Each generator object has fields: name id (string), bus id (string), power rating (MW), operating cost (€/MWh) [if uncertain, defined at each scenario tree node m]
External input	Discount factor	Assumption around discount factor for capital investments	Real number [if uncertain, defined at each scenario tree node m]

External input	Cost assumptions - lines	Cost assumptions for each candidate line project	List of candidate line expansion project objects. Each object has fields: name (string), line id (string), capacity addition (MW), reactance (p.u.), CAPEX (e) [if uncertain, defined at each scenario tree node m], building time (integer)
External input	Cost assumptions - storages	Cost assumption for each candidate storage project	List of candidate storage project objects. Each object has fields: name (string), bus id (string), plant power rating (MW), plant energy rating (MWh), CAPEX (e) [if uncertain, defined at each scenario tree node m], building time (integer), efficiency (real)

External input	Cost assumptions - demand-side schemes	Cost assumption for each candidate DSR scheme	List of candidate DSR scheme objects. Each object has fields: name (string), bus id (string), power rating (MW), maximum energy shift time (hours), CAPEX (€) [if uncertain, defined at each scenario tree node m], OPEX (€/MWh shifted), building time (integer)

Model results - Outputs

Table 9: Results of the transmission grid expansion model

Model	Output	Description	Format
Result	Optimal costs	Optimal cost across scenarios and time periods	For each scenario tree node $m \in \Omega_M|$: CAPEX (real number), OPEX (real number) (also presented per time period $t \in \Omega_T$)
Result	Optimal investment decisions	Optimal investment decisions at each scenario tree node	For each scenario tree node $m \in \Omega_M|$, list of the name (string) of assets built under the optimal strategy

3 Scenario valuation layer

The *Scenario valuation layer* is supposed to evaluate the investment decisions from the *Investment layer* by means of modeling the operation of the existing assets in the energy system.

This layer contains two distinct models, the first model will be referred to as the seasonal storage valuation and the second model will be the European unit commitment model.

Using lagrangian relaxation, system coupling constraints (e.g. electricity demand constraint) of the EUC can be relaxed. This makes it possible having several submodels representing the different generation technologies in the energy system. Via the lagrangian multipliers (acting as representatives of a electricity price) these models have an equal optimization criteria, thus beeing connected by these multipliers.

3.1 European unit commitment model

For a detailed overview of what unit commitment is, we refer to [7]. In the subsequel, we will refer to "unit commitment" as a collection of closely related optimization problems. The goal of these problems is to find an optimal (or near optimal) schedule satisfying the set of technical constraints. In view of this specification, it becomes clear that the "unit commitment" problem encompasses the set of "sub-problems". Furthermore, the problem can be seen separately from whatever methodology is employed to actually solve the problem at hand. To fix ideas, let us present a generic "formulation":

$$\begin{aligned} \min_{x_1,\dots,x_m} \quad & \sum_{i=1}^{m} f_i(x_i) \\ \text{s.t.} \quad & x_i \in X_i \\ & \sum_{i=1}^{m} h_i(x_i) \leq 0, \end{aligned} \tag{2}$$

where $X_i \subseteq \mathbb{R}^{n_i}$ is an arbitrary set, $f_i : \mathbb{R}^{n_i} \to \mathbb{R}$ a given cost function and $h_i : \mathbb{R}^{n_i} \to \mathbb{R}^p$ a given set of "coupling" constraints.

Practically, there are mainly five types of coupling constraints involved in $(h_i)_i$. Depending on which case study is under investigation, a set of these constrained is considered, while others might be excluded. As it won't be possible to represent the transmission network in whole details taking into account all nodes and transmission lines, a simplified modelling where nodes are aggregated into clusters will be used (see section 5.3).

Supply demand balance corresponding to power demand for each cluster, $n = 1, ..., N^{cluster}$, at each time step t of the optimization horizon:

$$D_{n,t} = g_{n,t} - d_{n,t} \ ,$$

where for each cluster, D_n, d_n and g_n are time series such that

- $(D_{n,t})_t$ is an input data corresponding to the non-flexible power demand, at cluster n, $D_{n,t} = (D_{n,t}^{trans}, D_{n,t}^{dist})$ where $(D_{n,t}^{trans})_t$ is the non-flexible demand connected to the transmission network at cluster n and $(D_{n,t}^{dist})$ is the non-flexible demand connected to the distribution network at cluster n;
- $(d_{n,t})_t$ is an output of the EUC model, $d_{n,t} = (d_{n,t}^{trans}, d_{n,t}^{dist})$ corresponds to the flexible demand at cluster n, which is either connected to the transmission network or connected to the distribution network.

- $(g_{n,t})_t$ is an output of the EUC model, $g_{n,t} = (g_{n,t}^{trans}, g_{n,t}^{dist})$ corresponds to the power injected in cluster n by (conventional or intermittent) power plants or storage devices, distinguishing $g_{n,t}^{trans}$, the power injected to the transmission network node of cluster n, with $g_{n,t}^{dist}$, the power injected to the distribution network node of cluster n;

The reason to distinguish devices connected to the transmission network with devices connected to the distribution network is motivated by specific constraints arising on the distribution grid

$$|D_{n,t}^{dist} + d_{n,t}^{dist} - p_{n,t}^{dist}| \leq \bar{P}_n \ ,$$

where $\bar{P}_n$ stands for the capacity of the distribution grid at cluster n and p_n^{dist} represents the power generated by power plants connected to the distribution grid at cluster n. Of course, this corresponds to an aggregated model of real constraints occurring on the distribution grid. Remark that in this framework, cluster n can be viewed as a couple of electrical nodes with

1. one *transmission electrical node* connected to the transmission network with load D_n^{trans}, d_n^{trans} and power generators providing p_n^{trans};
2. one *distribution electrical node* representing the distribution networks at cluster n with load D_n^{dist}, d_n^{dist} and power generators providing p_n^{dist}: the distribution electrical node is exclusively connected to the transmission electrical node n with a limited interconnection capacity given by $\bar{P}_n$.

Primary and secondary reserves constitute services provided by generators to the electrical system in order to support continuously the equilibrium between supply and demand. Some generators offer the ability to rapidly increase or decrease their production to meet fast changes in demand.
Primary and secondary reserves requirements may be stated specifically for each cluster, $n = 1, ..., N^{cluster}$, or more generally for each *reserve zone*, where a reserve zone may contain several clusters. Formally, a reserve zone is represented as a partition, $(Z_q)_q$, of $\{1, ..., N^{cluster}\}$. The reserve constraints are then stated for each time step t and each reserve zone Z_q as follows

$$\begin{cases} \sum_{n \in Z_q} r_{n,t}^1 \geq \underline{R}_{q,t}^1 \\ \sum_{n \in Z_q} r_{n,t}^2 \geq \underline{R}_{q,t}^2 \end{cases} ,$$

where

- $(\underline{R}_{q,t}^1, \underline{R}_{q,t}^2)_{q,t}$ is an input data corresponding to the primary and secondary reserve requirements for each reserve zone, q at each time step, t;

- $(r^1_{n,t}, r^2_{n,t})_{n,t}$ is an output from the EUC model corresponding to the primary and secondary reserve provided by the system to each cluster,n, at each time step, t.

Inertia is related to the dynamical properties of generators. It is a crucial characteristic of the electrical system which determines the time allowed for the grid operator before resorting to the primary reserve to control the stability of the system.
Inertia requirements may be stated for each *inertia zone*, $(Z_q)_q$ as follows

$$\sum_{n \in Z_q} \sum_{k} H_{n,k} p_{n,k,t} \geq \underline{H}_{q,t} \ ,$$

where

- $(\underline{H}_{q,t})_q$ is an input data corresponding to the inertia requirements for each inertia zone q at time t;
- $(H_{n,k})_n$ is an input data characterizing each power plant (n,k) (connected to cluster n and indexed by k) which represents the power plant ability to provide inertia.
- $(p_{n,k,t})_{n,k,t}$ is an output from the EUC model corresponding to the power provided by the power plant connected to cluster n and indexed by k at each time step t.

Heat demand balance corresponding to heat demand within the energy cells and connected to the generation units (heating-only as well as sector coupling technologies, e.g., power-to-heat, heat-pumps, CHP) by means of a heat-ID (see also section 1.4.1) which is defined as

$$D^{thermal}_{e,t,h} = \sum g^{thermal}_{e,t,h}$$

- $D^{thermal}_{e,t,h}$ describing the thermal demand in enegy cell e at time t that is identified by the heat-ID h.
- $g^{thermal}_{e,t,h}$ is an output from the EUC model representing the thermal generation provided by generation units with heat-ID h in energy cell e at time t.

DC power flow The generation schedule has to fulfil some grid constraints related to

- limited transmission capacity between clusters;
- physical laws which determines power flows through the grid.

The model will rely on the DC power flow model which is a linearization of the nonlinear AC power flow model. The DC flow model suppose a linear relationship between power injections at each node of the grid and active power flows through the transmission lines. This linear relationship will be represented by the power transfer distribution factor (PTDF) matrix that will constitute an important output of the clustering model. Then, the active power flows are limited by interconnection capacities between the clusters with constraints of the type $\underline{P}_\ell \leq p_\ell \leq \bar{P}_\ell$ for each line ℓ.

It requires two key elements to solve problem (2) or any of its variants (related to how data will be entered) in plan4res:

- Decomposition: based on the Lagrangian dual. This phase makes appear the notion of “sub-problem” and, if appropriate methodology is available, makes clear that f_i and X_i can be “relatively” arbitrary and specified “nearly” independently of (2).
- Primal Recovery. This optional phase is needed in order to retrieve a solution of decent quality following the Lagrangian dual phase. Note that this is already achieved when f_i is convex, h_i affine and X_i convex. Then, under these convexity assumptions, although a globally optimal solution is readily available by taking an appropriate combination of all produced iterates (the pseudo-schedule, compare for example [3]), it is not true that the sub-problems responses at optimal dual multipliers are optimal (for (2) (generally it is not even anywhere near optimality)). We refer to [1, 9, 10] for further information about this issue. Note that in the non-convex case, the pseudo-schedule is also a vital piece of information in order to retrieve a near optimal solution to (2). This optional phase will exploit information already produced and some knowledge of X_i, f_i and h_i. For the seasonal storage model, this phase will not be executed.

To get a deeper insight into this approach, the dual of problem (2) is formulated, which consists in solving

$$\sup_{\lambda \geq 0} \Theta(\lambda), \tag{3}$$

where $\Theta : \mathbb{R}^p_+ \to \mathbb{R}$ is given by:

$$\Theta(\lambda) := \sum_{i=1}^{m} \min_{x_i \in X_i} f_i(x_i) + \lambda^\mathsf{T} h_i(x_i). \tag{4}$$

In what follows we will call the optimization problem $\min_{x_i \in X_i} f_i(x_i) + \lambda^\mathsf{T} h_i(x_i)$ a subproblem; m subproblems need to be solved. Solving problem (3) will be called “maximizing the lagrangian dual" and is typically done by some iterative procedure such as subgradient method or bundle method (e.g., [5, 6]). This procedure will produce a series of Lagrangian multipliers or "price signals" converging to the optimal dual vector λ^*. At each price/Lagrange vector, each sub-problem will provide a candidate feasible solution w.r.t. the set X_i.

Primal recovery is all about combining this set of information to provide a near optimal solution. Indeed, since the decomposition does not guarantee the adherence to the relaxed constraints, this process can include a subsequent method determining the final solution (e.g., an economic dispatch with fixed integer variables). However, usually it is more beneficial to determine appropriate integer variables based on the information from the pseudo-schedule.

A feature that may warrant some further comments is the notion of time. In order to understand how this interacts with the model, we must first provide some more information on what the coupling constraints in (2) actually mean. They could cover various coupling constraints, but cover at least generation/demand balance conditions at different nodes in the transmission grid and for different instants of time. The natural notion of time for the EUC model is therefore this discretization. Furthermore in light of our earlier discussion on the modelling of temporal data in section 1.2, it becomes clear that subproblems need not follow exactly the same temporal discretization. Obviously in order to have some degree of consistency, at least one subproblem should match the temporal discretization of the European unit commitment problem (or the finest partition extracted from the set of subproblems should do so).

Model requirements - Inputs

The EUC model acts as central unit of the operation layer. As explained above, the relaxed constraints are attached to a price signal to which the submodel will react. The information about the assets, that are optimized within the subproblems are described in the respective submodels. Note that hydro valleys will be specified as a collection of assets connected through a set of reservoirs; then the terminology "asset" is meant to refer to an individual turbine/pump station.

Table 10: Required input data for the European unit commitment model

Model	Input	Description	Format
External input/SSV	Time horizon	Time horizon to be optimized	Value (number of hours)

External input/SSV	Time Steps	Discretization of time horizon (e.g. hourly)	Value or set of values if heterogeneous
External input	(Aggregated) electricity transmission grid nodes	Clusters representing an aggregated transmission grid	List of nodes
External input	(Aggregated) electricity transmission grid lines	Transmission grid lines connecting the definded clusters	From/to node list
External input	Electricity demand	Electricity demand assigned to the defined clusters	One timeseries per cluster
External input	Optional: Spinning reserve demand	Demand for primary/secondary spinning reserve assigned to the defined clusters	One timeseries per reserve zone and type (primary/secondary)
External input	Optional: Inertia demand	Demand for Inertia assigned to the defined clusters	One timeseries per inertia zone
External input	CO_2 Emission Cap	Constraint that defines the maximum allowed CO_2 emissions	Value
Seasonal storage valuation model	Cutting plane model	Associated with the last time step in the horizon, one/several convex cutting plane models for the future value of some aggregated state (representing for instance some “weighted sum” of final volume levels of associated reservoirs)	A set of triplets (value, subgradient, evaluation point)

External input	Initial parameters	• Initial cutting plane model (for the Lagrangian dual, optional) • Initial lagrangian multipliers • Initial parameters for lagrangian dual algorithm (e.g., stopping criteria, optional)	

Model results - Outputs

As a consequence of the iterative process, the European unit commitment model determines an adjusted price signal due to an over-/undersupply. After the final iteration the EUC delivers the final dispatch schedules needed for the transmission grid calculations.

Table 11: Results of the European unit commitment model

Model	Output	Description	Format
Result & Transmission grid operation model	Operation schedule	Generation of the power plant units (related to lagrangian multiplier/price signal)	One timeseries per power plant
Result & Transmission grid operation model	Operation schedule	Generation & consumption of the electric storages (related to lagrangian multiplier/price signal)	One timeseries per storage

Result & Transmission grid operation model	Operation schedule	Generation & consumption of the aggregated e-mobility storage (related to lagrangian multiplier/price signal)	One timeseries per aggregated e-mobility unit
Result & Transmission grid operation model	Operation schedule	Generation of the intermittent generation units (Wind, PV, Hydro) (related to lagrangian multiplier/price signal)	One timeseries per renewable generation unit
Result & Transmission grid operation model	Operation schedule	Electricity generation of distributed units (related to lagrangian multiplier/price signal)	One timeseries per distributed generation unit
Result & Transmission grid operation model	Operation schedule	Generation & consumption of distributed storages (related to lagrangian multiplier/price signal)	One timeseries per distributed storage unit
Result & Transmission grid operation model	Operation schedule	Electricity generation & consumption energy cells (related to lagrangian multiplier/price signal)	One timeseries per energy cell
Result & Transmission grid operation model	Operation schedule	Electricity consumption of the power-to-gas units (related to lagrangian multiplier/price signal)	One timeseries per power-to-gas unit

Submodels	Lagrangian multipliers (price signal) for the electricity demand	An indicator representing the electricity price that drives the operation of the submodels	One timeseries per cluster
Submodels	Optional: Lagrangian multipliers (price signal) for spinning reserve demand	An indicator representing the price for spinning reserve (primary/secondary) that drives the operation of the submodels	One timeseries per type of spinning reserve and reserve zone
Submodels	Optional: Lagrangian multipliers (price signal) for inertia demand	An indicator representing the price for inertia demand that drives the operation of the submodels	One timeseries per inertia zone
Submodels	Optional: Lagrangian multipliers (price signal) for the CO_2 constraint	An indicator representing the CO_2 price that drives the operation of the submodels	One timeseries per CO_2 zone
Submodels	Optional: Lagrangian multipliers (price signal) for the heat-constraints	An indicator representing the heating price that drives the operation of the generation units supplying heating energy	One timeseries per heat-ID (see section 1.4.1)
SSV	A cutting plane model for the Lagrangian dual of the model	A cutting plane model of the concave dual. It can be employed for quickly hot-starting another EUC run, whenever few changes are made.	A set of triplets for the CP model

Result	Estimated Optimality gap	Estimated Optimality gap	Value
Result	Indicators of success	Provides an indication if maximizing the Lagrangian dual has been succesful or failed for some reason (perhaps the process was stopped early due to max. number of iterations)	Value

3.2 Seasonal storage valuation

The objective of the seasonal storage valuation tool is to provide an accurate account of "the value" that seasonal storage can bring to the system. Indeed such seasonal storage (e.g., cascaded reservoir systems) can be used to store energy over large spans of time and use this "stored" energy when most needed. The actual use may in particular depend on adverse climatic situations (intense cold). But the ability to store the energy may in turn also depend on climatic conditions (e.g., draught). It is therefore clear that such a vision of value should be transferred in an appropriate way to shorter time span tools, such as the EUC model. In turn computing an accurate value intrinsically depends on the value of substitution, and thus ultimately on the EUC tool as well.

The purpose of the seasonal storage valuation tool is to compute an accurate value while accounting at best of whatever vision the EUC model may have.

Model requirements - Inputs

The main input for the seasonal storage valuation tool is the fine description of the set of "seasonal storage assets" to valuate as well as the discretization of time. Since a precise interaction with the EUC layer will take place, the latter tool will also require appropriate data.

Table 12: Required input data for the seasonal storage valuation model

Model	Input	Description	Format
External input	Time horizon	Time horizon to be optimized	Value (number of hours)
External input	Time Steps	Discretization of time horizon (e.g. hourly)	Value or set of values if heterogeneous

External input	Stage information	A finite selection of time instants in the above set of time steps considered to be the beginning of a stage; N.B. The evolution of information is assumed to be such that full knowledge of uncertainty is obtained over the stage $[t_i, t_{i+1})$ at $t = t_i$	A set of selected time instants
External input	Uncertainty information (option 1)	A set of possible realizations for uncertainty covering the time horizon, • Load • Inflows • Intermittent generation Implicitly these uncertainty factors will be assumed to be Markovian	As a collection of time series (indexed by node and scenario number)

External input	Uncertainty information (option 2)	A scenario lattice for the uncertainty (Seasonal storage valuation will employ stochastic dual dynamic programming (SDDP)). This means, that for each uncertainty factor, we dispose of a model from which we can generate the independent increments	An implementable model description
EUC	EUC - data	All data required to run the EUC model	See section 3.1
EUC	EUC - Lagrangian value	Optimal Lagrangian dual value. This information is recovered from a run of EUC triggered by SSV.	A value (indexed by stage and by scenario)
EUC	EUC - Lagrangian Multipliers	Optimal Lagrangian multipliers. This information is recovered from a run of EUC triggered by SSV.	A vector (indexed by stage and by scenario)
EUC	EUC - CP model of Lagrangian dual	Cutting plane model of Lagrangian dual. This information is recovered from a run of EUC triggered by SSV.	A collection of triplets for each cutting plane model (indexed by stage and by scenario)

Model results - Outputs

The main output of the SSV is an estimate of the cost-to-go function, i.e., an estimate of future expected cost given current storage levels.

Table 13: Output of the seasonal storage valuation model

Model	Output	Description	Format
Result & EUC	Cost to go functions	For each stage the SSV model computes a representation of the cost to go functions. This function allows to gauge the balance of consuming resources now against keeping them for the future.	For each stage a set of triplets storing the cutting plane model.
Result	Lower bound	Lower bound on optimal value	Value
Result	Upper bound	Estimated upper bound on optimal value (involves accounting for some confidence interval)	Value
Result	Indicators of success	Provides an indication if solving the SSV has been succesful or failed for some reason (perhaps the process was stopped early due to max. number of iterations)	Value
CEM	Optimal value	Indication of the optimal value at currently invested capacities	Value

CEM	Subgradient	Sensitivity of the optimal value at currently invested capacities (associated with certain dual variables)	Vector

4 Submodels

4.1 Thermal power plants

The *thermal power plant* model describes the operation of (large) conventional power plants directly connected to the transmission grid. This includes:

- Nuclear power plants
- Hard coal power plants
- Lignite power plants
- Gas turbine power plants
- Gas power plants
- Combined cycle power plants
- Oil power plants

An imbalance unit (generating any lack of energy at high cost) will be considered a special case of “Power plant”. This optional unit can be interpreted as a special kind of slack variable. Note that the main difference resides in a potentially special (i.e., non-linear) cost function. For a detailled description of thermal power plants in unit comittment we refer to [2].

Essential for modeling conventional power plants is detailled knowledge on the generation fleet including technical information. This data are collected within a powerplant database which needs to be constructed beforehand. This database includes information on

- Technology/fuel type
- Location indicator (cluster)
- Efficiency (in %)
- Electricity generated when offline. This fixed data can be arbitrary, but would usually be less than or equal to zero
- Minimal generation limit (time series)
- Maximal generation level (time series), e.g. to account for maintenance/failure

- Possibility to generate primary/secondary spinning reserve
- Amount of inertia provided by the generator when online (time series; in MWs/MW)
- Ramp-up/ramp-down slopes (time series; in MW/min or MW/h)
- Start-up/shut-down cost (time series; in €)
- Min up/min down time (minutes or hours)
- Optional: Emission rates for an arbitrary set of pollutants (e.g., CO_2, NO_x, SO_2, small-particles). Each specific pollutant is indicated by a specific identifier (time series; in tonne/MWh)
- Optional: Coefficient of heat (ratio of electric and thermal output) and connection to district heat or process heat via heat-demand-ID (see section 4.3)

The generation costs of the powerplants depend on the plant-specific parameters and might be calculated via:

- A cost function of the type $a+c^T p+0.5p^T Qp$, with Q a diagonal matrix, considering a proportional cost of generation (in €/ MWh), a fixed cost of generation (in €/ min) and potentially a quadratic cost term (in €/ ($(MW)^2 h$))
- A cost function defined as $\max_{i=1,\dots,k} \left\{a_i + c_i^{\mathsf{T}} p\right\}$ (fixed cutting plane model of a convex cost function), with a set of linear functions specified as $a_i + c_i^{\mathsf{T}} p$ (with a_i in €/ h and c_i in €/ MWh). This modelling options allows us to account for arbitrary convex cost functions, with an *a priori* approximation. The (convex) quadratic cost function is a special case of a convex function but does not require any approximations. In particular this option can be used to finely model imbalance costs (e.g., [3, Fig 2.2.]).

Besides electricity, power plants might also deliver heat (e.g., for supplying district heating or industrial process heating). Thus power plants might also have a heat-ID as a parameter, connecting the power plant model with the heat submodel (Section 4.3). Moreover the connection to a certain district heating grid or industrial heat demand implies that all power plants (and generation units/thermal storages from the heat submodel) connected to the same process (determined by the heat-ID) should be considered jointly.

Model requirements - Inputs

The *power plant model* is a submodel for unit commitment, and thus needs the lagrangian multipliers as input (which can be seen as 'price signals').

Table 14: Required input data for the power plant model

Model	Input	Description	Format
External input	Plant specific parameters	See above description	Powerplant "Matrix"
EUC	Lagrangian multipliers (price signal) for electricity demand	An indicator representing the price for electricity demand that drives the operation of the power plants	One timeseries per cluster
EUC	Lagrangian multipliers (price signal) for spinning reserve demand (primary/secondary)	An indicator representing the price for spinning reserve demand (primary/secondary) that drives the operation of the power plants	One timeseries for each type of spinning reserve (primary/secondary) and reserve zone
EUC	Lagrangian multipliers (price signal) for inertia	An indicator representing the price for providing inertia (at the node of the plant)	One timeseries per inertia zone
EUC	Optional: Lagrangian multipliers (price signal) for the CO_2 constraint	An indicator representing the CO_2 price (at the node of the plant) that drives the operation of the power plants	One timeseries per CO_2 zone

EUC	Optional: Lagrangian multipliers (price signal) for the heat-constraints	An indicator representing the heating price that drives the operation of the generation units supplying heating energy	One timeseries per heat-ID (see section 1.4.1)

Model results - Outputs

The model determines the operation schedules per power plant based on the above price signals.

Table 15: Results of the power plant model

Model	Output	Description	Format
EUC	Power plant operation	Electricity generation provided by the power plants	One timeseries per power plant
EUC	Power plant reserve	Reserve contribution provided by the power plant (primary/secondary)	One timeseries per power plant and per type of spinning reserve
EUC	Power plant inertia	Contribution to intertia provided by the power plant	One timeseries per power plant
EUC	Optional: Power plant heat generation	Amount of thermal energy provided by the power plant	One timeseries per power plant
EUC	Optional: Power plant CO_2 emission	Amount of CO_2 emitted by the power plant	One timeseries per power plant

4.2 Storages

The *storage model* describes the operation of electric storages within plan4res. This includes cascaded reservoir systems as well as batteries. Since these differ in the way they are modelled, they are described in their respective subchapters below.

4.2.1 Hydro storages

To model complex reservoir systems several technical parameters have to be considered. These are divided into reservoir-specific parameters, the hydro links connecting the reservoirs and finally the turbine/pump parameters. The values are collected within a reservoir database, a hydro-link database and a turbine/pump-database.

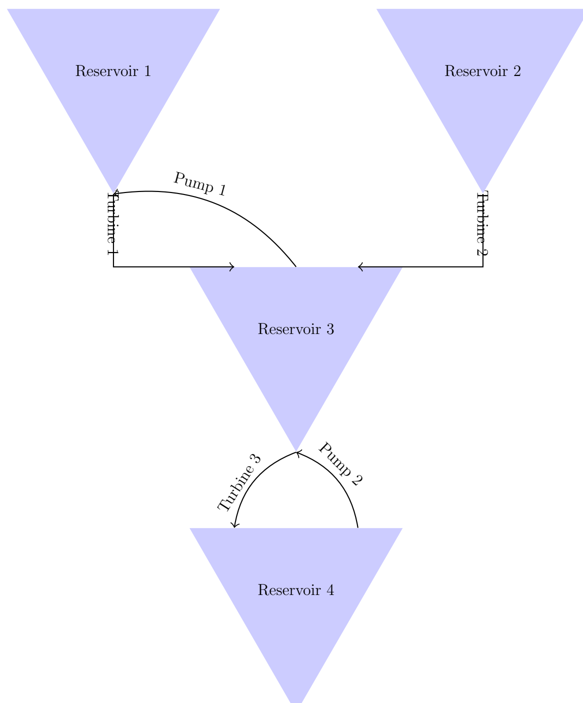


Figure 7: Example of cascading system

The *reservoir database* contains information about:

- Initial volume (single value; in m^3)
- Minimal/maximal volume level (timeseries; in m^3)

- Set of hydro inflows (timeseries; in m^3/s)
- Cutting plane model to value the final volume in the reservoir (timeseries; in €/m^3)

The *hydro link database* contains information about:

- Directed arc from a reservoir to another (the ocean = a reservoir of infinite volume)
- Uphill flow delay (single value; in hours)
- Downhill flow delay (single value; in hours)
- Assigned turbines/pumps (one or several)

The *turbine/pump database* contains information about:

- Cluster where the turbine / pump is located (important for the power balance equations)
- Initial flow rate (might be negative for pumps; single value; in m^3/s)
- Ramp-up/ramp-down slopes (timeseries; in m^3/s)
- Contribution to inertia (single value in MWs/MW)
- Percentage of generation for power supply, primary reserve and secondary reserve. This is considered as external data, without being modeled endogenously. More elaborate models would make the cascaded reservoir subproblems quite challenging and are beyond the scope of plan4res (for details see [4] or [8]).
- Piecewise linear concave cutting plane model linking flow rate to power output

Model requirements - Inputs

Required input data for the hydro storage model are the technical parameters characterizing the assets within the hydro system. Since the hydro storage model is a submodel of the EUC it also gets the lagrangian multipliers as input.

Table 16: Required input data for the hydro storage model

Model	Input	Description	Format
External input	Reservoir database	Technical parameters	Parameter “Matrix”
External input	Hydro link database	Technical parameters	Parameter “Matrix”
External input	Turbine/pump database	Technical parameters	Parameter “Matrix”
EUC	Lagrangian multipliers (price signal) for electricity demand	An indicator representing the price for electricity demand that drives the operation of the turbines/pumps	One timeseries per cluster
EUC	Lagrangian multipliers (price signal) for spinning reserve demand (primary/secondary)	An indicator representing the price for spinning reserve demand (primary/secondary) that drives the operation of the turbines/pumps	One timeseries for each type of spinning reserve (primary/secondary) and reserve zone
EUC	Lagrangian multipliers (price signal) for inertia	An indicator representing the price for inertia	One timeseries per inertia zone

Model results - Outputs

The Model determines the operation schedules per hydro storage.

Table 17: Results of the hydro storage model

Model	Output	Description	Format
EUC	Hydro operation	Electricity generation provided by the turbines/pumps	One timeseries per asset (might be negative due to pumping)
EUC	Hydro spinning reserve	Spinning reserve contribution provided by the turbines/pumps (primary/secondary)	One timeseries per asset (might be negative due to pumping) and per type of spinning reserve
EUC	Hydro inertia	Contribution to inertia provided by the asset	One timeseries per asset (might be negative due to pumping)
EUC	Storage trajectory	The evolution of volumetric contents in each reservoir. In particular, one can thus access the final storage level	One timeseries per reservoir

4.2.2 Battery storages

Battery storages are described by a battery storage database, containing the following information for every storage unit. Small storages are allocated to households and businesses and further aggregated within the energy cells (see section 1.4.1). Central battery storages are directly located within one cluster.

- Initial storage level (single value; in MWh)
- Minimal/maximal storage level (timeseries; in MWh)
- Maximal power intake/outtake (timeseries; in MW)

- Ramp-up/ramp-down slopes for change in power intake/outtake (timeseries; in MW/min)
- Optional: Amount of inertia provided by the battery (timeseries; in MWs/MW)
- Optional: Operational costs proportional to intake/outtake (timeseries; in €/MW)

Model requirements - Inputs

The battery storage model needs technical parameters as input data. Furthermore the lagrangian multipliers representing the electricity price are input data as the battery storages optimize with respect to this price.

Table 18: Required input data for the battery storage model

Model	Input	Description	Format
External input	Battery storage database parameters	Technical parameters (see above)	Parameter "matrix"
EUC	Lagrangian multipliers (price signal) for electricity demand	An indicator representing the price for electricity demand that drives the operation of the battery storages	One timeseries per cluster
EUC	Lagrangian multipliers (price signal) for spinning reserve demand (primary/secondary)	An indicator representing the price for spinning reserve demand (primary/secondary) that drives the operation of the battery storages	One timeseries for each type of spinning reserve (primary/secondary) and reserve zone
EUC	Optional: Lagrangian multipliers (price signal) for inertia	An indicator representing the price for inertia that drives the operation of the battery storages	One timeseries per inertia zone

Model results - Outputs

The battery storage model determines the schedules of the battery storages with respect to the lagrangian multipliers.

Table 19: Results of the battery storage model

Model	Output	Description	Format
EUC	Battery storage operation	Electricity generation/consumption of the battery storages	One timeseries per battery (might be negative due to consumption)
EUC	Battery spinning reserve	Spinning reserve contribution provided by the battery (primary / secondary)	One timeseries per battery (might be negative due to consumption) and per type of spinning reserve
EUC	Optional: Battery inertia	Contribution to inertia provided by the battery	One timeseries per battery (might be negative due to consumption; "zero" timeseries if not used)
EUC	Storage trajectory	The evolution of the energetic contents in the battery. In particular, one can thus access the final storage level	One timeseries per battery

4.3 Heat

Future energy system will use sector coupling technologies to connect different energy technologies (electricity, heat, electric mobility). Especially surplus energy from RES can be used for heating purposes and stored by heating storages. The heat submodel covers the interconnection of the electricity and the heat sector by considering electricity and heat demand as well as electricity generation technologies, heating technologies and thermal storages. Both are pre-defined within the energy cells (Section 1.4.1) that are input to the heat submodel and include:

- Combined heat and power units (Gas, biomass)
- Heatpumps (low temperature, high temperature)
- Peak load boiler (Gas, biomass, oil)
- Coal furnance
- High temperature furnance (Gas, biomass, hard coal, electricity)
- Solar thermal units
- Power-to-heat units
- Thermal storages

As described in the power plant model (Section 4.1) heat demand might also be supplied by power plants, that are allocated by means of the heat-ID defined in the power plant database.

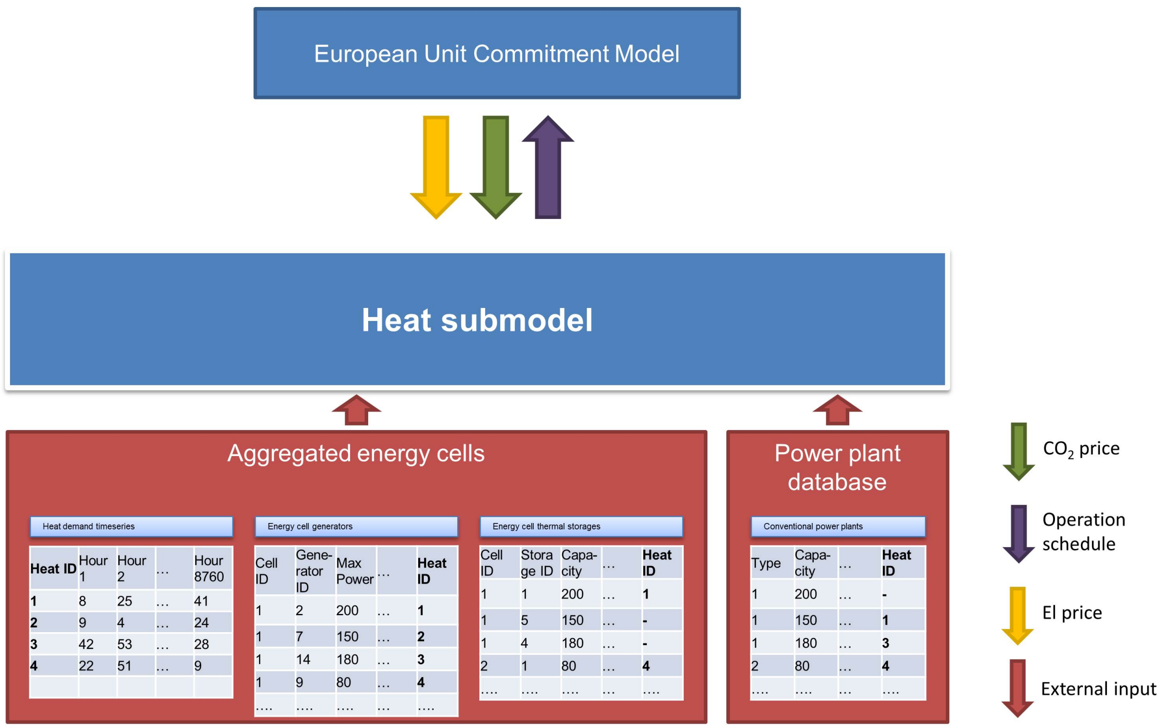


Figure 8: Model interconnection of the heat submodel

Model requirements - Inputs

The *heat submodel* is a submodel of the EUC model. Getting a price signal as input from the EUC the model optimizes the operation of energy cells against this price. Additionally, a CO_2 price as well as a heating price given by the EUC can be used as optimization criteria.

Table 20: Required input data for the heat model

Model	Input	Description	Format
External input	Energy cells (Section 1.4.1)	Pre-defined energy cells consisting of an aggregated heat-demand timeseries and assigned generation technologies (distributed heating-units, powerplants, thermal storages)	Energy cell matrix
EUC	Lagrangian multipliers (price signal) for electricity demand	An indicator representing the price for electricity demand that drives the operation of the energy cells	One timeseries per cluster
EUC	Optional: Lagrangian multipliers (price signal) for the CO_2 constraint	An indicator representing the CO_2 price that drives the operation of the energy cells	One timeseries per CO_2 zone
EUC	Optional: Lagrangian multipliers (price signal) for the heat-constraints	An indicator representing the heating price that drives the operation of the generation units supplying heating energy	One timeseries per heat-ID (see section 1.4.1)

Model results - Outputs

The Model determines the schedules of the aggregated distributed generation units within each energy cell and delivers these to the EUC, which coordinates the price adjustment process.

Table 21: Results of the heat model

Model	Output	Description	Format
EUC	Aggregated energy cell operation	Residual electricity generation within each energy cell (might be negative due to operation of heatpumps and power-to-heat units that need electricity to operate)	One timeseries per aggregated energy cell
EUC	Optional: Energy cell CO_2 emission	Amount of CO_2 emitted by the energy cell	One timeseries per energy cell

4.4 E-mobility

The *E-mobility model* provides the aggregated operation of the e-mobility fleet based on a predefined scenario framework. Electric vehicles can be seen as flexible loads (power-to-vehicle) and as storage technologies (vehicle-to-grid). A controlled recharging of batteries and grid feed-in can offer additional flexibilities to the energy system. Electric vehicles are allocated to the households and businesses of the registeres and defined as aggregated storages per energy cell.

Model requirements - Inputs

The e-mobility model acts as an aggregated electric storage that optimizes against the price signal from the unit commitment. The flexibility is limited by driving profiles (a predefined battery discharge that must be guaranteed to be available) as well as minimal and maximal power.

Table 22: Required input data for the e-mobility model

Model	Input	Description	Format
EUC	Lagrangian multipliers (price signal) for electricity demand	An indicator representing the price for electricity demand that drives the operation of the electric mobility storages	One timeseries per cluster
External input	Energy cells - driving profiles	Aggregated electric mobility driving profiles	One timeseries per energy cell
External input	Energy cells - maximum power	Aggregated electric mobility maximum power input/outtake	Two values per energy cell (maxmimum input/outtake)

Model results - Outputs

The Model determines the consumption and provision of electricity by the electric vehicles per energy cell. Relevant output for the EUC is the residual electric demand (might be negativ due to a surplus of electricity supply by electric vehicles).

Table 23: Results of the e-mobility model

Model	Output	Description	Format
EUC	Residual electric demand	Residual electric demand of consumption and provision of electricity	One timeseries per aggregated energy cell

4.5 Centralized demand response

Model Overview

The *centralized demand response model* consists of the adjustment of a flexible consumption resulting from the aggregation of various appliances connected to the transmission grid such as to minimize the system costs.
We consider two types of flexible consumption models arising at each cluster i. In the sequel, we will omit the index i corresponding to a specific cluster such as to simplify the notations.

1. *Shifting electricity consumption model*: inside each stage of the short-term problem, $[t_p, t_{p+1}]$ (corresponding for instance to one week), we specify some given periods, where a given volume of energy demand is flexible in the sense that the load profile can be chosen in order to optimize the system costs as long as the total energy consumption on each specified period remains fixed. This is considered as a deterministic storage problem, on each short-term stage $[t_p, t_{p+1}]$.

2. *Erasing electricity consumption model*: inside the mid-term horizon, $[\theta, T+\theta]$ (corresponding for instance to one year), a given quantity of energy can simply be removed from the demand profile all along the mid-term horizon $[\theta, T+\theta]$. This is considered as a stochastic storage problem.

More precisely, in the *shifting model* we have to specify the periods on which some fixed energy consumptions are required:

- we define n_I periods corresponding to non overlapping time intervals $I_j = [\tau_j^{init}, \tau_j^{final})$ inside a stage $[t_p, t_{p+1}]$, for $j = 1, \cdots, n_I$;

- for each $j = 1, \cdots n_I$, we define an energy need, E_j that should be consumed on the time interval, I_j;

The aim is to compute on each period $j = 1, \cdots, n_I$, the optimal load allocation $\ell_j := (\ell_{j,q})$ of the energy E_j over time steps $t \in I_i$ such that the energy constraint

$$\sum_q \ell_{j,q} = E_j \ ,$$

is fulfilled while minimizing the cost of electricity induced by the price signal (provided by the Lagrangian coordinator) and satisfying some power bounds $\underline{\ell}_{j,q} \leq \ell_{j,q} \leq \overline{\ell}_{j,q}$.

The *erasing model* is simply modeled as a seasonal storage. It is characterized by a given quantity of energy that could be erased from the demand on the mid-term horizon and some bounds on power that could be injected in the system at each time step.

The demand response is not supposed to provide inertia nor ancillary services to the system.

Model requirements - Inputs

This constitutes a sub-model for the unit commitment. The demand response model only takes into account the price signal related to power supply, since no ancillary services can be provided. The set of time intervals and related energy $(I_j, E_j)_{j=1,\cdots n_I}$, should be carefully fitted on historical data, for each cluster.

Table 24: Required input data for the centralized demand response model

Model	Input	Description	Format
EUC	Lagrangian multipliers (price signal) for electricity demand	An indicator representing the price for electricity demand that drives the operation of the centralized demand response	One timeseries per cluster
External input	*Shifting model* Periods definition	Definition of intervals $I_j = [\tau_j^{init}, \tau_j^{final})$ for $j = 1, \cdots, n_I$	Two n_I scalars $(\tau_j^{init})_j$ and $(\tau_j^{final})_j$, for each cluster
External input	*Shifting model* Maximal and minimal adjustment	Maximal possible demand adjustment per time step, for each time interval, for $j = 1, \cdots, n_I$	Two n_I timeseries $(\underline{\ell}_{j,q})_{j,q}$ and $(\overline{\ell}_{j,q})_{j,q}$, for each cluster,
External input	*Shifting model* Energy need	Energy E_j that should be delivered on the time interval I_j, for $j = 1, \cdots, n_I$	n_I scalars $(E_j)_j$ per cluster

External input	*Erasing model* Energy volume	Energy that could be erased from the demand	One scalar per cluster
External input	*Erasing model* Erasing bounds	Upper and lower bounds for power reduction by the *erasing model*	Two time series per cluster

Model results - Outputs

The *shifting model* determines the electricity demand shift answering to the price incentives, while fulfilling the energy constraint on each interval. The *erasing model* determines the electricity demand erased answering to the price incentives, while fulfilling the energy volume constraint on the whole period.

Table 25: Results of the centralized demand response model

Model	Output	Description	Format
EUC	*Shifting model* Residual demand	Demand adjusted by shifting electric demand	One timeseries per cluster
EUC	*Erasing model* Residual demand	Demand adjusted by erasing electric demand	One timeseries per cluster

4.6 Intermittent generation

Model Overview

The *intermittent generation* model provides the operation of intermittent renewable energy resources which are connected to the transmission or subordinate distribution grids (Wind farms, solar parks) within the plan4res framework.
This model relies mainly on historical data of local generation of wind and solar at each node of the transmission or subordinate distribution grid. These data are used to develop normalized generation profiles for wind and solar generation.

Furthermore, data is required concerning the ability of wind generators to contribute to the system inertia. Indeed, the unit commitment model has to ensure that a minimum inertia level is provided by the set of running units.

Intermittent generation could also contribute to ancillary services by providing only a proportion, say 80%, for instance, of the possible generation as power and the 20% left as primary or secondary reserve to compensate underproduction. This ratio of production allocated to power and to reserve would result from an optimization procedure based on dual power prices and dual reserve prices provided by the *Lagrangian coordinator*. A basic version of the model could consider this ratio to be fixed to 100% for power and 0% for reserve. In this latter version, the actual generation corresponds to the possible generation due to a given generation profile. More generally, we may also add the possibility of curtailing renewable generation in our model when the system balance requires it. The available production would then be shared into four proportions: power, primary reserve, secondary reserve and curtailment. The sub-problem will then consist in providing the ratio, $(\alpha_t, \beta_t^1, \beta_t^2, \gamma_t)_t$. Of course, one has only to determine three elements, since the fourth one is automatically given by the condition $\alpha_t + \beta_t^1 + \beta_t^2 + \gamma_t = 1$. We may also add some lower and upper bound conditions on this ratio $\underline{\alpha}_t \leq \alpha_t \leq \bar{\alpha}_t$, $\underline{\beta}_t^1 \leq \beta_t^1 \leq \bar{\beta}_t^1$, $\underline{\beta}_t^2 \leq \beta_t^2 \leq \bar{\beta}_t^2$, $\underline{\gamma}_t \leq \gamma_t \leq \bar{\gamma}_t$. For instance, setting $\underline{\gamma}_t = \bar{\gamma}_t = 0$ will eliminate any possibility of curtailment.

Model requirements - Inputs

The *Intermittent generation model* poses a sub-model for the EUC.
Units optimize their gains against price signals (to determine the proportion of power, primary and secondary reserves and curtailment) given by the EUC.

Table 26: Required input data for the intermittent generation model

Model	Input	Description	Format
External input	Generation profile	The generation profile per generation unit	One timeseries per generation unit
External input	Upper and Lower bounds on the ratio	$(\underline{\alpha}_t)_t$, $(\bar{\alpha}_t)_t$, $(\underline{\beta}_t^1)_t$, $(\bar{\beta}_t^1)_t$, $(\underline{\beta}_t^2)_t$, $(\bar{\beta}_t^2)_t$, $(\underline{\gamma}_t)_t$, $(\bar{\gamma}_t)_t$	Eight timeseries per generation unit
EUC	Inertia Provision	Possibility and amount of inertia that can be provided by each generation unit	One value per generation unit
EUC	Lagrangian multipliers (price signal) for electricity demand	An indicator representing the price for electricity demand that drives the operation of the renewable generation units	One timeseries per cluster
EUC	Lagrangian multipliers (price signal) for spinning reserve demand (primary/secondary)	An indicator representing the price for spinning reserve demand (primary/secondary) that drives the operation of the renewable generation units	One timeseries for each type of spinning reserve (primary/secondary) and reserve zone

EUC	Lagrangian multipliers (price signal) for inertia	An indicator representing the price for providing inertia	One timeseries per inertia zone

Model results - Outputs

The Model determines the aggregated schedules of the intermittent generation units given by the input data. Furthermore the model provides the ratio of power supply, reserve supply and curtailment for each timestep and generation unit.

Table 27: Results of the intermittent generation model

Model	Output	Description	Format
EUC	Aggregated generation profie	Aggregated generation profile per generation unit	One timeseries per generation unit
EUC	Generation partitioning	Ratio of generation for power supply, primary/secondary reserve and curtailment	Four values [%] per generation unit
EUC	Inertia	Contribution of the renewable generation units for providing inertia	One timeseries per generation unit

4.7 Distributed generation

Model Overview

The *distributed generation model* provides the operation of distributed electricity generation units connected to the distribution grid within the plan4res framework. This includes mainly PV roof systems, but, if not considered within the heat submodel (Section 4.3), biomass might be modeled here too as a thermal unit which does not have CO_2 emissions. Just like generation units connected to the transmission network, these distributed units may also provide ancillary services (primary and secondary reserve) besides power supply. The model for distributed generation is similar to the one described for intermittent generation connected to the transmission network. The model provides at each step of time the ratio of four elements $(\alpha_t, \beta_t^1, \beta_t^2, \gamma_t)$, such that $\alpha_t + \beta_t^1 + \beta_t^2 + \gamma_t = 1$ and corresponding to the proportion of the production profile allocated respectively to power demand, primary reserve, secondary reserve and curtailment.

However, distributed electricity generation units are strongly constrained by the electricity distribution network (Section 5.2) because of the limited capacity of the distribution grid. This additional constraint will be integrated into the model via two specific features

1. At the investment level: to evaluate the cost of increasing the distributed generation capacity in a given cluster i, one needs to evaluate the cost of increasing the maximum capacity, $\bar{P}_i$, of the distribution grid in cluster i.

2. At the operational level: the absolute value of the distributed production injected in cluster i, p_i^{dist}, minus the distributed flexible and non-flexible demand consumed in cluster i, $d_i^{dist} + D_i^{dist}$ should fulfill the power constraint from the distribution grid capacity at each time step t:

$$|p_{i,t}^{dist} - (d_{i,t}^{dist} + D_{i,t}^{dist})| \leq \bar{P}_i \ . \tag{5}$$

In this simple model, the integration of distributed generation requires either to invest into the distribution network, either to increase the capacity of flexible demand, in order to satisfy the grid distribution constraint (5).

Model requirements - Inputs

The *distributed generation model* is a submodel of the European unit commitment model. Getting a price signal as input from the EUC the model optimizes the unit operation with respect to this electricity price.

Table 28: Required input data for the intermittent generation model

Model	Input	Description	Format
External input	Generation profile	The generation profile per generation unit	One timeseries per generation unit
External input	Upper and Lower bounds on the ratio	$(\underline{\alpha}_t)_t$, $(\bar{\alpha}_t)_t$, $(\underline{\beta}^1_t)_t$, $(\bar{\beta}^1_t)_t$, $(\underline{\beta}^2_t)_t$, $(\bar{\beta}^2_t)_t$, $(\underline{\gamma}_t)_t$, $(\bar{\gamma}_t)_t$	Eight timeseries per generation unit
EUC	Inertia Provision	Possibility and amount of inertia that can be provided by each generation unit	One value per generation unit
EUC	Lagrangian multipliers (price signal) for electricity demand	An indicator representing the price for electricity demand that drives the operation of the renewable generation units	One timeseries per cluster
EUC	Lagrangian multipliers (price signal) for spinning reserve demand (primary/secondary)	An indicator representing the price for spinning reserve demand (primary/secondary) that drives the operation of the renewable generation units	One timeseries for each type of spinning reserve (primary/secondary) and reserve zone
EUC	Lagrangian multipliers (price signal) for inertia	An indicator representing the price for providing inertia	One timeseries per inertia zone
External input	Maximal grid capacity	Maximal grid capacity at each node; $(\bar{P}_i)_i$	One value at each node

Model results - Outputs

The Model determines the schedules of the aggregated distributed generation units for the EUC.

Table 29: Results of the distributed generation model

Model	Output	Description	Format
EUC	Distributed generation units schedules	Generation schedules of the distributed generation units	One timeseries per generation unit
EUC	Generation partitioning	Ratio of generation for power supply, primary/secondary reserve and curtailment	Four values [%] per generation unit
EUC	Inertia	Contribution of the renewable generation units for providing inertia	One timeseries per generation unit

4.8 Distributed load management

Model Overview

One motivation of integrating this model is to be able to compare the interest of investing into the distribution network rather than to develop distributed load management facilities in order to relax the distribution constraint (5) and to be able to integrate more distributed intermittent generation. More specifically, the distributed production injected at cluster i, P_i^{dist} (see Section 4.7), minus the distributed flexible demand consumed at cluster i, d_i^{flex}, should fulfill the power constraint from the distribution grid capacity at each time step t:

$$|p_{i,t}^{dist} - (d_{i,t}^{flex} + D_{i,t}^{dist})| \leq \bar{P}_i \ .$$

The model for distributed load management is very similar to the *shifting electricity consumption model* introduced for the centralized demand management (Section 4.5). However, the model will differ mainly because the shifted demand at cluster i will constitute a flexible demand, d_i^{flex}, that will be involved in the distribution network capacity constraint at cluster i, and eventually contribute to relax this constraint.

Model requirements - Inputs

The distributed load management constitutes a sub-model for the unit commitment. It consists of the adjustment of a flexible consumption resulting from the aggregation of various appliances (e.g. air conditioning, water heaters, electric vehicles charging) connected to the distribution grid such as to minimize the system costs. The distributed load management model only takes into account the price signal related to power supply, since no ancillary services can be provided. The set of periods and related energy $(P_j, E_j)_{j=1,\cdots n_I}$ should be carefully fitted on historical data.

Table 30: Required input data for the distributed load management model

Model	Input	Description	Format
European unit commitment	Lagrangian multipliers (price signal) for electricity demand	An indicator representing the price for electricity demand that drives the usage of distributed load management	One timeseries per cluster

External input	*Shifting model* Periods definition	Definition of intervals $I_j = [\tau_j^{init}, \tau_j^{final})$	Two n_I time series $(\tau_j^{init})_{j=1,\cdots,n}$ and $(\tau_j^{final})_{j=1,\cdots,n}$, for each cluster
External input	*Shifting model* Maximal and minimal adjustment	Maximal possible demand adjustment per time step	Two n_I timeseries $(\underline{\ell}_{j,q})_{j,q}$ and $(\bar{\ell}_{j,q})_{j,q}$, for each cluster
External input	*Shifting model* Energy need	Energy E_j that should be delivered on time interval I_j, for j=1,...,n_I	n_I scalars $(E_j)_j$ per cluster

Model results - Outputs

The *shifting model* determines the electricity demand shift answering to the price incentives, while fulfilling the energy constraint on each interval.

Table 31: Results of the distributed load management model

Model	Output	Description	Format
EUC	*Shifting model* Residual demand	Demand adjusted by shifting electric demand	One timeseries per cluster

4.9 Distributed storage

Model Overview

The *distributed storage model* differs mainly from the “centralized” storage model in view of the considered storage cycle. Indeed one of the main assumptions is, that distributed storages need not be optimized over long time horizons and thus can be considered locally in time on one stage $[t_p, t_{p+1}]$ (e.g. one week) of the short-term problem. Consequently, it constitutes a deterministic storage problem.

Moreover the distributed storage has to fulfill the local constraints related to the distribution grid on each cluster i, at each time step t:

$$|P_{i,t}^{dist} + s_{i,t}^{dist} - (d_{i,t}^{dist} + D_{i,t}^{dist})| \leq \bar{P}_i \ ,$$

where $s_{i,t}^{dist}$ is the injection into the grid of the storage at cluster, i, at time step, t. This energy storage is characterized by the volume of the energy storage S_i and bounds on power injected or withdrawn.

$$-\underline{S}_{i,t}^{dist} \leq s_{i,t}^{dist} \leq \overline{S}_{i,t}^{dist} \ .$$

Model requirements - Inputs

The distributed storage model constitutes a sub-model for the unit commitment. The distributed storage model only takes into account the price signal related to power supply, since no ancillary services can be provided.

Table 32: Required input data for the distributed storage

Model	Input	Description	Format
EUC	Lagrangian multipliers (price signal) for electricity demand	An indicator representing the price for electricity demand that drives the usage of distributed load management	One timeseries per cluster
External input	Storage volume	Volume S_i for each cluster	One scalar per cluster

External input	Power bounds	Bounds $(\bar{S}_{i,t}, \underline{S}_{i,t})$ of injected and withdrawn power for each time step and for each cluster	Two timeseries per cluster

Model results - Outputs

The distributed storage model determines the electricity injected or withdrawn from the storage into the distribution grid at each cluster answering to the price incentives, while fulfilling the energy constraint S_i and the power bound constraints $(\bar{S}_{i,t}, \underline{S}_{i,t})$.

Table 33: Results of the distributed storage model

Model	Output	Description	Format
EUC	Distributed storage operation	Injected and withdrawn power $((s_{i,t})_{i,t})$ at each time step and cluster	One timeseries per cluster (might be negative due to withdrawal)

4.10 Power-to-gas

Model Overview

The *power-to-gas model* provides the operation of central power-to-gas units. Power-to-gas provides flexibility to the electricity system by connecting the electricity sector with the gas sector. Thus the power-to-gas units optimize with regard to the lagrangian multipliers (representing a price signal) given by the EUC. The operation of power-to-gas units is limited by technical parameters given by a power-to-gas database, including:

- Location within the transmission grid (since they also provide flexibility for transmission grid redispatch calculations)
- Maximum power
- Efficiency
- Optional: Minimum power (e.g. due to industrial demand)

Model requirements - Inputs

The *power-to-gas model* is a submodel of the EUC an thus optimizes against the given price signal.

Table 34: Required input data for the power-to-gas model

Model	Input	Description	Format
External input	Power-to-gas database (see above)	Power-to-gas units and parameters	Parameter "Matrix"
EUC	Lagrangian multipliers (price signal) for electricity demand	An indicator representing the price for electricity demand that drives the operation of the electric mobility storages	One timeseries per cluster

Model results - Outputs

The Model determines the operational schedules of the power-to-gas units with respect to the given electricity price.

Table 35: Results of the power-to-gas model

Model	Output	Description	Format
EUC & Transmission grid calculation model	Power-to-gas operation schedule	Electricity used by power-to-gas units	One timeseries per unit

5 Supplemental Models

Besides the investment layer and the scenario valuation layer, there are additional models, that cover further aspects of the energy system.

These models include:

- Transmission grid clustering
- Transmission grid calculations (Powerflow, Redispatch)
- Gas grid calculations
- Distribution grid cost curves generation

Since these models are also strongly connected to specific (sub-)models of the scenario valuation layer, they are also described here.

5.1 Gas network

Model Overview

The gas network model is used to determine the transport capacity of the European gas network. In the context of this project, the connection of the electricity sector with the gas sector by power-to-gas is of particular interest as the gas grid allows the storage and transport of energy. It is expected that this connection adds flexibility to the energy system as a whole. The gas network model can serve two main purposes:

- The verification of gas transport requests, e.g., resulting from the *power-to-gas model*.
- The computation of limits to the application of power-to-gas implied by transport capacities of the gas network. These limits could then be used as input to the *power-to-gas model*.

In traditional power-to-gas models, the capacity of the gas network is typically not considered and assumed infinitely large. The output of this model is the in- and outflow vector, flow values through pipelines, configurations of the active network devices such as compressors, and the pressure distribution in the network.

The gas network model will be a stationary model, i.e., it does not consider dynamic effects of gas transport and in particular the delay of input of gas by power-to-gas and the output at some later point. One consequence is that the resulting gas inflow and gas extraction will always be in balance. We therefore assume the availability of gas storages capacities at the consumers to temporally decouple the generation of gas by power-to-gas and the consumption, e.g., by gas power plants.

Generally, apart from the network model, a demand profile in terms of limits on in- and outflows of gas into the network has to be given. The transport volume induced by power-to-gas is expected to be rather small compared to the base load already present in the network without considering power-to-gas. To account for this, other sources and consumers of gas should be included in the demand profiles. Regarding power-to-gas, the model can either be used to verify that a certain demand profile which results from an upstream model can be realized in the gas network or it can be used to determine maximum capacities the gas network can provide at power-to-gas facilities. In the latter case, the demand profiles will not be fixed at the sources and sinks effected by power-to-gas.

Optionally, the model can be extended to included investment decisions in power-to-gas units (as sources) or gas power plants (as sinks) which are placed at candidate locations.

The output of the model is then an optimal investment strategy (given a certain cost function) and the resulting gas flow.

Model requirements - Inputs

Table 36: Required input data for the gas network model

Model	Input	Description	Format
External input	Description of gas network	Technical parameters of pipelines, compressors, etc.	Format from the GasLib `gaslib.zib.de`
External input	Base demand profile	Gas transport situation independent of power-to-gas	Demand vector
Power-to-gas model	Transport request from power-to-gas	The additional transport request by power-to-gas	Demand vector
External input	Flow limits and objective values	Models the flexibility the gas network can provide to power-to-gas facilities	Vectors
External input	Power-to-gas database	Technical parameters of power-to-gas units	Parameter "Matrix"
External input	Power plant database	Technical parameters of gas power plants	Parameter "Matrix"

Model results - Outputs

As a result the gas network model determines the gas flow and pressure within the gas grid. Additionally limitations regarding the schedules of the power-to-gas units and the gas fired power plants can be provided to the upstream transmission grid operation model and considered there as additional operational constraints.

Table 37: Results of the gas network model

Model	Output	Description	Format
Result	Flow distribution across the network	Resulting gas flow within the gas grid	One value per node in the network
Result	Pressure distribution across the network	Resulting pressure within the gas grid	One value per node in the network
Transmission grid operation model	Power-to-gas limitations	Limitation for gas feedin at each power-to-gas unit	One value per power-to-gas unit
Transmission grid operation model	Gas power plant limitations	Limitation for gas consumption at each gas power plant	One value per gas power plant
Result	Potentially decision which facilities to open	Idicator if facility is opened	One value per condidate location

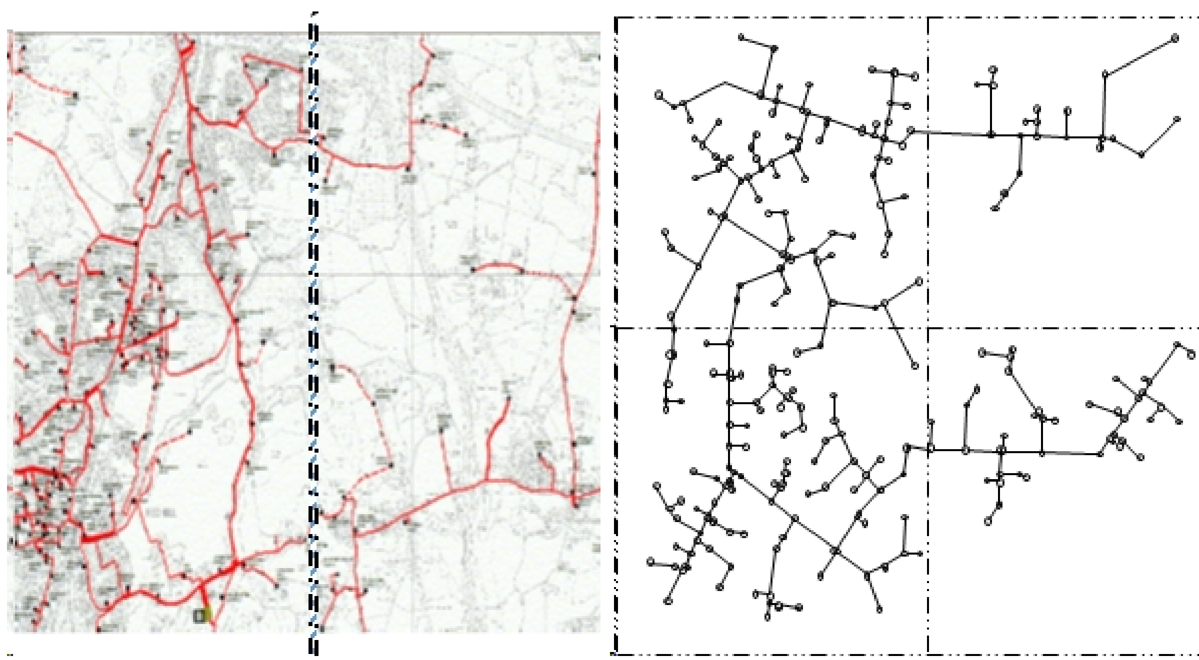
5.2 Electricity distribution model

Model Overview

Increased penetration of distributed energy ressources (DER) and participation of distributed flexibility in enhancing and providing balancing services for national electricity systems requires the impact of deploying these resources on the local distribution networks to be measured. In this context, the electricity distribution modelling work in plan4res has the primary objective to provide the reinforcement cost function of electricity distribution networks. This cost function is used to measure the impact of the increased installed capacity of RES and the use of demand response or distributed energy storage. The reinforcement cost functions will then be used by plan4res optimisation models, e.g. the EUC model to optimize and coordinate the deployment of various scales of flexibility and energy resources in the system. More precisely, as explained in section 2.1.2 these cost functions, related to investments on the distribution grid, are used to achieve a tradeoff between investments in the distribution grid capacity and operation costs. Indeed, as described at Section 3.1 each cluster (n) can be viewed as a couple of two electrical nodes:

- A *transmission electrical node*
- A *distribution electrical node*

Each node is characterized by a specific demand and some specific generator units. The specificity of the distribution electrical node relies on the fact that it is exclusively connected to the transmission electrical node n. The capacity of this connection constitutes a constraint involved in the EUC problem that can be relaxed by investments in the distribution grid. This will enable the optimisation module to find the whole-system solution which balances the national and local objectives when deploying the distributed resources. Modelling a large number of specific real distribution networks in Europe would be impractical and inefficient as distribution networks vary in topology, capacity, technologies (e.g. different types of transformers, circuits, etc.), the spatial distribution of demand and generation, etc. Moreover, the availability of such detailed and granular data is also limited particularly for the Low voltage (LV) networks where a substantial share of the distributed flexibility resources will be connected to. In order to address this problem, the work focuses on developing statistically resemblance of distribution network models. The representative network approach that delivers various types of generic electricity distribution network models resembling the topology, load density, branching intensity of urban, semi-urban, semi-rural, and rural distribution systems in different parts of Europe will be used to generate a few networks that statistically represent the regional distribution network characteristics. The main functionality of this module is to enable coordinated actions and the use of DER connected at the distribution level for the national/pan-European system objectives (e.g. the use of load management, assessment of electrification and/or deployment of distributed generation such as onshore wind farms or PV). The functionality specification captures a range of critical parameters that need to be considered during the modelling of the systems, e.g. the voltage levels and the use of smart voltage control. They are then mapped to match the overall number of customers, network length and number of transformers. An example of representative distribution network model is given in figure 9. The left diagram shows the real topology of the network capturing both high-density urban systems and low-density rural system; the right diagram shows the topology of the representative network model which also covers both urban and rural systems as in the real networks.



(a) Real network (b) Representative network

Figure 9: Comparison between the topology of real network and the associated representative network model

The EUC requires the distribution network reinforcement cost to be expressed as a function of peak load, driven by electrification of transport and heat sectors, or peak reverse flow contributed by distributed generation and distributed storage in a given distribution system. The cost function will also depend on the flexibility of the distribution system to maximise its latent capacity, e.g. by the use of smart grid approaches as well as different levels of penetration of electric vehicles and heat pumps as well as their operating regimes. This is informed by detailed modelling of representative networks and the control flexibility (for example, the use of active voltage control to solve the voltage problems and maximise the utilisation of system capacity). First, different types of distribution networks based on the statistical models of urban, semi-urban, semi-rural and rural networks will be developed using this module, which is based on a multi-stage fractal network modelling approach. Second, optimal power flow or load-flow studies will be carried out exploring the possible operating conditions of the system. Optimal system reinforcement, if necessary, will be modelled and determined to enable the development of the cost function. Given that the network reinforcement cost function is likely to be non-linear and lumpy, curve-fitting approaches can be used to generate a piecewise, linear approximation of the cost function proposed for the whole-system model formulation. Simulations for different load will be conducted to calculate network flows and voltages that are used to identify the assets that need upgrading which defines the overall upgrade cost of electricity distribution network. The reinforcement cost curves as a function of peak demand can be derived as a result. In order to implement cost curves into the EUC model linear piecewise cost curves are derived with the desired number of linear segments. Cost curves will also consider reverse power flow. An example is given in Figure 10. Sensitivity studies on the distribution network cost function can be carried out by taking different input assumptions on the cost of distribution circuits, characteristics of networks (rural/semi-urban/urban), operation mode (active or passive voltage control), etc. In Figure 10, the network reinforcement cost is modelled as a function of peak demand. The reinforcement cost at low-voltage (0.4 kV) and medium/high voltage can be calculated as a function of peak demand or reverse power flow driven by distribution grid (DG) output. The impact of implementing smart-grid technologies such as Voltage control (VC) can also be assessed and quantitatively analysed.

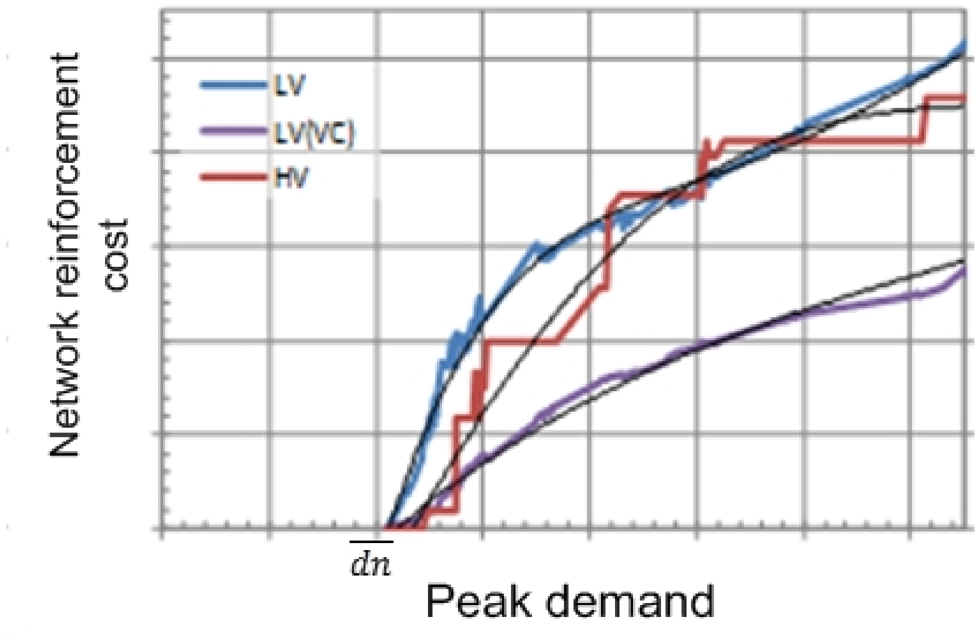


Figure 10: Illustration of distribution reinforcement cost curve functions

Model requirements - Inputs

An overview of the input required by the generic electricity distribution network model is given in Table 38.

Table 38: Required input data for the generic electricity distribution model

Model	Input	Description	Format
External input	Region abbreviation	Specifies the country region from an agreed list	Set of string
External input	Reference peak load	The current peak load of countries selected above	Array
External input	DG characteristics	Types, distribution and profiles of DG outputs	Array

External input	Network control strategy	Passive or active network voltage management	Array
External input	Network data	Data include: total length of circuits, number of transformers, etc. at different voltage levels	Array
External input	Network unit cost data	Data include: unit cost of different types of circuits (cables, overhead lines), unit cost of different transformers, etc.	Array
External input	Control parameters to define the cost function	These include: the range of peak demand that need to be evaluated by the cost curve, type of curve (linear, piece-wise linear, quadratic), number of segmentation	Array

Model results - Outputs

An overview of the output data produced by the generic electricity distribution network model is given in Table 39.

Table 39: Results of the generic electricity distribution model

Model	Output	Description	Format
EUC	List of reinforcement costs	Specifies the costs for upgrading at different levels of peak demand	Array
EUC	List of distribution network cost reinforcement function coefficients	Coefficients for the distribution network cost functions that will be used by other optimisation modules	Array

5.3 Clustering transmission grid

Model Overview

As it will not be possible to represent the transmission network in whole details taking into account all the nodes and transmission lines while conducting yearly European-wide power system simulations with hourly granularity and detailed production units constraints, the *clustering transmission grid model* provides a coarse vision of the network by aggregating some nodes into *nodes clusters* which will be considered as nodes in the simplified model of the grid and is shown exemplary in figure 11. For instance, eHighway clusters could be used to define some *nodes clusters*.

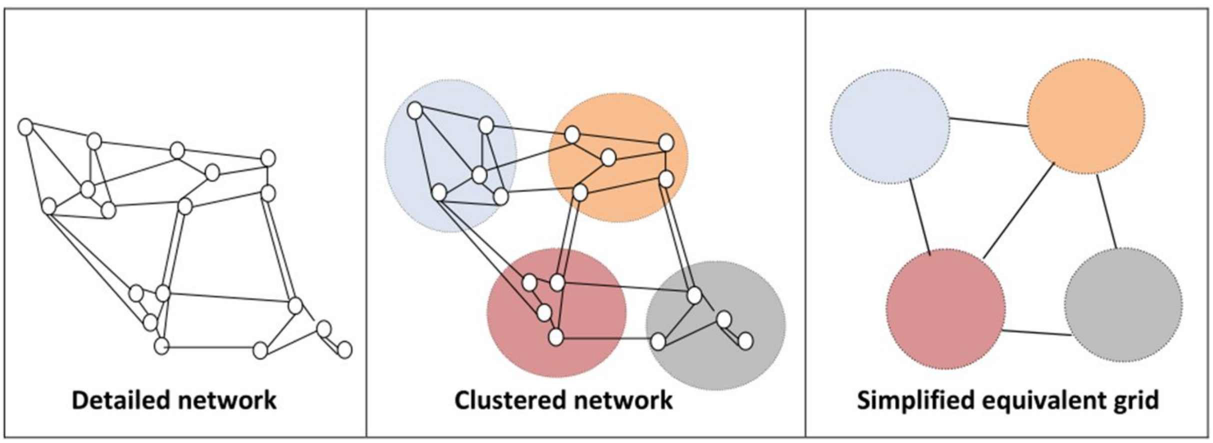


Figure 11: Process of transmission grid clustering

The generation scheduling has to fulfill some limited transmission capacity constraints between the different generation units and loads. One way to take this limitation into account in the EUC model is to use the DC power flow model which is a linearization of the nonlinear AC power flow with a reasonable level of accuracy. In the DC power flow model, a linear relationship between power injections at each node of the grid and active power flows through the transmission lines is established. For instance, this linear relationship could be represented via the PTDF matrix that will constitute an important output of the clustering model. Then, the active power flows are limited by the transmission capacities of the lines between the clusters with constraints of the type $\underline{P}_\ell \leq p_\ell \leq \bar{P}_\ell$ for each line ℓ.

Model requirements - Inputs

The clustering transmission grid model relies on two main input data:

1. The detailled transmission grid on which the clustering is based on (set of nodes and lines linking the nodes including electrotechical parameters)
2. Optional: External boundaries/areas the clusters should be aligned with. This might be further geographical data (e.g., non flexible and flexible consumption, intermittent generation, conventional generation). For instance the NUTS (Nomenclature des unités territoriales statistiques) regions could be used as input.

Table 40: Required input data for the clustering transmission grid model

Model	Input	Description	Format
External input	Detailed transmission grid (nodes, lines, capacities, all relevant data for DC power flow)	Characteristics of the detailed grid needed to run the DC power flow model	List of nodes and lines including technical parameters
External input	Load and generation fleet description (all relevant data for DC OPF on multiple scenario)	Characteristics of the generation plants and demand needed for the DC power flow model	Technical parameters of generation units
External input	Optional: Geographical boundaries	Geographical boundaries/areas the clusters should be aligned with	One polygon per boundary

Model results - Outputs

The Model determines the characteristics of the simplified and aggregated grid (nodes, lines, capacities, and PTDF matrix) needed to consider the DC power flow constraints in the EUC model.

Table 41: Results of the clustering transmission grid model

Model	Output	Description	Format
EUC	Aggregated network (clusters, PTDF matrix, aggregated lines, capacities)	Transmission grid characteristics needed consider DC power flow constraints in EUC	List of nodes and lines of aggregated network

5.4 Transmission grid operation model

Model Overview

The *transmission grid operation model* provides the power flows in the transmission grid resulting from the generation and load patterns determined by the EUC model using an AC formulation of the power flow equations. Based on these power flows congestions will be identified by the N-1 criterion using the line outage distribution factor (LODF) approach. There are different options to clear those congestions like redispatching of power plants or pump storages as well as the curtailment of RES generation. Other options are employing the flexibilities resulting from the coupling of the electricity sector with other energy sectors, e.g. using power-to-gas units to shift the energy transport from the electricity grid to the gas grid. Within plan4res the use of power-to-gas units is restricted by the operational limits of the gas grid, which are a result of the gas network model (Section 5.1).

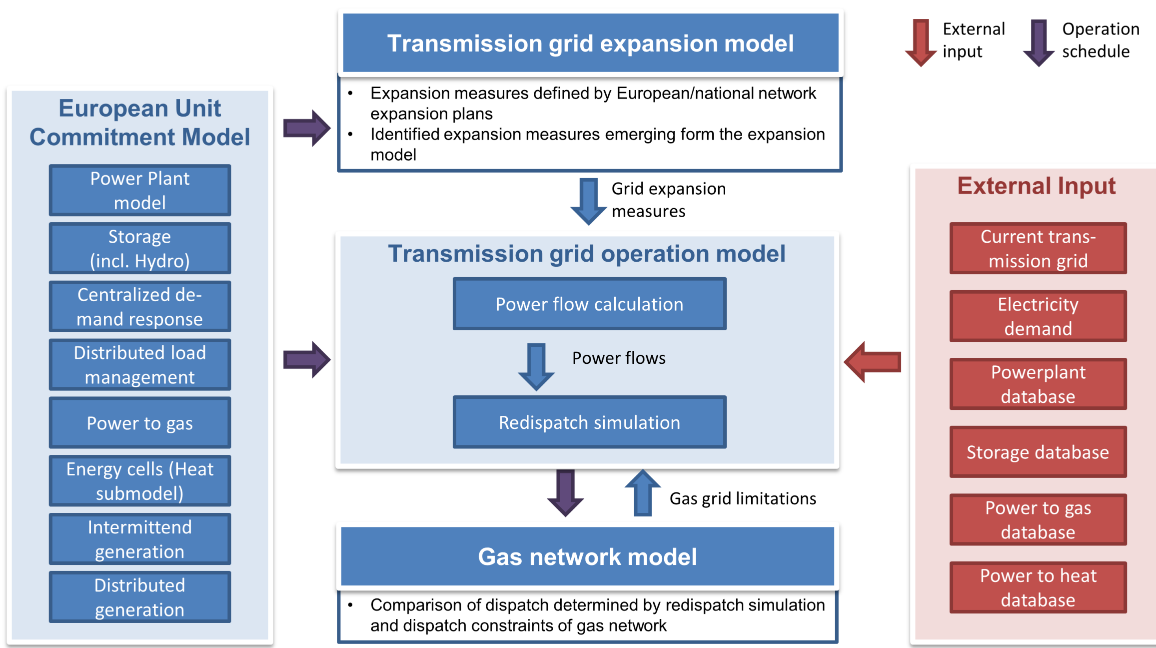


Figure 12: Model interconnection of the transmission grid operation model

Model requirements - Inputs

The operation of the transmission grid is simulated using network topologies including identified network expansion measures, hourly dispatch schedules provided by the sub-

modules of the EUC, limitations given by the gas network model as well as external input data like technological parameters and electricity demand timeseries.
The network expansion measures are determined by a detailed description of the connected stations/cluster of the topology and the chosen transmission corridors as well as the used technology and required technical parameters. The transmission expansion measures are either described by European or national network expansion plans or identified by the transmission grid expansion model.
The input data provided by the EUC model and the external input can be divided in data used for the power flow simulation and the data required for congestion management. The data mentioned first include dispatch schedules of power plants, storages, central/decentral intermittend generation units, power-to-gas units, the aggregated energy cell schedules (Heat submodel) as well as the electrical load. Redispatch calculations require additional information about the technical and operational constraints of power plants, pump storages and other generation facilities considered in the simulation of congestion management (e.g. power-to-gas, power-to-heat). Finally, the gas network model provides limitations in terms of restricted operating schedules for the power transfer between the electricity sector and the gas sector.

Table 42: Required input data for the transmission grid operation model

Model	Input	Description	Format
EUC	Dispatch schedules of power generation facilities	Hourly operation schedules to describe feed-in of generation facilities and power consumption used in grid simulation • Power plants • Storages • Centralised demand response • Distributed load management • Power-to-gas • Energy cells (Heat submodel) • Intermittend generation • Distributed generation	One time series per generation facility
EUC	Electricity demand	Electricity demand	One timeseries per cluster
External input	Databases	Technical parameters of generation facilities • Power plants • Storages • Power-to-gas units • Power-to-heat (Energy cells/Heat submodel)	Parameter “Matrices”

Gas network model	Power transfer schedules between electricity and gas sector	Power transfer between electricity and gas network	One time series for each coupling locations between gas network and electricity sector
Transmission grid expansion model	Network expansion measures	Required information (technology, voltage level, rating, etc.) to describe network expansion measures adequately	Matrix/table with relevant information
External input	Status quo transmission grid	Substations, lines including technical parameters and connectedt stations, transformers, etc.	List of nodes and lines including technical parameters

Model results - Outputs

The simulation of transmission grid operation provides the power flows, losses and the resulting congestions in the network based on the generation and load patterns. Furthermore, the redispatch volumes of each facility considered in the redispatch simulation and the curtailment of renewable generation units are determined as well the schedules of power flow controlling devices. The simulation of the grid operation results in new schedules of the generation facilities and in adjusted exchange time series between the electricity sector and other sectors to ensure N-1 secure grid operation and to balance the occurred network losses.

Table 43: Results of the transmission grid operation model

Model	Output	Description	Format
Gas network model	Power transfer schedules between electricity and gas sector	Power transfer between electricity and gas network	One time series for each coupling locations between gas network and electricity sector
Result	Adjusted schedules of generation facilities and adjusted exchange times series between electricity and other sectors	Hourly operation schedules to describe feed-in of generation facilities and exchange	One time series per generation facility

References

[1] A. Borghetti, A. Frangioni, F. Lacalandra, and C.A. Nucci. Lagrangian heuristics based on disaggregated bundle methods for hydrothermal unit commitment. *IEEE Transactions on Power Systems*, 18:313–323, 2003.

[2] M. Carrión and J.M. Arroyo. A computationally efficient mixed-integer linear formulation for the thermal unit commitment problem. *IEEE Transactions on Power Systems*, 21(3):1371–1378, 2006.

[3] L. Dubost, R. Gonzalez, and C. Lemaréchal. A primal-proximal heuristic applied to french unitcommitment problem. *Mathematical Programming*, 104(1):129–151, 2005.

[4] E.C. Finardi and E.L. Da Silva. Solving the hydro unit commitment problem via dual decomposition and sequential quadratic programming. *IEEE Transactions on Power Systems*, 21(2):835–844, 2006.

[5] A. Frangioni. Generalized bundle methods. *SIAM Journal on Optimization*, 13(1):117–156, 2002.

[6] W. de Oliveira, C. Sagastizábal, and C. Lemaréchal. Convex proximal bundle methods in depth: a unified analysis for inexact oracles. *Math. Prog. Series B*, 148:241–277, 2014.

[7] M. Tahanan, W. van Ackooij, A. Frangioni, and F. Lacalandra. Large-scale unit commitment under uncertainty: a literature survey. *4OR*, 13(2):115–171, 2015.

[8] R. Taktak and C. d'Ambrosio. An overview on mathematical programming approaches for the deterministic unit commitment problem in hydro valleys. *Energy Systems*, To appear, 2016.

[9] W. van Ackooij, N. Lebbe, and J. Malick. Regularized decomposition of large-scale block-structured robust optimization problems. *Computational Management Science*, 14(3):393–421, 2017.

[10] W. van Ackooij and J. Malick. Decomposition algorithm for large-scale two-stage unit-commitment. *Annals of Operations Research*, 238(1):587–613, 2016.