FlexPlan

Advanced methodology and tools taking advantage of storage and FLEXibility in transmission and distribution grid PLANning

Probabilistic optimization

of T&D systems planning

with high grid flexibility

and its scalability

D1.2

Distribution Level	RE
Responsible Partner	N-SIDE
Checked by WP leader	Hakan Ergun (KU Leuven) – WP1 Leader Date: 24/03/2021 (V2.0 – 01/08/2022)
Verified by the appointed Reviewers	Andrea Tognon and Alessio Siviero (Terna) – Date: 05/03/2021 Marko Kolenc and Klemen Dragaš (ELES) – Date: 05/03/2021
Approved by Project Coordinator	Gianluigi Migliavacca (RSE) Date: 25/03/2021 (V2.0 – 01/08/2022)

Issue Record

Planned delivery date	31/03/2021
Actual date of delivery	25/03/2021 (V2.0 – rev. 01/08/2022)
Status and version	FINAL

Version	Date	Author(s)	Notes
0.1	26.01.2021	Hakan Ergun (KUL), Iver Bakken Sperstad (SINTEF), Espen Flo Bødal (SINTEF), Dario Siface (RSE), Guido Pirovano (RSE), Marco Rossi (RSE), Matteo Rossini (RSE), Benedetta Marmiroli (RSE), Valentina Agresti (RSE), Matteo Paolo Costa (RSE), Alessandra Balzarini (RSE), Guillaume Leclercq (N-SIDE), Maxime Hanot (N-SIDE), Alberte Bouso (N-SIDE), Damien Lepage (N-SIDE), Tobias Patzwald (TUDO), Reinhilde D'hulst (VITO)	Cleaning, formatting and review of ID1.2 into D1.2.
0.2	27.01.2021	Hakan Ergun (KUL), Iver Bakken Sperstad (SINTEF), Espen Flo Bødal (SINTEF), Dario Siface (RSE), Guido Pirovano (RSE), Marco Rossi (RSE), Matteo Rossini (RSE), Benedetta Marmiroli (RSE), Valentina Agresti (RSE), Matteo Paolo Costa (RSE), Alessandra Balzarini (RSE), Guillaume Leclercq (N-SIDE), Maxime Hanot (N-SIDE), Alberte Bouso (N-SIDE), Damien Lepage (N-SIDE), Tobias Patzwald (TUDO), Reinhilde D'hulst (VITO)	Update layout to new FlexPlan deliverable template (logo, table styles, footer, table of content).
0.3	09.02.2021	Hakan Ergun (KUL), Iver Bakken Sperstad (SINTEF), Espen Flo Bødal (SINTEF), Dario Siface (RSE), Guido Pirovano (RSE), Marco Rossi (RSE), Matteo Rossini (RSE),	Merged updates from all partners following comments from N-SIDE. New Section 5.2.2: Hydropower stations modelling. New Section 3.8: Net Present Value formulation.

		Benedetta Marmiroli (RSE), Valentina Agresti (RSE), Matteo Paolo Costa (RSE), Alessandra Balzarini (RSE), Guillaume Leclercq (N-SIDE), Maxime Hanot (N-SIDE), Alberte Bouso (N-SIDE), Damien Lepage (N-SIDE), Tobias Patzwald (TUDO), Reinhilde D'hulst (VITO)	Summarized Chapter 7 (Monte Carlo scenario reduction) to avoid repetitions with D1.1 Cleaning of Data Model (Chapter 2).
0.4	10.02.2021	Hakan Ergun (KUL), Iver Bakken Sperstad (SINTEF), Espen Flo Bødal (SINTEF), Dario Siface (RSE), Guido Pirovano (RSE), Marco Rossi (RSE), Matteo Rossini (RSE), Benedetta Marmiroli (RSE), Valentina Agresti (RSE), Matteo Paolo Costa (RSE), Alessandra Balzarini (RSE), Guillaume Leclercq (N-SIDE), Maxime Hanot (N-SIDE), Alberte Bouso (N-SIDE), Damien Lepage (N-SIDE), Tobias Patzwald (TUDO), Reinhilde D'hulst (VITO)	Incorporation of updates from Dario Siface. Cleaning of objective function. Moving of T&D decomposition to Chapter 8 + creation of placeholder for Benders decomposition.
0.5	10.02.2021	Hakan Ergun (KUL), Iver Bakken Sperstad (SINTEF), Espen Flo Bødal (SINTEF), Dario Siface (RSE), Guido Pirovano (RSE), Marco Rossi (RSE), Matteo Rossini (RSE), Benedetta Marmiroli (RSE), Valentina Agresti (RSE), Matteo Paolo Costa (RSE), Alessandra Balzarini (RSE), Guillaume Leclercq (N-SIDE), Maxime Hanot (N-SIDE), Alberte Bouso (N-SIDE), Damien Lepage (N-SIDE), Tobias Patzwald (TUDO), Reinhilde D'hulst (VITO)	New Chapter 10: Proof-of-concept testing
0.6	11.02.2021	Hakan Ergun (KUL), Iver Bakken Sperstad (SINTEF), Espen Flo Bødal (SINTEF), Dario Siface (RSE), Guido Pirovano (RSE), Marco Rossi (RSE), Matteo Rossini (RSE), Benedetta Marmiroli (RSE), Valentina Agresti (RSE), Matteo Paolo Costa (RSE), Alessandra Balzarini (RSE),	Incorporation of updates from Guido Pirovano on air quality modelling Incorporation of Benders decomposition (Section 8.1.1, written by Dario Siface).

		Guillaume Leclercq (N-SIDE), Maxime Hanot (N-SIDE), Alberte Bouso (N-SIDE), Damien Lepage (N-SIDE), Tobias Patzwald (TUDO), Reinhilde D'hulst (VITO)	
0.7	12.02.2021	Hakan Ergun (KUL), Iver Bakken Sperstad (SINTEF), Espen Flo Bødal (SINTEF), Dario Siface (RSE), Guido Pirovano (RSE), Marco Rossi (RSE), Matteo Rossini (RSE), Benedetta Marmiroli (RSE), Valentina Agresti (RSE), Matteo Paolo Costa (RSE), Alessandra Balzarini (RSE), Guillaume Leclercq (N-SIDE), Maxime Hanot (N-SIDE), Alberte Bouso (N-SIDE), Damien Lepage (N-SIDE), Tobias Patzwald (TUDO), Reinhilde D'hulst (VITO)	Incorporation of tests on demand flexibility, reliability, distribution model, landscape impact model, scenario generation and reduction and stochastic optimization.
0.8	15.02.2021	Hakan Ergun (KUL), Iver Bakken Sperstad (SINTEF), Espen Flo Bødal (SINTEF), Dario Siface (RSE), Guido Pirovano (RSE), Marco Rossi (RSE), Matteo Rossini (RSE), Benedetta Marmiroli (RSE), Valentina Agresti (RSE), Matteo Paolo Costa (RSE), Alessandra Balzarini (RSE), Guillaume Leclercq (N-SIDE), Maxime Hanot (N-SIDE), Alberte Bouso (N-SIDE), Damien Lepage (N-SIDE), Tobias Patzwald (TUDO), Reinhilde D'hulst (VITO)	Incorporation of the list of abbreviations and executive summary.
0.9	16.02.2021	Hakan Ergun (KUL), Iver Bakken Sperstad (SINTEF), Espen Flo Bødal (SINTEF), Dario Siface (RSE), Guido Pirovano (RSE), Marco Rossi (RSE), Matteo Rossini (RSE), Benedetta Marmiroli (RSE), Valentina Agresti (RSE), Matteo Paolo Costa (RSE), Alessandra Balzarini (RSE), Guillaume Leclercq (N-SIDE), Maxime Hanot (N-SIDE), Alberte Bouso (N-SIDE), Damien Lepage (N-	Incorporation of tests on storage model for distribution.

		SIDE), Tobias Patzwald (TUDO), Reinhilde D'hulst (VITO)	
1.0	12.03.2021	Hakan Ergun (KUL), Iver Bakken Sperstad (SINTEF), Espen Flo Bødal (SINTEF), Dario Siface (RSE), Guido Pirovano (RSE), Marco Rossi (RSE), Matteo Rossini (RSE), Benedetta Marmiroli (RSE), Valentina Agresti (RSE), Matteo Paolo Costa (RSE), Alessandra Balzarini (RSE), Guillaume Leclercq (N-SIDE), Maxime Hanot (N-SIDE), Alberte Bouso (N-SIDE), Damien Lepage (N-SIDE), Tobias Patzwald (TUDO), Reinhilde D'hulst (VITO)	Addressed comments from Gianluigi Migliavacca (Project Coordinator) and Hakan Ergun (WP1 leader). Added conclusion to Chapter 9: Proof-of-concept testing. Addressed comments from Andrea Tognon and Alessio Sivier (Terna).
1.1	18.03.2021	Hakan Ergun (KUL), Iver Bakken Sperstad (SINTEF), Espen Flo Bødal (SINTEF), Dario Siface (RSE), Guido Pirovano (RSE), Marco Rossi (RSE), Matteo Rossini (RSE), Benedetta Marmiroli (RSE), Valentina Agresti (RSE), Matteo Paolo Costa (RSE), Alessandra Balzarini (RSE), Guillaume Leclercq (N-SIDE), Maxime Hanot (N-SIDE), Alberte Bouso (N-SIDE), Damien Lepage (N-SIDE), Tobias Patzwald (TUDO), Reinhilde D'hulst (VITO)	Addressed comments from Marko Kolenc and Klemen Dragaš (ELES).
1.2	24.03.2021	Hakan Ergun (KUL), Iver Bakken Sperstad (SINTEF), Espen Flo Bødal (SINTEF), Dario Siface (RSE), Guido Pirovano (RSE), Marco Rossi (RSE), Matteo Rossini (RSE), Benedetta Marmiroli (RSE), Valentina Agresti (RSE), Matteo Paolo Costa (RSE), Alessandra Balzarini (RSE), Guillaume Leclercq (N-SIDE), Maxime Hanot (N-SIDE), Alberte Bouso (N-SIDE), Damien Lepage (N-SIDE), Tobias Patzwald (TUDO), Reinhilde D'hulst (VITO)	Update of Sections 6.1.1and 6.1.2 (Air quality impact model) following reviews. New test sections on Air quality impact (9.3.5.2 and 9.3.5.3).Update of figures in Section 5.2.2 (Modelling the flexibility of hydropower plants): bigger text sizes. Addressed further comments from Gianluigi Migliavacca (Project Coordinator) and Hakan Ergun (WP1 leader). Removed time dependence of investment decisions.

1.3	01.08.2022	Hakan Ergun (KUL)	Updated modelling assumptions based on regional cases performance, and documentation of the open-source implementations FlexPlan.jl and OptimalTransmissionRouting.jl.

About FlexPlan

The FlexPlan project aims at establishing a new grid planning methodology considering the opportunity to introduce new storage and flexibility resources in electricity transmission and distribution grids as an alternative to building new grid elements. This is in line with the goals and principles of the new EC package *Clean Energy for all Europeans,* which emphasizes the potential usage of flexibility sources in the phases of grid planning and operation as alternative to grid expansion. In sight of this, FlexPlan creates a new innovative grid planning tool whose ambition is to go beyond the state of the art of planning methodologies, by including the following innovative features: integrated T&D planning, full inclusion of environmental analysis, probabilistic contingency methodologies replacing the N-1 criterion as well as optimal planning decision over several decades. However, FlexPlan is not limited to building a new tool but it also uses it to analyse six regional cases covering nearly the whole European continent, aimed at demonstrating the application of the tool on real scenarios as well as at casting a view on grid planning in Europe till 2050. In this way, the FlexPlan project tries to answer the question of which role flexibility could play and how its usage can contribute to reduce planning investments yet maintaining (at least) the current system security levels. The project ends up formulating guidelines for regulators and for the planning offices of TSOs and DSOs. The consortium includes three European TSOs, one of the most important European DSO group, several R&D companies and universities from 8 European Countries (among which the Italian RSE acting as project coordinator) and N-SIDE, the developer of the European market coupling platform EUPHEMIA.

Partners

Table of Contents

List of Abbreviations and Acronyms

Abbreviation/Acronym	Meaning
AC	Alternating Current
AQ	Air Quality
BESS	Battery Energy Storage System
CAMx	Comprehensive Air quality Model
CAES	Compressed Air Energy Storage
CENS	Cost of Energy Not Supplied
CHP	Combined Heat and Power
CTM	Chemistry and Transport Model
CO_2	Carbon Dioxide
DC	Direct Current
DDM	Decoupled Direct Method
DER	Distributed Energy Resources
DNEP	Distribution Network Expansion Problem
ESS	Energy Storage System
EV	Electric Vehicles
GWP	Global Warming Potential
HVDC	High-Voltage Direct Current
HV	High-Voltage
HVL	High-Voltage Lines
kVA	Kilovolt-Ampere
kWh	Kilowatt-hour
LAES	Liquified Air Energy Storage
LCC	Line-Commuted Converter
LFP	Lithium Ferrous Phosphate
MC	Monte Carlo
MILP	Mixed-Integer Linear Program
MINLP	Mixed-Integer Non-Linear Program

MJ	Megajoule
MVA	Megavolt-Ampere
MV	Medium Voltage
MW	Megawatt
MWh	Megawatt-Hour
OHL	Over Head Line
OPF	Optimal Power Flow
PCA	Principal Component Analysis
PM	Particulate Matter
PP	Power Plant
PST	Phase-Shifting Transformer
PV	Photovoltaic
PSH	Pumped Storage Hydro
RE	Renewable Energy
RR	Relative Risk
TNEP	Transmission Network Expansion Problem
UGC	Under Ground Cable
VRES	Variable Renewable Energy Source
VSC	Voltage-Source Converter
VSL	Value of a Statistical Life
WHO	World Healthcare Organization
WTP	Willing to Pay
YOLL	Years of Life Lost

Executive Summary

Recent advances in computing power allow to include complex optimization problems in the core operations of multiple sectors: logistic, automotive, or energy. If designed properly, an accurate model can boost processes' efficiency or decrease costs, which is reflected in the end consumer bills, and with an overall increase in the social welfare. In the electricity sector, network expansion activities account for a big share of Transmission System Operators (TSOs) and Distribution System Operators (DSOs) budget and are activities for which advanced optimization is identified as a potential tool to reduce the system costs. Nevertheless, challenges arise when modelling the uncertain system needs and the growing number of available technologies at the time when the investments take place. Indeed, technologies enabling demand flexibility and energy storage at various horizons are gaining interest as alternatives to classical grid investments (e.g., additional lines) to solve congestion issues. At the same time, in the context of increasingly connected grids as well as intermittent and unpredictable energy sources, the impact of an investment has to be computed at supra-national level and must account for the system variability, resulting in large-scale problems. Thus, one faces a trade-off between a sufficiently accurate representation of the system to optimize and the computational burden of reaching an optimal solution of the large-scale planning problem at hand.

This document aims to provide the generic design guidelines for the FlexPlan planning tool, aiming to overcome the challenges mentioned above. As such, it provides the necessary mathematical modelling details with respect to the optimization target function, network flow modelling, reliability modelling and flexibility modelling. Additionally, it gives an overview of the data requirements to ensure a proper alignment with the ongoing activities within the project, such as the scenarios generation and the proper flexibility characterization and valorisation. Optimization techniques to improve the computational efficiency of the tool are investigated and our first findings in terms of scalability of the model are presented using a proof-of-concept implementation.

One of the critical aspects when designing a large-scale optimization problem as the one presented here is the adequate definition of the target function. For such a task, the document presents a formulation for the social welfare, which accounts for a long planning horizon covering multiple decades, considering the effects of both the new installations and the recurrent operational costs. This characteristic, also known as dynamic optimization, allows to precisely model multiple factors in a cost-quantification fashion, such as environmental impact, reliability of supply, investment and operational costs. For all costs, the net present value formulation of the objective allows to take into account the present value of costs incurred at different years in the future. Furthermore, the stochastic formulation of the objective function allows to account for several scenarios modelling the uncertainty in load and generation profiles (e.g., due to renewable generation) at the different planning horizons.

Simplifications have to be made in order to keep the tractability of the problem, and this is reflected in the linear formulation of the power flow model, both for transmission and distribution networks. This way, the usage of efficient mixed-integer linear programming solvers is leveraged. In particular, starting from the non-linear representation of the transmission and distribution grids, the document derives the DC and the octagonal approximations, respectively, pointing out the assumptions for each one. Both formulations account for potential investments in new assets through sets of binary variables and adapted network constraints. Moreover, the model for transmission allows to model mixed AC and DC grids, taking among others DC interconnectors into account, both as existing and investment candidate assets. One of the novelties introduced in this work is the analytical formulation on the interface between TSOs and DSOs for the planning problem to enable combined optimization of transmission and distribution systems. Planning at distribution level can for instance be considered as a potential investment to enable upward or downward flexibility at transmission level. As such, four approaches are presented, with emphasis on the ad-hoc heuristic developed to integrate the simultaneous planning process.

As mentioned above, flexibility resources such as batteries or flexible loads are considered as alternative or complementary investment possibilities to the commissioning of new branches or their reinforcement. Hence, including their characteristics in a more generic and versatile way later allows to reduce the total system cost, for instance by investing in demand flexibility, enabling load decrease or load shifting when needed. The developed generic storage model also allows the representation of large flexibility sources such as Nordic hydro power as storage facilities.

The environmental impact model, which is included under the target function, elaborates on three parallel categories: (i) air quality, (ii) carbon footprint, and (iii) landscape impact. Whereas the first category only accounts for existing thermal generators, the carbon footprint generalizes to both investments and operations through the life cycle of each asset. Finally, through an optimal routing algorithm, the best routes for the new branches are evaluated, determining the minimum costs of the candidates taking into account their landscape impact costs.

In order for the network expansion planning tool to determine investment decisions which allow the network to operate reliably for a range of uncertain future conditions, scenarios are generated and subsequently reduced to a limited number of representative time-series to reduce the computational burden. The key uncertainties identified are the presence of renewable generation resources, temperature-dependent loads and hydro-condition dependent storage and production. Whereas the scenarios generation is out of the scope of the planning tool itself, we briefly present the methodology for reducing the number of such scenarios.

As mentioned earlier, the biggest challenge the FlexPlan planning tool faces is the computational efficiency for the large-scale optimization problem. Having this in mind, the document present four different techniques that can be applied to increase the computational efficiency if proven necessary. In particular, the

decomposition of the planning and operational problems using Benders decomposition and the combined T&D optimization are identified as the most promising ones.

Finally, the aforementioned models and methodologies are tested through proof-of-concept tests, based on the Garver transmission expansion test system and the CIGRE medium voltage distribution network. These validation tests present the key features and potential limitations of the model as well as our first findings on its scalability to large-scale systems. A proof-of-concept package, *FlexPlan.jl*, was implemented to conduct those tests. *FlexPlan.jl* will serve as a reference for the implementation of the FlexPlan planning tool and will be made publicly available at the end of the project activities.

1 Introduction

The main goal of FlexPlan is to develop and implement a grid expansion optimization tool able to incorporate flexible grid elements: conventional network assets on one hand and flexibility sources (such as storage and demand side management) on the other. The tool should be applicable to both transmission and distribution systems alike, providing the possibility to optimize investments in both networks at the same time.

Figure 1-1 shows the outline of the optimization model and the input parameters. A number of candidate grid investments, flexibility and storage options are provided as an input for the tool, which will be provided by the pre-processor developed within WP2 of FlexPlan. These expansion candidates are characterised both technically, e.g., power ratings and economically, e.g., CAPEX and OPEX. The network planning is carried out for a number of generation and demand time series. Transmission networks data (based on the Ten Years Network Development Plan – TYNDP) and distribution networks data (synthetic or real ones) are needed in order to provide grid constraints for the optimization problem.

As a first step, grid expansion and flexibility candidates are analysed in order to quantify their costs based on environmental impact (air quality, life-cycle assessment and landscape). These additional costs are included into the objective function of the expansion optimisation, such that the best trade-off between T&D system investments and operational costs is found by also considering environmental externalities.

The optimization is carried out in parallel for three Pan-EU scenarios, elaborated for the years 2030-2040-2050 and based on well-established EU and national "visions" as well as on the ENTSO-E's TYNDP. "National Trends" is the central policy scenario of the TYNDP 2020 report, designed to reflect the most recent EU member state National Energy and Climate Plans. In addition, ENTSO-E and ENTSO-G have created two scenarios in line with the COP 21 targets: "Distributed Energy" and "Global Ambition" with the objective to understand the impact on infrastructure needs against different pathways reducing EU-28 emissions to net-zero by 2050. More information on these scenarios can be found in deliverable D4.1 of the FlexPlan project [1]. A Monte-Carlo approach accounts for yearly climate variations in the planning optimisation framework.
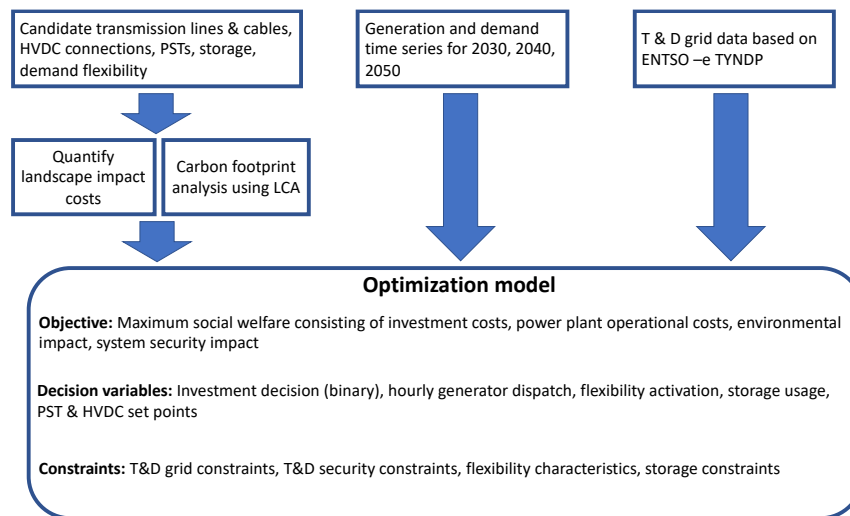
Figure 1-1 – High level outline of the optimization model

The objective of the optimization is to maximize the system social welfare. This is obtained by minimizing the sum of T&D grid investments, operational costs bound to system dispatch and environmental impact costs, while maximizing the benefits achieved by the use of the flexibility sources and storage. The optimization is performed jointly for three target years, namely 2030, 2040 and 2050, and each year is characterised by a continuous time series of ideally 8760 hours, which is necessary to model storage and flexibility activation accurately. As a result, a step-wise investment plan for new grid connections and flexibility investments is obtained.

Binary investment decision variables are used for grid and flexibility investments, whereas continuous variables are used for generation dispatch, and the dispatch of flexibility and storage sources. Considering the three target decades and the detailed characterisation of each planning year, a large-scale mixed integer problem optimization is obtained.

The power flow equations and technical constraints for flexibility sources and storage are formulated in a linear way, in order to maintain tractability of the model notwithstanding its huge dimensions. Security constraints for critical contingencies are included into the model. Possible re-dispatch and load curtailment costs stemming from these contingencies are weighted probabilistically. The weighted costs are added into the objective function of the optimization, in order to find the best trade-off between additional grid and flexibility investments to avoid congestions during outages versus the expected impact of such grid outages.

2 Data model

This section describes the data model used for the planning tool. First all sets, entities and indices are provided for consistent use throughout the document. For each set, detailed description of used variables, optimization parameters (as directly used in the optimization model), additional parameters for pre-processing and visualisation are provided. For all variables and parameters, their cardinality, typical ranges (where applicable) and mathematical symbols are provided.

The grid data model is provided for the nonlinear formulation of the power flow equations, which is further linearized in Section 4.1. Regardless of the actual implementation, different sets have been defined for candidate and existing assets, in order to have clear and concise notation.

For the sake of generality and considering that pre-solvers (CPLEX, Gurobi, …) can eliminate potentially redundant variables, some of the variables have been defined over a larger number of sets, e.g., the bounds on the maximum absorption and injection power for storage have been defined for every asset and every planning horizon (although the bounds may be the same in all planning horizons). Optional elements of the data model have been marked with a red colour.

2.1 Sets, entities and indices used in planning tool

Set / Entity	Symbol	Indices
Set of planning horizons	S_y	$y \in S_y$
Set of periods in the planning horizon	S_t	$t \in S_t$
Set of existing storage elements	S_j	$j \in S_j$
Set of candidate storage elements	S_{jc}	$jc \in S_{jc}$
Set of generators	S_g	$g \in S_g$
Set of ac nodes	S_n^{ac}	$m, k \in S_n^{ac}$
Set of loads	S_u	$u \in S_u$
Set of time windows for balancing of demand shifting	S_τ	$\tau \in S_\tau$
Set of AC branches	S_l^{ac}	$l \in S_l^{ac}$
Set of candidate AC branches	S_{lc}^{ac}	$lc \in S_{lc}^{ac}$
AC grid topology	T^{ac}	$l, m, k \in T^{ac}, (S_l^{ac} + S_{lc}^{ac}) \times S_n^{ac} \times S_n^{ac}$ Defined as tuple: (ac line id, from node, to node)
Set of PSTs	S_b	$b \in S_b$

Set of candidate PSTs	S_{bc}	$bc \in S_{bc}$
PST connectivity	T^b	$b, m, k \in T^b, (S_b + S_{bc}) \times S_n^{ac} \times S_n^{ac}$ Defined as tuple: (PST, from node, to node)
Set of dc nodes	S_n^{dc}	$e, f \in S_n^{dc}$
Set of DC branches	S_l^{dc}	$d \in S_l^{dc}$
Set candidate of DC branches	S_{lc}^{dc}	$dc \in S_{lc}^{dc}$
DC grid topology	T^{dc}	$d, e, f \in T^{dc}, (S_l^{dc} + S_{lc}^{dc}) \times S_n^{dc} \times S_n^{dc}$ Defined as tuple: (dc line id, from node, to node)
Set of AC/DC converters	S_z	$z \in S_z$
Set of candidate AC/DC converters	S_{zc}	$zc \in S_{zc}$
AC / DC converter connectivity	T^{acdc}	$zme \in T^{acdc}, (S_z + S_{zc}) \times S_n^{ac} \times S_n^{dc}$ Defined as tuple: (converter id, ac node, dc node)
Generator connectivity	T^{gen}	$g, m \in T^{gen}, S_g \times S_n^{ac}$ Defined as tuple: (generator id, ac node connected)
Load connectivity	T^{load}	$u, m \in T^{load}, S_u \times S_n^{ac}$ Defined as tuple: (load id, ac node connected)
Storage connectivity	T^{st}	$j, m \in T^{load}, (S_j + S_{jc}) \times S_n^{ac}$ Defined as tuple: (storage id, ac node connected)
Set of contingencies	S_c	$c \in S_c$
Set of pollutants	S_p	$p \in S_p$
Set of countries	S_{cy}	$cy \in S_{cy}$
Set of grid cells for air quality modelling	S_{ijk}	$ijk \in S_{ijk}$
Set of meteorological variables	S_m	$m \in S_m$
Set of AQ impacts	S_{imp}	$imp \in S_{imp}$

2.2 AC bus data model

Variables	Symbol	Cardinality	Unit
AC bus voltage magnitude	$U_{m,t,y}$	$\forall m \in S_n^{ac}, \forall t \in S_t, \forall y \in S_y$	kV
AC bus voltage angle	$\theta_{m,t,y}$	$\forall m \in S_n^{ac}, \forall t \in S_t, \forall y \in S_y$	rad

Optimization parameters	Symbol	Cardinality	Unit	Typical value
Nominal AC bus voltage magnitude	$U_{m,t,y}^{nom}$	$\forall m \in S_n^{ac}, \forall t \in S_t, \forall y \in S_y$	kV	$\{110, 220, 380\}$ kV
Maximum operating voltage	$U_{m,t,y}^{max}$	$\forall m \in S_n^{ac}, \forall t \in S_t, \forall y \in S_y$	kV	$U_{m,t,y} + 10\%$
Minimum operating voltage	$U_{m,t,y}^{min}$	$\forall m \in S_n^{ac}, \forall t \in S_t, \forall y \in S_y$	kV	$U_{m,t,y} - 10\%$
Maximum voltage angle	$\theta_{m,t,y}^{max}$	$\forall m \in S_n^{ac}, \forall t \in S_t, \forall y \in S_y$	rad	2π
Minimum voltage angle	$\theta_{m,t,y}^{min}$	$\forall m \in S_n^{ac}, \forall t \in S_t, \forall y \in S_y$	rad	-2π

Additional parameters	Symbol	Cardinality	Unit	Typical value
Bus ID	[-]	$\forall m \in S_n^{ac}$	[-]	[-]
Bus location	[-]	$\forall m \in S_n^{ac}$	[lat, lon]	[-]

2.3 DC bus data model

Variables	Symbol	Cardinality	Unit
DC bus voltage magnitude	$U_{e,t,y}$	$\forall e \in S_n^{dc}, \forall t \in S_t, \forall y \in S_y$	kV

Optimization parameters	Symbol	Cardinality	Unit	Typical value
Nominal DC bus voltage magnitude	$U_{e,t,y}^{nom}$	$\forall e \in S_n^{dc}, \forall t \in S_t, \forall y \in S_y$	kV	$\{320, 525, 600\}$ kV

Maximum operating voltage	$U_{e,t,y}^{max}$	$\forall e \in S_n^{dc}, \forall t \in S_t, \forall y \in S_y$	kV	$U_{e,t,y} + 10\%$
Minimum operating voltage	$U_{e,t,y}^{min}$	$\forall e \in S_n^{dc}, \forall t \in S_t, \forall y \in S_y$	kV	$U_{e,t,y} - 10\%$

Additional parameters	Symbol	Cardinality	Unit	Typical value
Bus ID	[-]	$\forall e \in S_n^{dc}$	[-]	[-]
Bus location	[-]	$\forall e \in S_n^{dc}$	[lat, lon]	[-]

2.4 Generator data model

Variables	Symbol	Cardinality	Unit
Active power generation	$P_{g,t,y}$	$\forall g \in S_g, \forall t \in S_t, \forall y \in S_y$	MW
Reactive power generation	$Q_{g,t,y}$	$\forall g \in S_g, \forall t \in S_t, \forall y \in S_y$	Mvar

Optimization parameters	Symbol	Cardinality	Unit	Typical value
Active power reference	$P_{g,t,y}^{ref}$	$\forall g \in S_g, \forall t \in S_t, \forall y \in S_y$	MW	[-]
Maximum active power generation	$P_{g,t,y}^{max}$	$\forall g \in S_g, \forall t \in S_t, \forall y \in S_y$	MW	[-]
Minimum active power generation	$P_{g,t,y}^{min}$	$\forall g \in S_g, \forall t \in S_t, \forall y \in S_y$	MW	0
Generation cost	$C_{g,t,y}^{f}$	$\forall g \in S_g, \forall t \in S_t, \forall y \in S_y$	€/MWh	[-]
Maximum reactive power exchange	$Q_{g,t,y}^{max}$	$\forall g \in S_g, \forall t \in S_t, \forall y \in S_y$	MW	[-]
Minimum reactive power exchange	$Q_{g,t,y}^{min}$	$\forall g \in S_g, \forall t \in S_t, \forall y \in S_y$	MW	[-]
Generator status	$s_{g,y,t}$	$\forall g \in S_g, \forall t \in S_t, \forall y \in S_y$	[-]	{0, 1}

Additional parameters	Symbol	Cardinality	Unit	Typical value
Generator ID	[-]	$\forall g \in S_g$	[-]	[-]
AC bus connected	[-]	$\forall g \in S_g$	[-]	[-]

Generator fuel type	[-]	$\forall g \in S_g$	[-]	Solar PV, Coal, Hydro, ...
Emission factor	$G_{g,p}^{ef}$	$\forall g \in S_g, \forall p \in S_p$	kg/MJ	[-]
Stack height	G_g^{sh}	$\forall g \in S_g$	m	[-]
Stack diameter	G_g^{sd}	$\forall g \in S_g$	m	[-]
Plume velocity	G_g^{plv}	$\forall g \in S_g$	m/s	[-]
Plume temperature	G_g^{plt}	$\forall g \in S_g$	°K	[-]
Generator type	[-]	$\forall g \in S_g$	[-]	Open cycle, ...
Fuel price	θ_y^f	$\forall y \in S_y$	[€/MWh]	[-]
Specific fuel consumption	η_g^f	$\forall g \in S_g$	[MWh/MWh]	[-]

2.5 (Flexible) demand data model

Variables	Symbol	Cardinality	Unit
Active power consumption	$P_{u,t,y}^{flex}$	$\forall u \in S_u, \forall t \in S_t, \forall y \in S_y$	MW
Reactive power consumption	$Q_{u,t,y}^{flex}$	$\forall u \in S_u, \forall t \in S_t, \forall y \in S_y$	Mvar
Not consumed power	$\Delta P_{u,t,y}^{nce}$	$\forall u \in S_u, \forall t \in S_t, \forall y \in S_y$	MW
Upward demand shifted	$\Delta P_{u,t,y}^{ds,up}$	$\forall u \in S_u, \forall t \in S_t, \forall y \in S_y$	MW
Downward demand shifted	$\Delta P_{u,t,y}^{ds,dn}$	$\forall u \in S_u, \forall t \in S_t, \forall y \in S_y$	MW
Load curtailment	$\Delta P_{u,t,y}^{lc}$	$\forall u \in S_u, \forall t \in S_t, \forall y \in S_y$	MW
Investment decision (enabling flexibility)	$\alpha_{u,y} \in \{0,1\}$	$\forall j \in S_j, \forall y \in S_y$	[-]

Optimization parameters	Symbol	Cardinality	Unit	Typical value
Power factor angle	$\varphi_{u,t,y}^{flex}$	$\forall u \in S_u, \forall t \in S_t, \forall y \in S_y$	rad	[-0.45, 0.45]
Reference demand	$P_{u,t,y}^{ref}$	$\forall u \in S_u, \forall t \in S_t, \forall y \in S_y$	MW	[-]
Maximum energy not consumed (accumulated load reduction)	$E_{u,y}^{nc,max}$	$\forall u \in S_u, \forall y \in S_y$	MWh	[-]

Maximum energy (accumulated net demand) shifted	$E_{u,y}^{ds,max}$	$\forall u \in S_u, \forall y \in S_y$	MWh	[-]
Superior bound on not consumed power (demand reduction)	$\Delta_{u,t,y}^{nce,max}$	$\forall u \in S_u, \forall t \in S_t, \forall y \in S_y$	MW	[-]
Superior bound on upward demand shifted	$\Delta_{u,t,y}^{ds,up,max}$	$\forall u \in S_u, \forall t \in S_t, \forall y \in S_y$	MW	[-]
Recovery period for upward demand shifting	$\tau_{u,y}^{ds,up,rec}$	$\forall u \in S_u, \forall y \in S_y$	h	< 24 h
Superior bound on downward demand shifted	$\Delta_{u,t,y}^{ds,dn,max}$	$\forall u \in S_u, \forall t \in S_t, \forall y \in S_y$	MW	[-]
Recovery period for downward demand shifting	$\tau_{u,y}^{ds,dn,rec}$	$\forall u \in S_u, \forall y \in S_y$	h	< 24 h
Maximum energy (accumulated load) shifted downward	$E_{u,y}^{ds,max}$	$\forall u \in S_u, \forall y \in S_y$	MWh	[-]
Compensation for consuming less (i.e., voluntary demand reduction)	$C_{u,t,y}^{nce}$	$\forall u \in S_u, \forall t \in S_t, \forall y \in S_y$	€/MWh	[-]
Compensation for load curtailment (i.e., involuntary demand reduction)	$C_{u,t,y}^{lc}$	$\forall u \in S_u, \forall t \in S_t, \forall y \in S_y$	€/MWh	[-]
Compensation for flexibility (demand shifting)	$C_{u,t,y}^{ds}$	$\forall u \in S_u, \forall t \in S_t, \forall y \in S_y$	€/MWh	[-]
Specific interruption costs (involuntary load curtailment costs, or costs of energy not supplied, due to contingencies)	$C_{u,t,y}^{cens}$	$\forall u \in S_u, \forall t \in S_t, \forall y \in S_y$	€/MWh	[-]
Investment cost (for enabling potential demand flexibility)	$I_{u,y}$	$\forall u \in S_u, \forall y \in S_y$	€	[-]
Carbon footprint cost (for enabling potential demand flexibility)	$FP_{u,y}^{CO_2}$	$\forall u \in S_u, \forall y \in S_y$	€	[-]
Status	$S_{u,t,y}$	$\forall u \in S_u, \forall t \in S_t, \forall y \in S_y$	[-]	{0, 1}

2.6 Storage data model

Variables	Symbol	Cardinality	Unit
Normalized energy storage level	$x_{j,t,y}$	$\forall j \in S_j, \forall t \in S_t, \forall y \in S_y$	[-]
Power absorbed from grid	$P_{j,t,y}^{abs}$	$\forall j \in S_j, \forall t \in S_t, \forall y \in S_y$	MW
Power injected to grid	$P_{j,t,y}^{inj}$	$\forall j \in S_j, \forall t \in S_t, \forall y \in S_y$	MW
Exchanged reactive power	$Q_{j,t,y}$	$\forall j \in S_j, \forall t \in S_t, \forall y \in S_y$	Mvar
Investment decision	$\alpha_{j,y} \in \{0,1\}$	$\forall j \in S_j, \forall y \in S_y$	[-]

Optimization parameters	Symbol	Cardinality	Unit	Typical value
Maximum energy content	$E_{j,y}^{max}$	$\forall j \in S_j, \forall y \in S_y$	MWh	[-]
Minimum energy content	$E_{j,y}^{min}$	$\forall j \in S_j, \forall y \in S_y$	MWh	[-]
Initial energy content	$E_{j,y}^{init}$	$\forall j \in S_j, \forall y \in S_y$	MWh	[-]
Maximum absorbed energy over a year	$E_{j,y}^{abs,max}$	$\forall j \in S_j, \forall y \in S_y$	MWh	[-]
Maximum absorbed power	$P_{j,t,y}^{abs,max}$	$\forall j \in S_j, \forall t \in S_t, \forall y \in S_y$	MW	[-]
Maximum injected power	$P_{j,t,y}^{inj,max}$	$\forall j \in S_j, \forall t \in S_t, \forall y \in S_y$	MW	[-]
Maximum reactive power exchange	$Q_{j,y}^{max}$	$\forall j \in S_j, \forall y \in S_y$	Mvar	[-]
Minimum reactive power exchange	$Q_{j,y}^{min}$	$\forall j \in S_j, \forall y \in S_y$	Mvar	[-]
Absorption efficiency	$\eta_{j,y}^{abs}$	$\forall j \in S_j, \forall y \in S_y$	[-]	[-]
Injection efficiency	$\eta_{j,y}^{inj}$	$\forall j \in S_j, \forall y \in S_y$	[-]	[-]
Maximum absorption ramp rate	$r_{j,y}^{abs,max}$	$\forall j \in S_j, \forall y \in S_y$	MW/h	[-]
Maximum injection ramp rate	$r_{j,y}^{inj,max}$	$\forall j \in S_j, \forall y \in S_y$	MW/h	[-]
Power provided or demanded by external process	$\xi_{j,t,y}$	$\forall j \in S_j, \forall t \in S_t, \forall y \in S_y$	MW	[-]
Hourly discharge rate	$dr_{j,y}$	$\forall j \in S_j, \forall y \in S_y$	[-]	[0, 1]
Status	$s_{j,t,y}$	$\forall j \in S_j, \forall t \in S_t, \forall y \in S_y$	[-]	{0, 1}

Absorption cost	$C_{j,t,y}^{abs}$	$\forall j \in S_j, \forall t \in S_t, \forall y \in S_y$	€/MWh	[-]
Injection cost	$C_{j,t,y}^{inj}$	$\forall j \in S_j, \forall t \in S_t, \forall y \in S_y$	€/MWh	[-]
Investment cost	$I_{jc,y}$	$\forall jc \in S_{jc}, \forall y \in S_y$	€	[-]
Carbon footprint cost	$FP_{jc,y}^{CO_2}$	$\forall jc \in S_{jc}, \forall y \in S_y$	€	[-]

Additional parameters	Symbol	Cardinality	Unit	Typical value
Storage ID	[-]	$\forall j \in S_j$	[-]	[-]
AC bus connected	[-]	$\forall j \in S_j$	[-]	[-]

2.7 AC branch data model (AC lines, cables and transformers)

Variables	Symbol	Cardinality	Unit
Active power flow in from-direction, existing lines	$P_{l,t,y}^{fr}$	$\forall l \in S_l, \forall t \in S_t, \forall y \in S_y$	MW
Active power flow in to-direction, existing lines	$P_{l,t,y}^{to}$	$\forall l \in S_l, \forall t \in S_t, \forall y \in S_y$	MW
Reactive power flow in from-direction existing lines	$Q_{l,t,y}^{fr}$	$\forall l \in S_l, \forall t \in S_t, \forall y \in S_y$	Mvar
Reactive power flow in to-direction, existing lines	$Q_{l,t,y}^{to}$	$\forall l \in S_l, \forall t \in S_t, \forall y \in S_y$	Mvar
Active power flow in from-direction, candidate lines	$P_{lc,t,y}^{fr}$	$\forall lc \in S_{lc}, \forall t \in S_t, \forall y \in S_y$	MW
Active power flow in to-direction, candidate lines	$P_{lc,t,y}^{to}$	$\forall lc \in S_{lc}, \forall t \in S_t, \forall y \in S_y$	MW
Reactive power flow in from-direction, candidate lines	$Q_{lc,t,y}^{fr}$	$\forall lc \in S_{lc}, \forall t \in S_t, \forall y \in S_y$	Mvar
Reactive power flow in to-direction, candidate lines	$Q_{lc,t,y}^{to}$	$\forall lc \in S_{lc}, \forall t \in S_t, \forall y \in S_y$	Mvar
Investment decision	$\alpha_{lc,y} \in \{0,1\}$	$\forall lc \in S_{lc}, \forall y \in S_y$	[-]
Voltage transformation ratio	$\tau_{l,t,y}$	$\forall l \in S_l, \forall t \in S_t, \forall y \in S_y$	[-]

Voltage transformation ratio, candidate lines	$\tau_{lc,t,y}$	$\forall lc \in S_l, \forall t \in S_t, \forall y \in S_y$	[-]	

Optimization parameters (Existing and candidate)	Symbol	Cardinality	Unit	Typical value
Resistance	$r_{l(c)}$	$\forall l(c) \in S_{l(c)}$	Ohm	[-]
Reactance	$x_{l(c)}$	$\forall l(c) \in S_{l(c)}$	Ohm	[-]
Susceptance	$b_{l(c)}$	$\forall l(c) \in S_{l(c)}$	1 / Ohm	[-]
Thermal rating	$S_{l(c)}^{rated}$	$\forall l(c) \in S_{l(c)}$	MVA	[-]
Emergency rating	$S_{l(c)}^{rated,em}$	$\forall l(c) \in S_{l(c)}$	MVA	[-]
Maximum voltage transformation ratio	$\tau_{l(c)}^{max}$	$\forall l(c) \in S_{l(c)}$	[-]	[-]
Minimum voltage transformation ratio	$\tau_{l(c)}^{min}$	$\forall l(c) \in S_{l(c)}$	[-]	[-]
Maximum angle difference	$\Delta\theta_{l(c)}^{max}$	$\forall l(c) \in S_{l(c)}$	Rad	[-]
Minimum angle difference	$\Delta\theta_{l(c)}^{min}$	$\forall l(c) \in S_{l(c)}$	Rad	[-]
Failure rate	$\lambda_{l(c)}$	$\forall l(c) \in S_{l(c)}$	1 / year	[-]
Mean time to repair	$t_{l(c)}^{MTTR}$	$\forall l(c) \in S_{l(c)}$	h	[-]
Mean time between failures	$t_{l(c)}^{MTBF}$	$\forall l(c) \in S_{l(c)}$	h	[-]
Status	$s_{l(c),t,y}$	$\forall l(c) \in S_{l(c)}, \forall t \in S_t, \forall y \in S_y$	[-]	{0, 1}
Investment cost	$I_{lc,y}$	$\forall lc \in S_{lc}, \forall y \in S_y$	€	[-]
Carbon footprint cost	$FP_{lc,y}^{CO_2}$	$\forall lc \in S_{lc}, \forall y \in S_y$	€	[-]
Landscape impact cost	$LS_{lc,y}$	$\forall lc \in S_{lc}, \forall y \in S_y$	€	[-]
Asset lifetime	$t_{l(c)}^{life}$	$\forall l(c) \in S_{l(c)}$	year	[-]

2.8 PST data model

Variables	Symbol	Cardinality	Unit
Active power flow in from-direction, existing PSTs	$P_{b,t,y}^{fr}$	$\forall b \in S_b, \forall t \in S_t, \forall y \in S_y$	MW

Active power flow in to-direction, existing PSTs	$P_{b,t,y}^{to}$	$\forall b \in S_b, \forall t \in S_t, \forall y \in S_y$	MW
Reactive power flow in from-direction, existing PSTs	$Q_{b,t,y}^{fr}$	$\forall b \in S_b, \forall t \in S_t, \forall y \in S_y$	Mvar
Reactive power flow in to-direction, existing PSTs	$Q_{b,t,y}^{to}$	$\forall b \in S_b, \forall t \in S_t, \forall y \in S_y$	Mvar
Active power flow in from-direction, candidate PSTs	$P_{bc,t,y}^{fr}$	$\forall bc \in S_{bc}, \forall t \in S_t, \forall y \in S_y$	MW
Active power flow in to-direction, candidate PSTs	$P_{bc,t,y}^{to}$	$\forall bc \in S_{bc}, \forall t \in S_t, \forall y \in S_y$	MW
Auxiliary active power flow in from-direction, candidate PSTs	$\tilde{P}_{bc,t,y}^{fr}$	$\forall bc \in S_{bc}, \forall t \in S_t, \forall y \in S_y$	MW
Auxiliary active power flow in to-direction, candidate PSTs	$\tilde{P}_{bc,t,y}^{to}$	$\forall bc \in S_{bc}, \forall t \in S_t, \forall y \in S_y$	MW
Reactive power flow in from-direction, candidate PSTs	$Q_{bc,t,y}^{fr}$	$\forall bc \in S_{bc}, \forall t \in S_t, \forall y \in S_y$	Mvar
Reactive power flow in to-direction, candidate PSTs	$Q_{bc,t,y}^{to}$	$\forall bc \in S_{bc}, \forall t \in S_t, \forall y \in S_y$	Mvar
Phase shift existing PST	$\varphi_{b,t,y}$	$\forall bc \in S_b, \forall t \in S_t, \forall y \in S_y$	Rad
Phase shift candidate PST	$\varphi_{bc,t,y}$	$\forall bc \in S_{bc}, \forall t \in S_t, \forall y \in S_y$	rad
Investment decision	$\alpha_{bc,y} \in \{0,1\}$	$\forall bc \in S_{bc}, \forall y \in S_y$	[-]

Optimization parameters (Existing and candidate)	Symbol	Cardinality	Unit	Typical value
Resistance	$r_{b(c)}$	$\forall b(c) \in S_{b(c)}$	Ohm	[-]
Reactance	$x_{b(c)}$	$\forall b(c) \in S_{b(c)}$	Ohm	[-]
Susceptance	$b_{l(c)}$	$\forall b(c) \in S_{b(c)}$	1 / Ohm	[-]
Thermal rating	$S_{b(c)}^{rated}$	$\forall b(c) \in S_{b(c)}$	MVA	[-]
Emergency rating	$S_{b(c)}^{rated,em}$	$\forall b(c) \in S_{b(c)}$	MVA	[-]
Maximum phase shift	$\varphi_{b(c)}^{max}$	$\forall b(c) \in S_{b(c)}$	Rad	0
Minimum phase shift	$\varphi_{b(c)}^{min}$	$\forall b(c) \in S_{b(c)}$	Rad	[-]

Variables	Symbol	Cardinality	Unit	
Maximum voltage transformation ratio	$\tau_{b(c)}^{max}$	$\forall b(c) \in S_{b(c)}$	[-]	0
Minimum voltage transformation ratio	$\tau_{b(c)}^{min}$	$\forall b(c) \in S_{b(c)}$	[-]	[-]
Maximum angle difference	$\Delta\theta_{bc}^{max}$	$\forall bc \in S_{bc}$	Rad	[-]
Minimum angle difference	$\Delta\theta_{bc}^{min}$	$\forall bc \in S_{bc}$	Rad	[-]
Maximum voltage difference	ΔU_{bc}^{max}	$\forall bc \in S_{bc}$	kV	[-]
Minimum voltage difference	ΔU_{bc}^{min}	$\forall bc \in S_{bc}$	kV	[-]
Maximum active power difference	ΔP_{bc}^{max}	$\forall bc \in S_{bc}$	MW	[-]
Minimum active power difference	ΔP_{bc}^{min}	$\forall bc \in S_{bc}$	MW	[-]
Maximum reactive power difference	ΔQ_{bc}^{max}	$\forall bc \in S_{bc}$	Mvar	[-]
Minimum reactive power difference	ΔQ_{bc}^{min}	$\forall bc \in S_{bc}$	Mvar	[-]
Status	$s_{b(c),t,y}$	$\forall b(c) \in S_{b(c)}, \forall t \in S_t, \forall y \in S_y$	[-]	{0, 1}
Investment cost	$I_{bc,y}$	$\forall bc \in S_{bc}, \forall y \in S_y$	€	[-]
Carbon footprint cost	$FP_{bc,y}^{CO_2}$	$\forall bc \in S_{bc}, \forall y \in S_y$	€	[-]
Landscape impact cost	$LS_{bc,y}$	$\forall bc \in S_{bc}, \forall y \in S_y$	€	[-]
Asset lifetime	$t_{b(c)}^{life}$	$\forall b(c) \in S_{b(c)}$	year	[-]

2.9 DC branch data model

Variables	Symbol	Cardinality	Unit
Active power flow in from-direction, existing lines	$P_{d,t,y}^{fr}$	$\forall d \in S_d, \forall t \in S_t, \forall y \in S_y$	MW
Active power flow in to-direction, existing lines	$P_{d,t,y}^{to}$	$\forall d \in S_d, \forall t \in S_t, \forall y \in S_y$	MW
Active power flow in from-direction, candidate lines	$P_{dc,t,y}^{fr}$	$\forall dc \in S_{dc}, \forall t \in S_t, \forall y \in S_y$	MW

Active power flow in to-direction, candidate lines	$P_{dc,t,y}^{to}$	$\forall dc \in S_{dc}, \forall t \in S_t, \forall y \in S_y$	MW
Investment decision	$\alpha_{dc,y} \in \{0,1\}$	$\forall dc \in S_{dc}, \forall y \in S_y$	[-]

Optimization parameters (Existing and candidate)	Symbol	Cardinality	Unit	Typical value
Resistance	$r_{d(c)}$	$\forall d(c) \in S_{d(c)}$	Ohm	[-]
Admittance	$g_{d(c)}$	$\forall d(c) \in S_{d(c)}$	1/Ohm	[-]
Thermal rating	$S_{d(c)}^{rated}$	$\forall d(c) \in S_{d(c)}$	MVA	[-]
Emergency rating	$S_{d(c)}^{rated,em}$	$\forall d(c) \in S_{d(c)}$	MVA	[-]
Failure rate	$\lambda_{d(c)}$	$\forall d(c) \in S_{d(c)}$	1 / year	[-]
Mean time to repair	$t_{d(c)}^{MTTR}$	$\forall d(c) \in S_{d(c)}$	h	[-]
Mean time between failures	$t_{d(c)}^{MTBF}$	$\forall d(c) \in S_{d(c)}$	h	[-]
Status	$s_{d(c),t,y}$	$\forall d(c) \in S_{d(c)}, \forall t \in S_t, \forall y \in S_y$	[-]	{0, 1}
Investment cost	$I_{dc,y}$	$\forall dc \in S_{dc}, \forall y \in S_y$	€	[-]
Carbon footprint cost	$FP_{dc,y}^{CO_2}$	$\forall dc \in S_{dc}, \forall y \in S_y$	€	[-]
Landscape impact cost	$LS_{dc,y}$	$\forall dc \in S_{dc}, \forall y \in S_y$	€	[-]
Asset lifetime	$t_{d(c)}^{life}$	$\forall d(c) \in S_{d(c)}$	year	[-]

2.10 DC converter data model

Variables	Symbol	Cardinality	Unit
AC side active power injection existing, converter	$P_{z,t,y}^{ac}$	$\forall z \in S_z, \forall t \in S_t, \forall y \in S_y$	MW
AC side reactive power injection, existing converter	$Q_{z,t,y}^{ac}$	$\forall z \in S_z, \forall t \in S_t, \forall y \in S_y$	Mvar
DC side active power injection, existing converter	$P_{z,t,y}^{dc}$	$\forall z \in S_z, \forall t \in S_t, \forall y \in S_y$	MW
AC side active power injection, candidate converter	$P_{zc,t,y}^{ac}$	$\forall zc \in S_{zc}, \forall t \in S_t, \forall y \in S_y$	MW

AC side reactive power injection, candidate converter	$Q_{zc,t,y}^{ac}$	$\forall zc \in S_{zc}, \forall t \in S_t, \forall y \in S_y$	Mvar	
DC side active power injection, candidate converter	$P_{zc,t,y}^{dc}$	$\forall zc \in S_{zc}, \forall t \in S_t, \forall y \in S_y$	MW	
Investment decision	$\alpha_{zc,y} \in \{0,1\}$	$\forall zc \in S_{zc}, \forall y \in S_y$	[-]	

Optimization parameters (Existing and candidate)	Symbol	Cardinality	Unit	Typical value
Auxiliary converter losses	$L_{z(c)}^a$	$\forall z(c) \in S_{z(c)}$	MW	
Linear converter losses	$L_{z(c)}^b$	$\forall z(c) \in S_{z(c)}$	MW/MW	
Quadratic converter losses	$L_{z(c)}^c$	$\forall z(c) \in S_{z(c)}$	MW/MW^2	
Thermal rating AC	$S_{z(c)}^{rated,ac}$	$\forall z(c) \in S_{z(c)}$	MVA	[-]
Emergency rating AC	$S_{z(c)}^{rated,em,ac}$	$\forall z(c) \in S_{z(c)}$	MVA	[-]
Rated active power AC	$P_{z(c)}^{rated,ac}$	$\forall z(c) \in S_{z(c)}$	MW	[-]
Rated reactive power AC	$Q_{z(c)}^{rated,ac}$	$\forall z(c) \in S_{z(c)}$	Mvar	[-]
Emergency rating active power AC	$P_{z(c)}^{rated,em,ac}$	$\forall z(c) \in S_{z(c)}$	MW	[-]
Emergency rating reactive power AC	$Q_{z(c)}^{rated,em,ac}$	$\forall z(c) \in S_{z(c)}$	Mvar	[-]
Thermal rating DC	$P_{z(c)}^{rated,dc}$	$\forall z(c) \in S_{z(c)}$	MW	[-]
Emergency rating DC	$P_{z(c)}^{rated,em,dc}$	$\forall z(c) \in S_{z(c)}$	MW	[-]
Status	$S_{z(c),t,y}$	$\forall z(c) \in S_{z(c)}, \forall t \in S_t, \forall y \in S_y$	[-]	{0, 1}
Investment cost	$I_{zc,y}$	$\forall zc \in S_{zc}, \forall y \in S_y$	€	[-]
Carbon footprint cost	$FP_{zc,y}^{CO_2}$	$\forall zc \in S_{zc}, \forall y \in S_y$	€	[-]
Landscape impact cost	$LS_{zc,y}$	$\forall zc \in S_{zc}, \forall y \in S_y$	€	[-]
Asset life time	$t_{z(c)}^{life}$	$\forall z(c) \in S_{z(c)}$	year	[-]

2.11 Air quality cost model

Variables/Data	Symbol	Cardinality	Unit	Notes
Emission Factor	$G_{g,p}^{ef}$	$\forall g \in S_g, \forall p \in S_p$	[kg/MWh]	Can be expressed also in terms of emissions by fuel consumption (kg/fuel ton) and generator efficiency (MWh/fuel ton).
Emission	$EM_{g,t,y,p}$	$\forall g \in S_g, \forall t \in S_t, \forall y \in S_y, \forall p \in S_p$	[kg/h]	Hourly emission rate for each pollutant.
Sensitivity coefficient	$SC_{g,t,y,p}$	$\forall g \in S_g, \forall t \in S_t, \forall y \in S_y, \forall p \in S_p$	[μg/m³/Kg]	Linear relationship between emission and concentration.
Hourly air quality concentration	$AQ_{g,t,y,p}$	$\forall g \in S_g, \forall t \in S_t, \forall y \in S_y, \forall p \in S_p$	[μg/m³]	Hourly concentration of each pollutant.
Yearly mean air quality concentration	$\overline{AQ}_{g,y,p}$	$\forall g \in S_g, \forall y \in S_y, \forall p \in S_p$	[μg/m³]	Yearly mean concentration of each pollutant.
Impact coefficient	$IMC_{p,imp}$	$\forall p \in S_p, \forall imp \in S_{imp}$	[#/μg/m3]	Expresses the variation of a health indicator (e.g., number of years of life lost) for a pollutant concentration variation.
Impact	$\overline{IM}_{g,y,p,imp}$	$\forall g \in S_g, \forall y \in S_y, \forall p \in S_p, \forall imp \in S_{imp}$	[#]	Expresses the health impact (e.g. number of years of life lost).
Cost coefficient	CC_{imp}		[euro/#]	Expresses the monetary evaluation of the corresponding variation of an impact indicator.
Cost	$\overline{C}_{g,y,p,imp}^{AQ}$	$\forall imp \in S_{imp}$	[euro]	Expresses the monetary evaluation of a health impact.
Emission inventory	EI	$\forall cy \in S_{cy}, \forall p \in S_p, \forall y \in S_y$	[tons/y]	Requires also spatial, temporal and speciation profiles.
Meteorological fields	MF	$\forall ijk \in S_{ijk}, \forall m \in S_m$	[-]	Gridded (2D/3D) hourly meteorological fields of all variables needed by air quality models.

Population data	Pop	$\forall ijk \in S_{ijk}, \forall y \in S_y$	[-]	Spatially distributed; it may also include age classes and impact reference data.
Cost of CO2 in reference year y	$\theta_y^{CO_2}$	$\forall y \in S_y$	[euro/t]	

3 Optimization target function

This section focuses on the objective function of the FlexPlan planning tool. The rest of the model (variables, constraints, bounds) is detailed throughout the Chapters 4-7.

3.1 General structure of the optimization target function

The FlexPlan planning tool aims at seeking out an optimal combination of new grid investments, both in classical new connections and installation of flexibility devices, to achieve maximum social welfare. Thus, the first logical choice for the objective of the optimization model upon which the planning tool is based is the maximization of the global social welfare of the system, since it is able to capture positive effects for all the system participants. A general formulation of the social welfare is shown in equation (1)

$$\sum_t \left[\sum_i (E_{t,i} - C_{t,i}) + \sum_j \alpha_j (E_{t,j} - C_{t,j}) - \sum_n K_{t,n} \right] - \sum_j \alpha_j I_j, \tag{1}$$

where the index t refers to the hours of the year, i to existing devices in the system, j to the new devices, for which installation has to be evaluated via the optimization, and n to the nodes. The term E indicates any benefit for the social welfare derived from the operation of devices, while C indicates any cost of the system derived from the operation of devices. The term K indicates costs of the system related to single nodes derived from system operation. Finally, the term I indicates investment costs for new devices. α_j is the integer decision variable indicating if the new device j should be installed or not.

Equation (1) allows to highlight some important peculiarities of the objective function. First of all, two kinds of effects on the global welfare are considered: time dependent (the terms in square brackets in equation (1)), including what is related to energy production and consumption and to the operation of the system in general, and time independent, that is investment costs for new installations (the terms $\alpha_j I_j$).

Furthermore, new installed devices may entail benefits and costs to the system derived from their operation.

It must be remarked that by default the planning tool will use an hourly time resolution for the optimization of the grid investments. This resolution has the advantages of being able to detect correctly system congestion in most cases, while preserving numerical tractability which is of high importance in regard of the large-scale problems at hand. Also, for long-term models looking 30 years ahead a smaller time resolution is hardly justifiable, as the macro scenarios will yield much higher uncertainty.

3.2 Dynamic Optimization

The planning tool developed should consider a long planning horizon covering multiple decades accounting for the effects of both the new installations performed over time and of the evolution of the system. To that aim, a number $y \in S_y$ of simulation years are considered and simulated together as shown in equation (2)

$$\sum_{y \in S_y} \left\{ f_y^{d,o} \sum_{t \in S_t} \left[\sum_i (E_{t,i} - C_{t,i}) + \sum_j \alpha_j (E_{t,j} - C_{t,j}) - \sum_n K_{t,n} \right] - f_y^d \sum_j \alpha_j I_j \right\} \qquad (2)$$

Note that the effects of each element of the network are considered only in those reference years in which it is operational, and investment costs are referred only to the reference years during which the installation takes place.

The set of years S_y in the objective function (2) are referred to as target years. The operation of the system is simulated explicitly only for this set of years, but each target year represents N_y actual operational years. In the FlexPlan project, the target years will be S_y = {2030, 2040, 2050}[1] so that each of the target years represent N_y = 10 operational years and the objective function represents a full 30-year planning horizon. The pre-factors $f_y^{d,o}$ and f_y^d for the operational cost terms and the investment cost terms, respectively, are introduced to weight each target year according to the operational years they represent and to calculate the present value of costs incurred in different years in the future. These pre-factors are explained in more detail in Section 3.8.

3.3   The cost minimization

The explicit quantification of future benefits would be hard to model, requiring the usage of a detailed market model. Also, depending on the objective of the system operator or the flexibility resource provider, the benefits associated with that source might not necessarily coincide with the social welfare. As such, within the FlexPlan project, only system costs are considered and minimized, since almost all the benefits from the social welfare perspective will be due to the reduction of the operation costs due to a more efficient use of energy in the system. Thus, it is more useful to consider the total system cost as the objective of the optimization problem. In this case, equation (3), representing the system total costs is used instead of (1), and minimized in the optimization process.

$$\sum_{y \in S_y} \left\{ \sum_{t \in S_t} \left[\sum_i C_{t,i} + \sum_{j \in S_j} \alpha_j C_{t,j} + \sum_{n \in S_n} K_{t,n} \right] + \sum_{j \in S_j} \alpha_j I_j \right\} \qquad (3)$$

3.4   Complete formulation

Equation (4) presents the complete formulation of the objective function including the detailed cost components of existing and candidate devices.

---

[1] The generality of the model however allows to consider any set S_y of target years.

$$\sum_{y\in S_y}\left\{ f_y^{d,o}\sum_{t\in S_t}\left[\sum_{g\in S_g}\left[C_{g,y}^{aq}+(\theta^{CO_2}G^{pf}+\theta^f)\eta_g^f\right]P_{g,t,y}+C_{g,y}^{res,curt}\Delta P_{g,t,y}^{res}\right.\right.$$

$$+\sum_{j\in S_j}\left[C_{j,t,y}^{abs}P_{j,t,y}^{abs}+C_{j,t,y}^{inj}P_{j,t,y}^{inj}\right]+\sum_{j\in S_{jc}}\left[C_{jc,t,y}^{abs}P_{jc,t,y}^{abs}+C_{jc,t,y}^{inj}P_{jc,t,y}^{inj}\right]$$

$$+\sum_{u\in S_u}\left[C_{u,t,y}^{nce}\left(P_{u,t,y}^{ref}-P_{u,t,y}^{nce}\right)+C_{u,t,y}^{ds}\left(\Delta P_{u,t,y}^{ds,up}+\Delta P_{u,t,y}^{ds,dn}\right)+C_{u,t,y}^{lc}\Delta P_{u,t,y}^{lc}\right]$$

$$\left.+\sum_{n\in S_n}\left(C_{n,t,y}^{EE}EE_{n,t,y}+C_{n,t,y}^{LL}LL_{n,t,y}\right)\right] \qquad (4)$$

$$+f_y^d\left[\sum_{j\in S_{jc}}\alpha_{jc,y}\left(I_{jc,y}^E\left(E_j^{\max}\right)+I_{jc,y}^P\left(P_j^{\max}\right)+FP_{jc,y}^{CO_2}\right)+\sum_{u\in S_u}\alpha_{u,y}\left(I_{u,y}+FP_{u,y}^{CO_2}\right)\right.$$

$$+\sum_{lc\in S_{lc}^{ac}}\alpha_{lc,y}\left(I_{lc,y}+FP_{lc,y}^{CO_2}+LS_{lc,y}\right)+\sum_{dc\in S_{lc}^{dc}}\alpha_{dc,y}\left(I_{dc,y}+FP_{dc,y}^{CO_2}+LS_{dc,y}\right)$$

$$\left.\left.+\sum_{zc\in S_{zc}}\alpha_{zc,y}\left(I_{zc,y}+FP_{zc,y}^{CO_2}+LS_{zc,y}\right)+\sum_{bc\in S_{bc}}\alpha_{bc,y}\left(I_{bc,y}+FP_{bc,y}^{CO_2}+LS_{bc,y}\right)\right]\right\}$$

3.5 Objective function terms explanation

In this section the terms introduced in equation (4) will be described in detail. As mentioned previously, two kinds of costs are considered in the optimisation objective.

Firstly, time dependent costs, that result from the operation of the system and of the generation/consumption devices connected to it and so are labelled in the following as "operational costs".

They include the costs derived from the operation of the system that are not accounted for in other ways. For example, costs related to network losses are implicitly considered by the further generation costs due to the increase in generation needed to cover the load; so, no explicit costs related to losses will be present in the objective function. Regarding grid losses, in our network models (see Chapter 4), only storage assets and converters will be considered as lossy. Indeed, due to the need to retain model linearity (which is necessary in order to maintain numerical tractability for the large-scale problems to be solved), we use DC network approximations which do not allow to model line losses. The main objective of the planning tool is to find best trade-off between classical and flexibility systems (and not the minimisation of losses), and given the fact that only a limited set of expansions will be realised in the future grid, the effect on the overall system losses with respect to the reference grid will also be very limited.

Time independent costs, labelled as "fixed" in the following paragraphs, are considered only for candidate devices. Possible candidates include storage j, flexibility resources u and new lines l, both AC and DC, PSTs b and HVDC converter stations z. For each of them, fixed costs include the investment cost I_*, the carbon footprint costs $FP_*^{CO_2}$ and possible landscape impact costs LS_*.

3.5.1 Thermal generators

No new installation of conventional thermal generators is considered within FlexPlan project, so only operational costs are considered. These costs are considered proportional to the hourly production $P_{g,t,y}$, and a specific term $C_{g,t,y}$ is introduced, which includes the following contributions:

- fuel costs $C_{g,y}^f = \left(\theta_y^{CO_2} G^{pf} + \theta_y^f\right)\eta_g^f$, where θ_y^f is the price of fuel f for the reference year y [€/MWh], θ^{CO_2} is the price of CO_2 emissions for the reference year y [€/t$_{CO2}$], G^{pf} is the emission factor of fuel f [t$_{CO2}$/MWh] and η_g^f is the specific consumption of fuel f for generator g [MWh/MWh];

- environmental costs $C_{g,y}^{aq}$, due to the effects of other pollutants on the air quality - they depend on the fuel considered, on the generation technology and on the geographical localization of the generator - and will be described in Section 6.1.

Then

$$\left[C_{g,y}^{aq} + (\theta^{CO_2} G^{pf} + \theta^f)\eta_g^f\right]P_{g,t,y} \tag{5}$$

represents the generation costs of generator g at time t in planning year y.

3.5.2 Hydro generators

No costs are ascribed to hydro generators, neither operational ones nor fixed ones, since no new installation of hydro generation is considered for FlexPlan project.

3.5.3 Renewable generation

The renewable generation is presented using time series of reference generation values $P_{g,t,y}^{res,ref}$. As the aim of the planning model is the maximisation of renewable infeed into the system by means of investing into power networks and flexibility, in the objective function, possible curtailment of renewable energy (e.g. due to contingencies) has been penalized using the term $C_{g,y}^{res,curt}\Delta P_{g,t,y}^{res}$, where $C_{g,y}^{res,curt}$ corresponds to the curtailment cost of renewables and $\Delta P_{g,t,y}^{res}$ corresponds to the generation curtailment $\forall g \in S_g^{res}, \forall t \in S_t, y \in S_y$. $\Delta P_{g,t,y}^{res}$ is defined as the difference between the actual generation of the RE source and the reference generation, $\Delta P_{g,t,y}^{res} = P_{g,t,y}^{res,ref} - P_{g,t,y}^{res}$ and $0 \leq \Delta P_{g,t,y}^{res} \leq P_{g,t,y}^{res,ref} \forall g \in S_g^{res}, \forall t \in S_t, \forall y \in S_y$ The total amount of renewable energy curtailment can be limited using:

$\sum_{t \in S_t}\sum_{g \in S_g^{res}}\Delta P_{g,t,y}^{res} \leq P_{g,y}^{res,curt,max}$ $\forall y \in S_y$ here $P_{g,y}^{res,curt,max}$ the maximum total renewable curtailment allowed for each year of the planning horizon.

3.5.4 Storage devices

3.5.4.1 Operational Costs

Maintenance costs and other operational costs can be considered for energy storage devices. In this case, costs could be counted both for charging and for discharging. No environmental cost is considered for storage devices, since no polluting emissions are related to charging or discharging. Then

$$C_{j,t,y}^{abs} P_{j,t,y}^{abs} + C_{j,t,y}^{inj} P_{j,t,y}^{inj} \qquad (6)$$

are the terms to be considered in the objective function, where $P_{j,t,y}^{abs}$ is the energy charged in hour t and reference year y by storage device j and $P_{j,t,y}^{inj}$ is the energy discharged in hour t and reference year y by storage device s; $C_{j,t,y}^{abs}$ is the charging cost for storage device j in time t and reference year y [€/MWh] and $C_{j,t,y}^{inj}$ is the discharging cost for storage device j in time t and reference year y [€/MWh]. Both these costs are usually set to zero, since for this kind of devices operational costs are mostly reflected in the initial investment costs.

3.5.4.2 Fixed Costs

Investment costs for many kinds of storage devices may have two components: one being a function of the total installed capacity E_j^{max} d one being a function of the nominal power P_j^{max} lso, carbon footprint costs have to be considered, so that the total fixed costs for a storage device are given by:

$$I_{jc,y}^E\left(E_{jc}^{\max}\right) + I_{jc,y}^P\left(P_{jc}^{\max}\right) + FP_{jc,y}^{CO_2} \qquad (7)$$

where the $y \in S_y$ subscript indicates the particular year in the planning horizon, as some candidate elements can appear in multiple planning horizons. Note that the costs $I_{jc,y}^E$ are discounted accordingly for each $y \in S_y$. As the planning model is going to use a number of candidate storage devices with different power ratings and energy contents, in the actual model implementation the fixed costs for storage candidates will take the form

$I_{jc,y} + FP_{jc,y}^{CO_2}$ where $I_{jc,y}$ is the sum of the evaluated functions $I_{jc,y}^E\left(E_{jc}^{\max}\right)$ and $I_{jc,y}^P\left(P_{jc}^{\max}\right)$ at given power and energy ratings.

3.5.5 Demand flexibility resources

As currently the amount of existing demand flexibility is very limited and data about characteristics are not available, the objective function only considers flexibility elements as future candidates. Flexibility resources modelling is described thoroughly in section 5.1. Here we recall briefly only the equations needed to describe the terms included in the objective function (4).

3.5.5.1 Operational costs

Two kinds of load flexibility can be considered by the planning tool: load decrease and load shifting.

In the first case, the flexibility resource f simply reduces its consumption $P_{u,t,y}^{ref}$ in hour t of reference year y by the amount $\Delta P_{u,t,y}^{nce}$ and to do so it receives a remuneration proportional to the reduction of the consumption, with a compensation $C_{u,t,y}^{nce}$ [€/MWh] so that the corresponding cost for the system is given by:

$$C_{u,t,y}^{nce} \Delta P_{u,t,y}^{nce} \qquad (8)$$

In the second case, the flexibility resource changes its load profile shifting part of its consumption $P_{u,t,y}^{ref}$, by increasing it of an amount $\Delta P_{u,t,y}^{ds,up}$ or reducing it by an amount $\Delta P_{u,t,y}^{ds,dn}$ from hour t to another hour τ (it can be

as well $t < \tau$ and $\tau < t$) but maintaining the original total consumption within a given period T such that $t, \tau \in$ T, i.e

$$\sum_T P_{u,t,y}^{ds} = \sum_T P_{u,t,y}^{ref} \tag{9}$$

Here a remuneration proportional to load shifted is considered, with a compensation $C_{u,t,y}^{ds}$ [€/MWh], so that the cost for the system in each hour t is given by

$$C_{u,t,y}^{ds}\left(\Delta P_{u,t,y}^{ds,up} + \Delta P_{u,t,y}^{ds,dn}\right) \tag{10}$$

Note that a flexibility resource could be in principle able to participate to both the flexibility mechanisms (load shifting and load reduction).

3.5.5.2 Fixed Costs

Fixed costs for flexibility devices include initial investment costs and, depending on the technology considered, carbon footprint costs.

$$I_{u,y} + FP_{u,y}^{CO_2} \tag{11}$$

where the subscript y indicates the first year of installation.

3.5.6 Load curtailment

Load curtailment mechanism and modelling are described in section 5.1.3. Here only a brief recall useful to describe terms included in objective function (4) are presented.

3.5.6.1 Operational costs

Concerning flexibility resources already dealt with in section 3.5.5 since $C_{u,t,y}^{lc} \gg C_{u,t,y}^{ds}$ and $C_{u,t,y}^{lc} \gg C_{u,t,y}^{nce}$, we expect that load curtailment will be activated after those two other mechanisms are completely exploited, that is after the lower bound $P_{u,t,y}^{ref} - \Delta P_{u,t,y}^{ds,dn} - \Delta_{u,t,y}^{nce,max}$ reached. Then, the cost for the system in hour t is given by

$$C_{u,t,y}^{lc}\Delta P_{u,t,y}^{lc} \tag{12}$$

Note that the decision (slack) variable $\Delta P_{u,t,y}^{lc}$ is defined for any load, also the non-flexible loads that are not described by the demand flexibility model. This means that for any load bus, load can be curtailed if necessary, to avoid infeasibility (the curtailment costs can however vary between loads).

3.5.6.2 Fixed costs

For newly installed curtailable loads, since they are also flexible loads, their fixed costs are already considered by equation (11). Existing curtailable loads, instead, do not have installation costs.

3.5.7 Nodal slacks

To ensure the feasibility of the problem considered, two slack variables could be introduced in each nodal balance constraint, $EE_{n,t,y}$ [MWh], to take into account the situations in which there is an excess of generation in node n at hour t for reference year y, and $LL_{n,t,y}$ [MWh], to consider those situations in which it is not possible to fulfil the load in hour t in node n for reference year y.

These terms are valorised with very high specific costs $C_{n,t,y}^{EE}$ [€/MWh] and $C_{n,t,y}^{LL}$ [€/MWh], respectively, so that the costs for the System are given by

$$C_{n,t,y}^{EE} EE_{n,t,y} + C_{n,t,y}^{LL} LL_{n,t,y} \qquad (13)$$

This way it is ensured that these nodal slacks are not chosen by the optimisation solver as means of differing necessary investments. It would also mean that, if these slack variables become nonzero, additional candidates should be added to the set of candidates to resolve the infeasibilities.

Because the model formulation also includes load curtailment slack variables for all load demand buses and generation curtailment slack variables for all non-dispatchable generation buses, it is not foreseen that additional nodal slack variables will be needed. They are however described here as an optional model feature that could be implemented if it is found that additional means to avoid infeasibility are necessary.

3.5.8 AC and DC lines

Operational costs for lines are substantially related only to losses, so they are already implicitly considered by other operational costs. So only fixed costs for new installations are considered in the objective function (4). Then, for each candidate AC line lc and for each candidate DC line dc the investment cost, the carbon footprint cost and the landscape impact cost should be considered for the year of installation y:

$$I_{lc,y} + FP_{lc,y}^{CO_2} + LS_{lc,y} \qquad (14)$$

for AC lines and

$$I_{dc,y} + FP_{dc,y}^{CO_2} + LS_{dc,y} \qquad (15)$$

for DC lines.

3.6 Stochastic Formulation

It is also possible to implement a stochastic formulation of the optimization model. In this case what will be minimized are the expected costs deriving from different possible scenarios s, each with a probability π_s. The term "scenarios" here (sometimes also referred to as Monte Carlo variants) indicates different possible realizations of the parameters, in particular, but not only, renewable generation and load profiles. Thus, the expected costs are calculated as a weighted sum of the costs in each scenario, with the weights corresponding to the scenario probabilities. Note that the operational model provides the operational costs for each scenario, whereas the investment decisions are taken across all scenarios alike. This is shown in equation (16).

$$\sum_s \pi_s \left\{ \sum_{y \in S_y} f_y^{d,o} \left\{ \sum_{t \in S_t} \left[\sum_{g \in S_g} [C_{g,y}^{aq} + (\theta^{CO_2} G^{pf} + \theta^f)\eta_g^f] P_{g,t,y,s} + C_{g,y}^{res,curt} \Delta P_{g,t,y,s}^{res} \right. \right. \right.$$

$$+ \sum_{j \in S_j} [C_{j,t,y}^{abs} P_{j,t,y,s}^{abs} + C_{j,t,y}^{inj} P_{j,t,y,s}^{inj}]$$

$$+ \sum_{j \in S_{jc}} [C_{jc,t,y}^{abs} P_{jc,t,y,s}^{abs} + C_{jc,t,y}^{inj} P_{j,ct,y,s}^{inj}]$$

$$+ \sum_{u \in S_u} [C_{u,t,y}^{nce} (P_{u,t,y,s}^{ref} - P_{u,t,y,s}^{nce}) + C_{u,t,y}^{ds} (\Delta P_{u,t,y,s}^{ds,up} + \Delta P_{u,t,y,s}^{ds,dn})]$$

$$\left. + C_{u,t,y}^{lc} \Delta P_{u,t,y}^{lc}] + \sum_{n \in S_n} (C_{n,t,y}^{EE} EE_{n,t,y,s} + C_{n,t,y}^{LL} LL_{n,t,y,s}) \right]$$

$$+ f_y^d \left[\sum_{j \in S_{jc}} \alpha_{jc,y} (I_{jc,y}^E (E_{jc}^{max}) + I_{jc,y}^P (P_{jc}^{max}) + FP_{jc,y}^{CO_2}) \right.$$

$$+ \sum_{u \in S_u} \alpha_{u,y} (I_{u,y} + FP_{u,y}^{CO_2}) + \sum_{lc \in S_{lc}^{ac}} \alpha_{lc,y} (I_{lc,y} + FP_{lc,y}^{CO_2} + LS_{lc,y})$$

$$+ \sum_{dc \in S_{lc}^{dc}} \alpha_{dc,y} (I_{dc,y} + FP_{dc,y}^{CO_2} + LS_{dc,y})$$

$$+ \sum_{zc \in S_{zc}} \alpha_{zc,y} (I_{zc,y} + FP_{zc,y}^{CO_2} + LS_{zc,y})$$

$$\left. \left. \left. + \sum_{bc \in S_{bc}} \alpha_{bc,y} (I_{bc,y} + FP_{bc,y}^{CO_2} + LS_{bc,y}) \right] \right\} \right\}$$

(16)

As will be presented Section 8.1.1, the stochastic problem can be divided in an upper-level problem (the optimal investment problem) and lower-level problems (the operational problems in each scenario, given the investment decisions taken in the upper problem). This yields a Benders decomposition which, with a specific solving methodology and a potential parallelization of the lower-level problems, can reduce the computational burden to reach an optimal solution of the network expansion problem.

3.7 Costs related to reliability of supply

Reliability of supply and probabilistic reliability criteria are proposed to be accounted for in the planning tool primarily through the objective function. This can be done by introducing a term for the expected cost of energy not supplied (CENS), i.e., the expected interruption costs for the demand elements. These costs approximate the socio-economic costs of power supply interruptions due to contingencies (e.g., transmission line failures). The set of contingencies that are considered in the planning problem is denoted with S_c and is an input to the planning tool. We define this set to always include the non-contingency state $c = 0$, which represents the intact grid without any components in an outage state. If one considers a "trivial" contingency set $S_c = \{0\}$ as input to the planning tool, this means that no contingencies are considered.

If one formulates a separate term in the objective function for the cost of energy not supplied, it would take the following form:

$$f_{\text{ENS}} = \sum_{y \in S_y} \sum_{t \in S_t} \Delta t \sum_{c \in S_c \backslash \{0\}} \widetilde{U}_{c,t,y} \sum_{u \in S_u} C_{u,t,y}^{voll} \Delta P_{u,c,t,y}$$

Here, the summation goes over the three target years y (i.e., 2030, 2040, 2050), all time steps $t \in S_t$ in each year, all considered contingency states c (excluding the non-contingency state $\backslash \{0\} \in S_c$), and all demand elements $u \in S_u$. $\widetilde{U}_{c,t,y}$ is the unavailability probability of contingency c during time step t in year y. This quantity represents the probability that the components involved in the contingency c are all in an outage state at time step t. For an N-1 (first-order or single-component) contingency involving a component l, this unavailability probability is the forced outage rate of the component and can be calculated as $\widetilde{U}_{c,t,y} \approx \frac{\lambda_l t_l^{MTTR}}{8760}$. For an N-2 (second-order or double-component) contingency involving components l and m, $\widetilde{U}_{c,t,y} \approx \frac{\lambda_l \lambda_m t_l^{MTTR} t_l^{MTTR}}{8760^2}$. These approximate expressions are very good approximations when $\frac{\lambda_l t_l^{MTTR}}{8760} \ll 1$. Here the failure rates λ_l are measured in units of 1/year (i.e., failures per year) and the mean times to repair t_l^{MTTR} are measured in units of hours, and the factors of 8760 are included in the formulas to convert correctly between units (hours and years) so that $\widetilde{U}_{c,t,y}$ is unitless. The factor $\Delta t = 1$ h is included in the calculation of the objective function term f_{ENS} so that it expresses the expected cost of energy not supplied during one year as measured in €.
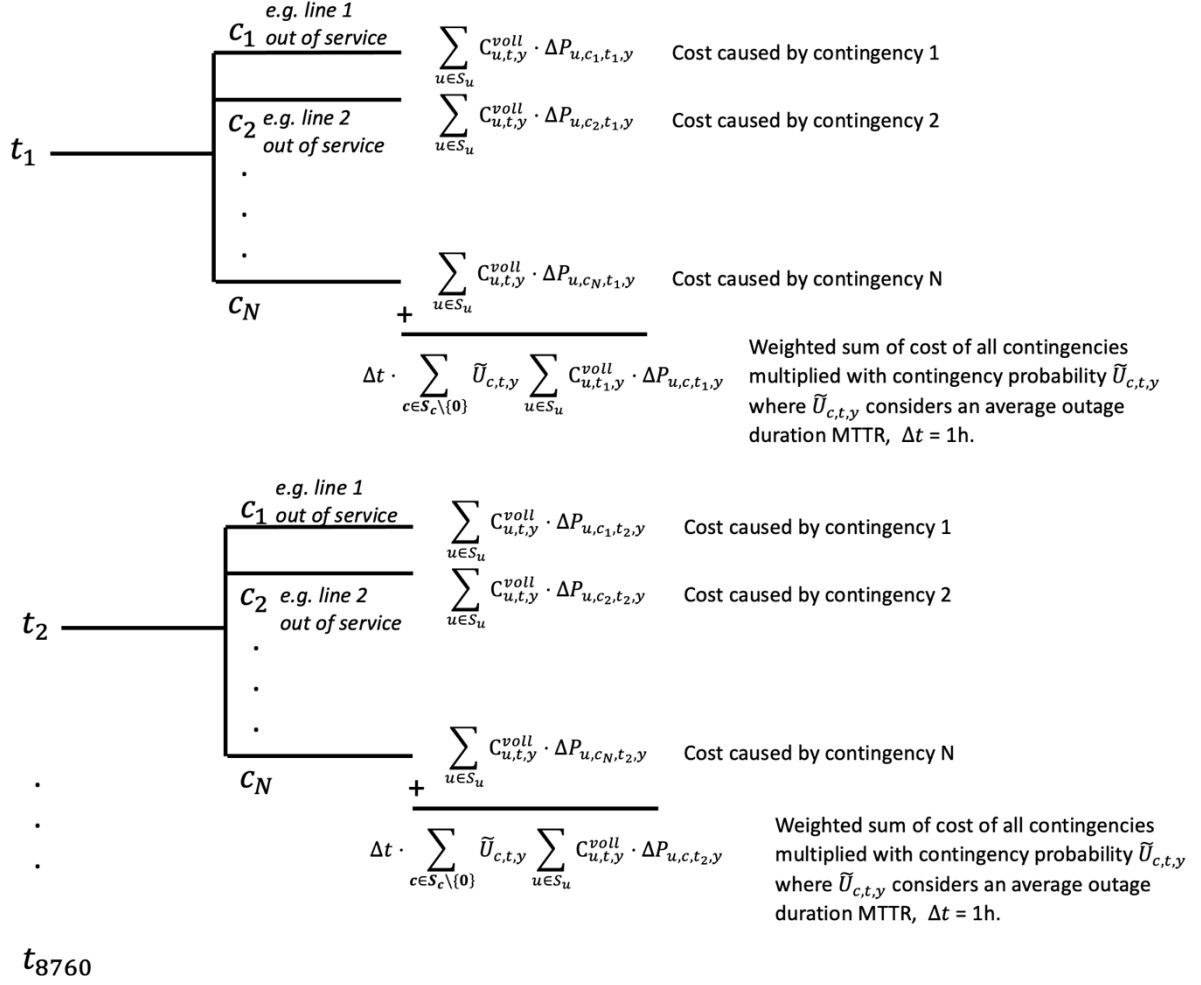
t_1 ——

c_1 e.g. line 1 out of service

$$\sum_{u \in S_u} C^{voll}_{u,t,y} \cdot \Delta P_{u,c_1,t_1,y}$$ Cost caused by contingency 1

c_2 e.g. line 2 out of service

$$\sum_{u \in S_u} C^{voll}_{u,t,y} \cdot \Delta P_{u,c_2,t_1,y}$$ Cost caused by contingency 2

.
.
.

c_N

$$\sum_{u \in S_u} C^{voll}_{u,t,y} \cdot \Delta P_{u,c_N,t_1,y}$$ Cost caused by contingency N

$+$

$$\Delta t \cdot \sum_{c \in S_c \setminus \{0\}} \widetilde{U}_{c,t,y} \sum_{u \in S_u} C^{voll}_{u,t_1,y} \cdot \Delta P_{u,c,t_1,y}$$

Weighted sum of cost of all contingencies multiplied with contingency probability $\widetilde{U}_{c,t,y}$ where $\widetilde{U}_{c,t,y}$ considers an average outage duration MTTR, Δt = 1h.

t_2 ——

c_1 e.g. line 1 out of service

$$\sum_{u \in S_u} C^{voll}_{u,t,y} \cdot \Delta P_{u,c_1,t_2,y}$$ Cost caused by contingency 1

c_2 e.g. line 2 out of service

$$\sum_{u \in S_u} C^{voll}_{u,t,y} \cdot \Delta P_{u,c_2,t_2,y}$$ Cost caused by contingency 2

.
.
.

c_N

$$\sum_{u \in S_u} C^{voll}_{u,t,y} \cdot \Delta P_{u,c_N,t_2,y}$$ Cost caused by contingency N

$+$

$$\Delta t \cdot \sum_{c \in S_c \setminus \{0\}} \widetilde{U}_{c,t,y} \sum_{u \in S_u} C^{voll}_{u,t,y} \cdot \Delta P_{u,c,t_2,y}$$

Weighted sum of cost of all contingencies multiplied with contingency probability $\widetilde{U}_{c,t,y}$ where $\widetilde{U}_{c,t,y}$ considers an average outage duration MTTR, Δt = 1h.

.
.
.

t_{8760}

Figure 3-1 - Structure of the cost of energy not supplied term in the objective function

Furthermore, $C^{voll}_{u,t,y}$ is the specific cost of energy not supplied for demand element (i.e., load point or delivery point) u due to power interruptions during a contingency (e.g., € per MWh of load lost). The (slack) variable $\Delta P_{u,c,t,y}$ is the power interrupted (in MW) for demand element u at a time step t during contingency c. This is a decision variable that is determined endogenously by the optimization model for each time step as defined by the nodal power balance equation in Section 4.2. The interrupted power can be expressed as the difference between the power actually supplied for the demand element at time t during the contingency, $P_{u,c,t,y}$, and the demand at that time in the non-contingency state, $P_{u,c=0,t,y}$. For the simplifying case that there is no load curtailment and no flexibility activation at the time step in the non-contingency case, one has $P_{u,c=0,t,y} = P^{ref}_{u,t,y}$ and $P_{u,c,t,y} = P^{flex}_{u,t,y}$ and thus $\Delta P_{u,c,t,y} = P^{ref}_{u,t,y} - P^{flex}_{u,c,t,y} \geq 0$. The logic behind the structure of the summation of the contribution from costs of energy to the objective function, according to the formula for f_{ENS} above, is illustrated in Figure 3-1 (here only the summation over all time steps for one of the target years is illustrated.)

In the general objective function formulation, the cost of energy not supplied would correspond to a term

$$K_{t,y}^{ens} = \Delta t \sum_{c \in S_c \setminus \{0\}} \widetilde{U}_{c,t,y} \sum_{u \in S_u} C^{voll}_{u,t,y} \Delta P_{u,c,t,y}$$

in the sum $K_{t,y} = K_{t,y}{}^{ds} + K_{t,y}{}^{nc} + K_{t,y}{}^{ens} + \cdots$ comprising the load shifting and curtailment contributions to the operational costs.

Note that the sum over contingencies in the formulas above could equivalently also have included the non-contingency state $c = 0$, i.e., running over $c \in S_c \backslash \{0\}$ instead of $c \in S_c$. However, in that case the contribution to the sum from $c = 0$ would have been zero assuming that the slack variable $\Delta P_{u,c,t,y} = 0$ in an intact grid without any contingencies.

If one wants the objective to capture that generation dispatch is modified during contingencies, this can be done in a somewhat simplified manner by modifying the generation cost term in the objective function in Section 3.4 (described for thermal generators in Section 3.5.1) as follows:

$$\sum_{c \in S_c} \widetilde{U}_{c,t,y} \sum_g [C_{g,y}^{aq} + (\theta^{CO_2} G^{pf} + \theta^f)\eta_g^f] P_{g,t,y}$$

Note that the sum here goes over the entire contingency set S_c including the non-contingency state $c = 0$. The non-contingency state corresponds to the intact grid, or in other words the state implicitly assumed in Section 3.4. The difference is that the modified term above accounts for the fact that the probability of occurrence for the non-contingency state is $\widetilde{U}_{c=0,t,y} = 1 - \sum_{c \in S_c \backslash \{0\}} \widetilde{U}_{c,t,y} < 1$. One simplification in the formula above is that it neglects that the contingency redispatch cost (for either increasing or decreasing generator output in contingency situations) can be higher than implied by the generator cost function for normal operation.

3.8   Details on present value calculation

3.8.1 Discounting operational costs and investment costs

In the objective function for the dynamic (planning) optimization problem in Section 3.4, the pre-factors $f_y^{d,o}$ and f_y^d were introduced for the operational cost terms and the investment cost terms, respectively. These factors are included to account for the facts that i) each of the target years $y \in S_y$ represent multiple (here: 10) operational years, and ii) that, simply put, costs and benefit incurred far into the future are worth less value than costs and benefits closer to the present time. In the objective function, all costs incurred during the predefined analysis horizon are referred to a reference time (e.g., present time), and the formula calculates the present value of cost elements at different points in the future. In FlexPlan it will be assumed that the reference time is the year $y_0 = 2030$. The principle of present value calculation applied to the socio-economic analysis represented by the FlexPlan objective function is illustrated in Figure 3-2.
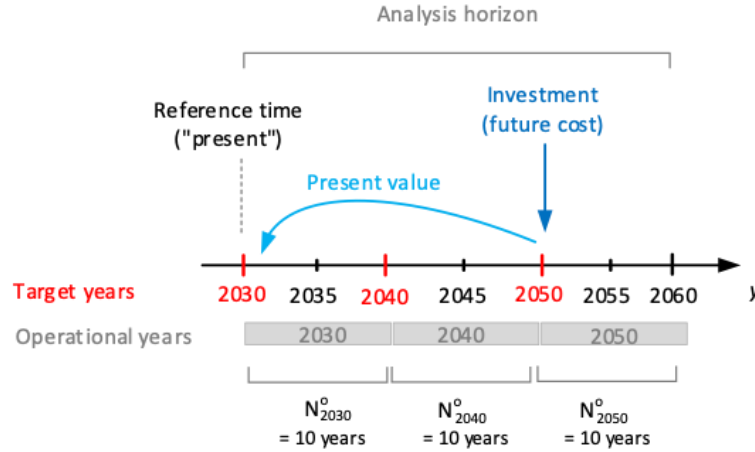
Figure 3-2 - Principle of present value applied to the socio-economic analysis represented by the FlexPlan objective function

The pre-factor for the investment cost terms is the discount factor, a simple form of which can be stated as

$$f_y^{\mathrm{d}} = \frac{1}{(1 + r_{\mathrm{d}})^{y - y_0}} \tag{17}$$

where r_{d} is the (real) discount rate. Using a real discount rate r_{d} means that the expected inflation is accounting for implicitly in the selection of the value of r_{d}. Inflation should therefore not be accounted for explicitly in addition. Recommended values for r_{d} for socio-economic cost-benefit analyses vary between countries and depend on government policy, regulation, and economic conditions. (For socio-economic cost-benefit analyses in Norway, for example, typical values of r_{d} can be around 4% to 5%. The European Agency for the Cooperation of Energy Regulators (ACER), the European Commission and ENTSO-E usually recommend to take $r_{\mathrm{d}} = 4\%$).

The pre-factor for the operational cost term needs to capture that each target year y represents N_y^{o} operational years $\{y, y + 1, \ldots, y + N_y^{\mathrm{o}} - 1\}$. It is assumed that the operational costs for each of these N_y^{o} operational years are identical but that the discount factors vary and decrease over time. The pre-factor $f_y^{\mathrm{d},0}$ for the operational cost terms thus becomes

$$f_y^{\mathrm{d},0} = \sum_{y'=y}^{y + N_y^{\mathrm{o}} - 1} f_y'^{\mathrm{d}} = \sum_{y'=y}^{y + N_y^{\mathrm{o}} - 1} \frac{1}{(1 + r_{\mathrm{d}})^{y' - y_0}}$$

3.8.2 Residual value of investments

In socio-economic analyses with a finite analysis horizon, it is opportune to account for the residual value at the end of the analysis horizon of the investments made within the analysis horizon. Otherwise, the impacts (i.e., operational benefits and costs) of investments towards the end of the analysis horizon are given disproportionately little weight compared to the investment costs, since some of these impacts are expected beyond the analysis horizon. One can say that residual values partly account for this inconsistency by effectively

reducing investment costs that are incurred relatively late in the analysis horizon. Residual values can be incorporated for instance by adding a term as follows to the objective function in Section 3.4:

$$\sum_{y \in S_y} \left\{ f_y^{d,o} \sum_{t \in S_t} \left[\sum_i C_{t,i} + \sum_{j \in S_j} \alpha_j C_{t,j} + \sum_{n \in S_n} K_{t,n} \right] + f_y^{d,i} \sum_{j \in S_j} \alpha_j I_j \right\} - f_{y_{end}}^d \sum_{j \in S_j} \alpha_j I_j^{res}$$

Here, I_j^{res} is the residual value of candidate (investment) j and $f_{y_{end}}^d$ is the discount factor defined previously, evaluated at the end of the analysis horizon. For the case that $S_y = \{2030, 2040, 2050\}$ and $N_y^o = 10$, $y_{end} = 2050 + 10 = 2060$. Furthermore, S_j is here used to denote the set of all investment candidates.

If one assumes linear depreciation, the time-dependent residual value $I_{j,y}^{res}$ can be calculated as

$$I_{j,y}^{res} = I_j \times \max\left\{ 1 - \frac{y - y_{inv}}{T^{life}}, 0 \right\}$$

where T^{life} is the economic lifetime of the investment (i.e., the asset). The residual value at the end of the planning horizon is defined as $I_j^{res} = I_{j,y_{end}}^{res}$. In the example in Figure 3-3, an investment in a transmission line is made at $y_{inv} = 2040$, and the transmission line is assumed to have an economic lifetime of $T^{life} = 30$ years. Since the end of the analysis horizon is $y_{end} = 2060$, $I_j^{res} = I_{j,y=2060}^{res} = I_j \times \frac{1}{3}$. This represents in a simplified manner that the transmission line has still approximately one third of its value "left" at the end of the analysis horizon explicitly considered in the socio-economic analysis. Impacts in terms of operational costs beyond the analysis horizon are still not being considered, since they can only be explicitly captured by extending the analysis horizon.
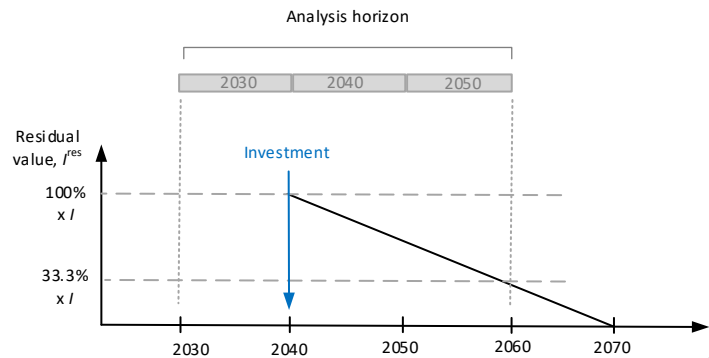
Figure 3-3 - Example calculation of residual value of investment

4 Network modelling

4.1 Power flow and equipment modelling for (sub)transmission networks

In the context of the FlexPlan approach, (sub)transmission networks refer to the <u>meshed operated part of the power network</u>, independent of the legal definition of transmission and distribution networks, which may refer to different voltage levels and operational rules in different countries.

The general AC/DC power flow model, based on [2, 3], is used for existing and newly built branches. The model has been extended with phase-shifting transformers (PSTs).

This section describes the power flow models used for transmission networks planning in the FlexPlan tool as well as the equipment models associated with HVDC links and phase shifting transformers (PSTs), which offer flexibility in terms of power flow control.

4.1.1 Nonlinear power flow and equipment modelling for (sub)transmission networks

Firstly, the mixed integer, nonlinear, nonconvex (MINLP) power flow model is introduced, from which, the linear mixed integer linear model (MILP) is derived, as large scale MINLP problems cannot be solved efficiently with current optimization solvers.

We first start with the modelling of AC nodes. Each AC node is characterized with a complex valued nodal voltage $U_k \angle \theta_k, k \in S_n^{ac}$ as shown in Figure 4-1 The operational limits of the AC nodal voltages are defined as follows:

$$U_{m,t,y}^{min} \le U_{m,t,y} \le U_{m,t,y}^{max} \qquad \forall m \in S_n^{ac}, \forall t \in S_t, \forall y \in S_y$$
$$\theta_{m,t,y}^{min} \le \theta_{m,t,y} \le \theta_{m,t,y}^{max} \qquad \forall m \in S_n^{ac}, \forall t \in S_t, \forall y \in S_y$$

where these operational limits need to hold for each time step of each planning horizon considered. An AC branch connects two different AC nodes k and m. As such, the AC branch model consists of AC lines and cables and power transformers, e.g., found in primary substations. A generic PI-model representation of AC branches is chosen, as depicted in Figure 4-1. g_l and b_l are the resistive and inductive series admittances, whereas bc_l is the shunt susceptance. For the representation of power transformers, the AC branch model is generalized with an ideal transformer with the voltage tap ratio τ_l. For an AC line or cable, $\tau_l = 1$ holds.
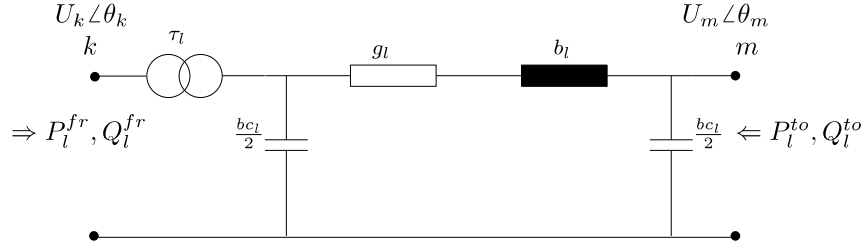
Figure 4-1 - AC branch model

The active and reactive power flow through an existing AC branch in the *from* direction $k \rightarrow m$ are defined as follows:

$$P_{l,t,y}^{fr} = g_l \left(\frac{U_{k,t,y}}{\tau_l}\right)^2 - g_l \frac{U_{k,t,y} U_{m,t,y}}{\tau_l} \cos(\theta_{k,t,y} - \theta_{m,t,y}) - b_l \frac{U_{k,t,y} U_{m,t,y}}{\tau_l} \sin(\theta_{k,t,y} - \theta_{m,t,y}) \quad \forall l \in S_l^{ac}, \forall t \in S_t, \forall y$$
$$\in S_y$$

$$Q_{l,t,y}^{fr} = -b_l \left(\frac{U_{k,t,y}}{\tau_l}\right)^2 + b_l \frac{U_{k,t,y} U_{m,t,y}}{\tau_l} \cos(\theta_{k,t,y} - \theta_{m,t,y}) - g_l \frac{U_{k,t,y} U_{m,t,y}}{\tau_l} \sin(\theta_{k,t,y} - \theta_{m,t,y}) \quad \forall l \in S_l^{ac}, \forall t$$
$$\in S_t, \forall y \in S_y$$

The active and reactive power flow through an existing AC branch in the *to* direction $m \rightarrow k$ are defined as follows:

$$P_{l,t,y}^{to} = g_l \left(\frac{U_{m,t,y}}{\tau_l}\right)^2 - g_l \frac{U_{k,t,y} U_{m,t,y}}{\tau_l} \cos(\theta_{m,t,y} - \theta_{k,t,y}) - b_l \frac{U_{k,t,y} U_{m,t,y}}{\tau_l} \sin(\theta_{m,t,y} - \theta_{k,t,y}) \quad \forall l \in S_l^{ac}, \forall t$$
$$\in S_t, \forall y \in S_y$$

$$Q_{l,t,y}^{to} = -b_l \left(\frac{U_{m,t,y}}{\tau_l}\right)^2 + b_l \frac{U_{k,t,y} U_{m,t,y}}{\tau_l} \cos(\theta_{m,t,y} - \theta_{k,t,y}) - g_l \frac{U_{k,t,y} U_{m,t,y}}{\tau_l} \sin(\theta_{m,t,y} - \theta_{k,t,y}) \quad \forall l \in S_l, \forall t$$
$$\in S_t, \forall y \in S_y$$

Candidate AC branches are defined between a pair of AC nodes. For candidate branches (lines, cables, primary substations), the power flow equations are extended by the binary decision variable $\alpha_{lc,t,y}$ such that the power flow is zero if $\alpha_{lc,t,y} = 0$ and the Kirchhoff equations are fulfilled if $\alpha_{lc,t,y} = 1$:

$$P_{lc,t,y}^{fr} = \left[g_{lc} \left(\frac{U_{k,t,y}}{\tau_{lc}}\right)^2 - g_{lc} \frac{U_{k,t,y} U_{m,t,y}}{\tau_{lc}} \cos(\theta_{k,t,y} - \theta_{m,t,y}) - b_{lc} \frac{U_{k,t,y} U_{m,t,y}}{\tau_{lc}} \sin(\theta_{k,t,y} - \theta_{m,t,y})\right] \alpha_{lc,t,y} \quad \forall lc$$
$$\in S_{lc}^{ac}, \forall t \in S_t, \forall y \in S_y$$

$$Q_{lc,t,y}^{fr} = \left[-b_{lc} \left(\frac{U_{k,t,y}}{\tau_{lc}}\right)^2 + b_{lc} \frac{U_{k,t,y} U_{m,t,y}}{\tau_{lc}} \cos(\theta_{k,t,y} - \theta_{m,t,y}) - g_{lc} \frac{U_{k,t,y} U_{m,t,y}}{\tau_{lc}} \sin(\theta_{k,t,y} - \theta_{m,t,y})\right] \alpha_{lc,t,y} \quad \forall lc$$
$$\in S_{lc}^{ac}, \forall t \in S_t, \forall y \in S_y$$

$$P_{lc,t,y}^{to} = \left[g_{lc} \left(\frac{U_{m,t,y}}{\tau_{lc}}\right)^2 - g_{lc} \frac{U_{k,t,y} U_{m,t,y}}{\tau_{lc}} \cos(\theta_{m,t,y} - \theta_{k,t,y}) - b_{lc} \frac{U_{k,t,y} U_{m,t,y}}{\tau_{lc}} \sin(\theta_{m,t,y} - \theta_{k,t,y})\right] \alpha_{lc,t,y} \quad \forall lc$$
$$\in S_{lc}^{ac}, \forall t \in S_t, \forall y \in S_y$$

$$Q_{lc,t,y}^{to} = \left[-b_{lc} \left(\frac{U_{m,t,y}}{\tau_{lc}}\right)^2 + b_{lc} \frac{U_{k,t,y} U_{m,t,y}}{\tau_{lc}} \cos(\theta_{m,t,y} - \theta_{k,t,y}) - g_{lc} \frac{U_{k,t,y} U_{m,t,y}}{\tau_{lc}} \sin(\theta_{m,t,y} - \theta_{k,t,y})\right] \alpha_{lc,t,y} \quad \forall lc$$
$$\in S_{lc}^{ac}, \forall t \in S_t, \forall y \in S_y$$

The power flow limits for existing and candidate AC branches are defined as follows:

$$\left(P_{l,t,y}^{fr}\right)^2 + \left(Q_{l,t,y}^{fr}\right)^2 \leq \left(S_l^{rated}\right)^2 \qquad \forall l \in S_l^{ac}, \forall t \in S_t, \forall y \in S_y$$

$$\left(P_{l,t,y}^{to}\right)^2 + \left(Q_{l,t,y}^{to}\right)^2 \leq \left(S_l^{rated}\right)^2 \qquad \forall l \in S_l^{ac}, \forall t \in S_t, \forall y \in S_y$$

$$\left(P_{lc,t,y}^{fr}\right)^2 + \left(Q_{lc,t,y}^{fr}\right)^2 \leq \left(S_l^{rated}\right)^2 \alpha_{lc,t,y} \qquad \forall lc \in S_{lc}^{ac}, \forall t \in S_t, \forall y \in S_y$$

$$\left(P_{lc,t,y}^{to}\right)^2 + \left(Q_{lc,t,y}^{to}\right)^2 \leq \left(S_l^{rated}\right)^2 \alpha_{lc,t,y} \qquad \forall lc \in S_{lc}^{ac}, \forall t \in S_t, \forall y \in S_y$$

Additionally, a maximum allowable voltage angle deviation along AC existing and candidate branches is defined, which is used as a proxy dynamic stability:

$$\theta_{k,t,y} - \theta_{m,t,y} \leq \Delta\theta_l^{max} \qquad \forall l \in S_l^{ac}, \forall t \in S_t, \forall y \in S_y$$

$$\theta_{m,t,y} - \theta_{k,t,y} \leq \Delta\theta_l^{max} \qquad \forall l \in S_l^{ac}, \forall t \in S_t, \forall y \in S_y$$

$$\theta_{k,t,y} - \theta_{m,t,y} \leq \Delta\theta_{lc}^{max} \qquad \forall lc \in S_{lc}^{ac}, \forall t \in S_t, \forall y \in S_y$$

$$\theta_{m,t,y} - \theta_{k,t,y} \leq \Delta\theta_{lc}^{max} \qquad \forall lc \in S_{lc}^{ac}, \forall t \in S_t, \forall y \in S_y$$

A phase shifting transformer (PST) can be represented by a series impedance connected in series to an existing transmission line as shown in Figure 4-2.
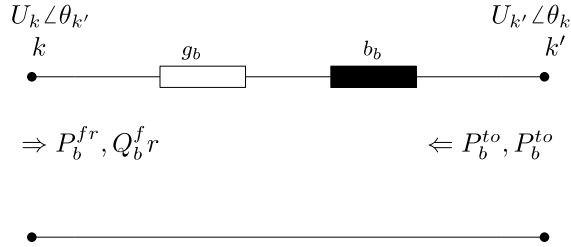
Figure 4-2 - PST Model

The PST is located between the nodes k and k' connected in series with a line $k' \rightarrow m$. A PST introduces a voltage angle shift $\varphi_{b,t,y}$ along the PST impedance, such that the power flow through an existing PST becomes,

$$P_{b,t,y}^{fr} = g_b\left(U_{k,t,y}\right)^2 - g_b U_{k,t,y} U_{k',t,y} \cos\left(\theta_{k,t,y} - \theta_{k',t,y} + \varphi_{b,t,y}\right) - b_b U_{k,t,y} U_{k',t,y} \sin\left(\theta_{k,t,y} - \theta_{k',t,y} + \varphi_{b,t,y}\right) \quad \forall b \in S_b, \forall t \in S_t, \forall y \in S_y$$

$$Q_{b,t,y}^{fr} = -b_b\left(U_{k,t,y}\right)^2 + b_b U_{k,t,y} U_{k',t,y} \cos\left(\theta_{k,t,y} - \theta_{k',t,y} + \varphi_{b,t,y}\right)$$
$$- g_b U_{k,t,y} U_{k',t,y} \sin\left(\theta_{k,t,y} - \theta_{k',t,y} + \varphi_{b,t,y}\right) \quad \forall b \in S_b, \forall t \in S_t, \forall y \in S_y$$

$$P_{b,t,y}^{to} = g_b\left(U_{k',t,y}\right)^2 - g_b U_{k,t,y} U_{k',t,y} \cos\left(\theta_{k',t,y} - \theta_{k,t,y} - \varphi_{b,t,y}\right)$$
$$- b_b U_{k,t,y} U_{k',t,y} \sin\left(\theta_{k',t,y} - \theta_{k,t,y} - \varphi_{b,t,y}\right) \quad \forall b \in S_b, \forall t \in S_t, \forall y \in S_y$$

$$Q_{b,t,y}^{to} = -b_b\left(U_{k',t,y}\right)^2 + b_b U_{k,t,y} U_{k',t,y} \cos\left(\theta_{k',t,y} - \theta_{k,t,y} - \varphi_{b,t,y}\right) - g_b U_{k,t,y} U_{k',t,y} \sin\left(\theta_{k',t,y} - \theta_{k,t,y} - \varphi_{b,t,y}\right) \quad \forall b \in S_b, \forall t \in S_t, \forall y \in S_y.$$

The phase angle shift is bounded with the maximum and minimum phase angle shift of the PST,

$$\varphi_b^{min} \leq \varphi_{b,t,y} \leq \varphi_b^{max} \quad \forall b \in S_b, \forall t \in S_t, \forall y \in S_y$$ also, the apparent power across the PST is bounded as follows,

$$\left(P_{b,t,y}^{fr}\right)^2 + \left(Q_{b,t,y}^{fr}\right)^2 \leq \left(S_b^{rated}\right)^2 \qquad \forall b \in S_b, \forall t \in S_t, \forall y \in S_y$$

$$\left(P_{b,t,y}^{to}\right)^2 + \left(Q_{b,t,y}^{to}\right)^2 \leq \left(S_b^{rated}\right)^2 \qquad \forall b \in S_b, \forall t \in S_t, \forall y \in S_y$$

For a candidate PST, only putting active and reactive power flows to zero if the PST is not built (as for candidate AC branches), is not an option, because the PST is connected in series with existing lines. As such the power flow equations are multiplied with the investment decision variable $\alpha_{lc,t,y}$ on both sides.

$$\alpha_{bc,y} P^{fr}_{bc,t,y} = \left[g_{bc}\left(U_{k,t,y}\right)^2 - g_{bc}U_{k,t,y}U_{k',t,y}\cos\left(\theta_{k,t,y} - \theta_{k',t,y} + \varphi_{bc,t,y}\right) \right.$$
$$\left. - b_{bc}U_{k,t,y}U_{k',t,y}\sin\left(\theta_{k,t,y} - \theta_{k',t,y} + \varphi_{bc,t,y}\right) \right] \alpha_{bc,y} \quad \forall bc \in S_{bc}, \forall t \in S_t, \forall y \in S_y$$

$$\alpha_{bc,y} Q^{fr}_{bc,t,y} = \left[-b_{bc}\left(U_{k,t,y}\right)^2 + b_{bc}U_{k,t,y}U_{k',t,y}\cos\left(\theta_{k,t,y} - \theta_{k',t,y} + \varphi_{b,t,y}\right) \right.$$
$$\left. - g_{bc}U_{k,t,y}U_{k',t,y}\sin\left(\theta_{k,t,y} - \theta_{k',t,y} + \varphi_{bc,t,y}\right) \right] \alpha_{bc,y} \quad \forall bc \in S_{bc}, \forall t \in S_t, \forall y \in S_y$$

$$\alpha_{bc,y} P^{to}_{bc,t,y} = \left[g_{bc}\left(U_{k',t,y}\right)^2 - g_{bc}U_{k,t,y}U_{k',t,y}\cos\left(\theta_{k',t,y} - \theta_{k,t,y} - \varphi_{bc,t,y}\right) \right.$$
$$\left. - b_{bc}U_{k,t,y}U_{k',t,y}\sin\left(\theta_{k',t,y} - \theta_{k,t,y} - \varphi_{bc,t,y}\right) \right] \alpha_{bc,y} \quad \forall bc \in S_{bc}, \forall t \in S_t, \forall y \in S_y$$

$$\alpha_{bc,y} Q^{to}_{bc,t,y} = \left[-b_{bc}\left(U_{k',t,y}\right)^2 + b_{bc}U_{k,t,y}U_{k',t,y}\cos\left(\theta_{k',t,y} - \theta_{k,t,y} - \varphi_{bc,t,y}\right) \right.$$
$$\left. - g_{bc}U_{k,t,y}U_{k',t,y}\sin\left(\theta_{k',t,y} - \theta_{k,t,y} - \varphi_{bc,t,y}\right) \right] \alpha_{bc,y} \quad \forall bc \in S_{bc}, \forall t \in S_t, \forall y \in S_y$$

The equations above result in the expression $0 = 0$ if the binary decision variable $\alpha_{bc,y} = 0$ and in the power flow equations as expected for $\alpha_{bc,y} = 1$. In order to avoid slack in power injections, the voltage magnitudes and angles of the nodes k and k' are set to equal values using

$$\Delta\theta^{\min \alpha_{bc,y}}_{bc} \leq \theta_{k,t,y} - \theta_{k',t,y} \leq \Delta\theta^{\max \alpha_{bc,y}}_{bc} \quad \forall bc \in S_{bc}, \forall t \in S_t, \forall y \in S_y$$

and

$$\Delta U^{\min \alpha_{bc,y}}_{bc} \leq U_{k,t,y} - U_{k',t,y} \leq \Delta U^{\max \alpha_{bc,y}}_{b} \quad \forall bc \in S_{bc}, \forall t \in S_t, \forall y \in S_y,$$

where $\Delta\theta^{min}_{bc}$d $\Delta\theta^{max}_{bc}$und sufficiently large voltage angle differences, and ΔU^{min}_{bc}d ΔU^{max}_{bc}rk a sufficiently large voltage magnitude range for the difference in nodal voltages of nodes k and k'. If the binary decision variable is zero, the voltages of nodes k and k' are equal in magnitude and angle and are determined by the power flows in case $\alpha_{bc,y} = 1$. In a similar way, the power flows of candidate PSTs need to be bound using,

$$\Delta P^{min}_{bc} \leq P^{fr}_{bc,t,y} + P^{to}_{bc,t,y} \leq \Delta P^{max}_{bc}\alpha_{bc,y} \quad \forall bc \in S_{bc}, \forall t \in S_t, \forall y \in S_y$$
$$\Delta Q^{min}_{bc}\alpha_{bc,y} \leq Q^{fr}_{bc,t,y} + Q^{to}_{bc,t,y} \leq \Delta Q^{max}_{b}\alpha_{bc,y} \quad \forall bc \in S_{bc}, \forall t \in S_t, \forall y \in S_y$$
$$\left(P^{fr}_{bc,t,y}\right)^2 + \left(Q^{fr}_{bc,t,y}\right)^2 \leq \left(S^{rated}_{bc}\right)^2 \quad \forall b \in S_b, \forall t \in S_t, \forall y \in S_y$$
$$\left(P^{to}_{bc,t,y}\right)^2 + \left(Q^{to}_{bc,t,y}\right)^2 \leq \left(S^{rated}_{bc}\right)^2 \quad \forall b \in S_b, \forall t \in S_t, \forall y \in S_y$$
$$\alpha_{bc,y}\varphi^{min}_{bc} \leq \varphi_{bc,t,y} \leq \varphi^{max}_{bc}\alpha_{bc,y} \quad \forall bc \in S_{bc}, \forall t \in S_t, \forall y \in S_y$$

where ΔP^{min}_{bc}, ΔP^{max}_{bc}, ΔQ^{min}_{bc} and ΔQ^{max}_{b} mark a sufficiently large ranges of active and reactive power flows through the PST.

According to Kirchhoff's current law, all active and reactive power injections in a node need to balance to zero,

$$\sum_{gm \in T^{gen}} P_{g,t,y} + \sum_{z(c)me \in T^{acdc}} P^{ac}_{z(c),t,y} + \sum_{jm \in T^{st}} P^{inj}_{j,t,y} - \sum_{jm \in T^{st}} P^{abs}_{j,t,y} - \sum_{lmk \in T^{ac}} P^{fr}_{l(c),t,y} - \sum_{bmk \in T^{b}} P^{fr}_{b(c),t,y}$$
$$- \sum_{um \in T^{load,flex}} P^{flex}_{u,t,y} - \sum_{um \in T^{load}} P_{u,t,y} = 0 \quad \forall m \in S_m, t \in S_t, y \in S_y$$

$$\sum_{gm \in T^{gen}} Q_{g,t,y} + \sum_{zme \in T^{acdc}} Q_{z,t,y}^{ac} + \sum_{jm \in T^{st}} Q_{j,t,y}^{inj} - \sum_{jm \in T^{st}} Q_{j,t,y}^{abs} - \sum_{lmk \in T^{ac}} Q_{l,t,y}^{fr} - \sum_{bmk \in T^{b}} Q_{b,t,y}^{fr}$$

$$- \sum_{um \in T^{load,flex}} Q_{u,t,y}^{flex} - \sum_{um \in T^{load}} Q_{u,t,y} = 0 \; \forall m \in S_m, t \in S_t, y \in S_y \,,$$

where,

- $P_{g,t,y}$ and $Q_{g,t,y}$ are the active and reactive power injections of generators connected to node m,
- $P_{z(c),t,y}^{ac}$ and $Q_{z(c),t,y}^{ac}$ are the AC side active and reactive power injections of HVDC converters connected to node m (both existing and candidate),
- $P_{l(c),t,y}^{fr}$ and $Q_{l(c),t,y}^{fr}$ are the active power flows leaving from node m (both existing and candidate),
- $P_{b(c),t,y}^{fr}$ and $Q_{b(c),t,y}^{fr}$ are the active flows and reactive power flows of PSTs leaving from node m (both existing and candidate),
- $P_{u,t,y}^{flex}$ and $Q_{u,t,y}^{flex}$ are the active and reactive power demands of flexible loads connected to node m,
- $P_{u,t,y}$ and $Q_{u,t,y}$ are the active and reactive power demands of non - flexible loads connected to node m,
- $P_{j,t,y}^{inj}$ and $P_{j,t,y}^{abs}$ are the active power injection and absorption of storage elements connected to node m,
- $Q_{j,t,y}^{inj}$ and $Q_{j,t,y}^{abs}$ are the reactive power injection and absorption of storage elements connected to node m.

The detailed model for the flexible demand is provided in Section 5.1 and the detailed model of storage elements is provided in Section 5.2.

Both dispatchable and renewable generators are subject to operational limits,

$$P_{g,t,y}^{min} \leq P_{g,t,y} \leq P_{g,t,y}^{max} \quad \forall g \in S_g, t \in S_t, y \in S_y$$

$$Q_{g,t,y}^{min} \leq Q_{g,t,y} \leq Q_{g,t,y}^{max} \; \forall g \in S_g, t \in S_t, y \in S_y,$$

where the active and reactive power limits of the renewable generators are defined by the climatic conditions at time instance t.

An HVDC converter links an AC grid node $m \in S_n^{ac}$ with a DC grid node $e \in S_n^{dc}$. For the HVDC converter model it is assumed that voltage source converter (VSC) technology is used. Using VSC technology, active and reactive power injections and absorptions can be controlled independently, in contrary to line commuted converters (LCC), where the absorbed reactive power is dependent on the active power injection. The HVDC converter is represented as a pair of AC and DC grid active power injections, which are linked via the converter losses

$$P_{z,t,y}^{ac} + P_{z,t,y}^{dc} = P_{z,t,y}^{loss} \quad \forall z \in S_z, t \in S_t, y \in S_y.$$

The converter losses are function of the AC side converter current $I_{z,t,y}^{ac}$:

$$P_{z,t,y}^{loss} = L_z^a + L_z^b I_{z,t,y} + L_z^c I_{z,t,y}^2 \quad \forall z \in S_z, t \in S_t, y \in S_y,$$

where the converter current and power are linked with the following equations:

$$3I_{z,t,y}^2 U_{m,t,y}^2 = \left(P_{z,t,y}^{ac}\right)^2 + \left(Q_{z,t,y}^{ac}\right)^2 \quad \forall z \in S_z, \; t \in S_t, y \in S_y,$$

$U_{m,t,y}$ being the AC nodal voltage of node m to which the converter z is connected to. The converter is subject to the following operational limits in terms of current and power:

$$\left(P_{z,t,y}^{ac}\right)^2 + \left(Q_{z,t,y}^{ac}\right)^2 \leq \left(S_z^{rated,ac}\right)^2 \quad \forall z \in S_z, t \in S_t, y \in S_y$$

$$-P_z^{rated,ac} \le P_{z,t,y}^{ac} \le P_z^{rated,ac} \quad \forall z \in S_z, t \in S_t, y \in S_y$$

$$-Q_z^{rated,ac} \le Q_{z,t,y}^{ac} \le Q_z^{rated,ac} \quad \forall z \in S_z, t \in S_t, y \in S_y$$

$$-P_z^{rated,dc} \le P_{z,t,y}^{dc} \le P_z^{rated,dc} \quad \forall z \in S_z, t \in S_t, y \in S_y.$$

The power flow equations are valid also for the candidate HVDC converters:

$$P_{zc,t,y}^{ac} + P_{zc,t,y}^{dc} = P_{zc,t,y}^{loss} \quad \forall zc \in S_{zc}, t \in S_t, y \in S_y.$$

$$P_{zc,t,y}^{loss} = L_{zc}^a + L_{zc}^b I_{zc,t,y} + L_{zc}^c I_{zc,t,y}^2 \quad \forall zc \in S_{zc}, t \in S_t, y \in S_y,$$

$$3 I_{zc,t,y}^2 U_{m,t,y}^2 = \left(P_{zc,t,y}^{ac}\right)^2 + \left(Q_{zc,t,y}^{ac}\right)^2 \quad \forall zc \in S_{zc}, t \in S_t, y \in S_y,$$

For candidate HVDC converters $zc \in S_{zc}$, the power injections can be bounded using the binary decision variable $\alpha_{zc,y}$ such that both AC and DC side power injections become zero if $\alpha_{zc,y} = 0$ and the power injections are set by the optimizer if $\alpha_{zc,y} = 1$:

$$\left(P_{zc,t,y}^{ac}\right)^2 + \left(Q_{zc,t,y}^{ac}\right)^2 \le \left(S_{zc}^{rated,ac}\right)^2 \alpha_{zc,y} \quad \forall zc \in S_{zc}, t \in S_t, y \in S_y$$

$$-\alpha_{zc,y} P_{zc}^{rated,ac} \le P_{zc,t,y}^{ac} \le P_{zc}^{rated,ac} \alpha_{zc,y} \quad \forall zc \in S_{zc}, t \in S_t, y \in S_y$$

$$-\alpha_{zc,y} Q_{zc}^{rated,ac} \le Q_{zc,t,y}^{ac} \le Q_{zc}^{rated,ac} \alpha_{zc,y} \quad \forall zc \in S_{zc}, t \in S_t, y \in S_y$$

$$-\alpha_{zc,y} P_{zc}^{rated,dc} \le P_{zc,t,y}^{dc} \le P_{zc}^{rated,dc} \alpha_{zc,y} \quad \forall zc \in S_{zc}, t \in S_t, y \in S_y.$$

Each DC branch $d \in S_l^{dc}$ connects a DC grid node $e \in S_n^{dc}$ with a DC grid node $f \in S_n^{dc}$. The power flow over a DC branch is defined according to Ohm's law in both the *from* $(e \rightarrow f)$ and the *to* $(f \rightarrow e)$ directions, respectively:

$$P_{d,t,y}^{fr} = p_d g_d U_{e,t,y}\left(U_{e,t,y} - U_{f,t,y}\right) \quad \forall d \in S_l^{dc}, t \in S_t, y \in S_y$$

$$P_{d,t,y}^{to} = p_d g_d U_{f,t,y}\left(U_{f,t,y} - U_{e,t,y}\right) \quad \forall d \in S_l^{dc}, t \in S_t, y \in S_y,$$

where $p_d \in \{1,2\}$ is the number of DC poles[2] and $g_d = \frac{1}{r_d}$ is the series admittance of the DC branch. All DC branches are subject to power flow limits,

$$-P_{d,t,y}^{rated} \le P_{d,t,y}^{fr} \le P_{d,t,y}^{rated} \quad \forall d \in S_l^{dc}, t \in S_t, y \in S_y$$

$$-P_{d,t,y}^{rated} \le P_{d,t,y}^{to} \le P_{d,t,y}^{rated} \quad \forall d \in S_l^{dc}, t \in S_t, y \in S_y.$$

The power flow equations need to be valid for all candidate DC branches:

$$P_{dc,t,y}^{fr} = p_{dc} g_{dc} U_{e,t,y}\left(U_{e,t,y} - U_{f,t,y}\right) \quad \forall dc \in S_{lc}^{dc}, t \in S_t, y \in S_y$$

$$P_{dc,t,y}^{to} = p_{dc} g_{dc} U_{f,t,y}\left(U_{f,t,y} - U_{e,t,y}\right) \quad \forall dc \in S_{lc}^{dc}, t \in S_t, y \in S_y.$$

The power flows through candidate branches can be bound using the binary decision variable $\alpha_{dc,y}$ which forces the power flows to zero if $\alpha_{dc,y} = 0$ and keeps them between operational bounds if $\alpha_{dc,y} = 1$:

$$-\alpha_{dc,y} P_{dc,t,y}^{rated} \le P_{dc,t,y}^{fr} \le P_{dc,t,y}^{rated} \alpha_{dc,y} \quad \forall dc \in S_{lc}^{dc}, t \in S_t, y \in S_y$$

$$-\alpha_{dc,y} P_{dc,t,y}^{rated} \le P_{dc,t,y}^{to} \le P_{dc,t,y}^{rated} \alpha_{dc,y} \quad \forall dc \in S_{lc}^{dc}, t \in S_t, y \in S_y.$$

Finally, all injections from DC converters and DC branches can be linked in DC nodal power balance equation similar to the AC nodes:

---

[2] $p_d = 1$ for monopoles and $p_d = 2$ for symmetrical monopoles or bipolar HVDC connections.

$$\sum_{zcme \,\in\, T^{acdc}} P_{zc,t,y}^{dc} + \sum_{dcef \,\in\, T^{dc}} P_{dc,t,y}^{fr} = 0 \ \forall e \in S_n^{dc}, t \in S_t, y \in S_y.$$

4.1.2 Linear power flow and equipment modelling for (sub)transmission networks

In the following paragraphs, the general nonlinear, nonconvex model is approximated with a linear model as the current MINLP solvers are not capable of efficiently solving such large-scale problems. The underlying assumptions for the linearization are that the magnitudes of the nodal AC and DC grid voltages can be assumed to be equal such that $U_{m,t,y} = U_{k,t,y} = U_{ac} \ \forall m, k \in S_n^{ac}$ and $U_{e,t,y} = U_{f,t,y} = U_{dc} \ \forall e, f \in S_n^{dc}$. Another assumption is that the AC grid voltage angle differences are so small such that $\sin(\theta_{m,t,y} - \theta_{k,t,y}) \approx \theta_{m,t,y} - \theta_{k,t,y}$ and $\cos(\theta_{m,t,y} - \theta_{k,t,y}) \approx 1$ and that the resistive part of the series admittance $g_l = 0$ is neglected, resulting for AC branches in:

$$P_{l,t,y}^{fr} = b_l \frac{U_{ac}^2}{\tau_l} \left(\theta_{m,t,y} - \theta_{k,t,y}\right) \ \forall l \in S_l^{ac}, \forall t \in S_t, \forall y \in S_y$$

$$P_{l,t,y}^{to} = b_l \frac{U_{ac}^2}{\tau_l} \left(\theta_{k,t,y} - \theta_{m,t,y}\right) = -P_{l,t,y}^{fr} \ \forall l \in S_l^{ac}, \forall t \in S_t, \forall y \in S_y$$

$$Q_{l,t,y}^{fr} = Q_{l,t,y}^{to} = 0 \ \forall l \in S_l^{ac}, \forall t \in S_t, \forall y \in S_y,$$

which means that all reactive power related terms disappear in the used model. The power flow equations for candidate branches are defined as

$$P_{lc,t,y}^{fr} = b_{lc} \frac{U_{ac}^2}{\tau_{lc}} \left(\theta'_{m,t,y} - \theta'_{k,t,y}\right) \ \forall lc \in S_{lc}^{ac}, \forall t \in S_t, \forall y \in S_y$$

$$P_{lc,t,y}^{to} = b_{lc} \frac{U_{ac}^2}{\tau_{lc}} \left(\theta'_{k,t,y} - \theta'_{m,t,y}\right) \ \forall lc \in S_{lc}^{ac}, \forall t \in S_t, \forall y \in S_y$$

where $\theta'_{m,t,y}$ and $\theta'_{k,t,y}$ are auxiliary variables. Branch flows are bound by the maximum voltage angle differences and the thermal limits of the AC branches (lines, cables, primary substations) for existing and candidate branches:

$$\theta_{k,t,y} - \theta_{m,t,y} \le \Delta\theta_l^{\max} \quad \forall l \in S_l^{ac}, \forall t \in S_t, \forall y \in S_y$$

$$\theta_{m,t,y} - \theta_{k,t,y} \le \Delta\theta_l^{\max} \quad \forall l \in S_l^{ac}, \forall t \in S_t, \forall y \in S_y$$

$$\theta'_{k,t,y} - \theta'_{m,t,y} \le \Delta\theta_{lc}^{\max} \quad \forall lc \in S_{lc}^{ac}, \forall t \in S_t, \forall y \in S_y$$

$$\theta'_{m,t,y} - \theta'_{k,t,y} \le \Delta\theta_{lc}^{\max} \quad \forall lc \in S_{lc}^{ac}, \forall t \in S_t, \forall y \in S_y$$

$$-P_l^{rated} \le P_{l,t,y}^{fr} \le P_l^{rated} \quad \forall l \in S_l^{ac}, \forall t \in S_t, \forall y \in S_y$$

$$-P_l^{rated} \le P_{l,t,y}^{to} \le P_l^{rated} \quad \forall l \in S_l^{ac}, \forall t \in S_t, \forall y \in S_y$$

$$-\alpha_{lc,y} P_{lc}^{rated} \le P_{lc,t,y}^{fr} \le P_{lc}^{rated} \alpha_{lc,y} \quad \forall lc \in S_{lc}^{ac}, \forall t \in S_t, \forall y \in S_y$$

$$-\alpha_{lc,y} P_{lc}^{rated} \le P_{lc,t,y}^{to} \le P_{lc}^{rated} \alpha_{lc,y} \quad \forall lc \in S_{lc}^{ac}, \forall t \in S_t, \forall y \in S_y.$$

The following constraints need to be added to link the auxiliary voltage angle variables to the nodal voltage angles,

$$-(1 - \alpha_{lc,y}) \cdot M \le \theta'_{m,t,y} - \theta_{m,t,y} \le (1 - \alpha_{lc,y}) \cdot M \quad \forall lc \in S_{lc}^{ac}, \forall t \in S_t, \forall y \in S_y$$

$$-(1 - \alpha_{lc,y}) \cdot M \le \theta'_{k,t,y} - \theta_{k,t,y} \le (1 - \alpha_{lc,y}) \cdot M \quad \forall lc \in S_{lc}^{ac}, \forall t \in S_t, \forall y \in S_y$$

such that the auxiliary voltage angles correspond to the nodal voltage angles in case candidate branches are built ($\alpha_{lc,y} = 1$) and they are freely chosen by the solver if the candidate branch is not built ($\alpha_{lc,y} = 0$), while

still ensuring that there is no power flow on the not built candidate branch. M is a sufficiently large angle and can be fixed to π, such that the auxiliary voltage angles remain in a range of $\pm\pi$.

The nodal voltage bounds are only enforced for the voltage angles as the voltage magnitude becomes a parameter:

$$\theta_{m,t,y}^{min} \leq \theta_{m,t,y} \leq \theta_{m,t,y}^{max} \quad \forall m \in S_n^{ac}, \forall t \in S_t, \forall y \in S_y.$$

Similar to AC branches, also the linearized equations for existing PSTs can be written using $U_{k,t,y} =$

$$U_{k',t,y} = U_{ac} \; \forall k', k \in S_n^{ac},$$
$$P_{b,t,y}^{fr} = b_b U_{ac}^2 \left(\theta_{k,t,y} - \theta_{k',t,y} + \varphi_{b,t,y} \right) \forall b \in S_b, \forall t \in S_t, \forall y \in S_y$$
$$P_{b,t,y}^{to} = b_b U_{ac}^2 \left(\theta_{k',t,y} - \theta_{k,t,y} - \varphi_{b,t,y} \right) \forall b \in S_b, \forall t \in S_t, \forall y \in S_y$$
$$-P_b^{rated} \leq P_{b,t,y}^{fr} \leq P_b^{rated} \qquad \forall l \in S_l^{ac}, \forall t \in S_t, \forall y \in S_y$$
$$-P_b^{rated} \leq P_{b,t,y}^{to} \leq P_b^{rated} \qquad \forall l \in S_l^{ac}, \forall t \in S_t, \forall y \in S_y$$
$$\varphi_b^{min} \leq \varphi_{b,t,y} \leq \varphi_b^{max} \quad \forall b \in S_b, \forall t \in S_t, \forall y \in S_y$$

or candidate PSTs two new variables $\tilde{P}_{bc,t,y}^{fr}$, $\tilde{P}_{bc,t,y}^{to}$ need to be defined, in order to avoid a non-linear formulation. The power flow through the PST in the *from* and *to* directions are separated in two terms as shown in the equation below. The first term is active if the investment decision is taken ($\alpha_{bc,y} = 1$). In this case $\tilde{P}_{bc,t,y}^{fr}$ and $\tilde{P}_{bc,t,y}^{to}$ become zero, using the variable bound inequality. In the case $\alpha_{bc,y} = 0$, the variables $\tilde{P}_{bc,t,y}^{fr}$ and $\tilde{P}_{bc,t,y}^{to}$ can vary between the active power limits of the PST. As in the case of the nonlinear formulation the nodal voltage angles of nodes k and k' are made equal, to match the power flow of the in series connected branch.

$$P_{bc,t,y}^{fr} = b_{bc} U_{ac}^2 \left(\theta_{k,t,y} - \theta_{k',t,y} + \varphi_{bc,t,y} \right) + \tilde{P}_{bc,t,y}^{fr} \; \forall bc \in S_{bc}, \forall t \in S_t, \forall y \in S_y$$
$$P_{bc,t,y}^{to} = b_{bc} U_{ac}^2 \left(\theta_{k',t,y} - \theta_{k,t,y} - \varphi_{bc,t,y} \right) + \tilde{P}_{bc,t,y}^{to} \; \forall bc \in S_{bc}, \forall t \in S_t, \forall y \in S_y$$
$$-P_{bc}^{rated} \leq P_{bc,t,y}^{fr} \leq P_{bc}^{rated} \qquad \forall bc \in S_{bc}, \forall t \in S_t, \forall y \in S_y$$
$$-P_{bc}^{rated} \leq P_{bc,t,y}^{to} \leq P_{bc}^{rated} \qquad \forall bc \in S_{bc}, \forall t \in S_t, \forall y \in S_y$$
$$\alpha_{bc,y} \varphi_{bc}^{min} \leq \varphi_{bc,t,y} \leq \varphi_{bc}^{max} \alpha_{bc,y} \quad \forall bc \in S_{bc}, \forall t \in S_t, \forall y \in S_y$$
$$\theta_{bc}^{min} \alpha_{bc,y} \leq \theta_{k,t,y} - \theta_{k',t,y} \leq \theta_{bc}^{max} \alpha_{bc,y} \quad \forall bc \in S\_bc, \forall t \in S\_t, \forall y \in S\_y$$
$$-(1 - \alpha_{bc,y}) P_b^{rated} \leq \tilde{P}_{bc,t,y}^{fr} \leq (1 - \alpha_{bc,y}) P_b^{rated} \; \forall bc \in S_{bc}, \forall t \in S_t, \forall y \in S_y$$
$$-(1 - \alpha_{bc,y}) P_b^{rated} \leq \tilde{P}_{bc,t,y}^{to} \leq (1 - \alpha_{bc,y}) P_b^{rated} \; \forall bc \in S_{bc}, \forall t \in S_t, \forall y \in S_y$$

In order to avoid slackness in the active power injections, following constraint is needed for the auxiliary active power variables $\tilde{P}_{bc,t,y}^{fr}$ and $\tilde{P}_{bc,t,y}^{to}$:

$$\tilde{P}_{bc,t,y}^{fr} + \tilde{P}_{bc,t,y}^{to} = 0 \; \forall bc \in S_{bc}, \forall t \in S_t, \forall y \in S_y$$

The AC nodal power balance is only written in active power:

$$\sum_{gm \in T^{gen}} P_{g,t,y} + \sum_{zme \in T^{acdc}} P_{z(c),t,y}^{ac} + \sum_{jm \in T^{st}} P_{j,t,y}^{inj} - \sum_{jm \in T^{st}} P_{j,t,y}^{abs} - \sum_{lmk \in T^{ac}} P_{l(c),t,y}^{fr} - \sum_{bmk \in T^{ac}} P_{b(c),t,y}^{fr}$$
$$- \sum_{um \in T^{load}} P_{u,t,y}^{flex} - \sum_{um \in T^{load}} P_{u,t,y} = 0 \; \forall m \in S_m, t \in S_t, y \in S_y$$

The lossy HVDC converter model is approximated by its linear components and the converter current is substituted by the AC side active power as the voltage magnitude becomes a parameter. Note that for our linear model neglecting reactive power both VSC and LCC type of converters can be modelled in the same way.

$$P_{z,t,y}^{ac} + P_{z,t,y}^{dc} = P_{z,t,y}^{loss} \quad \forall z \in S_z, t \in S_t, y \in S_y.$$

$$P_{z,t,y}^{loss} = L_z^a + L_z^b P_{z,t,y}^{ac} \quad \forall z \in S_z, t \in S_t, y \in S_y,$$

$$P_{zc,t,y}^{ac} + P_{zc,t,y}^{dc} = P_{zc,t,y}^{loss} \quad \forall zc \in S_{zc}, t \in S_t, y \in S_y,$$

$$P_{zc,t,y}^{loss} = L_{zc}^a + L_{zc}^b P_{zc,t,y}^{ac} \quad \forall zc \in S_{zc}, t \in S_t, y \in S_y.$$

Note that the converter losses need to be positive at all times, $P_{z,t,y}^{loss} = L_z^a + L_z^b P_{z,t,y}^{ac} \geq 0$, $P_{zc,t,y}^{loss} = L_{zc}^a + L_{zc}^b P_{zc,t,y}^{ac} \geq 0$.

The AC and DC side active power injections are bound as follows, for the existing and candidate DC converters:

$$-P_z^{rated,ac} \leq P_{z,t,y}^{ac} \leq P_z^{rated,ac} \quad \forall z \in S_z, t \in S_t, y \in S_y$$

$$-P_z^{rated,dc} \leq P_{z,t,y}^{dc} \leq P_z^{rated,dc} \quad \forall z \in S_z, t \in S_t, y \in S_y$$

$$-\alpha_{zc,y} P_{zc}^{rated,ac} \leq P_{zc,t,y}^{ac} \leq P_{zc}^{rated,ac} \alpha_{zc,y} \quad \forall zc \in S_{zc}, t \in S_t, y \in S_y$$

$$-\alpha_{zc,y} P_{zc}^{rated,dc} \leq P_{zc,t,y}^{dc} \leq P_{zc}^{rated,dc} \alpha_{zc,y} \quad \forall zc \in S_{zc}, t \in S_t, y \in S_y.$$

As $U_{e,t,y} = U_{f,t,y} = U_{dc} \; \forall e, f \in S_n^{dc}$ is assumed, the power flow model reduces to a network flow model for existing and candidate DC branches where,

$$P_{d,t,y}^{fr} + P_{d,t,y}^{to} = 0 \; \forall d \in S_l^{dc}, t \in S_t, y \in S_y$$

$$P_{dc,t,y}^{fr} + P_{dc,t,y}^{to} = 0 \; \forall d \in S_{lc}^{dc}, t \in S_t, y \in S_y.$$

The power flows of the DC branches are bound by the thermal rating for existing and candidate branches:

$$-P_{d,t,y}^{rated} \leq P_{d,t,y}^{fr} \leq P_{d,t,y}^{rated} \quad \forall d \in S_l^{dc}, t \in S_t, y \in S_y$$

$$-P_{d,t,y}^{rated} \leq P_{d,t,y}^{to} \leq P_{d,t,y}^{rated} \quad \forall d \in S_l^{dc}, t \in S_t, y \in S_y$$

$$-\alpha_{dc,y} P_{dc,t,y}^{rated} \leq P_{dc,t,y}^{fr} \leq P_{dc,t,y}^{rated} \alpha_{dc,y} \quad \forall dc \in S_{lc}^{dc}, t \in S_t, y \in S_y$$

$$-\alpha_{dc,y} P_{dc,t,y}^{rated} \leq P_{dc,t,y}^{to} \leq P_{dc,t,y}^{rated} \alpha_{dc,y} \quad \forall dc \in S_{lc}^{dc}, t \in S_t, y \in S_y.$$

Finally, the DC grid nodal balance equation remains unchanged:

$$\sum_{zcme \in T^{acdc}} P_{zc,t,y}^{dc} + \sum_{dcef \in T^{dc}} P_{dc,t,y}^{fr} = 0 \; \forall e \in S_n^{dc}, t \in S_t, y \in S_y.$$

4.1.3 Investment decision constraints

To ensure that selected candidate lines, converters and PSTs cannot be deactivated after the investment decision is taken for a certain planning year, following set of constraints is required:

$$\alpha_{lc,y-1} \leq \alpha_{lc,y} \quad \forall lc \in S_{lc}^{ac}, \forall y \in S_y: y > 1$$

$$\alpha_{dc,y-1} \leq \alpha_{dc,y} \quad \forall dc \in S_{lc}^{dc}, \forall y \in S_y: y > 1$$

$$\alpha_{bc,y-1} \leq \alpha_{bc,y} \quad \forall lc \in S_{bc}, \forall y \in S_y: y > 1$$

$$\alpha_{zc,y-1} \leq \alpha_{zc,y} \quad \forall zc \in S_{zc}, \forall y \in S_y: y > 1$$

The described model allows to define parallel candidate AC and DC lines, HVDC converters and PSTs[3], which all might have different power ratings, impedances and costs. As such, the optimizer chooses the best possible combination between the defined candidates. If the number of selected investments in a certain corridor, corresponding to a pair of nodes in the described model, needs to be limited, e.g., due to spatial constraints, the

---

[3] Investment decisions constraints for demand flexibility and storage are described in Chapter 0.

following constraint can be used, where $n_{x,y}^{max}$ maximum number of investments that can be connected to a specific node in each year of the planning horizon.

$$\sum_{zcme \, \in \, T^{acdc}} \alpha_{zc,y} \leq n_{zc,y}^{max} \quad \forall m \in S_m^{ac}, \forall e \in S_m^{dc}, y \in S_y$$

$$\sum_{dcef \, \in \, T^{dc}} \alpha_{dc,y} \leq n_{dc,y}^{max} \quad \forall e,f \in S_m^{dc}, y \in S_y$$

$$\sum_{lcmk \, \in \, T^{ac}} \alpha_{lc,y} \leq n_{lc,y}^{max} \quad \forall m,k \in S_m^{ac}, y \in S_y$$

$$\sum_{bcmk \, \in \, T^{b}} \alpha_{bc,y} \leq n_{bc,y}^{max} \quad \forall m,k \in S_m^{ac}, y \in S_y$$

4.2   Modelling of security constraints

Broadly speaking, one can in general account for reliability of supply and reliability criteria using the following approaches (or combinations thereof):

1. Accounting for reliability in the objective function (through energy not supplied)
2. Accounting for reliability in the constraints
    a. through security (contingency) constraints in a security-constrained OPF model
    b. through pre-determined reliability/security margins in the constraints for the base case power flow model (through heuristics or security analyses in a pre-processing phase)

In FlexPlan we propose to primarily account for reliability of supply through the objective function of the optimization model, and the modelling details were described in section 0. In addition to these objective function terms, a set of power flow equations and other technical restrictions (thermal ratings of branches etc.) as described in the previous section must be included in the constraints of the model for each contingency $c \in S_c$.

One way to specify and mathematically represent the contingencies is to use state matrices for the components subject to outage. For instance, if one considers AC branch contingencies, a state outage matrix L for all N-1 AC branch contingencies can be written as

$$L = \begin{bmatrix} 1 & 0 & \dots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & & 0 \end{bmatrix}.$$

An element $L_{l,c}$ equals zero if branch l is in an outage state in contingency c. The dimensionality of L is in general $N_l \times N_c$ where $N_c = |S_c|$ is the number of contingencies and $N_l = |S_l^{ac} \cup S_{lc}^{ac}|$ is the number of AC branches (both existing branches and candidate branches can be included for the contingency set). The first column of L corresponds to the non-contingency state $c = 0$ in which no components are in the outage state. For the example of all N-1 AC branch contingencies above, the dimensionality of L is $N_l \times N_c = N_l \times (N_l + 1)$. Including a set of power flow constraints for each contingency is necessary for the model to be able to

determine the value of the slack variables $\Delta P_{u,c,t,y}$, which depends on the power flow in the contingency-state network model. Including these contingency constraints amount to including a copy of all constraints in section 4.1 for each contingency $c \in S_c$ but with the connectivity tuples $T^{.}$ appropriately modified to reflect the component outage states for the contingency. For instance, if AC branch contingencies are considered, the AC grid topology tuple T^{ac} acquires a contingency subscript and becomes T_c^{ac}. The nodal power balance equation in section 4.1 is extended by a sum over variables $\Delta P_{u,c,t,y}$ for all demand elements $u \in S_u$:

$$\sum_{gm\ \in T^{gen}} P_{g,c,t,y} + \sum_{z(c)me\ \in T^{acdc}} P_{z(c),c,t,y}^{ac} + \sum_{jm \in T^{st}} P_{j,c,t,y}^{inj} - \sum_{jm \in T^{st}} P_{j,c,t,y}^{abs} - \sum_{lmk \in T_c^{ac}} P_{l(c),c,t,y}^{fr} - \sum_{bmk \in T^b} P_{b(c),c,t,y}^{fr}$$

$$- \sum_{um \in T^{load,flex}} P_{u,c,t,y}^{flex} - \sum_{um \in T^{load}} P_{u,t,y} - \sum_{um \in T^{load}} \Delta P_{u,c,t,y} = 0 \; \forall m \in S_m, t \in S_t, y \in S_y, c \in S_c.$$

Note that compared to section 4.1, an additional contingency subscript c is added for the power variables, since the power flow for each time step will in general be different for each contingency.

When accounting for contingencies in the AC branch flow constraints in section 4.1, they take the form

$$-L_{l,c}P_l^{rated} \le P_{l,c,t,y}^{fr} \le L_{l,c}P_l^{rated} \qquad \forall l \in S_l^{ac}, \forall t \in S_t, \forall y \in S_y, \forall c \in S_c$$
$$-L_{l,c}P_l^{rated} \le P_{l,c,t,y}^{to} \le L_{l,c}P_l^{rated} \qquad \forall l \in S_l^{ac}, \forall t \in S_t, \forall y \in S_y, \forall c \in S_c$$

for existing branches and the form

$$-L_{lc,c}\alpha_{lc,y}P_{lc}^{rated} \le P_{lc,c,t,y}^{fr} \le P_{lc}^{rated}\alpha_{lc,y}L_{lc,c} \qquad \forall lc \in S_{lc}^{ac}, \forall t \in S_t, \forall y \in S_y$$
$$-L_{lc,c}\alpha_{lc,y}P_{lc}^{rated} \le P_{lc,c,t,y}^{to} \le P_{lc}^{rated}\alpha_{lc,y}L_{lc,c} \qquad \forall lc \in S_{lc}^{ac}, \forall t \in S_t, \forall y \in S_y.$$

for candidate branches. $L_{l,c}$ and $L_{lc,c}$ are the lth (lc^{th}) line and c^{th} columns of the state outage matrix L for both existing and candidate lines, respectively. Analogous contingency constraints apply for AC branch voltage angle differences $\theta_{k,c,t,y} - \theta_{m,c,t,y}$ and for the other equipment described in the network and power flow modelling (section 4.1). It should be noted that contingency analysis can be computationally very heavy. One way of speeding it up when using a DC power flow formulation is to calculate line outage distribution factors. The line outage distribution factor $LODF_{l,c}$ gives the overload on line l, due to contingency c. These factors can quickly be calculated using the inverse matrix modification lemma as presented in [4]. The advantage of this approach is that the susceptance matrix only has to be inverted for the non-contingency case ($c = 0$).

To represent branch power transfer ratings that are higher during contingencies (i.e., emergency operation) than during normal operation, one can replace the thermal ratings $P_{l(c)}^{rated}$ in these constraints by emergency ratings $P_{l(c)}^{rated,em}$.

Note that these contingency constraints have similarities with the security constraints of a security-constrained OPF (cf. approach 2a above). However, since the slack variables $\Delta P_{u,c,t,y}$ are included in the problem in our case, the model does not strictly speaking *secure* the system against the contingencies, and it is thus more appropriate to refer to the constraints as contingency constraints than as security constraints.

It is also possible to follow approach 2b to representing security constraints (cf. the overview above) without introducing additional modelling features to the model. In that case, the limit parameters for technical constraints can be set to a more conservative value by including a security margin (or reliability margin). For instance, one can use a lower value for the thermal rating $S_{l(c)}^{rated}$ than the actual thermal rating of the branch.

This representation can be used also when no contingencies are included in the case ($S_c = \{0\}$). However, the approach requires that security margins are determined in a pre-processing step.

Apart from the contingency constraints described above, the only constraints in the model that are particular to the representation of security constraints are the boundaries on the slack variable $\Delta P_{u,c,t,y}$ that were implied in Section 3.1, namely

$$0 \leq \Delta P_{u,c,t,y} \leq P_{u,t,y}^{ref}.$$

Note that these constraints on the slack variable have the same mathematical structure as the constraints on the demand flexibility presented in Section 5.1 below. Expressed in these terms, one has $P_{u,t,y}^{max} = P_{u,t,y}^{ref}$d $P_{u,t,y}^{min} = 0$r contingency states (i.e., contingency index $c > 0$). We assume that that there are no separate decision variables for demand flexibility (i.e., voluntary load reduction) in contingency states (i.e. these are only relevant degrees of freedom for the non-contingency state $c = 0$).

In order to account for generator contingencies in an implicit way, following constraint is used,

$$\sum_{g \in S_g} P_g^{max} - P_g \geq \max(P_g^{max}) \; \forall t \in S_t, \forall y \in S_y$$

where the power rating of the largest generator in the system. This way, it is ensured that there is always enough reserve generation available to compensate generator contingencies.

4.3   Distribution system modelling

4.3.1 On the definition of distribution system

The FlexPlan planning tool could be applied with network models including transmission systems, distribution systems, or both. In the context of the network model used in the planning tool, a distribution system is understood as a part of the power system that is radial or radially operated. The network model for meshed operated networks has been presented in Section 4.1. The ambition of FlexPlan is to be able to include the highest voltage level of the radially operated system in the integrated transmission and distribution system planning problem.

The planning tool is to be applied to power systems of different countries in the regional case studies. Which parts of the power system that is regarded as distribution system will depend on the country. To use Italy as an example, distribution systems typically are at voltage levels around 20 kV and below. This example is illustrated in Figure 4-3 with a radially operated 20 kV distribution system and a meshed transmission system represented by voltage levels 380 kV and (a sub-transmission network at) 150 kV. In other countries, parts of what is regarded as distribution systems are at higher voltage levels and meshed. To use Norway as an example, networks at around 60 kV and some parts of the 132 kV sub-transmission network are regarded as part of distribution systems. However, 60 kV is the highest voltage level that is usually operated radially.
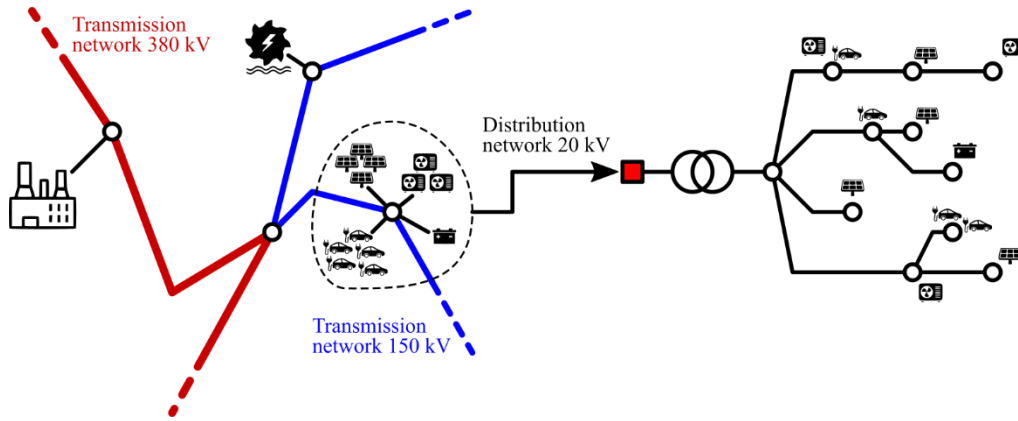
Figure 4-3 - Example of a network model including the transmission network and a distribution network.

4.3.2 Power flow modelling of radial distribution system

Distribution systems refer to the radial (or radially operated) parts of the grid operated at medium/low voltages. The distribution system is composed by the same physical elements as transmission networks. Lines and transformers can be modelled in the same way (see **Figure 4-1**), however the characteristics of radial grids equipped with medium/low voltage devices make the assumptions taken for transmission systems not entirely
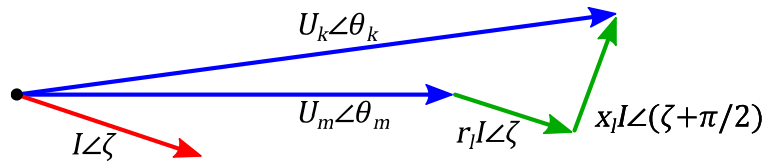
Figure 4-4 Example of phasor diagram for an AC branch

applicable on distribution systems.

Having considered the AC branch model reported in Figure 4-4**Figure *4-1***, the phasor diagram of the electrical quantities can be represented as depicted in Figure 4-5**Errore. L'origine riferimento non è stata trovata.**, where $I\angle\zeta$ is assumed to be the current flowing from node k to node m through the branch resistance r_l and reactance x_l.

In the literature [5], several methods for carrying out optimal power flow calculation. The most common practices for distribution networks are based on a non-linear (and rigorous) representation of the network physical behaviour. Some others [6] adopted linear formulations which, thanks to opportune strategies, can be reconducted to the exact physics of the electrical network. Having considered the normal operation of a distribution grid, there are some reasonable approximations that can be performed in order to formulate the power flow problem linearly.

One of the main assumptions adopted for distribution networks consists of considering negligible the voltage phase-angle difference among two neighbouring buses. In this condition, the magnitude of the line voltage drop can be approximated as:

$$U_k - U_m \cong \sqrt{3}r_l I \cos(\varphi) + \sqrt{3}x_l I \sin(\varphi)$$

where $\varphi = \theta_{km} - \zeta$ and θ_{km} is the phase-angle of the median voltage of the branch (see Figure 4-5). Thanks to this definition, the active power flowing through the branch is calculated as $P_l = \sqrt{3}U_{km}I\cos(\varphi)$, while the reactive power as $Q_l = \sqrt{3}U_{km}I\sin(\varphi)$.
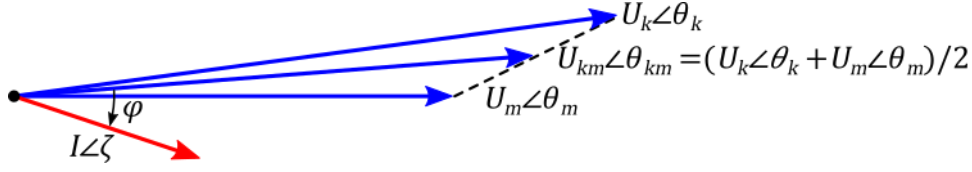
Figure 4-5 - Definition of the median voltage of the branch and power factor angle

φ

For small phase-angle displacements, it can be recognized that $U_{km} \cong (U_k + U_m)/2$, therefore:

$$U_k^2 - U_m^2 \cong 2r_l P_l + 2x_l Q_l$$

The obtained approximation can be compared with the exact formulation of the squared-voltage difference, which can be demonstrated to be equal to:

$$U_k^2 - U_m^2 = 2r_l P_l + 2x_l Q_l + (r_l^2 + x_l^2)\frac{P_l^2 + Q_l^2}{U_m^2}$$

It can be noticed that the adopted approximation consists of the linearization of the exact formula around the working point $P_l = 0$ and $Q_l = 0$, which is making the model neglecting power losses occurring in lines. According to that, the proposed formulation can be considered to be acceptable if the following assumptions are validated:

- low voltage phase-angle deviation between consecutive buses;
- low power losses in network lines. lines if compared to power transit.

Real systems are always affected by (technical) power losses and this introduces a systematic error within the model. The energy efficiency of distribution networks normally decreases with the voltage level, spreading from ~98% for medium voltage grids to ~92% for low voltage grids [7]. Limiting the planning problem to medium voltage, the model can be considered affected by

- a systematic error of about 2% within the power balance;
- a similar error magnitude within the branch voltage drop.

Finally, in order to consider the presence of power transformers, the equation modelling the generic AC branch can be rewritten as:

$$\frac{U_{k,t,y}^2}{\tau_l^2} - U_{m,t,y}^2 \cong 2r_l P_{l,t,y}^{fr} + 2x_l Q_{l,t,y}^{fr} \quad \forall l \in S_l^{ac}, \forall t \in S_t, \forall y \in S_y$$

where τ_l is the voltage transformation ratio. In case the voltage transformation ratio is controllable, it can be noticed that the selected model results to be non-linear for a variable τ_l. However, for a typical distribution network, controllable transformation ratios are experienced only for on load tap changers, which are directly connected to the swing bus (point of common coupling between transmission and distribution). In this case, since the proposed transmission network model is not managing the voltage variable, the generic voltage $U_{k,t,y}^2$

can be assumed to be equal to the nominal voltage $(U_k^{nom})^2$, and removed from the optimization variables so that the consideration of τ_l (which becomes $\tau_{l,t,y}$) does not jeopardize the linearity of the model:

$$\frac{1}{\tau_{l,t,y}^2}(U_k^{nom})^2 - U_{m,t,y}^2 \cong 2r_l P_{l,t,y}^{fr} + 2x_l Q_{l,t,y}^{fr} \quad \forall l \in S_l^{ac}, \forall t \in S_t, \forall y \in S_y$$

Having noticed that reactive power is required for the processing of the branch model, the power balance needs to take it into account. . For this reason, in addition to the same active power balance adopted for transmission system (see Section 4.1.1)

$$\sum_{gm \, \in T^{gen}} P_{g,t,y} + \sum_{jm \in T^{st}} P_{j,t,y}^{inj} - \sum_{jm \in T^{st}} P_{j,t,y}^{abs} - \sum_{lmk \in T^{ac}} P_{l(c),t,y}^{fr} - \sum_{um \in T^{load}} P_{u,t,y}^{flex} - \sum_{um \in T^{load}} P_{u,t,y} = 0,$$

$$\forall m \in S_m, t \in S_t, y \in S_y,$$

also reactive power needs to be balanced

$$\sum_{gm \, \in T^{gen}} Q_{g,t,y} + \sum_{jm \in T^{st}} Q_{j,t,y}^{inj} - \sum_{jm \in T^{st}} Q_{j,t,y}^{abs} - \sum_{lmk \in T^{ac}} Q_{l,t,y}^{fr} - \sum_{um \in T^{load,flex}} Q_{u,t,y}^{flex} + \sum_{um \in T^{load}} Q_{u,t,y} = 0,$$

$$\forall m \in S_m, t \in S_t, y \in S_y.$$

With respect to transmission network power balance, the contribution of DC lines and phase-shifter is excluded since they are not considered possible candidates for radial distribution networks.

Concerning the other optimization constraints, in the first instance, only voltage and loading limitations of the network are considered for distribution system management and planning. These constraints can be formulated as:

$$U_{m,t,y}^{min} \leq U_{m,t,y} \leq U_{m,t,y}^{max} \quad \forall m \in S_n^{ac}, \forall t \in S_t, \forall y \in S_y,$$

$$\tau_{l(c)}^{min} \leq \tau_l \leq \tau_{l(c)}^{max} \quad \forall l \in S_l^{ac}, \forall t \in S_t, \forall y \in S_y,$$

$$\sqrt{\left(P_{l,t,y}^{fr}\right)^2 + \left(Q_{l,t,y}^{fr}\right)^2} \leq S_{l,t,y}^{rated} \quad \forall l \in S_l^{ac}, \forall t \in S_t, \forall y \in S_y.$$

In order to keep the model linear, the following reformulation are adopted:

$$\left(U_{m,t,y}^{min}\right)^2 \leq U_{m,t,y}^2 \leq \left(U_{m,t,y}^{max}\right)^2 \quad \forall m \in S_n^{ac}, \forall t \in S_t, \forall y \in S_y,$$

$$\left(\frac{1}{\tau_l^{max}}\right)^2 \leq \left(\frac{1}{\tau_l}\right)^2 \leq \left(\frac{1}{\tau_l^{min}}\right)^2 \quad \forall l \in S_l^{ac}, \forall t \in S_t, \forall y \in S_y,$$

$$\left.\begin{aligned}
-\cos\left(\frac{\pi}{8}\right) S_l^{rated} \leq P_{l,t,y}^{fr} \leq \cos\left(\frac{\pi}{8}\right) S_l^{rated} \\
-\cos\left(\frac{\pi}{8}\right) S_l^{rated} \leq Q_{l,t,y}^{fr} \leq \cos\left(\frac{\pi}{8}\right) S_l^{rated} \\
-\sqrt{2}\cos\left(\frac{\pi}{8}\right) S_l^{rated} \leq P_{l,t,y}^{fr} + Q_{l,t,y}^{fr} \leq \sqrt{2}\cos\left(\frac{\pi}{8}\right) S_l^{rated} \\
-\sqrt{2}\cos\left(\frac{\pi}{8}\right) S_l^{rated} \leq P_{l,t,y}^{fr} - Q_{l,t,y}^{fr} \leq \sqrt{2}\cos\left(\frac{\pi}{8}\right) S_l^{rated}
\end{aligned}\right\} \quad \forall l \in S_l^{ac}, \forall t \in S_t, \forall y \in S_y$$

where the overloading constraint is expressed by using an octagonal approximation of the circular power flow capability of branches. In order to guarantee that the actual power transit limits are not overtaken the octagon approximation inscribed within the actual capability is adopted.

Similarly to the (sub)transmission network model, the line candidates (having $\tau_l = 1$) can be modelled by using auxiliary variables that, for the case of distribution network, consist of the auxiliary voltage magnitude $\left(U_{m,t,y}^2\right)'$:

$$U_{k,t,y}^2 - \left(U_{m,t,y}^2\right)' = 2r_{lc}P_{lc,t,y}^{fr} + 2x_{lc}Q_{lc,t,y}^{fr} \quad \forall lc \in S_{lc}^{ac}, \forall t \in S_t, \forall y \in S_y$$

or, for voltage regulating transformers connected to the swing bus:

$$\frac{1}{\tau_{lc,t,y}^2}(U_k^{nom})^2 - \left(U_{m,t,y}^2\right)' = 2r_{lc}P_{lc,t,y}^{fr} + 2x_{lc}Q_{lc,t,y}^{fr} \quad \forall lc \in S_{lc}^{ac}, \forall t \in S_t, \forall y \in S_y$$

where, having defined a sufficiently large voltage magnitude M, the related constraints can be expressed as:

$$-\left(1 - \alpha_{lc,y}\right) \cdot M \le \left(U_{m,t,y}^2\right)' - U_{m,t,y}^2 \le \left(1 - \alpha_{lc,y}\right) \cdot M \quad \forall lc \in S_{lc}^{ac}, \forall t \in S_t, \forall y \in S_y,$$

$$\left.\begin{array}{l} -\alpha_{lc,y}\cos\left(\frac{\pi}{8}\right)S_{lc}^{rated} \le P_{lc,t,y}^{fr} \le \alpha_{lc,y}\cos\left(\frac{\pi}{8}\right)S_{lc}^{rated} \\[2mm] -\alpha_{lc,y}\cos\left(\frac{\pi}{8}\right)S_{lc}^{rated} \le Q_{lc,t,y}^{fr} \le \alpha_{lc,y}\cos\left(\frac{\pi}{8}\right)S_{lc}^{rated} \\[2mm] -\alpha_{lc,y}\sqrt{2}\cos\left(\frac{\pi}{8}\right)S_{lc}^{rated} \le P_{lc,t,y}^{fr} + Q_{lc,t,y}^{fr} \le \alpha_{lc,y}\sqrt{2}\cos\left(\frac{\pi}{8}\right)S_{lc}^{rated} \\[2mm] -\alpha_{lc,y}\sqrt{2}\cos\left(\frac{\pi}{8}\right)S_{lc}^{rated} \le P_{lc,t,y}^{fr} - Q_{lc,t,y}^{fr} \le \alpha_{lc,y}\sqrt{2}\cos\left(\frac{\pi}{8}\right)S_{lc}^{rated} \end{array}\right\} \quad \forall lc \in S_{lc}^{ac}, \forall t \in S_t, \forall y \in S_y.$$

4.3.3 Candidate management at distribution level

The proposed formulation is applicable only to radial grids, conditions for which the number of equations balances the number of variables (voltage magnitudes and active/reactive power flows). For this reason, candidate branches and related decision variables need to be formulated in order to avoid the possible creation of meshes.

When a generic candidate line lc is aimed at replacing an existing line l, the model needs to consider their mutual exclusivity. Taking as reference the equations reported in the previous section, lc is included within the model as usual:

$$U_{k,t,y}^2 - \left(U_{m,t,y}^2\right)' = 2r_{lc}P_{lc,t,y}^{fr} + 2x_{lc}Q_{lc,t,y}^{fr} \quad \forall t \in S_t, \forall y \in S_y,$$

$$-\left(1 - \alpha_{lc,y}\right) \cdot M \le \left(U_{m,t,y}^2\right)' - U_{m,t,y}^2 \le \left(1 - \alpha_{lc,y}\right) \cdot M \quad \forall t \in S_t, \forall y \in S_y,$$

$$\left.\begin{array}{l} -\alpha_{lc,y}\cos\left(\frac{\pi}{8}\right)S_{lc}^{rated} \le P_{lc,t,y}^{fr} \le \alpha_{lc,y}\cos\left(\frac{\pi}{8}\right)S_{lc}^{rated} \\[2mm] -\alpha_{lc,y}\cos\left(\frac{\pi}{8}\right)S_{lc}^{rated} \le Q_{lc,t,y}^{fr} \le \alpha_{lc,y}\cos\left(\frac{\pi}{8}\right)S_{lc}^{rated} \\[2mm] -\alpha_{lc,y}\sqrt{2}\cos\left(\frac{\pi}{8}\right)S_{lc}^{rated} \le P_{lc,t,y}^{fr} + Q_{lc,t,y}^{fr} \le \alpha_{lc,y}\sqrt{2}\cos\left(\frac{\pi}{8}\right)S_{lc}^{rated} \\[2mm] -\alpha_{lc,y}\sqrt{2}\cos\left(\frac{\pi}{8}\right)S_{lc}^{rated} \le P_{lc,t,y}^{fr} - Q_{lc,t,y}^{fr} \le \alpha_{lc,y}\sqrt{2}\cos\left(\frac{\pi}{8}\right)S_{lc}^{rated} \end{array}\right\} \quad \forall t \in S_t, \forall y \in S_y.$$

In case the candidate lc is selected ($\alpha_{lc,y} = 1$), the existing line l needs to be excluded. By using the same decision variable $\alpha_{l,y}$, the potentially replaceable line l model can be written as follows:

$$U_{k,t,y}^2 - \left(U_{m,t,y}^2\right)'' = 2r_l P_{l,t,y}^{fr} + 2x_l Q_{l,t,y}^{fr} \quad \forall t \in S_t, \forall y \in S_y,$$

$$-\alpha_{lc,y} \cdot M \le \left(U_{m,t,y}^2\right)'' - U_{m,t,y}^2 \le \alpha_{lc,y} \cdot M \quad \forall t \in S_t, \forall y \in S_y,$$

$$\left.\begin{array}{r} -\left(1-\alpha_{lc,y}\right)\cos\left(\frac{\pi}{8}\right)S_l^{rated} \leq P_{l,t,y}^{fr} \leq \left(1-\alpha_{lc,y}\right)\cos\left(\frac{\pi}{8}\right)S_l^{rated} \\ -\left(1-\alpha_{lc,y}\right)\cos\left(\frac{\pi}{8}\right)S_l^{rated} \leq Q_{l,t,y}^{fr} \leq \left(1-\alpha_{lc,y}\right)\cos\left(\frac{\pi}{8}\right)S_l^{rated} \\ -\left(1-\alpha_{lc,y}\right)\sqrt{2}\cos\left(\frac{\pi}{8}\right)S_l^{rated} \leq P_{l,t,y}^{fr} + Q_{l,t,y}^{fr} \leq \left(1-\alpha_{lc,y}\right)\sqrt{2}\cos\left(\frac{\pi}{8}\right)S_l^{rated} \\ -\left(1-\alpha_{lc,y}\right)\sqrt{2}\cos\left(\frac{\pi}{8}\right)S_l^{rated} \leq P_{l,t,y}^{fr} - Q_{l,t,y}^{fr} \leq \left(1-\alpha_{lc,y}\right)\sqrt{2}\cos\left(\frac{\pi}{8}\right)S_l^{rated} \end{array}\right\} \quad \forall t \in S_t, \forall y \in S_y.$$

The same model can be adopted for the substitution of transformers. In this case, the voltage drop equations of the candidate transformers lc and the existing one l the can be respectively encoded as:

$$\frac{1}{\tau_{lc,t,y}^2}(U_k^{nom})^2 - \left(U_{m,t,y}^2\right)' = 2r_{lc}P_{lc,t,y}^{fr} + 2x_{lc}Q_{lc,t,y}^{fr} \quad \forall t \in S_t, \forall y \in S_y,$$

$$\frac{1}{\tau_{l,t,y}^2}(U_k^{nom})^2 - \left(U_{m,t,y}^2\right)'' = 2r_l P_{l,t,y}^{fr} + 2x_l Q_{l,t,y}^{fr} \quad \forall t \in S_t, \forall y \in S_y$$

The substitution of a line preserves the radiality of the network and it does not represent a modelling issue. On the contrary, when a candidate line lp is aimed at reinforcing the existing l, the creation of a physical mesh is unavoidable (two lines in parallel are defining a loop). In order to preserve the radiality condition required by the model, the two lines in parallel can be coded as a single fictitious branch lc having opportunely selected impedance (r_{lc},x_{lc}) and power limit S_{lc}^{rated}:

$$r_{lc} = \frac{r_l\left(r_{lp}^2 + x_{lp}^2\right) + r_{lp}(r_l^2 + x_l^2)}{\left(r_l + r_{lp}\right)^2 + \left(x_l + x_{lp}\right)^2},$$

$$x_{lc} = \frac{x_l\left(r_{lp}^2 + x_{lp}^2\right) + x_{lp}(r_l^2 + x_l^2)}{\left(r_l + r_{lp}\right)^2 + \left(x_l + x_{lp}\right)^2},$$

$$S_{lc}^{rated} = \frac{1}{\sqrt{r_{lc}^2 + x_{lc}^2}}\min\left(S_l^{rated}\sqrt{r_l^2 + x_l^2}, S_{lc}^{rated}\sqrt{r_{lp}^2 + x_{lp}^2}\right).$$

Thanks to this definition the number of variables of the model is not increased and its consistency is preserved. In fact, the fictitious line lc can be managed as a line aimed at substituting the existing one l, by simply using the model described above. The same approach can be adopted for voltage regulating transformers, for which operation in parallel is theoretically possible and sometimes adopted. However, this solution is leading to several drawbacks (re-circulating currents, higher power losses, higher short-circuit currents, etc.) and it is rarely considered as a planning option.

5 Generic flexibility modelling

5.1 Load modelling

The model intends to describe flexibility options in the most generic way within the planning tool. The modelling approach and the data model have been chosen to be as generic as possible such that they are independent of the type of the flexibility option and the used technology. The model parameters themselves reflect the type and technology.

This model can be used for these flexibility options:

- Electric Vehicles (EV)
- Industrial Demand Response
- Residential Demand Response
- Thermal Loads
- Hydrogen production as industrial load

Management of distributed energy sources is part of the generator modelling and is excluded from the general flexibility model. Load flexibility is referred to those resources that are able to change their hourly consumption profile in order to meet the needs of the System. If $P_{u,t,y}^{ref}$ is the expected hourly consumption, it can be both reduced up to $P_{u,t,y}^{min}$ d increased to $P_{u,t,y}^{max}$ or any load typology, power factor angle is assumed to be a constant, making reactive power varying proportionally with the active power:

$$Q_{u,t,y}^{flex} = P_{u,t,y}^{flex} \tan(\varphi_{u,t,y}^{flex}) \quad \forall u \in S_u, \forall t \in S_t, \forall y \in S_y$$

Two kinds of load flexibility can be considered by the planning tool: load reduction and load shifting. In addition, if these two mechanisms are not sufficient to meet generation, load curtailment can be considered as well.

Note that a flexibility resource is in principle able to participate to both the flexibility mechanism. Fixed costs for enabling flexibility for a load include initial investment costs and, depending on the technology considered, carbon footprint cost. Those costs were described in detail in Section 3.5.5.

5.1.1 Load decrease

In the first case, the flexibility resource simply reduces its consumption $P_{u,t,y}^{ref}$ in hour t of reference year y by an amount $\Delta P_{u,t,y}^{nce}$. When the flexibility is enabled by an investment, this load reduction must be positive and cannot be higher than a bound $\Delta_{u,t,y}^{nce,max}$:

$$0 \leq \Delta P_{u,t,y}^{nce} \leq \alpha_{u,y} \Delta_{u,t,y}^{nce,max} \tag{18}$$

To do so it will receive a remuneration proportional to the reduction of the consumption, with a compensation $C_{u,t,y}^{nce}$ [€/MWh] so that the corresponding cost for the system is given by:

$$C_{u,t,y}^{nce} \Delta P_{u,t,y}^{nce} \tag{19}$$

Moreover, the total amount of energy not consumed over a planning year can be constrained using,

$$0 \leq \sum_{t \in S_t} \Delta t \cdot \Delta P_{u,t,y}^{nce} \leq \alpha_{u,y} E_{u,y}^{nc,max}, \qquad \forall u \in S_u, y \in S_y.$$

5.1.2 Load shifting

In the second case, the flexibility resource changes its load profile shifting part of its consumption $P_{u,t,y}^{ref}$, i.e. increasing by an amount $\Delta P_{u,t,y}^{ds,up}$ or reducing by an amount $\Delta P_{u,t,y}^{ds,dn}$ from hour t_1 to another hour t_2 (it can be as well $t_1 < t_2$ and $t_2 < t_1$) but maintaining the total consumption within a given period τ:

$$\sum_{t \in \tau} \Delta P_{u,t,y}^{ds,up} = \sum_{t \in \tau} \Delta P_{u,t,y}^{ds,dn} \tag{20}$$

Also, the activation of the shifting flexibility is bounded and forced to be zero if no investment is performed,

$$0 \leq \Delta P_{u,t,y}^{ds,up} \leq \alpha_{u,y} \Delta P_{u,t,y}^{ds,up,max}$$

$$0 \leq \Delta P_{u,t,y}^{ds,dn} \leq \alpha_{u,y} \Delta P_{u,t,y}^{ds,dn,max}$$

Here a remuneration proportional to load shifted is considered, with a compensation $C_{u,t,y}^{ds}$ [€/MWh], so that the cost for the system in each hour t is given by

$$C_{u,t,y}^{ds} \Delta P_{u,t,y}^{ds,dn} \tag{21}$$

As demand shifting can only be performed for a short period of time, e.g., a number of hours, the upward and downward flexibility is limited for a number of hours τ_u^{rec} based on the activation of the flexibility as:

$$0 \leq \Delta P_{u,t,y}^{ds,up} \leq \Delta_{u,t,y}^{ds,up,max} - \sum_{\tau \in \left\{t - \tau_{u,y}^{ds,up,rec},\ldots,t-1\right\}} \Delta P_{u,\tau,y}^{ds,up}, \qquad \forall u \in S_u,\, t \in S_t, y \in S_y$$

$$0 \leq \Delta P_{u,t,y}^{ds,dn} \leq \Delta_{u,t,y}^{ds,dn,max} - \sum_{\tau \in \left\{t - \tau_{u,y}^{ds,dn,rec},\ldots,t-1\right\}} \Delta P_{u,\tau,y}^{ds,dn}, \qquad \forall u \in S_u,\, t \in S_t, y \in S_y$$

Moreover, the total amount of energy shifted over a planning year can be constrained using,

$$0 \leq \sum_{t \in S_t} \Delta t \cdot \Delta P_{u,t,y}^{ds,dn} \leq \alpha_{u,y} E_{u,y}^{ds,max}, \qquad \forall u \in S_u, y \in S_y.$$

5.1.3 Load curtailment

Load curtailment is similar to the flexibility resources load decrease described in section 3.5.5, but it is "activated" by the system when flexibility from flexibility resources has been fully exploited but the system still needs to reduce load in hour t of reference year y. Only particular resources can be curtailed and their remuneration $C_{u,t,y}^{lc}$, which is much higher than the price for simple load reduction, is decided by regulation. It could be also possible to neglect completely the consumption of the considered resources.

Load curtailment is modelled considering that the reference load $P_{u,t,y}^{ref}$ can be curtailed by an amount $\Delta P_{u,t,y}^{lc}$ such that:

$$0 \le \Delta P_{u,t,y}^{lc} \le P_{u,t,y}^{ref}, \qquad \forall u \in S_u, \; t \in S_t, y \in S_y$$

For what concerns flexibility resources, since $C_{u,t,y}^{lc} \gg C_{u,t,y}^{ds}$ and $C_{u,t,y}^{lc} \gg C_{u,t,y}^{nce}$, we expect that load curtailment will be activated after those two other mechanisms are completely exploited, that is after the lower bound $P_{u,t,y}^{min}$ reached. Then, the cost for the system in hour t is given by

$$C_{u,t,y}^{lc} \Delta P_{u,t,y}^{lc} \tag{22}$$

The hourly demand value as seen from the grid and used in the nodal balance equations need to be defined based on the demand flexibility model explained above. Figure 5-1 shows the relationship between the flexible demand and the reference demand expected over a 24-hour period. The flexible power demand seen at grid nodes $P_{u,t,y}^{flex}$ is be defined as:

$$P_{u,t,y}^{flex} = P_{u,t,y}^{ref} - \Delta P_{u,t,y}^{nce} + \Delta P_{u,t,y}^{ds,up} - \Delta P_{u,t,y}^{ds,dn} - \Delta P_{u,t,y}^{lc}$$
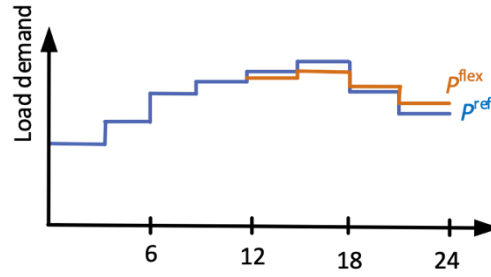
Figure 5-1 - Reference demand versus flexible demand

The bounds on $P_{u,t,y}^{flex}$ can be derived from the variable bounds on load reduction ($\Delta_{u,t,y}^{nce,max}$ nd load shifting ($\Delta_{u,t,y}^{ds,up,max}$_$(u,t,y)^\wedge(ds,down,max)$)) and do not need to be specified explicitly in the planning tool. The only additional constraint that is needed is to require that $P_{u,t,y}^{flex} \ge 0$ to ensure that load curtailment or simultaneous activation of load reduction and load shifting do not turn the load into a net power producer. The variable bounds can be derived using expressions provided previously. Figure 5-2**Figure 5-2** illustrates the bounds $P_{u,t,y}^{flex}$ considering for simplicity only the variable bounds on demand shifting.
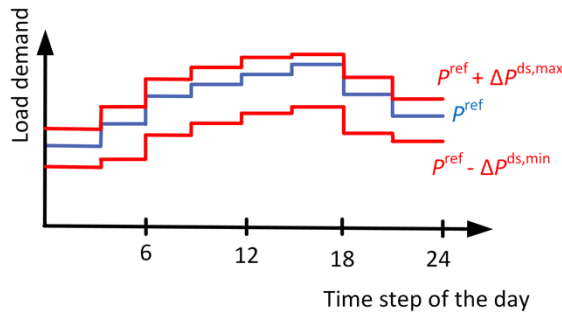
Figure 5-2 - Flexible demand bounds considering load shifting

5.1.4 Demand flexibility investments

Demand flexibility elements can be investment candidates in the sense that an investment is needed to enable the potential demand flexibility at a load point. To make sure that if an investment is made to enable demand flexibility at a load u in horizon y, the demand flexibility is enabled at load u for the next planning horizons we add the following constraint:

$$\alpha_{u,y-1} \leq \alpha_{u,y} \quad \forall u \in S_u, \forall y \in S_y : y > 1$$

Investment costs associated with enabling demand flexibility as well as operational costs were described with more details in Section 3.5.5.

5.2 Storage modelling

We define a generic model which applies for all kind of storage. Examples of different storage technologies are:

- Reservoir Hydro
- Pumped Storage Hydro (PSH)
- Battery energy storage system (BESS)
- Compressed Air Energy Storage (CAES)
- Liquified Air Energy Storage (LAES)
- Hydrogen as an energy storage system
- Thermo-electric storage

5.2.1 Storage constraints

The storage model is composed using following assumptions:

- Impact of reactive power on storage energy efficiency is assumed to be negligible.
- Minimum charging and discharging times are assumed to be zero.
- Times to start charging/discharging are assumed to be zero.

It is also worth mentioning that, as for other grid assets and coherently with our planning perspective, we consider an hourly resolution in our storage model (and so, by default $\Delta t = 1$ hour). Here, the benefits from storage devices can be seen as performing an "arbitrage" between the hours, thus contributing to an overall dispatching cost reduction. Other advantages like support to system balancing are not considered because it would require a much denser time resolution which is not tractable for the size of the problems at hand.

As such, the operational benefits of a storage asset as estimated in the planning tool can be seen as a lower bound on the real benefits that would result from the operation of the storage. This means that in cases where storage investments are chosen by the optimisation, we are on the "safe side" of that investment decision, as the actual benefits will be higher. The other direction, when certain storage assets are not chosen, can always be analysed by means of a sensitivity analysis on the storage costs (e.g., how much do the additional benefits need to be (or the costs be decreased) in order to make the investment profitable at grid level).

We can define the dynamic storage equation using the normalized energy storage level $x_{j,t,y}$:

$$E_{j,y}^{\max} x_{j,t,y} = \left(1 - dr_{j,y}\right)^{\Delta t} E_{j,y}^{\max} x_{j,t-\Delta t,y} + \Delta t \left(\eta_{j,y}^{abs} P_{j,t,y}^{abs} - \frac{P_{j,t,y}^{inj}}{\eta_{j,y}^{inj}} + \xi_{j,t,y} \right), \qquad \forall j \in S_j, t \in S_t, y$$

$$\in S_y \tag{23}$$

where two different variables $P_{j,t,y}^{abs}$ and $P_{j,t,y}^{inj}$ for charging and discharging have been used as the charging and discharging efficiencies can be different. The parameter $\xi_{j,t,y}$ accounts for the potential power demanded or provided by external processes (e.g., water inflow or evaporation in PSH).

The parameter $dr_{j,y}$ is the hourly self-discharge rate of the storage asset j in horizon y. For long-term storage (Pumped Hydro, CAES, etc.), self-discharge is mostly not relevant. For those storage technologies losses occur mainly when charging and discharging and are modelled through the absorption/injection efficiencies. The parameter $dr_{j,y}$ can then be set to 0.

For more short-term storage technologies, self-discharge can however be non-negligible [8], certainly over long periods (e.g., thermal storage, electrical batteries). A nonzero value of the parameter $dr_{j,y}$ then allows to take into account self-discharging of the storage asset (i.e., static losses). Modelling static losses as proportional to the energy level ensure that no losses are accounted for when the storage is empty (i.e., $x_{j,t,y} = 0$), without the need for additional binary variables.

The normalized energy storage level is bound as follows:

$$E_{j,y}^{\min} \leq E_{j,y}^{\max} x_{j,t,y} \leq E_{j,y}^{max}, \qquad \forall j \in S_j, \ t \in S_t, y \in S_y.$$

Then, the charging and discharging power is bound as follows:

$$0 \leq P_{j,y}^{abs,min} \leq P_{j,t,y}^{abs} \leq P_{j,y}^{abs,max}, \qquad \forall j \in S_j, t \in S_t, y \in S_y,$$
$$0 \leq P_{j,y}^{inj,min} \leq P_{j,t,y}^{inj} \leq P_{j,y}^{inj,max}, \qquad \forall j \in S_j, t \in S_t, y \in S_y$$

Typically, $P_{j,y}^{abs,min} = 0$ and $P_{j,y}^{inj,min} = 0$ for all j, y as otherwise there would be constant charging and discharging of the storage present, resulting in the following equations:

$$0 \leq P_{j,t,y}^{abs} \leq P_{j,y}^{abs,max}, \qquad \forall j \in S_j, t \in S_t, y \in S_y,$$
$$0 \leq P_{j,t,y}^{inj} \leq P_{j,y}^{inj,max}, \qquad \forall j \in S_j, t \in S_t, y \in S_y.$$

For existing storage devices, we make sure that charging or discharging is only available when the status of the storage is equal to 1 (to model the unavailability of existing storage):

$$0 \leq P_{j,t,y}^{abs} \leq P_{j,y}^{abs,max} \cdot s_{j,t,y}, \qquad \forall j \in S_j, t \in S_t, y \in S_y,$$
$$0 \leq P_{j,t,y}^{inj} \leq P_{j,y}^{inj,max} \cdot s_{j,t,y}, \qquad \forall j \in S_j, t \in S_t, y \in S_y.$$

For, charging or discharging exclusivity, two classical formulations exist:

$$P_{j,t,y}^{abs} \cdot P_{j,t,y}^{inj} = 0, \qquad \forall j \in S_j, t \in S_t, y \in S_y$$

or

$$V_{j,t,y} P_j^{abs,min} \leq P_{j,t,y}^{abs} \leq V_{j,t,y} P_j^{abs,max}, \qquad \forall j \in S_j, t \in S_t, y \in S_y,$$
$$(1 - V_{j,t,y}) P_j^{inj,min} \leq P_{j,t,y}^{inj} \leq (1 - V_{j,t,y}) P_j^{inj,max}, \qquad \forall j \in S_j, t \in S_t, y \in S_y,$$

with $V_{j,t,y}$ being binary variables defining if the storage asset is charging ($V_{j,t,y} = 1$) or discharging ($V_{j,t,y} = 0$).

Both formulations are either non-linear or use binary variables. Therefore, in the FlexPlan model charging/discharging exclusivity will be considered implicitly via the efficiency. As the general objective is the minimization of operational costs in combination with investments, a solution with simultaneous charging and discharging would indeed be sub-optimal (as both charge and discharge losses would be accounted for – see eq. (23)). Formally, if one sets $P_{j,t,y}^{tot} = P_{j,t,y}^{abs} - P_{j,t,y}^{inj}$ and both $P_{j,t,y}^{abs} > 0$ and $P_{j,t,y}^{inj} > 0$ (simultaneous charging and discharging) then, assuming that the storage charging and discharging efficiencies are not set to 1 (lossless storage asset):

- If $P_{j,t,y}^{tot} > 0$, setting $P_{j,t,y}^{abs} = P_{j,t,y}^{tot}$ and $P_{j,t,y}^{inj} = 0$ yields solution with a lower objective function
- If $P_{j,t,y}^{tot} < 0$, setting $P_{j,t,y}^{abs} = 0$ and $P_{j,t,y}^{inj} = -P_{j,t,y}^{tot}$ yields solution with a lower objective function

Moreover, we add the following limiting constraints (which preserve the linearity of the model and do not add binary variables the problem):

$$P_{j,t,y}^{abs} + P_{j,t,y}^{inj} \leq \max\left(P_{j,t,y}^{abs,max}, P_{j,t,y}^{inj,max}\right), \qquad \forall j \in S_j, \, t \in S_t, y \in S_y.$$

Moreover, constraints will not eliminate the risk of having simultaneous absorption and injection but will limit it, in the case where $\eta_{j,y}^{abs}$ and $\eta_{j,y}^{inj}$ would both be set to 1. Indeed, when there is overgeneration, $P_{j,t,y}^{abs}$ should be maximized and so $P_{j,t,y}^{inj}$ is pushed to zero.

Generally, ramping constraints can be considered for storage, although for most of the storage technologies, the ramping rate will be less than one hour, and those constraints will be omitted:

$$P_{j,t,y}^{abs} - P_{j,t-\Delta t,y}^{abs} \leq \Delta t \cdot r_{j,y}^{abs,max}, \qquad \forall j \in S_j, t \in S_t, y \in S_y,$$
$$P_{j,t,y}^{inj} - P_{j,t-\Delta t,y}^{inj} \leq \Delta t \cdot r_{j,y}^{inj,max}, \qquad \forall j \in S_j, t \in S_t, y \in S_y.$$

Here, we also have to ensure it is the case only if the storage is available (status=1):

$$s_{j,t,y} \cdot (P_{j,t,y}^{abs} - P_{j,t-\Delta t,y}^{abs}) \leq \Delta t \cdot r_{j,y}^{abs,max}, \qquad \forall j \in S_j, t \in S_t, y \in S_y,$$
$$s_{j,t,y} \cdot (P_{j,t,y}^{inj} - P_{j,t-\Delta t,y}^{inj}) \leq \Delta t \cdot r_{j,y}^{inj,max}, \qquad \forall j \in S_j, t \in S_t, y \in S_y.$$

Initial and final conditions for the energy content are needed, e.g., to avoid storage being emptied towards the end of the planning year considered:

$$E_{j,y}^{\max x_{j,0,y}} = E_{j,y}^{init}, \qquad \forall j \in S_j, y \in S_y,$$
$$E_{j,y}^{\max x_{j,T,y}} \geq E_{j,y}^{init}, \qquad \forall j \in S_j, y \in S_y.$$

Finally, we also add an integral constraint on the maximum amount of energy which can be absorbed over a year in order to avoid having some storages which are over-used during the simulated years. As for the other storage parameters, the value of $E_{j,y}^{abs,max}$ provided as an input of the planning tool to allow the user to account for the type technology that is used, its expected lifetime, etc.

$$\sum_{t \in S_t} \Delta t \cdot P_{j,t,y}^{abs} \leq E_{j,y}^{abs,max}, \quad \forall j \in S_j, y \in S_y.$$

Simmilarly, for candidate storage technologies, the dynamic storage equations remain the same except that losses and external exchanges are not accounted if the storage is not invested in:

$$E_{jc,y}^{max} x_{jc,t,y} = \left(1 - \alpha_{jc,y} dr_{j,y}\right)^{\Delta t} E_{jc,y}^{max} x_{jc,t-\Delta t,y} + \Delta t \cdot \left(\eta_{jc,y}^{abs} P_{jc,t,y}^{abs} - \frac{P_{jc,t,y}^{inj}}{\eta_{jc,y}^{inj}} + \alpha_{jc,y} \xi_{jc,t,y}\right), \qquad \forall jc \in S_{jc}, t \in S_t, y$$

$$\in S_y.$$

The normalized energy storage level and charging and discharging power should be bound using the binary investment decision variable:

$$E_{jc,y}^{min} \alpha_{jc,y} \leq E_{jc,y}^{max} x_{jc,t,y} \leq E_{jc,y}^{max} \alpha_{jc,y}, \qquad \forall jc \in S_{jc}, t \in S_t, y \in S_y,$$
$$0 \leq P_{jc,t,y}^{abs} \leq \alpha_{jc,y} P_{jc,y}^{abs,max}, \qquad \forall jc \in S_{jc}, t \in S_t, y \in S_y,$$
$$0 \leq P_{jc,t,y}^{inj} \leq \alpha_{jc,y} P_{jc,y}^{inj,max}, \qquad \forall jc \in S_{jc}, t \in S_t, y \in S_y.$$

Finally, initial and final conditions for the energy contents are also updated depending on the investment variables:

$$E_{jc,y}^{\max x_{jc,0,y}} = \alpha_{jc,y} E_{jc,y}^{init}, \qquad \forall jc \in S_{jc}, y \in S_y,$$
$$E_{jc,y}^{\max x_{jc,T,y}} \geq \alpha_{jc,y} E_{jc,y}^{init}, \qquad \forall jc \in S_{jc}, y \in S_y.$$

To ensure that selected candidate storage cannot be deactivated after the investment decision is taken for a certain planning year, following constraint is required:

$$\alpha_{jc,y-1} \leq \alpha_{jc,y} \quad \forall lc \in S_{jc}, \forall y \in S_y: y > 1.$$

Reactive power is another control variable of a generic storage device and it is included within the model since it can be beneficial for distribution network planning. As stated above, reactive power is assumed to not have any impact on the energy efficiency (Joule losses due to higher currents are neglected) and a reasonable approximation of the storage capability can be formulated as follows:

$$Q_{j,y}^{min} \leq Q_{j,t,y} \leq Q_{j,y}^{max}, \qquad \forall j \in S_j, t \in S_t, y \in S_y.$$

For candidates, the same constraint can be defined as:

$$Q_{j,y}^{min} \cdot \alpha_{jc,y} \leq Q_{j,t,y} \leq Q_{j,y}^{max} \cdot \alpha_{jc,y}, \qquad \forall jc \in S_{jc}, t \in S_t, y \in S_y.$$

Finallly, fixed (investment) costs and operational costs for storage assets were described in Section 0.

5.2.2 Modelling the flexibility of hydropower plants

Realistic modelling of hydropower plants for the purpose of hydropower production scheduling and hydro-thermal market analyses involves complex stochastic optimisation models [9, 10]. Among other factors, such models need to account for the stochastic inflow over the scheduling horizon, the value of having energy stored at the end of the scheduling horizon, hydrological coupling in water courses linking different power plants and reservoirs, limitations within the water courses such as environmental constraints. Such level of detail is outside of the scope of the FlexPlan model, and compromises are made for representing hydropower plants in a way that is deemed sufficiently accurate for the purpose of the model. An important issue is to represent the availability and marginal cost of flexibility from hydropower in a reasonable way.

In the FlexPlan model, the models for VRES-based power plants and energy storage devices can be used as building blocks for several different representations of hydropower plants. Two very simple representations are i) representing a hydropower plant as a generic energy storage device ("a big battery") with a scenario-specified inflow time series $\xi_{j,t,y}$, or ii) representing hydropower generation as a VRES-based generator with a fixed scenario-specified generation schedule $P_{g,t,y}^{res,ref}$. The latter representation (ii) is most applicable to run-of-river hydropower plants but are unable to capture the flexibility of regulated, dispatchable hydropower plants with reservoirs. The former representation (i) includes storage balance equations representing the "state of charge" dynamics through the year, energy capacity constraints $E_{j,y}^{max}$ and $E_{j,y}^{min}$ and power injection/absorption capacities $P_{j,t,y}^{inj,max}$ and $P_{j,t,y}^{abs,max}$t is more applicable to dispatchable hydropower plants but will typically overestimate the flexibility it can provide in practice. This overestimation is due to neglecting the uncertainty and variability of inflow, environmental constraints as well as the hydraulic coupling in water courses. The two representations (i, ii) with exemplary power injection curves are illustrated below.
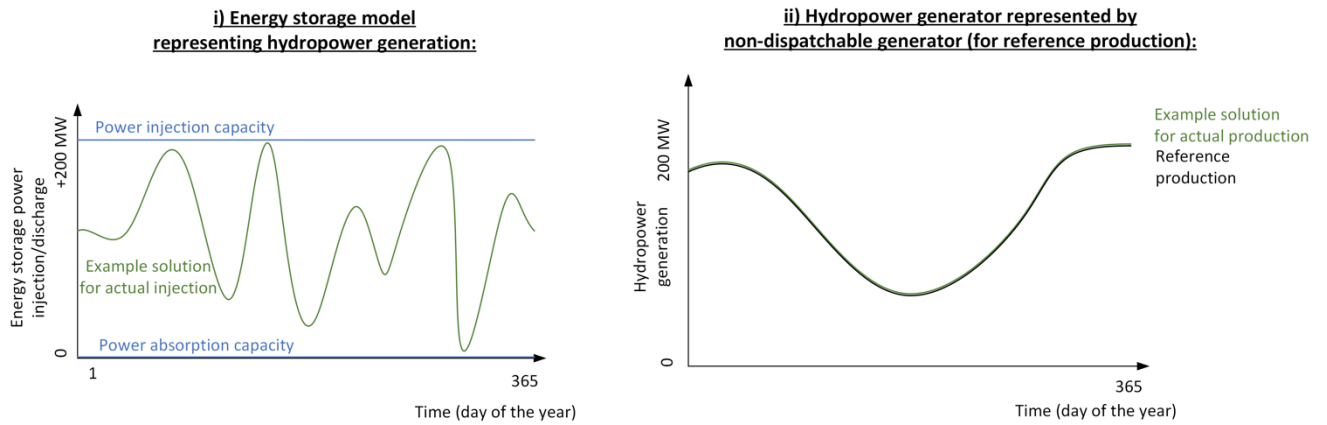
Figure 5-3 - Representing hydropower generation by i) generic ESS model or ii) non-dispatchable

These two alternatives represent two extremes in terms of how flexible the hydropower plants are modelled to be. Although all the possible representations that have been considered in the FlexPlan project have their drawbacks in terms of fidelity and input data requirements, a combination of the two (i + ii) is proposed as a fair compromise: A non-dispatchable generator with a reference production time series $P_{g,t,y}^{res,ref}$ (i) is combined with (ii) a generic ESS model with absorption and injection limits (power capacities) $P_{j,t,y}^{abs,min}$
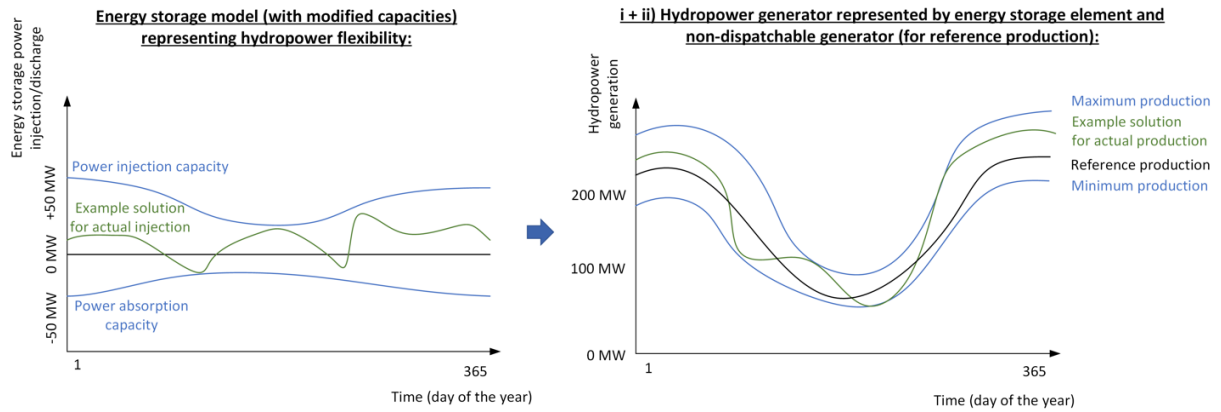
Figure 5-4 - Representing hydropower generation with reservoirs by a generic ESS model combined with a VRES-based generator for the reference production time series.

and $P_{j,t,y}^{inj,max}$ respectively. These power capacities could be specified as time series with seasonal variation to represent that the flexibility potential typically varies over the year. The two elements (i) and (ii) could be connected to the same bus of the grid model and appear as a single element from a power flow perspective. The modified ESS model (i) is illustrated on the left-hand side of the Figure *5-4*, while the right-hand side illustrates the combination (i + ii). This allows hydropower generation within a restricted flexibility band around the reference production time series.

Hydropower generators with reservoirs and pumping capabilities (pumped-hydro energy storage) can be represented using representation (i) or representation (i+ii): In representation (i), one can set power absorption capacity $P_{j,t,y}^{abs,max} > 0$ include the pumping capability ($P_{j,t,y}^{abs,max} = 0$ is assumed in the illustration of representation (i) above). In representation (i+ii), pumping capabilities can be represented by setting the $P_{g,t,y}^{res,ref}$ and $P_{j,t,y}^{abs,max}$ such that the minimum power production $P_{g,t,y}^{res,ref} - P_{j,t,y}^{abs,max} < 0$ at least during parts of the year.

Which representation is most appropriate to use when running the tool for a case could be decided based on the hydropower generators to be represented in the case and the availability of data. For many hydropower plants, the combined representation (i + ii) is likely to give more realistic generation schedules than a pure ESS model that might overestimate the flexibility of the generator. Another advantage of this representation is that input data for the inflow is not needed directly in the FlexPlan tool. A related advantage may be that it is less dependent on accurate estimates of the reservoir energy capacity than representation. The drawback is the input data for the power capacity time series need to be judiciously chosen to represent the available flexibility correctly. These power capacities are not "fundamental" parameters of the hydropower generators that can be specified directly but instead need to be estimated more indirectly through calibration against existing, fundamental hydro-thermal market models.

6 Environmental impact modelling

6.1 Air quality impact modelling

6.1.1 Conceptual formulation of the cost function

The evaluation of air quality impact is limited to thermal generation and is carried out through a modelling approach whose conceptual scheme is reported in Figure *6-1*.

Figure 6-1 - Conceptual model for air quality impact evaluation.

The main input variable of the air quality model is one of the outputs of the optimization, namely the hourly energy production

$$P_{g,t,y} \qquad \forall g \in S_g, t \in S_t, y \in S_y$$

On the basis of specific emission factors G^{ef}, that express the amount of pollutant mass emitted by the energy production unit, depending on pollutant p and generator g

$$G^{ef}_{g,p} \qquad \forall g \in S_g, p \in S_p,$$

the hourly emission rate EM can be computed

$$EM_{g,t,y,p} \qquad \forall g \in S_g, t \in S_t, y \in S_y, p \in S_p.$$

For each generator a simplified air quality model is derived in order to define a linear relationship between emission rates and air quality concentration. The simplified model, whose definition is discussed below, depends on:

- Generator features: Stack geometry and plume properties (e.g., flow rate, temperature and velocity)
- Background concentrations representing the air quality concentrations due to all the other sources affecting the domain

- Meteorology

The simplified air quality model is expressed in terms of a linear relationship, therefore on the basis of a simple coefficient

$$SC_{g,t,y,p} \qquad \forall g \in S_g, t \in S_t, y \in S_y, p \in S_p$$

that represent the link between the emission rate and the corresponding air quality concentration. By means of the simplified model the air quality concentration, expressed as:

$$AQ_{g,t,y,p} \qquad \forall g \in S_g, t \in S_t, y \in S_y, p \in S_p$$

can be derived for each hour, pollutant, year and generator.

The air quality concentration represents the main input for the impact functions that allow to compute the effect of air pollution on health (and optionally on vegetation). Such impact is expressed through a set of specific indicators (such as years of life lost, number of hospitalisation days, etc.). The evaluation of the air quality impacts requires yearly indicators, such as the yearly mean concentration

$$\overline{AQ}_{g,y,p} \qquad \forall g \in S_g, y \in S_y, p \in S_p.$$

Once the yearly air quality indicator is computed, health impacts

$$\overline{IM}_{g,y,p,imp} \qquad \forall g \in S_g, y \in S_y, p \in S_p, imp \in S_{imp}$$

can be computed through a linear relationship expressed by an impact coefficient

$$IMC_{p,imp} \qquad \forall \, p \in S_p, imp \in S_{imp}.$$

IMC quantify the effect of a specific air quality level on a health indicator. The final step of the modelling chain concerns the evaluation of the environmental costs due to air pollution. Such cost is expressed as

$$\overline{C}^{AQ}_{g,y,p,imp} \qquad \forall g \in S_g, y \in S_y, p \in S_p, imp \in S_{imp}$$

and it can be computed on the basis of the corresponding air quality impact through a cost function that expresses a linear relationship between impacts and cost. Such relationship is represented by a coefficient

$$CC_{imp} \qquad \forall \, imp \in S_{imp}$$

that allows to express the monetary cost of an air quality impact. The total cost related to each generator is then derived as a sum of all costs due to each pollutant p and impact imp. In conclusion the air quality cost can be expressed as:

$$\overline{C}^{AQ}_{g,y} = \sum_{p \in S_p} \sum_{imp \in S_{imp}} \overline{C}^{AQ}_{g,y,p,imp} =$$

$$\overline{C}^{AQ}_{g,y} = \sum_{p \in S_p} \sum_{imp \in S_{imp}} CC_{imp} \cdot \overline{IM}^{AQ}_{g,y,p,imp} =$$

$$\overline{C}^{AQ}_{g,y} = \sum_{p \in S_p} \sum_{imp \in S_{imp}} CC_{imp} \cdot IMC_{p,imp} \cdot \overline{AQ}_{g,y,p} =$$

$$\overline{C}^{AQ}_{g,y} = \sum_{p \in S_p} \sum_{imp \in S_{imp}} CC_{imp} \cdot IMC_{p,imp} \cdot \frac{1}{8760} \sum_{t \in S_t} AQ_{g,t,y,p} =$$

$$\overline{C}^{AQ}_{g,y} = \sum_{p \in S_p} \sum_{imp \in S_{imp}} CC_{imp} \cdot IMC_{p,imp} \cdot \frac{1}{8760} \sum_{t \in S_t} SC_{g,t,y,p} \cdot EM_{g,t,y,p} =$$

$$\overline{C}^{AQ}_{g,y} = \sum_{p \in S_p} \sum_{imp \in S_{imp}} CC_{imp} \cdot IMC_{p,imp} \cdot \frac{1}{8760} \sum_{t \in S_t} SC_{g,t,y,p} \cdot G^{ef}_{g,p} \cdot P_{g,t,y}$$

The simplified air quality model is derived through a Taylor expansion approach on the basis of a full 3D chemistry and transport model (CTM) able to reproduce all chemical and physical processes which air pollutants undergo in the atmosphere. The CAMx model (Comprehensive Air Quality model with extensions, [11]) is applied in this project together with the embedded DDM algorithm (Decoupled Direct Method for sensitivity analysis in a three-dimensional air quality model, [12]) able to derive the sensitivity coefficients *SC* to be introduced in the simplified model. It is important to note that SC coefficients are derived as a perturbation with a reference case. This means that a reference working profile for each generator to be modelled in the optimization function. A schematic representation of the methodology is represented in Figure *6-2*.
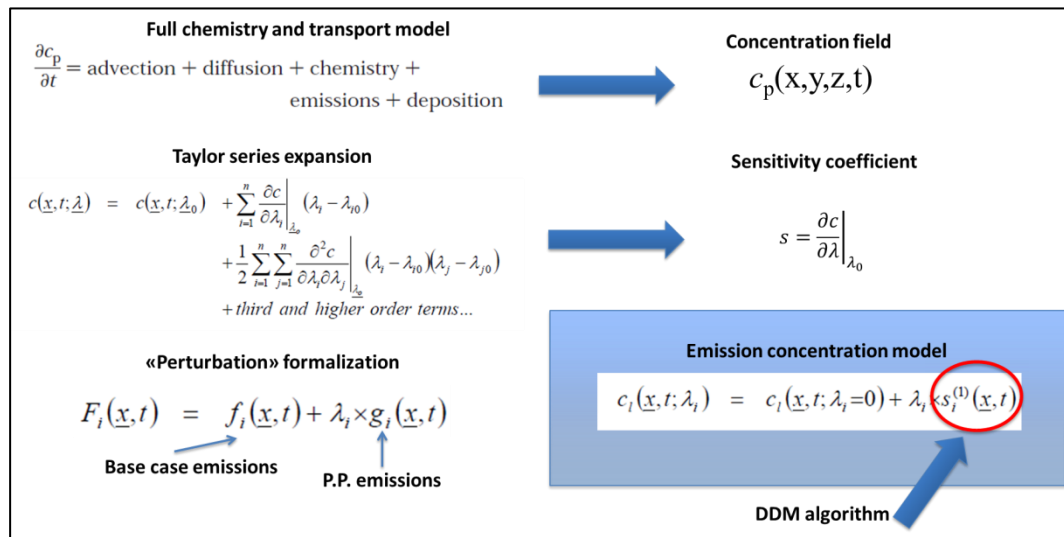
Figure 6-2 - Derivation of the simplified air quality model

A second simplification concerns the definition of a unique air quality indicator for each generator, rather than a concentration field that is the usual output of air quality model. To this aim a specific algorithm is developed and applied in order to define a specific "spatial fingerprint" of each generator (see Figure *6-3*). The definition of the fingerprint is essential in order to establish an area of influence of each generator over which computing the corresponding impacts.
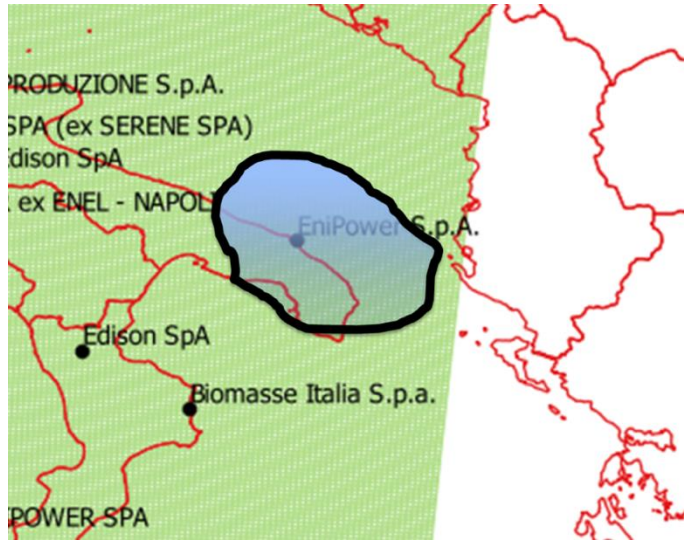
Figure 6-3 - Example of a power generator "spatial
fingerprint"

As already mentioned, on the basis of air quality concentrations, expressed in terms of yearly indicators,
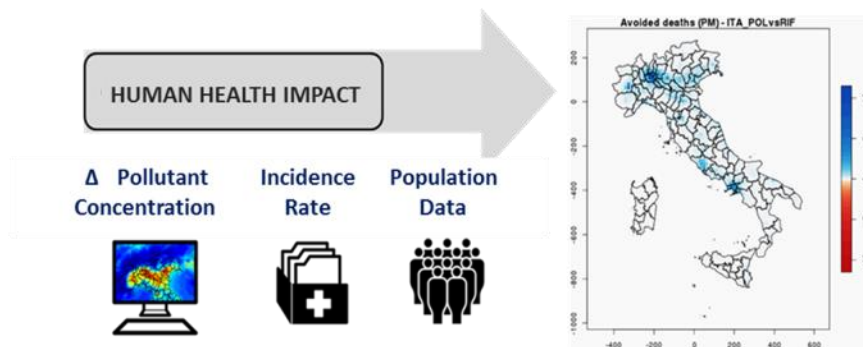
Figure 6-4 - Schematic description of the methodology implemented to evaluate air
quality impacts

corresponding impacts and costs are evaluated. In the following Figure *6-4* a schematic description of the methodology is depicted.

Following the well-established European Environment Agency methodology [13], a log-linear model can be used to derive the health impact function, which results in the following equation:

$$HI_{ijk} = P_{ij} \cdot r_{ijk} \left(1 - \frac{1}{e^{\beta_k \Delta C_{ij}}} \right)$$

where HI_{ijk} is the impact on health endpoint k (e.g. avoided mortality) due to change in PM$_{2.5}$ concentration ΔC_{ij} in the grid cell (i,j), P_{ij} is the population in grid cell (i,j), β_k is a parameter which accounts for the sensitivity of impact on health endpoint k on a concentration change ΔC_{ij} and r_{ijk} is the baseline incidence rate

of endpoint k in cell (i, j). In this work only on one health endpoint has been taken into account (i.e., $k = \{1\}$), namely the number of premature deaths, as it is classified as reliable by the most up-to-date European literature on PM2.5 health impact assessments [13]. The value of the parameter β_k is directly linked to the relative risk (RR_k), which is defined in the log-linear model as:

$$RR_k = \frac{r_{k,BC}}{r_{k,SC}} = e^{\beta_k \Delta C}$$

Where $r_{k,BC}$ stands for the baseline incidence rate whereas $r_{k,SC}$ is the incidence rate under a defined scenario. As RR_k values and their 95% confidence interval for different health endpoints k have been estimated during the European HRAPIE project (WHO, 2013) for a 10 μg m-3 change in PM2.5, β_k can be simply computed as:

$$\beta_k = \frac{\ln(RR_k)}{10}$$

The total impact for the endpoint k is finally the sum of I_{ijk} over each grid cell $(i.j)$.

It should be noted that for premature deaths, ΔC_{ij} is expressed in terms of the change in the annual average PM2.5 concentration, as premature mortality is supposed to be an effect of a long-term exposure. We assume, coherently with WHO, 2013, that premature mortality affects only 30+ adults. Following [13], in this work we use a value of RR per 10 μg/m3 equal to 1.062 (1.040-1.083, 95% confidence interval) for premature mortality. In order to allow a more robust economical evaluation of the premature mortality impact, health impacts are expressed as Years of Life Lost (YOLL) through the equation

$$YOLL = \sum_{k=1}^{max} M_k \cdot LE_k$$

where M_k and LE_k are, respectively, the number of premature death and life expectancy at age k.

The health impact of air pollution can be expressed also in terms of monetary costs. The key issue to derive a monetary cost for air pollution is the definition of the so-called **Value of a Statistical Life** (VSL) that is the amount of money that a community of people are **willing to pay** (WTP) to lower the risk of an anonymous instantaneous premature death within that community (cost of mortality at a level of society as a whole). Starting from the VLS values is then possible to define, respectively:

- *Cost of mortality from air pollution*: VSL * number of premature deaths
- *Economic benefit of a mitigating action*: VLS * number of prevented deaths.

Once the VSL is defined, it can be easily adapted in terms of other impact metrics than number of premature deaths, such as YOLL, therefore expressing the cost in terms of years of life lost. In literature are available several studies aiming at providing a monetary cost for impact studies. Particularly, in this project we refer to [14] and [15].

6.1.2 Operational implementation of the Cost function

Starting from the general formulation of the cost function

$$\overline{C}_{g,y}^{AQ} = \sum_{p \in S_p} \sum_{imp \in S_{imp}} CC_{imp} \cdot IMC_{p,imp} \cdot \frac{1}{8760} \sum_{t \in S_t} SC_{g,t,y,p} \cdot G_{g,p}^{ef} \cdot P_{g,t,y}$$

and focusing on a specific year, impact and pollutant for the sake of simplicity, we can express the cost function as:

$$\overline{C}_g^{AQ} = V \cdot K \cdot \frac{1}{8760} \sum_{t \in S_t} SC_{g,t} \cdot G_{g,p}^{ef} \cdot P_{g,t}$$

where V = CC_{imp} and K = $IMC_{p,imp}$

Particularly, in the framework of this project we consider as air pollutant particulate matter (PM$_{2.5}$) and as impact indicator the yearly metric Years of Life Lost (YOLL) [14], widely used in health impact studies concerning air pollution.

The key aspect of the simplified model design is the quantification of $SC_{g,t}$, representing the overall delta concentration impact due to a unitary emission load for plant g at time t.

Particularly $SC_{g,t}$ can be expressed as:

$$SC_{g,t} = \sum_{i,j \in N_g} \Delta C_{uni}(g,t,i,j) \cdot pop(i,j)$$

where N_g is the set of cells impacted by power plant g (i.e., the power plant fingerprint).

As already mentioned, the simplified relationship between emission and concentration, expressed by $SC_{g,t}$, is derived thanks to DDM tool, that allows to compute the so called sensitivity coefficients (s_r), quantifying the concentration variation deriving from an emission variation corresponding to the whole perturbation.

The sensitivity coefficients are computed by CAMx/DDM for each cell and each hour, therefore they should be expressed as $s_r(i,j,t)$, but for sake of simplicity we refer them as s_r.

In the framework of this project the "perturbation" consists in a "reference" temporal profile, at hourly resolution, of the power plant (PP) emission load, due to a corresponding hourly production "reference" profile. In order to reduce the errors related to the linearization of the simplified model, the modelled profile should correspond, as far as possible, to a realistic operational functioning close to the maximum load.

The "base case" consists in a simulation where all main emission sources (road transport, heating, …) are considered, in order to compute "perturbations" close to a realistic air pollution burden.

In the test phase, described below, we computed sensitivity coefficients for three different perturbation strengths, with respect to the "reference" load, corresponding to a fraction of the total PP emissions: 25, 50 and 100%. This means that the first run computed DDM sensitivity coefficients (s_r) representing an emission perturbation corresponding to 25% of the reference PP emissions; the second one corresponding to a 50% of the reference PP emissions and the last one corresponding to the total reference PP emissions.

We performed three runs to investigate the possible influence of non-linear processes in the computation of the coefficients.

Indeed, in case of fully linear behaviour $s_{100}=2*s_{50}=4*s_{25}$, only one run would be needed, while in case of non-linearity sensitivity coefficients should depend on perturbation strength. The analysis reported below pointed out that in most cases the model shows an almost linear behaviour and, particularly, that s_{100} coefficients provide the most robust results when compared to the full 3D model.

In order to provide simple but realistic coefficients to be used to test the objective function independently from a specific meteorological year, the hourly s_r values were averaged on hourly basis over two shorter periods (January and June) in order to derive two "mean days" (i.e., 24 mean hourly values), one for winter and one for summer. Such sets of 48 sensitivity coefficients were computed for each grid cell. We refer to these coefficients as $s_{W,r}(h,i,j)$ and $s_{S,r}(h,i,j)$, where "W" and "S" respectively refer to January (winter) and June (summer) monthly averages, r to the perturbation strength, h to the hour of the day (0 to 23), i,j to each grid cell. $r = 100$ was finally selected as perturbation strength, according to the performance evaluation described below.

For example:

$S_{W,100}(4,24,35)$ quantifies the variation of the hourly PM10 concentration in cell (24, 35) at 04:00 on a January day, due to a 100% variation of the PP emissions.

Therefore, supposing that $C_0(4,24,35)$ is the $PM_{2.5}$ concentration in cell (24, 35) at 04:00 when the PP is off, the $PM_{2.5}$ concentration corresponding to an emission load equal, for example, to 60% of the reference load is expressed as:

$C_{0.6}(4,24,35) = C_0(4,24,35) + 0.6 * S_{W,100}(4,24,35)$ in winter (and $C_{0.6}(4,24,35) = C_0(4,24,35) + 0.6 * S_{S,100}(4,24,35)$ in June).

This implies that the concentration variation can be simply expressed as:

$\Delta C_{0.6}(4,24,35) = 0.6 * S_{W,100}(4,24,35)$ in January (and $\Delta C_{0.6}(4,24,35) = 0.6 * S_{S,100}(4,24,35)$ in June).

Following the approach previously discussed, in order to compute the air quality delta concentration related to each PP the following indicator is introduced:

$$Imp_{ref}(g, seas, h) = \sum_{i,j \in N_g} Sseas_{100}(h, i, j) \cdot pop(i, j)$$

where: $Imp_{ref}(g, seas, h)$ represents the whole PM_{10} delta concentration impact of the power plant g, when it works at 100% emission load at hour h (0...23) of season $seas$ (winter, summer). The summation includes all cells belonging to the area of influence of PP g (i.e., the PP fingerprint). On the basis of a few preliminary tests, we consider as PP fingerprint an area having a radius of 30 km around the PP stack.

Then $Imp_{ref}(g,d,h)$, representing the PM_{10} delta concentration impact of the power plant g, when it works at 100% emission load at hour h (0...23) for a generic day d can then be expressed as:

$$Imp_{ref}(g, d, h) = wg(w, d) \cdot Imp_{ref}(g,w,h) + wg(s, d) \cdot Imp_{ref}(g,s,h)$$

where $wg(seas,d)$ is reported in Table 1.

Month	Day of the year (d)	wg(winter)	wg(summer)
1	1-31	1.0	0.
2	32-59	0.8	0.2
3	60-90	0.6	0.4
4	91-120	0.4	0.6
5	121-151	0.2	0.8
6	152-181	0.	1.0
7	182-212	0.	1.0
8	213-243	0.2	0.8

9	244-273	0.4	0.6
10	274-304	0.6	0.4
11	305-334	0.8	0.2
12	335-365	1.0	0.

Table 1- daily values of seasonal wights to be used to modulate Production and Emission reference values for each day of the year.

We can now express $SC_{g,t}$ in terms of $Imp_{ref}(g,d,h)$ as:

$$SC_{g,t} = \frac{Imp_{ref}(g,d,h)}{Emi_{ref}(g,d,h)}$$

Where $emi_{ref}(g,d,h)$ [kg/h]is the reference hourly emission load introduced in CAMX/DDM as perturbation term.

Introducing the double summation on d and h, instead of t, we obtain:

$$\overline{C}_g^{AQ} = V \cdot K \cdot \frac{1}{8760} \sum_{d=1}^{365} \sum_{h=0}^{23} \frac{Imp_{ref}(g,d,h)}{Emi_{ref}(g,d,h)} \cdot G_g^{ef} \cdot P_{g,d,h}$$

In order to simplify the introduction of the cost term into the objective function we can perform a final conversion introducing an additional term, namely:

$$Prod_{\text{ref}}(g,d,h) = \frac{Emi_{ref}(g,d,h)}{G_g^{ef}}$$

$Prod_{\text{ref}}(g,d,h)$ is computed on the basis of proper Emission Factors G_g^{ef} [kg/MWh], expressing the emission load for unit of energy produced:

Likewise $Imp_{ref}(g,d,h)$, $Prod_{\text{ref}}(g,d,h)$ as can be computed as a weighted average of the corresponding seasonal values, i.e.:

$$Prod_{\text{ref}}(g,d,h) = wg(\text{w},d) \cdot Prod_{\text{ref}}(g,\text{w},h) + wg(\text{s},d) \cdot Prod_{\text{ref}}(g,\text{s},h)$$

Substituting G_g^{ef} we finally obtain:

$$\overline{C}_g^{AQ} = V \cdot K \cdot \frac{1}{8760} \sum_{d=1}^{365} \sum_{h=0}^{23} Imp_{ref}(g,d,h) \cdot \frac{P_{g,d,h}}{Prod_{ref}(g,d,h)}$$

where $\overline{C}_g^{AQ}$ is the yearly cost due to the air quality impact of power plant g, working with an energy production profile $P_{p,d,h}$ [MWh], representing is the decision variable value for plant g, on day d, hour h.

The final step of the linearization process concerns the cost function.

As previously introduced health impacts are generally related to concentration variations through a logarithmic model, such as:

$$HI_{ijk} = P_{ij} \cdot r_{ijk} \left(1 - \frac{1}{e^{\beta_k \Delta \overline{C}_{ij}}}\right)$$

where $P(i,j)$ is population in cell (i,j) and $\Delta \tilde{C}(i,j)$ is the variation of the yearly mean concentration of PM$_{2.5}$ in cell i,j, r_{ijk} and β_k parameters specific for each health impact metric. In order to develop a fully linear simplified model the health impact (HI) function can be linearized; this implies that the general formulation

$$HI_{\text{ori}}(g) = \sum_{i,j \in N_g} f\big(\Delta \bar{C}(i,j)\big) \cdot pop(i,j)$$

should be expressed as

$$HI_{\text{lin}}(g) = \sum_{i,j \in N_p} \text{K} \cdot \Delta \bar{C}(i,j) \cdot pop(i,j) = \text{K} \cdot \frac{1}{8760} \sum_{d=1}^{365} \sum_{h=0}^{23} Imp(g,d,h)$$

Considering that the hourly concentration delta is linearly related to *Sseas100(h,i,j)* through $\lambda(p,d,h)$ the previous formula can be easily rearranged in order to obtain the *Cost(p)* formulation previously introduced.

If the hypothesis of a linear behaviour for *HI(g)* holds we can write:

$$HI_{\text{ori}}(g) = HI_{\text{lin}}(g)$$

And then easily compute K.

Indeed, if we consider that, for example, the yearly mean delta concentration corresponds to a specific *Sseas100(h,i,j)* field, (e.g. the winter one for hour 3), *HIlin(g)* can be computed as:

$$HI_{\text{lin}}(g) = \text{K} \cdot \sum_{i,j \in N_p} Sw_{100}(3,i,j) \cdot pop(i,j)$$

And then

$$HI_{\text{lin}}(g) = \text{K} \cdot Imp(g,\text{w},3)$$

Computing *HI(g)* also with the original formulation

$$HI_{\text{ori}}(g) = \sum_{i,j \in N_g} f\big(Sw_{100}(3,i,j)\big) \cdot pop(i,j)$$

K can be expressed as

$$K = \frac{HI_{\text{ori}}(g)}{Imp(g,\text{w},3)}$$

Performing such calculation a few times for different *Imp(g,s,h)* fields will allow to estimate a rather robust value for K. Considering that the range of $\Delta \bar{C}(i,j)$ due to PP functioning is rather limited our tests confirmed that the health impact function can be linearized. Similar results were obtained by [14].

Finally, considering as exposure response function the Years of Life Lost (YOLL), widely used in environmental studies to evaluate PM impact, the V coefficient can be set equal to 57 510 EUR$_{2013}$ per YOLL, according to [15].

In summary, in order to evaluate the air quality cost of each power plant, the following parameters will be provided as an input to the model:

a) *wg(seas,d)*: daily weights [2,365]

b) *Imp$_{ref}$(g, seas, h)*: represents the whole PM$_{10}$ delta concentration impact of the power plant *g*, when it works at 100% of the reference emission load at hour *h* (0..23) of season *seas* (winter, summer).

c) *Prod$_{ref}$(g, seas, h)*: represents the energy production of the power plant *g*, when it works at 100% of the reference emission load at hour *h* (0..23) of season *seas* (winter, summer).

d) *V*: coefficient expressing the linearized relationship between delta concentration and YOLL

e) *K*: cost per YOLL [EUR$_{2013}$/YOLL]

6.2 Carbon footprint impact modelling

Evaluating the carbon footprint of a product or a service means calculating all the emission of greenhouse gasses occurring during the entire life cycle of the analysed product/service. Emissions are accounted for in terms of their potential effect on climate, called Global Warming Potential, which is measured as kg of carbon dioxide equivalent (kg CO_2 eq). In Figure *6-5* a schematic representation of the carbon footprint of electricity
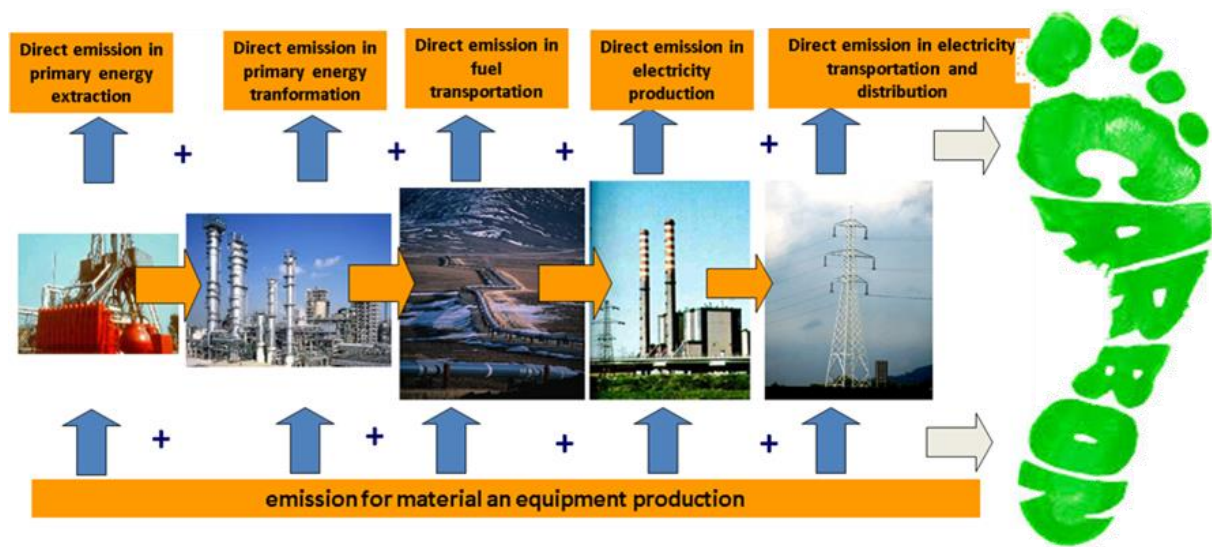
Figure 6-5 - Schematic description of the electricity production carbon footprint

production is represented.

Keeping the general goal of the FlexPlan project in mind, the approach for the inclusion of carbon footprint of candidates in the planning tool will be limited to the differences among investment options. The carbon footprint of the existing network will not be considered, as it remains the same for all different grid expansion options. Obviously, the effect of CO_2 emitting generators in the system are accounted for, which has the largest carbon footprint impact. The carbon footprint of each candidate can be expressed as the sum of the carbon footprint of electric production and the carbon footprint of grid components. The considered grid components in this framework are new lines, new storage systems, new HVDC converters and phase shifting transformers (PSTs). The annual carbon footprint of each candidate can be expressed as:

$$CO_2 tot = \sum_{i=1}^{n} kWh_i * CO_2 eq_i + \sum_l km_l * CO_2 eq_{HVL} + \sum_S kWh_S * CO_2 eq_S + \sum_{HVDC} MVA_{HVDC} * CO_2 eq_T + \sum_{PST} MVA_{PST} * CO_2 eq_T,$$

where:

- $CO_2 eq_i$ is the Life Cycle CO_2eq emission of generation technology *i* per kWh;
- $CO_2 eq_{HVL}$ is the CO_2eq emission for HV lines construction per km and per year;
- $CO_2 eq_S$ is the CO_2eq emission for Li-ion Batteries construction per kWh of capacity and per year;
- $CO_2 eq_T$ is the CO_2eq emission for HV transformers construction per MVA and per year.

Since no installation of new generators is considered by the optimization tool, for the sake of simplicity, the carbon footprint evaluation will not consider the power plant construction and decommissioning. This means that carbon footprint of enabled energy production will be limited to the electricity produced by thermal power plants. The carbon footprint of electricity production from non-thermal renewable power plants (wind, solar, hydro) is mainly due to power plant construction and decommissioning and thus out of scope. Keeping the Life Cycle perspective of the carbon footprint calculations in mind, emissions due to energy source extraction (including biomass cultivating), to fuel production and to fuel combustion in the power plant are considered:

$$CO_2 eq_{Elprod} = \sum_{i=1}^{n} kWh_i * eff_i * (CO_2 eq_{FiP} + CO_2 eq_{FiC})$$

where:

- $CO_2 eq_{Elprod}$ is the carbon footprint of electricity production;
- kWh_i is the production from each generator i;
- eff_i is the efficiency of the generator i in MJ/kWh;
- $CO_2 eq_{FiP}$ is the greenhouse gasses emission (in kg of CO_2eq/MJ) due to fuel F production (including primary energy source extraction) used in generator i;
- $CO_2 eq_{FiC}$ is the greenhouse gasses emission (in kg of CO_2eq/MJ) due to fuel F combustion used in generator i;

Regarding the carbon footprint of grid components, in general terms, they are due to construction and to energy losses during component use due to the Joule effect. The carbon footprint of such losses is already accounted for implicitly by the additional generation needed in the system (for which we compute the kg of CO_2eq generated as presented above[4]). The carbon footprint of grid components (lines, storage, transformers) will therefore consider only the construction phase and dismantling, excluding losses. Moreover, a rough simplification will be made for the carbon footprint quantification of grid components (as discussed below, a detailed Life Cycle Assessment of each particular new grid component is out of the scope of the FlexPlan project).

Concerning new lines, the carbon footprint will consider materials for conductors, insulators and pylons. Data will refer to a generic HV line, and will be taken from existing literature (and in particular from Ecoinvent database [16]):

$$CO_2 eq_L = \sum_{i=1}^{n} km_i * \frac{CO_2 eq_{THVL}}{a}$$

where:

- $CO_2 eq_L$ is the carbon footprint of candidate new lines (in kg of CO_2eq/year);
- km_i is the length of new line i in the candidate options (in km);
- $CO_2 eq_{THVL}$ is the total CO_2eq emission for a generic HV line construction (in kg of CO_2eq/km);
- a is the estimated line lifetime (in years).

---

[4] For the assets which are not modelled as lossless

Regarding storage, for the carbon footprint calculation, Li-ion stationary batteries will be considered. For the sake of simplicity, only Lithium Ferrous Phosphate (LFP) batteries will be considered in modelling. In fact, they are considered as the most promising technology for the near future scenarios [17]. The carbon footprint of candidate storage can be calculated as

$$CO_2eq_S = \sum_S kWh_S * \frac{CO_2eq_{TS}}{a}$$

where:

- CO_2eq_S is the carbon footprint of candidate storage capacity (in kg of CO_2eq/year);
- kWh_S is installed capacity of the storage system S (in kWh);
- CO_2eq_{TS} is the total CO_2eq emission for a LFP battery construction (in kg of CO_2eq/kWh);
- a is the estimated battery lifetime (in years);

Regarding other flexibility options, the analysis will be limited to HVDC converters and to phase shifting transformers. Despite the complexity and the variability of this equipment, the carbon footprint of PSTs and HVDC converters can be attributed mainly to the transformer [18]. For this reason, for the FlexPlan carbon footprint modeling both devices will be considered as generic HV transformers. Hence the carbon footprint of an HVDC converter can be modeled as

$$CO_2eq_{HVDCc} = \sum_{HVDC} MVA_{HVDC} * \frac{CO_2eq_T}{a}$$

where:

- CO_2eq_{HVDCc} is the carbon footprint of candidate installed HVDC converters (in kg of CO_2eq/year);
- MW_{HVDC} is the installed power of each HVDC converter (in MW);
- CO_2eq_T is the greenhouse gas emission due to 1MW of a generic HV converter production (in kg of CO_2eq/MW);
- a is the estimated transformer lifetime (in years);

Similarly, for Phase Shifter Transformers (PSTs) the carbon footprint of the candidate options can be expressed as:

$$CO_2eq_{PST} = \sum_{PST} MVA_{PST} * \frac{CO_2eq_T}{a}$$

where:

- CO_2eq_{PST} is the carbon footprint of candidate installed PSTs (in kg of CO_2eq/year);
- MVA_{PST} is the installed power of each PST (in MVA);
- CO_2eq_T is the greenhouse gasses emissions due to 1MVA of transformer production (in kg of CO_2eq/MVA);
- a is the estimated PST lifetime (in years).

6.2.1 Results of candidate carbon footprints

The most common LCA database, Ecoinvent [16] does not detail electrical grid components to the required level of this project. While a dataset for overhead lines is available, there are no specific inventories for underground cables or transformers. Thus, for assessing the carbon footprint of the candidates and to solve the total carbon footprint equation, a literature search is required. The literature search was based on a keyword-based search on Scopus and Web of Science, and it resulted in a limited number of scientific papers on the topic. The search string results are: 5 pertinent results for transformers, 4 for underground cables (see Table 2).

Literature results	Authors	Titles
Transformers	Manshila et al. 2018 [19]	Life cycle assessment of electrical distribution transformers: Comparative study between aluminium and copper coils
	Mouhamad and Lauzevis 2013 [20]	Life cycle assessment and inrush currents measurement of amorphous transformers
	Santos Jorge et al. 2012 [21]	Life cycle assessment of electricity transmission and distribution-part 2: Transformers and substation equipment
	Borghetto et al. 2009 [22]	Comparative life cycle assessment of a MV/LV transformer with an amorphous metal core and a MV/LV transformer with a grain oriented magnetic silicon steel core
	Berti et al. 2009 [23]	Comparison of the ecoprofiles of superconducting and conventional 25 MVA transformers using the life cycle assessment methodology
Lines and cables	Jones and McManus 2010 [24]	Life-cycle assessment of 11 kV electrical overhead lines and underground cables
	Bumby et al. 2010 [25]	Life cycle assessment of overhead and underground primary power distribution
	Bumby et al 2009 [26]	Comparative Life Cycle Assessment (LCA) of overhead and underground medium voltage power distribution
	Santos Jorge et al. 2012 [27]	Life cycle assessment of electricity transmission and distribution-part 1: power lines and cables

Table 2 - Result of literature review for carbon footprint computation of candidate assets

The studies have been sorted through to find relevant properties of the components and to harmonize the results where enough information is available. The studies by Santos Jorge et al. [27] [21] have been identified as the most promising source for the current analysis, since they cover the widest spectrum of grid components among the literature, namely 7 types of lines and cables, 9 transformers of different size and various substation equipment (switchgears, disconnectors, circuit brakers, etc. …). Using the same source for various components allows to have homogeneous datasets and system boundaries, data with the same level of detail, thus removing inconsistencies when comparing different candidates.

The selected studies have been analysed in order to harmonize the results to the need of the current approach. Firstly, the system boundary has been analysed. All studies include the energy losses due to transmission in lines and cables and due to conversion for transformers in their result. These losses represent a relevant part of the total impact of the components and are highly dependent on the energy mix. In the present study, however, only the construction, installation and dismantling of lines and transformers is required.

The impact due to energy losses are subtracted from the results. The system boundaries of overhead lines, underground cables and transformers are depicted in Figure 6-6 and Figure 6-7 based on the study of Santos Jorge et al. 2012, from which operational losses have been excluded.
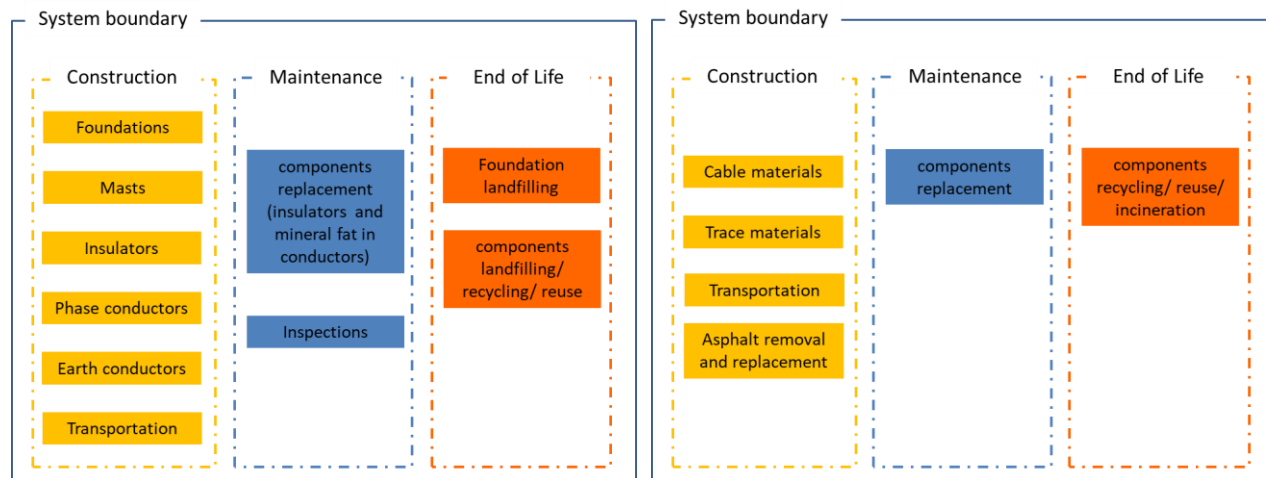
Figure 6-6 -System boundary of overhead lines (left) and underground cables (right), with losses excluded, based on Santos Jorge et al. 2012.
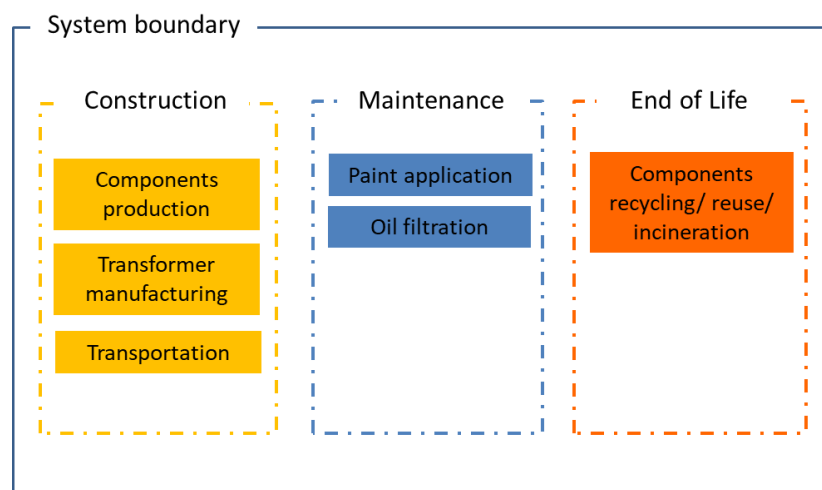
Figure 6-7- System boundary of transformers, with losses excluded, based on Santos Jorge

et al. 2012.

The results by [27] are expressed in kg CO_2eq/km for lines and cables, and in kg CO_2eq/equipment for transformers. After removing the impact due to losses, these results have been normalized to kg CO_2eq/(km*year) and kg CO_2eq/(MVA*year) for lines and transformers respectively. In Table 3 and Table 4 the results of Santos Jorge et al. 2012 are reported, along with the properties of the components and the normalized results.

Component Type	Impacts excluding losses		Functional unit	Characteristics of the component	Normalized impacts [kg CO$_2$/(km*year] (losses excluded)
underground cables 150 kV AC	140	ton CO$_2$eq/km	1 km operating for 40 y	150 kV AC oil land cable cross section 310 mm^2 conductor material: Cu	3500
underground cables 400 kV DC	130	ton CO$_2$eq/km	1 km operating for 40 y	400 kV HVDC land cable cross section 800 mm^2 conductor material: Cu	3250
overhead lines 400 kV AC	260	ton CO$_2$ eq/km	1 km operating for 40 y	400 kV overhead lines; Cross section 772 mm^2; conductor material: FeAl	6500
overhead lines 150 kV AC	110	ton CO$_2$ eq/km	1 km operating for 40 y	150 kV overhead lines cross section 454 mm^2 conductor material: FeAl	2750
overhead lines 350 kV DC	130	ton CO$_2$ eq/km	1 km operating for 40 y	350 kV HVDC overhead lines; cross section 772 mm^2 conductor material: FeAl	3250

Table 3 - Carbon footprint of underground tables and overhead lines from Santos Jorge et al 2012, with losses excluded and normalized to kg CO2eq/(km*year).

Transformer Type	Impacts excluding losses		Functional unit	Characteristics of the component	Normalized impacts [kg CO$_2$/(MVA*year)] (losses excluded)
Distribution transformer 315 kVA	0.27	kton CO$_2$ eq/component	1 equipment operating during its lifetime at 50% average load condition (30 y)	315 kVA, expected lifetime 30 y; converts power of 11-22 or 33 kV to 250 or 422 V	857
Large distribution transformer 9.6 MVA	4.61	kton CO$_2$ eq/component	1 equipment operating during its lifetime at 50% average load condition (30 y)	9.6 MVA, expected lifetime 30 y; transforms power to 6.3 kV	480
Large distribution transformer 16 MVA	6.2	kton CO$_2$ eq/component	1 equipment operating during its lifetime at 50% average load condition (30 y)	16 MVA, expected lifetime 30 y; converts power to 22 kV	517
Large distribution transformer 20 MVA	8.5	kton CO$_2$ eq/component	1 equipment operating during its lifetime at 50% average load condition (30 y)	20 MVA, expected lifetime 30 y; converts power to 22 kV	425
Power transformer 40 MVA	16.22	kton CO$_2$ eq/component	1 equipment operating during its lifetime at 50% average load condition (35 y)	40 MVA, expected lifetime 35 y	232

Power transformer 50 MVA	21.9	kton CO_2 eq/component	1 equipment operating during its lifetime at 50% average load condition (35 y)	50 MVA, expected lifetime 35 y	250
Power transformer 63 MVA	23.86	kton CO_2 eq/component	1 equipment operating during its lifetime at 50% average load condition (35 y)	63 MVA, expected lifetime 35 y	216
Power transformer 250 MVA	51.62	kton CO_2 eq/component	1 equipment operating during its lifetime at 50% average load condition (35 y)	250 MVA, expected lifetime 35 y	118
Power transformer 500 MVA	88.23	kton CO_2 eq/component	1 equipment operating during its lifetime at 50% average load condition (35 y)	500 MVA, expected lifetime 35 y	101

Table 4 - Carbon footprint of transformers from Santos Jorge et al 2012, with losses excluded and normalized to kg CO2eq/(MVA*year).

For batteries, data are derived from the RSE project [28], where LCA of LFP batteries has been considered. The carbon footprint expressed in terms of kg CO_2 per kWh of capacity are 114 kg CO_2 eq/kWh. This result includes impact deriving from extraction and manufacturing of raw materials, battery manufacturing and transportation, maintenance, and end of life.

6.3 Landscape impact modelling

Figure *6-8* shows the workflow which will be used to assess the landscape impact related costs of investment candidates. The set of candidates obtained by the FlexPlan pre-processor will be fed-in to the optimal routing algorithm described in the remainder of this section. The optimal routing algorithm determines the minimum cost of the candidates, taking the spatial properties of the installation into account. This helps to quantify the landscape impact on one hand, and to determine the exact technical details, such as impedance and level of partial undergrounding, especially for candidate line and cable connections for both AC and DC technology. Eventually, the cost of candidates and their technical parameters are fed into the planning model described in the previous sections. The detailed methodology is described in [29].
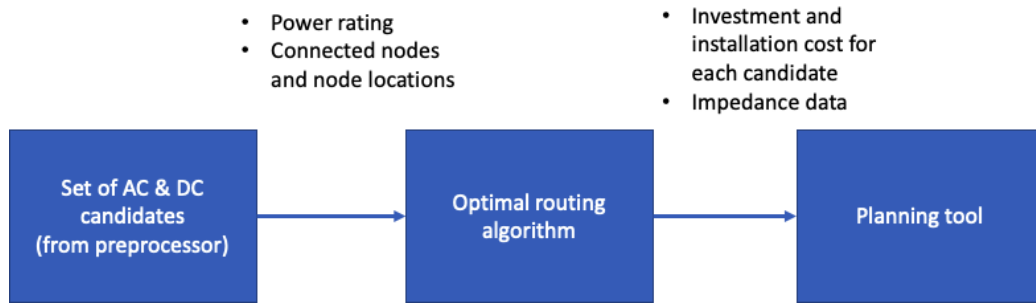
Figure 6-8 - Workflow of optimal routing algorithm
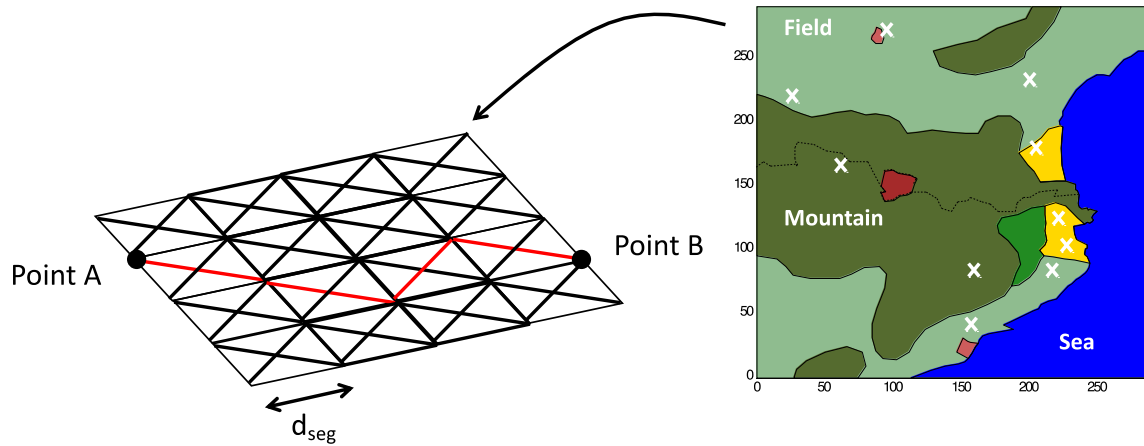
Figure 6-9 - Discretization of a spatial map to weighted graph

In the first step of the optimal routing algorithm, a spatial map of the installation area is discretized into a weighted graph as shown in Figure *6-9*. d_{seg} corresponds to the geographical resolution that has been chosen. Each node of the weighted graph is connected via a number of horizontal and vertical edges with the distance d_{seg}, as well as diagonal connections with the length of $\sqrt{2d_{seg}}$.

Thus, by assigning appropriate weights to the edges, an optimal routing algorithm can find the shortest "weighted" distance between two points A and B as indicated in Figure *6-9*. If the weights are chosen to be the cost of the transmission equipment including the landscape related costs, the shortest path obtained by the algorithm equals to the total investment cost of the transmission asset.

As the FlexPlan planning model considers both AC and DC transmission candidates, the discretization will be done separately for both technologies. For new transmission connections, both overhead line and underground cable systems can be used. In order to find the best combination of overhead line and underground cable sections for a given connection, two separate weighted graphs are generated, whose edge weights have different values depending on the combination of technology and spatial area located (Figure *6-10*). These two weighted graphs are connected at each node (x, y) with an additional edge, reflecting the costs required to switch from overhead line connections to cables and vice versa, as depicted for three nodes Figure *6-10*.
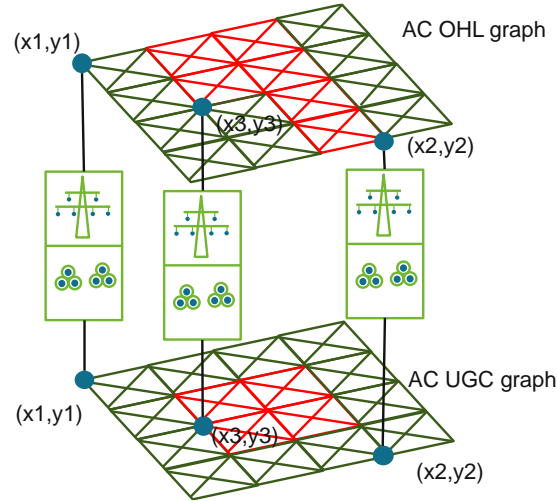
Figure 6-10 – Connection of graphs for OHL and UGC sections

This way the optimal routing algorithm finds the "shortest" or least cost path being able to switch between the technologies at any possible point. The geographic distance of the technology switch edges is defined as $d_{seg}^{tsw} = 0$. The edge weight ω_i for any edge i of the connected graph is defined as follows,

$$\omega_i = [C^{eq}(\alpha_i, \tau^{tech}) + C^{inst,ref}(\tau^{tech}) \cdot w(\alpha_i)] \cdot d_{seg}(i) + w^{switch}(i),$$

where $C^{eq}(\alpha_i, \tau^{tech})$ is the cost related to the installation area α and technology $\tau^{tech} \in \{OHL, UGC\}$ of the equipment. For instance, the equipment cost of a land cable is different than the cost of a submarine cable. Thus, depending on which area the edge i is located in, a different equipment cost is assigned to i. $C^{inst,ref}(\tau^{tech})$ is the technology dependent installation cost of overhead lines and underground cables for a reference area, e.g., installation in an open field. Depending on the area α_i where the edge i is located, a specific cost multiplier $w(\alpha_i)$ is multiplied to the installation cost, in order to account for the landscape impact. Both $C^{eq}(\alpha_i, \tau^{tech})$ and $C^{inst,ref}(\tau^{tech})$ are given in €/km. By multiplying these costs with $d_{seg}(i)$ of the edge i, which is dependent on the chosen spatial resolution, the total cost of equipment and installation on a particular edge i is obtained. For edges representing a technology switch, $w^{switch}(i)$ is added to the weight function. As d_{seg}^{tsw} for technology switch is defined to be zero, only the cost of the technology switch remains for these edges.

Once the graph weights are all assigned, the shortest, thus least cost path is obtained by applying off the shelf shortest path algorithms such as the Dijkstra or A* algorithms [29]. The Dijkstra algorithm is fully deterministic and guarantees optimality [30], whereas the A* algorithm uses a heuristic approach, while still

guaranteeing optimality, if the graph edge weights chosen are smaller or equal to the Euclidian distances between the vertices [31].

7 Monte Carlo scenario reduction

7.1 Background

The objective of the network expansion planning tool is to find the optimal network expansion measures that will allow the network to operate reliably for a range of uncertain future conditions, spanning several decades (2030-2040-2050). These future conditions are characterized by several long-term visions, describing possible developments of the energy system as well as divergent European energy policies.

Uncertainties in the planning problem are introduced by the presence of renewable generation resources, temperature-dependent loads and hydro-condition dependent storage and production. With the robust approach in mind, the expanded network found as solution by the planning problem must be able to supply the demand in most, if not all, possible situations. The approach chosen within FlexPlan is to provide a representative set of inputs, referred to as Monte Carlo years (MC years), to the planning tool to make sure that the solution is calculated based on a representative set of possible uncertain input scenarios, mainly characterized by the weather conditions.

However, due to the fact that the FlexPlan tool aims at covering a time horizon of multiple decades, including many technologies and spanning a large geographical area, solving this problem is a computationally demanding task. Therefore, the chosen set of MC-years must be limited as to allow for a computationally tractable planning problem. However, it is difficult to determine which scenarios are relevant a priori.

Before renewable generation became a substantial part of grid operations, it was often considered sufficient to evaluate new transmission lines only for the hour of the year with the highest and lowest loads. Any transmission network that allowed for adequate operation during this hour was considered likely to operate at least as well during any other demand scenario. This "worst hour" approach to selecting a test scenario is still deceptively appealing, but in modern systems, it is not obvious which hour is most likely to cause issues that severely threaten the reliability and/or security of supply. Also, the worst hour may vary by region or change as new lines are added and reliability issues are resolved. The inclusion of storage (and other flexible technologies) into the transmission expansion plan, require representative time series as input scenarios to the planning problem to enable a correct modeling of the intertemporal constraints of the flexible assets.

In this project clustering is chosen as methodology for scenario reduction, based on a literature review. The result of the scenario reduction has to be a limited set of representative timeseries to be used as input for the network expansion optimization problem. With clustering-based scenario reduction techniques, the aim is to group the overall space of possible scenarios into a (smaller) number of clusters, based on a metric that characterizes the scenario or its effect on the solution of the optimization problem. Then, one scenario from each cluster is selected and fed into the planning problem. Additionally, a probability can be assigned to each scenario to enable a stochastic approach as described in section 0.

The scenario reduction methodology is described in depth in deliverable D1.1 of the FlexPlan project [32]. A summary of the methodology is repeated below.

7.2 Methodology

Clustering is a machine learning technique that is used to find groups, i.e., clusters, of data objects within a dataset. Within the context of scenario reduction for transmission network expansion planning, the objective of the clustering is to find a limited, yet representative, set of input scenarios of the planning problem. A clustering algorithm detects clusters of 'similar scenarios' within the overall dataset of possible scenarios. It is then assumed that solving the planning problem for the set of input scenario's that consists of one scenario per cluster will lead to a solution that is the same as or close to the solution that would be calculated if the complete set of possible input scenarios would be used.

To use clustering, the data objects must have common features by which they are compared. The network state, or stated differently, the need for transmission capacity or flexibility, is defined by the load and/or renewable generation at each network node. Similar network states are thus defined by a similar combination of nodal load and generation. Both load and generation vary over time, as such, load/generation time series can be considered similar when a similar variation in the load/generation time series is observed. Within the network expansion planning problem, one data object consists of (all of) the time series of load and (renewable) generation at all of the nodes of the network. The data objects thus have features in two dimensions:

- Node dimension: the value of load/generation at each separate node, or the power of each demand/generation element can be considered as separate features. The number of node dimension features equals $n(S_g) + n([(S]_u)$.
- Time dimension: the node values at each separate time step of the time series can be considered as separate features. The number of time dimension features equals $n(S_t)$.

In total, one observation or one data object thus reaches $\left(n(S_g) + n([(S]_u)\right) \cdot n([(S]_t)$ features. Given the size of the considered networks, as well as a minimal required time series length, this number of features become very large, and feature reduction techniques should be used to make sure sensible clusters are produced from the dataset. Three feature reduction techniques have been proposed:

1. Principal component analysis (PCA). PCA is a mathematical technique that is used to transform the data while preserving as much variability as possible.
2. Feature selection by defining significant features: with this technique each scenario is characterized by a (small) number of significant features, and clustering is done based on the selected features alone.
3. Clustering the node dimension: with this method only the time-dimension features are used, while the node dimension is reduced.

More information on these techniques can be found in deliverable D1.1 [32]. The final choice of which technique to use will mainly depend on the actual dataset. Also, tests are performed to compare the different feature reduction techniques.

The dataset to be clustered, consists of 40 year-long time series of hourly time steps (i.e., 8760 time steps) for each network node. The result of the clustering are y time series (of length $n[\![(S]\!]_t)$, with $n[\![(S]\!]_t)$ equal to 8760 or less) of load and generation power values. These y time series are chosen as members of each cluster. As clustering algorithm, K-means clustering was chosen. The size of each cluster is an indication of the probability of occurrence of that combination of load and generation present in that cluster. These probabilities can also be provided as input to the planning optimization tool, e.g., if a stochastic planning approach is envisaged.

The overall clustering-based scenario reduction methodology can be outlined as follows:

1. (if required) Split the yearly scenarios into scenarios with a (time-dimension) length equal to the setting required by the planning tool. The length is set to 24 timesteps if representative daily scenarios are required, to 168 timesteps, if weekly scenarios are required, etc.

2. Normalize the load and generation time series as produced by the scenario generation, by the nominal capacity of the respective loads and generators.

3. (If required) apply feature reduction by (a) applying principal component analysis (PCA), (b) selecting and computing significant features, or (c) reducing the node dimension by clustering the node-dimension features.

4. Perform K-means clustering on the overall, but possibly feature-reduced, dataset. The requested number of clusters *s* is defined by the computational efficiency of the planning tool and is provided as an input to the scenario reduction.

5. Choose one representative scenario for each cluster. The computed cluster centroid is an option, but also a randomly chosen scenario belonging to that cluster can be used. If feature reduction was applied, choosing the cluster centroid is not an option, as the centroid in this case will not have the same dimensions as the original scenarios.

6. Rescale the chosen scenarios back to their original values.

7. (if required) Identify the importance of the representative scenario's according to their cluster size.

As stated, more information on the scenario reduction approach, and an in-depth discussion on each step outlined above, can be found in deliverable D1.1 of the FlexPlan project [32].

8 Further modelling decisions based on implementation challenges and computational efficiency

This document outlines the general workflow of the FlexPlan planning tool. The flexible planning model which is the core of the planning tool has been defined in a generic way to capture a variety of grid investments for radial and meshed grids, storage and demand flexibility models. The developed model includes grid constraints, a reliability model and models to assess the environmental impact. Combining all aspect results in a large scale linear mixed-integer problem, which needs to be implemented and solved efficiently for large systems as envisaged in the regional case analysis of FlexPlan.

A number of techniques have been identified in order to solve the problem more efficiently. Based on the computational performance of the implemented model and tests on small test cases, a combination of these techniques can be used in order to further increase performance if proven necessary.

(1) Reduction of the number of timesteps considered in the problem: Using the Monte Carlo sampling and clustering approaches briefly introduced in the previous section of this document, the number of considered time steps can be reduced. By doing so, representative days/ weeks / months of the planning problem can be identified capturing a variety of grid states. By solving the model only for this reduced set of hours instead of the full planning year, reduces the problem size accordingly. Nevertheless, before applying such a reduction, the accuracy with respect to temporal modelling of especially storage and demand flexibility as well as the gain in computation time needs to be validated against the full -year representation on small scale test networks.

(2) Decomposition of the planning and operational model: There exist a number of applications of Benders decomposition to solve the planning problem and a number of operational problems in a sequential and iterative way. This can achieve computational gains, as a number of smaller problems can often be solved more efficiently than a single large optimisation problem. If proven necessary, such a decomposition can be applied to the FlexPlan model, where its efficiency can be tested on small test networks.

(3) In a similar way, decomposition can be applied in the form of reliability cuts. The current formulation of the reliability model relies on the solution of the network equations for each possible contingency of a defined set, each time point of the planning year, and each year of the planning horizon. By applying reliability cuts, the reliability modelling can be solved as a collection of subproblems. If necessary, such a decomposition can be tested on small test networks and its efficiency can be verified.

(4) Decomposition of network levels: Solving the optimal investment problem for integrated transmission and distribution networks can be computationally demanding or even intractable. To overcome this challenge and improve the computational trackability of the problem, a decomposition of radial and meshed networks is suggested, which considers investments into radial networks as a possible extension candidate of the meshed network. As there is a large number of systems and feeders

operated radially, such an approach can be applied to reduce the size of the planning problem and solve the optimisation models for the radial and meshed networks sequentially. As the FlexPlan model can be generally applied to meshed and radial networks alike, the performance of such a decomposition can be assessed on a small test network.

(5) Implementing so-called hierarchic models which decompose the optimization problem into sequential optimization problems having different time resolutions that are typical of the different entities making up the system. A typical hierarchic decomposition is between two separate optimization problems: the first, a yearly hydro optimization problem characterized by a weekly granularity, sets the initial and final conditions for 52 decoupled hourly problems calculating the dispatch for each week of the year. This kind of decomposition is appealing in case of huge problems which can't be numerically treated differently, but the process introduces an important set of artificial rigidities which could prevent to obtain the optimal solution.

In particular, among these five techniques, <u>two of them: (2) and (4) are identified as the most promising and will be implemented and tested in priority</u>. Approach (1) and (5) (reducing the time resolution or implementing hierarchic models) would result in a reduction of the model precision, whereas decomposition techniques allow to preserve details. In particular, approach (5) would allow to represent accurately most hydro plant but probably not seasonal ones (like in Scandinavia) and additionally, it would introduce limitations in the range of actions of what we called storage "arbitrage" between hours (see Section 5.2).

The theorical concepts behind techniques (2) and (4) are described in the two following sections of this chapter.

8.1.1 Benders decomposition

The mathematical model described in the previous sections of this document consists in a Mixed-Integer Linear Programming model. If applied to the planning of a realistic national or multi-national (regional) power system, with a network consisting of hundreds of interconnections and nodes, to which a high number of generation and consumption units of different technologies are connected, the computational burden could be excessive, so that the model could prove uncapable to provide the optimal solution in a reasonable amount of time.

However, the structure of the problem (see in particular Section 3), with the integer variable dedicated to the decision on investments and the clear identification of a given reference year for each cost term, is well suited for decomposition techniques.

This structure is clearly evident giving a look at the objective function formulation (4), where the terms related to operational costs are clearly separated from the terms related to investment cost. Furthermore, the former ones depend on the hour, while the latter do not. It is then clear to which reference year y each cost is referred.

This happens also when a stochastic approach is considered, as in equation (16), where different scenarios s are considered altogether, each with a probability of occurrence π_s. Here, investment decisions are independent of scenario realizations; operational costs, on the contrary, vary depending on the scenario s.

y_1			y_2			y_N		
Investment Costs			**Investment Costs**			**Investment Costs**		
Operational Costs s_1, t_1	Operational Costs s_1, t_2	Operational Costs s_1, t_T	Operational Costs s_1, t_1	Operational Costs s_1, t_2	Operational Costs s_1, t_T	Operational Costs s_1, t_1	Operational Costs s_1, t_2	Operational Costs s_1, t_T
Operational Costs s_2, t_1	Operational Costs s_2, t_2	Operational Costs s_2, t_T	Operational Costs s_2, t_1	Operational Costs s_2, t_2	Operational Costs s_2, t_T	Operational Costs s_2, t_1	Operational Costs s_2, t_2	Operational Costs s_2, t_T
Operational Costs s_S, t_1	Operational Costs s_S, t_2	Operational Costs s_S, t_T	Operational Costs s_S, t_1	Operational Costs s_S, t_2	Operational Costs s_S, t_T	Operational Costs s_S, t_1	Operational Costs s_S, t_2	Operational Costs s_S, t_T

Figure 8-1 - Graphical representation of dependence of variables

This sort of "block structure" is represented in Figure *8-1***Figure 8-1**. Investment costs, as highlighted above, are obtained only by considering integer variables, while some operational costs depend on integer variables (since operational costs are present only if a device has been installed). Furthermore, operational costs for different scenarios and for different years are independent.

Thus, one can try to solve separately the problems related to investment decision for each year and the problems related to operational costs for each year and scenario. Obviously, the separation is not complete, since integer variables expressing the decision on investment influence also the operational costs for the considered year. Therefore, a recursive procedure can be set up, in which the decision on the investments is taken by taking into account the operational costs and the decision on the operation is taken given the

y_1		
Investment Costs		
Operational Costs s_1, t_1	Operational Costs s_1, t_2	Operational Costs s_1, t_T
Operational Costs s_2, t_1	Operational Costs s_2, t_2	Operational Costs s_2, t_T
Operational Costs s_S, t_1	Operational Costs s_S, t_2	Operational Costs s_S, t_T
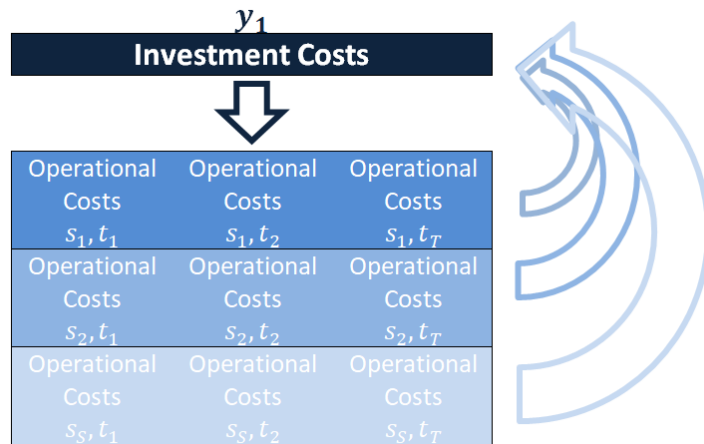
Figure 8-2 - Recursive procedure representation

investment decision, as depicted in Figure *8-2*.

Investment decisions referred to different years should however be taken altogether, firstly because an investment could be postponed or anticipated from one year to another depending on the operation of the

system and then also because if an investment takes place in one year it may have effects also in the following years.

The problem related to the decision on integer variables constitutes the upper level problem, while the decisions on operation in dependence of the year and the scenario realization are taken in the lower level problems. Lower level problems are solved with the values of the decision variable of the upper level problem considered as parameters. Upper level decision are taken taking into account the results of the lower level problems by means of Lagrange multipliers of some particular constraints, called "cuts".

Decomposition techniques with the characteristics described here above are called Benders Decomposition Techniques [33]. This section will not provide the theoretical background of these techniques; in their application to the optimization model presented in the previous sections only Benders optimality cuts are considered, in a way similar to what has been performed in [34].

8.1.1.1 Upper level problem

As mentioned above, the upper level problem aims at taking an optimal decision on the investment considering only the investment costs, which are minimized, as shown in the objective function (24)

$$
\begin{aligned}
\min \psi^{down} = \sum_{y \in S_y} \Bigg\{ f_y^d \Bigg[&\sum_{jc \in S_{jc}} \alpha_{jc,y} \big(I_{jc,y}^E (E_j^{\max}\ ) + I_{jc,y}^P (P_j^{\max}\ ) + FP_{jc,y}^{CO_2} \big) \\
&+ \sum_{u \in S_u} \alpha_{u,y} \big(I_{u,y} + FP_{u,y}^{CO_2} \big) + \sum_{lc \in S_{lc}^{ac}} \alpha_{lc,y} \big(I_{lc,y} + FP_{lc,y}^{CO_2} + LS_{lc,y} \big) \\
&+ \sum_{dc \in S_{lc}^{dc}} \alpha_{dc,y} \big(I_{dc,y} + FP_{dc,y}^{CO_2} + LS_{dc,y} \big) \\
&+ \sum_{zc \in S_{zc}} \alpha_{zc,y} \big(I_{zc,y} + FP_{zc,y}^{CO_2} + LS_{zc,y} \big) \\
&+ \sum_{bc \in S_{bc}} \alpha_{bc,y} \big(I_{bc,y} + FP_{bc,y}^{CO_2} + LS_{bc,y} \big) \Bigg] \Bigg\} + \sum_s \pi_s \gamma_s
\end{aligned}
\tag{24}
$$

The effects of the operation, that is of the lower level problems, are considered by means of the Lagrange multipliers $\phi_{*,y,s}$ of their optimality cuts with the constraint (25) which defines the term γ_s in (24)

$$
\begin{aligned}
\gamma_s \geq \sum_y \Bigg\{ &\psi_{y,s} + \sum_{j \in S_{jc}} \phi_{jc,y,s}^{(j)} \big(\alpha_{jc,y} - \alpha_{jc,y}^{(j)} \big) + \sum_{u \in S_u} \phi_{u,y,s}^{(j)} \big(\alpha_{u,y} - \alpha_{u,y}^{(j)} \big) \\
&+ \sum_{lc \in S_{lc}^{ac}} \phi_{lc,y,s}^{(j)} \big(\alpha_{lc,y} - \alpha_{lc,y}^{(j)} \big) + \sum_{dc \in S_{lc}^{dc}} \phi_{dc,y,s}^{(j)} \big(\alpha_{dc,y} - \alpha_{dc,y}^{(j)} \big) \\
&+ \sum_{zc \in S_{zc}} \phi_{zc,y,s}^{(j)} \big(\alpha_{zc,y} - \alpha_{zc,y}^{(j)} \big) + \sum_{bc \in S_{bc}} \phi_{bc,y,s}^{(j)} \big(\alpha_{bc,y} - \alpha_{bc,y}^{(j)} \big) \Bigg\}
\end{aligned}
\tag{25}
$$

where the superscript (j) indicates the values calculated in the previous step of the iteration, i.e. $j = i - 1$.

The upper level problem includes also all the original constraints on the integer variables $\alpha_{*,y}$ and the following constraint

$$\gamma_s \geq \gamma^{down} \tag{26}$$

8.1.1.2 Lower level problems

Lower level problems represent the cost minimization (27) for the operation of the system for a given year y and a given scenario s under all the operation constraints presented in the previous sections 4 and 5 and the optimality cuts (28).

$$
\begin{aligned}
\min \psi_{y,s} = f_y^{d,o} \Bigg\{ \sum_{t \in S_t} \Bigg[&\sum_{g \in S_g} \left[C_{g,y}^{aq} + (\theta^{CO_2} G^{pf} + \theta^f) \eta_g^f \right] P_{g,t,y,s} + C_{g,y}^{res,curt} \Delta P_{g,t,y,s}^{res} \\
&+ \sum_{j \in S_j} \left[C_{j,t,y}^{abs} P_{j,t,y,s}^{abs} + C_{j,t,y}^{inj} P_{j,t,y,s}^{inj} \right] + \sum_{j \in S_{jc}} \left[C_{jc,t,y}^{abs} P_{jc,t,y,s}^{abs} + C_{jc,t,y}^{inj} P_{j,ct,y,s}^{inj} \right] \\
&+ \sum_{u \in S_u} \left[C_{u,t,y}^{nce} \left(P_{u,t,y,s}^{ref} - P_{u,t,y,s}^{nce} \right) + C_{u,t,y}^{ds} \left(\Delta P_{u,t,y,s}^{ds,up} + \Delta P_{u,t,y,s}^{ds,dn} \right) + C_{u,t,y}^{lc} \Delta P_{u,t,y}^{lc} \right] \\
&+ \sum_{n \in S_n} \left(C_{n,t,y}^{EE} EE_{n,t,y,s} + C_{n,t,y}^{LL} LL_{n,t,y,s} \right) \Bigg] \Bigg\}
\end{aligned} \tag{27}
$$

$$
\begin{aligned}
\alpha_{jc,y} &= \alpha_{jc,y}^{(i)} : \phi_{jc,y,s} \\
\alpha_{u,y} &= \alpha_{u,y}^{(i)} : \phi_{u,y,s} \\
\alpha_{lc,y} &= \alpha_{lc,y}^{(i)} : \phi_{lc,y,s} \\
\alpha_{dc,y} &= \alpha_{dc,y}^{(i)} : \phi_{dc,y,s} \\
\alpha_{zc,y} &= \alpha_{zc,y}^{(i)} : \phi_{zc,y,s} \\
\alpha_{bc,y} &= \alpha_{bc,y}^{(i)} : \phi_{bc,y,s}
\end{aligned} \tag{28}
$$

The symbols $\phi_{*,y,s}$ in (28) refers to the Lagrange multipliers of those constraints.

Since the optimality of the iterative process is granted only if the lower level problems are convex, all the integer variables that are not referred to the investment decisions of the upper level problem must be relaxed to continuous variables.

8.1.1.3 The iterative process

The value of the variable ψ^{down} in the objective function (24) of the upper level problem represents a lower bound to the optimal solution of the original problem.

An upper bound to this optimal solution can be defined by means of

$$\psi^{up} = \psi^{down} - \sum_s \pi_s \gamma_s + \sum_s \pi_s \sum_y \psi_{y,s} \tag{29}$$

Thus, it is possible to define an iterative process solving alternatively the upper level problems and the lower level problems till the convergence of ψ^{up} and ψ^{down}, thus to optimality. In practice, this cannot be achieved; instead, an optimality gap ϵ has to be provided, so that the process can stop if

$$\left| \frac{\psi^{up(i)} - \psi^{down(i)}}{\psi^{down(i)}} \right| \le \epsilon \tag{30}$$

The iterative process can be set up as follows:

1. set $\psi^{up(0)} = +\infty$, $\psi^{down(0)} = -\infty$ and arbitrary values for all the other parameters (for instance, set all to 0);

2. set $i = 1$;

3. solve the upper level problem and set $\alpha_{*,y}^{(i)} = \alpha_{*,y}$ and $\psi^{down(i)} = \psi^{down}$

4. solve the lower level problems and set $\phi_{*,y,s}^{(i)} = \phi_{*,y,s}$;

5. compute $\psi^{up(i)}$ by means of (29);

6. evaluate the stop condition (30):

      i.   if it is satisfied, then go to step 7;

      ii.   if it is not satisfied set $i = i + 1$ and go to step 3;

7. if some integer variables have been relaxed, solve the lower level problems eliminating the continuous relaxation and fixing $\alpha_{*,y}^{(i)}$; provide the results and exit.

8.1.2 Combined modelling of transmission and distribution systems

The planning tool specified in this report is in principle equally applicable to 1) transmission system planning, 2) distribution system planning and 3) integrated transmission and distribution system planning. In other words, the two linear mathematical models (one for meshed and one for radial grids) can be merged to represent both distribution and transmission systems alike. The difference between the three "use cases" is the network model used as input. In principle, the tool can also be applied to integrated transmission and distribution system planning (use case 3) by using the transmission system plus distribution system models as input. However, for large power system models as those expected for the FlexPlan regional case studies, the size of these network models could make this application of the model computationally intractable. This is most likely the case also when restricting the distribution system definition in the model to only the highest voltage level of the radially operated system, as described in earlier sections. An approach to overcome this problem in case it materializes in the application to real cases is therefore outlined in this section. To first give an overview, one can distinguish between the following four use cases of the planning tool:

1) Transmission system planning
   o Only the transmission system is included in the network model.
   o Distribution systems are represented in a simplified manner, with distribution-level load (or generation) and flexibility aggregated to the connecting node in the transmission system.

2) Distribution system planning
   o Only the distribution system is included in the network model.

- The transmission system is represented in a simplified manner as a generator at the node (swing or slack bus) connecting to the distribution system. As a simplifying assumption, we assume that the voltage on the HV side of the transformer does not change with the loading of the transformer, and as such the transformer is connected directly to the slack bus.

3) Integrated transmission and distribution system planning
    - Both the transmission and distribution system[5] are included in the network model.
    - Could possibly only be considered for relatively small power system models and would be computationally intractable for larger networks.

4) Integrated transmission and distribution system planning (coupling approach)
    - Step 1: The planning tool is applied to each distribution system (as in use case 2), with modifications to be specified below, to establish the costs of re-planning the distribution system to increase the potential to provide flexibility to the transmission system.
    - Step 2: The planning tool is then applied using only the transmission system network model (as in use case 1) and with the re-planning of the distribution systems as flexibility investment candidates with the costs established in step 1.

Each use case, and in particular the proposed coupling approach of use case 4, will be explained in more detail below. Note that in terms of modelling, the approach for integrated transmission and distribution system planning (use case 3) will reduce to use case 1 if there is no difference between the mathematical representation of the transmission and distribution parts of the grid model (both systems operated with a meshed structure). For all use cases, it is however important that it is encoded in the network model which nodes are part of distribution systems, which nodes are parts of the transmission system, and which nodes define interfaces between the transmission system and a distribution system.

As the methodology outlined for use case 4 would be, to the best of our knowledge, a novel and untested approach to transmission and distribution system planning, its effectiveness and efficiency needs to be verified before applying it to large-scale power network models. After specifying and implementing the methodology, one can build up experience with its application by first considering small- and medium-scale case studies where it is possible to compare with the fully integrated transmission and distribution system planning (use case 3). This would also give insights into the potential benefits of integrated transmission and distribution system planning before considering use case 4 in applications to e.g., regional case studies. Before this point, use case 3 is the recommended application of the planning tool to large-scale case studies (e.g., the regional case studies).

Radial distribution systems, as understood in the context of the FlexPlan planning tool, differ from transmission (or high voltage distribution) systems in being radial or radially operated. However, one must ensure that this radiality condition is respected for all possible solutions of the medium voltage distribution system planning problem. This will be done at the pre-processor stage where the set of planning candidates

---

[5] Here distribution refers mainly to the first radially operated voltage level of the grid (see Section 4.3).

are selected. Here one simply needs to check that this candidate set does not include any distribution lines that would make the distribution network meshed. This implies that dynamic grid reconfiguration is not represented in the optimization model. In other words, meshed distribution networks that are radially operated are represented by the same radial distribution network for each hour of the year. (Including dynamic grid reconfiguration as a flexibility option would imply additional integer decision variables for each hour of the year.) It also implies that reserve connections between adjacent radials are not considered neither as candidate lines nor non-candidate (existing) lines. In turn, this implies that distribution system reliability analysis considering post-contingency network reconfiguration is left out of scope for the model. Possible redundancy in the distribution system, e.g., the usage of parallel HV/MV transformers is considered in the reliability analysis as long as the radiality assumption holds. The condition of radiality is most easily enforced by only including reinforcement of existing distribution lines and not the creation of new distribution lines in the pre-processing step selecting the distribution planning candidates.

Illustration of use case 1: Transmission system planning

Figure *8-3***Figure *8-3*** illustrates the application of the planning tool to transmission system planning (use case 1). For this use case, the distribution system is not explicitly represented in the network model, but it is represented in a simplified manner by an aggregated flexibility element at the transmission node interfacing with the distribution system (D+E). The method for aggregation is assumed to be provided by the pre-processing tool, which will elaborate the potential of each specific distribution network in providing flexibility services to the transmission system. It is assumed that this flexibility element can be represented using the generic flexibility models described in section 0. The transmission system planning candidates in this example are new transmission (A) and sub-transmission (B) lines and investment in a flexibility (storage) element (C). Possibly, also distribution flexibility elements (storage units D and E) can be considered candidates for transmission planning.
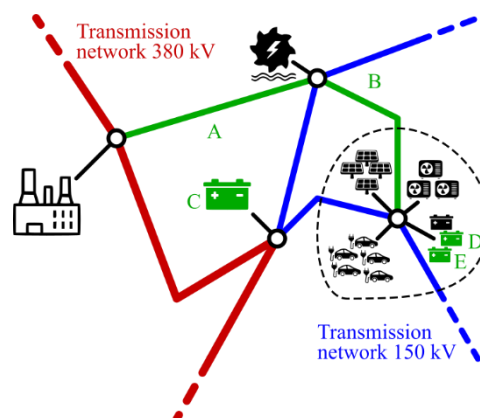
Figure 8-3 - Illustration of transmission system planning (use case 1).

Illustration of use case 2: Distribution system planning
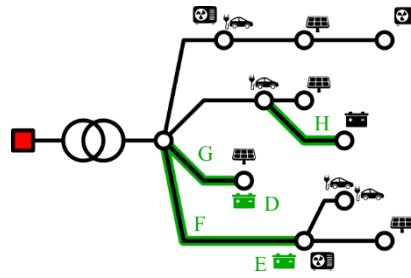
Figure 8-4 - Illustration of distribution system planning (use case 2).

Figure *8-4* illustrates the application of the planning tool to distribution system planning (use case 2). Here the transmission system is not explicitly included in the network model but is represented in a simplified manner by the interface (swing bus) node marked with a red square. The generation cost for the aggregated generator at the interface must be determined in a pre-processing step. The distribution system planning candidates in this example are the reinforcement of some distribution lines (F, G and H) and the investment in flexibility (storage) elements (D and E).

Illustration of use case 3: Integrated transmission and distribution system planning

Figure *8-5***Figure *8-5*** illustrates the application of the planning tool for integrated transmission and distribution system planning (use case 3). This use case is referred to as integrated approach since the full transmission and distribution system models are used as input in the optimization model, which for realistic cases could be computationally intractable. In this application of the tool, planning candidates A to C (in the transmission system) and D to H (in the distribution system) are all considered jointly in the optimization problem.
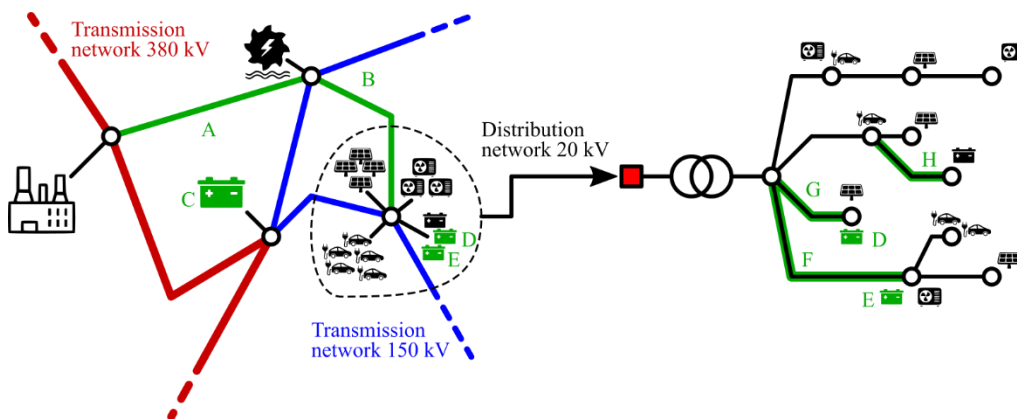
Figure 8-5 - Illustration of naïve approach to transmission and distribution system planning (use case 3).

Outline of methodology for coupling transmission and distribution system planning problem (use case 4)

This use case consists of running planning processes for transmission and distribution systems separately. With respect to use case 1 and 2, distribution planning is further investigated in order to support also transmission system planning (final step of the considered use case). In practice:

1. Different planning options are computed for each distribution system, beginning with the least expensive option with the lowest total socio-economic cost (see use case 2)

2. In addition, the procedure returns additional planning solutions aimed at maximising the exploitation of distribution system flexibility for transmission system planning too.

3. As (an) optional step(s), intermediate planning options for the distribution system can be computed that have costs lying between that of the least expensive option (step 1) and the option maximizing the distribution system flexibility made available to the transmission system (step 2).

4. The transmission system planning is executed as for use case 1, but the different solutions returned by the previous step are included within the optimization problem as alternative candidates (and each of those provides different flexibility options for the transmission network).

Step 1 – Definition of least expensive planning option for distribution system

The first step of the planning procedure investigates different options for distribution system planning, having considered the fact that distribution system resources and candidates can support the optimal planning of transmission system too. The first iteration of this procedure consists of solving the planning problem for the distribution system at the minimum cost (same problem of use case 2). In this case the optimization is formulated with the model of the considered distribution network and the planning costs are minimized according to the objective function described in section 3:

$$\min \sum_{s} \pi_s \left\{ \sum_{y \in S_y} \left\{ \sum_{t \in S_t} \left[\sum_i C_{t,i} + \sum_{j \in S_j} \alpha_j C_{t,j} + \sum_{n \in S_n} K_{t,n} \right] + \sum_{j \in S_j} \alpha_j I_j \right\} \right\}.$$

Thanks to this procedure, the solver returns a possible planning solution, capable of solving all the expected congestion issues at the minimum cost (opportunely weighting the scenario probability). Figure 8-6 reports the candidates (E and G) for the exemplificative distribution network which has been selected by the optimization algorithm.
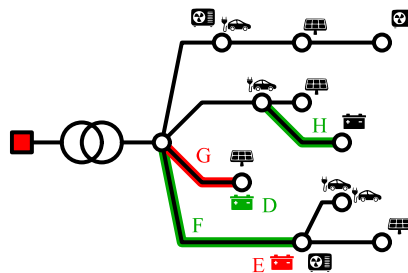
Figure 8-6 - Illustration of the optimized planning for distribution system

However, the amount of flexible power/energy that can be delivered to the transmission system (aimed at supporting its planning) might be limited since it is not considered in the objective function. By looking at the

selected candidates reported in Figure *8-6*, the solution at distribution level determines the remaining flexibility candidates available for transmission services (Figure *8-7*):

- Candidate E cannot be used for transmission services (since its dimensioning is optimized for distribution services).
- The optimal sizing of the line candidate G considers only the power flow necessities of the connected PV and the storage candidate D (if activated for transmission services) can be exploited at maximum 50% of its power capacity.
- The storage device connected at the end of the line candidate H can be exploited at maximum 70% of its capacity, because of grid bottlenecks.
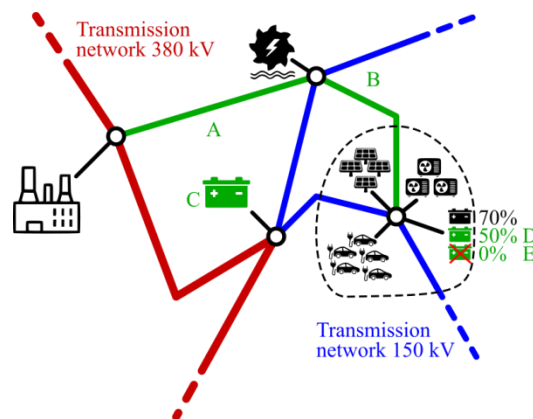
Figure 8-7 - Illustration of the remaining flexibility for transmission planning
(having optimized distribution planning for local congestions only)

The amount of distribution flexibility that can be exploited for transmission services/planning can be calculated by running two optimal power flow processes. The first one is aimed at calculating the maximum upward flexibility and features the following objective function:

$$\min \sum_{s} \pi_s \left\{ \sum_{y \in S_y} \sum_{t \in S_t} P_{l,t,y}^{fr} - \sum_{y \in S_y} \sum_{t \in S_t} \sum_{j \in S_j} P_{j,t,y}^{inj} - \sum_{y \in S_y} \sum_{t \in S_t} \sum_{j \in S_{jc}} P_{j,t,y}^{inj} \right\}$$

where l corresponds to the AC branch index of the transformer connecting the considered distribution network to the transmission system. In addition to $P_{l,t,y}^{fr}$, the (existing and candidate) storage power injection has been included, since their integral constraint does not impact the total energy exchange over the considered period of time. In this case, in order to calculate the flexibility contribution to transmission system also for energy units that have not been selected:

- All the line candidates selected by the planning engine are supposed to be in place.
- All the flexible energy candidates (selected or non-selected) are supposed to be in place.

The optimization process returns power profiles for all the flexible energy resources (existing and candidates) of the considered distribution network. By observing and analysing these power profiles, the maximum amount of exploitable flexibility for transmission services can be deduced for each energy resource.

The second process repeats the same calculation (with the same assumptions in terms of model) with a different objective function. In this case the downward flexibility is maximized by identifying the solution that optimizes:

$$\max \sum_s \pi_s \left\{ \sum_{y \in S_y} \sum_{t \in S_t} P_{l,t,y}^{fr} + \sum_{y \in S_y} \sum_{t \in S_t} \sum_{j \in S_j} P_{j,t,y}^{abs} + \sum_{y \in S_y} \sum_{t \in S_t} \sum_{j \in S_{jc}} P_{j,t,y}^{abs} \right\}$$

Step 2 – Definition of planning option for distribution system which guarantee the maximum flexibility for transmission services

Having assumed that distribution candidates and resources are requested to contribute to the transmission planning objective, another objective function for distribution planning consists of maximising the amount of flexible power/energy that can be exploited by the transmission system. For this purpose, the same optimization problem described above with all the candidates to be selected can be considered, but the objective function is modified as follows:

$$\min \sum_s \pi_s \left\{ \sum_{y \in S_y} \sum_{t \in S_t} P_{l,t,y}^{fr} - \sum_{y \in S_y} \sum_{t \in S_t} \sum_{j \in S_j} P_{j,t,y}^{inj} - \sum_{y \in S_y} \sum_{t \in S_t} \sum_{j \in S_{jc}} P_{j,t,y}^{inj} \right\}$$

where l corresponds to the AC branch index of the transformer connecting the considered distribution network to the transmission system. In practice, the solution of this optimization corresponds to the planning option for which the maximum upward power flexibility is guaranteed by the distribution resources to transmission network. The same optimization problem can be processed for the maximum downward power flexibility, by rewriting the objective function as follows:

$$\max \sum_s \pi_s \left\{ \sum_{y \in S_y} \sum_{t \in S_t} P_{l,t,y}^{fr} + \sum_{y \in S_y} \sum_{t \in S_t} \sum_{j \in S_j} P_{j,t,y}^{abs} + \sum_{y \in S_y} \sum_{t \in S_t} \sum_{j \in S_{jc}} P_{j,t,y}^{abs} \right\}$$

By merging the solutions of these two optimization processes, a possible result for the considered distribution network can be the one represented in Figure *8-8*.
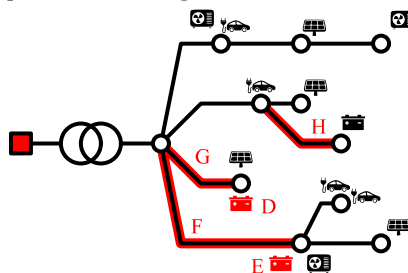
Figure 8-8 - Illustration of the optimized planning for distribution system

For this exemplificative case of distribution network planning (Figure *8-9*), larger volumes of transmission services can be guaranteed (Figure *8-10*). From the analysis of the resulting time series of flexible devices, it can be deduced that:

- The candidate F solves local congestions and allows the exploitation of the storage E (for transmission services) at 90% of its capacity.
- The candidate G (which is dimensioned for the power supply of both the connected PV and storage units) solves local congestions and allows the exploitation of the storage D (for transmission services) at 70% of its capacity.
- The candidate H allows the full exploitation of the existing storage connected to it, which was limited to 70% of its capacity.
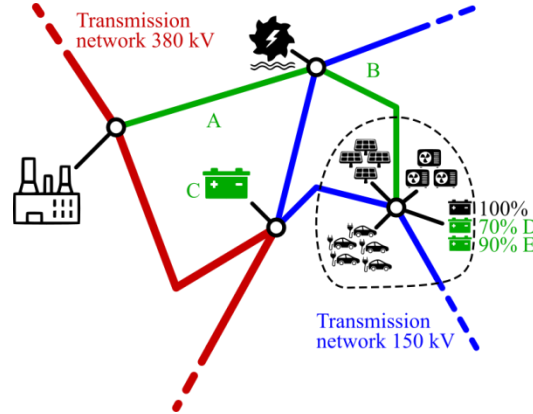
Figure 8-9 - Illustration of the remaining flexibility for transmission planning (having optimized distribution planning for the provision of flexibility services to transmission)

Step 3 – Definition of intermediate planning options for distribution system

The distribution planning returned by step 1, guarantees the minimization of the total cost which is the result of the objective function. From this function it is possible to extract the costs related to new lines investments and defined as:

$$C_{step1} = \sum_{j \in S_{lc}^{ac}} \alpha_j I_j$$

The same term can be calculated for the selected new-line candidates resulting from step 2, which returns C_{step2} (expected to be higher than C_{step1}). In case further planning options need to be investigated, the optimization problems proposed for step 2 can be reprocessed by adding the following constraint:

$$C_{step3} \geq \sum_{j \in S_{lc}^{ac}} \alpha_j I_j$$

where C_{step3} is a user-defined value within the interval $[C_{step1}, C_{step2}]$. In this case, an exemplificative result is reported in Figure 8-10. Of course, this step can be repeated for any new-line investments budget in order to explore a larger spectrum of planning possibilities.
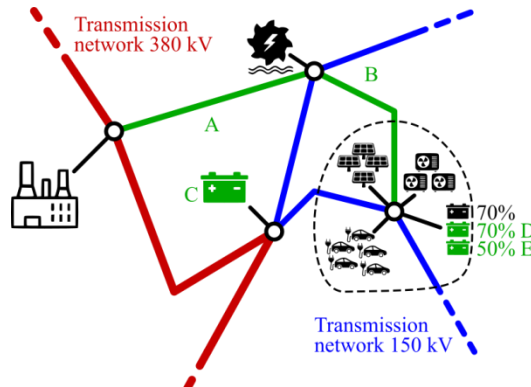
Figure 8-10 - Illustration of the remaining flexibility for transmission planning

Step 4 – Planning of the transmission system

Once (some of) the possible planning options for distribution planning have been listed, the transmission planning can be carried out by considering:

- The candidates available at transmission level (A, B and C).
- The candidates (energy units) available at distribution level (D and E) which might be limited by local constraints.

The participation of distribution resources to transmission planning depends on the selected option for distribution planning. Having assumed that three cases have been processed for the considered distribution network, the planning tool needs to select the best option by considering the following alternatives:

Distribution planning alternative	Availability of storage capacity from distribution network			Additional investments costs
	Existing	Candidate D	Candidate E	
I	70%	50%	0%	C_{step1}
II	100%	70%	90%	C_{step2}
III	70%	70%	50%	C_{step3}

Table 5 - Distribution network alternatives to be considered for the transmission planning

At this point, the planning tool is launched for transmission network, having preliminary provided as input the table of alternatives reported above. The objective function, which sees all the distribution (limited) candidates and flexible resources, selects the best planning option which, for the considered network, could be (Figure *8-11*):

- The installation of a new transmission line (candidate B).
- The exploitation of the existing distribution-level storage at 100% of its capacity (alternative II of distribution planning), which can be achieved with an additional investment cost of C_{step2}.
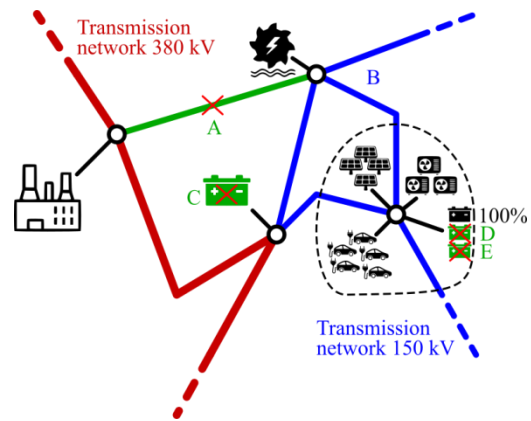
Figure 8-11 - Illustration of the optimal transmission planning

9 Proof-of-concept testing

As the last part of this document, models and methodologies developed earlier are validated through a proof-of-concept to demonstrate their key features, identify their potential limitations, and present the first findings on their scalability to large networks.

9.1 The FlexPlan.jl package

The FlexPlan.jl package is created in Julia/JuMP language[6] [35] in order to validate and test the mathematical model developed within WP1. In the course of FlexPlan, FlexPlan.jl will be used as the design reference for the planning tool, and to conduct validation tests. The software package uses the PowerModels.jl [36] and PowerModelsACDC.jl [37] packages as a basis to formulate the grid planning model, and to represent the AC and DC grid equations. Within FlexPlan.jl the aforementioned packages have been extended with the following functionalities so far:

- Demand flexibility modelling according to Section 5.1.

- Storage modelling according to Section 5.2.

- A stochastic planning model as introduced in Section 0.

- New network formulations for radial networks and combined transmission and distribution system planning (Sections 4.3 and 8.1.2 in this document)

- Addition of carbon footprint and emission costs in the "classical" planning objective (Section 6.2)

The main advantage of FlexPlan.jl is that it allows to test certain building blocks of the comprehensive FlexPlan model independently, without having to solve the full planning model. In this way, the model sensitivities with respect to certain parameters can be analysed in greater depth. Additionally, using the design guidelines of PowerModels.jl, the developed software package can be extended in a flexible way, and be used to analyse specific networks, planning objectives and other grid extension options, currently not in the scope of the FlexPlan model. Last but not least, FlexPlan.jl allows to use any commercial and open-source optimisation solver with an interface to Julia/JuMP, such that the computational performance of the planning model can be verified for a large range of solvers. FlexPlan.jl will be made available as an open-source software package in the course of FlexPlan.

---

[6] JuMP is a modelling language for mathematical optimization embedded in Julia supporting multiple solvers for a variety of problem classes, including linear, mixed-integer, second-order conic, semidefinite, and nonlinear programming.

9.2 Test systems

The model developed in this document should be applicable to both meshed grids (i.e., transmission systems) and radial grids (i.e., distribution systems) and is thus tested on both types of networks.

9.2.1 Transmission test system

The transmission test system used is based on the IEEE 6-bus system as defined in [38]. The original system has been augmented with DC branches, DC buses and AC/DC converters, in addition to the AC buses and AC branches of the original test system. The topology of the test system is illustrated on Figure 9-1, the right-hand side illustration presenting also the branches and converters investment candidates in dashed lines.
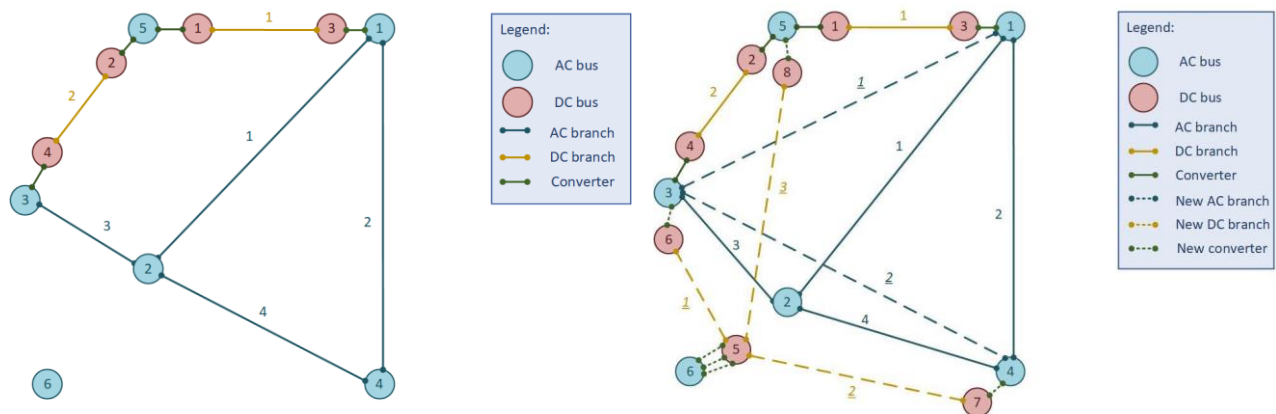
Figure 9-1 - FlexPlan version of IEEE 6-bus system with DC branches and converters. The left-hand side of the figure shows the base case network and the right-hand side of the figure also shows the network expansion candidates.

As in the original IEEE 6-bus system, AC bus 6 is not yet connected to the network and it is for the planning optimization to find the best candidate to connect the bus to the grid. Similarly, to [38], generation units are connected to the AC buses 1, 3 and 6. The main difference with the original system is that some generators at AC bus 3 and 6 have been assigned variable generation profiles to model vRES (variable Renewable Energy Sources).

Variable load profiles have also been used to characterize the loads at AC buses 1-5 by mapping the IEEE 6-bus system on Italy, with each node corresponding to a market zone (the network topology however does not reflect the actual Italian grid). This is illustrated on Figure 9-2. Load profiles can then be extracted from ENTSO-E Transparency Platform for each node[7]. More details on load and generation profiles are given in Section 9.3.7 (Tests on scenario generation and reduction).
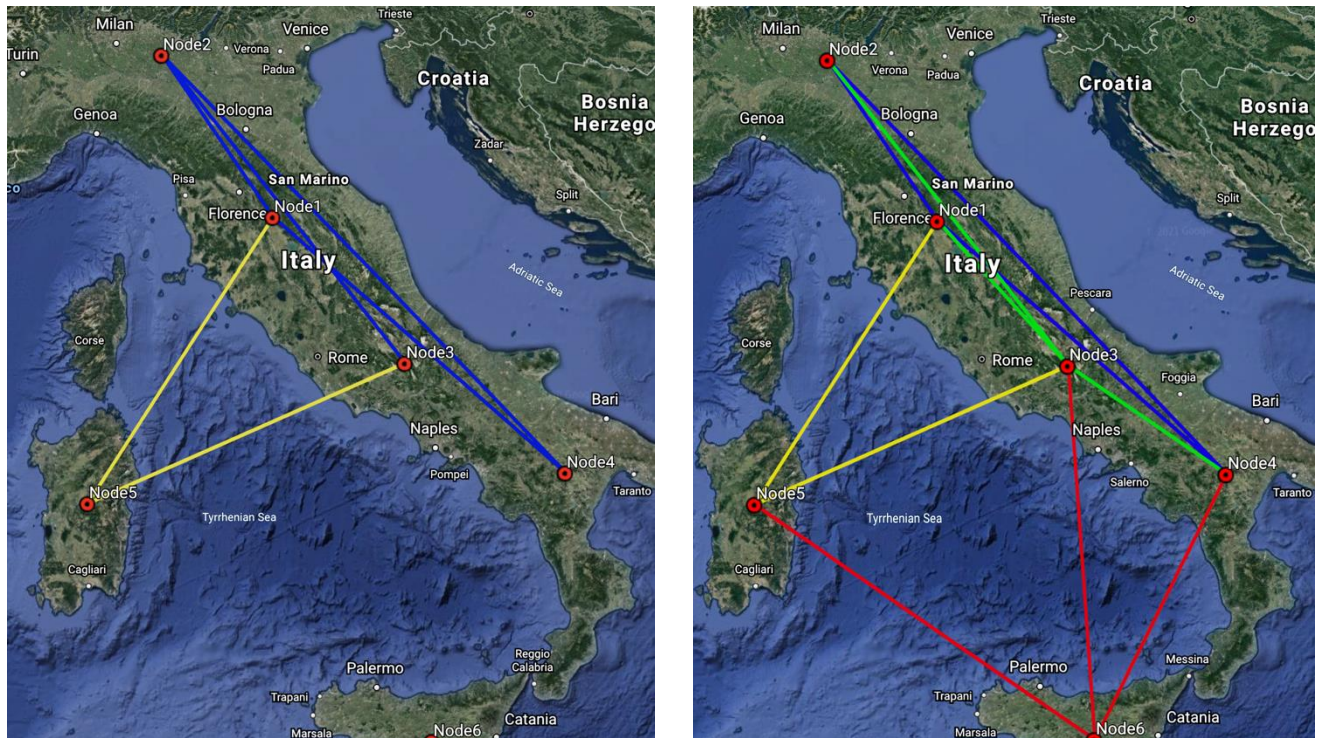
Figure 9-2 - IEEE 6-bus system mapped on Italy. The left-hand side shows the base case network and the right-hand side also shows the grid expansion candidates.

The technical specifications of the transmission test system are given in Section 11.1.1 (Appendix A).

9.2.2 Distribution test system

The chosen test distribution system is the CIGRE Medium Voltage (MV) distribution network benchmark (European configuration) described in [39]. The CIGRE benchmark networks are aimed at analysing the integration of Distributed Energy Resources (DER) at high, medium and low voltage levels. In particular, "the medium voltage (MV) distribution network benchmark is derived from a physical MV network in southern Germany, which supplies a small town and the surrounding rural area. Compared with this original network, the number of nodes for the benchmark network was reduced to enhance user friendliness and flexibility while fully maintaining the realistic character of the network" [39].

---

[7] https://transparency.entsoe.eu/load-domain/r2/totalLoadR2/show

The topology of the MV test distribution system is shown on Figure 9-3. The system consists of two feeders (shown by dashed lines) operating at 20 kV and connecting a total of 14 buses (plus the HV bus). The two feeders are connected to the 220 kV (sub)-transmission network via separate transformers. (The HV voltage
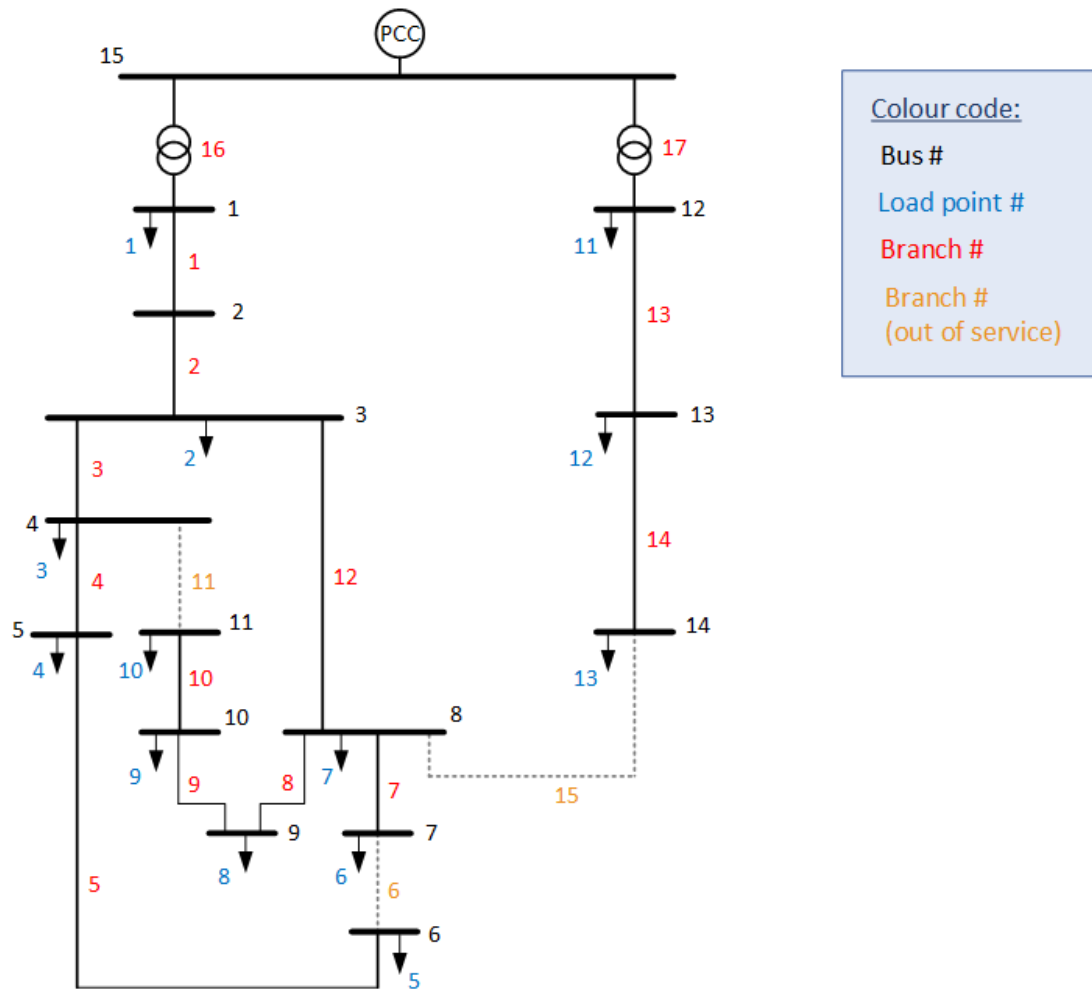
Figure 9-3 - CIGRE MV benchmark network (European configuration) with bus, load and branch numbers

level is 110 kV in the default version of the original benchmark network.) Three switches offer flexibility to consider either a meshed or a radial network. In our case, since we want to test the FlexPlan model on a radial network, all the switches will be considered as open, and the corresponding branches are marked as "out of service".

The test system features 13 DER of different types: PVs, residential fuel cells, wind turbines, CHP diesel and CHP fuel cells. The generators and load parameters are reported in Section 11.1.2 (Appendix A). For some tests, variable load and generation profiles from Italy and Norway are used, those are described in the corresponding tests.

9.3 Test cases

The test cases presented below aim at validating the different parts of the planning tool in an independent way and as such are presented in separate subsections. For the sake of conciseness, the results presented here focus on the tests that identify limitations of the model, bottlenecks for extrapolation to large-scale problems or challenges for the implementation of the planning tool. The tests that are less important in regard of the aspects above are described in the Appendices (Chapter 11). Finally, it is worth noting that this section does not provide the complete set of testing activities, as these testing activities will be continued in the course of FlexPlan. The additional tests that will be carried out are listed in the Section 9.3.9, the results of those tests will be published in a future update of this deliverable.

9.3.1 Tests on the storage model

9.3.1.1 Test specifications

The tests on storage are divided int base case tests and sensitivity tests. The base case tests aim at performing the optimization on the small test networks described above to check if the returned solution coincides with an expected solution of the power system model described in this document. Besides a general check on the solution compliance with the model constraints, the base case solution should answer the following questions on the existing storage assets:

1. Are the existing storage assets used? Does the observed usage make sense: power absorption when low loads in the system and injection when high or peak loads?

2. Are absorption and injection indeed exclusive (i.e., at a given time step a storage asset either absorbs or injects power but not both)?

3. Are the constraints on maximal energy content, maximal absorbed and injected power (over a time step and over a whole period) respected?

4. Is the final energy content at least as high as the initial energy content for each of the storage assets?

5. Verify the correct application of the dynamic storage equations between time steps: are absorption, injection and storage losses correctly taken into account to update storage state?

Regarding the candidate storage assets, the following questions should be answered:

6. Are investments in candidate storage assets made? If so, do they make sense (e.g., generation + existing storage cannot meet system demand during peak loads)? If not, why are the investments not made (e.g., other investments more interesting, load curtailment is used)?

7. If investments are made, are the corresponding investment costs correctly added in the objective function?

8. Are the new storage assets used (i.e., energy absorption and injection)?

9. As for the existing storage assets, are the constraints on the maximal energy content, maximal absorbed and injected power, injection/absorption exclusivity, initial and final energy levels as well as the dynamic energy level equations respected?

The transmission test system described above features one existing storage asset connected at AC bus 5 and two candidate storage assets connected at AC buses 2 and 5. The specifications of the storage assets are given in Appendix A (Section 11.1.3). For the distribution test system, we will use one storage asset connected at bus 5, similarly to [39][8]. The specifications of the storage asset are similar to those of electrochemical storage systems that can be found at distribution level [40], the full specifications are given in Table 35 (Annex A). For the tests on the distribution system, we use load and generation time series from Italy, the peak loads and maximum generation capacities being specified in Table 31 and Table 32.

Sensitivity tests are carried by changing the value of one parameter of the storage assets given above (while keeping the other parameters at their base value) and checking how the variation in the target parameter affects the global solution of the TNEP/DNEP. This section details the different important parameters that will be tested regarding storage assets and the tests that can be performed.

Parameter	Symbol	Description
Maximum energy content	$E_{j,y}^{max}$	Check that maximum energy content becomes a limiting constraint as this value decreases significantly.
Maximum energy absorbed over a year	$E_{j,y}^{abs,max}$	Check that maximum energy absorbed over a year becomes a limiting constraint as this value decreases significantly.
Maximum absorbed power	$P_{j,y}^{abs,max}$	Check that storage is inhibited as this value goes to zero. The other way around, if this value is very high can the storage asset charge completely in 1 time step?
Maximum injected power	$P_{j,y}^{inj,max}$	Check that storage is inhibited as this value goes to zero. The other way around, if this value is very high can the storage asset discharge completely in 1 time step?

[8] In [38], an additional storage asset is available at bus 10.

Absorption and injection efficiency	$\eta_{j,y}^{abs}, \eta_{j,y}^{inj}$	Check that storage is inhibited as these values go to zero. What is the efficiency threshold below which the system considers that it is not profitable to store energy? Check also if both charging and discharging can happen when the values are set to 1.
Self-discharge rate (storage static losses)	$dr_{j,t,y}$	Check that storage is inhibited as this value gets closer to 1.
Storage investment costs (equipment, installation, environmental)	$I_{jc}^{eq}, I_{jc}^{inst}, I_{jc}^{env}$	Check that storage is inhibited as these values rise significantly. The other way around, if investment costs are zero, are all storage asset candidates implemented?
Status	$s_{j,t,y}$	Setting status to 0 should make sure the corresponding storage assets are not used (i.e., no power injection and absorption). Are investments in new assets then made?
Power provided by external process (e.g., water supply in Pumped Storage Hydro)	$\xi_{j,t,y}$	Check that the energy balance for each asset is respected when this value is nonzero.

The storage ramping constraints materialized by the maximum absorption ramp rate ($r_{j,y}^{abs,max}$ nd the maximum injection ramp rate ($r_{j,y}^{abs,max}$ ave not been implemented as part of the WP1 testing prototype.

9.3.1.2 Test results

Tests on the transmission system

As a first test we run the <u>base case</u> for storage on the transmission test system over 96 hours using only 1 scenario (based on 2019 data) to have a first intuitive look at the results. We then run the base case test over a longer period, namely 1000 hours. The most interesting results, describing key features or limitations of the model, are detailed below while, for the sake of brevity, other test results are described in Appendix B (Section 10.2.1).

Figure 9-4 shows that the observed grid storage usage makes sense: power is absorbed when the total generation capacity of the system (red curve) exceeds the total load in the system (blue curve) (e.g., because of high renewable energy production during that hour). The storage asset then injects power into the grid mainly when the generation capacity of the network cannot meet the demand (i.e., when the blue curve passes over the red curve). We also see that with lossy storage assets ($\eta_{j,y}^{abs} < 1$, $\eta_{j,y}^{inj} < 1$), absorption and injection are observed to be exclusive as required (even if the nonlinear exclusivity constraint is not enforced as such, see Section 5.2.1 for details). In the base case, storage charging and discharging is also observed to be exclusive over the longer test period (1000 hours).

Regarding the test over a longer time period, it is useful to know that, even if the optimization is conducted
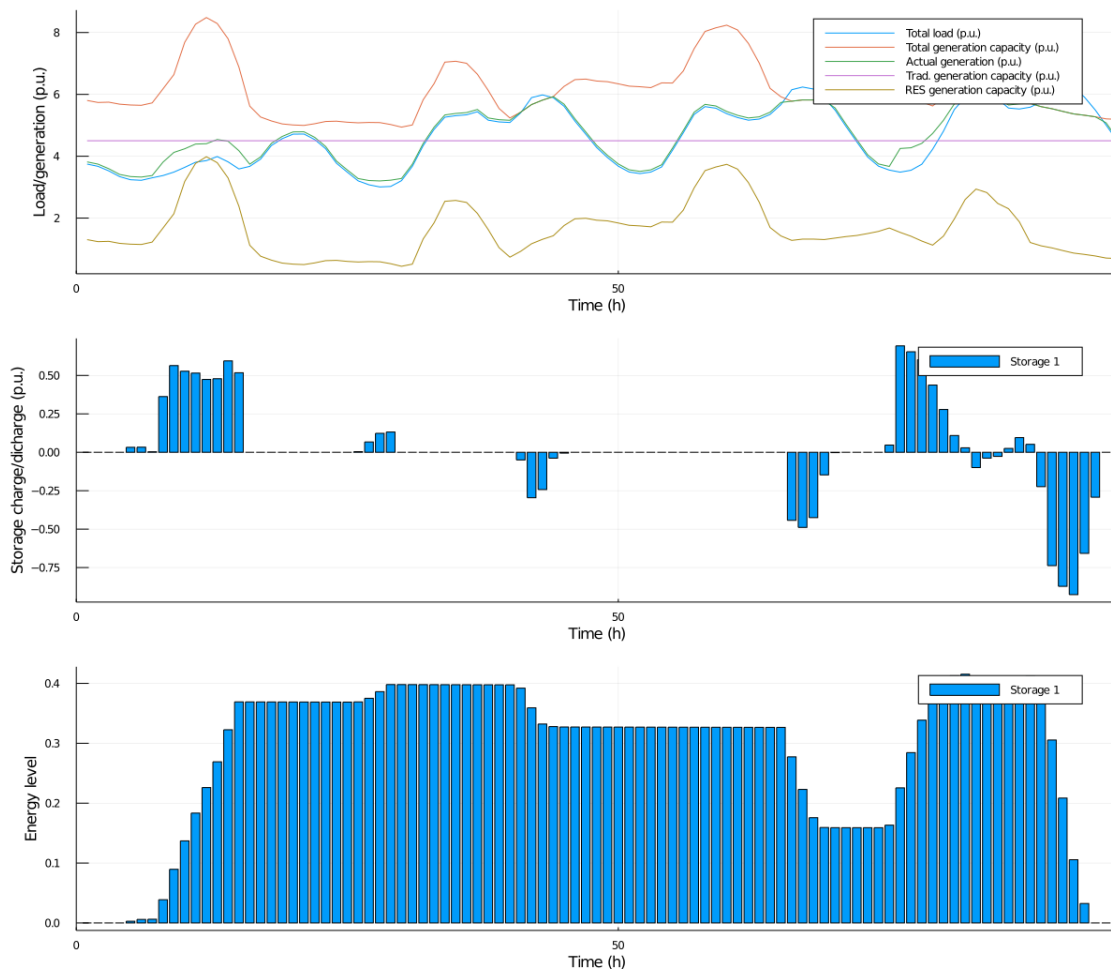
Figure 9-4- Results of the base case test on the transmission test system for storage assets (first 96 hours). In the upper figure the curves of interest are the system load (blue), the system generation capacity (blue) and the system actual generation (green).

over a limited period (1000 hours here), the FlexPlan.jl module scales the operational costs (i.e., generation and emission costs) to the whole year. The costs can also be scaled further to a given number of planning years. Scaling the operational costs allows to take into account that the investments made will be operational for several years (and thus to scale the investment costs over this horizon). In our case we scale the operational costs to 10 years.

In the TNEP solution, the optimal binary decision values show that investments are made in both candidate storage assets of the transmission system: the one connected at AC bus 2 (candidate storage 1) and the one AC bus 5 (candidate storage 2). By analysing the optimal solution for the 1000-hours problem, we find that the DC lines connecting AC bus 5 (Sardinia) with mainland Italy are often congested (see Table 6). It therefore makes sense that the planning tool invests in a storage asset in mainland Italy (candidate storage 1) as the connections to the existing storage in Sardinia are often congested.

Line	Congestion rate (rate of congested hours/year)
AC line 1 (AC bus 1 to AC bus 2)	0 %
AC line 2 (AC bus 1 to AC bus 4)	3 %
AC line 3 (AC bus 2 to AC bus 3)	23.5 %
AC line 4 (AC bus 2 to AC bus 4)	23.2 %
DC line 1 (DC bus 1 to DC bus 3)	35.8 %
DC line 2 (DC bus 2 to DC bus 4)	51.6 %
Candidate DC line 2 (DC bus 5 to DC bus 7)	0 %

Table 6 - Congestion rates of lines in the transmission test system (1000 hours optimization). The congestion rate is the number of hours during which the line is operated at its maximal (nominal) capacity over the total number of hours.

As for the additional storage capacity which is invested at AC bus 5 (Sicily), it can first be ruled out that the investment was made due to the bounds on maximum injection and absorption power as the existing and new storage assets together never absorb or inject more than 160 MW of power (whereas the bound is at 250 MW per storage asset). However, regarding the energy rating, the existing and new storage assets at AC bus 5 store more than 1500 MWh together at some point (which is more than the 1000 MWh energy rating of a single storage asset). It is thus likely that the investment in additional capacity at AC bus 5 was made to increase the bounds on the total storage capacity.

The only line investment which is made is the implementation of a DC line between DC bus 7 (South Italy) and DC bus 5 (Sicily), to connect the latter to the network. It therefore turns out that with the parameters of candidate assets specified in Appendix A (Section 11.1.3), investment in storage is chosen over investment in new lines (despite the fact that most line candidates are cheaper than the storage candidates). One of the main reasons is probably that the load and generation profiles are such that at some point the total load exceeds the generation capacity of the system. As we do not consider demand flexibility in this test case (i.e., load shifting and load reduction), the additional power needed either has to be injected through storage assets or costly load curtailment has to be performed. If investments in new lines can solve congestion issues, they do not provide

additional power in those hours of generation scarcity. Of course, the investments made depend also on the parameters which are set for the storage candidates, further study of the importance of these parameters is performed in the sensitivity analysis.

We also perform the base case optimization test using the stochastic approach with two equiprobable scenarios. The first scenario uses load and generation data from 2019, the second scenario uses data from 2018. As above, we run the simulation over a thousand hours but with the costs scaled to 10 years. Both storage candidates are again invested in, but their usage depends on the scenario - as load and generation profiles are different (see Figure *9-5* for an example of different storage usage depending on the scenario for a 96h-period). The value of the objective function is slightly lower in the stochastic problem than in the problem with a single scenario (2019), which could be explained by slightly lower load values/higher RES production in 2018 than
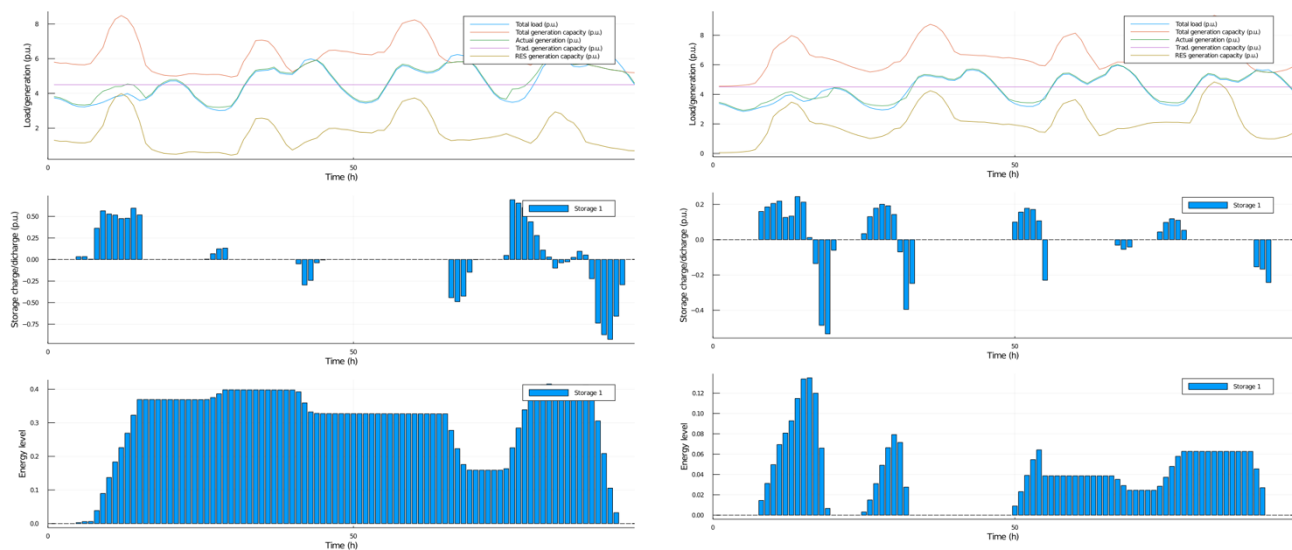
Figure 9-5 - Illustration of two different storage usage profiles depending on the scenario: 2019 (left) and 2018 (right) (96 hours stochastic optimization). In the upper figures the curves of interest are the system load (blue), the system generation capacity (blue) and the system actual generation (green).

in 2019.

For the sensitivity tests**,** we first increase the cost of the candidate storage assets (compared to the base case) so that no investment in storage is made at all (i.e., no candidate is built, and load curtailment is used instead). We then change the values of the storage assets' parameters and analyse the potential changes in the TNEP solution (e.g., are different investment decisions taken).

Most sensitivity test are detailed in the Appendix B (Section 11.2.1) as they result in solutions that make sense and do not highlight limitations of the model developed in this deliverable. The only result that we will detail here is the sensitivity analysis on the absorption and injection efficiencies as we find that setting both parameters to 1 can lead to inconsistent solutions with storage charge and discharge happening simultaneously.

As we see for hour 9 in the Figure 9-6 below, storage charge and discharge can then happen simultaneously since the mutual exclusion of power absorption and injection is not enforced as such in the model. As simultaneous charging and discharging comes at no cost (since there are no losses), some optimal solutions may indeed contain such hours where injection and absorption are not exclusive. It should therefore be noted that using the planning tool with storage assets having 100% round-trip efficiency may lead to inconsistent solutions to be returned
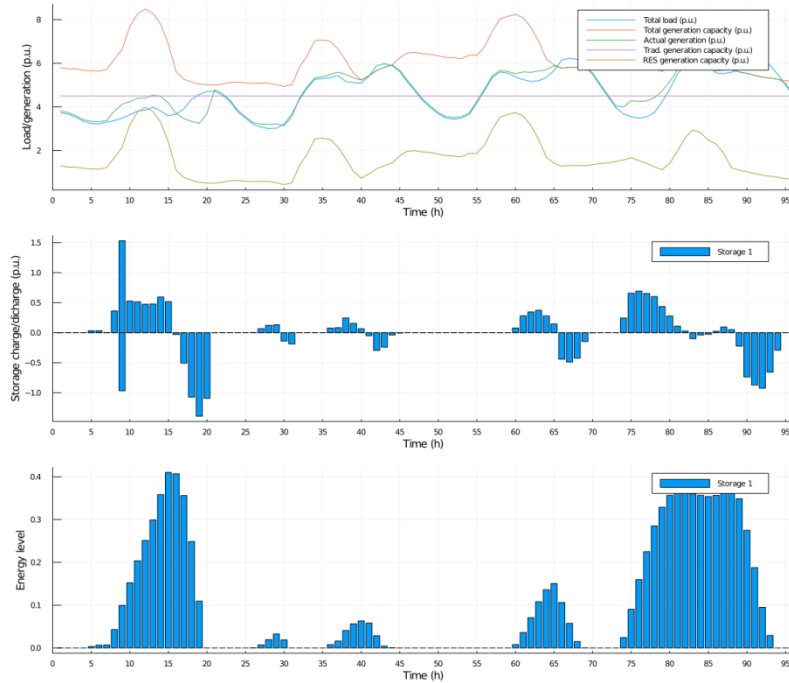
Figure 9-6 - Sensitivity test: 100% absorption and injection efficiencies.

In the opposite way we also simulate very low efficiencies with $\eta_{j,y}^{abs} = \eta_{j,y}^{inj} = 0,1$. As depicted below, the storage asset then starts charging much earlier as it has to cope with important absorption losses. The amount of power injected into the grid is also severely reduced due to losses at injection, and the system has to perform costly load curtailment at load 2 and load 5.
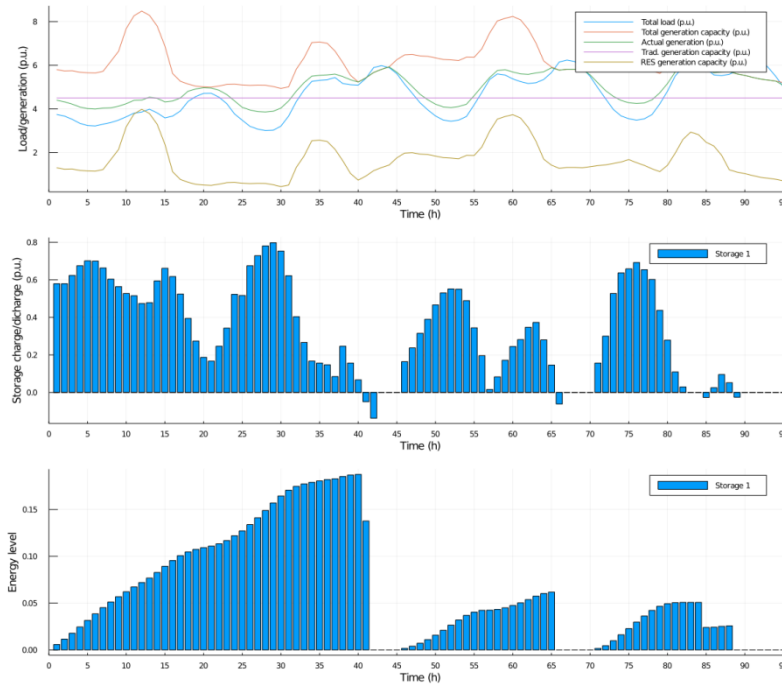
Figure 9-7 - Sensitivity test: 10% absorption and injection efficiencies

Tests on the distribution system

The storage model for distribution systems differs from the model for transmission only by the presence of reactive power, and we will thus focus on how reactive power absorptions/injections from storage assets can provide operational support at distribution level. As presented on Figure 9-8 below, in the first 48 hours of the simulation we see little active power absorptions/injections (only once at hours 8 and 9). On the other hand, the storage asset injects reactive power at almost each time step. As such, in these hours the storage asset acts similarly to a Var compensator, providing reactive power to the grid.

Reactive power injections allow to smooth the voltage profile at bus 5 during those hours, as presented on Figure 9-10 and Figure 9-9. In Figure 9-10, the storage asset at bus 5 is deactivated and the lower bound on voltage magnitude (0.95 p.u.) is reached at several timesteps. With the storage asset at bus 5 activated, the voltage profile stays much closer to the baseline (Figure 9-9).

In the simulation above, we use load and generation profiles from Italy, the load profile being multiplied by a scaling factor of 2.20 to trigger congestion issues in the grid. With an increased scaling factor of 2.30, the case without storage asset becomes infeasible: since branches 1 and 2 are congested no additional reactive power can be pulled from the HV connection and the voltage magnitude drops below 0.95 p.u. The case with the storage asset activated however remains feasible with a load scaling factor of up to 2.40.

The model developed in Section 5.2 does not take into account losses associated with reactive power absorptions and injections, as storage losses are mainly accounted for through active power absorption and injection efficiencies.

If one would like to reflect that using storage to improve the voltage profile is lossy, the costs for reactive power need to be added to the objective. This might however be tricky for two reasons:

- The need to account for the cost independently of the sign of Q requires absolute values, or two separate variables as for the active power injections and absorptions coupled with an exclusivity constraint.

- The cost for reactive power services is not really defined as such, so it may be hard to put a realistic number on them.

One way could be to add active power loss terms in dependence of the reactive power (in a linear way), to account for the fact that providing reactive power services is of course lossy.
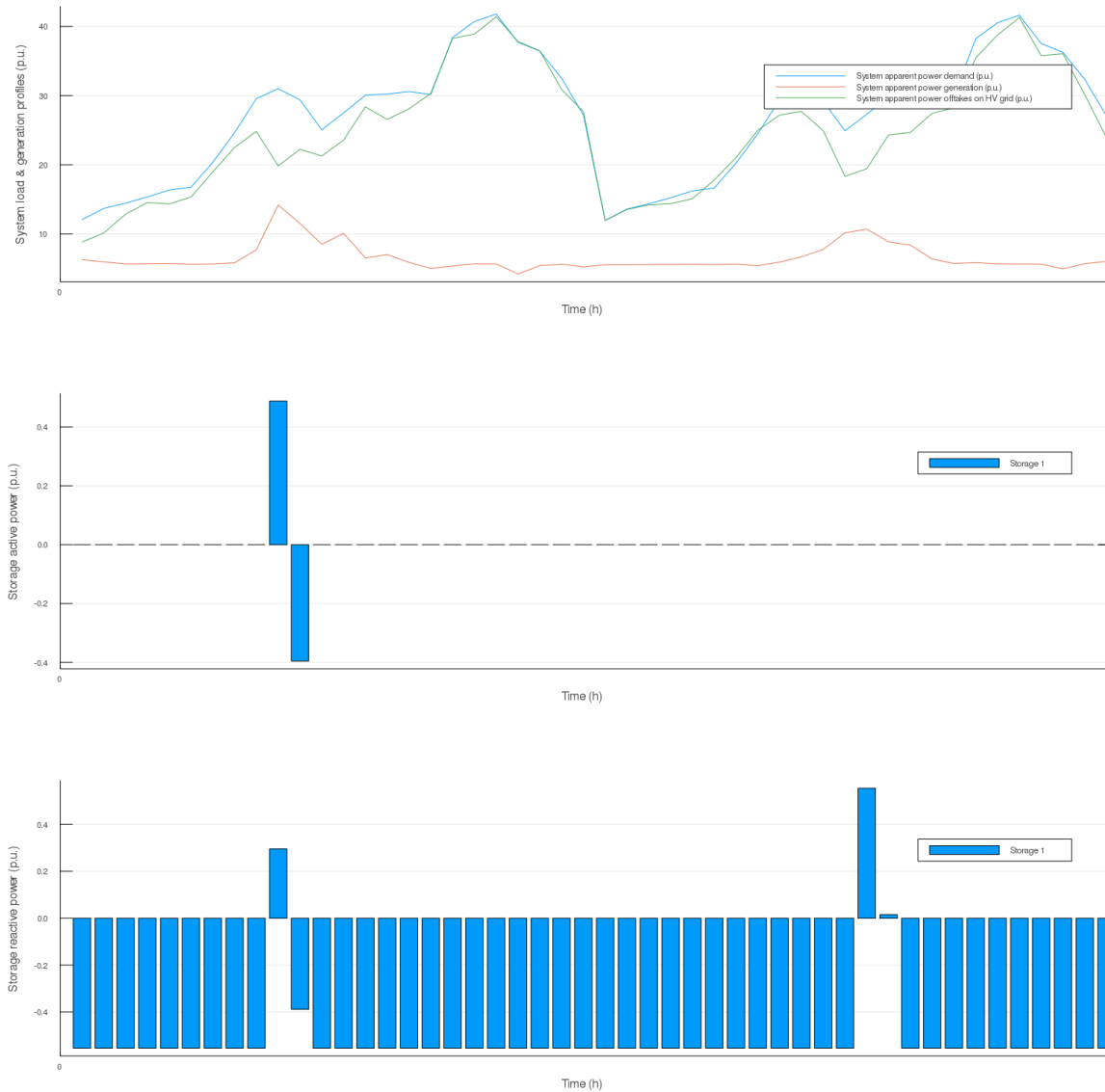
Figure 9-8 - CIGRE MV system: system load and generation profiles (1), storage active power absorptions/injections (2) and storage reactive power absorptions/injections (3). In the figure (1) the curves of interest are the system apparent power demand (blue), the system apparent generation (red) and the apparent power offtakes on the HV grid (green).
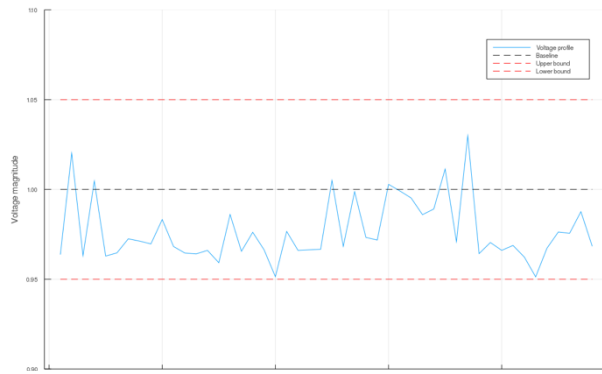
Figure 9-10 - Voltage profile at bus 5 with storage asset deactivated
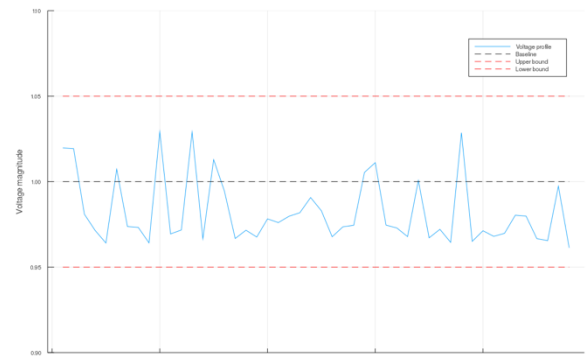
Figure 9-9 - Voltage profile at bus 5 with storage asset activated

9.3.2 Tests on the demand flexibility model

9.3.2.1 Test specifications

For the tests on demand flexibility at the transmission level, it is assumed that only load at AC bus 5 can provide demand flexibility. The characteristics of this flexible load are given in Table 36 (Appendix A). Note that these parameter values are chosen for a simple and illustrative demonstration of the model functionalities, and that more realistic values are presented for the distribution system test case.

For the test of the flexible demand model in the distribution test system, it is assumed that all loads in the network are residential and that they all have the same demand flexibility characteristics. The base case parameter values are specified in Table 37 (Appendix A) and therefore apply for all loads in the network.

Load demand data from Norway is used to test the demand flexibility model for distribution systems, based on the original CIGRE MV benchmark network data [39]. Table 38 (Appendix A) gives the baseline load parameters for the demand flexibility tests. Since the load demand time series are all for residential loads, the original data set is modified to represent only residential load for all buses in the network. The power factors values are the same as for the residential component of the loads in [39].

The Norwegian data are hourly residential load demand time series for a year measured at household level and aggregated to the distribution substation level for 14 distribution substations. These data are then distributed on the buses of the CIGRE MV benchmark system. Since there is no load at bus 2 in the original CIGRE MV benchmark system, load 2 is at bus 3, load 3 is at bus 4, etc., and only 13 of the 14 load time series are used in the system. The original load demand time series are normalized (as p.u. values) and scaled with
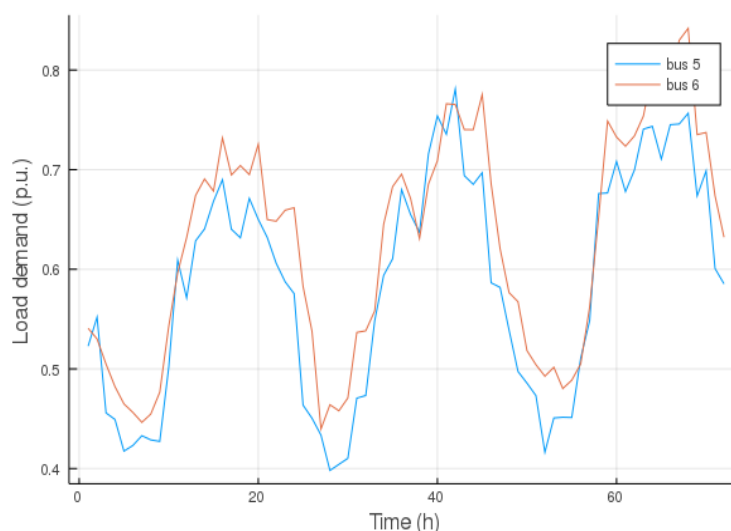
Figure 9-11 – Example of load profiles at bus 5 and 6 in the

the base case load demand at the respective buses in the CIGRE MV benchmark case. Examples for the first three days of load demand for buses 5 and 6 are shown below.

The tests are specified as a 1) base case test for a set of base case input parameters specified in above and in Appendix A (Section 11.1.4), and 2) and a set of sensitivity tests that checks the implications of changing values of the parameters defining the base case.

The base case load and generation data should be such that an investment needs to be made to reduce overload on the power lines in the system for a number of time steps during the year. This can be obtained by having a load peak that gives power flows exceeding the power line limits for a few hours for at least one of the representative days in the year. The base case load and generation data should be such that an investment needs to be made to reduce overload on the power lines in the system for a number of time steps during the year. This can be obtained by having a load peak that gives power flows exceeding the power line limits for a few hours for at least one of the representative days in the year. The base case should be specified in such a way that the optimal solution is to activate demand flexibility. At the same time, it should be easily tuned to trigger a power line investment for parameter values where sufficient demand flexibility activation is not possible or cost-effective.

For the transmission test system, the base case test involves net power imported to bus 5 in such a way that there are hours with overload of DC branches 2 and 3 feeding bus 5. This overload can either be relieved with activation of demand flexibility at bus 5 or by investing in an additional DC line feeding bus 5 (candidate DC line 3).

For the distribution test system, the base case test involves undervoltage problems along the 1st radial of the MV distribution network (buses 1 to 11) on times when the load demand for these buses become too large. The grid reinforcement alternative is to install new distribution lines (underground cables) along the segments of the radial network with the largest voltage drop (from bus 1 to bus 3). The specifications of these lines are given in Table 39 (Appendix A). The lines are assumed to be installed in parallel with the existing lines. For simplicity there is assumed to be no on-load tap changers in the network and no distributed generation. With the voltage limits set for the distribution network (±0.05 p.u.), this means that undervoltage problems appear in the network even with lower load demand than the baseline peak load demand specified in Table 38 unless grid investments are made or demand flexibility is utilized. A load demand scaling factor is applied to reduce the load demand for all loads proportionally.

Regarding sensitivity tests, the table below gives a brief description of the possible checks to perform:

Parameter	Variable name	Description
Reference demand	(Load demand scaling factor introduced in the test data)	Can test with different load profiles, and check e.g., if line investment becomes necessary as reference demand gets more severe peak. (What happens if there are multiple peaks per day?)
Superior bound on not consumed power	$\Delta^{nce,max}$	Check that voluntary load demand
Superior bound on upward demand shifted	$\Delta^{ds,up,max}$	Check that load shifting is inhibited as the value goes to zero.
Recovery period for upward demand shifting	$\tau^{ds,up,rec}$	Check that load shifting is inhibited as the value goes to zero.
Superior bound on downward demand shifted	$\Delta^{ds,dn,max}$	Check that load shifting is inhibited as the value goes to zero.
Recovery period for downward demand shifting	$\tau^{ds,dn,rec}$	Check that load shifting is inhibited as the value goes to zero.
Investment costs	I	Check that flexibility alternative is not optimal solution when this value is set too high.

In addition, sensitivity tests could be specified for the line investment alternative, e.g., to check that line investment is included in the optimal solution as the line investment cost goes to zero.

9.3.2.2 Test results

Test results on the transmission test system

We use the IEEE 6-bus transmission test system and a synthetical load demand time series for load 5 to test the response of the flexible load and run the model for 96 hours. Demand peaks are introduced at day 2 and 3 where more electricity demand is required than can be served through DC links 1 and 2. As a result demand has to be either curtailed, reduced or shifted. The resulting energy balance without load shifting is shown in the Figure *9-12* below. The accumulated energy not consumed due to load reductions is shown at the bottom at the Figure *9-12*, and one can see how load reductions are scheduled to meet the upper boundary (1000 MWh) during the 96-hour planning horizon.
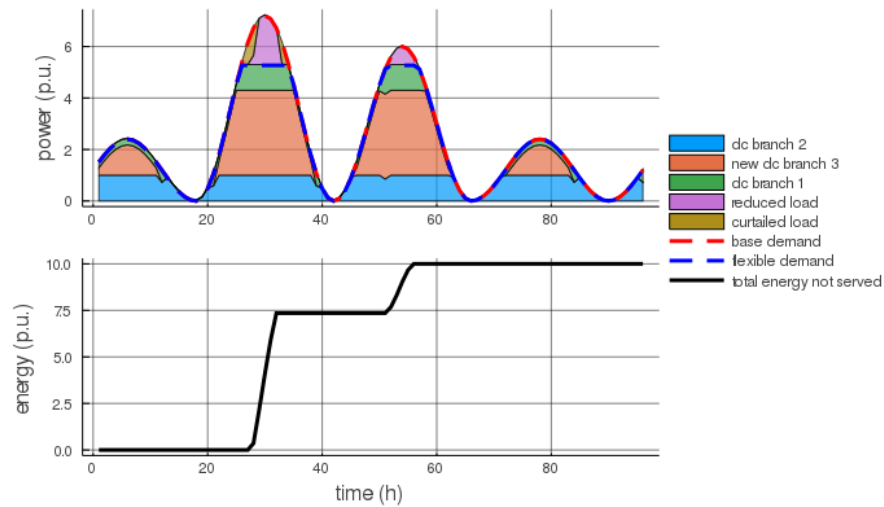
Figure 9-12 - Demand profile at load 5 (base case parameters)

Allowing load shifting with no restrictions related to the recovery period removes all load reduction and curtailment as seen in the Figure *9-13*:
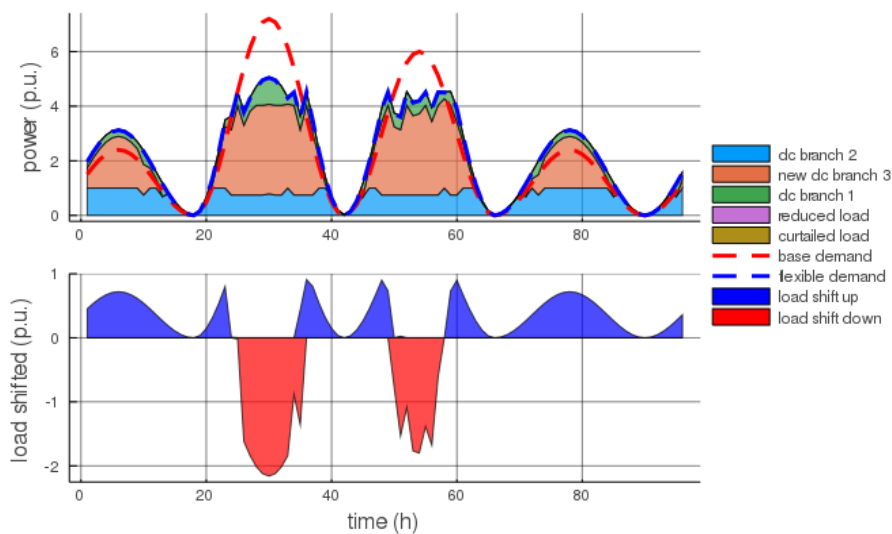
Figure 9-13 - Demand profile at load 5 (no restrictions on recovery period)

Setting the maximum load shifting to 10% of the baseline load demand reduces the amount of load shifting: When the recovery period is set to 10 hours and the maximum load shifting to 100%, this reduces the amount of load shifting further and increases demand reduction.
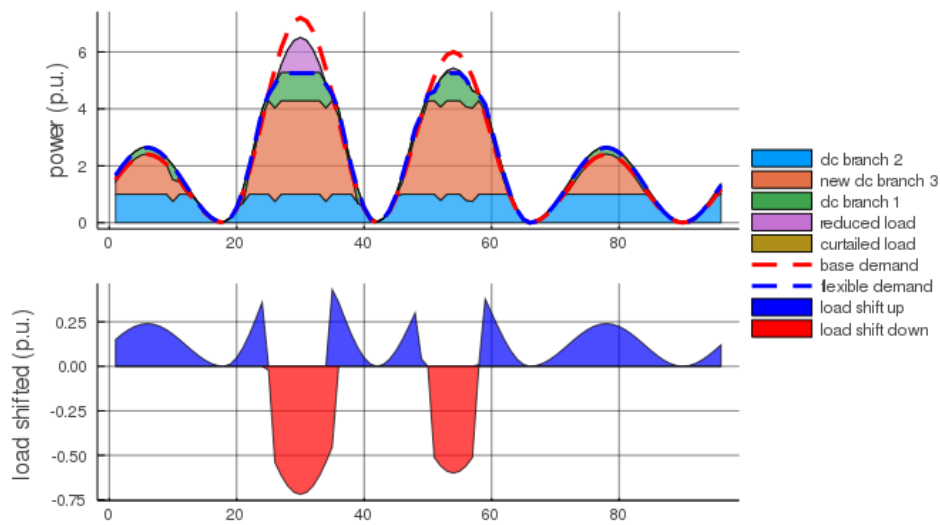
Figure 9-14 - Demand profile at load 5 (upper bound on load shifting at 10% of reference load)

The recovery period is changed to 20 hours which lead to significantly less load shifting as the demand cannot recover between the peaks.
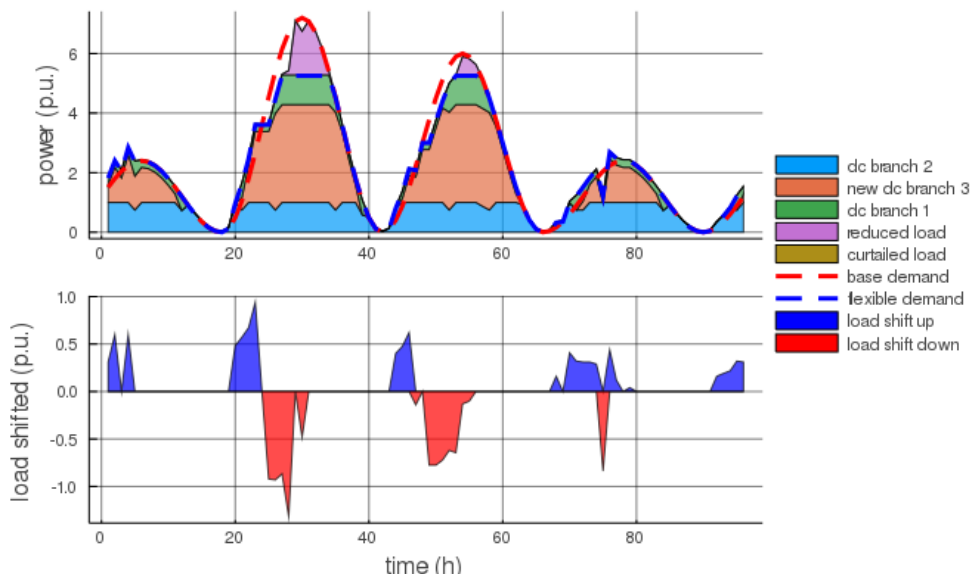
Figure 9-15 - Demand profile at load 5 (recovery period set to 10 hours)
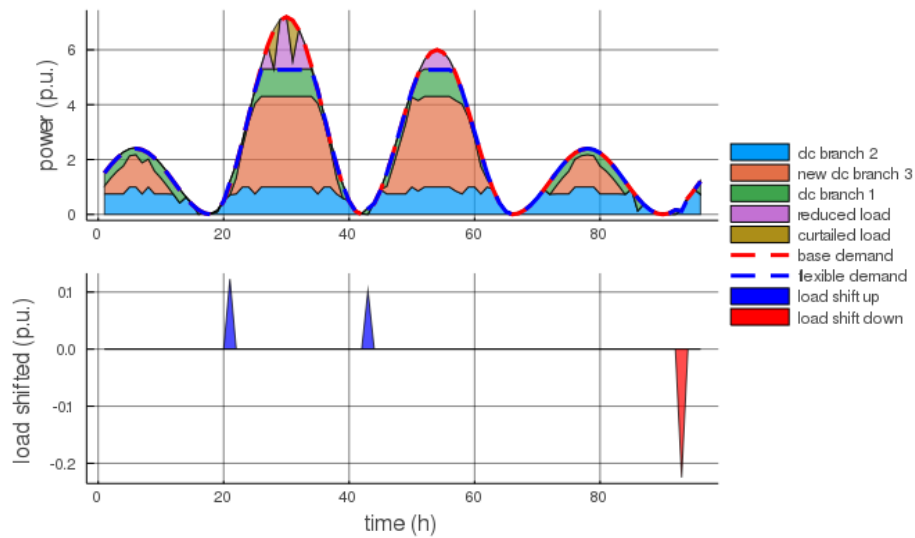
Figure 9-16 - Demand profile at load 5 (recovery period set to 20 hours)

The cost of DC branch 3 is increased to 1000 such that this branch is no longer built. This results in large amounts of curtailed energy and maximum load reduction as shown in the energy balance below.

We reset the investment cost of DC line 3 and run a marginal analysis for the flexible load investment, i.e., the costs of enabling flexibility for the load. The investment cost is increased from 0 to 10^4 with an increment of 10 as shown in the figure below. For an investment cost between 100 and 1000 it is not profitable to invest in flexible loads anymore.
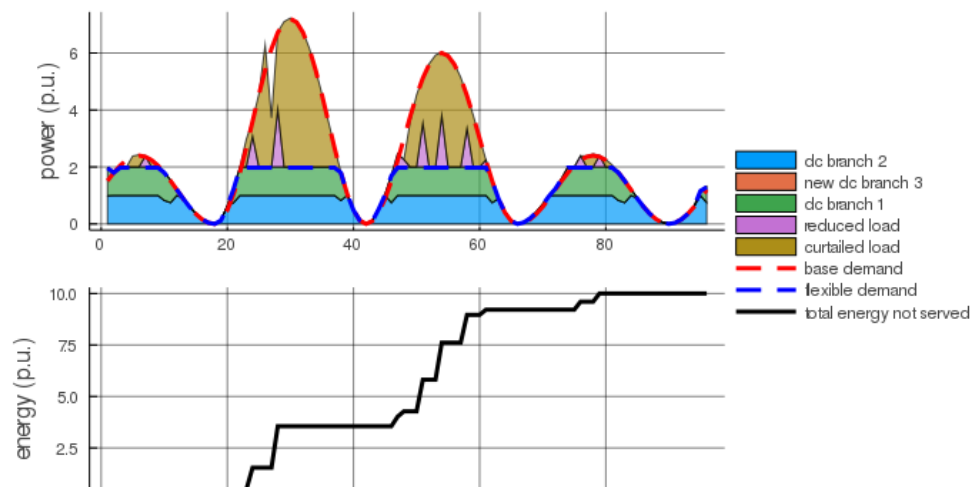
Figure 9-17 - Demand profile at load 5 (increased cost of candidate DC branch 3)
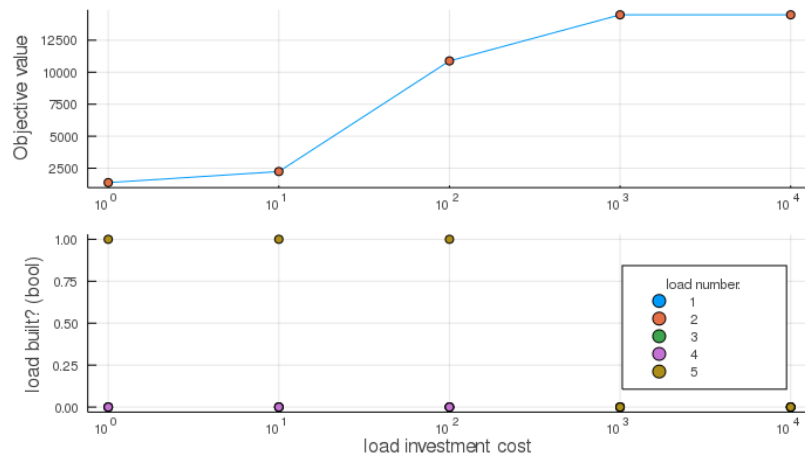
Figure 9-18 - Marginal analysis on flexible load investment costs

Test results on the distribution test system

Figure 9-19 shows results for the bus voltages in radial 1 for the case that no demand flexibility is enabled in the distribution test system. This means that the demand shifting potential and voluntary load reduction potential are both set to zero. Figure 9-20 shows the corresponding load flow along the radial network from bus 1 and outwards. To make it simple to illustrate and interpret the results, only the three first days of the load demand time series are used. A load scaling factor of 0.8 relative to the baseline load demand has been applied to the time series.

For all hours, the voltage drop is greatest from bus 1 to bus 2 (branch 1) and from bus 2 to bus 3 (branch 2). These are also the buses between which candidate branches exist in the test case. In this case, investment is made in the second candidate branch (from bus 2 to bus 3) to avoid undervoltage for buses 3 and outwards. One can see from Figure 9-19 that the voltage magnitudes at these buses are barely above the limit of 0.95 p.u. for the hours with highest load demand. The colour of the area under the graph in Figure 9-20 indicates that a
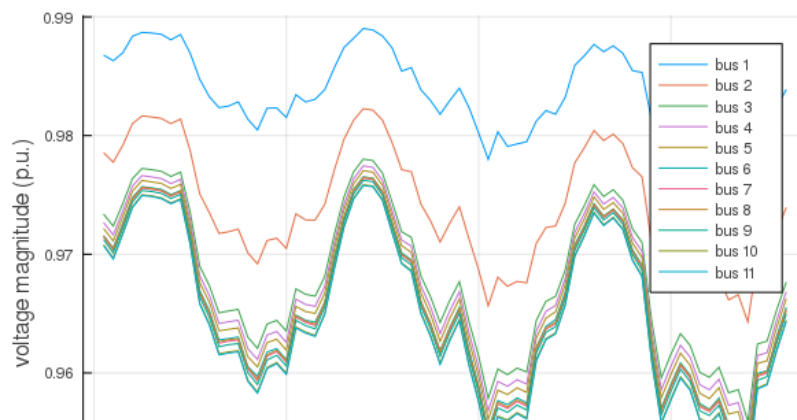
Figure 9-19 - Time development of bus voltage magnitudes in the
CIGRE MV benchmark system without demand flexibility.

new branch is being built and that the new and old branches are treated as a single new branch in the model (representing two branches in parallel) that replaces the old one.
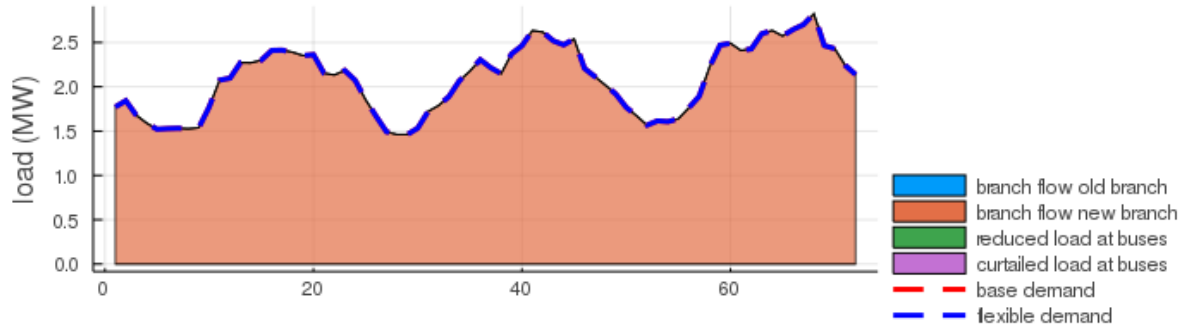
Figure 9-20 - Load balance in radial 1 of the CIGRE MV benchmark system beyond bus 1 without demand flexibility.

Figure 9-21 and Figure 9-22 show the results as demand flexibility is enabled and the demand shifting potential (the fraction of load demand that can be shifted, $\Delta^{ds,dn,max}$ ^$(ds,up,max)= \Delta^{ds,max}$s set to 0.1 for all load buses. The fraction of the load that can be reduced voluntarily ($\Delta^{nce,max}$s set to the baseline value of 0.05. The colour of the area under the upper graph in Figure 9-21 indicates that there are no new line investments in this case. One can moreover see that load is being shifted for all three days, and as the load demand becomes more severe in day 2 and 3, the more costly options of voluntary and involuntary load reductions are also progressively activated. The bus voltage time series in Figure 9-22 shows how undervoltage is the phenomenon driving investments or flexibility activation. For peak load periods, the voltage at bus 11 (in the end of the radial) is forced down to the lower limit of 0.95 p.u.
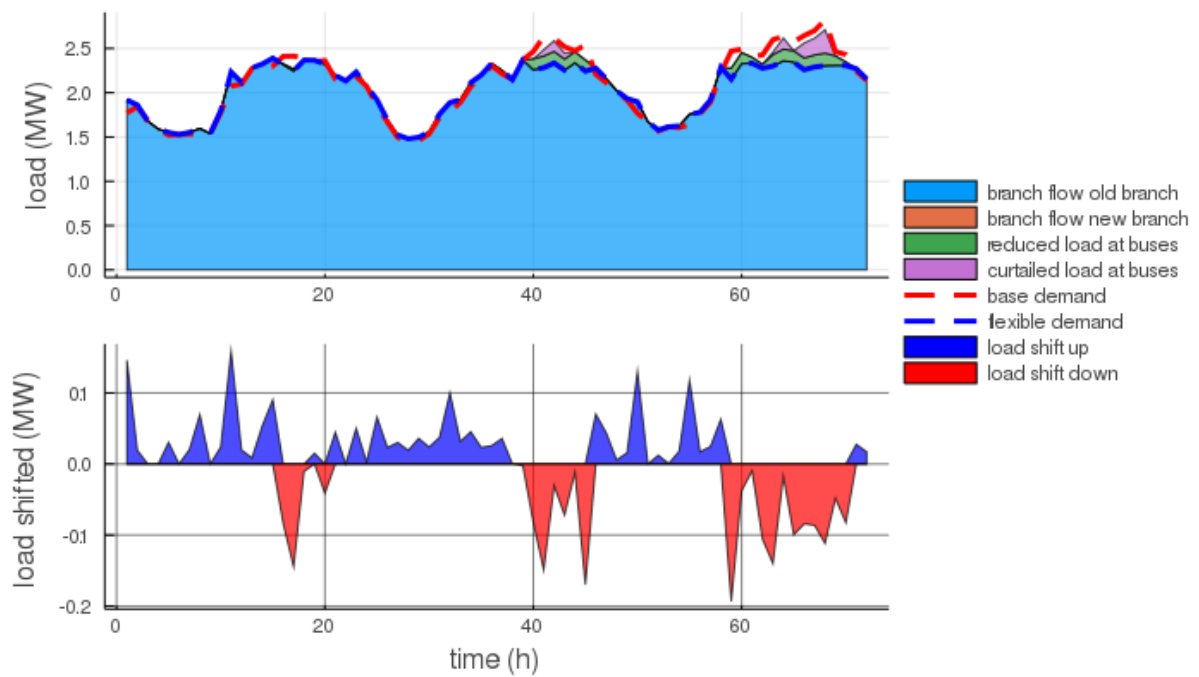
Figure 9-21 - Load balance (top) and demand shifting (bottom) in radial 1 of the CIGRE MV benchmark system beyond bus 1 with demand flexibility.
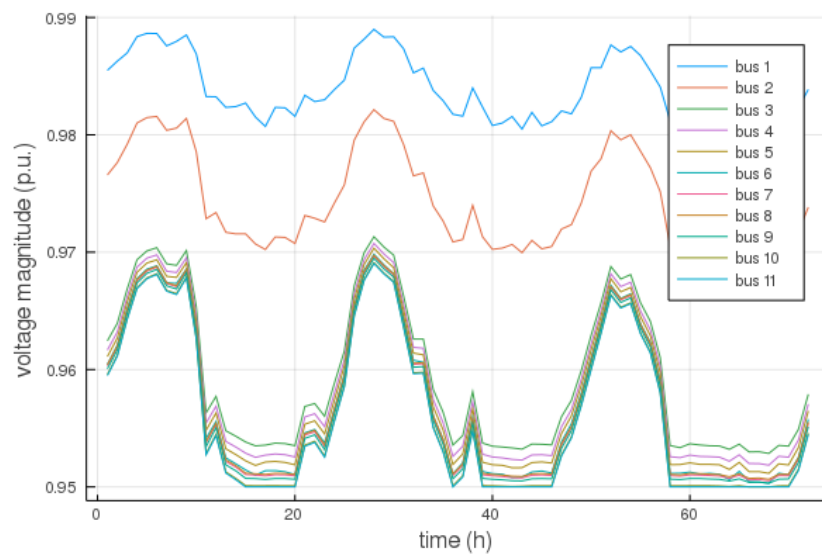
Figure 9-22 - Time development of bus voltage magnitudes in the CIGRE MV benchmark system including demand flexibility.

The sensitivities of the line investment decisions are tested for the demand flexibility model considering i) increases in the total load demand in the network using a proportional load scaling factor, ii) the demand shifting potential, iii) and the recovery period for load shifting. The results are shown in Figure 9-23. Here, the number of new lines built is shown for each combination of the parameter values. As the load demand increases, more line investments have to be made, but the need for line investments are reduced if the demand shifting potential increases and the recovery period decreases. The column marked with $\Delta^{ds,max}.0*$ contains results for the case with no demand flexibility, i.e., there is no potential for voluntary load reduction reduction ($\Delta^{nce,max} = 0.0$; for all other sensitivity cases $\Delta^{nce,max} = 0.05$ is kept fixed).

9.3.3 Tests on the reliability model

9.3.3.1 Test specifications

The input parameters specific to the reliability modelling in FlexPlan are listed in the table below together for the values assumed for the test. The data and sources underlying these estimates are stored in Appendix A (Table 40).

The following assumptions were made in these estimates: Only a single VOLL value is used for all load points, and it is assumed that they represent the same customer group composition. The value is representative for a Norwegian city with most load demand being due to residential and commercial loads. Typical values for other countries will differ. The reliability data are based on the Norwegian standardised system FASIT for collection, calculation and reporting of disturbance and reliability data (FASIT). The values are for permanent

Recovery period: 2 h   **Recovery period: 6 h**   **Recovery period: 10 h**

Load scaling factor	p_shift_max 0.0*	p_shift_max 0.0	0.1	0.2	0.3	p_shift_max 0.0	0.1	0.2	0.3	p_shift_max 0.0	0.1	0.2	0.3
0.800	1	1	0	0	0	1	1	1	0	1	1	1	0
0.825	1	1	1	0	0	1	1	1	1	1	1	1	1
0.850	1	1	1	1	0	1	1	1	1	1	1	1	1
0.875	1	1	1	1	1	1	1	1	1	1	1	1	1
0.900	1	1	1	1	1	1	1	1	1	1	1	1	1
0.925	2	1	1	1	1	1	1	1	1	1	1	1	1
0.950	2	1	1	1	1	1	1	1	1	1	1	1	1
0.975	2	2	1	1	1	2	1	1	1	2	2	1	1
1.000	2	2	2	1	1	2	2	2	1	2	2	2	2

Figure 9-23 - Sensitivity results for line investments (the number of new branches) for the flexible demand model considering the load scaling factor for the total load demand in the network, the limit on the fraction of load demand that can be shifted ($\Delta^{ds,dn,max} = \Delta^{ds,up,max} = \Delta^{ds,max}$) and the recovery period.

faults of overhead transmission lines at voltage levels 300 - 220 kV, and the average values over the past 20 years is estimated by SINTEF. The MTTR value is strictly speaking based on the outage time rather than the repair time but is in any case associated with large uncertainties.

Parameter	Variable name	Base case value
Value Of Lost Load, VOLL (€/MWh)	VOLL	6000[9]
Failure rate for 240 kV transmission lines (per year per km)	λ	0.0007
Mean Time To Repair for 240 kV transmission lines (h)	t^{MTTR}	60

Table 7- Specifications of parameters for reliability modelling

The tests are specified as a base case test and a set of sensitivity tests that check the implications of changing values of the parameters defining the base case.

Setting up a base case

First the load time series and/or parameters of the 6-bus transmission test system should be modified in such a way that one can easily provoke load shedding in a deliberate and controlled manner. Network expansion could be neglected at first, making the system effectively a 5-bus system. No storage elements or flexible demand elements need to be included in the base case test. One approach can be to ensure that there is a net power deficit at AC bus 4 and AC bus 5 throughout the considered time period, which should be the case since there are load demand at those buses but no generation. The base case test only needs to consider a single time step (e.g., one hour), making it a single-period rather than a multi-period OPF problem.

Load shedding test

The first test will verify that the slack variable used for representing load shedding in an intact grid (i.e., without considering contingencies) works as intended. This variable should get a non-zero value only if the generation is not sufficient to supply the load demand, and for these hours a cost of energy not supplied should be incurred. The model specification currently specifies multiple decision (slack) variables with associated cost parameters that could be used to represent this: Compensation for load curtailment $C_{u,t,y}{}^{nce}$, compensation for consuming less $C_{u,t,y}{}^{nce}$ and nodal load slack cost associated non-supplied load $C_{u,t,y}{}^{LL}$. The purpose of this test is to test functionality that is similar to load shedding due to contingencies before extending the implementation to actually represent contingencies. (If not correctly implemented, one could end up with load

---

[9] ENTSO-E rather uses a value of 10 000 €/MWh.

curtailment due to contingencies causing non-zero values for the wrong slack variables.) The compensation for consuming less $C_{u,t,y}^{nce}$ should only be relevant when flexible demand is present and could therefore be neglected here.

To provoke load shedding without contingencies, the easiest approach may be to reduce the power transfer capacity of the lines feeding into the bus that for which one will provoke load shedding. Bus 4 and bus 5 are connected to the rest of the network by only two branches each, making it relatively easier to control the power supply to these buses. When the power transfer capacities are reduced below a certain threshold, the variable one should get $L_{u,t,y}^{ref} - L_{u,t,y}^{lc}$ and an additional operational cost $C_{u,t,y}^{lc} \llbracket (L)_{u,t,y}^{ref} - L_{u,t,y}^{lc} \rrbracket$.

A second step in the test can be to put one of the two branches feeding the selected load bus in an outage state. This should give the same result as the test above if the power transfer limit of the remaining branch feeding the bus is set to an appropriate value for getting load shedding at the bus.

Load shedding due to contingencies

After implementing contingency constraints and the objective function term for costs of energy not supplied due to contingencies, one could replicate the load shedding tests above where two branches feeding load bus 4 (or 5) are sufficient to supply the load demand but a single branch is not. In the "base case" (intact grid) both branches feeding the load bus should be in an up state, and one of the branches should be included in the contingency list. In this case, load shedding should be represented in the solution by non-zero values for the slack variable $\Delta P_{u,c,t,y}$ in the contingency case $c = 1$ and not by any of the other slack variables, and there should be no load shedding in the non-contingency case $c = 0$.

Sensitivity tests

The following is a preliminary and roughly described list of sensitivity tests and combined tests that may be interesting to carry out on the transmission test system:

- Increasing the failure rate or mean time to repair:
  - Increasing the value of the failure rate of a branch or the mean time to repair for a branch should give changes in the objective value that can be verified analytically.
- Branches subject to outage:
  - One can specify which branches are subject to outage or in other words included in the contingency list. E.g., one could include both branches feeding the load bus under study rather than just one, or all 4 branches feeding load buses 4 and 5, or all branches in the network.
- Combining reliability modelling with transmission expansion:
  - It should be possible to combine contingency constraints with the specification of candidate branches to solve a transmission expansion problem considering contingencies. If only existing branches are subject to outage and if no branches connected to load buses 4 and 5 are candidates, this should not influence the results significantly(?)
- Reliability-driven transmission expansion:

- - If the value of lost load at load bus 4 or 5 is set sufficiently high, this should incentivize the investment in a transmission line in parallel with the lines already supplying these load buses. This should also be possible to achieve by increasing the failure rate and/or the mean time to repair for the branches.
- Candidate branches subject to outage:
  - It should be possible to specify that a candidate branch is subject to outage and included in the contingency set. If the failure rate is sufficiently high, this should disincentivize the investment in this line, requiring the investment cost to be lower for the line to be build according to the optimization model.
- Combining reliability modelling with multi-period OPF:
  - It should be possible to straightforwardly combine contingency constraints with going from a single-period OPF to a multi-period OPF formulation, e.g., a 24-period problem for a single day. If the load demand at the buses of interest (4 and 5) varies significantly during the day, it should be possible to see a corresponding variation in the solution for the load shedding (slack) variable and in the contributions to the cost of energy not supplied in the objective function.
- Combining reliability modelling with flexibility modelling:
  - It should be possible to straightforwardly combine contingency constraints with flexibility elements e.g., at the load buses 4 or 5. The interactions may be easiest to investigate by first considering a pre-installed flexibility element (i.e. not a candidate). A storage element at the bus at which we are provoking load shedding should reduce the costs of energy not supplied, but it has to be considered more closely how these interactions will play out. Similarly, a demand flexibility at these load buses should give solutions with (voluntary) curtailment of load (and possibly shifting of load) rather than (involuntary) shedding of load. (This requires that the constraints and objective function terms for the flexibility elements are replicated in the model formulation for all contingencies.) There are probably also more subtle interactions that are not anticipated at the test planning stage.

9.3.3.2 Test results

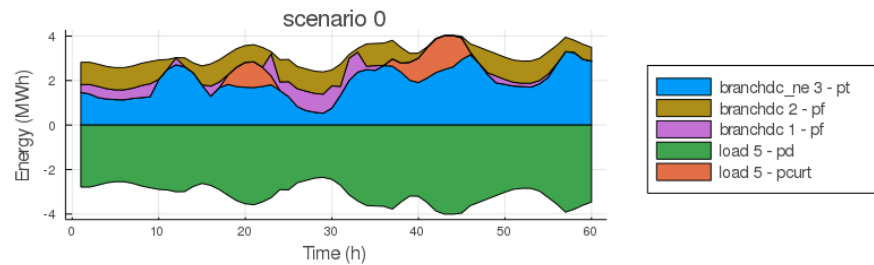| Contingency name	Outage element	Weight
Scenario 0 (base)	-	1.0

Table 8 -Base case: non-contingency state

The base case is defined such that both DC lines to AC bus 5 (DC line 1 and 2) are needed to supply the electricity demand at this bus.
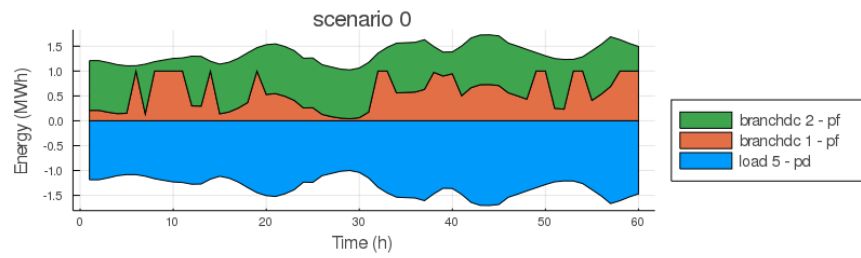
Figure 9-24 – Load profile at AC bus 5 and power flows on DC branches 1 and 2 (base case)

Doubling load results in investment in a new DC branch nr. 3 and some curtailed load.

We define two outage scenarios (plus the base case) for the lines connected to AC bus 5, as shown in the table below:
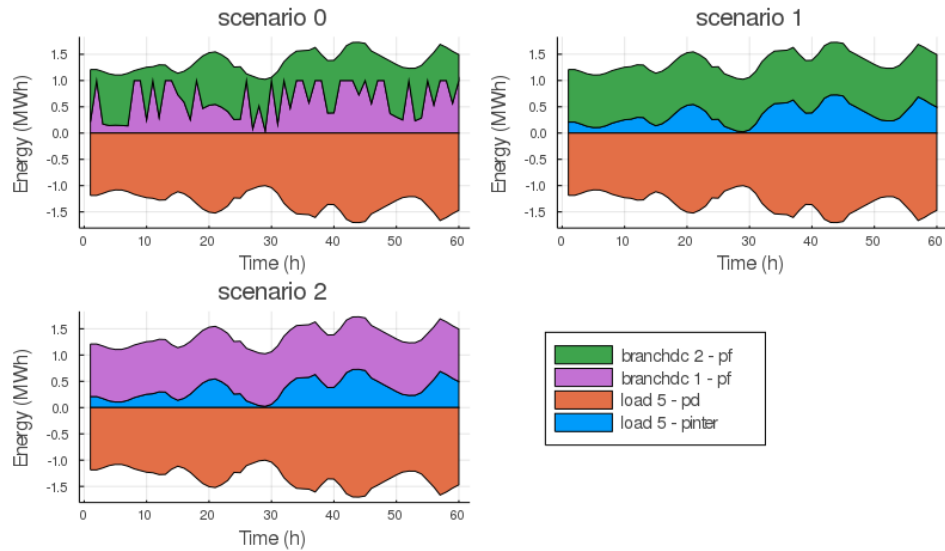

Figure 9-25 – Load profile at AC bus 5 and power flows on DC branches 1 - 3 (power demand doubled)

Contingency name	Outage element	Weight
Scenario 0 (base)	-	0.98
Scenario 1	DC branch 1	0.01
Scenario 2	DC branch 2	0.01

Table 9 - Reliability model tests: contingency states

The investment costs for new DC branch nr. 3 is set high such that the contingencies in scenario 2 and 3 cause the load slack variable ΔP^{inter} (interrupted electricity supply to the load bus during contingencies) to be activated as opposed to load curtailment (the ΔP^{lc} variable). This shows how this interrupted load demand is directly related to contingencies and serves a different purpose than load curtailment.

Figure 9-26 - Load profile at AC bus 5 and power flows on DC branches 1 and 2

(contingency scenarios – high DC branch candidate cost)

The investment cost for new DC branch nr. 3 is reduced to the original value which results in investment in this branch. The load is no longer interrupted due to the contingencies defined in scenario 1 and 2.

The load demand at AC bus 5 is doubled to check that the load slack variables work as intended simultaneously for contingencies. In the base case, load curtailment is activated as all line investments are not sufficient to supply the electricity demand. Load curtailment in the two contingency scenarios is the same as in the base case since it is related to investments and not a result of an outage of a component. No load
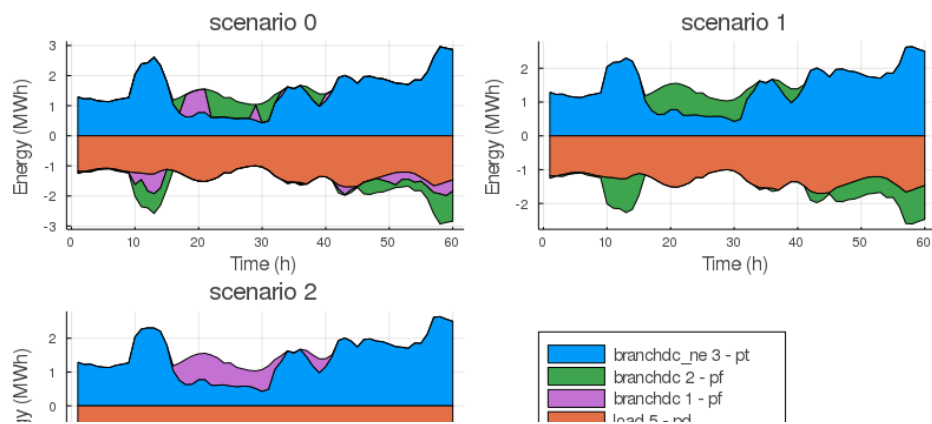
Figure 9-27 - Load profile at AC bus 5 and power flows on DC branches 1 - 3

(contingency scenarios – reduced DC branch candidate cost)

interruption is present in the base scenarios as it does not include any contingencies. Some activation of load interruption is observed for both contingency scenarios after 22-23 hours.
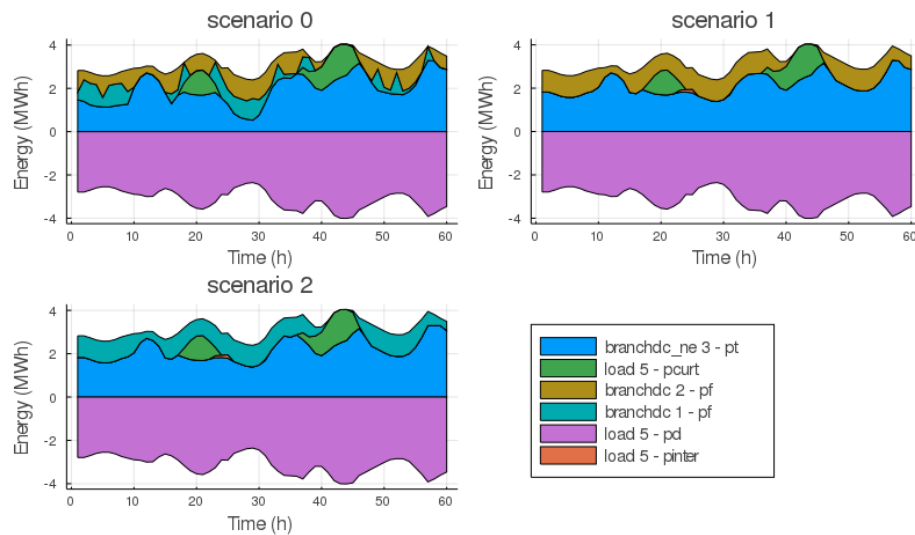
Figure 9-28 - Load profile at AC bus 5 and power flows on DC branches 1 - 3
(contingency scenarios – doubled power demand at load 5)

We define a third scenario to check that contingencies can be included for branch candidates (branches that are not built).

Contingency name	Outage element	Weight
Scenario 0 (base)	-	0.98
Scenario 1	DC branch 1	0.01
Scenario 2	DC branch 2	0.01
Scenario 3	DC branch 3 (candidate)	0.01

Table 10 - Reliability model tests: contingency states including candidate branches

If the new DC branch is subject to an outage this would lead to significant amounts of interrupted load as shown in scenario 3.
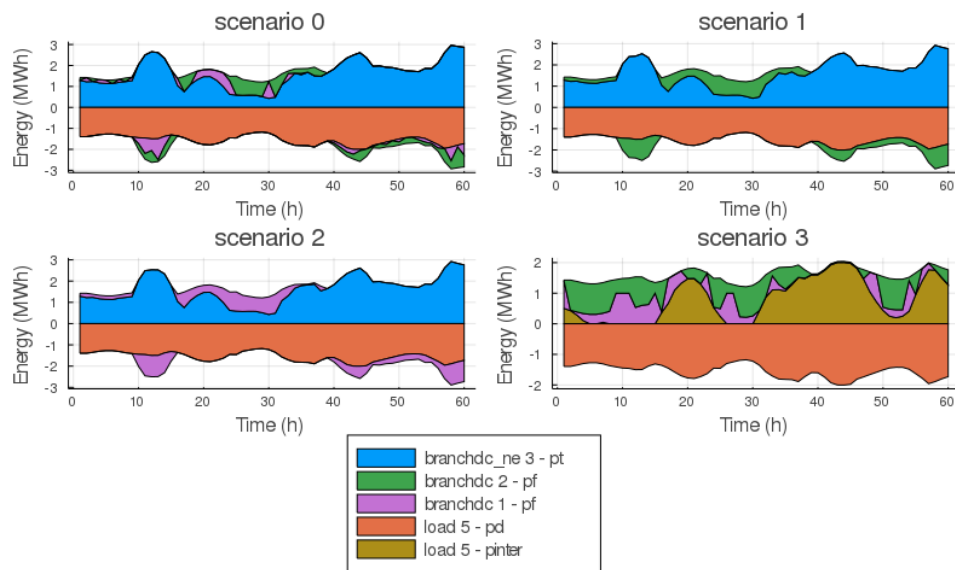
Figure 9-29- Load profile at AC bus 5 and power flows on DC branches 1 - 3
(contingency scenarios including outage of candidate DC line 3)

9.3.4 Tests on the distribution network model

9.3.4.1 Test specifications

The tests on the distribution network model aim at assessing that power flow equations are correctly implemented, in both cases where expansion candidates are present or not.

Since the time variable is not involved in power flow equations, these tests are performed through single-period optimizations, unless otherwise stated. Furthermore, storage and flexible loads are not present since specific tests are planned for these components.

Base case test

The base case test aims at running the optimization on the small distribution benchmark network described above to check if the returned solution coincides with an expected solution of the power system model described in Section 4.3.

The base case should be specified in such a way that the investment in new branches is not needed; at the same time, it should be easily tuned to trigger investments by changing the load demand at one or more buses.

The base case solution should answer the following questions:

- Does the solver return a feasible solution?
- Is the active/reactive power flowing in each branch the same at both ends?
- Is the bus active/reactive power balance satisfied?
- Are the net sums of active and reactive power generated/adsorbed in the whole network equal to 0?

Concerning the investments in new branches, the comparison of the solutions of the base case and of similar cases in which the load demand has been increased should answer the following questions:

- Are candidate branches built only if necessary?
- If an investment is made and if multiple candidates are present that satisfy the technical constraints, is the chosen candidate the least expensive one?

Sensitivity tests

The table below outlines interesting sensitivity tests that can be performed by changing the value of a parameter, or a pair of parameters, while keeping the other parameters at their base value, and verifying that the expected results are compatible with those actually obtained.

Parameter	Symbol	Description
Bus min/max operating voltage	$U_{m,t,y}^{min}, U_{m,t,y}^{max}$	Check that, if these parameters are set to the same value U, either the voltage of the involved bus also assumes the value U and the setpoints of the generators are affected by this change, or the problem becomes infeasible.
Generator min/max reactive power exchange	$Q_{g,t,y}^{min}, Q_{g,t,y}^{max}$	Check that, if these parameters are set to the same value Q, either the reactive power setpoint of the involved generator also assumes the value Q and the voltage of the nearby buses is affected by this change, or the problem becomes infeasible.
(Candidate) branch resistance/reactance	$r_{l(c)}, x_{l(c)}$	Check that, as both these parameters get closer to 0, the voltages of the two buses that are connected by the branch approach the same value.
(Candidate) branch thermal rating	$S_{l(c)}^{rated}$	Check that power flow is inhibited as this value goes to 0, possibly making the problem infeasible.
(Candidate) OLTC min/max tap voltage ratio	$\tau_{l(c)}^{min}, \tau_{l(c)}^{max}$	Check that, if these parameters are set to the same value τ, either the tap ratio of the involved OLTC also assumes the value τ and the voltage of the nearby buses is affected by this change, or the problem becomes infeasible.
Candidate branch investments costs	$I_{lc,y}$	Check that, if multiple candidates are present that satisfy the technical constraints, by increasing the cost of the built candidate enough, another candidate is chosen instead.

Table 11- Distribution model: sensitivity tests

9.3.4.2 Test results

The base case consists in the CIGRE European MV distribution benchmark network as specified in Section 9.2.2, to which a set of candidate AC branches is added, as detailed in Appendix A. When using baseline loads, a feasible solution is obtained without the need for investments in candidate branches.

All sensitivity tests were performed on this base case and produced the expected results. Where the infeasibility of the problem was among the expected results, by adequately choosing another component (bus, generator or branch) and repeating the test it was always possible to obtain a feasible solution and observe the result described in Table 11.

Tests on power flows

The obtained solutions are compliant with the model constraints: the nodal active/reactive power balances are satisfied, and the active/reactive power has the same value at both ends of the branches – as expected, since the model is lossless.

To understand the results, it is useful to remember that reactive power does not affect the cost of the solutions. The reasons for this behaviour are twofold. First, there are no terms associated with reactive power in the objective function. Second, while in reality the reactive power flowing in a line affects grid losses (therefore energy production, and therefore the overall cost), the same does not happen with the adopted model, which is lossless. So, in general, infinite optimal solutions exist to a given instance of the planning problem at hand. These infinite solutions differ in reactive power and bus voltage.

Among the practical implications, the following worth be mentioned:

● By solving the same instance with different solvers, or on different computing architectures, different solutions may be obtained.

● The bus voltage and the reactive power of the returned solutions are only one of the possible combinations of values that make the planning problem feasible: they are not necessarily similar to those values that a DSO would set in the daily operation of their own distribution network. Such values are however acceptable for planning oriented optimization.

Tests on candidate branches

By solving the problem with increasing values of specific loads (loads 1, 11, 12, and 13; one at a time) it has been verified that the candidate branches are built only if necessary (not always with the purpose of increasing the rating, but sometime as a last resort to reduce the voltage drop in heavy load conditions) and that the least expensive candidate among those that satisfy the technical constraints is chosen.

Furthermore, the same tests related to power flow as described above have been carried out in presence of built candidate branches. Finally, it has been verified that in multi-period problems, if the replacement[10] of a

---

[10] As an alternative to replacement, the addition of a new branch in parallel to the existing one is supported too (except for OLTCs, for which no parallel cases are known). The addition in parallel to an existing branch is implemented as the replacement with the parallel equivalent, so the activation/deactivation mechanism is the same.

branch is necessary, then the candidate branch is activated – and the replaced branch is deactivated – in all the optimization periods; for this specific test, the base case of the flexible loads tests was used.

9.3.5 Tests on the air quality impact model

9.3.5.1 Emission – concentration model

Before computing the *Imp(p,seas,h)* and *Prod*ref*(p,seas,h)* tables, *V* and *K*, we performed an evaluation of the robustness of the DDM results over a simplified test case including 6 power plants belonging to the Italian electrical network. The list of the 7 plants is reported in Table 12. Considering that Porto Marghera and Fusina are very close, with respect to the spatial scale of the power plant fingerprint, they will be considered as a single plant (Venezia), whose barycentre is placed in the middle of the two plants.

In order to identify the parameters of the simplified emission-concentration model, three CAMx/DDM runs were carried out. In all runs the "perturbation" was based on a variation of the emissions of the 6 PP (all stacks).

Simulations were performed over the computational domain shown in Figure 9-30, adopting a 4 km grid cell size. Simulations covered the whole 2017, but then only January and June month were used as representative of "winter" and "summer" seasons. All the following discussion refers to PM_{10}, the only air pollutant considered to evaluate health impacts and corresponding costs.

	Plant	City	Region	X [m]	Y [m]	lat	long	Fuel
1	ENEL PRODUZIONE S.p.A. (Porto Marghera)	Venezia	VEN	-25607	381468	45.45	12.26	Coal
	ENEL PRODUZIONE S.p.A. (Porto Marghera)	Venezia	VEN	-25607	381468	45.45	12.26	Coal
	ENEL PRODUZIONE S.p.A. (Porto Marghera)	Venezia	VEN	-25607	381468	45.45	12.26	Coal
	ENEL PRODUZIONE S.p.A. (Fusina)	Venezia	VEN	-26370	379324	45.43	12.25	Coal
	ENEL PRODUZIONE S.p.A. (Fusina)	Venezia	VEN	-26370	379324	45.43	12.25	Coal
2	BIOENERGIE	Argenta	EMR	-54264	292541	44.62	11.89	Biomass
	BIOENERGIE	Argenta	EMR	-54264	292541	44.62	11.89	Biomass
3	ENEL PRODUZIONE Spa - TORVALDALIGA NORD	Civitavecchia	LAZ	-67006	25094	42.13	11.76	Coal
4	ENEL PRODUZIONE S.p.A.	Brindisi	PUG	444275	-129219	40.56	18.03	Coal
	ENEL PRODUZIONE S.p.A.	Brindisi	PUG	444275	-129219	40.56	18.03	Coal
	ENEL PRODUZIONE S.p.A.	Brindisi	PUG	444275	-129219	40.56	18.03	Coal

	ENEL PRODUZIONE S.p.A.	Brindisi	PUG	444275	-129219	40.56	18.03	Coal
5	ENDESA ITALIA S.p.A.	Sassari	SAR	-350243	-103634	40.85	8.3	Coal
	ENDESA ITALIA S.p.A.	Sassari	SAR	-350243	-103634	40.85	8.3	Coal
	ENDESA ITALIA S.p.A.	Sassari	SAR	-350243	-103634	40.85	8.3	Coal
	ENDESA ITALIA S.p.A.	Sassari	SAR	-350243	-103634	40.85	8.3	Coal
	ENDESA ITALIA S.p.A.	Sassari	SAR	-350243	-103634	40.85	8.3	Coal
	ENDESA ITALIA S.p.A.	Sassari	SAR	-350243	-103634	40.85	8.3	Coal
6	Biomasse Italia S.p.a.	Strongoli	CAL	377229	-278242	39.22	17.11	Biomass
	Biomasse Italia S.p.a.	Strongoli	CAL	377229	-278242	39.22	17.11	Biomass

Table 12 – List of selected Power Plants. Some plants include more than one stack.
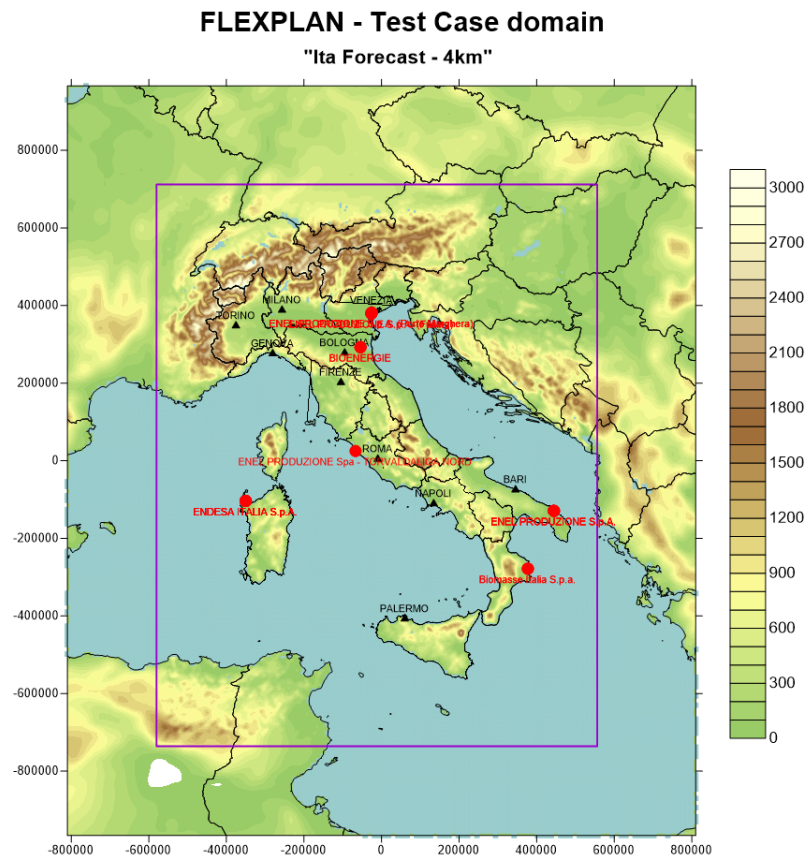
Figure 9-30 - Computational domain and position of the selected PP.

In order to validate the simplified model, we performed an additional simulation based on a brute force approach, i.e., computing concentration difference as an actual difference between the base case and a scenario simulation. The latter was defined introducing a different emission load with respect to the base case.

Particularly we considered 4 different temporal modulations, that we applied on hourly basis over the different days of the month according to Table 13. Daytime covers hours 4:00 to 18:00 included, while night-time includes the remaining hours. According to the day of month an emission variation of ± 20% either ± 60% was introduced, in order to simulate a simplified but reasonable functioning of a PP.

day	modulation	daytime	nighttime
1	P20D_M20N	0.2	-0.2
2	P20D_M20N	0.2	-0.2
3	M20D_P20N	-0.2	0.2
4	M20D_P20N	-0.2	0.2
5	P60D_M60N	0.6	-0.6
6	P60D_M60N	0.6	-0.6
7	M60D_P60N	-0.6	0.6
8	M60D_P60N	-0.6	0.6
9	P20D_M20N	0.2	-0.2
10	P20D_M20N	0.2	-0.2
11	M20D_P20N	-0.2	0.2
12	M20D_P20N	-0.2	0.2
13	P60D_M60N	0.6	-0.6
14	P60D_M60N	0.6	-0.6
15	M60D_P60N	-0.6	0.6
16	M60D_P60N	-0.6	0.6
17	P20D_M20N	0.2	-0.2
18	P20D_M20N	0.2	-0.2
19	M20D_P20N	-0.2	0.2
20	M20D_P20N	-0.2	0.2
21	P60D_M60N	0.6	-0.6
22	P60D_M60N	0.6	-0.6
23	M60D_P60N	-0.6	0.6
24	M60D_P60N	-0.6	0.6
25	P20D_M20N	0.2	-0.2
26	P20D_M20N	0.2	-0.2
27	M20D_P20N	-0.2	0.2
28	M20D_P20N	-0.2	0.2
29	P60D_M60N	0.6	-0.6
30	P60D_M60N	0.6	-0.6
31	M60D_P60N	-0.6	0.6

Table 13 - Daytime/Night-time emission variation coefficients (λ_h) applied over the different days of month

Then the full model concentration difference, computed as $\Delta C(s,h) = C_{base}(s,h) - C_{BF}(s,h)$ was compared to the corresponding concentration difference computed by means of DDM coefficients as

$$SIM_{DDM_r}(s,h) = \lambda_h \cdot Sseas_r(s,h) \cdot 100/r$$

where $Sseas_r (s,h)$ is the sensitivity coefficient for perturbation r (25, 50, 100%) computed for season *seas*, hour h, at site *s*. $\boxed{}_h$ is the emission variation coefficient defined according to Table 13.

For each power plant three sites were defined (see Table 14 and Figure 9-31) corresponding to the power plant barycentre; a cell affected by high concentration due to the plume; a cell with significant population exposed.

#	Name	Lon	Lat	Group
1	VEN1	12.059	45.264	ENEL_VEN
2	VEN2	11.867	45.416	ENEL_VEN
3	EMR1	12.201	44.407	BIO_EMR
4	EMR2	11.622	44.821	BIO_EMR
5	LAZ1	12.091	42.097	ENEL_LAZ
6	LAZ2	11.765	42.34	ENEL_LAZ
7	SAR1	8.532	40.733	ENDESA_SAR
8	SAR2	8.237	40.779	ENDESA_SAR
9	PUG1	18.131	40.393	ENEL_PUG
10	PUG2	17.806	40.558	ENEL_PUG
11	CAL1	17.039	39.159	BIO_CAL
12	CAL2	16.955	39.297	BIO_CAL
13	ENEL_VEN	12.255	45.44	ENEL_VEN
14	BIO_EMR	11.89	44.62	BIO_EMR
15	BIO_CAL	17.11	39.22	BIO_CAL
16	ENEL_LAZ	11.76	42.13	ENEL_LAZ
17	ENEL_PUG	18.03	40.56	ENEL_PUG
18	ENDESA_SAR	8.3	40.85	ENDESA_SAR

Table 14 – Receptor sites used to compare full 3D and simplified model results.
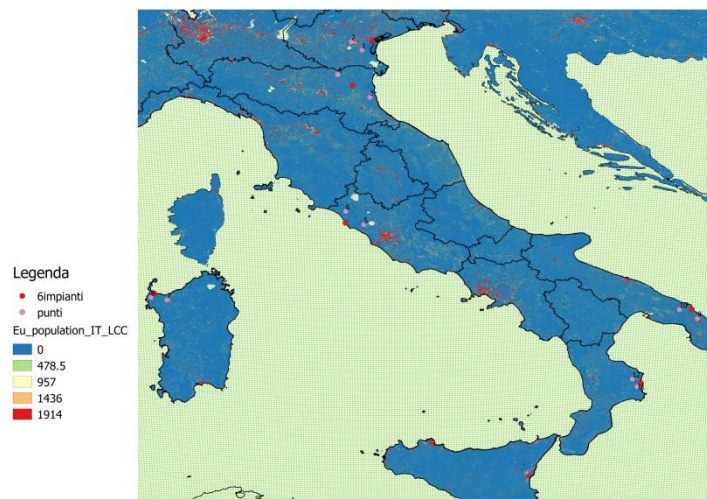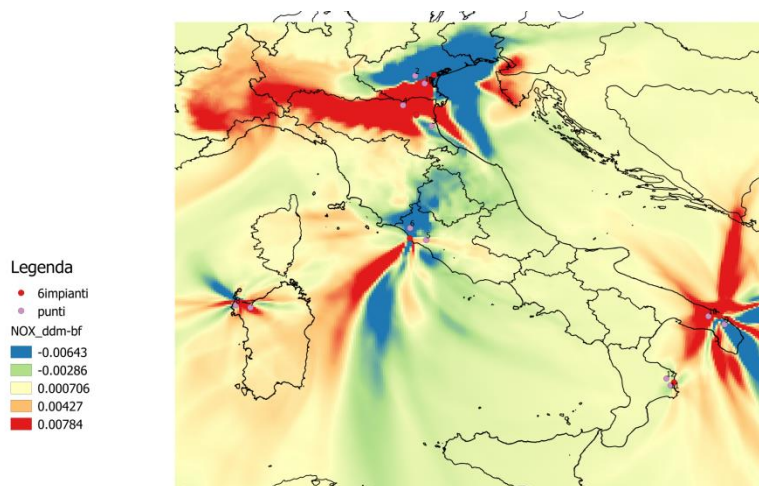
Figure 9-31 - Position of receptor sites related to population (top) and plume (bottom).

In the following graphs (Figure 9-32) simplified model results are compared to the corresponding concentration difference computed by the full 3D model that represents the "true value" to be reproduced. Figure 9-32 and Figure 9-33 refer to Venezia power plants, while additional graphs are available in Appendix B (Section 11.2.2) covering whole Italy and different technologies. For each plant, the mean day concentration differences from CAMx and SIM_DDM are compared for all perturbation coefficients, for January and June. Results refer to each site for one of the 4 temporal modulations, while for the remaining ones, results are averaged over the three sites belonging to each group.

SIM_DDM results generally show a good agreement with CAMx results, being able to capture the order of magnitude of the concentration difference as well as the hourly profile. Of course, there are several cases in which SIM_DDM shows a discrepancy with respect to the full model, but in all cases, results can be considered reasonable and robust for the inclusion in the objective function.

Finally, as already pointed out, best performance is obtained when r=100% coefficients are used.
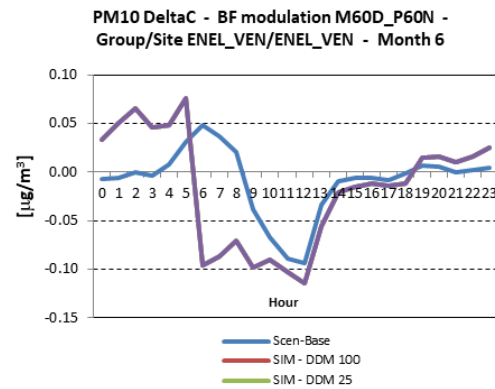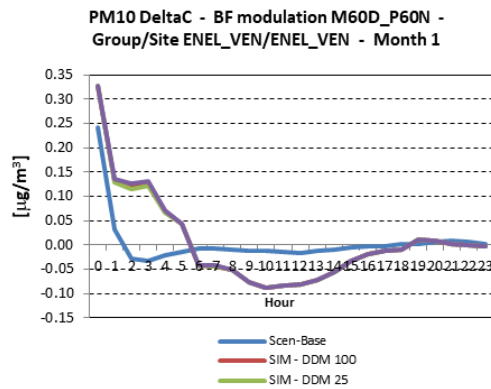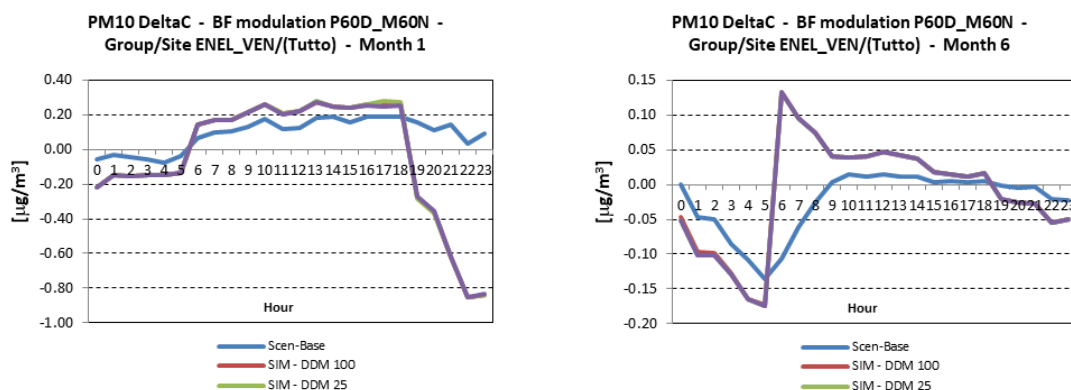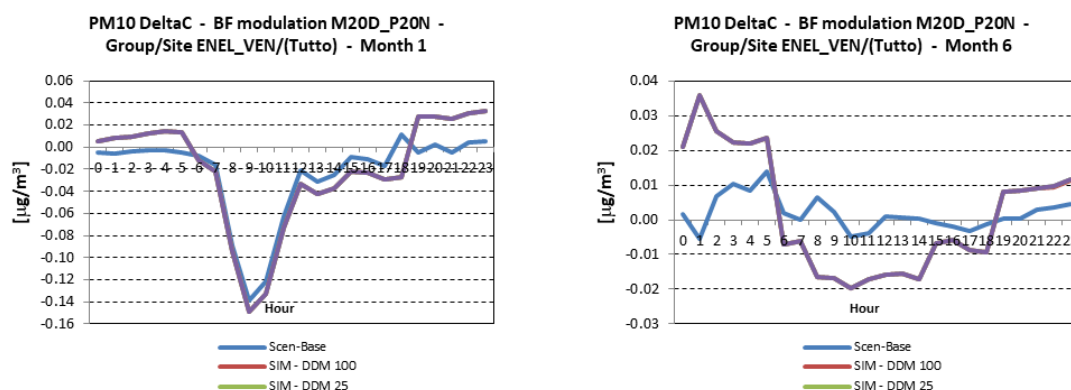
a)
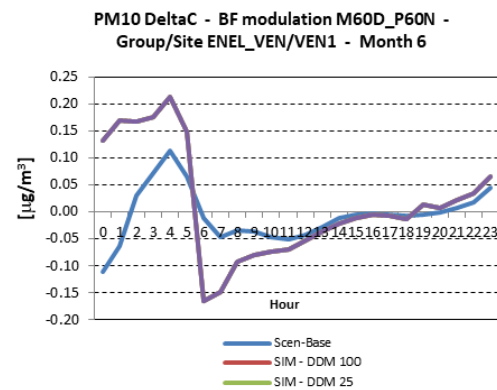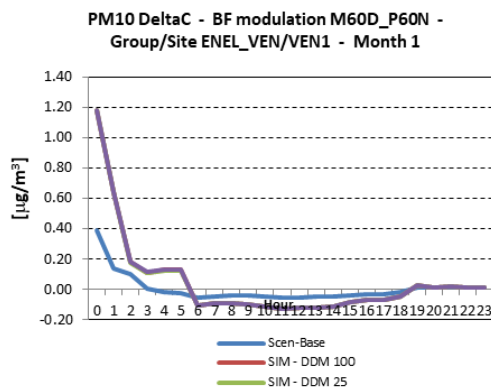
b)
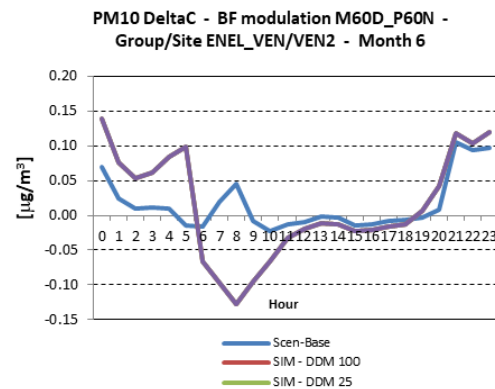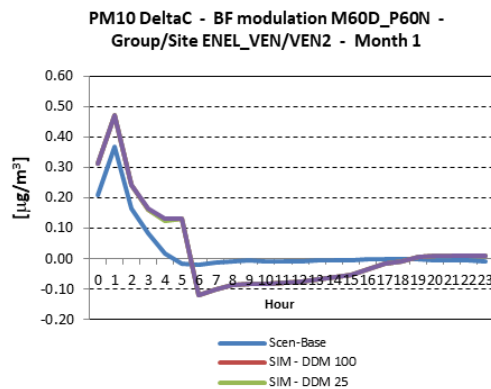
c)
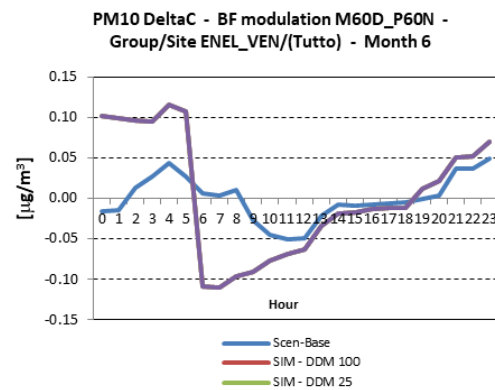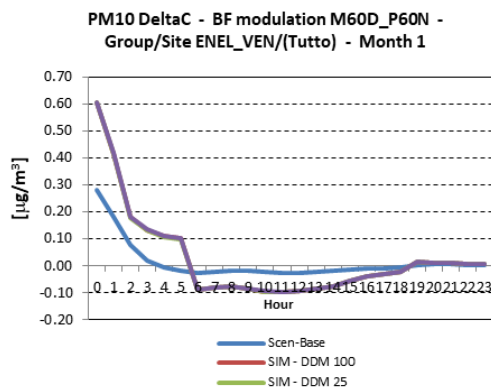
d)

Figure 9-32 – Comparison of SIM results (red, green and purple lines) against Brute Force scenarios obtained with the full 3D model (light blue line) for January (left) and June (right) mean day concentrations for ENEL PRODUZIONE S.p.A. (Porto Marghera/Fusina) power plant. Results refer to receptors: a) ENEL_VEN; b) VEN1; c) VEN2; d) average of all VEN receptors and are computed for days implementing a M60D_P60N emission modulation (see text for details)
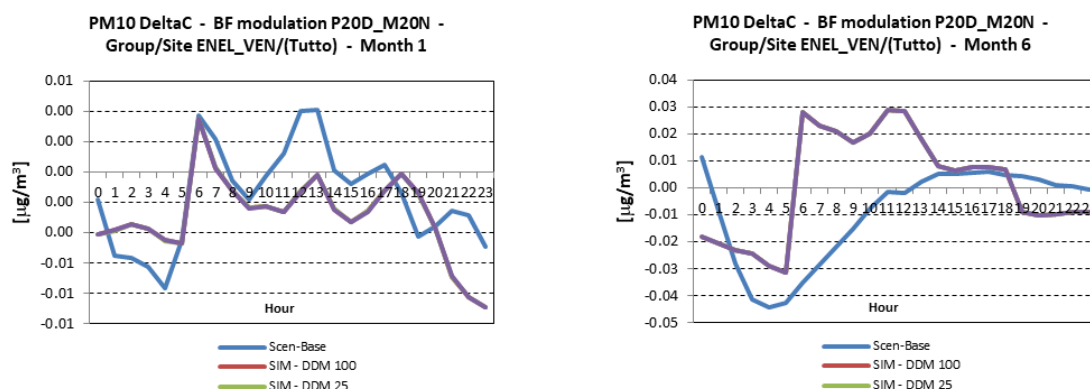
a)

b)

c)

Figure 9-33 – Comparison of SIM results (red, green and purple lines) against Brute Force scenarios obtained with the full 3D model (light blue line) for January (left) and June (right) mean day concentrations for

ENEL PRODUZIONE S.p.A. (Porto Marghera/Fusina) power plant. Results refer to scenario: a) P60D_M60N; b) M20D_P20N; c) P20D_M20N emission modulation (see text for details) and are computed as average of all VEN receptors.

9.3.5.2 Estimation of reference production coefficients

Reference production coefficients, defined in Section 6.1.2 as

$$Prod_{\text{ref}}(g,d,h) = \frac{Emi_{ref}(g,d,h)}{G_g^{ef}}$$

are computed based on proper Emission Factors G_g^{ef} [kg/MWh], representing the emission load for unit of energy produced. More generally, for each power plant g and for each pollutant p, reference hourly emission values and the corresponding emission factors used to convert those into production values should be set so that they describe the same operating conditions:

$$Prod_{\text{ref}}(g,d,h) = \frac{Emi_{ref}(g,p,d,h)}{G_{g,p}^{ef}}$$

In order to derive realistic values for $G_{g,p}^{ef}$, the annual net production data and the annual emission data for the main pollutant species (SO_2, NO_X, $PM10$) were found in the plants' mandatory environmental declarations for the reference year 2017. For each pollutant, it was then possible to estimate the average annual emission factor for each plant (Table 15).

Power plant	Production in year 2017 [MWh]	Emissions in reference year [t/year]			Average emission factors [kg/MWh]		
		SO_2	NO_X	$PM10$	SO_2	NO_X	$PM10$
ENEL_VEN	3 769 560	1 035	1 663	4	0.275	0.441	0.0011
BIO_EMR	178 850	0.33	235	0.85	0.0018	1.31	0.0048
ENEL_LAZ	11 055 000	1 889	2 519	52	0.171	0.228	0.0047
ENEL_PUG	6 068 067	1 641	3 009	92	0.270	0.496	0.0152
BIO_CAL	368 791	0.4	209	3.95	0.0011	0.567	0.011
ENDESA_SAR	3 565 000	1 765	2 271	129	0.495	0.637	0.0362

Table 15 – Production, emissions and emission factors for selected power plants.

Due to how they are derived, the emission factors refer to the actual operation of the plants in the reference year of the air quality model simulations. In contrast, the model is fed with inventory emission data, which do not necessarily refer to the same year or match the actual data. In both cases, the energy output of the individual plant is a constant that does not depend on the pollutant species. For actual emissions the production value is known, whereas in the case of inventory emissions it can only be estimated from the available data.

Power plant	SO_2	NO_X	$PM10$
ENEL_VEN	1 700.8	2 896.7	74.73
BIO_EMR	0	7.5	0.36
ENEL_LAZ	3 113.9	4 733.1	129.34
ENEL_PUG	4 141.7	5 812.6	161.04

|			
BIO_CAL	0	273.0	13.13
ENDESA_SAR	743.8	1 078.7	30.15

Table 16 – Total annual emissions (t/year) from the ISPRA 2015 emission inventory.

The annual totals were taken from the emissions database used for the simulations (ISPRA 2015 national inventory, Table 16), in order to investigate the electricity production values obtained by dividing the inventory emissions by the emission factors related to actual operation in the reference year. Values are comparable, with the exception of PM10 for ENEL_VEN power station. Due to a particularly low declared annual emission, the emission factor is much smaller than those obtained for the other plants, and consequently the production estimated from that emission data – around 70 GWh – is significantly higher than the figures based on the emissions of the other two pollutants. It was therefore decided to use the ENEL_PUG emission factor for PM10, so that the production value of ENEL_VEN plant is in line with the other calculated values (Table 17, Figure 9-34).

Power plant	Location	SO$_2$	NO$_X$	PM10	Average	Actual production
ENEL_VEN	Venezia	6 194	6 566	4 929*	5 897	3 770
BIO_EMR	Argenta		6	75	40	179
ENEL_LAZ	Civitavecchia	18 224	20 772	27 496	22 164	11 055
ENEL_PUG	Brindisi	15 315	11 722	10 621	12 553	6 068
BIO_CAL	Strongoli		482	1 226	854	369
ENDESA_SAR	Sassari	1 502	1 693	833	1 343	3 565

*Figure obtained using ENEL_PUG emission factor

Table 17 – Annual production estimates (GWh) obtained from emission inventory values by means of emission factors.
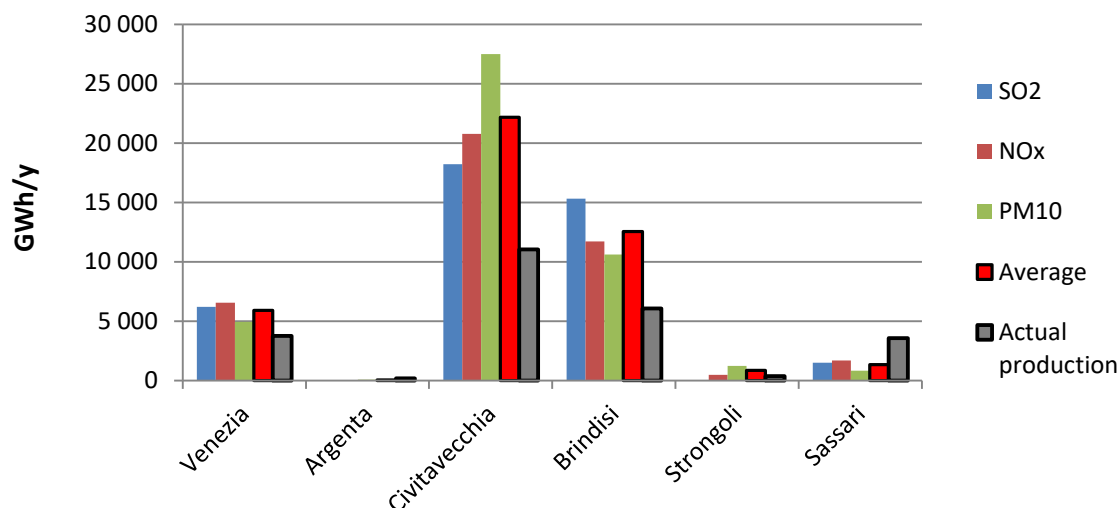
Figure 9-34 – Annual production estimates (GWh) obtained from emission inventory values by means of emission factors.

The results show that, especially in the case of coal-fired power stations, it is possible to consider the average of the three production values obtained by means of the SO_2, NO_X and PM10 emission factors as representative, in order of magnitude, of the plant's actual production. Such average values will therefore be used for $Prod_{ref}(g,d,h)$.

This allows a further step in the conceptual scheme of the relationship between a variation in energy production and the impact that variation causes on air quality around each power plant. If the emissions $Emi_{ref}(g,p,d,h)$ used for the reference simulation of the dispersion model, as it is the case in the present test, can be set so that they match the same realistic and representative production value $Prod_{ref}(g,d,h)$ of each plant, the sensitivity coefficients that this simulation computes can be used to calculate the impacts directly from the variations in production, without the need to express them explicitly in terms of the (pollutant-dependent) emission values and emission factors.

The last operation needed to obtain the reference hourly production coefficients $Prod_{ref}(g,d,h)$ is to apply temporal modulation profiles to the annual values. A common practice in emission processing for dispersion models is to use three different coefficients in chain, the first picked from a 12-values profile to disaggregate months, the second from a 7-values profile to disaggregate days of week, and the third from a 24-values profile to disaggregate hours of day. Frequently, the profiles do not depend on the pollutant as they describe the overall activity of the emitting source, so they can be validly used to disaggregate annual production values in the same way. Values used for the power plants used in the test are reported in Table 18, Table 19 and Table 20 .

J	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
0.0899	0.0899	0.0874	0.0849	0.0774	0.0758	0.0783	0.0758	0.0791	0.0891	0.0874	0.0849

Table 18 – Coefficients of the year-to-month time profile.

Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
0.149	0.150	0.153	0.155	0.152	0.126	0.116

Table 19 – Coefficients of the week-to-weekday time profile.

Hour	Weekday	Sat-Sun
1	0.0405	0.0401
2	0.0405	0.0401
3	0.0405	0.0401
4	0.0405	0.0401
5	0.0405	0.0401
6	0.0405	0.0401
7	0.0417	0.0417
8	0.0417	0.0417
9	0.0428	0.0432
10	0.0428	0.0432
11	0.0428	0.0432
12	0.0428	0.0432

13	0.0428	0.0432
14	0.0428	0.0432
15	0.0428	0.0432
16	0.0428	0.0432
17	0.0428	0.0432
18	0.0428	0.0432
19	0.0417	0.0417
20	0.0417	0.0417
21	0.0405	0.0401
22	0.0405	0.0401
23	0.0405	0.0401
24	0.0405	0.0401

Table 20 – Coefficients of the day-to-hour time profile.

As discussed in Section 6.1.2, $Prod_{ref}(g,d,h)$ as can be computed as a weighted average of the corresponding seasonal values, i.e.:

$$Prod_{ref}(g, d, h) = wg(\text{w}, d) \cdot Prod_{ref}(g, \text{w}, h) + wg(\text{s}, d) \cdot Prod_{ref}(g, \text{s}, h)$$

where w stands for winter and s stands for summer, corresponding to the two chosen reference months of January and June respectively, as defined in Section 9.3.5.1. For each power plant, two sets of 24 seasonal $Prod_{ref}(g,seas,h)$ values ($seas$ = w,s) were computed, representing the average day of the corresponding month, i.e. the mean of 31 (Jan) or 30 (Jun) $Prod_{ref}(g,d,1)$ values associated to hour 1, the mean of 31 or 30 $Prod_{ref}(g,d,2)$ values associated to hour 2, and so on.

The resulting values are shown in Table 21 for Venezia power plants, while additional tables are available in Appendix B (Section 11.2.2) referring to the remaining plants.

Hour	Winter	Summer
1	711.52	591.51
2	711.52	591.51
3	711.52	591.51
4	711.52	591.51
5	711.52	591.51
6	711.52	591.51
7	733.55	609.70
8	733.55	609.70
9	755.59	627.90
10	755.59	627.90
11	755.59	627.90
12	755.59	627.90
13	755.59	627.90
14	755.59	627.90
15	755.59	627.90

16	755.59	627.90
17	755.59	627.90
18	755.59	627.90
19	733.55	609.70
20	733.55	609.70
21	711.52	591.51
22	711.52	591.51
23	711.52	591.51
24	711.52	591.51

Table 21 – $Prod_{ref}(g,seas,h)$ values [MWh] for ENEL_VEN power plants.

9.3.5.3 Estimation of reference impact coefficients

This section presents the results of the impact estimation $Imp_{ref}(g, seas, h)$ for the six plants listed in Table 12, selected for the present case study starting from the population data and the sensitivity coefficients in a neighbourhood of each plant according to equation (31).

$$Imp_{ref}(g, seas, h) = \sum_{i,j \in N_g} Sseas_{100}(h, i, j) \cdot pop(i, j) \tag{31}$$

Where $Sseas_{100}(h, i, j)$ is the hourly change in concentration in cell i, j caused by the 100% change in plant emission, pop(i, j) is the interpolated resident population in cell i,j. The impact $Imp_{ref}(g, seas, h)$ was estimated for two typical days, one in winter (January) and the other in summer (June). The area of influence was defined on the basis of a preliminary assessment of the actual impact of the thermoelectric plants and in this first application case a neighbourhood of 91 cells (9 x 9) of the CAMx model with 4 km of resolution was chosen, equal to an area of approximately 1300 km^2. Table 22 contains the coordinates of the center of the CAMx cell in which the plant falls and the population residing in the defined area surrounding each plant. The most populated plant in its surroundings is that of Porto Marghera / Fusina located in the city of Venice, the least populated one is Biomasse Italia S.p.a. located in the city of Strongoli (KR).

PLANT	Resident population	CITY	REGION	LAT	LON
ENEL PRODUZIONE S.p.a. (Porto Marghera/Fusina)	583912	Venezia	VEN	45.43	12.25
BIOENERGIE	168244	Argenta	EMR	44.62	11.89
Biomasse Italia S.p.a.	86027	Strongoli	CAL	39.22	17.11
ENEL PRODUZIONE Spa - TORVALDALIGA NORD	94101	Civitavecc	LAZ	42.13	11.76
ENEL PRODUZIONE S.p.A.	238272	Brindisi	PUG	40.56	18.03
ENDESA ITALIA S.p.A.	112277	Sassari	SAR	40.85	8.3

**Table 22** –Total population living within the area of influence of each power plant.

Figure 9-35 shows the resident population data interpolated on the CAMx grid in the area of influence. With the exception of the Bioenergie plant located in Argenta (FE), the other plants are located in coastal areas, therefore the impact calculation takes place only in the land cells and not on the sea (grey cells).

Venezia

Argenta

Strongoli

Civitavecchia

Brindisi

Sassari
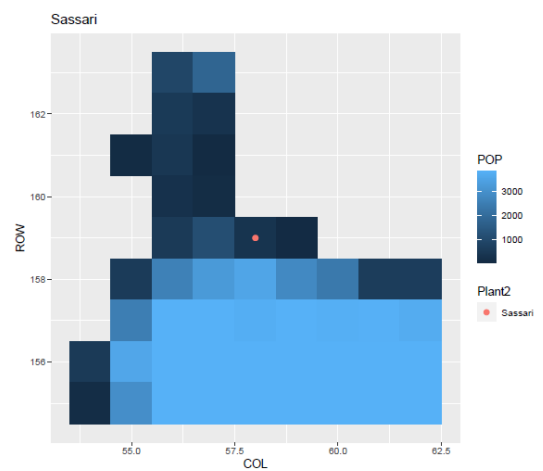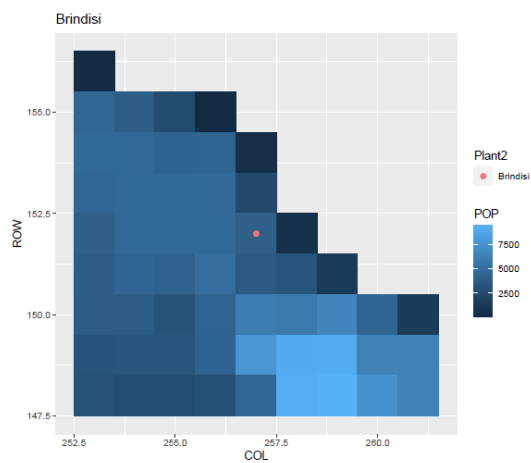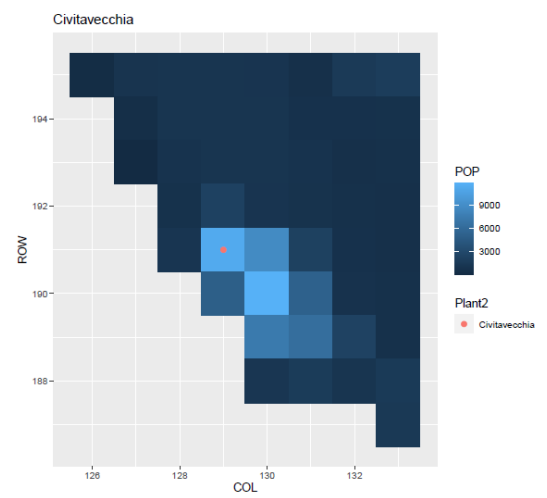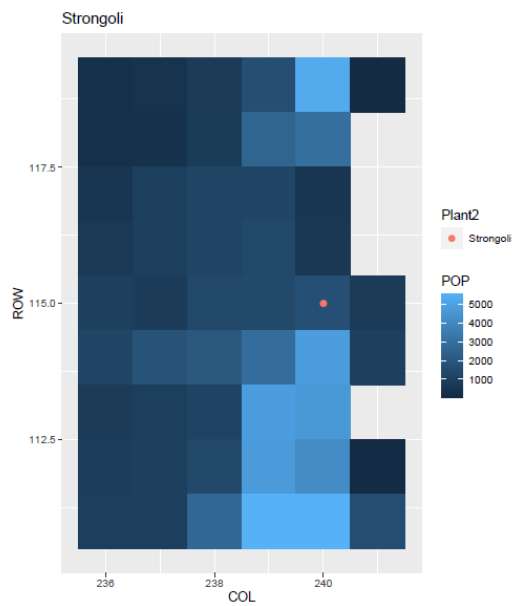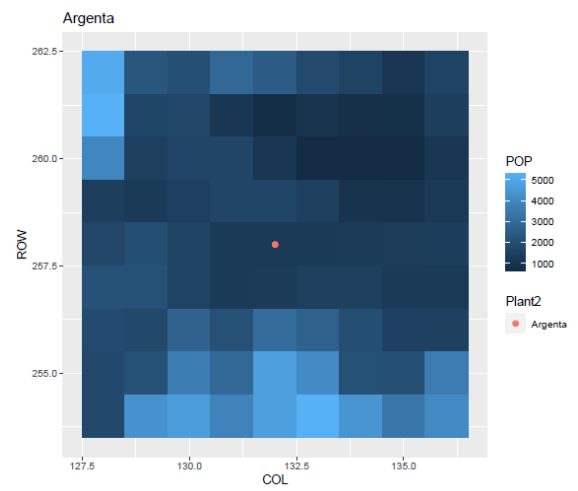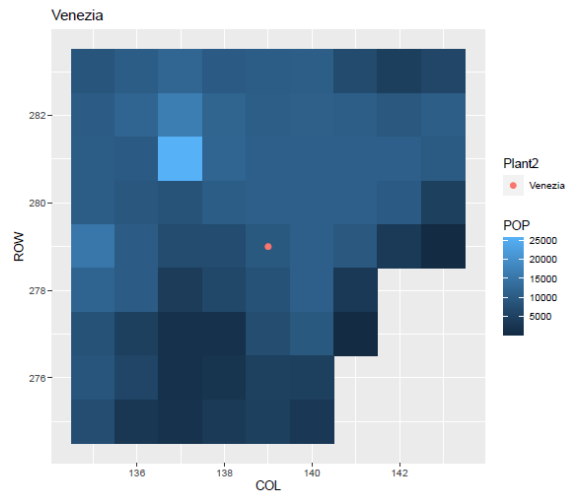
Figure 9-35 – Resident population in the area of influence of each plant interpolated on the CAMx model grid at 4 km.

Figure 9-35 shows the hourly impact coefficients calculated around the six plants, in the two typical days of January and June. The values obtained seem reasonable overall and reflect both the operating profiles of the plants (eg double winter peak in the Argenta plant) and the influence of meteorology. It can in fact be noted that for the two plants located in Northern Italy (Venice and Argenta) the impact is greater in January, due to the typical stable weather conditions in this area. The trend is opposite in the power plants located in central-southern Italy and located near the coast. Probably also in this case the meteorological factor can be decisive. Similarly to the reference production data ($Prod_{ref}$), the values of the typical impact days (Imp_{ref}) are then used to feed the impact model of the six plants selected for the case study, included in the objective function.
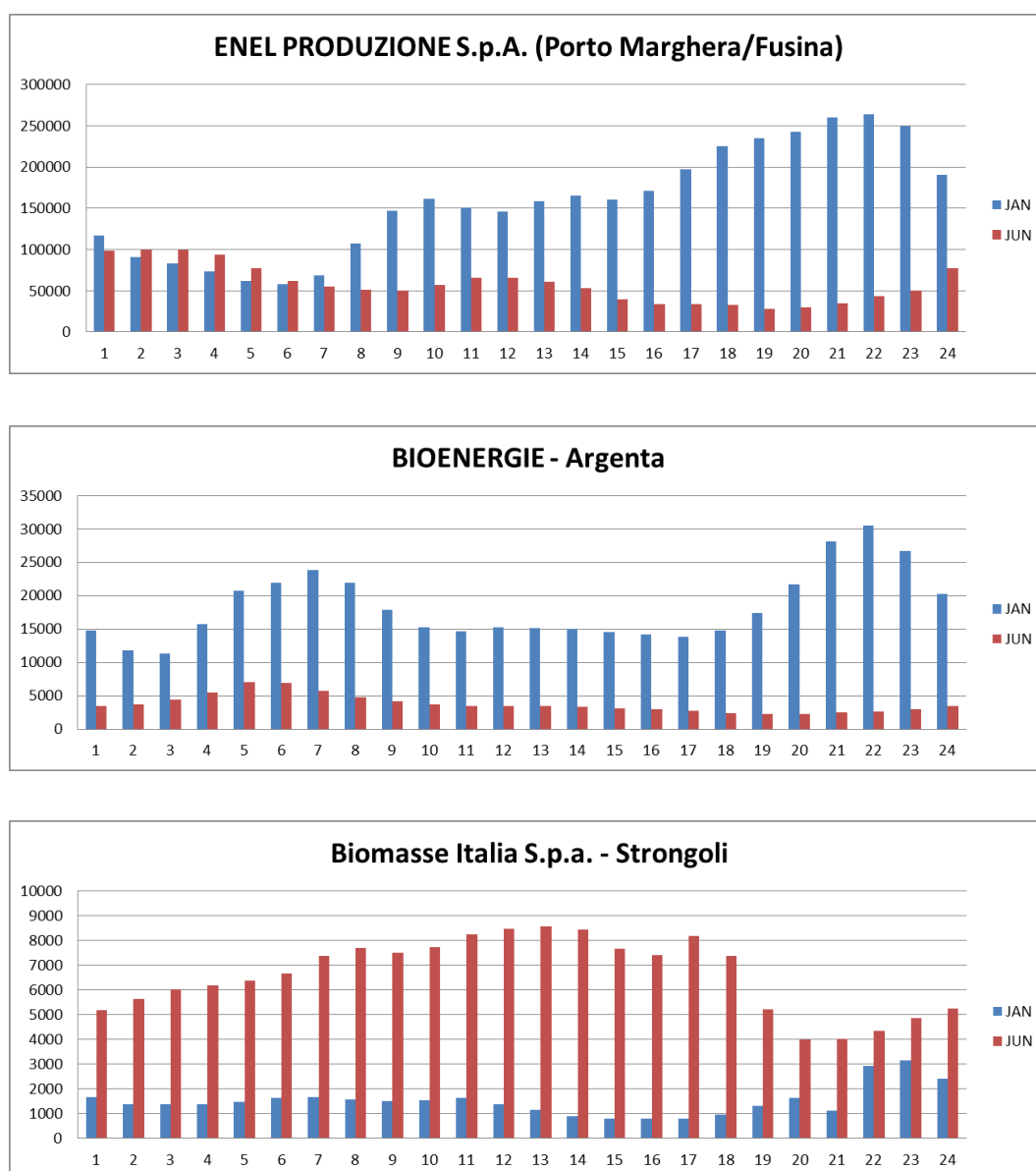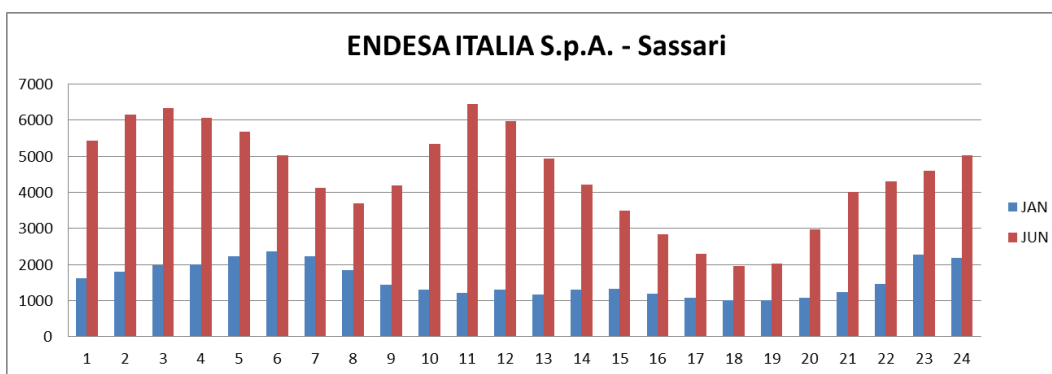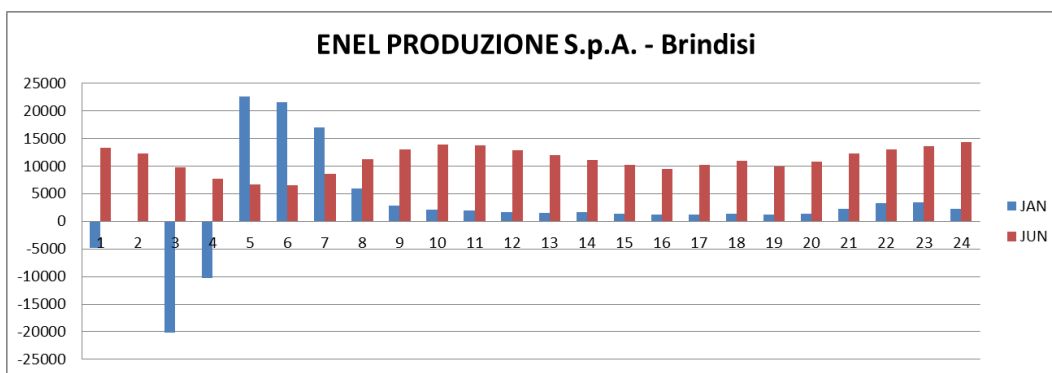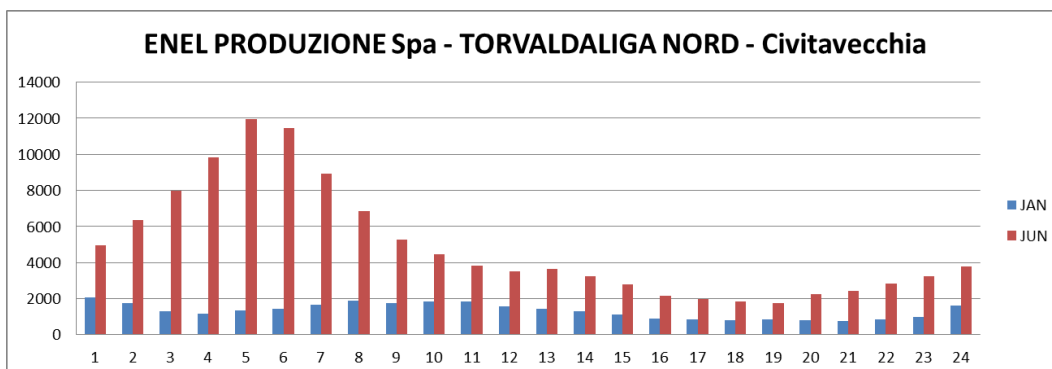
Figure 9-36 – Hourly impact Imp_{ref} [µg/m³ x pop] estimated for each of the six plants, for the months of January (blue) and June (red).

9.3.6 Tests on landscape impact model

9.3.6.1 Test data

The optimal routing model as shown in section 6.3 has been implemented in Matlab software for the proof-of-concept testing. For the quantification of the landscape impact of the transmission assets, the spatial

information as shown in Figure 9-37 is used. The blue crosses mark locations of substations and the coloured areas mark areas with different landscape characteristics. The spatial information is provided in a resolution of 2,5 km and contains the categorization into the areas as shown in Table 23. The used spatial weights to represent area dependent installation cost factors (ICF) of transmission equipment are provided in Table 23 per technology option considered. Note that a specific point on the map can belong to multiple areas at the same time, such as an already existing highway crossing a natural area or city and so on. Given the characteristics of the data, following sensitivities are considered in the testing process:

- Representation of multiple overlapping areas
  - o Using the average ICF of a certain location
  - o Using minimum ICF of a certain location
- Possibility of system expansion with overhead lines
  - o Strategy 1: Allowed in all spatial areas
  - o Strategy 2: Allowed only on existing infrastructure corridors
  - o Strategy 3: Not allowed (only underground cables allowed)

As such, there are for each candidate connection 6 possible ways of determining the landscape impact, due to the characteristics of the data. The tests are performed for the set of candidate AC and DC connections provided in Table 46 (Appendix A). It is noted that these values are arbitrarily chosen to demonstrate the working principle of the optimal routing model.
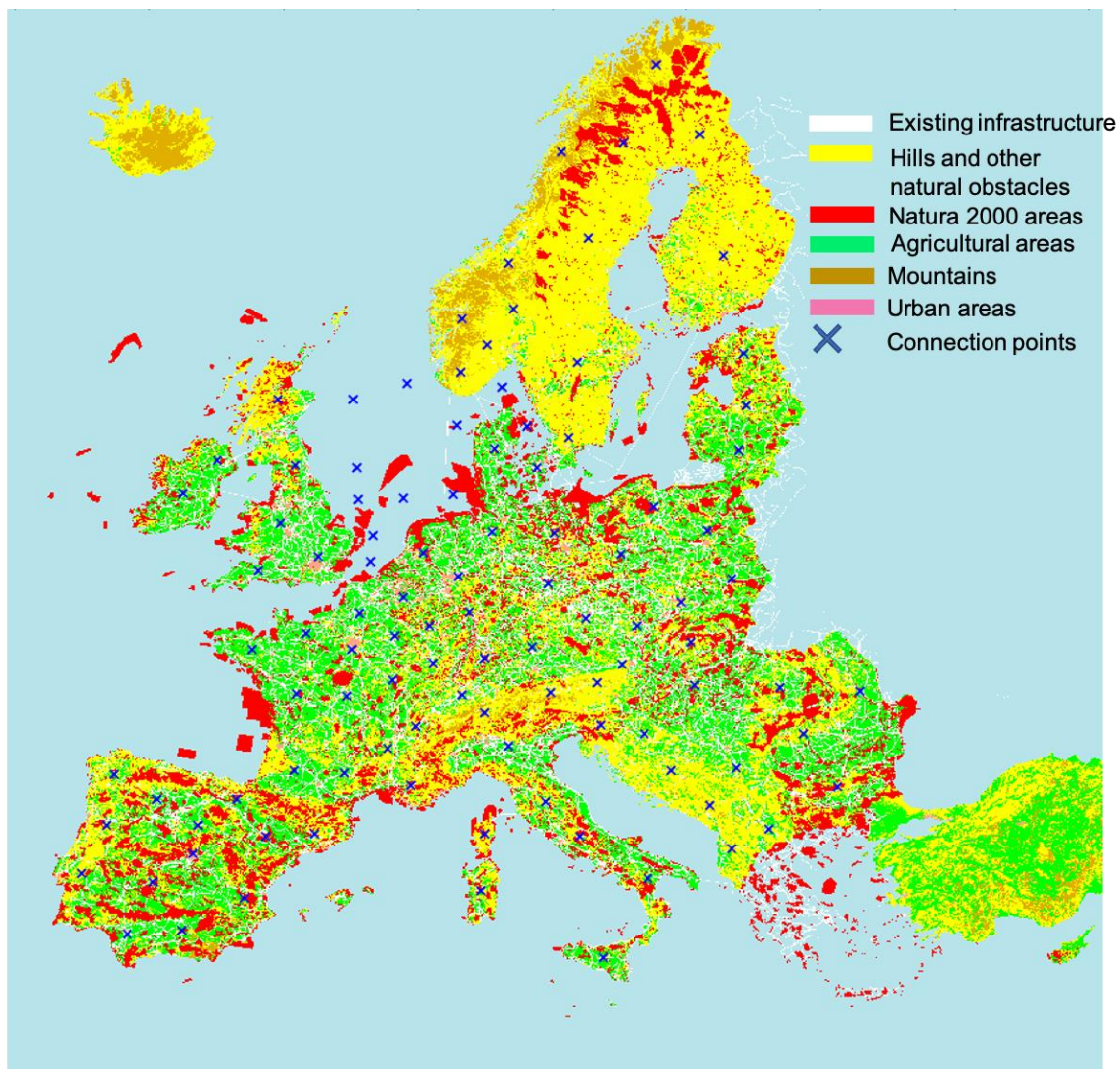
Figure 9-37 - Spatial installation categories for landscape impact costs

Spatial category	**ICF AC overhead line**	**ICF AC underground cable**	**ICF DC overhead line**	**ICF DC underground cable**
Agricultural area	1	1	1	1
Existing grid infrastructure	1	1	1	1
Road infrastructure	1	1	1	1
Railroad infrastructure	1	1	1	1

|				
Hills and other natural obstacles	2.5	1	2.5	1
Mountain areas	10	4.5	10	4.5
Urban areas	40	2.5	40	2.5
Natura 2000 protection areas	40	2.5	40	2.5
Submarine	N.A	0.75	N.A	0.75

Table 23 - Installation cost factors per technology per spatial category

9.3.6.2 Test results

In the following paragraphs the sensitivities with respect to the choice of the ICF (minimum vs average) and the possibility of the expansion with overhead lines (all areas allowed vs only on existing corridors) are analysed. Figure 9-38 shows the optimal transmission route obtained if the minimum spatial weight for overlapping installation areas is used. The methodology is applied to the candidate AC connection between nodes 1 and 3 according to Table 46 (Appendix A). Strategy 1 is used as an expansion option, which states that overhead lines can be installed in all areas, according to the spatial weights defined in Table 23. We observe that in this case a pure overhead line solution (circular markers) is chosen, which can also cross nature protected areas (red zones), as the spatial data provided in this region is overlapping. Figure 9-39 shows the change in the transmission route, if the average spatial weight of the overlapping areas is used. We can observe in this case that the transmission route is different and avoids nature protected areas.
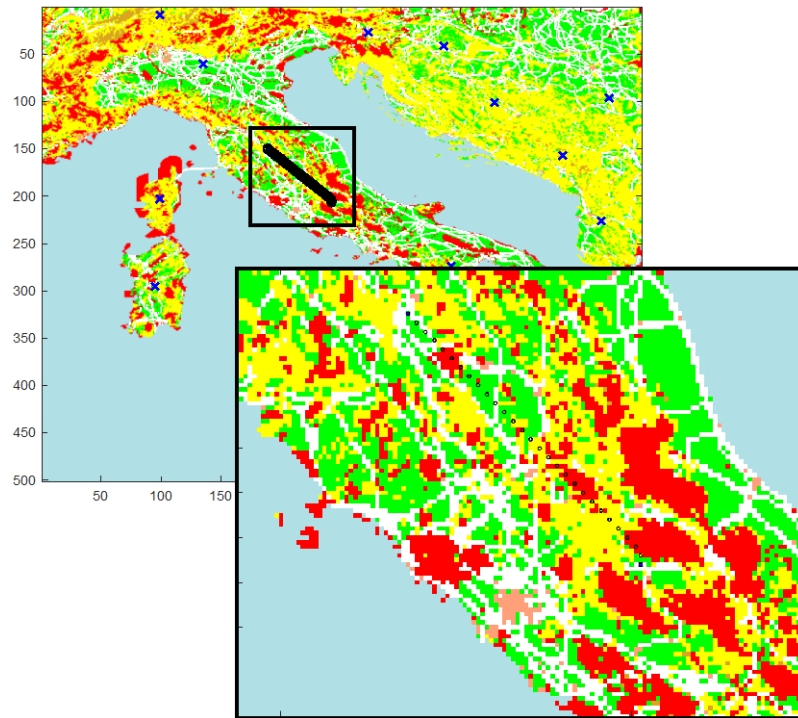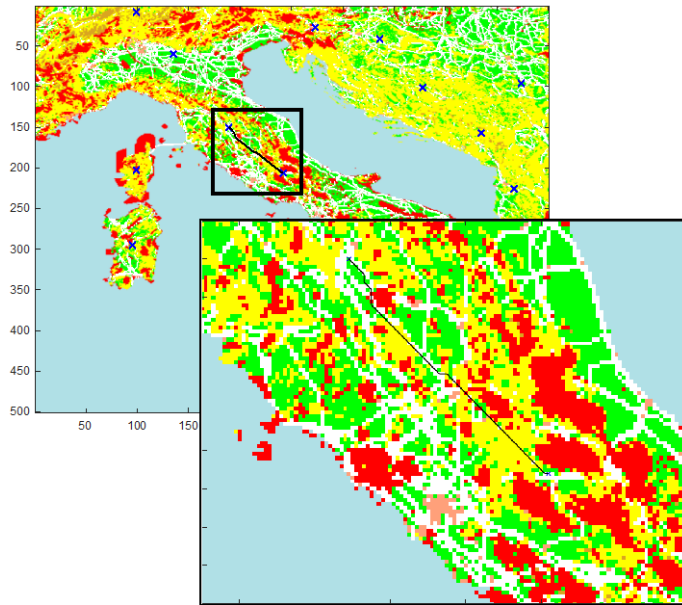
Figure 9-38 - Optimal transmission route for candidate AC connection between node 1 and 3, with usage of minimum spatial weights.
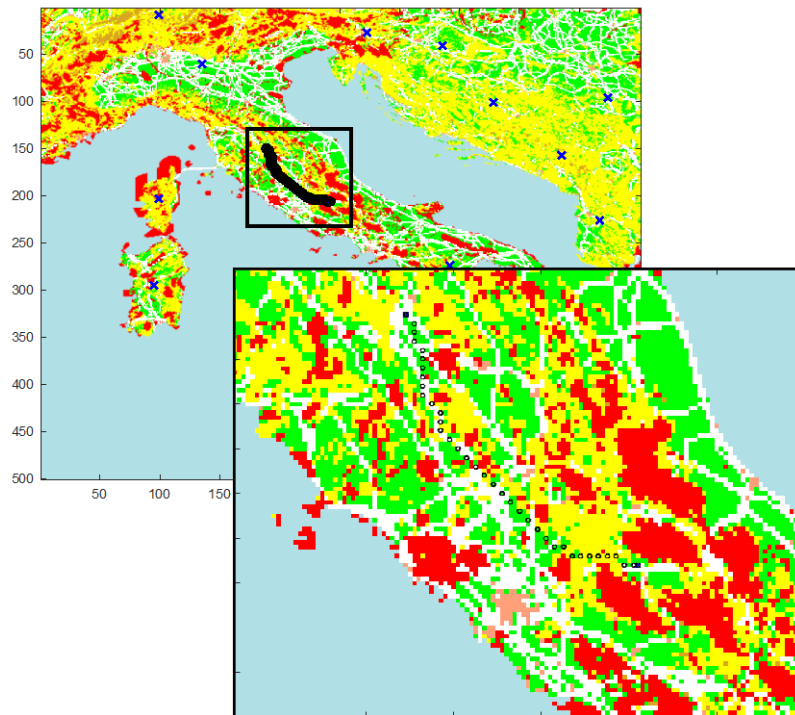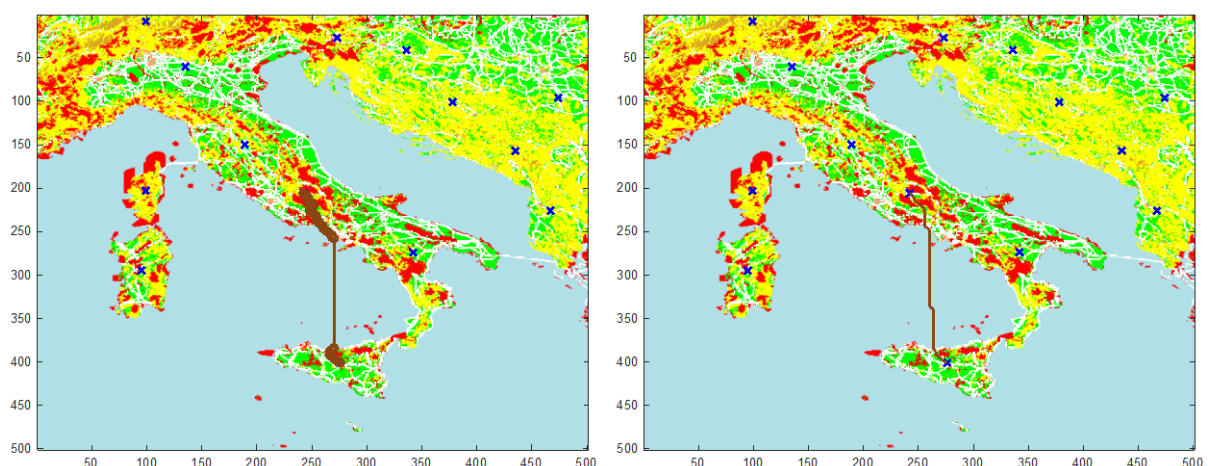

Figure 9-39 - Optimal transmission route for candidate AC connection between node 1 and 3, with usage of average spatial weights.

If strategy 2 (overhead lines only allowed at existing infrastructure corridors) is used in combination with the minimum spatial weight approach, the same transmission route as in Figure 9-38 is obtained, as in this particular region existing infrastructure areas are already present in protected or urban areas. Nevertheless, if the average spatial weight is chosen, the solution in Figure 9-40 is obtained, which favours underground cables (solid lines) instead of overhead lines.

Figure 9-40 - Optimal transmission route for candidate AC connection between node 1 and 3, with usage of average spatial weights if overhead lines only allowed at existing infrastructure corridors.

A similar effect is observed in the case of candidate HVDC connections as shown in Figure 9-41. For the candidate link between nodes 3 and 6, a combination of HVDC overhead lines and submarine cables is used for

Figure 9-41 - Optimal transmission routes for the HVDC candidate between nodes 3 and 6, using average spatial weight for strategy 1 (left) and strategy 2 (right).

strategy 1, whereas a fully HVDC cabling option is chosen for strategy two, as the landscape impact of overhead lines is estimated to be high.

Table 24 provides an overview of the total costs for the different candidates, calculated for a power rating of 1000 MVA / MW for all possible candidates, where for the HVDC options, the costs of HVDC converters is included. All costs are provided in M€. We can clearly observe that for the total costs of the candidates the choice of overhead line vs underground cabling options has a major influence. Depending on the choice of the spatial weights of overlapping areas, a shift from overhead lines to underground cables is observed, reflected in a large increase in the costs of AC candidates. For HVDC candidates this effect is less pronounced, mainly due to the fact that submarine cables need to be used in any case.

Candidate	Strategy 1		Strategy 2		Strategy 3	
	min. weight	avrg. weight	min. weight	avrg. weight	min. weight	avrg. weight
AC node 1-3	177.4	214.0	177.4	569.9	556.5	569.9
AC node 3-4	286.1	329.0	286.1	915.5	905.0	915.5
AC node 4-6	460.4	577.8	460.4	1331.1	1275.8	1331.1
DC node 3-6	602.1	711.1	602.1	851.9	915.2	851.9
DC node 4-6	453.6	515.7	453.6	692.4	727.9	692.4
DC node 5-6	779.1	753.5	779.1	892.1	969.3	892.1

Table 24 - Summary of the total costs for the expansion candidates for the different expansion strategies and choice of spatial weights for overlapping installation areas.

9.3.7 Tests on the scenario generation and reduction

9.3.7.1 Test specifications

Deliverable D1.1 describes the methodology on how scenarios will be generated within the FlexPlan project. The developed methodology is used to first generate a large variety of nodal generation and demand scenarios in terms of hourly time series, respectively. The generated time series are further reduced to a representative set of time series which are used as input for the planning tool itself.

The intermittent generation from variable renewable energy sources and the electricity demand are considered as stochastic inputs with respect to the grid expansion planning problem. The planning approach incorporates storage and demand flexibility as alternatives to classical grid expansion. Thus, time series data is required as input for the advanced planning tool in order to accurately represent the intertemporal constraints linked to the operational characteristics of demand flexibility and storage. As such, scenarios are provided as hourly time series for all stochastic inputs of the planning problem at hand.

As described in D1.1, the methodology uses pan-EU macro-scenarios as an input to generate operational scenarios for the stochastic inputs, namely vRES and electricity demand. Subsequently, the generated

operational scenarios are reduced by a clustering approach. The developed methodology consists of the following steps:

- A time series generator which samples a broad variety of possible realizations of vRES and hydropower generation based on historical meteorological and hydrological data
- A time series generator which samples temperature-sensitive load profiles based on sensitivity factors depending on outdoor ambient temperature
- A clustering algorithm, identifying and selecting a reduced subset of relevant operational scenarios.

This methodology was applied to the IEEE 6-bus test system, applied to Italy. Wind, PV and demand profiles were generated for each network node for 40 climate years. Of the 40 generated years, there are 35 years with data on demand and PV as well as wind, all correlated on time. Those 35 correlated years are used as input for the scenario reduction.

To test the impact of scenario reduction as well as shorter/longer scenario inputs on the planning tool results and decisions, the overall set of 35 yearly time series was reduced using different cluster sizes (leading to a different number of Monte Carlo samples for the planning problem), as well as different clustering methodologies and different scenario lengths. An overview of the generated scenarios is shown in Table 25. The obtained scenarios are fed into the planning problem, the impact of the different scenarios on the resulting optimisation decisions is discussed below.

Timeseries length	# Monte Carlo samples/ #clusters	Note
year	1	Random pick
	5	Features reduced to 'main characteristics'
	35	Full scenario set
month	2	Features reduced to 'main characteristics'
		Features reduced with PCA
	4	Features reduced to 'main characteristics'
		Features reduced with PCA
	6	Features reduced to 'main characteristics'
		Features reduced with PCA
week	2	Features reduced to 'main characteristics'
		Features reduced with PCA
	4	Features reduced to 'main characteristics'
		Features reduced with PCA
	6	Features reduced to 'main characteristics'
		Features reduced with PCA
day	4	Features reduced with PCA
	6	Features reduced with PCA

	8	Features reduced with PCA

Table 25 - Overview of generated reduced scenarios: length of timeseries, number of Monte Carlo samples, and information on how the scenarios have been generated from the overall scenario set.

To test the scenario reduction approach, the different load/generation profiles, as stated above, were applied to the test system. The questions the tests need to solve are:

- Are the optimal decisions as calculated by the planning tool the same for the full scenario set as the reduced scenario set?
- What is the minimal reduced scenario size for which equally reliable results are obtained?
- Which cluster size gives best results, or what is the minimal to use cluster size to obtain reliable results?
- Which feature reduction technique performs best, w.r.t. being able to capture the necessary characteristics of the scenarios to compute.

9.3.7.2 Test results

The testcase that was at a first instance considered in the scenario reduction tests is similar to the base test case, except for the investment candidates, these are:

- 3 storage candidates
- 1 AC branch candidate
- 2 DC branch candidates
- 4 converter candidates.

More information on the testcase, and more specifically the candidates is given in Appendix A (Section 11.1.6). In this testcase the loads are not considered as being flexible.

The testcase is such that not enough generation is available at all times to cover all load, especially not during night-time hours (when there is no PV generation). At least one storage element needs to be installed to cover the load during these hours. It is the aim of the optimisation to find out which storage candidate(s) to install best, and at the same time to decide which other infrastructure is required to dispatch all generation to all loads (which branches/converters).

Figure 9-43 shows the resulting objective for each scenario year of the full scenario set of 35 weather years. In this figure, the investment costs are indicated in blue, the generator operational costs are indicated in green, and load shedding costs are indicated in orange. It is clear from this figure that load shedding is only applied in 2 scenario-years (year 2 and year 31). Also, generation operational costs are much higher than the investment costs.
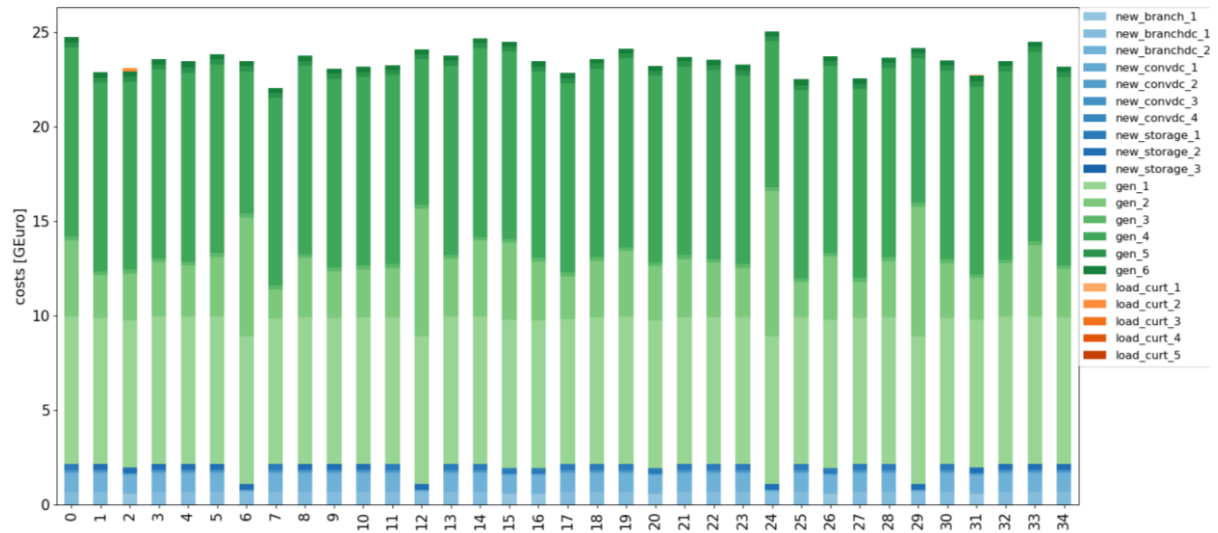
Figure 9-43 - Overall calculated optimal objective for the full scenario set of 35 years. Investment costs are indicated in blue, operational generation costs are indicated in green, and load shedding costs are indicated in orange.
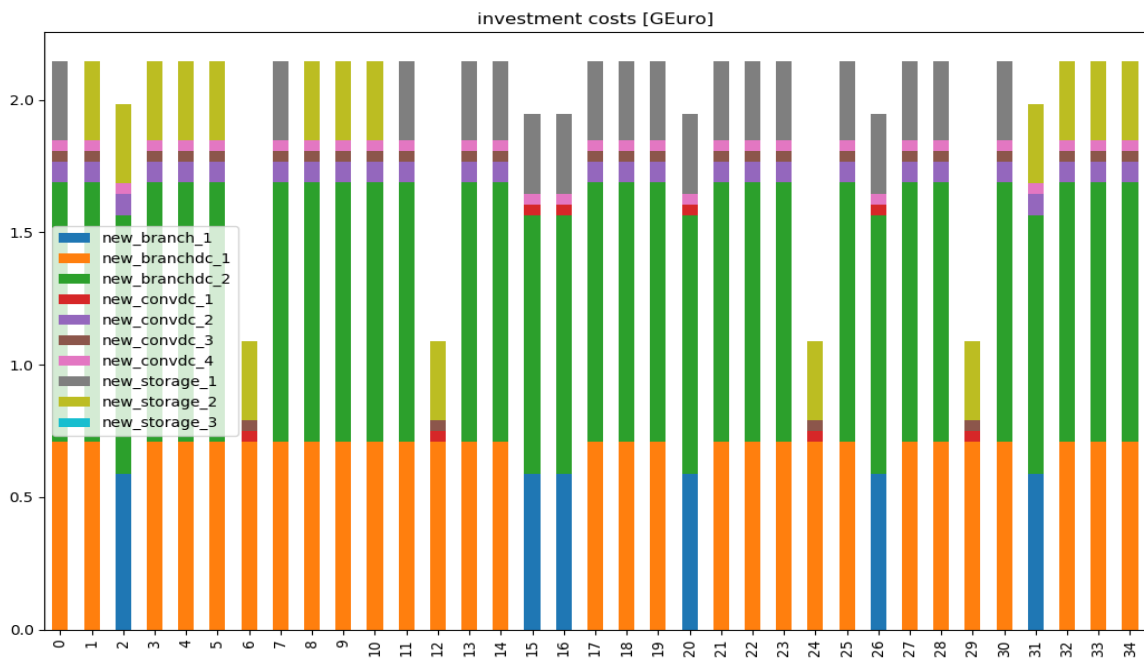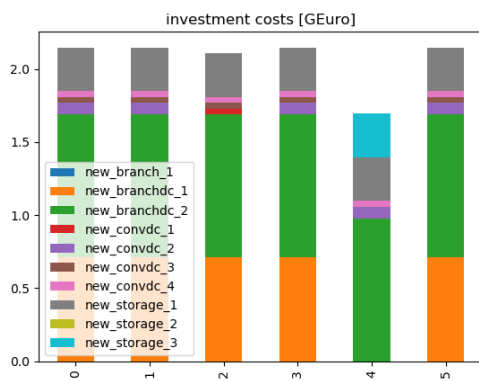

Figure 9-42 - Optimal investment decisions, and their respective costs, for the full scenario set of 35 years.

An overview of the optimal investment decisions for each scenario-year is shown in Figure 9-42. These results show that 5 different variants are chosen over these 35 years. The variant in which is invested in storage 1, combined with the 2 DC branches and converters 2,3 and 4 is chosen most of the time, followed by a variant with storage 2 and the same branch and converter investments.

Next, the same testcase was run for different reduced scenario sets. Results of the optimal decisions obtained from a reduced set of 6 monthly scenarios, 6 weekly scenarios and 6 daily scenarios are shown in Figure 9-45 to Figure 9-46.

Monthly scenarios

Figure 9-45 and Figure 9-44 show the investment decisions when monthly scenarios are used. In Figure 9-45 the clustering was done after feature reduction by PCA, in Figure 9-44 clustering was done on scenarios reduced by describing them by their main characteristics. The results show that different variants are again obtained for each cluster, and that results differ depending on the feature reduction technique used. However, the variant that occurs most in the monthly scenarios is the same variant occurring mostly in the yearly scenarios, for both feature reduction techniques. This indicates that reducing to monthly scenarios gives good indications of providing reliable results.
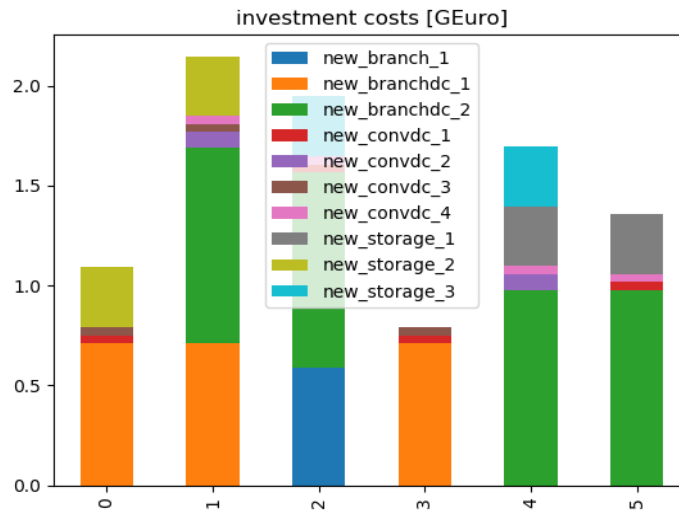

Figure 9-45 - Optimal investment decisions, and their respective costs, for the reduced scenario set of 6 monthly scenarios. These scenarios were reduced with PCA as feature reduction technique.
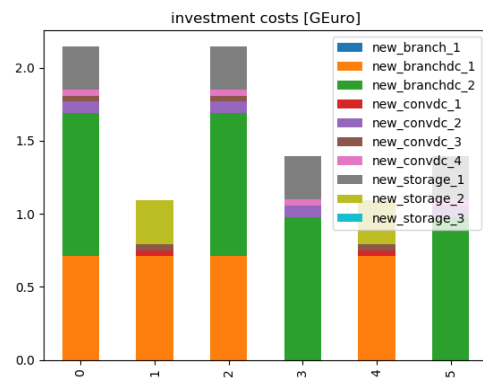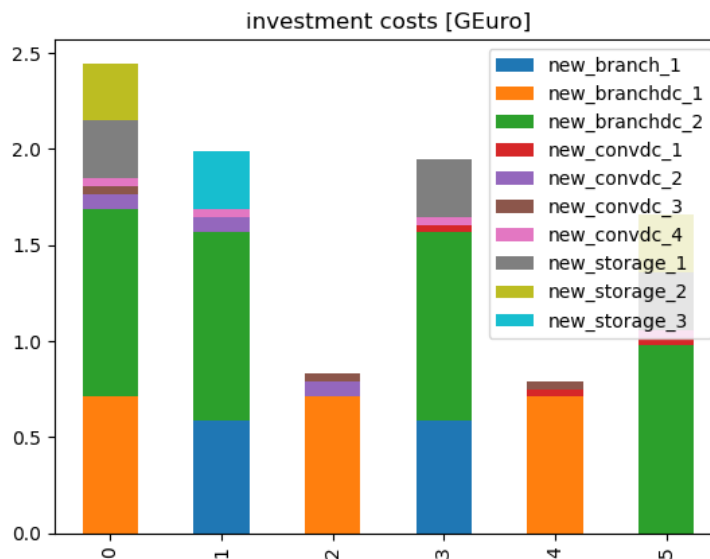
Figure 9-44 - Optimal investment decisions, and their respective costs, for the reduced scenario set of 6 monthly scenarios. These scenarios were reduced by describing the scenarios by their main characteristics.

Weekly and daily scenarios

Figure 9-47 and Figure 9-46 show the investment decisions when weekly and daily scenarios are used respectively. It is clear from these figures that completely different investment decisions are obtained, when compared with the yearly results. (Only the results for scenarios reduced using PCA are shown here, but similar conclusions can be drawn for the other feature reduction techniques). Reducing to these timeframes does not seem to be a good choice for this testcase.


Figure 9-47 - Optimal investment decisions, and their respective costs, for the reduced scenario set of 6 weekly scenarios. These scenarios were reduced with PCA as feature reduction technique.


Figure 9-46 - Optimal investment decisions, and their respective costs, for the reduced scenario set of 6 daily scenarios. These scenarios were reduced with PCA as feature reduction technique.

The reason for this discrepancy lies most likely with the size of the storage candidates present in the testcase. The size of the storage candidates is set at 400 MW/10000MWh, which makes it takes about 25 hours to charge the storage at full power. The full capacity is however never available in the system, which makes that it would rather take a few days to fully charge or discharge. This is illustrated in Figure 9-48, where the optimised storage state of charge in one monthly scenario is shown: the state of charge never reaches a value beyond 30%. The storage charge and discharge power of the same scenario is shown in Figure 9-49: charge and discharge powers above 1 pu (or 100MW) are rare.

To capture the large storage time constant in the scenarios, these scenarios must be long enough, which is clearly not the case for the weekly and daily scenario sets.
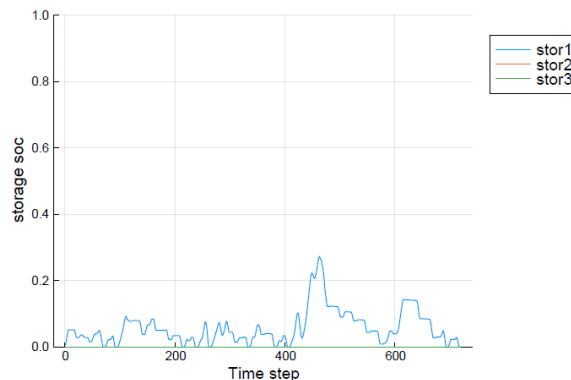
Figure 9-48 - Storage state of charge in one monthly scenario.
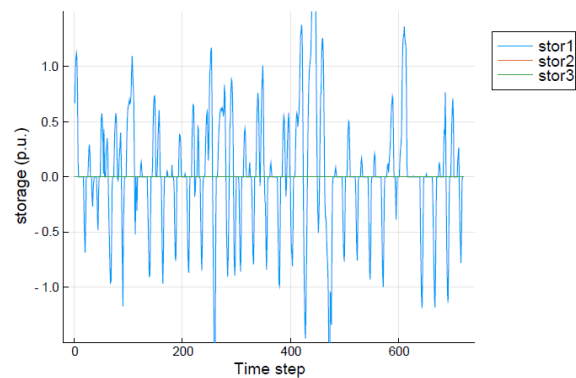
Figure 9-49 - Storage charge/discharge power in one monthly scenario.

In contrast with the test case applied, time constants for realistic storage systems will probably not be that large. Therefore, it was decided to repeat the scenario reduction test on a testcase with storage candidates with smaller time constants. Results of this test case will be provided in an update of this test report.

9.3.8 Tests on the stochastic optimisation model formulation with multiple Monte Carlo time series

This particular test analyses how the choice of the Monte Carlo time series affects the overall planning decision taken by the optimisation model. Therefore, two different Monte Carlo time series have been used as input for the planning optimisation. The demand time series is obtained from the ENTSO-E transparency platform for the six market zones in Italy, in correspondence with the test case depicted in Section 9.2.1. For the renewable generation (PV, wind), the data is obtained from Renewables Ninja[11], using the coordinates of the nodes hosting a renewable generator (nodes 3 - Centre South - and node 6 -Sicily). The time series data is obtained for the years 2018 and 2019 for 8760 hours. For both planning years, in the optimisation only the

---

[11] https://www.renewables.ninja/

first month (720 hours) of the time series is used. Cost information for the generation, candidates and demand are provided in the Appendix A (Section 11.1.7). Additionally, the renewable generation has been upscaled/downscaled for the years 2019 and 2018 with a factor of 2 and 0.5, respectively, in order to have distinct scenarios with high and low renewable generation availability. In the low renewable generation scenario, the missing renewable generation is replaced by conventional generation. The generation costs have been scaled to 10 years, in order to provide a planning decision covering a decade.

The planning problem is solved using the stochastic objective as described in Section 0 using both time series, e.g., scenarios. This means that the resulting optimisation problem considers 2 times 720 hours, and finds the necessary investments making a probabilistic weighing of both scenarios. As planning options, AC and DC connections, storage investments and demand flexibility investments are considered, for which the list is given in Appendix A (Section 11.1.7).

The probabilities (π) of the scenarios have been varied as follows, resulting in 11 different calculations:

$$\pi_{2019} = 0, 0.1, 0.2, \dots, 1 \rightarrow \pi_{2018} = 1 - \pi_{2019}.$$

Figure 9-50 shows the total system cost in dependence of the scenario probabilities. We can observe that the total system cost increases with the probability of the high renewable generation scenario.
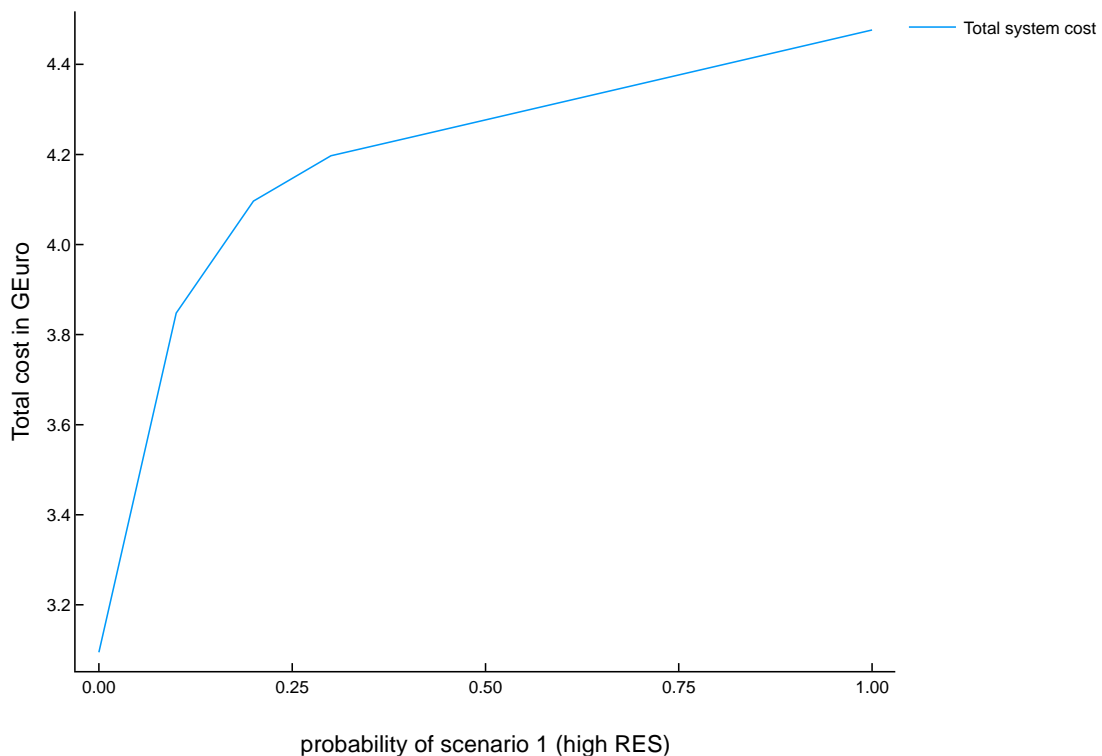
Figure 9-50 - Total system cost in dependence of the scenario probability for high RES generation

Figure 9-51 shows the optimal system topology in the case of the low renewable generation scenario ($p\_1 = 0 \rightarrow p\_2 = 1$). We can see that in this case only the HVDC link between node 4 and 6 is built, as the amount of renewable generation that needs to be evacuated from node 6 is minimal and the demand can be satisfied largely by the conventional generation.
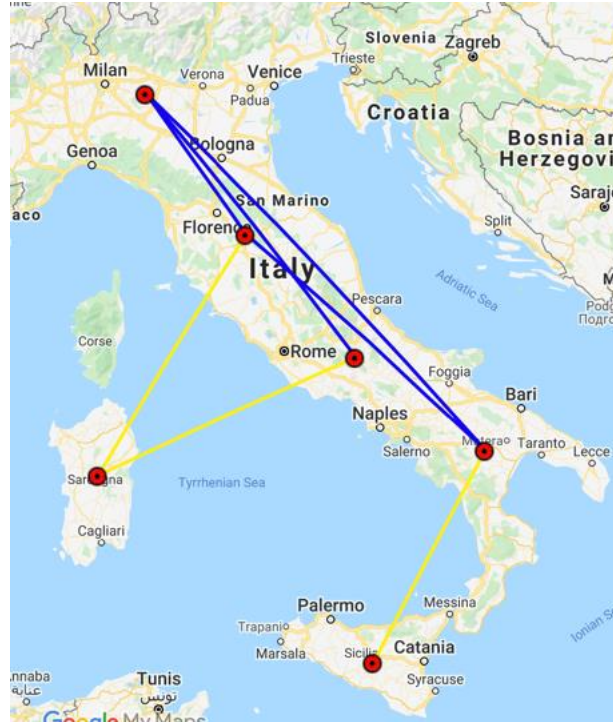
Figure 9-51 - Optimal system layout for the low RES generation scenario

Figure 9-52 shows the demand flexibility activations for the demand on node 1 in the low renewable generation scenario ($p_1 = 0 \rightarrow p_2 = 1$). We can observe that demand shifting actions both for upwards and downwards actions are needed along with voluntary demand reduction actions (brown lines). The cost of the voluntary demand reduction actions in this case does not justify the instalment of storage devices, as the probability of the high renewable scenario is zero and is thus not represented in the objective function.
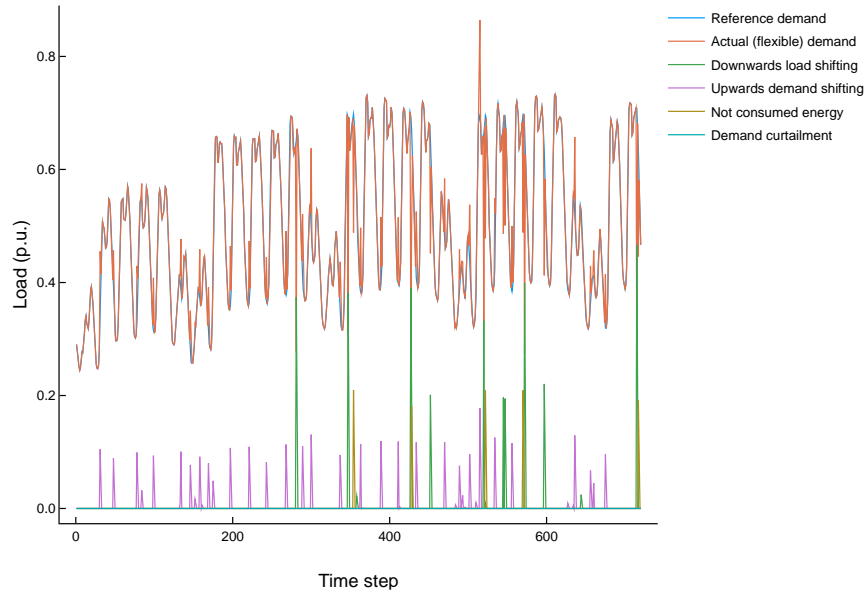
Figure 9-52 - Demand flexibility activations for the low RES generation scenario
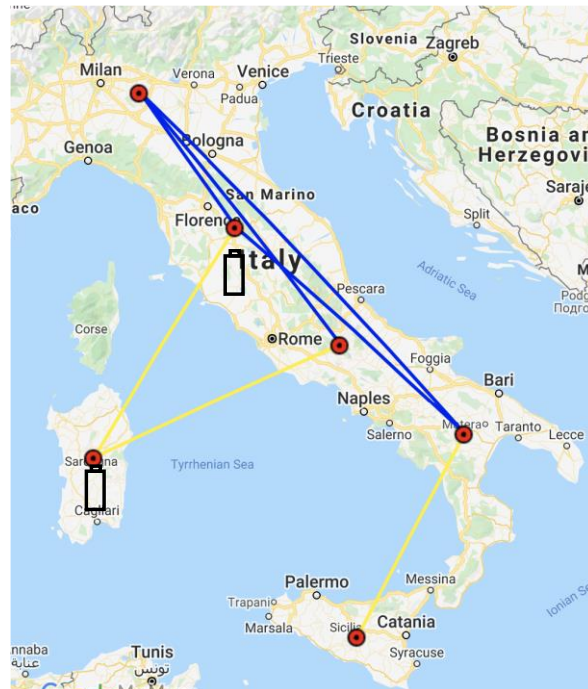
Figure 9-53 - Optimal system layout for the hi1gh RES generation scenario probability of 10%

Figure 9-53 shows the system layout if the probability of the high renewable generation scenario is increased to 10% ($p_1 = 0.1 \rightarrow p_2 = 0.9$). We can observe that in this case the candidate storage investments on nodes 2 (Center North) and 5 (Sardinia) are chosen as the expected cost of demand reduction and possible curtailment justifies the investment costs into these storage assets. If we look at the demand flexibility actions on node 1 of both scenarios for this scenario combination, we can observe that the installation of the storage devices is sufficient to avoid demand actions for the low RES scenario (Figure 9-54), whereas actions for the

high RES scenario are still largely utilised (Figure 9-55), due to the fact that this scenario is only weighed with 10% in the objective function, and thus not affect the overall objective to a large extend.
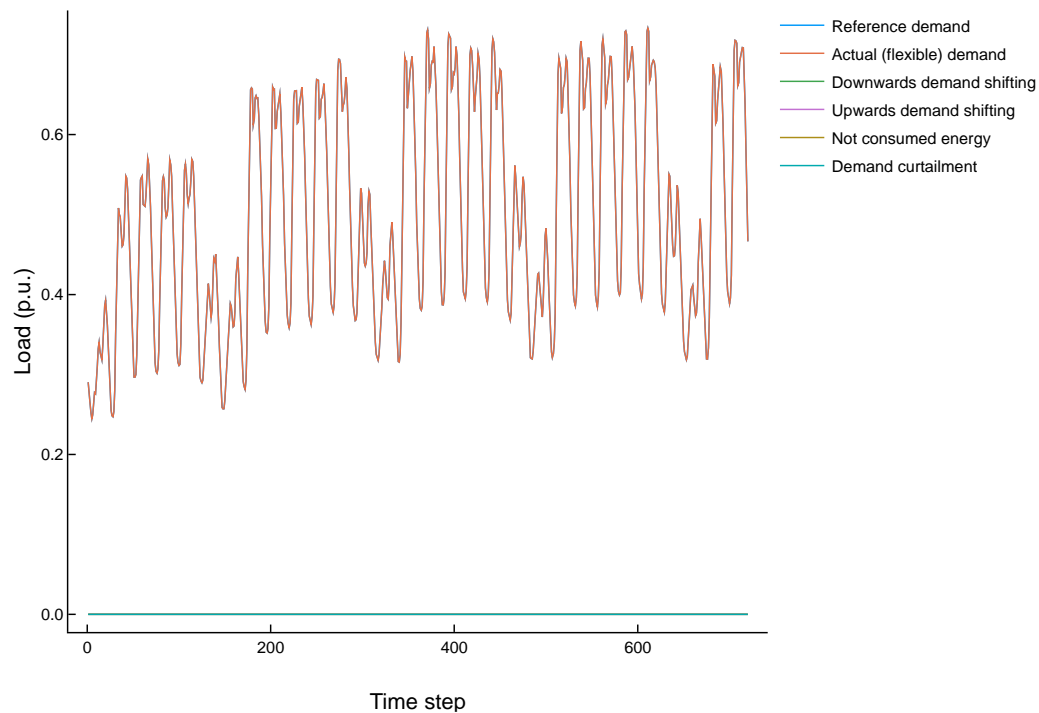
Figure 9-54 - Demand flexibility actions of the low RES generation scenario for scenario probability combination ($p_1 = 0.1 \rightarrow p_2 = 0.9$)
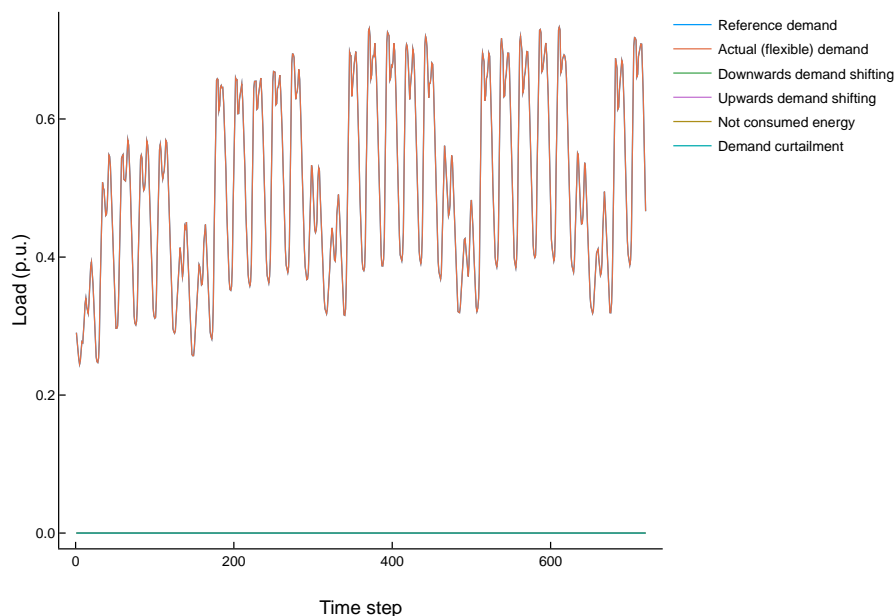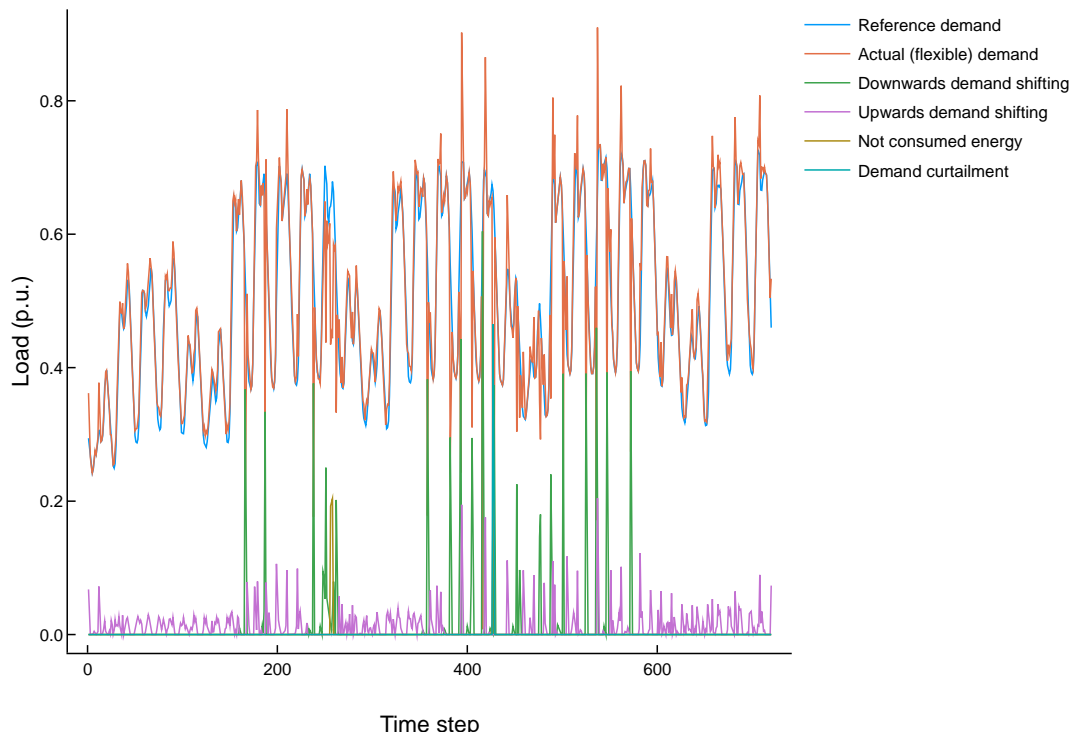

Figure 9-55 - Demand flexibility actions of the high RES generation scenario for scenario probability combination ($p_1 = 0.1 \rightarrow p_2 = 0.9$)

As reflected in the total system costs (Figure 9-50) the next large change in the objective function occurs at the probability combination ($p_1 = 0.3 \rightarrow p_2 = 0.7$), where instead of the storage candidate on node 5 (Sardinia), the candidate on node 6 (Sicily) is chosen, which has a higher charge and discharge power rating but also higher costs. The following figures show again the demand flexibility actions on node 1 for both scenarios. For the low RES generation scenario weighed with a factor of 0.7 (Figure 9-56), we still do not observe any demand flexibility actions, as the use of storage assets is sufficient to cover the generation fluctuations coming from the RES. For the high RES generation scenario, now weighed with a factor of 0.3 (Figure 9-57) in the objective function we can observe that the number of necessary demand flexibility actions is highly reduced compared

Figure 9-56 - Demand flexibility actions of the low RES generation scenario for

to Figure 9-55. If the probability of the high RES generation scenario is further increased, we neither observe any changes in the investment candidates chosen, nor in the pattern of the demand flexibility actions. The increase in the total system costs solely stems from the higher scenario weight of the high RES scenario in the objective function of the optimisation, as the demand flexibility action costs gain a higher weight in the total objective.
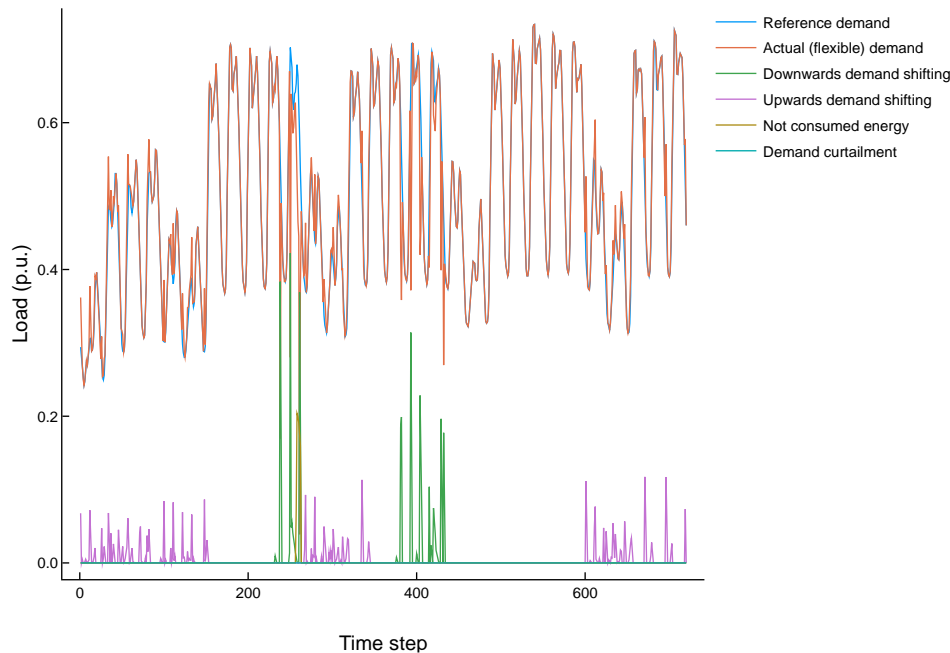
Figure 9-57 - Demand flexibility actions of the high RES generation scenario for

9.3.9 Proof-of-concept conclusions

The proof-of-concept implemented through the *FlexPlan.jl* package and the tests conducted using it have shown that the models developed in this deliverable provide coherent results for network expansion problems, both on transmission and distribution grids. Moreover, the different building blocks of the model have been shown to interact in a consistent way, including the demand flexibility and storage models (Chapter 5) and the scenario reduction methodology (Chapter 7), whose inclusion in network expansion problems is one of the main innovations of the FlexPlan project.

Dedicated tests assessing separately each feature of the model have allowed to present their key characteristics and potential limitations. It was for instance highlighted that modelling storage assets as lossless could result in an optimal solution with simultaneous charging and discharging of the asset. As a second example, it was also stressed that the optimal investment decisions can depend significantly on the set of operational scenarios, their probabilities and the reduction process. Daily scenarios were for instance shown to poorly capture the flexibility that can be delivered by large storage assets (e.g. PHS) due to their larger charge and discharge times. The potential impact of those limitations on the optimal TNEP/DNEP solution was discussed and possible solutions were investigated.

These considerations are of high importance for the implementation of the planning tool. As such, the *FlexPlan.jl* package serves as a reference design for the development of the tool. Results of the small-case tests performed on the proof-of-concept are also directly compared to the results delivered by the planning tool and

serve as a validation of the correct implementation of the latter. The main points of attention underlined in the tests above are closely monitored in the development of the planning tool, for instance in order to correctly calibrate the granularity of the scenario reduction.

In parallel to the development of the planning tool, the *FlexPlan.jl* package will continue to be used to explore and test additional features of the model such as optimisation techniques that could significantly reduce the computation time of the tool (see Chapter 8). The features that are still being tested on the proof-of-concept implementation is detailed below:

- Benders decomposition
- Combined transmission and distribution optimization

As soon as they are finalized, those tests will give rise to a revision of this deliverable, along with possible other feed details to be set up during the planning tool implementation and its testing phase.

10 Model implementation in the final planning tool

Based on the experiences from the performance of the planning tool on the large-scale test cases within the regional case studies, in the final planning model a number of simplifications and relaxations have been applied, in order to obtain manageable computation times. The following sections give an overview of the model simplifications with respect to the planning model as described in the previous sections.

10.1 Simplifications on the demand flexibility and storage modelling

The computational requirements to model demand flexibility and storage are very high, due to the introduction of time coupling constraints within the optimisation model. The time coupling constraints break the block nature of the constraint matrices which are passed to the chosen optimisation solver. As such, by introducing relaxations and temporal decompositions based on engineering assumptions, the overall efficiency of the optimisation model can be improved. The following paragraphs outline a number of such relaxation and decomposition methods, which have been applied to the planning model.

10.1.1 Temporal decomposition of large hydro storage modelling

The challenge in the temporal decomposition is to account for seasonal variability of inflow based on the utilisation of large hydro reservoirs. As the idea of the decomposition is to divide the optimisation problem into a number of subproblems without a direct connection, it is necessary to determine and define the seasonal distribution and utilisation of inflow exogenously, e.g., as an input parameter to the problem. Here it should be noted that the stochastic optimisation of long-term hydropower dispatch (optimisation of seasonal arbitrage potential) is outside the scope of the FlexPlan tool.

The aim of the temporal decomposition is to have a consistent representation of subproblems, e.g., to have consistency between the time series for inflow, reservoir handling and hydro production. Additionally, available data, e.g. the reference production data provided by the MILES tool, should be utilized as much as possible, without having to gather new data.

Let us first start with the generic storage model as defined in section 5.2:

$$E_{j,y}^{max} x_{j,t,y} = E_{j,y}^{max} x_{j,t-\Delta t,y} + \Delta t \cdot \left(\eta_{j,y}^{abs} P_{j,t,y}^{abs} - \frac{P_{j,t,y}^{inj}}{\eta_{j,y}^{inj}} + \xi_{j,t,y} \right),$$

where:

- $E_{j,y}^{max}$ is the energy rating (energy volume) of the storage, e.g., the capacity of the basin in MWh,
- $x_{j,t,y}$ is the storage level,
- $P_{j,t,y}^{abs}$ is the absorption power, e.g., the power of the pump in MW and $\eta_{j,y}^{abs}$ its efficiency, respectively for each hour. The absorption power is bound by the power rating of the pump, $0 \leq P_{j,t,y}^{abs} \leq P_{j,y}^{abs,max}$
- $P_{j,t,y}^{inj}$ is the injection power, e.g., the power of the turbine in MW and $\eta_{j,y}^{inj}$ its efficiency, respectively for each hour. The turbine power is bound by the power rating of the turbine $0 \leq P_{j,t,y}^{inj} \leq P_{j,y}^{inj,max}$ and
- $\xi_{j,t,y}$ is the natural inflow.

- Δt is the time resolution considered, e.g., 1

In the above equation, $E_{j,y}^{max}, \Delta t,\ \eta_{j,y}^{abs},\ \eta_{j,y}^{inj}, \xi_{j,t,y}$ are known parameters, $P_{j,t,y}^{abs}$ and $P_{j,t,y}^{inj}$ are the decision variables determined by the planning tool, and $x_{j,t,y}$ is a state variable as a direct result of the choice of the power injection / absorption decision of each hour. The storage model in section 5.2 states that $E_{j,y}^{max} x_{j,T,y} \geq E_{j,y}^{init}$, which means that the energy level of the storage at the end of a given period (T) needs to be larger or equal to the starting value ($E_{j,y}^{init}$). Also, the stored energy is bound by $E_{j,y}^{min} \leq E_{j,y}^{max} x_{j,t,y} \leq E_{j,y}^{max}$, $E_{j,y}^{min}$ would be minimum storage level to be considered.

10.1.2 Storage modelling with explicit modelling of the inflow

In order to decompose the storage model in a set of monthly or weekly models, we need to determine a priori:
- $E_{j,y}^{init,m}$ for the beginning of each period
- $\xi_{j,t,y}$ the natural inflows for each hour during that period

For the example of a monthly decomposition, by setting $T = 720$ hours (= 1 Month), we can use $E_{j,y}^{\max x_{j,720,y}} \geq E_{j,y}^{init,m+1}$, e.g., the energy stored at the end of the month needs to be larger or equal than the initial value of the storage in the following month.

For the first month we can assume:
- $E_{j,y,m=1}^{init} = E_{j,y}^{max} \cdot 0,5$ meaning that at the beginning of the year, the storage is half-full

For the consecutive months we can calculate the initial values:
- $E_{j,y,m=2}^{init} = E_{j,y,m=1}^{init} + \sum_{h=1\ to\ 720} \xi_{j,h,y} - \sum_{h=1\ to\ 720} E_h^{ref}$, where
  - E_h^{ref} is the hourly reference production provided by MILES, e.g., $E_h^{ref} = P_h^{inj,ref} - P_h^{abs,ref}$ where $P_h^{inj,ref}$ is the hourly injection into the network and $P_h^{abs,ref}$ is the hourly pumping power
    - $P_h^{abs,ref}$ can be obtained from the ***"pumped storage"*** production of the MILES output
    - $P_h^{inj,ref}$ can be obtained from ***"storage"*** production of the MILES output

The only parameter which is not determined in the above equations is the hourly inflow $\xi_{j,h,y}$. This means that if we can approximate the hourly inflows, we can describe the full model. To that end we use an average weekly inflow value which approximates the hourly inflow, e.g.

- $\xi_{j,h=1-168,y} = \dfrac{\left(\sum_{h=1\ to\ 168}\left(P_h^{inj,ref} - P_h^{abs,ref}\right)\right)}{168}$

Hence, the inflow values for the first 168 hours are calculated directly as a sum from the reference hydro storage production $E_h^{ref} = P_h^{inj,ref} - P_h^{abs,ref}$ and averaged over 168 hours. This is repeated for all remaining hours throughout the year:

- $\xi_{j,h=169-336,y} = \dfrac{\left(\sum_{h=169\ to\ 336}\left(P_h^{inj,ref} - P_h^{abs,ref}\right)\right)}{168}$

- …….

The temporal decomposition of the (pumped) hydro-storage does not require any changes to chosen data model, as only the constraint $E_{j,y}^{max} x_{j,T,y} \geq E_{j,y}^{init}$ is altered by means of a different parametrisation of $E_{j,y}^{init}$ and $E_{j,y}^{max}$.

10.1.3 Relaxation of integrality constraints for storage and demand flexibility modelling

As mentioned in the previous paragraphs, the time dependent and especially the integrality constraints which are used to model demand flexibility and storage, break the sparsity and the block structure of the constraint matrices, making the problem much harder to solve for optimisation solvers. As such, a number of relaxations w.r.t. the models presented in sections 5.1 and 5.2 have been made in the implemented planning tool as listed below:

- Relaxations of the demand flexibility model
  - In order to simplify the demand flexibility model, the constraint on the total amount of energy shifted over a given planning year, $0 \leq \sum_{t \in S_t} \Delta t \cdot \Delta P_{u,t,y}^{ds,dn} \leq \alpha_{u,y} E_{u,y}^{ds,max}$, has been removed within the implemented planning tool, as each upwards or downwards demand shifting action is decided based on the cost of such actions and as such could be performed as many times as economically viable.

  - Instead of using the formulation introduced in section 5.1.2 for modelling the demand recovery, a new constraint is introduced which maintains the energy balance between upwards and downwards demand shifting over a period of 24 hours, e.g., $\sum_{t=1\ to\ 24}\left(\Delta P_{u,t,y}^{ds,up} - \Delta P_{u,t,y}^{ds,dn}\right) = 0$. Within the planning tool, such 24-hour blocks are implemented making using special ordered sets.

- Relaxations of storage model
  - The minimum energy constraint on storage, as in $E_{j,y}^{min} \leq E_{j,y}^{max} x_{j,t,y} \leq E_{j,y}^{max}$ can introduce infeasibilities in combination with the inflow modelling and as such it assumed that a minimum storage level of 0 can be achieved, e.g. $0 \leq E_{j,y}^{max} x_{j,t,y} \leq E_{j,y}^{max}$.

  - The constraint for approximating the charging and discharging exclusivity, e.g., $P_{j,t,y}^{abs} + P_{j,t,y}^{inj} \leq \max\left(P_{j,t,y}^{abs,max}, P_{j,t,y}^{inj,max}\right)$, as described in section 5.2.1 has been dropped in the actual implementation. Considering that the storage devices have an efficiency of smaller than 1, simultaneous charging and discharging should be prevented by the optimisation objective as the additional losses need to be covered by additional generation. Thus, simultaneous charging and discharging behaviour would only be expected in circumstances where energy needs to be dissipated to keep the feasibility of the model. By dropping the afore mentioned constraint we introduce additional slack in the optimisation model which will get active in such hard instances requiring energy dissipation. This slack can be inspected in the post-processing stage in order to find structural weaknesses in the grid, requiring energy dissipation.

○ The storage model in section 5.2.1 includes a constraint on the total amount of energy that can be absorbed by a storage device, as a proxy of the cycle lifetime of storage system, e.g., $\sum_{t \in S_t} \Delta t \cdot P_{j,t,y}^{abs} \leq E_{j,y}^{abs,max}$. Owing it to its integral nature, this constraint decreases the computational efficiency of the implementation. As such, in the planning tool, this particular constraint has been omitted, under the assumption that through inspection of the optimisation results the cycling behaviour can be obtained, and if proven to excessive, additional costs for the replacement of such assets can be taken into account.

10.2 Time series and scenario reduction using representative weeks

The objective function of the planning model is defined in section 0 as follows:

$$
\sum_s \pi_s \left\{ \sum_{y \in S_y} f_y^{d,o} \left\{ \sum_{t \in S_t} \left[\sum_{g \in S_g} [C_{g,y}^{aq} + (\theta^{CO_2} G^{pf} + \theta^f) \eta_g^f] P_{g,t,y,s} + C_{g,y}^{res,curt} \Delta P_{g,t,y,s}^{res} + \sum_{j \in S_j} [C_{j,t,y}^{abs} P_{j,t,y,s}^{abs} + C_{j,t,y}^{inj} P_{j,t,y,s}^{inj}] \right. \right. \right.
$$

$$
+ \sum_{j \in S_{jc}} [C_{jc,t,y}^{abs} P_{jc,t,y,s}^{abs} + C_{jc,t,y}^{inj} P_{j,ct,y,s}^{inj}]
$$

$$
+ \sum_{u \in S_u} [C_{u,t,y}^{nce}(P_{u,t,y,s}^{ref} - P_{u,t,y,s}^{nce}) + C_{u,t,y}^{ds}(\Delta P_{u,t,y,s}^{ds,up} + \Delta P_{u,t,y,s}^{ds,dn}) + C_{u,t,y}^{lc} \Delta P_{u,t,y}^{lc}]
$$

$$
\left. + \sum_{n \in S_n} (C_{n,t,y}^{EE} EE_{n,t,y,s} + C_{n,t,y}^{LL} LL_{n,t,y,s}) \right]
$$

$$
+ f_y^d \left[\sum_{j \in S_{jc}} \alpha_{jc,y}(I_{jc,y}^E(E_{jc}^{max}) + I_{jc,y}^P(P_{jc}^{max}) + FP_{jc,y}^{CO_2}) + \sum_{u \in S_u} \alpha_{u,y}(I_{u,y} + FP_{u,y}^{CO_2}) \right.
$$

$$
+ \sum_{lc \in S_{lc}^{ac}} \alpha_{lc,y}(I_{lc,y} + FP_{lc,y}^{CO_2} + LS_{lc,y}) + \sum_{dc \in S_{lc}^{dc}} \alpha_{dc,y}(I_{dc,y} + FP_{dc,y}^{CO_2} + LS_{dc,y})
$$

$$
\left. \left. \left. + \sum_{zc \in S_{zc}} \alpha_{zc,y}(I_{zc,y} + FP_{zc,y}^{CO_2} + LS_{zc,y}) + \sum_{bc \in S_{bc}} \alpha_{bc,y}(I_{bc,y} + FP_{bc,y}^{CO_2} + LS_{bc,y}) \right] \right\} \right\}
$$

The objective function consists of two major parts. The green part consists of the investment costs (denoted with I), cost of landscape impact (denoted with LS) and cost of carbon footprint impact (denoted with FP), which re are summed up over the set of planning horizons $y \in S_y = \{2030, 2040, 2050\}$. Additionally, the costs are weighted with the probability π_s of the different scenarios, e.g., climate variants s. Note that $\sum_s \pi_s = 1$.

The red part of the objective function consists of the operational costs of the system, such as generation dispatch, costs related to demand flexibility and potentially related to operating storage systems. As these costs are different for each hour of the considered planning year $t \in S_t = \{1, 2, \ldots, 8760\}$, they are summed up. As in the case of investment costs, additionally, these operational costs are summed up over the different planning horizons and weighed and summed according to the climate variants.

In order to reduce the dimensionality of the problem, representative operational weeks can be used, also in line with the decomposed storage modelling approach shown in the previous section. In this approach each operational week represents the demand and RES generation conditions representative of a larger number time steps, e.g. a certain number of weeks.

In this case, the objective function can be re-written as:

$$\sum_s \pi_s \left\{ \sum_{y \in S_y} f_y^{d,o} \left\{ \sum_{w \in S_w} K_w \sum_{t \in S_{wt}} \left[\sum_{g \in S_g} \left[C_{g,y}^{aq} + (\theta^{CO_2} G^{pf} + \theta^f) \eta_g^f \right] P_{g,wt,y,s} + C_{g,y}^{res,curt} \Delta P_{g,wt,y,s}^{res} \right. \right. \right.$$

$$+ \sum_{j \in S_j} \left[C_{j,wt,y}^{abs} P_{j,wt,y,s}^{abs} + C_{j,t,y}^{inj} P_{j,wt,y,s}^{inj} \right] + \sum_{j \in S_{jc}} \left[C_{jc,wt,y}^{abs} P_{jc,wt,y,s}^{abs} + C_{jc,t,y}^{inj} P_{j,wt,y,s}^{inj} \right]$$

$$+ \sum_{u \in S_u} \left[C_{u,wt,y}^{nce} \left(P_{u,wt,y,s}^{ref} - P_{u,wt,y,s}^{nce} \right) + C_{u,wt,y}^{ds} \left(\Delta P_{u,wt,y,s}^{ds,up} + \Delta P_{u,wt,y,s}^{ds,dn} \right) + C_{u,wt,y}^{lc} \Delta P_{u,wt,y}^{lc} \right]$$

$$\left. + \sum_{n \in S_n} \left(C_{n,wt,y}^{EE} EE_{n,wt,y,s} + C_{n,wt,y}^{LL} LL_{n,wt,y,s} \right) \right]$$

$$+ f_y^d \left[\sum_{j \in S_{jc}} \alpha_{jc,y} \left(I_{jc,y}^E \left(E_{jc}^{max} \right) + I_{jc,y}^P \left(P_{jc}^{max} \right) + FP_{jc,y}^{CO_2} \right) + \sum_{u \in S_u} \alpha_{u,y} \left(I_{u,y} + FP_{u,y}^{CO_2} \right) \right.$$

$$+ \sum_{lc \in S_{lc}^{ac}} \alpha_{lc,y} \left(I_{lc,y} + FP_{lc,y}^{CO_2} + LS_{lc,y} \right) + \sum_{dc \in S_{lc}^{dc}} \alpha_{dc,y} \left(I_{dc,y} + FP_{dc,y}^{CO_2} + LS_{dc,y} \right)$$

$$\left. \left. \left. + \sum_{zc \in S_{zc}} \alpha_{zc,y} \left(I_{zc,y} + FP_{zc,y}^{CO_2} + LS_{zc,y} \right) + \sum_{bc \in S_{bc}} \alpha_{bc,y} \left(I_{bc,y} + FP_{bc,y}^{CO_2} + LS_{bc,y} \right) \right] \right\} \right\}$$

Here, the operational hours considered within the planning year are replaced by operational hours occurring in a representative week $wt \in S_{wt} = \{1, 2, \ldots, 168\}$. The operational costs are summed up over the hours of the representative weeks, which are in turn weighed with the factor K_w over the chosen set of representative weeks $w \in S_w = \{1, \ldots, N_w\}$, where N_w denotes the number of chosen representative weeks. The factor K_w can be calculated as $K_w = \frac{8760}{168} R_w$, where R_w is the measure of representativeness of each particular week w, with $0 \leq R_w \leq 1$ and $\sum_{w \in S_w} R_w = 1$.

Form the point of view of the planning tool, this means that operational hours would need to be grouped into representative weeks, and for each representative week, the measure of representativeness R_w would need to be provided.

In order to avoid defining two separate parameters for the probability of the climate variants as well as the weight of each representative week, both can be combined to a single parameter defined for each representative week of each climate variant as:

$$L_{s,w} = \pi_s K_w \quad \forall s, w \in S_w.$$

**Implications for storage and demand flexibility modelling:**

As mentioned in the previous section, explicit parameters for the starting and final energy content of the storage devices have been defined for being able to decompose the storage model in operational weeks, which could directly be used as the starting and final energy contents of the representative weeks. As such from the

perspective of storage modelling, no additional input parameters would be needed w.r.t. to the original data model (or the input data file of the actual implementation). With respect to demand flexibility modelling, this change would also be transparent to the users, as the recovery of the demand shifting is assumed to take place within a 24-hour period, as outlined within the previous section.

Scenario variants and representative weeks as used for the calculations on the regional cases:

For the practical implementation within the regional cases the number of lengths of the input time series have been reduced using k-means clustering. For the regional case calculations, two climatic variations have been used, where for each climatic variation consists of 4 representative weeks. The representative weeks consist of 84 hours each, by using averaging two consecutive hours to represent both. The next paragraphs show the scenario and time series reduction procedure used. Using k-means clustering, first, 2 climatic variations of 8760 hours are generated out of the 35 climatic variations (Figure 10-1).
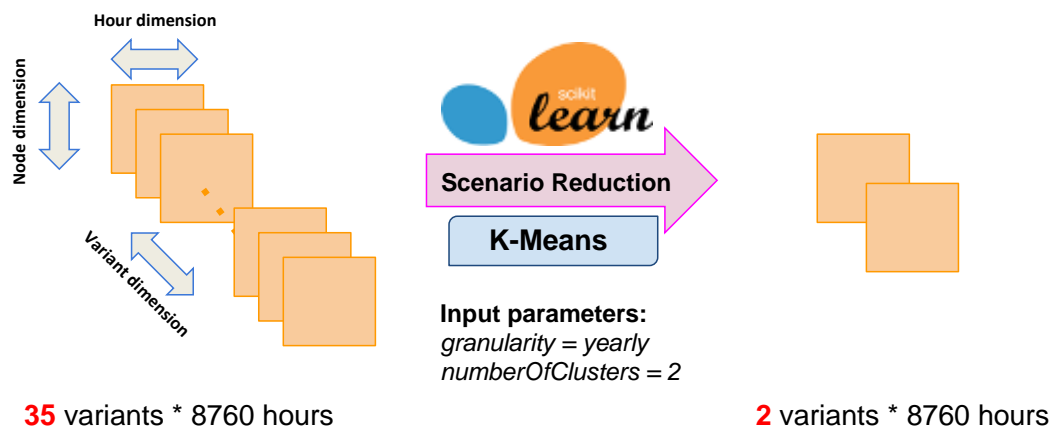
Figure 10-1 Clustering of climatic variations

In the next step, out of each time series of 8760 hours, 12 clusters of 168 hours are generated as shown in Figure 10-2.
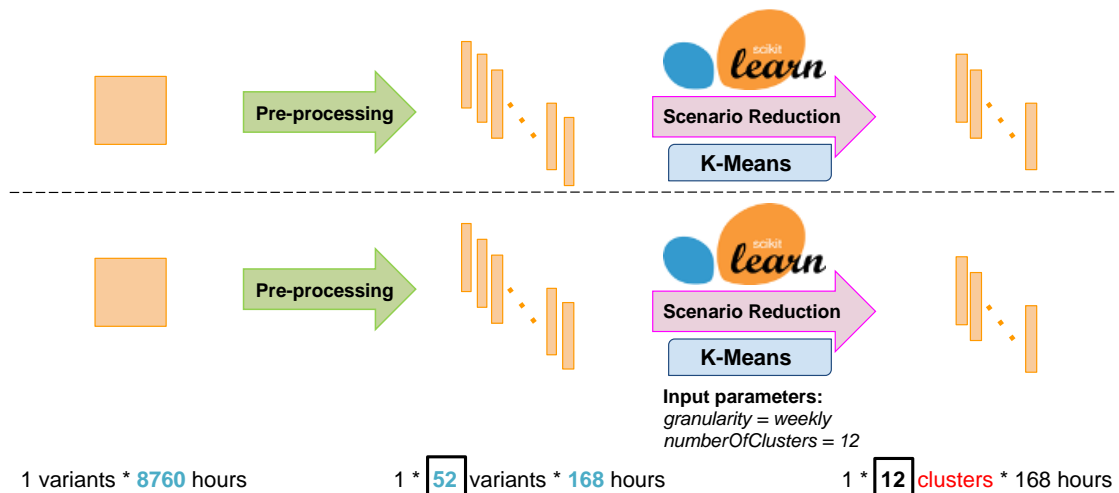
Figure 10-2 Generation of weekly clusters

In the final step, 4 out of the 12 clusters are selected which are distributed among the different seasons to account for seasonality. Eventually, within each chosen representative week, two consecutive hours are averaged to further reduce the length of the time-series used.
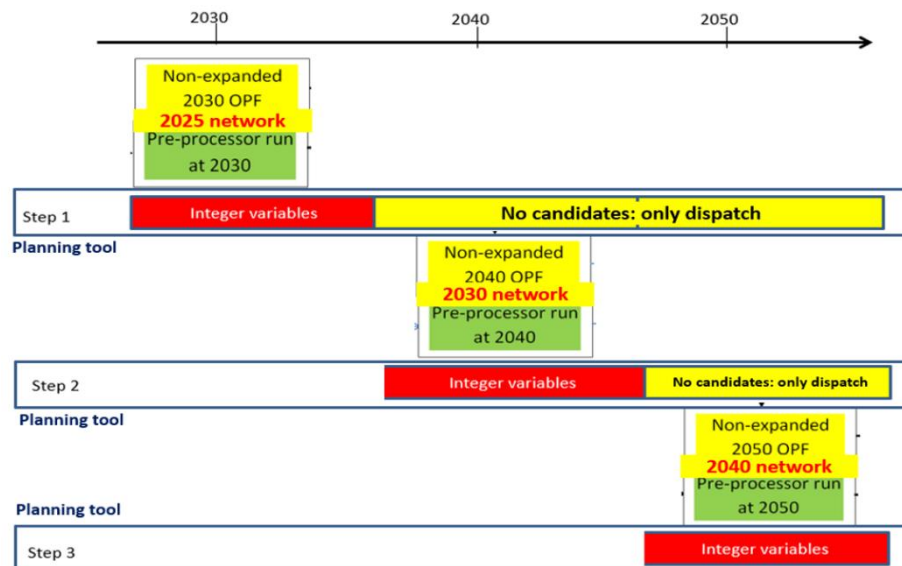
10.3 Pre-processor and planning tool interaction

The expansion candidates for the planning tool are delivered by the pre-processor unit as briefly described in section 1. As the application of the pre-processor routine to the non-expanded network for the planning years 2040 and 2050 would result in unrealistically large congestions, as the network should be expanded after the first planning year, e.g. 2030. Generally, such problems are solved by iteratively applying the pre-processor routine and the planning until a result is found for which the pre-processor routine does not provide a different set of candidates (and thus the planning models obtains the same solution). This can potentially result in a high number of iterations between the pre-processor routine and the planning tool. In order to reduce the number of iterations between the pre-processor routine, delivering the expansion candidates, and the planning tool, the procedure shown in Figure 10-3 is applied. In the first step, the pre-processor tool is applied to the non-expanded network using the expected generation and demand time series for 2030. Then the planning problem is solved using the candidates for 2030, whereas for the remaining two planning years, only the operational problem is solved. In the second step, the expanded 2030 network is used as the starting point for the pre-processor using the 2040 generation and demand time series. Then the planning problem is solved for the year 2040 as well as the operational problem for 2050. In the last step, the expanded 2040 network is used as the starting point for the pre-processor using 2050 generation and demand time series and subsequently the planning model is solved for the planning year 2050.
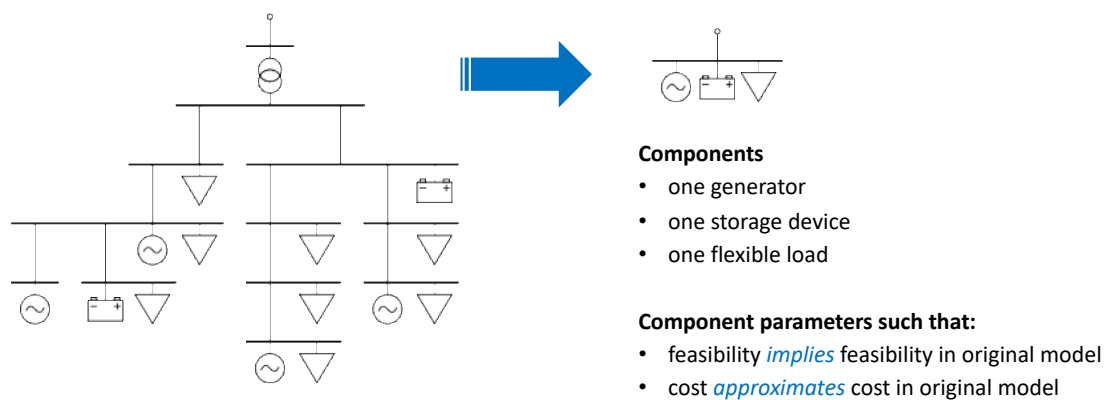

Figure 10-3 Interaction between the pre-processor and the planning tool

10.4 Decomposition of the transmission and distribution system planning models

In order to decrease the number of evaluations of the distribution system planning problem, the transmission and distribution system decomposition model has been simplified. For this purpose, a three-stage procedure is applied in contrast to the method described in section 8.1.2:

1. For each distribution network, a surrogate model is computed, which reduces the distribution networks to a single bus representation consisting of aggregated generation, flexible demand and storage as shown in Figure 10-4, where the parameters of the storage model can be optimised using the model described in section 5.2 and the simplifications applied in section 10.1.

2. In this stage the transmission planning problem with the connected surrogate distribution grid model is solved. This provides the optimal solution from the perspective of the transmission grid and provides the necessary costs of the expansion and the expected power exchanges between the transmission and the distribution grids.

3. Using the expected power exchange between the transmission and distribution grids as an input, and using the full distribution network with a number of defined candidates, the optimal expansion of the distribution grid is determined using the planning model, and the power flow formulation as described in section 4.3. This provides eventually the optimal solution from the distribution grid point of view, while ensuring the necessary flexibility provision towards the transmission system.

Components
- one generator
- one storage device
- one flexible load

Component parameters such that:
- feasibility *implies* feasibility in original model
- cost *approximates* cost in original model

Figure 10-4 Surrogate distribution grid model

11 Appendices

11.1 Appendix A – Tests input data

11.1.1 Transmission test system specifications

All AC buses are characterized by a nominal voltage magnitude of 240 kV as well as a maximal and a minimal operating voltage of 1.05 and 0.95 p.u. respectively.

AC branches in the test system (in blue) are characterized by their resistance, reactance, rating, maximum and minimum angle difference. The characteristics of the existing AC branches are summarized in the table below.

ID	From AC bus	To AC bus	R (ohm)	X (ohm)	Rating (MVA)	Min angle difference (degrees)	Max angle difference (degrees)
AC line 1	1	2	0.04	0.4	100	-60	60
AC line 2	1	4	0.06	0.6	80	-60	60
AC line 3	2	3	0.02	0.2	100	-60	60
AC line 4	2	4	0.04	0.4	100	-60	60

Table 26 - Existing AC branches characteristics

One of the objectives of FlexPlan is to consider high voltage DC lines (HVDC) in addition to traditional AC lines. The original IEEE 6-bus system is therefore extended by adding a DC bus at each bus to allow for potential DC branches. DC buses have nominal voltage magnitude of 320 kV, maximal and minimal operating voltages are 1.1 and 0.9 p.u. respectively.

The lines from buses 1 and 3 to bus 5 (Sardinia) are thereby implemented using DC branches (in yellow) for which the specifications are given in Table 27. At those nodes, the AC and DC buses are connected through converters whose specifications are given in Table 28.

ID	From DC bus	To DC bus	R (ohm)	Rating (MVA)
DC line 1	1	3	0.01	100
DC line 2	2	4	0.01	100

Table 27- Existing DC lines characteristics

Auxiliary converter losses (MW)[12]	1.1033
Linear converter losses (MW/MW)	0.887
Rated active power AC (MW)	110

Table 28- Specifications of existing converters

As in the original test system, AC bus 6 (Sicily) is initially not connected to the network. Line candidates are therefore made available to connect AC bus 6 to the network but also to potentially reinforce the existing grid. The following AC and DC branch candidates are made available. For AC lines an approximated construction cost of 1M€/km is considered for overhead lines and 4M€/km for underground cables. For candidate DC lines, the construction costs are 2.1 M€/km on land and 1.7M€/km for submarine cables.

ID	From AC bus	To AC bus	R (ohm)	X (ohm)	Rating (MVA)	Min angle difference (degrees)	Max angle difference (degrees)	Construction costs (M€)
Candidate AC line 1	1	3	0.02	0.2	100	-60	60	250
Candidate AC line 2	3	4	0.02	0.2	100	-60	60	247
Candidate AC line 3	4	6	0.02	0.2	100	-60	60	447
Candidate AC line 4	2	3	0.02	0.2	100	-60	60	508

Table 29 - Candidate AC lines characteristics

---

[12] It is to note that the losses arising from the transformer, reactor and filter impedances are included in the auxiliary and linear converter losses.

ID	From bus	To bus	R (ohm)	Rating (MVA)	Construction costs (M€)
Candidate DC line 1	3	6	0.01	330	891
Candidate DC line 2	4	6	0.01	330	710.2
Candidate DC line 3	5	6	0.01	330	977.4

Table 30 - Candidate DC lines characteristics

Finally, candidate converters are made available at AC buses 3, 4, 5 and 6 with the same specifications as the existing converters (see Table 3) except for the rated reactive power which is either 330 MW (constructions costs are then 40 M€) or 660 MW (constructions costs 80 M€).

11.1.2 Distribution test system specifications

The parameters of generators and loads of the MV distribution system are reported in Table 31 and Table 32 respectively.

Generator #	Bus #	Generator type	Pmin [MW]	Pmax [MW]	Qmin [MVar]	Qmax [MVar]
1	3	Photovoltaic	0	0.020	-0.005	0.005
2	4	Photovoltaic	0	0.020	-0.005	0.005
3	5	Photovoltaic	0	0.030	-0.008	0.008
4	5	Residential fuel cell	0	0.033	-0.006	0.006
5	6	Photovoltaic	0	0.030	-0.008	0.008
6	7	Wind turbine	0	1.500	-0.654	0.654
7	8	Photovoltaic	0	0.030	-0.008	0.008
8	9	Photovoltaic	0	0.030	-0.008	0.008
9	9	CHP diesel	0	0.310	-0.135	0.135
10	9	CHP fuel cell	0	0.212	-0.093	0.093
11	10	Photovoltaic	0	0.040	-0.011	0.011
12	10	Residential fuel cell	0	0.014	-0.003	0.003
13	11	Photovoltaic	0	0.010	-0.003	0.003

Table 31 - Parameters of generators of CIGRE European MV distribution benchmark network. The original source reports only the type and the maximum active power. The other parameters are estimates based on the available data.

Load #	Bus #	Apparent Power [MVA]		Power factor	
		Residential	Commercial / Industrial	Residential	Commercial / Industrial
1	1	15.300	5.100	0.98	0.95
2	3	0.285	0.265	0.97	0.85
3	4	0.445		0.97	
4	5	0.750		0.97	
5	6	0.565		0.97	
6	7		0.090		0.85
7	8	0.605		0.97	
8	9		0.675		0.85
9	10	0.490	0.080	0.97	0.85
10	11	0.340		0.97	
11	12	15.300	5.280	0.98	0.95
12	13		0.040		0.85
13	14	0.215	0.390	0.97	0.85

Table 32 - Parameters of loads of CIGRE European MV distribution benchmark network.

A set of 5 AC candidate branches has been added to the benchmark network, as reported in the table below.

Candidate branch #	Buses	Resistance [Ω]	Reactance [Ω]	Rating [MVA]	Investment type	Component type	Scaled investment cost
1	(15, 1)	0.095	0.955	50.0	replacement	OLTC (0.9-1.1 p.u. regulation)	0.80
2	(15,12)	0.190	1.910	25.0	addition in parallel	transformer	0.75
3	(12,13)	1.225	0.895	15.0	replacement	line	0.73
4	(12,13)	1.225	0.895	15.0	replacement	line	0.78
5	(13,14)	1.498	1.094	7.5	addition in parallel	line	0.45

Table 33 - Specifications of candidate AC branches added to the benchmark network.

11.1.3 Storage assets specifications for test cases

For the tests on the modified IEEE 6-bus system (transmission test system), we considered an existing storage device connected to AC bus 5 whose specifications are given below. The specifications of the storage assets are similar to those of a Pumped Hydro power plant. Note that aside from the "base" values given in the

table, the model has been tested for several values of the storage asset parameters (maximum energy content, absorption/injection efficiency, cost, etc.) as described in Section 9.3.1.1 (storage sensitivity analysis).

Maximum energy content (MWh)	1000
Maximum absorbed active power over a time step (MW)	200
Maximum injected active power over a time step (MW)	250
Absorption efficiency	0.9
Injection efficiency	0.9
Maximum absorbed active power over a year (MWh)	350 000 (∼ max 1 cycle/day)
Discharge rate (storage losses)	10^{-4}
Stationary energy inflow (storage gains) (MW)	0

Table 34 - Specifications of existing storage assets for tests on transmission system

Candidate storage assets are also made available in the transmission test system at AC bus 2 and 5 to potentially increase the storage capacity of the system. The parameters of the candidate assets are identical to the existing one (see Table 34), and the implementation costs are computed using data from [41]: 400€ / kW (power) + 40€/kWh (energy). For the storage asset specified above this means a construction cost of 140 M€ per candidate.

Maximum energy content (MWh)	2.4
Maximum absorbed power over a time step (MW)	0.6
Maximum injected power over a time step (MW)	0.6
Maximum absorbed reactive over a time step (MW)	0.6

Maximum injected reactive over a time step (MW)	0.6
Absorption efficiency	0.9
Injection efficiency	0.9
Maximum absorbed active power over a year (MWh)	1500 (~ max 2 cycles/day)
Discharge rate (storage losses)	10^{-4}
Stationary energy inflow (storage gains) (MW)	0

Table 35 - Specifications of existing storage assets for tests on distribution system

11.1.4 Flexible loads specifications for tests on demand flexibility

Parameter	Symbol	Unit	Base case value
ID of load point	n/a	n/a	5
Maximum energy not consumed (accumulated load reduction)	$E^{nc,max}$	MWh	1000
Superior bound on not consumed power (demand reduction)	$\Delta^{nce,max}$	p.u.	0.3
Inferior bound on not consumed power (demand reduction)	n/a	MW	0
Superior bound on upward demand shifted	$\Delta^{ds,dn,max}$	p.u.	0.3
Superior bound on downward demand shifted	$\Delta^{ds,dn,max}$	p.u.	1

Maximum energy (accumulated load) shifted downward	$E^{ds,max}$	MWh	1000
Recovery period for upward demand shifting	$\tau^{ds,up,rec}$	h	10
Recovery period for downward demand shifting	$\tau^{ds,dn,rec}$	h	10
Compensation for consuming less (i.e., voluntary demand reduction)	C^{nce}	€/MWh	0.1
Compensation for demand shifting	C^{ds}	€/MWh	0
Compensation for load curtailment (i.e., involuntary demand reduction)	C^{lc}	€/MWh	10
Investment costs	I	€	0
Whether load is flexible	n/a	(bool)	1
Carbon footprint costs	FP^{CO_2}	€	0.5

Table 36 - Specifications of flexible loads parameters for tests on transmission system

Parameter	Symbol	Unit	Base case value
Maximum energy not consumed	$E^{nc,max}$	MWh	1000
Superior bound on not consumed power	$\Delta^{nce,max}$	p.u.	0.05

Inferior bound on not consumed power	n/a	MW	0
Superior bound on upward demand shifted	$\Delta^{ds,dn,max}$	MW	0.1
Superior bound on downward demand shifted	$\Delta^{ds,dn,max}$	MW	0.1
Maximum energy (accumulated load) shifted downward	$E^{ds,max}$	MWh	1000
Recovery period for upward demand shifting	$\tau^{ds,up,rec}$	h	2
Recovery period for downward demand shifting	$\tau^{ds,dn,rec}$	h	2
Compensation for consuming less	C^{nce}	€/MWh	1.0437
Compensation for demand shifting	C^{ds}	€/MWh	0.104
Compensation for load curtailment	C^{lc}	€/MWh	57.1428
Investment costs	I	€	0
Whether load is flexible	n/a	n/a	1
CO2 costs	FP^{CO_2}	€	0
Power factor angle θ, giving the reactive power as Q = P × tan(θ)	φ^{flex}	n/a	See Table 38

Table 37 - Specifications of flexible loads parameters for tests on distribution system

Load #	Bus #	Active power consumption (MW)	Power factor angle
1	1	19.839	0.200
2	3	0.502	0.246
3	4	0.432	0.246
4	5	0.728	0.246
5	6	0.548	0.246
6	7	0.077	0.246
7	8	0.587	0.246
8	9	0.574	0.246
9	10	0.543	0.246
10	11	0.330	0.246
11	12	20.010	0.200
12	13	0.034	0.246
13	14	0.540	0.246

Table 38 - Specifications of load demand parameters for testing of demand flexibility model on distribution system.

Candidate branch #	Buses	Resistance [Ω]	Reactance [Ω]	Rating [MVA]	Investment type	Component type	Scaled investment cost
1	(1,2)	0.072	1	18.0	addition in parallel	line (underground cable)	1.0
2	(2,3)	0.112	3.16	18.0	addition in parallel	line (underground cable)	1.0

Table 39 - Specifications of candidate AC branches added to the distribution test system for demand flexibility tests.

11.1.5 Data and sources for reliability modelling tests

Source	Description	NOK/MWh
GARPUR_D3.2	Example high	53,000
GARPUR_D3.2	Example low	13,250
GARPUR_D3.2	UK domestic	116,000
GARPUR_D3.2	UK SMEs	440,800
GARPUR_D3.1	RBTS example	108,120
GARPUR_D3.1	Norwegian residential	12,296
GARPUR_D3.1	Norwegian industrial	104,516
OPAL_4-area_network_DiB	Norwegian Commercial	220,300
OPAL_4-area_network_DiB	Norwegian Industry	132,600
OPAL_4-area_network_DiB	Norwegian Residential	23,500
OPAL_4-area_network_DiB	Norwegian Agriculture	21,400
OPAL_4-area_network_DiB	Norwegian Public service	194,500
OPAL_4-area_network_DiB	Norwegian Energy-intensive Industry	58,200
Cost-benefit_HILP	Average for Norwegian city	55,000

Table 40 - Data and sources for reliability modelling test specifications

11.1.6 Investment candidates for tests on scenario reduction

The test case that was used to evaluate the scenario reduction methodology is equal to the base case test system for transmission. The test system differs however in terms of candidates, for the scenario reduction tests, the candidates are listed below in Table 41 to Table 44.

ID	From node	To node	R (ohm)	X (ohm)	Rating (MVA)	Min angle difference (degrees)	Max angle difference (degrees)	Construction costs (M€)
Candidate AC line	4	6	0.02	0.2	100	-60	60	588

Table 41- Candidate AC lines characteristics

ID	From bus	To bus	R (ohm)	Rating (MVA)	Construction costs (M€)
Candidate DC line 1	4	6	0.01	330	710.2
Candidate DC line 2	5	6	0.01	330	977.4

Table 42 - Candidate DC lines characteristics.

ID	AC bus	rating (MVA)	Construction costs (M€)
Candidate converter 1	6	330	40
Candidate converter 2	6	660	80
Candidate converter 3	4	330	40
Candidate converter 4	5	330	40

Table 43 - Candidate AC-DC converter characteristics

ID	AC bus	power rating (MVA)	energy rating (MWh)	installation + equipment costs (M€)
Candidate storage 1	2	400	10000	300
Candidate storage 2	5	400	10000	300
Candidate storage 3	1	400	10000	300

Table 44 - Candidate storage characteristics

11.1.7 Generation cost data, and investment candidate cost data used in the testing of the stochastic optimisation model formulation.

Generator ID	Connected node	Type	Generation cost in €/MWh
1	Centre North	Natural Gas	49.6
2	Centre South	Coal	94,1
3	Centre South	Wind	38.6
4	Sicily	Natural Gas	49.6
5	Sicily	Wind	38.6
6	Sicily	PV	42.8

Table 45 - Generation cost data used for the stochastic optimisation model

Type of candidate	Node(s)	Power rating[13]	Energy rating	Investment costs
AC Branch	1-2	100 MVA	n.a.	250 M€
AC Branch	3-4	100 MVA	n.a.	247 M€
AC Branch	4-6	100 MVA	n.a	588 M€
AC Branch	2-3	100 MVA	n.a.	508 M€
DC Branch	3-6	330 MW	n.a.	891 M€
DC Branch	4-6	330 MW	n.a.	710.2 M€

---

[13] Transmission capacity for branches and DC converters, or charging / discharging power for storage candidate

DC Branch	5-6	330 MW	n.a.	977.4 M€
DC Converter	6	330 MW	n.a.	400 M€
DC Converter	6	330 MW	n.a.	400 M€
DC Converter	6	660 MW	n.a.	800 M€
DC Converter	3	330 MW	n.a.	400 M€
DC Converter	4	330 MW	n.a.	400 M€
DC Converter	5	330 MW	n.a.	400 M€
Storage	2	± 250 MW	n.a.	750 M€
Storage	5	± 250 MW	n.a.	750 M€
torage	6	±1000 MW	n.a.	1600 M€

Table 46 - List of investment candidates for the testing of the stochastic model formulation

11.2 Appendix B – Additional test results

11.2.1 Additional test results on the storage model tests

In addition to the test case results on storage for transmission presented in Section 9.3.1.2, we also check that the absorbed and injected power, the maximum energy content and the maximum energy absorbed over a year are fulfilled at any time step of the returned solution, for both existing and candidate storage assets. The results of those checks are summarized in Table 47 below.

a. The bound on absorbed and injected power of storage assets over an hour is set to 250 MW, which corresponds to 2.5 p.u. (since the base power is 100 MVA). The maximum values in the obtained solution are 1.610 p.u. 1.607 for absorption and injection power respectively.

b. The maximum energy content corresponds to an energy level of 1. This threshold is reached in one time step for the candidate storage asset at AC bus 2, for the other storage assets the maximum energy content is never reached.

c. The energy absorbed over a year is 80 p.u. for the existing storage asset, 92 and 61 p.u. for the candidate (and implemented) storage assets. The bound is set at 350 000 MWh which corresponds to 3500 in p.u. We note that the initial bound of 2400 MWh/year had to be raised to 350 000 MWh/year because the problem was infeasible for a thousand hours.

d. The initial energy level is zero and after the last hour the energy level in all storage assets is back to zero.

	Storage 1 (AC bus 5)	Candidate Storage 1 (AC bus 2)	Candidate Storage 2 (AC bus 5)
Investment made	/	yes	yes
Max injected power	1.539	1.469	1.607
Max absorbed power	1.077	1.610	1.080
Max energy level	0.897	1.0	0.627
Absorbed energy over 1000 hours	80.245	91.579	60.531

Table 47 - Results of base case test on transmission test system for storage assets (1000 hours)

To check that the investment costs of storage assets are correctly taken into account in the objective function we simply set their investment costs to zero and re-run the optimization. Both storage candidates are again chosen and the difference in the objective value corresponds to the cost of the two storage assets (8 M€).

The inspection of the storage energy levels at a few time steps shows that the storage dynamic equation works as intended: storage losses as well as charge and discharge are correctly taken into account (see example of energy levels in two sequential time steps in Table 48).

Energy level at t_{39}	0.635800
Self-discharge at t_{40}(0.01% of energy level at t_{39})[14]	- 0.000064
Storage charge at t_{40}(input power times charging efficiency)	+ 0.066323 (0.073692 ·0.9)
Energy level at t_{40}	= 0.702059

Table 48 - Energy level of a storage asset at two sequential timesteps

---

[14] The self-discharge losses are very low because for the base case we use a "realistic" self-discharge rate for an electrical battery, in the sensitivity tests the model is tested with higher discharge rates.

The different sensitivity tests conducted on the transmission test system are summarized per parameter below. The solution returned for each sensitivity test is compared with the solution of the base case (with increased storage investment costs). As a reminder, in this case no storage was built and only the existing storage asset at AC bus 5 was used.

- <u>Maximum energy content</u>: we decrease $E_{j,y}^{max}$ om 1000 MWh to 100MWh for both the existing and candidate storage assets. As a consequence, the two storage candidates are invested in (which was not the case with the base case parameters and increased candidate costs). Indeed, hours 89 to 94 have a cumulated residual load (total system load minus total system capacity) of 328 MWh (3.28 p.u.) which requires the three storage assets to be fully charged beforehand (plus some load curtailment which is used at load 5).
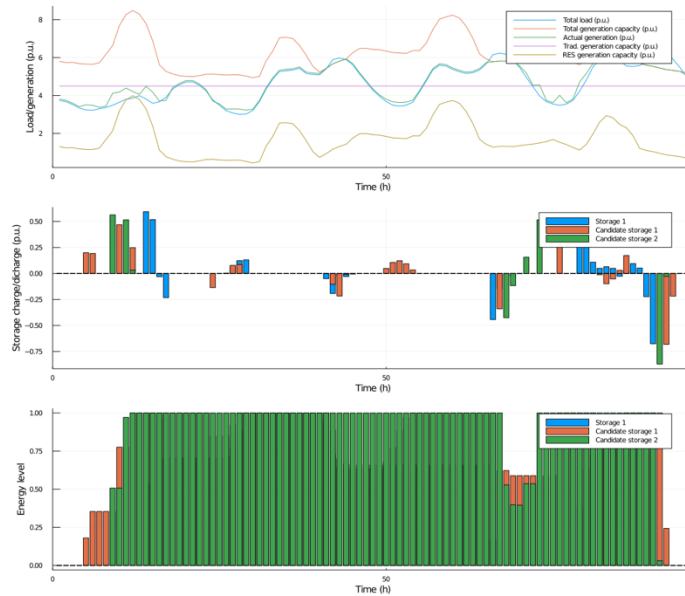
Figure 11-1 - Sensitivity test: maximum energy content

- <u>Maximum energy absorbed over a year</u>: we decrease $E_{j,y}^{abs,max}$ om 3500 p.u. to 5 p.u. In this case, a candidate storage is built as the bound on maximum energy absorbed over a year is reached for the existing storage asset.
- <u>Maximum absorbed power and maximum injected power</u>: in the base case test, the constraints on the maximum absorbed power ($P\_(j,y)^\wedge(abs, max) = 2,0\ p.u.$) and on the maximum injected power ($P\_(j,y)^\wedge(inj, max) = 2,5\ p.u.$) are never tight, i.e., the bounds are never reached. Increasing those bounds would therefore not change the solution but decreasing them below the maximum observed charge/discharge power will. We therefore limit both the maximum absorbed and injected power to 0,3 p.u. We see that in this case investments are made in the two candidate storage assets to increase the total power that can be injected from storage into the system. Indeed, the residual load of the system reaches 0,85 p.u. at hour 92, which can only be met (without curtailment) by using the three storage assets together.
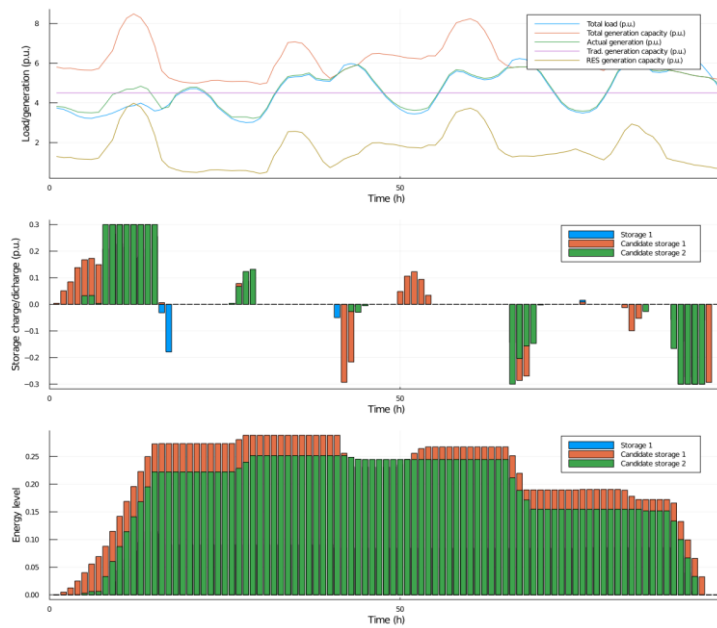
**Figure 11-2 -** Sensitivity test: maximum absorbed and injected power

- <u>Self-discharge rate (storage static losses):</u> we increase $dr_{j,t,y}$ from 10^{-4} to 0,9 which means that the storage asset loses 90% of its energy between each time step. As a result, the storage asset only absorbs power to inject it at the very next time step (no storage of power over more than one time step). The injected power is then much lower than the absorbed power due to the 90% losses in the energy level between the two time steps. As above, load curtailment therefore has to be used to reach a feasible solution (as we do not consider flexible loads for the tests on storage).
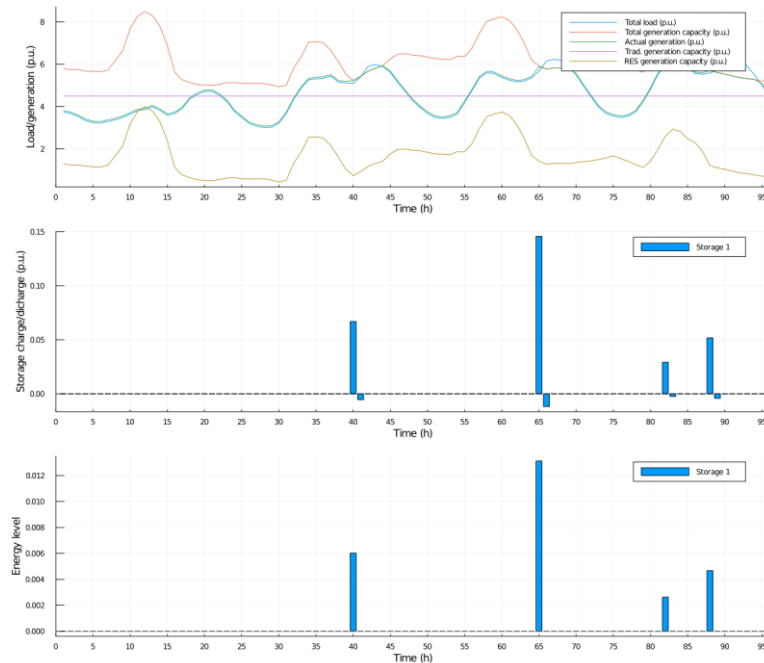
Figure 11-3 - Sensitivity test: self-discharge rate

- <u>Storage investment costs and status:</u>

- In the base case results, no investments are made in additional storage assets, only the existing storage was used.
- When we decrease the investment costs of storage candidates to 100 000 €, the candidate storage at AC bus 2 is invested in.
- When we deactivate the existing storage by setting its status to zero, an investment is performed in the candidate storage at AC bus 5.
- When we deactivate the existing storage while simultaneously increasing the costs of the candidate storage assets, no investment is made, and load curtailment was used instead.

- Power provided by external process: we can suppose that the existing pumped hydro storage is fed by a river by setting $\xi_{j,t,y}$ to nonzero values. If we therefore set a constant energy inflow of 100 MW per hour for the existing storage, we get the results below (Figure 11-4). The storage asset never absorbs power from the grid, it only discharges the energy received "for free" and as a consequence the actual system generation (green curve) is always below the total system load (blue curve).
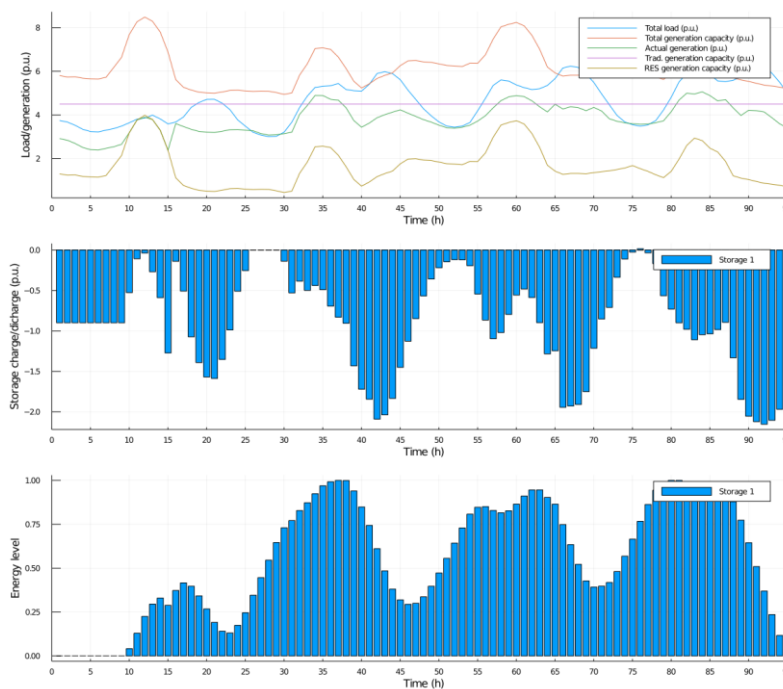
Figure 11-4 - Sensitivity test: stationary energy inflow

## 11.2.2 Additional results on the air impact model

The figures below present air quality impact results for 3 additional power plants in Italy using different technologies (on top of the results for the Enel power plant in Porto Marghera/Fusina already presented in Section 9.3.5.1). For each plant the mean day delta concentrations from CAMx and SIM_DDM are compared for all perturbation coefficients, for January and June. Results refer to each site for one of the 4 temporal modulations, while for the remaining ones, results are averaged over the three sites belonging to each group.
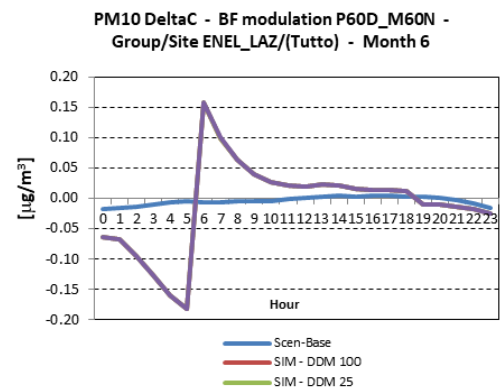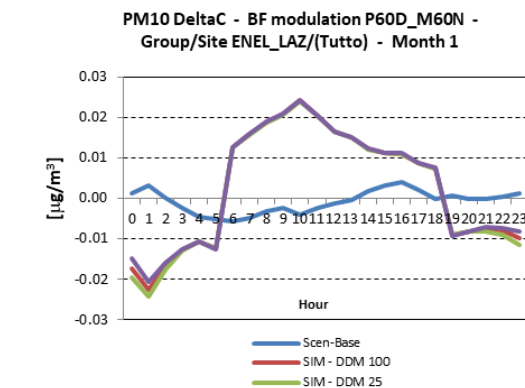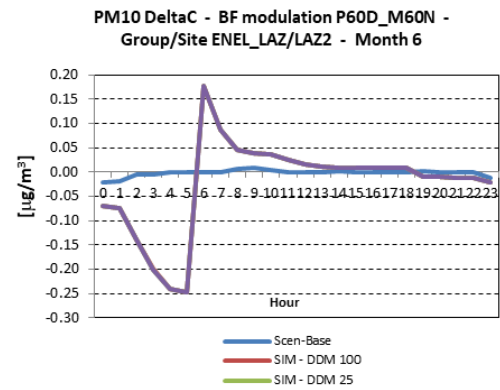
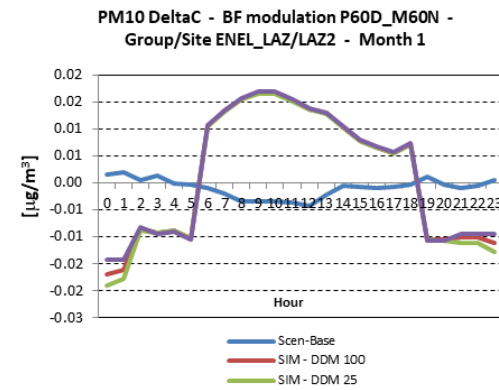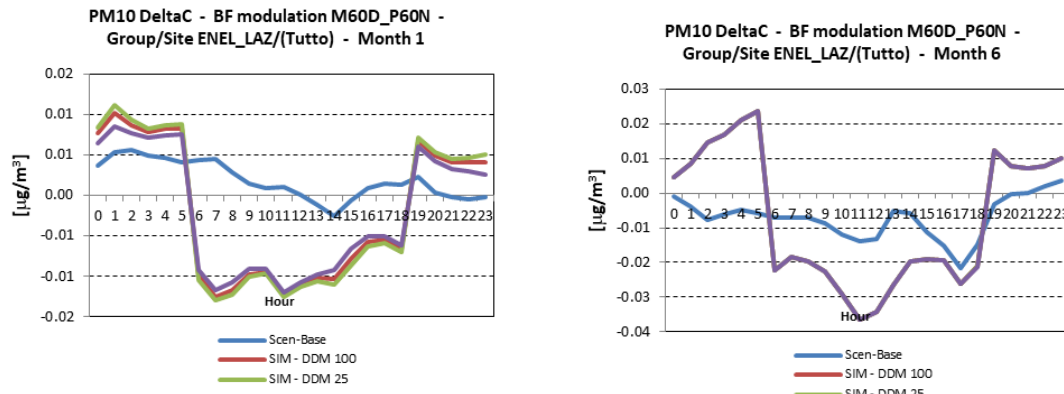**ENEL PRODUZIONE Spa - TORVALDALIGA NORD**
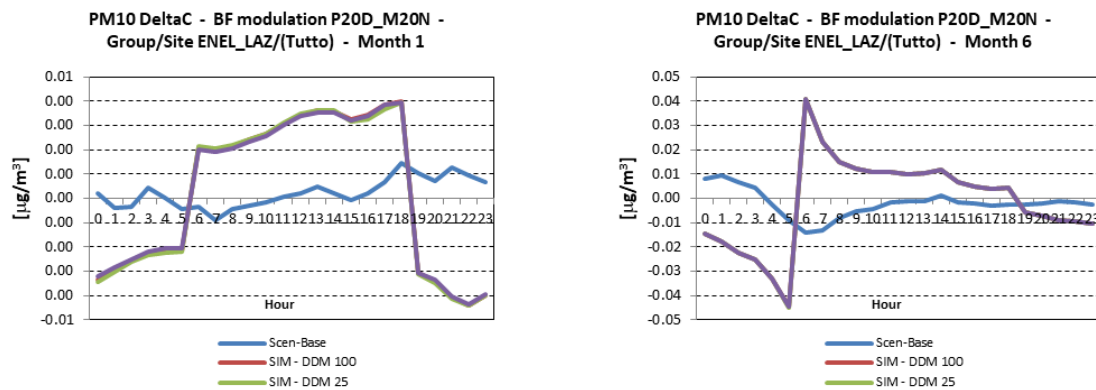
a)
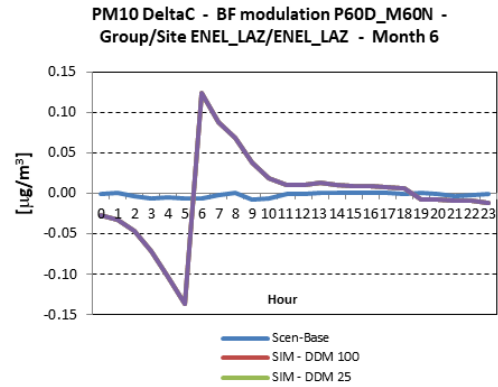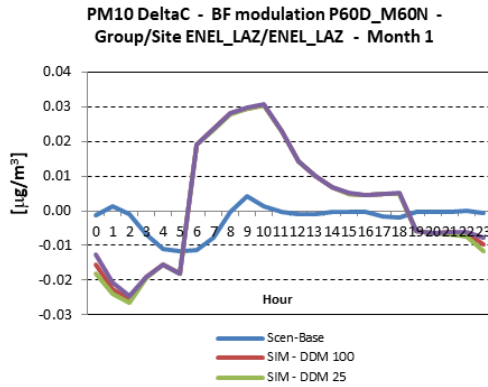
b)
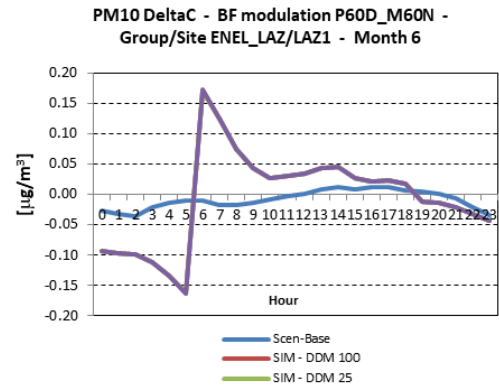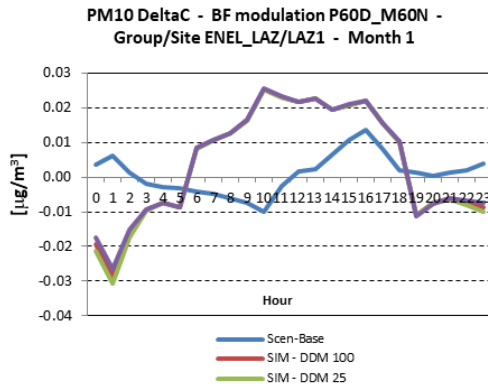
c)

d)

Figure 11-5 – Comparison of SIM results (red, green and purple lines) against Brute Force scenarios obtained with the full 3D model (light blue line) for January (left) and June (right) mean day concentrations for ENEL PRODUZIONE Spa - TORVALDALIGA NORD power plant. Results refer to receptors: a) ENEL_LAZ; b) LAZ1; c) LAZ2; d) average of all LAZ receptors and are computed for days implementing a P60D_M60N emission modulation (see text for details)

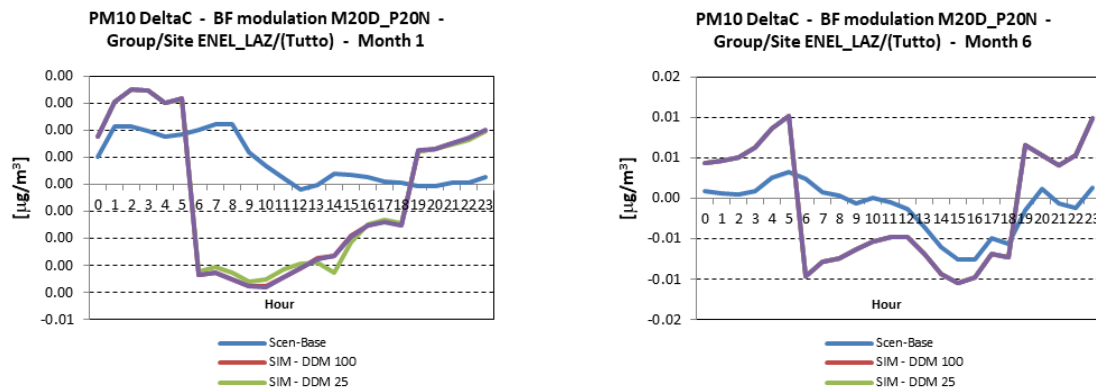a)

b)
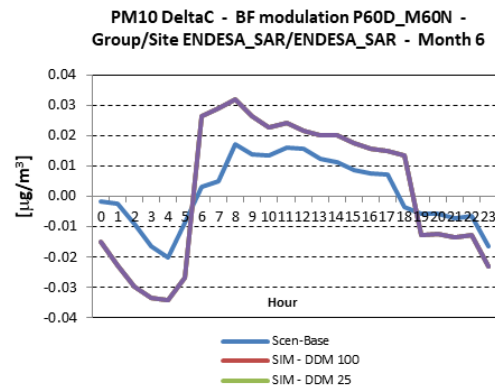

c)

Figure 11-6 – Comparison of SIM results (red, green and purple lines) against Brute Force scenarios obtained with the full 3D model (light blue line) for January (left) and June (right) mean day concentrations for

ENEL PRODUZIONE Spa - TORVALDALIGA NORD power plant. Results refer to scenario: a) M60D_P60N; b) P20D_M20N; c) M20D_P20N emission modulation (see text for details) and are computed as average of all LAZ receptors

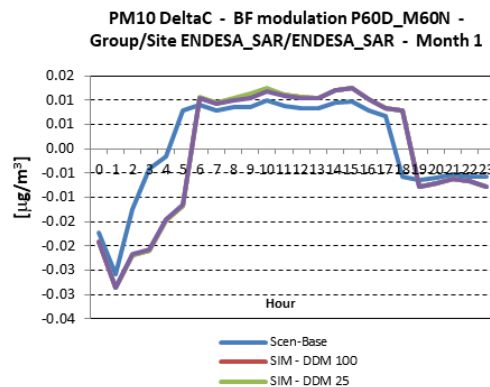**ENDESA ITALIA S.p.A**

a)

b)

c)

d)

Figure 11-7 – Comparison of SIM results (red, green and purple lines) against Brute Force scenarios obtained with the full 3D model (light blue line) for January (left) and June (right) mean day concentrations for ENDESA Italia (SpA) power plant. Results refer to receptors: a) ENDESA_SAR; b) SAR1; c) SAR2; d) average of all SAR receptors and are computed for days implementing a P60D_M60N emission modulation (see text for details)

a)

b)

c)
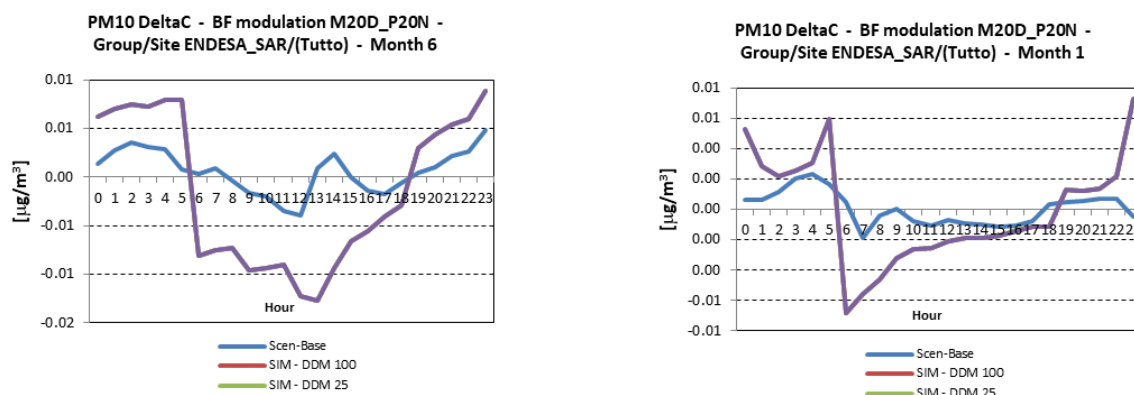
Figure 11-8 – Comparison of SIM results (red, green and purple lines) against Brute Force scenarios obtained with the full 3D model (light blue line) for Januar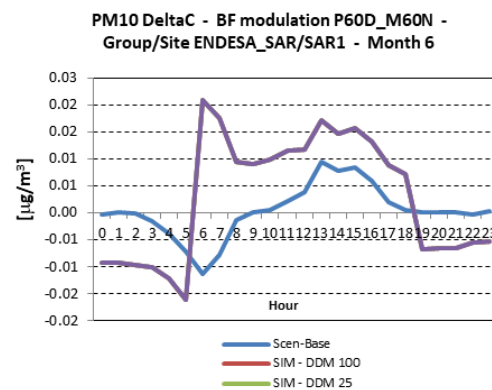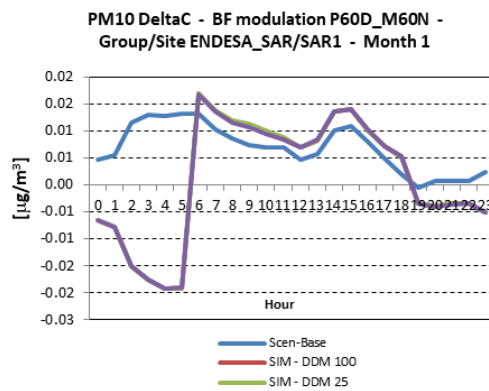y (left) and June (right) mean day concentrations for ENDESA Italia SpA power plant. Results refer to scenario: a) M60D_P60N; b) P20D_M20N; c) M20D_P20N emission modulation (see text for details) and are computed as average of all SAR receptors
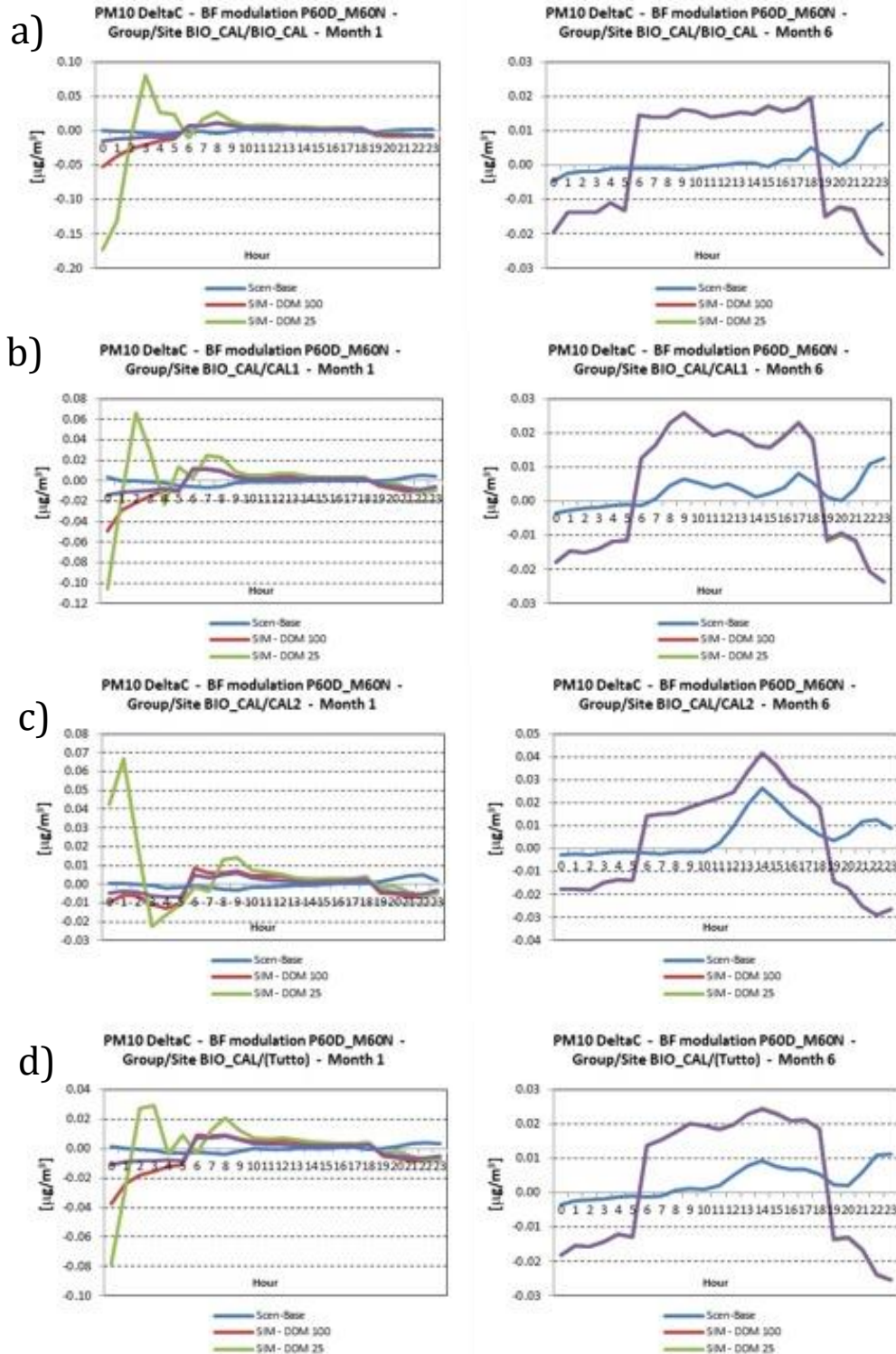
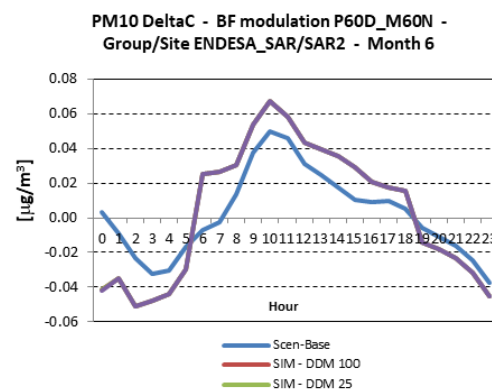**Figure 11-9 –** Comparison of SIM results (red, green and purple lines) against Brute Force scenarios obtained with the full 3D model (light blue line) for January (left) and June (right) mean day concentrations for BIOMASSE Italia (SpA) power plant. Results refer to receptors: a) BIO_CAL; b) CAL1; c) CAL2; d) average of all

CAL receptors and are computed for days implementing a P60D_M60N emission modulation (see text for details)
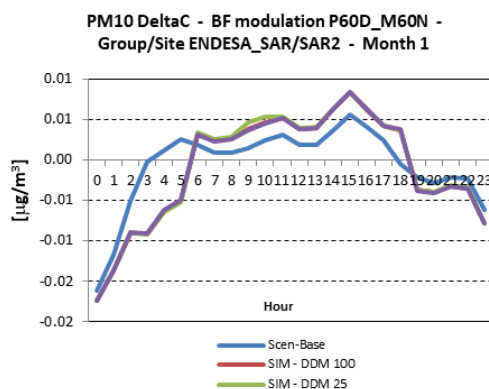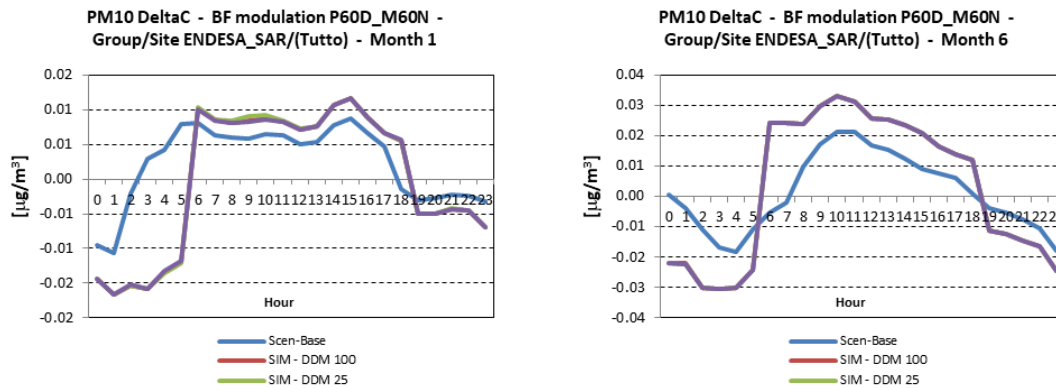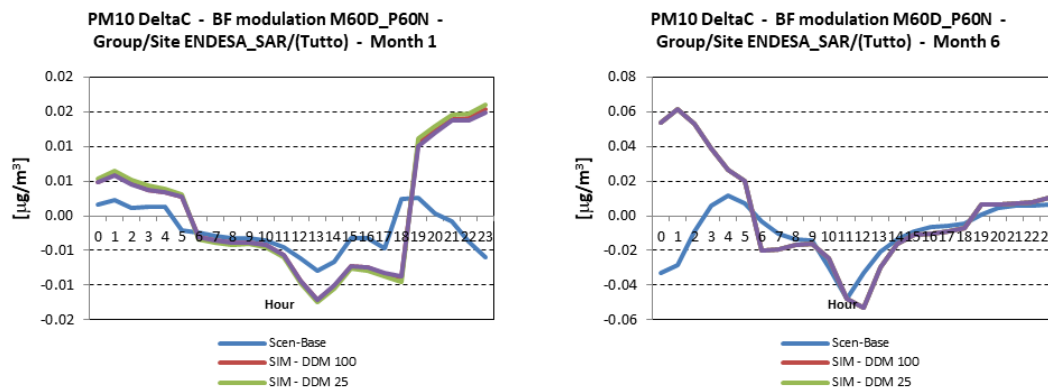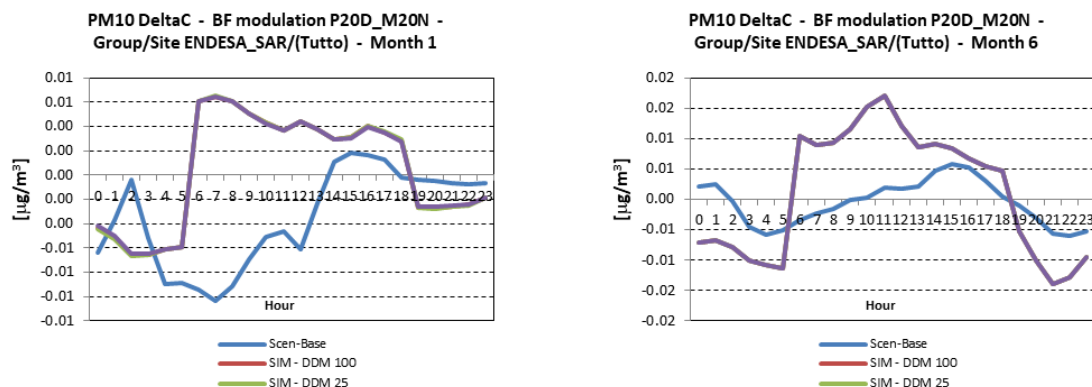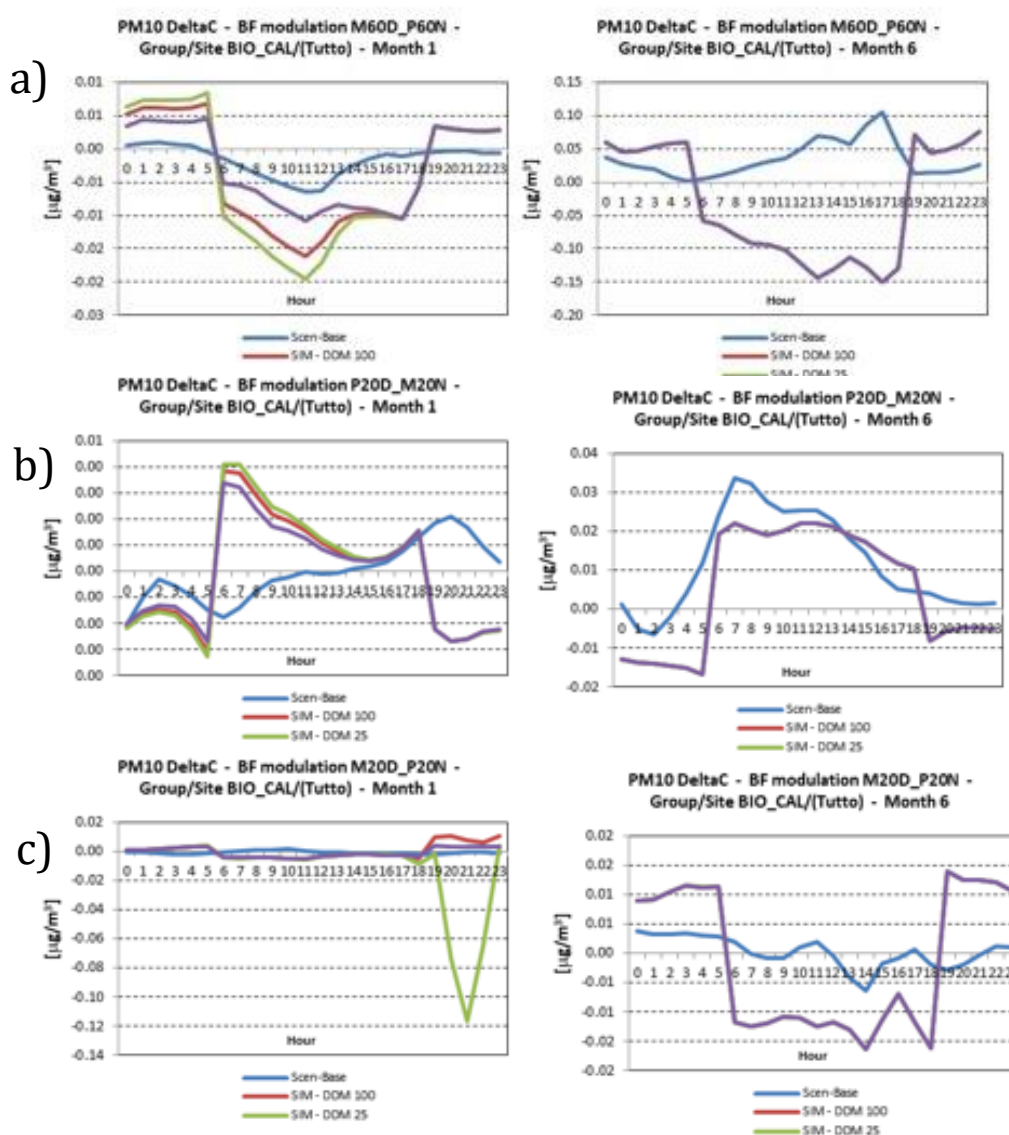

Figure 11-10 – Comparison of SIM results (red, green and purple lines) against Brute Force scenarios obtained with the full 3D model (light blue line) for January (left) and June (right) mean day concentrations for BIOMASSE Italia SpA power plant. Results refer to scenario: a) M60D_P60N; b) P20D_M20N; c) M20D_P20N emission modulation (see text for details) and are computed as average of all CAL receptors

Hour	Winter	Summer
1	3.29	2.74
2	3.29	2.74
3	3.29	2.74
4	3.29	2.74
5	3.29	2.74
6	3.29	2.74

7	3.39	2.82
8	3.39	2.82
9	3.49	2.90
10	3.49	2.90
11	3.49	2.90
12	3.49	2.90
13	3.49	2.90
14	3.49	2.90
15	3.49	2.90
16	3.49	2.90
17	3.49	2.90
18	3.49	2.90
19	3.39	2.82
20	3.39	2.82
21	3.29	2.74
22	3.29	2.74
23	3.29	2.74
24	3.29	2.74

Table 49 – $Prod_{ref}(g,seas,h)$ values [MWh] for BIO_EMR power plant.

Hour	Winter	Summer
1	2674.45	2223.36
2	2674.45	2223.36
3	2674.45	2223.36
4	2674.45	2223.36
5	2674.45	2223.36
6	2674.45	2223.36
7	2757.28	2291.75
8	2757.28	2291.75
9	2840.10	2360.14
10	2840.10	2360.14
11	2840.10	2360.14
12	2840.10	2360.14
13	2840.10	2360.14
14	2840.10	2360.14
15	2840.10	2360.14
16	2840.10	2360.14
17	2840.10	2360.14
18	2840.10	2360.14
19	2757.28	2291.75
20	2757.28	2291.75
21	2674.45	2223.36
22	2674.45	2223.36
23	2674.45	2223.36

| | 24 | 2674.45 | 2223.36 |

Table 50 – *Prod*ref(*g*,*seas*,*h*) values [MWh] for ENEL_LAZ ower plant.

Hour	Winter	Summer
1	1514.72	1259.24
2	1514.72	1259.24
3	1514.72	1259.24
4	1514.72	1259.24
5	1514.72	1259.24
6	1514.72	1259.24
7	1561.63	1297.97
8	1561.63	1297.97
9	1608.54	1336.70
10	1608.54	1336.70
11	1608.54	1336.70
12	1608.54	1336.70
13	1608.54	1336.70
14	1608.54	1336.70
15	1608.54	1336.70
16	1608.54	1336.70
17	1608.54	1336.70
18	1608.54	1336.70
19	1561.63	1297.97
20	1561.63	1297.97
21	1514.72	1259.24
22	1514.72	1259.24
23	1514.72	1259.24
24	1514.72	1259.24

Table 51 – *Prod*ref(*g*,*seas*,*h*) values [MWh] for ENEL_PUG power plant.

Hour	Winter	Summer
1	103.04	85.66
2	103.04	85.66
3	103.04	85.66
4	103.04	85.66
5	103.04	85.66
6	103.04	85.66
7	106.24	88.30
8	106.24	88.30
9	109.43	90.93
10	109.43	90.93
11	109.43	90.93

12	109.43	90.93
13	109.43	90.93
14	109.43	90.93
15	109.43	90.93
16	109.43	90.93
17	109.43	90.93
18	109.43	90.93
19	106.24	88.30
20	106.24	88.30
21	103.04	85.66
22	103.04	85.66
23	103.04	85.66
24	103.04	85.66

Table 52 – $Prod_{ref}(g,seas,h)$ values [MWh] for BIO_CAL power plant

Hour	Winter	Summer
1	162.06	134.72
2	162.06	134.72
3	162.06	134.72
4	162.06	134.72
5	162.06	134.72
6	162.06	134.72
7	167.07	138.87
8	167.07	138.87
9	172.09	143.01
10	172.09	143.01
11	172.09	143.01
12	172.09	143.01
13	172.09	143.01
14	172.09	143.01
15	172.09	143.01
16	172.09	143.01
17	172.09	143.01
18	172.09	143.01
19	167.07	138.87
20	167.07	138.87
21	162.06	134.72
22	162.06	134.72
23	162.06	134.72
24	162.06	134.72

Table 53 – $Prod_{ref}(g,seas,h)$ values [MWh] for ENDESA_SAR power plant

11.3 Appendix C – FlexPlan.jl Documentation

11.3.1 Overview

FlexPlan.jl is a Julia/JuMP package to carry out transmission and distribution network planning considering AC and DC technology, storage and demand flexibility as possible expansion candidates. Using time series input on renewble generation and demand, as well a list of candidates for grid expansion, a mixed-integer linear problem is constructed which can be solved with any commercial or open-source MILP solver. Some modelling features are:

- Multi-period, multi-stage formulation to model a number of planning years, and planning hours within years for a sequential grid expansion plan
- Stochastic formulation of the planning problem, based on scenario probabilities for a number of different time series
- Linearized DistFlow model considering reactive power and voltage magnitudes for radial distribution grids
- Extensive, parametrized models for storage, demand flexibility and DC grids
- Different decomposition methods for solving the large-scale MILP problem
- This package builds upon the PowerModels.jl[15] and PowerModelsACDC.jl[16] packages,and uses a similar structure.

The latest stable release of FlexPlan can be installed using the Julia package manager with:
] add "FlexPlan"

To install FlexPlan.jl in the development mode for further user extensions following command can be used:
] develop "https://github.com/Electa-Git/FlexPlan.jl.git"

The online user manual can be accessed through: https://electa-git.github.io/FlexPlan.jl/dev/

11.3.2 Running scripts within FlexPlan.jl

Some scripts have been provided in FlexPlan/test/scripts to test the package functionality. To run those scripts, an environment needs to be activated where import all the required packages, such as PowerModels.jl, PowerModelsACDC.jl or optimisation solvers can be added. This procedure is shown below:

---

[15] https://github.com/lanl-ansi/PowerModels.jl

[16] https://github.com/Electa-Git/PowerModelsACDC.jl

(1) In a Julia REPL, choose a directory where to create the environment:

```
julia> cd("path/to/env/dir")
```

(2) Enter the Pkg REPL by pressing] from the Julia REPL:

```
julia> ]
```

(3) Activate the environment:

```
pkg> activate .
```

(4) add the FlexPlan package:

```
pkg> add FlexPlan
```

(5) add every package required by the script. For example, if the script contains import Plots, then execute

```
pkg> add Plots
```

11.3.3 Structure of FlexPlan.jl scripts

A number of example scripts have been provided within FlexPlan.jl under "FlexPlan/examples". The general structure of the example scripts is as follows.

Declaration of the required packages and solvers

The required packages for FlexPlan.jl are PowerModels.jl and PowerModelsACDC.jl. One can declare the packages as follows, and use short names to access specific functions without having to type the full package name every time.

```
import PowerModels; const _PM = PowerModels
import PowerModelsACDC; const _PMACDC = PowerModelsACDC
import FlexPlan; const _FP = FlexPlan
```

Any other additional package that you might need, e.g., for printing, plotting, exporting results etc. can be declared in the same way. Also, the solution of the problem will require an MILP solver. As FlexPlan.jl is in the Julia / JuMP environment, it can be interfaced with any optimisation solver. You can declare and initialize the solver as follows:

```
import Cbc
optimizer = _FP.optimizer_with_attributes(Cbc.Optimizer, "logLevel"=>0)
```

Input data

The data model is very similar to the PowerModels.jl/PowerModelsACDC.jl data models. As such, a data dictionary containing all information is passed to the optimisation problem. The standard network elements such as generators, buses, branches, etc. are extended with the existing and candidate storage and demand flexibility elements. The multi-network modelling functionality of the PowerModels.jl package is used to represent the different number of scenarios, planning years and planning hours within

the year. The procedure is further explained under section 11.3.6 . The package contains some sample time-series as well as grid data, which is located under FlexPlan/test/data. This data has been used in for the validation of the model as presented in this document.

The data dictionary can be created by the user directly (see section Data model for complete description) but also provided as a Matpower[17] file, as within PowerModels.jl. Using the MatPower file, only the grid data dictionary will be created. In order to add time-series and scenario information to the data dictionary, a number of additional functions are required. An example of the process is illustrated below for the combined T&D planning model:

```
# Define number of hours
number_of_hours = 24
# Transmission network instance (all data preparation except for make_multinetwork() call)
t_file = "./test/data/case6/case6_2030.m" # Input case for transmission network
t_data = _FP.parse_file(t_file) # Parse input file to obtain data dictionary
_FP.add_dimension!(t_data, :hour, number_of_hours) # Add dimension, e.g. hours
_FP.add_dimension!(t_data, :scenario, Dict(1 => Dict{String,Any}("probability"=>1)), metadata = Dict{String,Any}("mc"=>true)) # Add dimension, e.g. scenarios
_FP.add_dimension!(t_data, :year, 1; metadata = Dict{String,Any}("scale_factor"=>1)) # Add_dimension, e.g. years
_FP.add_dimension!(t_data, :sub_nw, 1) # Add dimension, e.g. underlying networks
_FP.scale_data!(t_data) # Scale investment & operational cost data based on planning years & hours
t_data, t_loadprofile, t_genprofile = create_profile_data_italy!(t_data) # Load time series data based demand and RES profiles of the six market zones in Italy from the data folder
t_time_series = create_profile_data(number_of_hours, t_data, t_loadprofile, t_genprofile) # Create time series data to be passed to the data dictionary
```

Coupling of transmission and distribution networks

FlexPlan.jl provides the possibility to couple multiple radial distribution networks to the transmission system, for solving the combined T&D grid expansion problem. For the meshed transmission system the linearized 'DC' power flow formulation is used, whereas radial networks are modelled using the linearised DistFlow model (more information can be found under section 4).

To create the data for radial networks you can use following approach:

```
## Distribution network instance 1 (all data preparation except for make_multinetwork() call)
d_file    = "test/data/cigre_mv_eu/cigre_mv_eu.m" # Input case for distribution networks
scale_load = 1.0 # Scaling factor of loads
```

---

[17] https://matpower.org

```
scale_gen  = 1.0 # Scaling factor of generators

d_data_1 = _FP.parse_file(d_file) # Parse input file to obtain data dictionary

_FP.add_dimension!(d_data_1, :hour, number_of_hours) # Add dimension, e.g. hours

_FP.add_dimension!(d_data_1, :scenario, Dict(1 => Dict{String,Any}("probability"=>1))) # Add dimension, e.g. scenarios

_FP.add_dimension!(d_data_1, :year, 1; metadata = Dict{String,Any}("scale_factor"=>1)) # Add dimension, e.g. years

_FP.add_dimension!(d_data_1, :sub_nw, 1) # Add dimension, e.g. underlying networks

_FP.shift_ids!(d_data_1, _FP.dim_length(t_data)) # Shift network IDs to avoid overwriting those of transmission network

_FP.scale_data!(d_data_1) # Scale investment & operational cost data based on planning years & hours

d_time_series = create_profile_data_cigre(d_data_1, number_of_hours; scale_load, scale_gen) # Load time series data based
```

demand and RES profiles of the six market zones in Italy from the data folder

```
_FP.add_td_coupling_data!(t_data, d_data_1; t_bus = 1, sub_nw = 1) # Connect the first distribution network to bus 1 of
```

transmission network.

Note that a number of different distribution networks can be created in the same way. Eventually, all networks are coupled (for an example with one transmission and two distribution networks) using:

```
## Multinetwork data preparation

t_mn_data = _FP.make_multinetwork(t_data, t_time_series) # Merge transmission data & time series data

d_data_1 = _FP.make_multinetwork(d_data_1, d_time_series) # Merge data & time series data for distribution network 1

d_data_2 = _FP.make_multinetwork(d_data_2, d_time_series) # Merge data & time series data for distribution network 2

d_mn_data = _FP.merge_multinetworks!(d_data_1, d_data_2, :sub_nw) # Merge the two distribution networks in a single data
```

dictionary

Solving the optimisation problem

Finally, the problem can be solved using (example of planning problem with storage & demand flexibility candidates):

```
result = _FP.flex_tnep(t_mn_data, d_mn_data, _PM.DCPPowerModel, _FP.BFARadPowerModel, optimizer; setting=s)
```

Inspecting results

To obtain power flow results, you can use the standard `print_summary` function of PowerModels.jl. Further, there are number of possibilities to plot your time series results and also a .kml export, if you provide the latitude and longitude of the buses as an additional entry in your `data["bus"]` dictionary. Please consult `FlexPlan/examples` for different plotting possibilities.

11.3.4 Problem types

The FlexPlan.jl package contains the following problem types:

T(D)NEP problem with storage candidates

This problem solves the AC/DC grid TNEP problem considering existing and candidate storage candidates. As such, starting from an AC / (DC) network with existing storage devices, the optimisation problem finds the best AC and DC grid investments as well as storage investments. The problem is defined both for transmission networks, using the linearised 'DC' power flow model as well as radial distribution grids using the linearised 'DistFlow' formulation. The problem can be solved using the following functions for the transmission and distribution networks respectively:

```
result_tnep = FlexPlan.strg_tnep(data, PowerModels.DCPPowerModel, solver; setting)
result_dnep = FlexPlan.strg_tnep(data, FlexPlan.BFARadPowerModel, solver; setting)
```

TNEP problem with storage candidates and demand flexibility (Flexible T(D)NEP)

This problem solves the AC/DC grid TNEP problem considering existing and candidate storage candidates as well demand flexibility. As such, starting from an AC / (DC) network with existing storage devices, the optimisation problem finds the best AC and DC grid investments as well as storage and demand flexibility investments. The problem is defined both for transmission networks, using the linearised 'DC' power flow model as well as radial distribution grids using the linearised 'DistFlow' formulation. The problem can be solved using the following functions for the transmission and distribution networks respectively:

```
result_tnep = FlexPlan.flex_tnep(data, PowerModels.DCPPowerModel, solver; setting)
result_dnep = FlexPlan.flex_tnep(data, FlexPlan.BFARadPowerModel, solver; setting)
```

Additionally, this particular problem can also be solved for both transmission and distribution networks combined, using specific data for both the transmission and the distribution network:

```
result_t_and_d_nep = FlexPlan.flex_tnep(t_data, d_data, PowerModels.DCPPowerModel, FlexPlan.BFARadPowerModel, solver; setting)
```

Stochastic flexible T(D)NEP

This problem type extends the multi-year, multi-hour planning problem for a number of scenarios, e.g., variations of the planning year, and optimizes the investments taking into account the explicit scenario probabilities. As such, the objective is extended as described in section 0, w.r.t. to the flexible T(D)NEP problem. The problem is defined both for transmission networks, using the linearised 'DC' power flow model as well as radial distribution grids using the linearised 'DistFlow' formulation. The problem can be solved using the following function:

```
result_tnep = FlexPlan.stoch_flex_tnep(data, PowerModels.DCPPowerModel, solver; setting)
result_dnep = FlexPlan.stoch_flex_tnep(data, FlexPlan.BFARadPowerModel, solver; setting)
```

11.3.5 Network formulations

Two different network formulations have been used in the FlexPlan package:

- `PowerModels.DCPPowerModel` is a linearised 'DC' power flow formulation that represents meshed AC/DC transmission networks;
- `FlexPlan.BFARadPowerModel` is a linearised 'DistFlow' formulation that represents radial AC distribution networks.

For the comprehensive formulation of the network equations, along with the detailed model for storage and demand flexibility, the readers are referred to sections 4 and 5 of this document.

11.3.6 Modelling assumptions

When preparing data for problems spanning a multi-year horizon (here the word *year* indicates an investment period: different investment decisions can be made in different years), investment candidates must adhere to the following two assumptions:

- If a candidate exists in a year, then it exists in all subsequent years and is defined in the same row of the corresponding table in the input data files.
- Each candidate has the same parameters in all the years in which it exists, except for the cost which may vary with the years.

These assumptions are used not only when parsing input data files, but also in some variables/constraints where an investment candidate must be tracked along years.

Model dimensions

All the optimization problems modeled in FlexPlan are multiperiod and make use of the following dimensions:

- `hour`: the finest time granularity that can be represented in a model. During an hour, each continuous variable has a constant value.
- `year`: an investment period. Different investment decisions can be made in different years.
- `scenario`: one of the different possible sets of values related to renewable generation and consumption data.

These dimensions must be defined in each model by calling the function `add_dimension!` on single-period data dictionaries. The `add_dimension!` function takes the name of the dimension as input in form of key, as well as either integer values, e.g., for number of hours or years, or a dictionary, e.g., containing multiple scenarios. In the case of scenario input, probabilities and other meta data can be added. An example can be found below:

```
_FP.add_dimension!(data, :hour, number_of_hours) # Add dimension, e.g. number of hours
_FP.add_dimension!(data, :scenario, Dict(1 => Dict{String,Any}("probability"=>1)), metadata = Dict{String,Any}("mc"=>true)) # Add dimension, e.g. number of scenarios
_FP.add_dimension!(t_data, :year, 1; metadata = Dict{String,Any}("scale_factor"=>1)) # Add dimension of years, using cost scaling factors in metadata
```

Data model

FlexPlan.jl extends data models of the PowerModels.jl and PowerModelsACDC.jl packages by including candidate storage devices, `:ne_storage,` additional fields to parametrize the demand flexibility models which extend `:load`, some additional parameters to both existing and candidate storage devices to represent external charging and discharging of storage, e.g., to represent natural inflow and dissipation of water in hydro storage, some additional parameters extending `:gen` to include air quality impact and CO_2 emission costs for the generators. Section 2 of this document provides the full list of set, variables, and parameters of the optimisation problem.

11.4 Appendix D – OptimalTransmissionRouting.jl documentation

OptimalTransmissionRouting.jl is a Julia/JuMP package to determine the optimal transmission system route considering spatial information. The underlying principle is that spatial information coming from an image file is converted to an array of installation cost weights. To that end spatial information from http://www.eea.europa.eu/data-and-maps/data/corine-land-cover-2006-raster-2 is used in order to distinguish between different spatial areas. The created array represents a weighted graph connecting a number of nodes horizontally, vertically and diagonally with edge-weights reflecting the investment and installation costs for each spatial area of the input spatial map. Using the A-star algorithm, the shortest path in this weighted graph is found, which provides the least cost transmission path.

11.4.1 Usage

The first step is to define the locations of the sending (bus 1) and receiving (bus 2) end buses in latitude and longitude using a dictionary, e.g.,

```
bus1 = Dict{String, Any}()
bus1["longitude"] = 16.2487678
bus1["latitude"] = 40.358515
bus2 = Dict{String, Any}()
bus2["longitude"] = 14.1482998
bus2["latitude"] = 37.5900782
```

The spatial weights are be defined using the chosen underground cabling strategy as input as a string, using:

```
spatial_weights, voltages, resolution, impedances = OTR.define_weights_voltages(strategy)
```

Here, the three possible strategies are:

```
strategy = "all_permitted"
strategy = "OHL_on_existing_corridors"
strategy = "cables_only"
```

Furter the vertices of the graph are created, assigning each vertex to one or more spatial regions, e.g., urban areas, mountain, sea, existing infrastructure, open field, using:

```
rgb_values, nodes_lp, boundaries, plot_dictionary = OTR.convert_image_files_to_weights(bus1, bus2)
```

Eventually, the technical characteristics of the connection considered need to be provided in form of an input dictionary. This is used to calculate parameters such as the number of circuits and cross-sections of the overhead line and underground cable sections, as well as resistance, and reactance parameters to estimate losses and determine reactive power needs. The required and optional fields are shown below:

```
input_data = Dict{String, Any}()
input_data["resolution_factor"] = 2 # resolution_factor 1,2,3, ... to speed up algorithm
input_data["algorithm_factor"] = 1 # algorithm_factor 1.....1.3 to speed up Astar algorithm, goes at expense of accuracy
input_data["distance"] = 2.5  # do not change: this is the standard resolution of the environmental data
input_data["algorithm"] = "Astar"  # "Astar" or "Dijkstra"
input_data["voltages"] = voltages # transmission voltages
input_data["spatial_weights"] = spatial_weights # spatial weights
input_data["rgb_values"] = rgb_values # Spatial data as RGB values
input_data["boundaries"] = boundaries # VBoundaries of the area (to avoid using full European range)
input_data["overlapping_area_weight"] = "average" # "average" = average weight of the spatial weights; "minimum" = minimum of the overlapping weights
input_data["strategy"] = strategy # or "OHL_on_existing_corridors" or "cables_only"
input_data["losses"] = 0.01 # proxy for losses
input_data["lifetime"] = 30 # lifetime: NOT USED in FLEXPLAN
input_data["interest"] = 0.02 # Interest: NOT USED in FLEXPLAN
input_data["technology"] = "dc" # or "dc"
input_data["power_rating"] = 2000 # power rating
input_data["start_node"] = Dict{String, Any}()
input_data["start_node"]["x"] = nodes_lp["x1"]
input_data["start_node"]["y"] = nodes_lp["y1"]
input_data["end_node"] = Dict{String, Any}()
input_data["end_node"]["x"] = nodes_lp["x2"]
input_data["end_node"]["y"] = nodes_lp["y2"]
input_data["impedances"] = impedances # Provide look-up table for OHL & OGC impedances
```

The optimisation of the routing is carried out using following command:

```
spatial_data, spatial_data_matrices, cost_data, equipment_data, c_tot, optimal_path, ac_dc, ac_cab, dc_cab, route_impedance, route_legnth = OTR.do_optimal_routing(input_data)
```

During the route optimisation, following aspects are considered:

- Maximum offshore length for AC cables
- Costs for switching between overhead lines and underground cables, e.g., transition station costs
- In case of HVDC technology: Costs of the converter stations
- Chosen strategy:
    - OHL can be installed everywhere (depending on the chosen weights)
    - OHL can only be installed in existing infrastructure corridors
    - Only underground cables can be installed
- If certain vertices belong to multiple spatial areas following possibilities can be chosen to assign a unique weight:
    - The average of both weights
    - The minimum of both weights

The tool delivers following outputs:

- c_tot, the total cost of the transmission link
- cost_data, the per km cost of equipment
- equipment_data, number of circuits, bundles and cross-sections
- route_impedance, resistance, admittance, and capacitance in Ohm and in pu (base power = 100 MVA)
- route_legnth, the total length, as well as length of the OHL and UGC sections

Using the following command, a plot of the obtained solution is created which is saved a pdf-file:

`OTR.plot_result(plot_dictionary, input_data, spatial_data, spatial_data_matrices, optimal_path)`

12 References

[1]   FlexPlan Consortium, "D4.1 Pan-European scenario data.," 2020.

[2]   J. Dave, H. Ergun, T. An, J. Lu and D. Van Hertem, "TNEP of meshed HVDC grids: 'AC', 'DC' and convex formulations," *IET GENERATION TRANSMISSION & DISTRIBUTION,* vol. 13 (24), pp. 5523-5532, 2019.

[3]   H. Ergun, J. Dave, D. Van Hertem and F. Geth, " Optimal Power Flow for AC/DC Grids: Formulation, Convex Relaxation, Linear Approximation and Implementation," *IEEE TRANSACTIONS ON POWER SYSTEMS,* 2019.

[4]   J. Guo, Y. Fu, Z. Li and M. Shahidehpour, "Direct Calculation of Line Outage Distribution Factors," *IEEE Transactions on Power Systems,* vol. 24, no. 3, pp. 1633 - 1634.

[5]   F. Geth, "Convex Power Flow Models for Scalable Electricity Market Modelling," *CIRED - Open Access Proceedings Journal,* vol. 1, pp. 989-993, 2017.

[6]   L. Gan and S. H. Low, "Convex relaxations and linear approximation for optimal power flow in multiphase radial networks," *2014 Power Systems Computation Conference, Wroclaw, Poland,* pp. 1-9, 2014.

[7]   J. Romero Agüero, "Improving the efficiency of power distribution systems through technical and non-technical losses reduction," *PES T&D 2012, Orlando, FL, USA,* pp. 1-8, 2012.

[8]   A. Evans, V. Strezov and T. J. Evans, "Assessment of utility energy storage options for increased renewable energy penetration," *Renewable and Sustainable Energy Review,* vol. 16, pp. 4141-4147, 2012.

[9]   O. Wolfgang, A. Haugstad, B. Mo, A. Gjelsvik, I. Wangensteen and G. Doorman, "Hydro reservoir handling in Norway before and after deregulation," *Energy,* vol. 34, no. 10, p. 1642–1651, 2009.

[10]   A. Helseth and A. C. Geber de Melo, "Scheduling Toolchains in Hydro-Dominated Systems : evolution, Current Status and Future Challenges for Norway and Brazil," SINTEF Energy Research, Trondheim, Norway, 2020.

[11]   Ramboll, *CAMx User's Guide Version 7.00,* Novato, CA: Ramboll, 2020.

[12]   A. Dunker, G. Yarwood, J. Ortmann and G. Wilson, "The decoupled direct method for sensitivity analysis in a three-dimensional air quality model - implementation, accuracy, and efficiency," *Environmental Science & Technology,* vol. 36, no. 13, pp. 2965-2976, 2002.

[13]   WHO Regional Office for Europe, "Health Risk of air pollution in Europe - HRAPIE project. Recommendations for concentration-response function for cost-benefit analysis of particulate matter, ozone and nitrogen dioxide," 2013.

[14]   U. Im and et al, "Assessment and economic valuation of air pollution impacts on human health over Europe and the United States as calculated by a multi-model ensemble in the framework of AQMEII3," *Atmospheric Chemisttry and Phisycs,* no. 18, p. 5967–5989, 2018.

[15]   A. Alberini, A. Hunt and A. Markandya, "Willingness to pay to reduce mortality risks: Evidence from a three-country contingent valuation study," *Environ. Resour. Econ.,* vol. 33, p. 251–264, 2006.

[16]   G. Wernet, C. Bauer, B. Steubing, J. Reinhard, E. Moreno-Ruiz and B. Weidema, "The ecoinvent database version 3 (part I): overview and methodology," *The International Journal of Life Cycle Assessment,* vol. 21, no. 9, p. 1218–1230, 2016.

[17]   I. Tsiropoulos, D. Tarvydas and N. Lebedeva, "Li-ion batteries for mobility and stationary storage applications – Scenarios for costs and market growth," Publications Office of the European Union, Luxembourg, 2018.

[18]   M. Hegedić, T. Opetuk, G. Đukić and H. Drašković, "Life cycle assessment of power transformer-case study," *Management of Technology–Step to Sustainable,* 2016.

[19]   M. Mansilha, M. Brondani, F. Farret, L. Cantorski da Rosa and R. Hoffmann, "Life Cycle Assessment of Electrical Distribution Transformers: Comparative Study Between Aluminum and Copper Coils," *Environmental Engineering Science,* vol. 36, no. 1, 2018.

[20]   M. Mouhamad and P. Lauzevis, "Life cycle assessment and inrush currents measurement of amorphous transformers," in *Electricity Distribution (CIRED 2013), 22nd International Conference and Exhibition*, 2013.

[21]   R. Santos Jorge, T. R. Hawkins and H. E. G., "Life cycle assessment of electricity transmission and distribution—part 2: transformers and substation equipment," vol. 17, pp. 184-191, 2012.

[22]   J. Borghetto, F. Barberis, R. Berti and R. Passaglia, "Comparative life cycle assessment of a MV/LV transformer with an amorphous metal core and a MV/LV transformer with a grain-oriented magnetic silicon steel core," in *20th International Conference on Electricity Distribution* , Prague, 2009.

[23]   R. Berti, F. Barberis, V. Rossi and L. Martini, "Compararison of the ecoprofiles of superconducting and conventional 25 MVA transformers using the life cycle assessment methodology," in *20th International Conference and Exhibition on Electricity Distribution - Part 1*, Prague, 2009.

[24]   C. I. Jones and M. C. McManus, "Life-cycle assessment of 11 kV electrical overhead lines and underground cables," *Journal of Cleaner Production,* vol. 18, pp. 1464-1477, 2010.

[25]   S. Bumby, E. Druzhinina, R. Feraldi, D. Werthmann, R. Geyer and J. Sahl, "Life Cycle Assessment of Overhead and Underground Primary Power Distribution," *Environmental Science and Technology,* vol. 44, no. 14, p. 5587–5593, 2010.

[26]   S. Bumby, E. Druzhinina, R. Feraldi, D. Werthmann, R. Geyer and J. Sahl, "Comparative Life Cycle Assessment (LCA) of overhead and underground medium voltage power distribution," in *IEEE International Symposium on Sustainable Systems and Technology*, 2009.

[27]   R. Santos Jorge, T. R. Hawkins and E. G. Hertwich, "Life cycle assessment of electricity transmission and distribution—part 1: power lines and cables," *The Internationa Journal of Life Cycle Assessment,* vol. 17, pp. 9-15, 2012.

[28]   RSE, "LCA and LCC of lithium-ion stationary batteries in the Italian scenario," 2020.

[29]   H. Ergun, B. Rawn, R. Belmans and D. Van Hertem, "Optimization of Transmission Technology and Routes for Pan – European Electricity Highways Considering Spatial Aspects (Paper No. 467697)," PowerTech, Eindhoven, 2015.

[30]   E. Dijkstra, "A note on two problems in connexion with graphs," *Numer. Math.,* vol. 1, no. 1, pp. 269-271, 1959.

[31]   W. Zeng and R. L. Church, "Finding shortest paths on real road networks: the case for A*," *International Journal of Geographical Information Science,* vol. 23, no. 4, p. 531–543, 2009.

[32]   FlexPlan Consortium, "D1.1 Monte Carlo scenario generation and reduction," 2020.

[33]   J. F. Benders, "Partitioning procedures for solving mixed-variables programming problems," *Numer. Math.,* vol. 4, p. 238–252, 1962.

[34]   G. Micheli, M. T. Vespucci, M. Stabile, C. Puglisi and A. Ramos, "A two-stage stochastic MILP model for generation and transmission expansion planning with high shares of renewables," *Energy Syst,* 2020.

[35]   JuliaLang, "Julia Programming Language," [Online]. Available: https://julialang.org/. [Accessed 12 3 2021].

[36]   C. Coffrin, R. Bent, K. Sundar, Y. Ng and M. Lubin, "PowerModels. JL: An Open-Source Framework for Exploring Power Flow Formulations. In Proceedings of the 2018 Power Systems Computation Conference (PSCC)," Dublin, Ireland, 2018.

[37]   H. Ergun, J. Dave, D. van Hertem and F. Geth, "Optimal Power Flow for AC–DC Grids: Formulation, Convex Relaxation, Linear Approximation, and Implementation," *IEEE Trans. Power Syst,* no. 34, p. 2980–2990, 2019.

[38]   L. Garver, "Transmission Network Estimation Using Linear Programming," *IEEE TRANSACTIONS ON POWER APPARATUS AND SYSTEMS,* vol. 89, no. 7, pp. 1688-1697, 1970.

[39]   CIGRE Task Force C6.04.02, Benchmark Systems for Network Integration of Renewable and Distributed Energy Resources, CIGRE, 2014.

[40]   C. Das, O. Bass, G. Kothapalli, T. Mahmoud and D. Habibi, "Overview of energy storage systems in distribution networks: Placement, sizing, operation, and power quality," *Renewable and Sustainable Energy Reviews,* vol. 91, pp. 1205-1230, 2018.

[41] FlexPlan consortium, "D2.2 Flexibility elements identification and characterization," 2020.

[42] A. Gargiulo, P. Girardi and A. Temporelli, "LCA of electricity networks: a review," *The International Journal of Life Cycle Assessment,* vol. 22, no. 10, pp. 1502-1513, 2017.

[43] J. P. Bukenberger and M. D. Webster, "Approximate Latent Factor Algorithm for Scenario Selection and Weighting in Transmission Expansion Planning," *IEEE Transactions on Power Systems,* vol. 35, no. 2, pp. 1099 - 1108, March 2020.

[44] Garpur, "D4.2, Upgrading of the decision-making process for system development," *FP7-ENERGY,* 2013.

[45] K. Poncelet, H. Höschle, E. Delarue, A. Virag and W. D'haeseleer, "Selecting Representative Days for Capturing the Implications of Integrating Intermittent Renewables in Generation Expansion Planning Problems," *IEEE Transactions on Power Systems,* vol. 32, no. 3, pp. 1936 - 1948, May 2017.

[46] R. Alvarez, A. Moser and C. A. Rahmann, "Novel Methodology for Selecting Representative Operating Points for the TNEP," *IEEE Trans. on Power Systems,* vol. 32, no. 3, pp. 2234-2242, May 2017.

[47] P. Nahmmacher, E. Schmid, L. Hirth and B. Knopf, "Carpe diem: A novel approach to select representative days for long-term power system models with high shares of renewable energy sources," 2014. [Online]. Available: http://dx.doi.org/10.2139/ssrn.2537072.

[48] H. Yu, C. Y. Chung, K. P. Wong, H. W. Lee and J. H. Zhang, "Probabilistic Load Flow Evaluation With Hybrid Latin Hypercube Sampling and Cholesky Decomposition," *IEEE Transactions on Power Systems,* vol. 24, no. 2, pp. 661 - 667, May 2009.

[49] ENTSO-E, "2nd ENTSO-E Guideline For Cost Benefit Analysis of Grid Development," September 2018. [Online]. Available: https://eepublicdownloads.entsoe.eu/clean-documents/tyndp-documents/Cost%20Benefit%20Analysis/2018-10-11-tyndp-cba-20.pdf. [Accessed March 2021].