Development and Improvement of the power system module within the European Integrated Assessment Model MEDEAS.

Supervisor:
Quoilin Sylvain

Author:
Diffels Noé

Jury Members:
Cornélusse Bertrand
Wehenkel Louis
Solé Jordi

Master's thesis completed in order to obtain the degree of Master of Science in Energy Engineering, professional focus in Networks.

University of Liège - School of Engineering and Computer Science
Academic Year 2024-2025

Acknowledgements

This master's thesis marks the final step of my journey in engineering studies, and I would like to address a special word to the following individuals:

First and foremost, I would like to thank Sylvain Quoilin for introducing me to the fascinating world of integrated assessment modeling and MEDEAS. I am especially grateful for the opportunity to spend a week in Barcelona with the MEDEAS research team, allowing me to shape the present work.

Thank you to Jordi Solé for his insightful advice and engaging discussions. I am also deeply grateful to Enric Alcover Comas and Tareen Muhammad Umair for their valuable assistance and support throughout this project.

Lastly, a special thank you to my parents, my brother, my friends, and Marion. Your support was crucial, and without you, this journey would have been even more challenging to complete.

The work presented in this report navigates between multiple variables defined across different ecosystems associated to several models. For a sake of clarity, variables implemented within the MEDEAS model are denoted using `a special font`. Additionally, all the files implemented during this master's thesis are open-source and available in a GitHub repository.

The IA model *Chat GPT* was used during the writing of this report for proofreading, grammar checking, and clarity rephrasing.

In order to be consistent with previous works realized with MEDEAS model, the monetary value expressed in $ (USD) in this report, without explicit contrary mention, is the USD dollar value from 1995.

Abstract

This master's thesis is situated within the broader framework of global climate change and the European Green Deal, which sets the ambitious goal of achieving net-zero greenhouse gas (GHG) emissions in Europe by 2050. In particular, this research focuses on the Integrated Assessment Model (IAM) MEDEAS. This model aims to address the challenges of the energy transition within the European Union (EU) by providing comprehensive assessments of the potential impacts and mitigation strategies associated with various policy measures.

The master's thesis aims to propose a new version of MEDEAS which incorporates a machine learning-based surrogate model (SM) to improve the predictive potential of the IAM, particularly in simulating the European electrical power grid's curtailment and load shedding dynamics. This surrogate model was developed in previous works and is an efficient and flexible tool mirroring Dispa-SET unit commitment and economic dispatch model.

The other key advancements include the integration of additional data from PyPSA-EUR, enabling both the integration of the SM and new investment assessments of renewable energy sources (RES), grid reinforcement, and storage installations. Additionally, new feedback mechanisms inspired by PID control theory simulate instantaneous societal responses aimed at reducing energy curtailment and load shedding.

A comparative analysis against the previous MEDEAS version and a practical case study demonstrate the enhanced model's utility in exploring new energy scenarios and providing meaningful insights for policymakers.

Contents

List of abbreviations and acronyms

BAU	Business As Usual
CF	Capacity Factor
CoD	Coefficient of Determination (R^2)
CSP	Concentrated Solar Power (or Plant)
DS	Design Space
Dmnl	Dimensionless
EC	European Commission
ENTSO-E	European Network of Transmission System Operators for Electricity
EU	European Union
EV	Electric Vehicle(s)
FEFpc	Final Energy Footprint per capita
FFF	Fossil Fuels Fostering
GDP	Gross Domestic Product
HEV	Hybrid Vehicle(s)
IAM	Integrated Assessment Model
IEA	International Energy Agency
IOA	Input-Output Analysis
IPCC	Intergovernmental Panel on Climate Change
LEE	Low-Emission Electricity
LP	Linear Programming
MEDEAS	Modelling Energy Development under Environmental And Socioeconomic constraints
MILP	Mixed-Integer Linear Programming
ML	Machine Learning
MLP	Multi Layer Perceptron
NTC	Net Transfer Capacity
NZP	Net Zero Pathway
O&M	Operation and Maintenance
OT	Optimal Transition
PCHI	Piecewise Cubic Hermite Interpolation
PHS	Pumped Hydro Storage
PID	Proportional-Integral-Derivative (controller)
PV	Photovoltaic
RES	Renewable Energy Sources
RF	Random Forest
rNTC	Net Transfer Capacities ratio
SD	System Dynamics
SM	Surrogate Model
SSPs	Shared Socioeconomic Pathways
URR	Ultimate Recoverable Resource

1 Introduction

1.1 Context

This master's thesis is situated within the broader context of a global energy transition, from fossil fuels paradigm towards sustainable power systems, to address the pressing challenges of climate change and resource depletion. The crises faced by humanity are now widely recognized and emphasized across various domains, including public debate, media discourse, and scientific literature. In 2023, the Intergovernmental Panel on Climate Change (IPCC) published its Sixth Assessment Report on climate change [1], detailing the profound transformations occurring on Earth due to human activity. The IPCC report highlights numerous detrimental anthropogenic consequences of the climate change, including rising global temperatures, sea level increases, intensification of extreme weather events (such as droughts and floods), economic damages across various sectors, the degradation of the biosphere (affecting both living and non-living organisms), etc.

Nevertheless, the IPCC report also emphasizes that effective climate governance - enabled by political commitment, coordination, and robust laws, policies, and strategies - can facilitate both mitigation and adaptation to these deep changes. In alignment with this perspective, the European Union, committed to the goals of the Paris Agreement to limit global temperature increase to well below $+2\,^{\circ}\mathrm{C}$ [2], has established a comprehensive climate strategy. This strategy, known as the European Green Deal [3], sets an ambitious target to achieve net-zero greenhouse gas emissions within the EU by 2050. As a first step towards achieving climate neutrality by 2050, the European Commission (EC) has implemented a legislative and policy framework called the European Climate Law [4]. This set of laws codifies the objectives outlined in the European Green Deal but also introduces a new intermediate target for 2030: reducing net greenhouse gas emissions by at least 55% compared to 1990 levels.

This paradigm shift requires advanced tools to assess the impacts of various policies and compare their effectiveness in mitigating aforementioned environmental and societal damages. System modeling techniques have significantly evolved in the last decades, largely driven by advancements in computing power, enabling the simulation of large-scale system dynamics models. Study of such a scale must integrate the relationships among all societal sectors, such as economy, environment, energy but also the social aspects in order to accurately capture the systemic phenomena that shape our world. Consequently, since modeling the Earth system as a whole is inherently interdisciplinary, a new cross-disciplinary research stream was introduced to develop this type of models called Integrated Assessment Models (IAMs). These models enable researchers to assess the long-term impacts of policy decisions, and resource constraints on a global scale.

This interdisciplinary new modelling technique emerged in the early 1970s, with the creation of the Club of Rome. This nonprofit and non-governmental organization brought the pressing environmental global issues to public attention in 1972 with the publication of *The Limits to Growth* [5]. This pioneer scientific report used computer simulations to suggest that perpetual growth in production and consumption was unsustainable due to Earth's finite resources, pollution, and other constraints. Figure 1 illustrates the essence of the article by presenting the simulation results of a proposed integrated model, *World3*. The model clearly predicts collapse mechanisms occurring under a Business As Usual (BAU) scenario, which has attracted significant media attention.

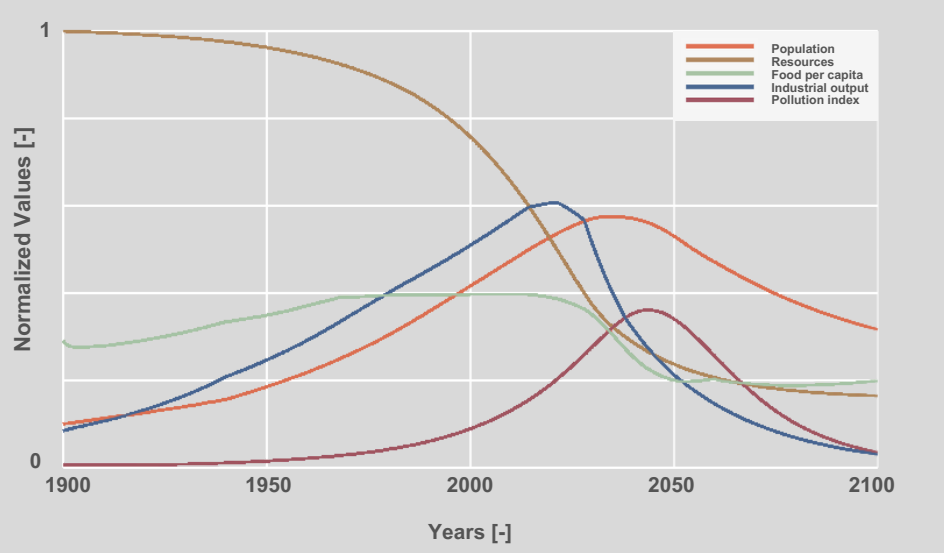


Figure 1: Standard run of *World3* model described in *Limit to Growth*, results extracted from [6].

Nowadays, integrated modeling remains highly relevant and is widely used by research institutes, intergovernmental agencies, and policymakers around the world. Additionally, the advance in artificial intelligence enables the implementation of hybrid models which combine system dynamics with Machine Learning (ML) algorithms, enhancing predictive accuracy and applicability to the system dynamics theory field.

1.2 Scope of this work

This master's thesis focuses on one IAM developed under *Horizon Europe*, a research and innovation funding program promoted by the EC [7]. The model, named MEDEAS (*Modelling Energy Development under Environmental And Socioeconomic constraints*) is defined as an economy-energy-environment model designed to advice policy makers and stakeholders with insights on the constraints and challenges of transitioning the EU towards a low-carbon and sustainable socio-economy.

The work presented in this report aims to propose an enhanced version of the MEDEAS model by integrating a machine learning-based surrogate model. This surrogate model, developed in previous master's theses ([8], [9]), serves as an efficient and adaptable tool for analyzing and optimizing power grid operations at the European level. It is designed to efficiently replicate the outputs of Dispa-SET [10], an economic dispatch and unit commitment optimization model. By reproducing Dispa-SET's outputs with significantly improved computational efficiency, the surrogate model introduces substantial added value to MEDEAS. Through this integration, the IAM gains the ability to evaluate the state of the European electrical network via two key outputs: curtailment and load shedding. This new modification addresses a limitation of the initial MEDEAS version, particularly in the implementation of curtailment, as discussed in Section 2.2.6. Furthermore, the surrogate model provides the capability to assess load shedding, absent from the initial implementation.

Subsequently, another external model is important in this master thesis: PyPSA-EUR. This power system optimization model is mainly used in this work by extracting external data from previous simulations of this model, further including new features in MEDEAS. Indeed, new assessments related to the required investments for integrating renewable energy sources into the European energy mix, associated grid reinforcement expenses, and new storage installations costs are incorporated into MEDEAS. Additionally, PyPSA-EUR plays a pivotal role in enabling the integration of the surrogate model, as it provides a crucial input: the Net Transfer Capacity ratio (rNTC). This variable measures the electrical network's capacity to transport energy and is introduced into MEDEAS, enabling the

IAM to evaluate this critical parameter during the simulations.

Ultimately, two society feedback mechanisms are proposed in this report. These processes assess the required investments and efforts to further reduce the undesirable effects computed by the surrogate model, namely the curtailment and the load shedding. Inspired by PID control theory, these feedback processes are user-configurable, allowing the adjustment of the feedback intensity and the activation year in pre-processing.

This master's thesis thus proposes a new version of MEDEAS, with several new features and processes implemented within the ecosystem. All variables integrated are listed in the Appendix A and the associated python files created during this work are available in a Github Repository. Additionally, in order to assess the relevance of this new configuration, this report presents both a comparative analysis with previous MEDEAS version, and a practical case study to assess the practical interest of the new version.

1.3 Structure of the report

This report begins with an exhaustive description of the MEDEAS model in Section 2 where the model's structure, its submodules, limitations, and examples of its outcomes are explored in detail. The subsequent Section 3 introduces all external models used in this study, including Dispa-SET (Section 3.2), the machine-learning surrogate model (Section 3.3), and the PyPSA-EUR model (Section 3.4). Following the presentation of these resources, the report describes the integration processes in Section 4 and the integration results in Section 5. Subsequently, the feedback mechanisms designed to simulate societal reactions based on the PID control theory are detailed in Section 6. The report continues with the a comparative analysis of the updated MEDEAS model against its initial version in Section 7, along with a case study that compares three energy scenarios using the new implementation, presented in Section 8. Finally, key results of this master's thesis are summarized, and further possible improvements are provided in the final conclusion, Section 9.

2 MEDEAS Model

2.1 Overview

This section describes MEDEAS, which serves as the central point of the work presented in this report. The MEDEAS project is a European Integrated Assessment Model that focuses on assessing the transition towards low-carbon energy systems while considering resource limitations, climate change impacts, and policy measures.

While focusing at the European level, the model needs to interact with the international (upper level) dynamics since continental interactions are strongly conditioned by the rest of the world. More specifically, MEDEAS is composed of three distinct geographical models, namely: MEDEAS-World, MEDEAS-EU, and MEDEAS-Country. Boundary conditions derived from the MEDEAS-World simulations are required by the European-level model. The work described in this report mainly focuses on the European framework.

The following Section 2.2 describes the model in more details, in particular main mechanisms that take place when the model is running (Section 2.2.2), how it is implemented and the different submodule interactions (Section 2.2.5). After its description, two practical case studies are proposed in order to illustrate the capabilities of MEDEAS (Section 2.3), but also limitations of the model (Section 2.4). Finally, the improvements and modifications proposed in this work are enumerated in Section 2.5.

2.2 Description of the model

2.2.1 System Dynamics framework

MEDEAS is an IAM developed using System Dynamics (SD) programming. This approach aims at modeling the nonlinear behavior of a complex system with a set of linear equations describing the interactions between system's components in the form of stocks, flows, feedback loops, etc. Through numerous assumptions, the complex interactions of a model can be approximated by linear relations in order to be resolved efficiently.

A basic system modelled using SD is derived from *Vensim* documentation [11]. The simple system modeled is the writing of a master's thesis. Figure 2 illustrates this system implemented in *Vensim*.

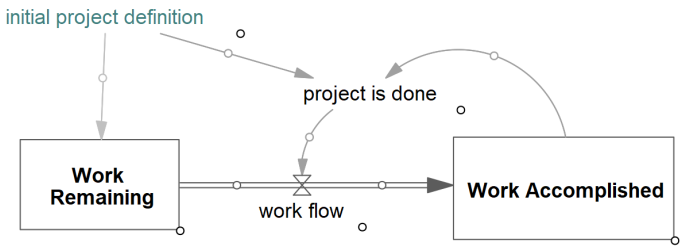


Figure 2: Basic example of System Dynamic modelling.

In the Figure 2 several components composing the writing process of a master's thesis are depicted. First, the initial project definition is set to 1000 [words] and the work flow to 200 [words/months], in order to represent the 5 months this thesis should take. Of course, the true work flow depends on numerous factors, but it is important to remind that this model only represents the aggregate representation of the thesis using mean values. The work accomplished is defined using a stock filling over time by the work flow. Once the initial project definition is reached, it means that the project is done, and the work flow stops.

The previous example presents a simple SD modeling process but provides insights on the further possibilities of large-scale complex systems modelling, as it is the case in MEDEAS.

2.2.2 Model's mechanisms in a nutshell

MEDEAS is a demand-driven European IAM which is based on Input-Output Analysis (IOA)[1] to define the economic system. This method positions MEDEAS as a general equilibrium model, accounting for the interactions between various economic sectors and examining the economy as an interconnected system where changes in one sector affect the others. More specifically, at each iteration, the global monetary demand from households and industry depending on the simulation scope (World, European, or Country level) is translated into final energy demand, by fuel and by sector. The energy needs are compared to the energy availability, considering factors such as the fossil fuels scarcity, or the RES integration, etc. If the demand exceeds the feasible supply, it adjusts to fit the available energy production through an energy-economy feedback, thus balancing the economic activity. The simulated energy mix influences the GHG emissions and impacts both social and climate change modules, that will be presented further in this report (Figure 4).

The energy-economic feedback implemented in MEDEAS is one of the major innovative differences this model presents over the other IAMs. This process thus constraints the economic sector by incorporating the impacts of environmental and energy availability on financial institutions, yielding a more realistic approach than most of the other alternatives. Indeed, many other IAMs assume an exogenous economical growth, without considering the critical role of resource scarcity as MEDEAS, which assumes that the health of the economy is inherently dependent on the energy sector.

The economic output of a sector is related to the energy demand through an important concept in MEDEAS: the energy intensity. It is defined in the MEDEAS documentation [13] such as: *"the ratio between the energy used in a process and its economic output"*. While considering 5 different final energy consumptions (electricity, heat, gases, liquids, and solids), along with 15 economic sectors (households, agriculture, construction, etc.), MEDEAS-EU models the final energy intensity of each economic actor considered by a statistical yearly mean indicator, resulting in 75 (15×5) energy intensities for each year.

2.2.3 Time horizon

MEDEAS model typically runs for a long term simulations from 1995 to 2050. Due to this large time scope and the large physical representation level (see all modules in Section 2.2.5), the time step is set at 0.03125 year, or 11.4 day. This time step is too large to capture short-term variability, and not suitable to model daily storage for example.

2.2.4 *PySD* library

The MEDEAS model was previously developed using *Vensim*, which is an interactive software environment that allows the development, exploration, analysis and optimization of simulation models [11]. For the sake of universality, new versions of the model were translated from Vensim into Python using a fork of the *PySD* module, described in the associated article [14]. All the work presented in this report is conducted in Python.

[1] IOA is an economic model developed by Wassily Leontief describing interdependencies of industries within the global economy. It describes the relationships between outputs of several industries being inputs of others, and the globalized network behind. The initial work regarding this approach is detailed in [12].

2.2.5 Structure of the model

The version used in this work is PyMEDEAS 2, the second Python version of this model. This Python version is composed of a main model class that represents the system dynamics which contains all model equations organized in submodules, as illustrated in Figure 4. On the one hand, these submodules contain several Python files where all endogenous variables are declared as functions. On the other hand, the model handles both historical data and exogenous parameters defined by the user in *Excel* files, integrating them as fixed variables of the model. Modifying these files, the user might change the model's behavior such as, for example, RES installations in order to compare different energy policies. A partial view of the implementation architecture is presented in Figure 3.

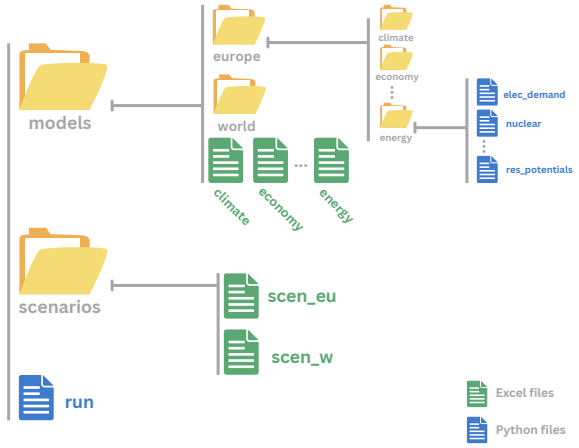


Figure 3: MEDEAS partial architecture implementation view.

Submodules containing relations and equations linked to a specific domain composing MEDEAS are presented in Figure 4.

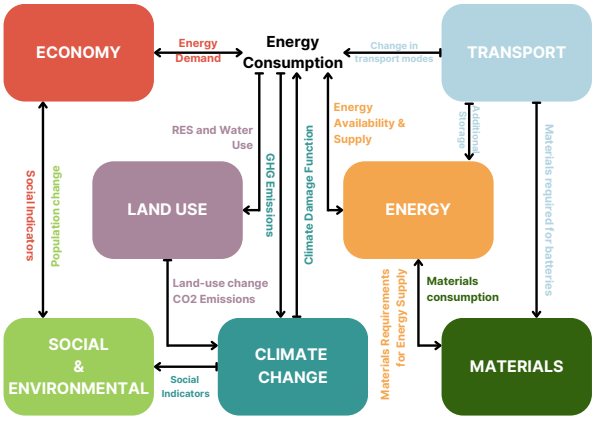


Figure 4: MEDEAS structure decomposed by modules and connections between them.

These submodules aim to represent a simplified version of worldwide interactions in various sectors, and are described below:

1. **Economy:** The economy is modelled assuming non-clearing markets. The economy is driven by the demand growth and considers energy supply constraints, through the aforementioned

economy-energy feedback. The economic structure is captured by the integration of IOA, containing 14 industrial sectors and households, considering a general equilibrium assumption.

2. **Energy:** This module includes both renewable and non-renewable energy resources potentials and availability taking into account biophysical and temporal constraints. A net energy approach has been followed, meaning that the final demand considers all the required energy, from the extraction, through the processing, to the delivery.

3. **Transport:** This module describes the transport sector composed of commercial and household transports. It is mainly based on two dynamics: enhancing the efficiency of liquid-based vehicles and transitioning them to electric-battery or natural-gas vehicles.

4. **Materials:** Several different materials are required by the economy to function. MEDEAS especially tracks the material requirements for the construction and the Operations and Maintenance (O&M) of the energy infrastructures. The materials extraction's demand is compared with the levels of available metrics of reserves and resources for fossil fuels, rare metals, etc.

5. **Land Use:** This module accounts for the land requirements of the RES, allowing the endogenization of built-up lands, and potential for both biomass and solar energy.

6. **Climate Change:** This module projects the climate change levels due to the GHG emissions generated by Human activity, modeled by a damage function representing the link between an increasing temperature and a GDP loss.

7. **Social & Environmental Impacts Indicators:** This module translates the "biophysical" results of the simulations into metrics related with well-being or social indicators and environmental impacts.

2.2.6 Current implementation of curtailment and load shedding

Since the main objective of this master's thesis is to enhance MEDEAS representation for curtailment and load shedding phenomena, this subsection introduces their definition in the model. A part of the energy module mentioned earlier is depicted in Figure 5, view through the *Vensim* software. The blue arrows represent interconnections between the concerned variables of the model. The *PySD* Python library mentioned in Section 2.2.4 is used to translate and solve the model in Python.

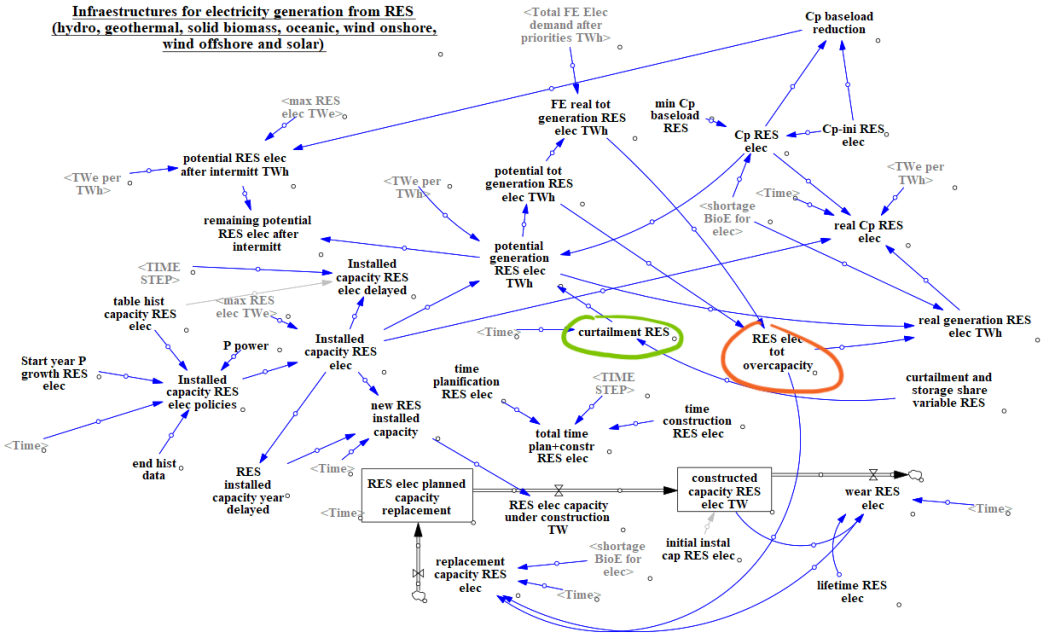


Figure 5: MEDEAS Energy module: RES capacities and generation submodule, view through *Vensim*.

MEDEAS is implemented in such a way that `curtailment_res` (highlighted in green, Figure 5) is an exogenous variable, meaning that it is set and fixed by the user before running the model. This implementation leads to a loss of accuracy as this percentage does not depend on any endogenous variables of the model and is just a fixed parameter. For the worldwide level model, curtailment is set at 0 %, while at the European level it is assumed to be 10 %, both representing a percentage of the whole electrical production which varies at each time step.

In this MEDEAS version, curtailment controls the electrical RES potential generation, thus being a cornerstone in the energy module. Equation 1 shows the relationship between curtailment and the potential generation. In other words, the blue arrow linking these two variables, actually means:

$$\texttt{potential\_tot\_generation\_res\_elec\_twh} = \texttt{installed\_capacity\_res\_elec} \cdot \texttt{Cp\_res\_elec} \cdot (1 - \texttt{curtailment\_res}) \cdot 8760 \quad (1)$$

Where `potential_tot_generation_res_elec_twh` is the potential electrical final energy generated by RES (expressed in TW h) based on the installed capacity `installed_capacity_res_elec` (expressed in TW), the RES dimensionless capacity factors `Cp_res_elec`, and the external parameter approximating the share in curtailment `curtailment_res`.

Regarding load shedding, it is currently neither computed nor to be fixed by the user in MEDEAS, which does not consider the unsatisfied electrical demand. The initial version considers the sectoral demand not covered in terms of monetary value by implementing the aforementioned economy-energy feedback mechanism. The difference between the desired demand and the real demand is represented by `demand_not_covered_by_sector_FD_EU`, which is expressed in M$/year.

2.3 Scenarios outcomes examples

This section presents two examples of MEDEAS policies analysis, detailing the altered parameters and providing a concise explanation of the corresponding simulations. The simulations are conducted at the world level and the baseline scenario used in both examples is detailed in the following Section 2.3.1.

2.3.1 World baseline scenario

The model used in this section is MEDEAS-World, describing the interactions between all previous sectors seen in Figure 4 (economy, energy, environment, etc.) at a global scale. The simulations ranges from 1995 to 2050, which is the typical time horizon used by MEDEAS to comprehensively described the energy transition.

Both historical data and projections are used to assess the population growth during the simulation, represented in Figure 6. The economic demand is captured by the integration of IOA representing 34 sectors and the households. This implementation produces a desired Gross Domestic Product (GDP) per capita growth, common for all further case studies and illustrated in $/per in Figure 7.

Both these variables are fixed while sensitivity analyses. The desired GDP per capita represents the maximum demand to fulfill if the economy-energy feedback allows it. However, it will be demonstrate in the next subsections that different policies might result in lower GDP.

Without any contrary mention, the scenario presented in this section will be placed in an energy transition context where the electrical generation from RES power plants increases, as shown in Figure 8.

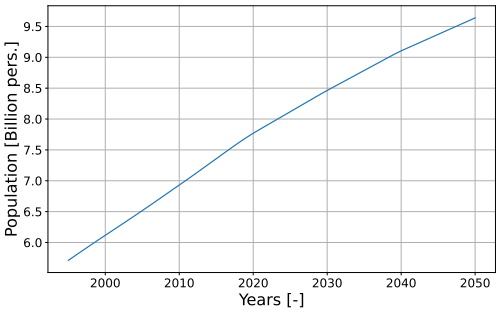


Figure 6: Population growth assessment.

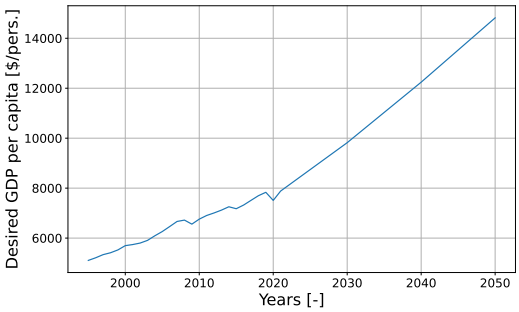


Figure 7: Desired GDP per capita.

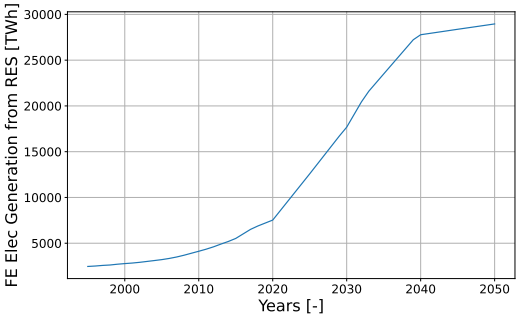


Figure 8: Final Energy generated by RES installations, baseline scenario.

The following subsections describe two sensitivity analyses, a first where different nuclear policies are compared (Section 2.3.2) and the second where two energy policies are presented (Section 2.3.3).

2.3.2 Nuclear policies

This subsection details an example of nuclear policy analysis at a global scale. The external parameter `selection_of_nuclear_scenario` can be configured to modify the nuclear policy used in the model. This input is translated by the model through a modification of the annual increase of new planned nuclear capacity. The nuclear policies - represented in Figure 9 - are enumerated below:

1. Nuclear growth policy: "Pro-nuclear" policy where the installed capacity of nuclear power plants grows.
2. Nuclear phase-out policy: The installed capacity of nuclear power plants is reduced and rapidly decommissioned to reach near 0 in 2050.
3. Nuclear decommissioning policy: Current European (moderate) nuclear policy where no more nuclear power plants are constructed, and the current capacity depreciates.

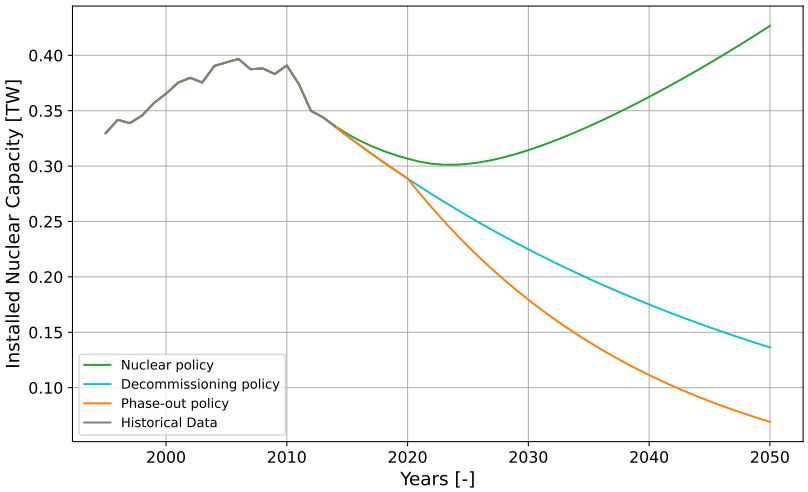


Figure 9: Installed Nuclear Capacity in TW, comparison between nuclear policies.

It is important to note that the model prioritizes the RES production over non-renewable energy sources, leading to a constant electricity generation from RES between all nuclear scenarios. However, the decision of whether the nuclear infrastructures are promoted or not, will still affect the CO_2 emissions and the uranium extraction, as it can be seen respectively in Figure 10 and Figure 11.

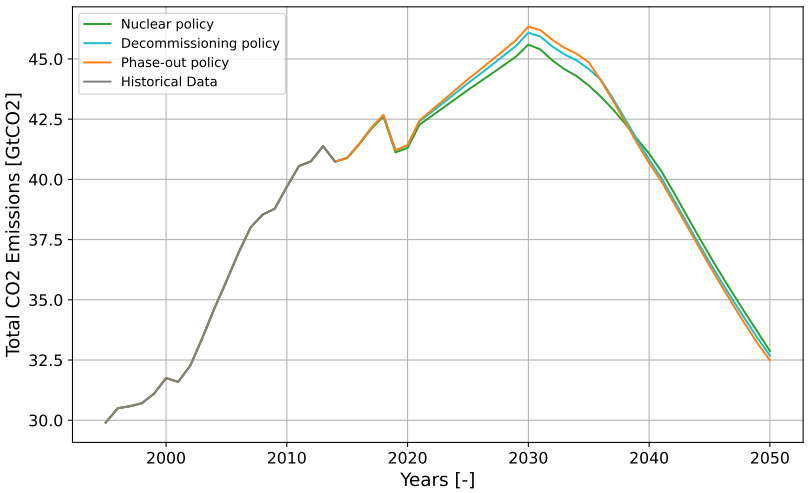


Figure 10: Total Carbon Dioxide Emissions in $GtCO_2$, comparison between nuclear policies.

Figure 10 shows that a nuclear-driven policy would help to handle faster anthropogenic CO_2 emissions reduction, since the CO_2 peak reached is smaller than both other policies. As previously mentioned, since the simulation is placed within a decarbonization context where the RES installations growth, the CO_2 emissions by 2050 decrease, falling around 32.5 $Gt\,CO_2\,year^{-1}$ for each scenario. It results to a strong decrease of more than 10 $Gt\,CO_2\,year^{-1}$ in a decade. It can also be noticed that a nuclear policy brings a slower decrease, which should be easier to handle than a strong and radical change considering the inertia of such large-scale system.

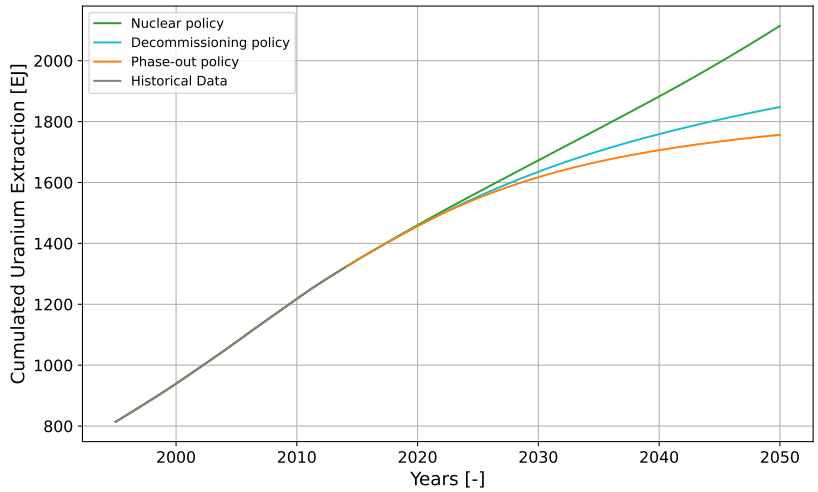


Figure 11: Cumulated Uranium Extraction in EJ, comparison between three nuclear policies.

Figure 11 highlights a notable difference in uranium extraction by 2050 under different policies. Under the pro-nuclear policy, approximately 2100 EJ of uranium is extracted, with a continued growing trend, whereas in the phase-out scenario, around 1750 EJ is extracted, and the curve begins to stabilize as almost no further uranium is required. These policy scenarios are considered realistic within the model's framework, as MEDEAS assumes an Ultimate Recoverable Resource (URR) for uranium between 3700 EJ [15] and 4000 EJ [16], both detailed in Table 12 of the documentation [13].

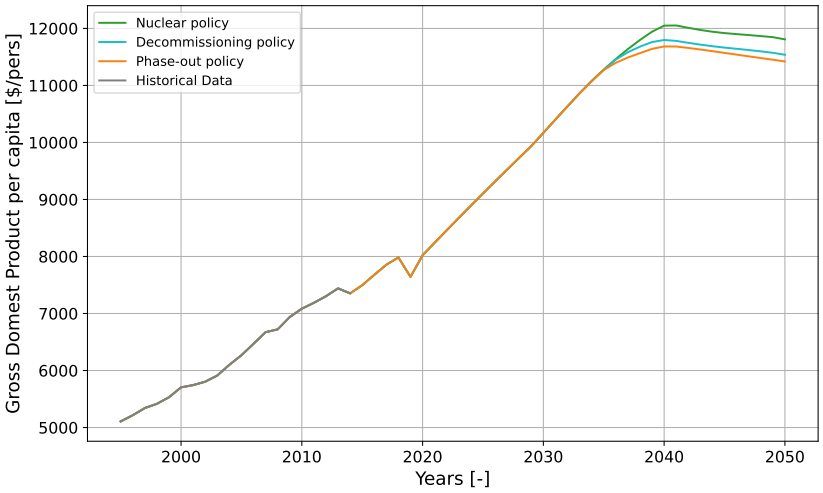


Figure 12: World Gross Domestic Product per capita in $/pers., comparison between three nuclear policies.

Figure 12 illustrates the potential impact of nuclear policies on the GDP per capita. It can be seen that the desired GDP per capita illustrated in Figure 7 is not followed by these curves suggesting that a part of the economic demand is not satisfied. Economically speaking, it means that the growth rate of the GDP decreases. This phenomenon results from the RES integration of the world baseline scenario (more precisely the associated intermittency), which leads to a gap between desired economic growth and economic demand satisfied. This is a clear demonstration of the economy-energy feedback

mentioned earlier. The same results is heightened while the second policies analysis. When comparing the nuclear policies, the graph shows a slightly higher trajectory under pro-nuclear scenario, suggesting that this policy supports more the economic growth in long terms through the installation and O&M nuclear investments.

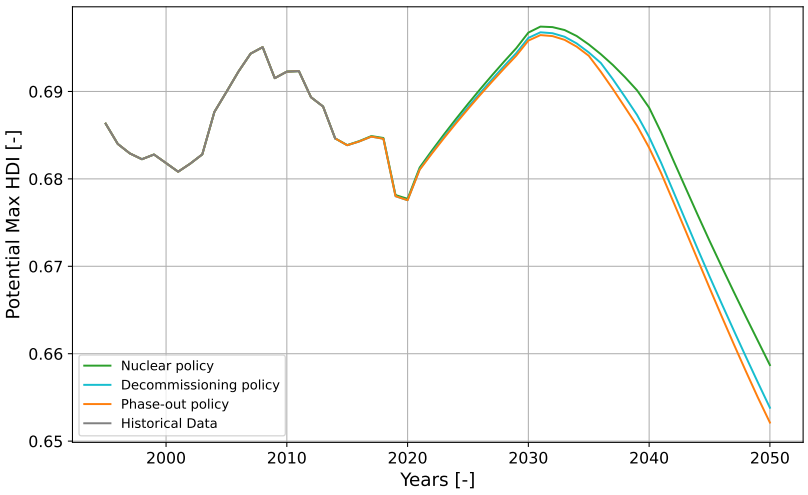


Figure 13: Potential Maximum HDI indicator.

Figure 13 describes the evolution of the Human Development Index (HDI), which is a social index aggregating several indicators such as life expectancy, education, and per capita incomes. It is used to categorize countries according to the associated human development and is defined as follows:

$$HDI = \sqrt[3]{LEI \cdot EI \cdot II} \tag{2}$$

Where LEI is the life expectancy index, EI is the education index, and II the income index. Among all these factors describing the social wellness level, only the income index is available in MEDEAS. The model rather suggests a regression between HDI and Final Energy Footprint per capita ($FEFpc$, in GJ/person), shown in Figure 66 (Section 2.7.2) of the MEDEAS user's manual [13]. This definition can explain why such a drop is observed with all policies in Figure 13. Indeed, since HDI is depending on $FEFpc$ and the latter being reduced (Figure 10), a steep drop is observed.

Both Figures 12 and 13 clearly indicate better trends in terms of social impacts with nuclear policies. However, it is important to keep in mind that such trends are only insights into the true physical phenomena taking place and cannot fully represent reality. Moreover it cannot prevent any black swan events, which are global-scale unpredictable events with low probability and high impacts on society. By extent, a failure in nuclear power plant is highly unexpected but could eventually have opposite effects on human wellness. It is important to keep in mind the complexity of society mechanisms and the objective of such indicators.

2.3.3 Fossil Fuel policies

This subsection details an energy policy analysis between renewable energy sources and Fossil Fuels (FF) at a global scale. Regarding the FF scenario, several variables are changed, namely: the installed RES powers are set as constants from 2020, fossil fuels are no longer restricted to remain underground, and the reduction policies of oil in heat and electricity contribution are removed. Regarding the nuclear parameter, it is set as phase-out policy.

Both analyzed energy policies can be described as,

1. FF policy: Fossil Fuel policy, the renewable energy sources are no longer promoted, neither nuclear power plants. The society extract fossil fuels without considering the resources scarcity, until the peak of fossil fuel extraction is reached.
2. RES policy: Baseline scenario where a growth of RES production in the context of the energy transition is observed. While the carbon-emissive power plants are decommissioned, RES capacity increases.

Figure 14 shows the radical difference between RES scenario and Fossil Fuels scenario, where the RES are set constant from 2020 until 2050. Obviously, the latter scenario is not considered realistic but is used here to show MEDEAS outputs example in a simple context.

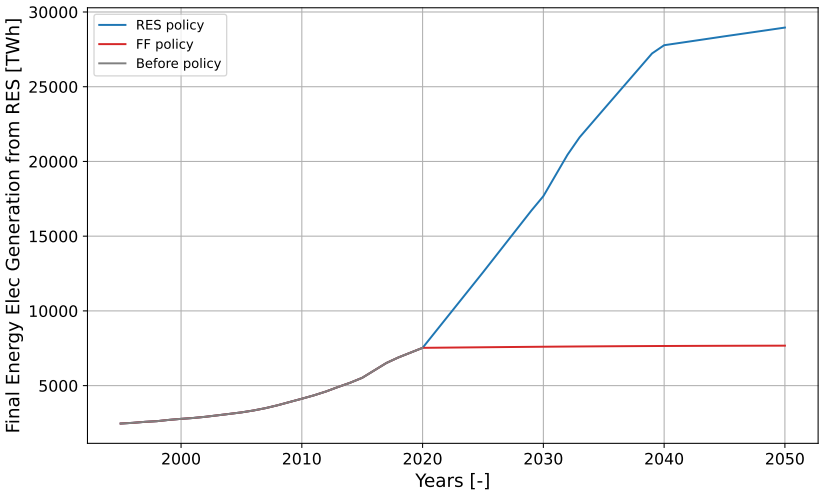


Figure 14: Total Final Energy Generation of Electricity from RES in TWh, comparison between energy policies.

Obviously, such an aggressive policy of fossil fuels highlights depletion effects on non-renewable resources. Figure 15 shows the reached peaks for all fossil fuels around 2027, indicating a too high extraction price leading to a forced reduction in numerous sectors as it will be shown in the following paragraphs. The drops illustrated in Figure 15 indicate that the fossil fuels paradigm cannot indefinitely be extended and should be abandoned in favour of innovative technologies.

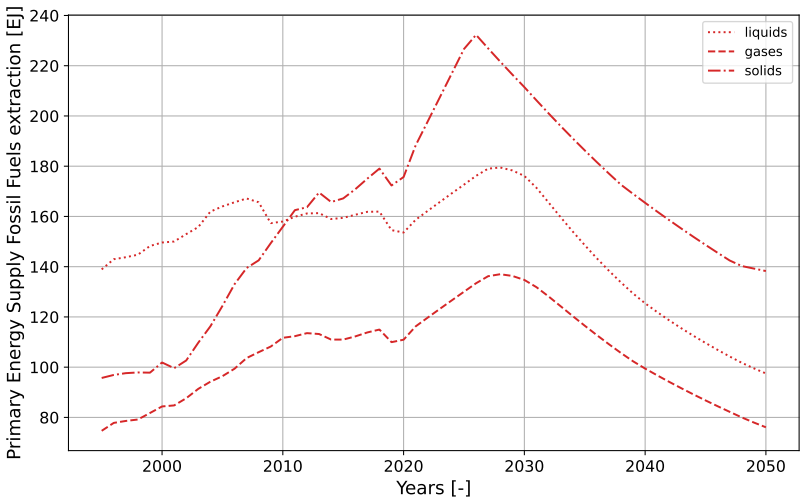


Figure 15: Primary Energy Supply from Fossil Fuels extraction, in FF scenario.

Figures 16-19 present significant and abrupt changes in each variable considered, being in line with the IPCC sixth Assessment Report on climate change [1] where predictions with high confidence are made regarding systemic unavoidable and irreversible turmoils if fossil fuels are still being fostered in the near future. As it was previously introduced at the beginning of this report, Humanity is currently facing an historical turnover, since a radical transition must be done regarding the fossil fuels paradigm in order to mitigate the risks of severe impacts on society [1]. Fossil fuels extraction and consumption not only exacerbate climate change through the emission of GHG, but also degrade the biosphere (livings and non-livings organisms) [17], and raise societal [18] and economy [19] instabilities.

As it can be read in the literature [18], [20], all tipping points of the complex system Earth-Society seem to be interconnected thus enhancing risks of cascading effects at a global scale. This is the reason why IAMs such as MEDEAS are crucial for evaluating systemic interactions and providing a comprehensive global perspective on complex bio-societal mechanisms, without limiting the decision-making process to a single sector.

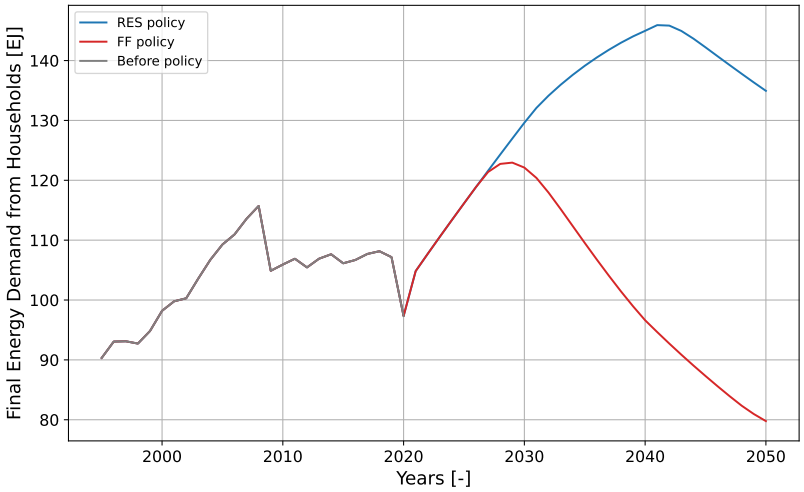


Figure 16: Total Final Energy Demand of HouseHolds in EJ, comparison between energy policies.

As shown in Figure 16, RES policy could, in the long term, lead to a higher households consumption where the final energy demand is not constrained by fossil fuel peaks and can even growth until 2040. The abrupt drop observed in red (Figure 16) may indicate severe impacts on the citizens' welfare, such as significant restrictions on energy consumption or higher energy prices, which could exacerbate social inequalities. This is not the case with the RES scenario where more time is allowed to address the consumption peak.

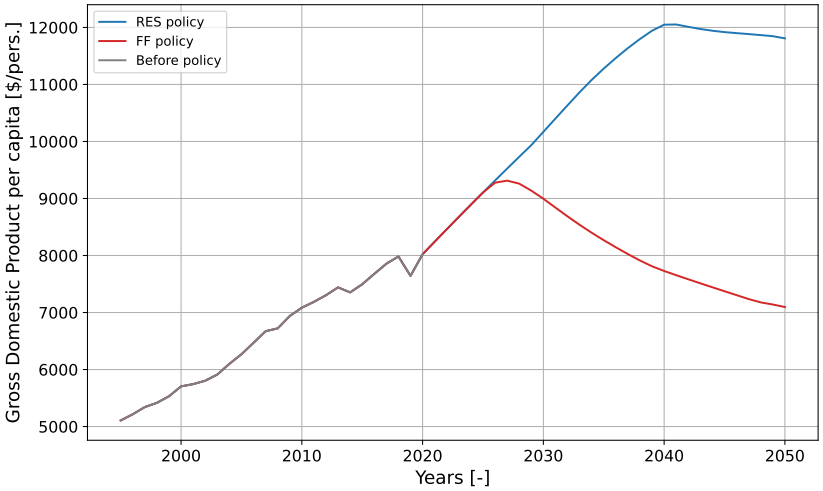


Figure 17: Gross Domestic Product per capita in $/pers., comparison between energy policies.

The difference between desired GDP per capita and the model's prediction, already observed in previous policies analysis (Figure 12), once again appears in Figure 17. However, with the fossil fuel resources depletion, the economy-energy feedback is heightened in this case. It results in a large difference by 2050 with the observed GDP per capita being the half of the desired output from the economy module illustrated in Figure 7.

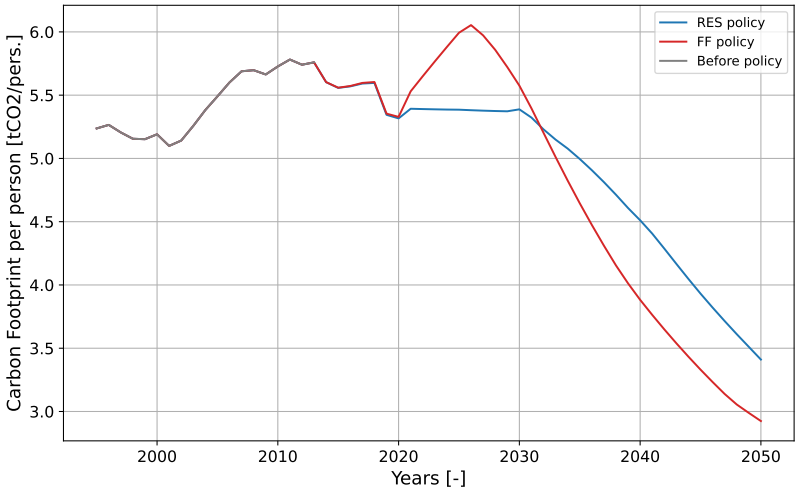


Figure 18: Carbon Footprint per person in $\mathrm{t\,CO_2\,per^{-1}}$, comparison between energy policies.

Fossil fuels policy shows a peak in carbon footprint per person in 2025 in Figure 18, where a large

drop is observed reflecting continued emissions due to the reliance on non-renewable energy sources until the tipping point is activated, where fossil fuels are no longer available. In contrast, the energy-transition policy demonstrates a steady state preceding a significant reduction in emissions, suggesting a smoother transition which will be easier to handle but also demonstrating the inertia related to a global scale paradigm transition.

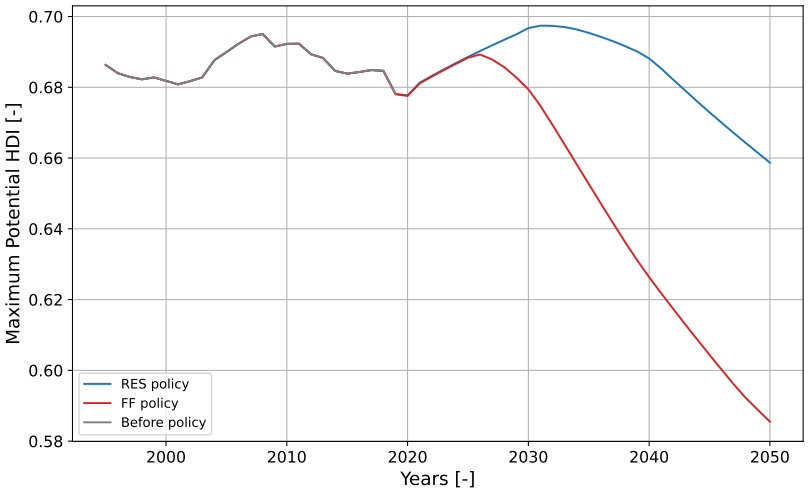


Figure 19: Potential Maximum Human Development Index, comparison between energy policies.

Despite the reduction in both scenarios observed in Figure 19, the RES policy proposes a larger maximum potential HDI emphasizing how reducing carbon emissions and investing in clean energy align with socio-economic development goals. Once again, the carbon reduction scenario shows a smoother transition in contrast of abrupt reduction imposed by the scarcity in fossil fuels resources.

The previous two case studies show a small part of all possibilities that MEDEAS present in terms of insights on future policies impacts. These scenarios were conducted at a world scale, but this report focuses on the European level as it will be explained in the next sections.

2.4 Limitations of the model

It is important to keep in mind that scientific modelling is not intended to perfectly predict the future but to present plausible scenarios providing an objective guidance in the associated field. As an example, MEDEAS tries to qualitatively advise policy makers to take the best decisions for the energy transition to a low carbon economy.

The second limitation is the tradeoff that needs to be made when exploring such a large scope of modeling. Obviously, not a single integrated assessment model can simulate a detailed and perfect version of the world and the infinite number of interactions that take place anywhere, anytime. Physical connections are modeled using assumptions in order to get an approximation of trends that could, possibly, occur. In addition, as seen in Section 2.2.3 the time step used is too large to perfectly translate the reality and all short term interactions. For example, regarding storage, MEDEAS only considers two technologies at a large scale: Pumped Hydro Storage (PHS) and Electric Vehicles (EV) batteries, which is a major assumption. Finally, the main part of the historical data used are yearly mean values and only offer insights regarding former global trends.

2.5 Improvements proposed in this work

This master thesis aims at proposing a new implementation for the MEDEAS power system module, by incorporating the following modifications:

1. Integration of a machine learning surrogate model. See Section 4.2.
2. Modification of curtailment definition. See Section 4.2.3.
3. Integration of load shedding. See Section 4.2.3.
4. Proposition and integration of a new artificial metric describing the capacity of the European grid to transfer electricity. See Section 4.3.1.
5. Integration of an assessment variable for the RES deployment investments. See Section 4.3.2.
6. Integration of an assessment variable for the grid reinforcement investments. See Section 4.3.2.
7. Integration of an assessment variable for the storage-related investments. See Section 4.3.2.
8. Integration of society reaction mechanisms which are adjustable by the user and aim at limiting curtailment and load shedding. See Section 6.

Section 7 is dedicated to the modifications made during this work, and a comparative analysis between the initial and the proposed version is made. These modifications required many new variables in MEDEAS ecosystem, which are listed in Appendix A with a brief description. Subsequently, Section 8 proposes a practical case study where three scenarios are compared using the new proposed implementation of MEDEAS. Finally, the Section 9 presents the take-home messages and the further improvements of this work.

3 External Models

3.1 Overview

In addition of the base framework from MEDEAS model depicted in Section 2, this master's thesis uses several other models detailed in this section.

The starting point of the workflow illustrated in Figure 20, is the unit commitment and economic dispatch optimization model Dispa-SET [10] (Section 3.2) which focuses on the balancing and flexibility problems in European grids. Dispa-SET was used in previous master's theses [8] and [9], to develop a machine learning-based surrogate model which aims to enhance the efficiency of simulating the European power system, in a very effective way. The proposed surrogate model was designed to be integrated into other large time scope models. Indeed, the gain in computation time allows the precision of a detailed model like Dispa-SET to be incorporated into models spanning larger time horizons, such as MEDEAS, where simulations are conducted over several decades.

Integrating Dispa-SET into MEDEAS can offer significant advantages, enhancing both the functionality and precision of MEDEAS. Dispa-SET is useful when analyzing the integration of renewable energy sources in the electrical mix, particularly assessing issues like load shedding and curtailment. Since MEDEAS aims at study the transition towards a low-carbon economy, integrating Dispa-SET can enhance its ability to evaluate the impact of intermittent renewable energy sources (such as wind and solar), on grid stability and operational efficiency in the short term.

Additionally, a third model is used in this work in order to extract necessary external data which are not available through the other models: PyPSA [21] (Section 3.4). This model is an open-source energy model for the European power system at the transmission level. It differs from Dispa-SET since it optimizes the energy generation and transmission considering the investments required for different technologies, using a multi-sectoral approach. The additional data is obtained from an in-house PyPSA-EUR model of the European sector-coupled energy system in the context of a study described in the following article [22].

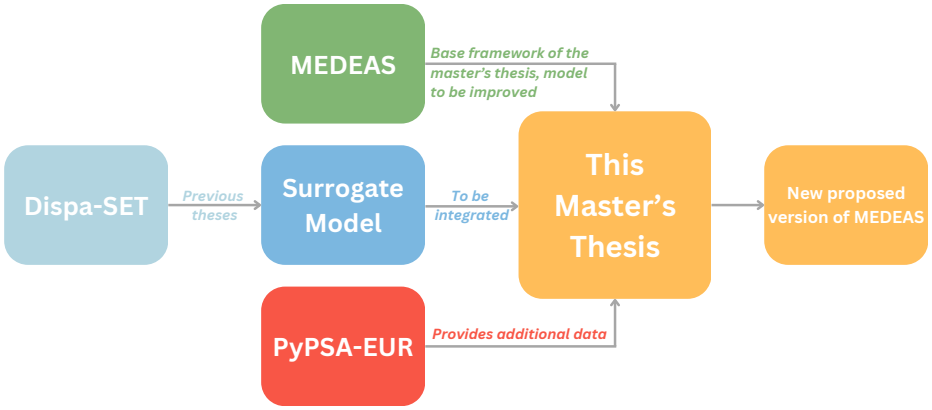


Figure 20: Overall workflow of this master's thesis and each model's contribution.

The following subsections provide a more detailed explanation of these three models.

3.2 Dispa-SET Model

3.2.1 Model in a nutshell

Dispa-SET is an open-source unit commitment and optimal dispatch model focused on the balancing and flexibility problems in European grids [10]. It is mainly written in Python, and is optimized through GAMS [23]. Dispa-SET is developed within the Joint Research Centre of the EU Commission, with the help of Liège University and the KU Leuven. The optimization problem translated by this model can be divided into two main parts, namely the Unit commitment and the Economic Dispatch.

1. Unit Commitment: This part focuses on whether each unit is in on or off state. (*i.e.*, determining whether to power up the plant or shut it down is a more optimal choice.) These power states are subject to startup, shutdown, and ramping constraints.
2. Economic Dispatch: This part focuses on the energy equilibrium required between generation and demand, in order to minimize the system's overall costs and determine the precise generation of each power unit at each time step.

This model could either be described using a Mixed-Integer Linear Programming (MILP) or relaxed into a Linear Programming (LP) optimization problem, which depends on the trade-off between computational cost and accuracy level.

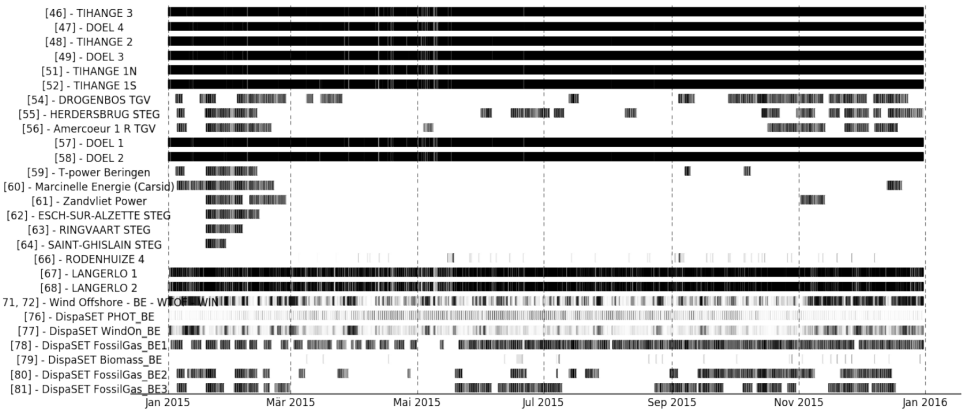


Figure 21: Visual output example of Dispa-SET model: Commitment and power level status of each unit in Belgium and Luxembourg, taken from [10].

3.2.2 Key inputs and outputs

On the one hand main inputs required by the model consists of the power plants properties, such as generation and ramping limits, minimum up and down times, weather data for non-dispatchable units (*e.g.* solar panels, wind turbines, etc.), startup and ramping costs. But also data on the network topology (capacity constraints for congestion computation), targets for carbon policies, yearly data for potential units outages, etc.

On the other hand, the model returns the commitment status and the power produced of each unit, the load shedding and the curtailment at each time step.

3.2.3 Spatial and temporal scopes

Dispa-SET model typically runs for a yearly simulation, but could become too expensive in computational terms when solving it with a high details level data. Therefore, the problem is divided into low level optimization problems that are solved recursively through the year, using rolling horizon principle.

Regarding the spatial scope, as previously mentioned, the model depicts large-scale (country or continent) power systems.

3.2.4 Applications

This project aims at providing with high-level of detail the short-term operations of large-scale power systems. The applications of the model in the scientific literature ranges from specific country studies for RES integration [24] (Belgium) to power systems adequacy and flexibility assessments in developing countries [25].

In the context of this master's thesis, Dispa-SET is the model which aims to be integrated in MEDEAS, through the surrogate model described in the following Section 3.3 and focuses on the European power system.

3.3 Surrogate Model

3.3.1 Model in a nutshell

Previous works [8] [9] aimed to create an efficient surrogate model used to approximate Dispa-SET simulations, which is too computationally expensive to be integrated in another environment. In this manner, the new machine learning model can mirror the key outcomes of Dispa-SET, thus creating a more efficient tool for analyzing and optimizing European power system mechanisms. This lighter model can thus be simulated efficiently and be integrated in a higher level model like MEDEAS, which is the purpose of the work presented in this report.

The implementation of the surrogate model was conducted in two main parts, dataset generation and machine learning algorithm implementation (related to [8] and [9]).

First, a large dataset must be generated in order to have an initial space for training the machine learning algorithm. The Latin-hypercube strategy developed in [9] was chosen in order to sample 2655 Dispa-SET simulations in a multidimensional input space. Table 1 shows the data obtained at each simulation while creating the machine learning dataset.

Variable	Unit	Variable	Unit
Cost	€/MWh	MaxLoadSheddingShare	%
Congestion	h	CurtailmentToRESGeneration	%
PeakLoad	MW	CF gas	[-]
MaxCurtailment	MW	CF nuc	[-]
MaxLoadShedding	MW	CF wat	[-]
Demand	TWh	CF win	[-]
NetImports	TWh	CF sun	[-]
Curtailment	%	Capacity ratio	[-]
Load Shedding	%	Share flex	[-]
LostLoad	TWh	Share sto	[-]
MaxRESGeneration	TWh	Share Wind	[-]
TotalGeneration	TWh	Share PV	[-]
ShareRESGeneration	%	rNTC	[-]

Table 1: Data computed during each simulation while creating the dataset, taken from [8]. Red variables are the outputs of the surrogate model while blue ones are the inputs.

Once the large dataset is created, a machine learning algorithm can be applied in order to predict both targets (in red) from the six features (in blue) presented in Table 1. The selection of variables depends on the inputs and outputs from Dispa-SET, as explained in previous work [9]. This problem is a typical ML supervised regression problem, which can be mathematically defined as:

From a learning sample derived from the Dispa-SET dataset $\{(\mathbf{x}_i, \mathbf{y}_i | i = 1, ..., N)\}$ *with* $x_i \in \mathcal{F}$ *(Features),* $\mathbf{y}_i \in \mathcal{T}$ *(Targets), and* $N = 1858$ *(70% of the 2655 data points, rounded down) the problem is to find a function* $f : \mathcal{F} \to \mathcal{T}$ *minimizing the expectation of some loss function* $l : \mathcal{T} \times \mathcal{T} \to \mathbb{R}$ *over the joint distribution of features-targets combinations:* $E_{\mathbf{x},\mathbf{y}}\{l(f(\mathbf{x}), \mathbf{y})\}$.

3.3.2 Machine Learning algorithms

Two machine learning algorithms are proposed in previous master's thesis [8], namely: Random Forest (RF) and Multi-Layer Perceptron (MLP). Both methods are briefly presented in the following paragraphs.

Random Forest

The Random Forest algorithm applies the concept of bagging (or Bootstrap Aggregating) into the decision trees paradigm: several random samples of the learning set are drawn and a different regression tree is trained on each subset. After the training, predictions on unseen data are computed by averaging predictions from each tree. A representation for the RF algorithm is presented in Figure 22.

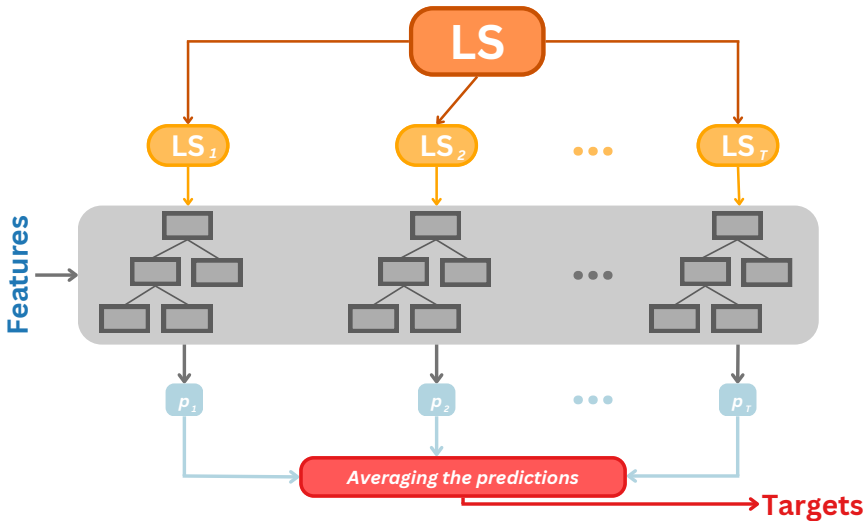


Figure 22: Representation of a Bagging Random Forest with T samples, each producing one prediction that will affect the mean outputs.

The proposed implementation is made of 500 trees, without any maximum depth for each tree and without a maximum number of features to consider for the best split.

Multi-Layer Perceptron

A Multi-Layer Perceptron is a type of Artificial Neural Network (ANN) consisting of many interconnected nodes (or neurons) organized by layers. Three different categories of layers compose the MLP, namely: the input layer, one or more hidden layers, and an output layer. The information flows through this architecture from inputs, through the hidden layers, to the outputs. MLPs use activation functions (*e.g.* ReLU or sigmoid) to introduce non-linearity and are trained using backpropagation by adjusting connection strengths (or weights) between neurons in order to minimize the error. A representation for the MLP algorithm is presented in Figure 23.

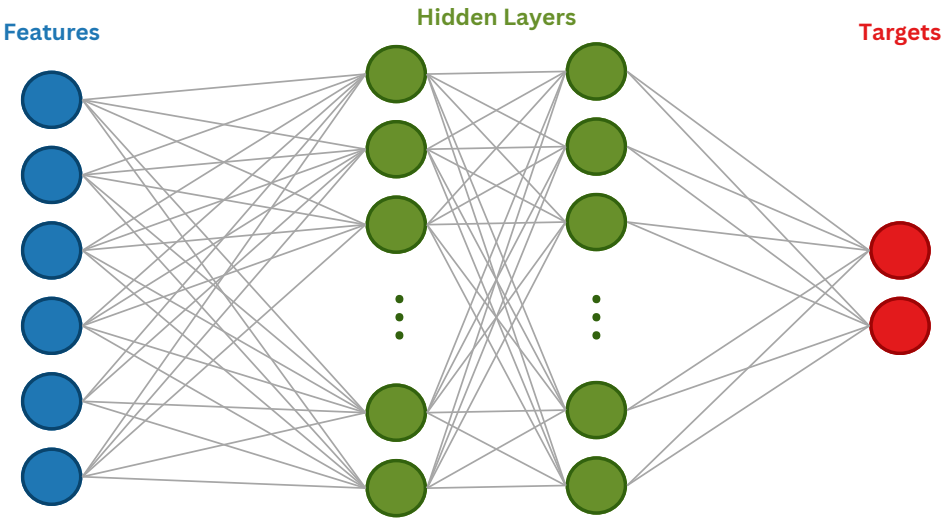


Figure 23: Representation of a Multi-Layer Perceptron with 6 inputs, 2 hidden layers, and 2 outputs.

The proposed implementation is made of two hidden layers composed by respectively 100 and 50 neurons, and using ReLU as activation function.

Algorithms comparison

Several loss functions were used in previous work [8] to assess both models on unseen data and thus evaluating the prediction ability on new data. Performance metrics are detailed in Table 2a for curtailment and in Table 2b for load shedding.

	RF	**MLP**
RMSE [-]	0.034	0.029
MAE [-]	0.023	0.021
Training Time [s]	57	2.7
Prediction Time [s]	0.02264	0.000174

(a) Curtailment

	RF	**MLP**
RMSE [-]	0.034	0.016
MAE [-]	0.002763	0.000125
Training Time [s]	6.24	6.27
Prediction Time [s]	0.014	0.063

(b) Load shedding

Table 2: Performance metrics assessing RF and MLP accurateness on unseen data for both targets.

Where **RMSE** and **MAE** denote respectively the Root Mean Squared Error and the Mean Absolute Error. Despite the fact that Tables 2a and 2b indicate better performance for the MLP rather than RF, as a first step, both surrogate models are implemented alongside and a practical comparison between them is conducted later in this report (Section 5.4).

3.3.3 Units classification

Both models described in Sections 3.2 (Dispa-SET) and 3.3 (Surrogate Model) classify power plants into two types, namely: flexible and slow units. According to IRENA's classification [26], flexible units are power plants that can maintain a low minimum operating level, ramp up quickly, and have fast start-up and shutdown times. Previous power plant classification are given in Table 3. This distinction separates slow from flexible units, thus defining $Share_{flex}$ input (Section 3.3.4). In addition, three other separated kinds are considered, namely: PV, Wind, and Storage Units.

Type	Condition
Flexible	PartLoadMin $< 0.5\%$, and TimeUpMin < 5 [h], and RampUpRate > 0.01 [MW/h]
Slow	PartLoadMin $\geq 0.5\%$, or TimeUpMin ≥ 5 [h], or RampUpRate ≤ 0.01 [MW/h]

Table 3: Units Classification: Flexible vs Slow, taken from [9].

Where the conditions PartLoadMin, TimeUpMin, and RampUpRate represent respectively the percentage of minimum nominal capacity, minimum up time, and the ramping up rate (*ie* how quickly the unit can increase its power generation).

3.3.4 Key inputs and outputs

The surrogate model was implemented using a design space composed of six dimensionless features aimed at predicting two dimensionless targets, as listed in Table 4.

Features	Targets
Capacity Ratio	Curtailment
Share of Flexibility	Load Shedding
Share of PV	
Share of Wind	
Share of Storage	
rNTC	

Table 4: List of dimensionless features and targets composing the surrogate model, taken from [8].

The six features are the dimensionless representations of the six inputs from the Dispa-SET simulations composing the dataset used to train the ML algorithms. Both curtailment and load shedding are extracted from Dispa-SET simulations and are thus the two targets of the surrogate model. The other outputs from the simulations are the commitment status and the power produced of each unit which are not relevant in MEDEAS context.

Obviously, each feature influences differently each target, leading to varying correlations between them. This is demonstrated in Section 3.3.7, where Figure 24a and Figure 24b illustrate, respectively, the impacts of the share of storage and share of wind on curtailment, as well as the effects of the capacity ratio and the Net Transfer Capacity ratio on load shedding.

The previous master's thesis [8] defines boundaries for each feature of the surrogate model based on a 2019 reference Dispa-SET simulation (current estimations) and external data from both PRIMES [27], [28] and PyPSA [21] (predictions). These boundaries are crucial for the design space definition since the six-dimensional space created by the features depends on it, shaping the surrogate model definition space. The next paragraphs present each pair of (dimensionless) boundaries alongside the associated feature definition in Dispa-SET nomenclature.

Capacity Ratio

The capacity ratio represents the ratio of the installed power of all units in the power system (flexible and slow units) to the peak load.

It is defined as

$$\text{CapacityRatio} = \frac{\text{PowerCap}_{\text{flex}} + \text{PowerCap}_{\text{slow}}}{\text{PeakLoad}} \tag{3}$$

Where $\text{PowerCap}_{\text{flex}}$ is the installed power of flexible units expressed in TW, $\text{PowerCap}_{\text{slow}}$ is the installed power of slow units expressed in TW, and PeakLoad represents the maximum possible demand expressed in TW.

The bounds for the capacity ratio in the surrogate model are **[0.4, 1.3]**. Based on the 2019 reference value **1.16** and the prediction **0.74**. The capacity ratio tends to decrease with the rise of VRES and storage capacity in the electrical mix.

Share Flex

The share of flexibility represents the proportion of power capacity generated by flexible power plants to the power capacity generated by all units (flexible and slow). It is defined as

$$\text{ShareFlex} = \frac{\text{PowerCap}_{\text{flex}}}{\text{PowerCap}_{\text{flex}} + \text{PowerCap}_{\text{slow}}} \tag{4}$$

Where all variables were previously defined. The bounds for the share of flexibility in the surrogate model are **[0.25, 0.9]**. Based on the 2019 reference value **0.42** and the prediction **0.46**. Despite the prediction being slightly larger from the reference value, the previous master's thesis [8] argues that technologies tend to be more and more flexible and thus assumes a large upper bound. This tendency is also verified with the following article [29], as discussed in Section 4.2.2.

Share Storage

The share of storage represents the ratio of the overall nominal storage power to the peak load.

It is defined as

$$\text{ShareSto} = \frac{\text{PowerCap}_{\text{sto}}}{\text{PeakLoad}} \tag{5}$$

Where $\text{PowerCap}_{\text{sto}}$ is the installed power of storage (*ie* the maximum rate at which the storage can charge or discharge) expressed in TW and PeakLoad was previously defined.

The bounds for the share of storage in the surrogate model are **[0, 3]**. Based on the 2019 reference value **0.001** and the prediction **2.6**. The necessary flexibility of the future electrical network naturally increases while VRES share rises. This flexibility is provided, among others, by storage capacity.

Share Wind

The share of wind represents the ratio of electricity generated by wind turbines (onshore and offshore) to the peak load.

It is defined as

$$\text{ShareWind} = \frac{\text{PowerCap}_{\text{wind}} \cdot CF_{\text{wind}}}{\text{PeakLoad} \cdot CF_{\text{load}}} \tag{6}$$

Where $\text{PowerCap}_{\text{wind}}$ is the installed power of wind units expressed in TW, CF_{wind} is the considered dimensionless capacity factor of the wind turbine, PeakLoad was previously defined, and CF_{load} represents the dimensionless capacity factor of the load from the reference simulation, defined in Romain's thesis [8] (Eq. 4.3) as follows:

$$CF_{load} = \frac{\text{Total Annual Energy Demand (MWh)}}{\text{Peak Demand (MW)} \cdot 8760 \text{ (h)}} = 0.736 \text{ [-]} \tag{7}$$

The bounds for the share of wind in the surrogate model are **[0, 0.55]**. Based on the 2019 reference value **0.2** and the prediction **0.45**, in line with increase in VRES.

Share PV

The share of solar PV represents the ratio of electricity generated by solar panels to the peak load.

It is defined as

$$\text{SharePV} = \frac{\text{PowerCap}_{\text{PV}} \cdot CF_{\text{PV}}}{\text{Peak Load} \cdot CF_{\text{load}}} \tag{8}$$

Where $\text{PowerCap}_{\text{PV}}$ is the installed power of solar units expressed in TW, CF_{PV} is the considered dimensionless capacity factor of solar units, and the other variables were previously defined.

The bounds for the share of solar PV in the surrogate model are **[0, 0.35]**. Based on the 2019 reference value **0.047** and the prediction **0.2**, in line with the increase in VRES.

rNTC

The Net Transfer Capacity Ratio is an artificial metric that depicts the global interconnection state of an electrical network. It increases along the network ability to share electricity between distant areas.

In order to define $rNTC$, first the Net Transfer Capacity is computed for each zone z to any other zone x over each N_h hour of the data:

$$NTC_{z \to x} = \frac{1}{N_h} \sum_h NTC_{z \to x,h} \tag{9}$$

The Net Transfer Capacity between two zones is the maximum amount of electricity (in MW) that can be transferred between two interconnected power (zonal) systems without compromising the security of both systems.

Secondly, Equation 10 defines the zonal NTC, providing a representative value of NTC for one zone z by summing all the NTC from this zone z to the others.

$$NTC_z = \sum_x NTC_{z \to x} \tag{10}$$

Finally the Net Transfer Capacity ratio (rNTC) is defined by the sum of each zonal NTC weighted by the sum of the peak load for each zone x.

$$rNTC = \sum_z \frac{NTC_z}{\sum_x \text{PeakLoad}_x} \tag{11}$$

The bounds for rNTC in the surrogate model are **[0, 0.75]**. Based on the 2019 reference value **0.28** and the prediction **0.55**. The transfer capacity of the electrical grid is expected to increase in order to balance growing share of VRES.

Curtailment

The curtailment target represents, for a given time step t, the ratio of total energy curtailed from VRES (*ie* "spoiled energy" that could have been produced by VRES but was deliberately not) to the maximum VRES generation from all units considering the availability factor.

It is defined as

$$\text{Curtailment} = \frac{\text{TotEnergyCurtailed}_t}{8760 \sum_{units} AF_u \cdot \text{PowerCap}_u} \tag{12}$$

Where TotEnergyCurtailed represent the amount of energy curtailed at time step t (expressed in [TWh]), AF represents the average capacity factor (or Availability Factor) for the unit u for one year, and PowerCap_u stands for the installed capacities of the unit u in [TW] with $units$ the set of all considered power plants u.

There are no specified bounds for curtailment. However, by definition, it cannot be negative.

Load Shedding

The load shedding target represents, for a given time step t, the ratio of the load that cannot be fulfilled to the total demand. The electrical demand might not be fully satisfied during the simulation since MEDEAS considers a non-equilibrium economy (*ie* non-clearing market). This unbalance typically occurs when the production is lower than demand.

It is defined as

$$\text{LoadShedding} = \frac{\sum \text{ShedLoad}_t}{\sum \text{TotDemand}_t} \tag{13}$$

There are no specified bounds for load shedding. However, by definition, it cannot be negative.

3.3.5 Scaling

A scaling approach was used during the surrogate model development in the previous master's thesis [8]. The approach consists to scale both inputs and outputs from the surrogate model in order to normalize data into a consistent scale thus improving the performance of the machine learning algorithm. Indeed, without the scaling procedures the features with high ranges might dominate the learning process, leading biased model performance. Moreover, the scaling improved the convergence speed during training in [8].

The proposed approach consists to use the *MinMaxScaler* from *Scikit-learn* Python package [30], which is defined as follows:

$$X_{\text{scaled}} = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \tag{14}$$

Where X is the original feature value, $X_{\max}$ and $X_{\min}$ are the maximum and minimum values of the feature in the training set, and X_{scaled} is the scaled feature between 0 and 1.

The inverse transformation is subsequently made to translate the scaled output from the surrogate model to the true physical value of either curtailment or load shedding. It is important to note that each feature is scaled individually to ensure that the relationships between features remain unaltered.

3.3.6 Spatial and temporal scopes

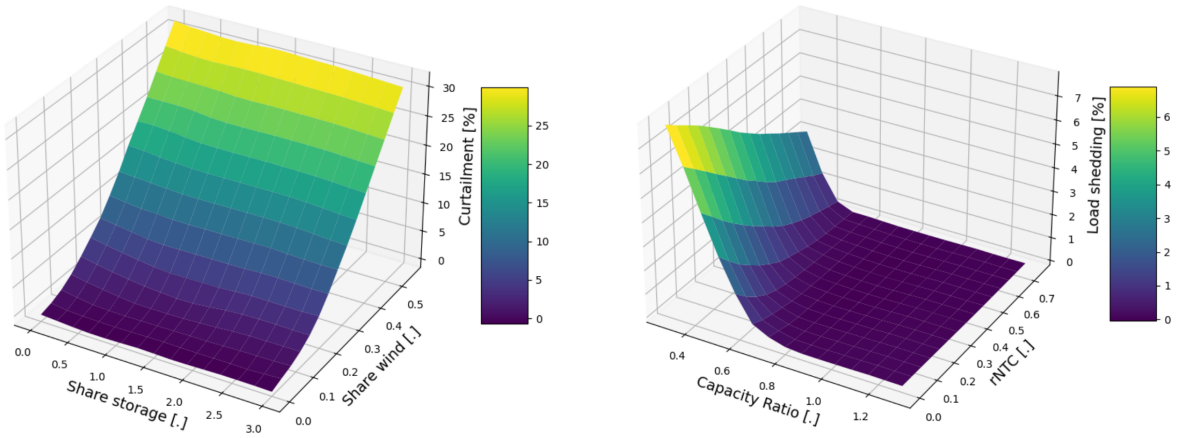
Since the surrogate model was trained to mimic Dispa-SET, it represents European-scale power systems but can be integrated in simulation of large time horizon (decades). Regarding the temporal scope, the computation of the two targets is remarkably fast: around 0.000 15 s (*resp.* 0.012 64 s) using MLP (*resp.*

RF) on a typical laptop, compared to up to 1.5 h for Dispa-SET. The surrogate model simulations were realized on a typical laptop with 8GB of RAM and a i5 processor, while the Dispa-SET computation time is extracted from previous master's thesis [8].

3.3.7 Applications

The proposed surrogate model was designed to be integrated into other models with larger time scopes. Indeed, the gain in computation time allows the precision of a detailed model like Dispa-SET to be incorporated into models spanning larger time horizons, such as MEDEAS, where simulations are conducted over several decades.

In order to provide a more practical perspective, the Figures 24a and 24b help to visualize the behavior of both SM targets while varying two of the six features, the others remaining constant.



(a) Curtailment while share of storage and share of wind vary.

(b) Load shedding while capacity ratio and $rNTC$ vary.

Figure 24: Results of the (MLP) surrogate model for curtailment and load shedding. Surface plots with two varying features, the others remaining constant. Figures taken from [8].

One the one hand, Figure 24a demonstrates the variation in curtailment when the share of storage and the share of wind vary. This figure shows that curtailment is highly sensitive to changes in the share of wind, while the share of storage appears to have little to no influence. Which, at a first sight, seems surprising since an increase in storage capacity should bring more flexibility to the grid, ultimately decreasing the curtailment. Previous master's thesis [8] argues that this phenomenon might be caused by the curtailment cost defined in Dispa-SET. Indeed, since the curtailment cost is set to zero, the model seems to not penalized it which results in a misconception of the true impact that storage capacity has on grid's flexibility. This aspect should be further investigated but is considered out of the scope for this master's thesis, which considers the provided surrogate model as it.

On the other hand, Figure 24b demonstrates the variation in load shedding when the capacity ratio and the Net Transfer Capacity of the grid vary. It can be seen that the worst case, *ie* the highest unsatisfied demand, occurs for low capacity ratio combined with a low transfer capacity, reaching to 7% of the demand that cannot be fulfilled. This is expected since a decrease in the capacity ratio occurs when either the peak demand increases or the maximum power production - excluding wind and solar - decreases. In either case, fulfilling the demand becomes more challenging. Furthermore, a grid with low transfer capacity struggles to distribute energy effectively, making it difficult to satisfy the demand across all nodes of the network.

3.3.8 Limitations

This model comes with several limitations, which are listed below.

1. Non-flexibility: The model is implemented in such a way that the features and the targets cannot be changed, making it highly specific and inflexible in situations where the user does not have access to some features or wants to compute another target.
2. Boundaries constraints: The design space is limited by boundaries for each feature (Section 3.3.4) which thus constraints the function definition limits and restricts the potential case studies.
3. Trade-off with accurateness: The proposed time reduction comes with a minor trade-off in precision. However, this slight loss in accuracy remains acceptable given the substantial gains in speed for more extensive scenario analysis.

3.4 PyPSA-EUR Model

3.4.1 Model in a nutshell

PyPSA-EUR [21] is an open-source energy system model for the European power system at the transmission network level. It includes detailed data on thousands of high-voltage lines and substations, but also on conventional power plants, renewable energy potentials, and time series for demand and renewable generation. The model focuses on optimizing energy generation and transmission while considering spatial and temporal variations in renewable resources in the grid.

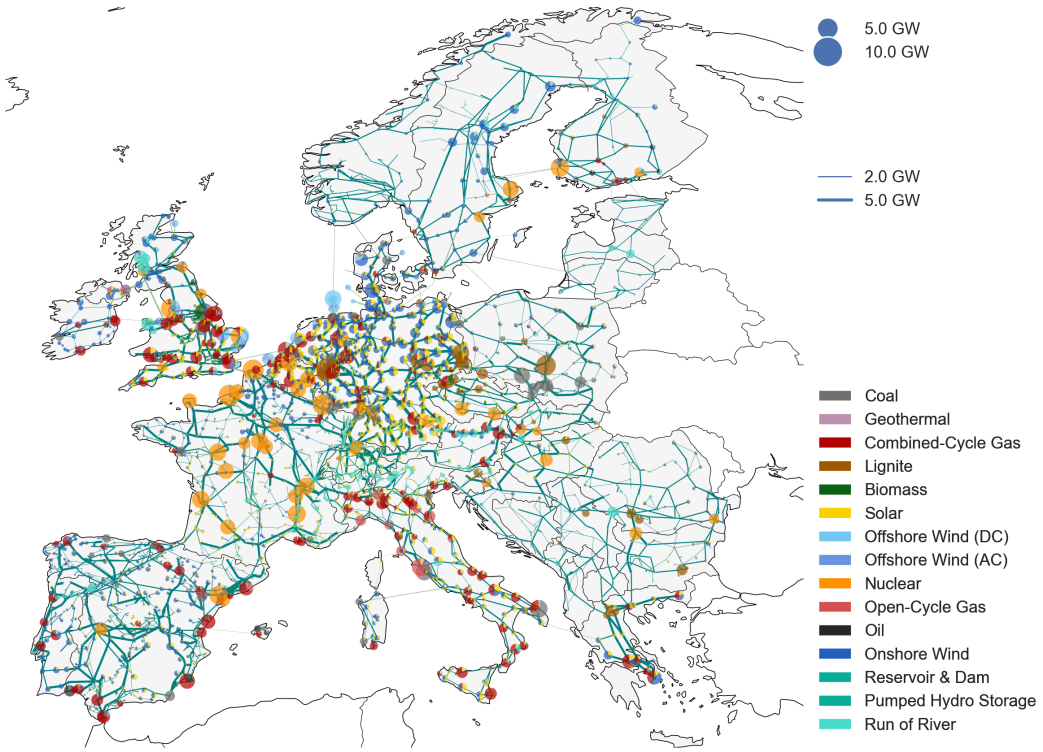


Figure 25: Visual output example of PyPSA-EUR model, taken from Github repository of the project.

This model was previously used in a study [22] conducted by the thermodynamics Laboratory at ULiège to explore the role of energy sufficiency in driving Europe's transition. Using this model, the

Laboratory analyzed how reducing energy demand will impact costs, emissions, and infrastructures in Europe through an optimization problem definition. Capacity and investments data derived from this study are incorporated in the MEDEAS ecosystem with a view of improving the energy transition representation of the model. Refer to Section 4.3 for further details.

3.4.2 Key inputs and outputs

The model considers several open-source datasets, including the grid topology (OpenStreetMap [31]), renewable generation potentials (*atlite* [32]), and others. For sector coupling, additional inputs including demand and supply profiles for heating, transport, industry, agriculture, residential, etc. can also be integrated. Once the simulation is done, the model provides optimal configurations for both power generation and transmission networks while considering costs, CO_2 emissions, grid congestion, etc.

3.4.3 Spatial and temporal scopes

This model covers the entire ENTSO-E area[2] but geographical scope is a user-specified option, alongside planning horizon resolution. It can be resolved on 5-years steps or for a decade, and the time resolution can range from one hour for the whole year or three, six hours. In the aforementioned sufficiency study [22], simulations were conducted for each decade from 2020 to 2050, resulting in four one-year simulations with hourly resolutions. It was solved for 28 countries with NIC5 cluster with 16 CPUs and 300 GB memory, solving a year in approximately 10 hours.

3.4.4 Applications

This model was primarily developed for studying both long and short-term planning and power system operations. PyPSA-EUR could be used for electricity-only studies, where only the power grid is considered (*e.g.*, power grid expansion), or for sector-coupled studies where additional sectors can also be included.

3.4.5 Limitations

Since this model is supposed to represent a large spatial system, many approximations are made due to the lack of data. Obviously, some assumptions regarding exact line impedances, or wind and solar distributions have been made. Further information is available in the related paper [21] (Section 4).

3.5 Summary

In conclusion, four different useful models are used during this work, properties and applications for each of them are summarized in Table 5.

[2]European Network of Transmission System Operators for Electricity is the association for the cooperation of the European transmission system operators. Interactive map available at: https://www.entsoe.eu/data/map/.

	MEDEAS	**Dispa-SET**	**Surrogate Model**	**PyPSA-EUR**
Type	Integrated Assessment Model	Unit Commitment & Optimal Dispatch (LP or MILP)	Either MLP or RF ML Algorithm	Energy System model including investments optimization (LP or MILP)
Spatial Scope	World, Europe, or Country level.	European Grid	European Grid	European Grid
Temporal Resolution	Simulation between 1995 and 2050, 11.4 days resolution.	Simulation for one year, and Hourly Resolution.	No predefined temporal resolution, instantaneous.	Simulation for one year, and Hourly Resolution.
Purpose	Assess Energy Transition with an integrated point of view.	Assess unit commitment and economic dispatch within European grid.	Mimic Dispa-SET but in a lighter version.	High detail level European grid description with optimal investments.
Role in this work	To be improved. (Section 7)	Replicated by the Surrogate Model. (Section 3.3)	To be integrated in MEDEAS. (Section 4.2)	Provides additional data. (Section 4.3)

Table 5: Summary of models used in this work.

Both the Dispa-SET and the Surrogate Model applications are straightforward, but one could ask why a third energy model (PyPSA-EUR) operates in this work while it seems inconvenient to have too many external models in the same ecosystem. However, the latter model brings information unavailable in the others since this model, as previously mentioned in Section 3.4, uses investment data in its optimization process, whereas Dispa-SET only focuses on optimizing the production costs. Additional data brought by PyPSA-EUR is detailed in Section 4.3.

4 External models integration into MEDEAS

4.1 Overview

The core objective of the work presented in this report is to integrate a machine learning surrogate model (Section 3.3) into the MEDEAS energy module (Section 2.2.5). This integration aims to enhance the model's representation of the electrical network by incorporating key targets identified in Section 3.3.4 (curtailment and load shedding) into the IAM. To do so, six features required by the surrogate model must be implemented within MEDEAS (Section 4.2.2).

As previously mentioned, both machine learning algorithms (RF and MLP) are integrated into the system. However, for the sake of clarity, only the outputs of the MLP are presented in Sections 5.2 and 5.3. A comparative analysis of both algorithm is provided in Section 5.4.

Moreover, additional data derived from PyPSA-EUR is integrated in the model, in order to:

1. Integrate rNTC feature within MEDEAS, since it is currently not taken into account in the model (Section 4.3.1);
2. Integrate installed capacities and investments data to better assess future grid costs (Section 4.3.2).

4.2 Surrogate Model integration

4.2.1 Unit classification

Since MEDEAS does not assess each unit parameter but works with mean values for each type of fuel, the classification described in Section 3.3.3 is not relevant in this context. This does not affect surrogate model features such as $Share_{\text{flex}}$, as the assumption of a linear increase is made, or Cap_{ratio}, since all units capacities are summed up regardless of their classification. Refer to associated paragraphs in following Section 4.2.2 for further information.

4.2.2 Links between MEDEAS variables and SM features

The work presented in Section 3.3, is used in this master's thesis in another environment from its initial purpose. Indeed, the machine learning surrogate model was developed to compute both targets (curtailment and load shedding) using six features described in Section 3.3.4 which are defined using Dispa-SET variables. However, since the purpose of this work is to integrate the surrogate model into MEDEAS, a list of relationships between these two environments can be outlined. New definitions of the surrogate model features using variables from MEDEAS are given below, in the following paragraphs.

All the following feature definitions are available in the file `features.py` (available in the GitHub repository), where all the surrogate model inputs are defined.

Peak Load

Since MEDEAS works with mean values for consumption and production (due to its large time scale, Section 2.2.3), there is no pre-existing variable for peak demand in electricity. However, it can be easily defined as follows:

$$\texttt{peak\_load} = \frac{\texttt{total\_fe\_elec\_demand\_twh}}{365 \cdot 24 \cdot CF_{\text{load}}} \tag{15}$$

Where `total_fe_elec_demand_twh` is a MEDEAS variable representing the annual electrical demand in TWh, and $CF_{\text{load}} = 0.736$ [-] describes the capacity factor of the load for a reference scenario, which is defined in Equation 4.3 in Romain Cloux's thesis [8]:

$$\text{CF}_{\text{load}} = \frac{\text{TotalAnnualEnergyDemand [TWh]}}{\text{PeakDemand [TW]} \cdot 365 \cdot 24 \text{ [h]}} = 0.736 \tag{16}$$

It means that the peak demand is assumed to be proportional to the mean electrical demand, more specifically with an increase of 35% of the average considered load.

Capacity Ratio

Previously defined in Equation 3 using Dispa-SET variables, the capacity ratio can be defined within MEDEAS such as:

$$\texttt{cap\_ratio} = \frac{\texttt{cap\_installed\_tot}}{\texttt{peak\_load}} \tag{17}$$

Where the new implemented variable `cap_installed_tot` contains the sum of all power capacities installed considered in the model expressed in TW. Regarding RES installations, a straightforward variable in MEDEAS is implemented and called `installed_capacity_res_elec`. However, this is not the case for non-renewable fuels. Indeed, for fossil fuel units, only yearly electrical generation is directly available in MEDEAS. The following assumptions (in Table 6) regarding Capacity Factors (CF) are thus made, to obtain an estimate of the installed capacity.

Fuel	Constant CF estimation [-]
Oil	0.19
Gas	0.39
Coal	0.51

Table 6: Capacity factor estimations used to assess the installed capacities of fossil fuel sources in MEDEAS, derived from [33].

On the contrary, the nuclear CF is already supported by MEDEAS, through `cp_nuclear` variable which is based on a 2017 study from Capellán-Pérez, de Castro [34] and is equal to 0.77 [-].

Once the capacity factor is determined, the installed nominal capacity can be directly estimated. Since the associated reference [33] provides historical values, these assumptions do not strictly reflect real-world conditions and future trends - since capacity factors vary by unit and fluctuate through the year - this methodology provides a valuable insight into the actual installed power capacity of non-renewable sources. A further better approach might be investigated, consisting to train another ML model based on Dispa-SET results which aims to predict the associated capacity factors related to fossil fuels power plants. However, it is not considered during this work.

Share of flexible generation capacity

As it was previously mentioned in Section 4.2.1, the share of flexible units cannot be determined endogenously within MEDEAS since the model neither incorporates the flexibility indicators described in Section 3.3.3 nor accounts for individual units. Instead, MEDEAS only considers the average yearly electrical generation by fuel type. Previously defined in Equation 4 using Dispa-SET variables, the share of flexibility is thus defined within MEDEAS such as:

$$\texttt{share\_flex} = m \cdot (\texttt{time} - 1995) + p \tag{18}$$

With Equation 18 a linear approximation for the share of flexible generation capacity, `time` the time step (expressed in year), m the slope, and p the intercept with y-axis. The last two being defined in Equations 19 and 20 using both following points shown in Table 7.

Explanation	Year	Share of flexibility [-]
Lower bound at the first time step, Section 3.3.4	1995	0.251
Reference value for October 2019, Table 3.1 in [8]	2019	0.41

Table 7: Share of flexible units numerical values used for the linear approximation of this feature.

Using Table 7, m and p are respectively defined such as:

$$m = \frac{0.41 - 0.251}{2019 - 1995} = 0.006625 \tag{19}$$

$$p = 0.41 - m \cdot (2019 - 1995) = 0.251 \tag{20}$$

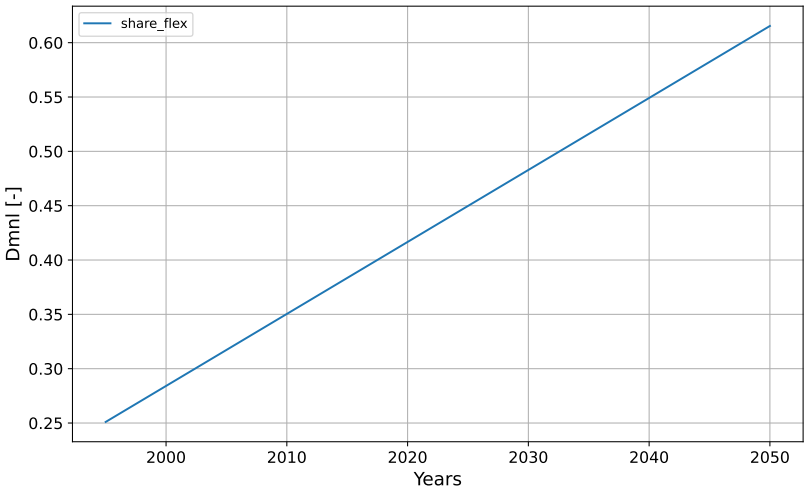


Figure 26: Share of flexible units feature, linear assumption.

The share of flexibility, as depicted in Figure 26, grows linearly starting from its lower bound of 25%. Despite its simplicity, this linear approximation offers a valuable insight into future trends for this feature. The share of flexible units is expected to approach 1, meaning that all new power plants in construction will be flexible and slow units may phase out in the future. Indeed, with advancements in technology, it is expected that minimum load, minimum time up and ramping rate will improve significantly, leading to a reduction in the share of slower units in the energy mix [29].

Share of Storage

Previously defined in Equation 5 using Dispa-SET variables, the share of storage can be defined within MEDEAS such as:

$$\texttt{share\_sto} = \frac{\texttt{total\_capacity\_elec\_storage\_tw}}{\texttt{peak\_load}} \tag{21}$$

Where `total_capacity_elec_storage_tw` describes total electrical capacity storage available in TW.

Share of Wind

Previously defined in Equation 6 using Dispa-SET variables, the share of wind can be defined within MEDEAS such as:

$$\texttt{share\_wind} = \frac{\texttt{installed\_capacity\_res\_elec(wind)} \cdot \texttt{cp\_res\_elec(wind)}}{\texttt{peak\_load} \cdot CF_{\text{load}}} \tag{22}$$

Where `installed_capacity_res_elec(wind)` describes the installed capacity of wind units in TW, `cp_res_elec(wind)` is the dimensionless capacity factor related to wind in MEDEAS, and the factor $CF_{\text{load}} = 0.736$ [-] once again comes from Equation 16. Since wind energy is divided in onshore and offshore in MEDEAS, the share of wind is defined as the addition of both shares.

Share of Solar PV

Previously defined in Equation 8 using Dispa-SET variables, the share of solar PV can be defined within MEDEAS such as:

$$\texttt{share\_pv} = \frac{\texttt{installed\_capacity\_res\_elec(solar)} \cdot \texttt{cp\_res\_elec(solar)}}{\texttt{peak\_load} \cdot CF_{\text{load}}} \tag{23}$$

Where `installed_capacity_res_elec(solar)` describes the installed capacity of solar units in TW, `cp_res_elec(solar)` is the dimensionless capacity factor related to solar energy in MEDEAS, and the factor $CF_{\text{load}} = 0.736$ [-] once again comes from Equation 16.

Net Transfer Capacity ratio

Previously defined in Equation 11 using Dispa-SET variables, $rNTC$ can be defined within MEDEAS such as:

$$\texttt{rNTC} = \texttt{pypsa\_rNTC} \tag{24}$$

Where `pypsa_rNTC` is the external time-dependent data from the PyPSA-EUR model. The implementation of this specific feature is described in Section 4.3.1.

4.2.3 New definitions for curtailment and load shedding

Both targets and associated variables are available in the file `targets.py` (available in the GitHub repository).

Curtailment

As mentioned in Section 2.2.6, curtailment is defined as a constant parameter (`curtailment_res`) in the current version of MEDEAS. It should therefore be replaced. However, a second variable highlighted in orange in Figure 5 is of interest: `res_elec_tot_overcapacity`. This dimensionless variable defines the renewable installations overcapacity within the system and influences the electricity generation from RES as follows:

$$\texttt{FE\_real\_tot\_generation\_res\_elec} = \frac{\texttt{potential\_tot\_generation\_res\_elec}}{1 + \texttt{res\_elec\_tot\_overcapacity}} \cdot \texttt{shortage\_bioe\_for\_elec} \tag{25}$$

Where `res_elec_tot_overcapacity` is the variable of interest, `potential_tot_generation_res_elec` is the potential electrical final energy generated by RES based on the installed capacity (expressed in TW h), `FE_real_tot_generation_res_elec` is the real electrical final energy generated by RES after accounting for the overcapacity (also expressed in TW h), and `shortage_bioe_for_elec` is the dimensionless shortage ratio induced by biogas electricity generation. The latter refers to a prioritization factor for biogas electricity production, which was implemented in the previous MEDEAS version and has remained unchanged since then.

A review of the overcapacity definition has been made since the implementation presented some inaccuracies. The following update resulting from a discussion with MEDEAS' research department is thus made in the initial MEDEAS version:

$$\begin{aligned} \texttt{res\_elec\_tot\_overcapacity} &= \frac{\texttt{potential\_tot\_generation\_res\_elec}}{\texttt{FE\_real\_tot\_generation\_res\_elec}} \\ \rightarrow \frac{\texttt{potential\_tot\_generation\_res\_elec} - \texttt{FE\_real\_tot\_generation\_res\_elec}}{\texttt{potential\_tot\_generation\_res\_elec}} \end{aligned} \tag{26}$$

Since the potential generation is affected by the capacity factor, this variable is nothing more (physically) than the curtailment variable computed by the surrogate model (Equation 12). Therefore, in the MEDEAS proposed version, overcapacity is trivially defined by the MLP output, with a delay to avoid recursion issues:

$$\texttt{res\_elec\_tot\_overcapacity} = \texttt{curtailment\_delayed} \tag{27}$$

In order to avoid a double accounting of curtailment induced by both variables `curtailment_res` and `res_elec_tot_overcapacity`, the first one is set to zero and the second one is represented by the delayed output from the MLP.

The following Figure 27 summarizes the update of curtailment definitions, from the initial model and the proposed version.

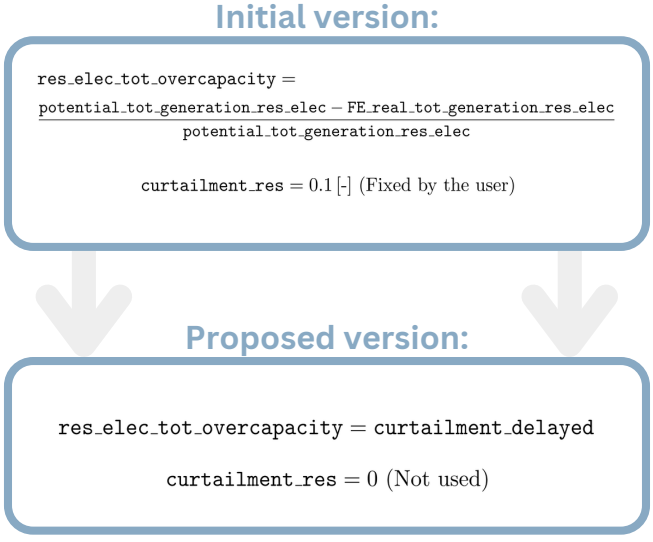


Figure 27: Summary of the curtailment definitions: initial MEDEAS model and proposed version.

Additionally, the Equation 25 is updated to match the definition of the curtailment presented earlier in Equation 12.

$$\begin{aligned}&\texttt{FE\_real\_tot\_generation\_res\_elec} = \\ &\quad \texttt{potential\_tot\_generation\_res\_elec} \cdot (1 - \texttt{curtailment\_delayed}) \cdot \texttt{shortage\_bioe\_for\_elec}\end{aligned} \tag{28}$$

Load Shedding

Regarding load shedding, since the initial version of MEDEAS does not consider this specific variable, the integration is straightforward. The total electrical demand from the economy module is trivially reduced by the energy not served.

$$\texttt{fe\_demand\_elec\_consum\_twh} = \texttt{required\_fed\_by\_fuel(electricity)} \cdot (1 - \texttt{load\_shedding}) \tag{29}$$

Where `fe_demand_elec_consum_twh` is the final demand of electricity after accounting the load shedding (expressed in TW h), and `required_fed_by_fuel(electricity)` is the demand derived from the economic module, representing the desired consumption of electricity considered in the simulation (expressed in TW h).

The Equation 29 constraints the production of electricity to apply the lost of load computed by the surrogate model, possibly heightening the difference between demand (economic module) and production (energy module). For further comparison, the value of lost load is also computed by the proposed version and is defined as follows:

$$\texttt{load\_shed\_twh} = \texttt{load\_shedding} \cdot \texttt{total\_fe\_elec\_demand\_twh} \tag{30}$$

Where `load_shed_twh` is the new assessment of lost load in expressed in TW h, `load_shedding` is the dimensionless output from the SM, and `total_fe_elec_demand_twh` is the total electricity demand in TW h.

4.3 PyPSA-EUR additional data integration

Some required variables to implement the surrogate model within MEDEAS are missing, more particularly to assess the Net Transfer Capacity ratio of the electrical network. The Equations 9, 10, and 11 indeed use the Net Transfer Capacity between zones of the network. This necessary input for DispaSET is not supported by MEDEAS, therefore an additional model is used to bridge the gap: PyPSA-EUR. Additionally, new investments assessments are proposed thanks to PyPSA, which also returns investment predictions for each considered technology during the simulation.

The additional data is obtained from an in-house PyPSA-EUR model of the European sector-coupled energy system in the context of a study described in the following article [22]. The aforementioned study uses PyPSA-EUR sector-coupled model to analyze 28 European countries (composing *ENTSO-E*). The article states that energy-sufficiency measures can significantly lower costs, emissions, and infrastructure demands compared to a business-as-usual scenario. This approach complements energy efficiency and renewable energy integration, highlighting its critical role in the energy transition.

The study was conducted on five full-year simulations, resolved on an hourly basis for the years 2020, 2030, 2040, and 2050, using a myopic pathway optimization approach. This methodology involves

solving an investment optimization problem for each of these years while considering the transition pathway connecting these decades.

In the next subsections new parameters and variables that are integrated in MEDEAS are listed, first to introduce rNTC feature in the model and then to assess investments variables.

4.3.1 rNTC ratio implementation

The initial version of MEDEAS does not considers the ability of the network to transfer electricity. However, as a feature of the surrogate model, this information must be assessed. Additionally, increasing the share of VRES in the electricity generation mix will require a grid adaptation [35], in order to be more efficient, flexible, and reliable to effectively flow power through it. As an important phenomenon in the future of the European grid, incorporating this variable provides a relevant way to enhance MEDEAS representation of the electrical network.

Since MEDEAS is an IAM, it cannot accurately integrate at high level of details whole transmission network and lines parameters as PyPSA-EUR. Indeed, it would definitely impact the resolution time, and anyhow it is not the purpose of an IAM. The method proposed in this work is to integrate a global variable, used as feature in the surrogate model, that depicts this grid evolution: $rNTC$.

For each simulation made using PyPSA-EUR, one $rNTC$ value is computed using the maximum transfer capacities between 28 considered countries (composing *ENTSO-E*), according to the definition which was presented in Section 3.3.4 using Equations 9, 10, and 11.

Since only four values associated to four years (2020, 2030, 2040, 2050) are computed, both extrapolation and interpolation must be done in order to represent the larger time horizon simulated by MEDEAS. A piecewise cubic Hermite interpolation (PCHI) is thus made[3] as presented in Figure 28.

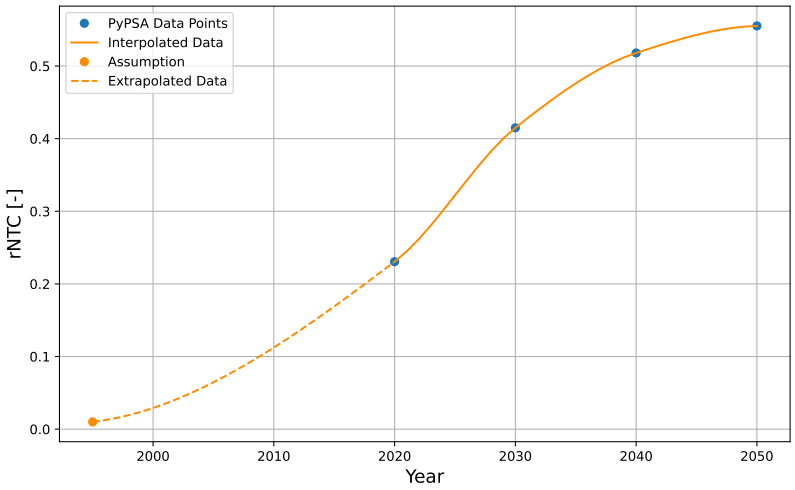


Figure 28: rNTC Feature definition in MEDEAS time horizon.

As it can be seen in Figure 28, an assumption is necessary alongside the PyPSA-EUR data in order to assess values outside the range 2020-2050. The hypothesis of 1% for the $rNTC$ is thus considered at the beginning of the simulation (1995) in order to avoid negative or zero values. It can also be

[3]The PCHI method is preferred over the basic cubic spline interpolation because the latter does not ensure function monotonicity.

observed that the curve stabilizes around 2050 at approximately 55%, indicating a sufficient level of electricity transportability reached.

This representation of $rNTC$ will be used as a "base load" meaning that the ratio cannot be lower than this evolution, representing the minimum increase in grid transfer capacity until 2050. However, as it will be seen in Section 6, larger values could emerge, indicating the system's reaction for potential decrease processes in curtailment and/or load shedding.

4.3.2 Investment data

One of the main objectives of this work is to assess the investments costs in the power sector during the whole energy transition. The initial version of MEDEAS provides also investment assessment, but as it is discussed in Section 7.3.4, the proposed predictions are considered more comprehensive.

Data is extracted from PyPSA-EUR in terms of installed capacity and monetary value. Table 8 displays selected outputs derived from the energy model for the year 2020. For each variable, PyPSA-EUR computes the installed capacity (in MW) along with the associated annualized costs for 2020, 2030, 2040, and 2050. They are divided into three categories in this work: RES, Storage, and Grid.

PyPSA variable	Type	Installed capacity [MW]	Annualized costs [$10^9\cdot$ EUR]
Solar PV	RES	152,364.6	10.485
Wind (onshore and offshore)	RES	209,038.5	30.110
Run-Of-River	RES	44,502.6	13.714
PHS	Storage	53,480	9.770
Hydro conventional	Storage	95,965.3	17.53
Batteries	Storage	18,549.3	0.429
AC lines	Grid	89,496.2	2.122
DC lines	Grid	40,982	1.079
Distribution Grid	Grid	402,915	20.254

Table 8: PyPSA-EUR extracted selected outputs: values for 2020.

The costs considered by PyPSA-EUR are annualized, meaning they are distributed evenly over the expected lifetime of the power plants on a per-year basis and encompass both capital costs (CAPEX) and operational costs (OPEX). This monetary data is expressed in *2020*-EUR, whereas MEDEAS is implemented using *1995*-USD. To reconcile these differences, a conversion is performed using the same inflation adjustment tool used by MEDEAS [36] and the 2020 exchange rate sourced from the European Central Bank's website on the 1st of January 2020: 1.1193 USD/EUR [37]. While MEDEAS does not consider annualized investments but overnight capital costs, the model provides the lifetime expectation for RES power plants. However, since this data is extracted from a 2011 study [38], more recent values from PyPSA datasheet are also considered. The assumptions made in this master's thesis are listed in the following Table 9.

RES	Expected lifetime [years]
Hydroelectricity	80
Geothermal and Biomass	30
Oceanic	40
Wind (Offshore and Onshore)	27
Solar	35
CSP	25

Table 9: Expected lifetime for RES considered in this report, either coming from MEDEAS assumptions (derived from IPCC, SRREN 2011 [38]) or from PyPSA datasheet.

Thanks to the PyPSA-EUR model, the following assessments on future investment costs expressed in *1995*-USD/W are now available: Solar PV, Wind (onshore and offshore), Run-Of-River, PHS, Classical Batteries, AC and DC electrical lines, and Distribution Grid additional lines. Consequently, by considering the future capacity installation for these variables, one could easily compute all associated investments costs. An example for solar and wind investment predictions is discussed in Section 5.5.

Some variables are not directly endogenous in MEDEAS such as the added electrical lines and classical batteries capacities. Nevertheless, PyPSA-EUR also computes the capacity evolution, so it can be considered as exogenous parameters within MEDEAS.

A major assumption is made regarding the assessment of the storage-related investments and installed capacities in the proposed MEDEAS version. Since the model only supports Pumped Hydro Storage (PHS) and flexible storage from Electric Vehicles (EVs) at a large scale, it is assumed that all additional storage corresponds to these technologies. This constitutes a significant assumption, as PyPSA's dispatch also accounts for classical batteries whose impact is considered non-negligible in the future power system. However, integrating a new large-scale storage technology within MEDEAS is considered out of this master's thesis scope, and this assumption is made, regardless of the potential bio-physical limitation of PHS or EVs.

To summarize, following additional properties are now integrated in the initial version of MEDEAS:

1. All six Surrogate Model features, including the new definition of *rNTC*.
2. Both Surrogate Model targets.
3. PyPSA-EUR additional data regarding capacities and investments, expressed in *1995*-USD.

All the new variables implemented in MEDEAS during this master's thesis are listed in the Appendix A and the associated python files are available in the Github repository. The next section presents a typical simulation with all these new variables integrated.

5 MEDEAS simulations with the new implementation

5.1 Overview

In this section, the results from the integration of the surrogate model and the additional data are presented. First, two scenarios are simulated within the proposed MEDEAS implementation: a baseline scenario (Section 5.2) and a high-RES (Section 5.3) scenario. Both sections present the features and targets evolving in MEDEAS ecosystem. For the sake of clarity, only the MLP surrogate model's outputs are presented in these two sections, however both pairs of outputs are compared in Section 5.4.

5.2 Baseline scenario

The baseline scenario used in this report is based on the Shared Socioeconomic Pathways (SSPs) which are the standard set of scenarios in current climate change research [39]. The SSPs are a set of five scenarios used to explore how global trends (energy, economic, social, and technological) might evolve and interact with climate change. They range from sustainability-focused development (*SSP1 - sustainability*) to fossil-fueled growth (*SSP5 - conventional development*), providing frameworks to assess climate impacts on the society. The baseline scenario is based on the paradigm described by *SSP1* where a strong energy transition is established, aiming to achieve a European NZP in 2050. A set of set of numerous policies starting in 2020 with the aim to reduce the carbon emissions are implemented. These policies include: higher RES integration, slight increase in nuclear power, higher society electrification (*e.g* shift towards HEVs), higher recycling rates of minerals, etc.[4]

Unfortunately, when *SSP1* is simulated within the new implementation of MEDEAS which embedded the surrogate model, the share of solar PV reaches its upper bound in 2043. This results in a special case which will be discussed in Section 5.3. In order to create a scenario staying within the space of definition from the SM, the reductions of solar capacity are presented in Table 10 are applied:

	2020	2030	2040	2050
SSP1	0.124	0.600	1	1.4
Baseline scenario (used in this report)	0.124	0.350	0.700	1

Table 10: Capacity of Solar PV installed during the simulation expressed in TW, *SSP1* and baseline scenario.

The baseline scenario can thus be considered as a transition towards NZP by 2050, but with lower RES penetration than SSP1. Indeed, the simulation of the baseline scenario provides a total emission of 0.27 Gt CO_{2e} in 2050, which can be considered as a NZP.

When the baseline scenario runs in MEDEAS, both the targets and the features presented in Section 3.3.4 are now evolving in the MEDEAS ecosystem as it can be seen in Figure 29 and Figure 30, where *Dmnl* on the y axis stands for Dimensionless.

[4]The scenario is further described in the SSPs overview [39], and the scenario implementation Excel file is available in the Github Repository of the master's thesis.

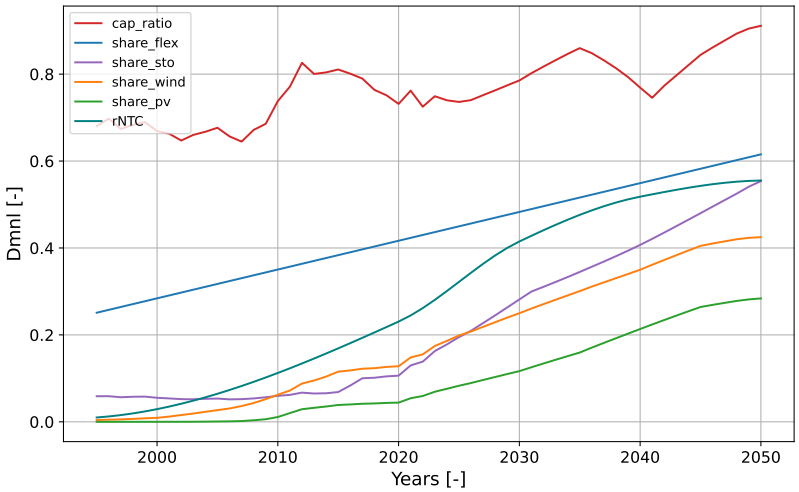


Figure 29: Evolution of the six inputs of the surrogate model in MEDEAS, for the baseline scenario.

Figure 29 highlights the energy transition followed by the baseline scenario. Indeed, while the shares of wind, solar, and storage capacities naturally increase, both main assumptions regarding the surrogate model features can also be observed. In particular, the linear interpolation of `share_flex` (Section 4.2.2) and $rNTC$ cubic Hermite interpolation (Section 4.3.1). Regarding the capacity ratio, it exhibits an oscillatory trajectory throughout the simulation while maintaining a generally increasing trend.

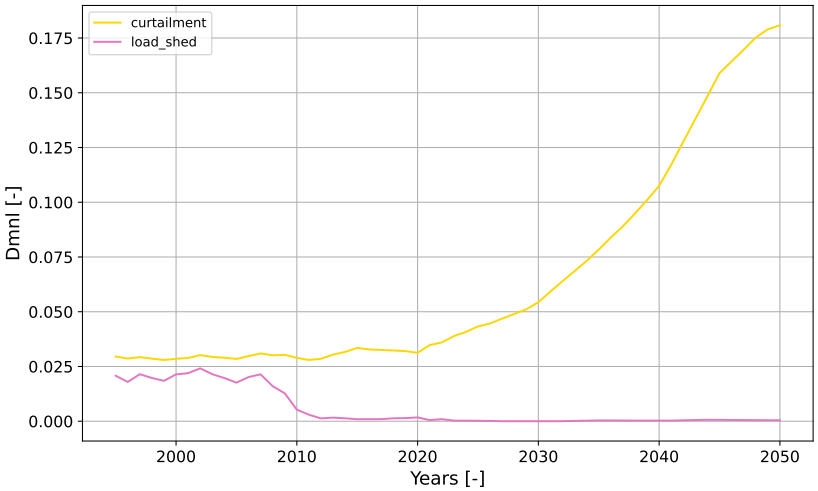


Figure 30: Prediction of the level of curtailment (Equation 12) and load shedding (Equation 13) for the baseline scenario.

Figure 30 shows that both outputs of the surrogate model follow expected trends for the baseline scenario. Indeed, while the curtailment increases along the integration of the VRES in the electrical mix, the level of load shedding remains moderately present at the beginning of the simulation and then negligible. It suggests that the current version of the model sufficiently invests in power generation capacity to avoid a high unsatisfied demand but lacks in flexibility resources at the end of the simulation to balance curtailment. The investment assessments provided by the model are presented below for

the RES integration (Figure 31), the improvement in grid transfer capacity and storage capacity (Figure 32).

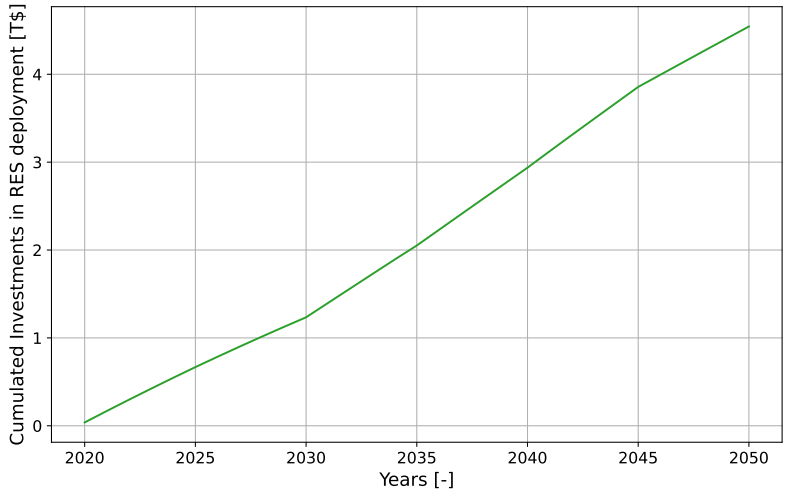


Figure 31: Cumulated investments assessment from 2020 for new RES installation in the European grid, baseline scenario.

As shown in Figure 31 and Figure 32, cumulated investments in all sectors naturally increase, reflecting the ongoing need for yearly investments and efforts to support the grid transition. The larger share of these investments is allocated to deploying RES, followed by new storage installations, and finally, the expansion of grid transfer capacities.

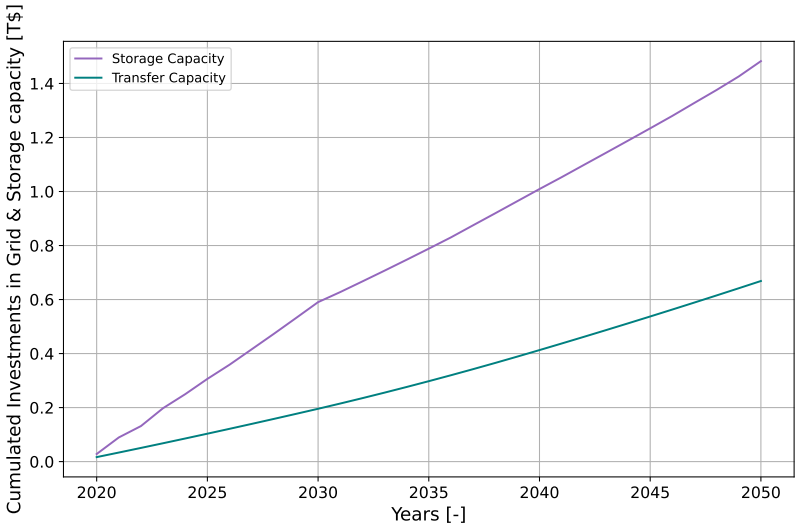


Figure 32: Cumulated investments assessment for both transfer capacity and new Storage installations in the European grid, baseline scenario.

In order to assess the reliability of these investments, Section 5.5 presents the results for a simulation of the *SSP1* NZP by 2050 scenario, and a comparison with external data derived from a 2020 EC communication is made.

5.3 High-RES scenario

In order to show the limitations of this new version, another more aggressive scenario is studied: while everything remains the same as in the previous case, much more wind and solar capacity are simultaneously deployed. Obviously, it results in higher levels of non-desired curtailment in Figure 34, but it can also be observed in Figure 33 that boundaries of the features design space are reached.

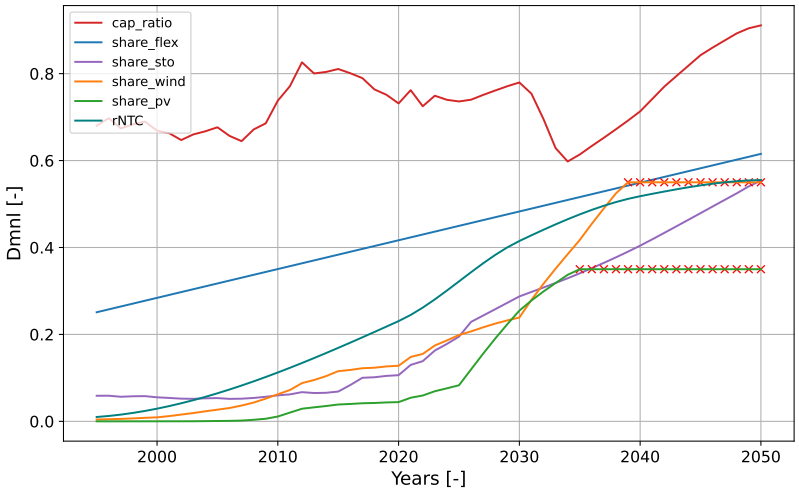


Figure 33: Evolution of the six inputs of the surrogate model in MEDEAS, for the high-RES scenario. Red crosses represent features values which are outside the surrogate model's design space.

In order to avoid unpredictable behavior from the surrogate model, all features must remain in the design space defined by the boundaries specified for each feature in Section 3.3.4. When a scenario presents one or more variables outside its boundaries, this variable is set to the nearest (upper or lower) definition bound to prevent any pathologic cases. However, when this phenomenon occurs, the surrogate model inputs differ from the actual simulation state, meaning that the values provided by the surrogate model cannot be considered accurate anymore.

When comparing Figure 30 and Figure 34, it can be observed that curtailment is indeed related to the increase of VRES penetration. The peak reached in Figure 34 is caused by the stabilization of both wind and solar shares within the SM while the other features such as the transfer capacity ratio, storage and flexibility shares, are still growing.

This subsection highlights the importance to wisely define the scenario to study and the limitation of the design space. As it was previously mentioned in Section 3.3.8, the design space is limiting the surrogate model horizon of definition, thus restricting the model to a tight simulations scope. However, the scope proposed by the surrogate model enables enough space for studying various interesting scenarios and policies, ranging from fossil fuels promoting policies, or business as usual scenarios, to near zero carbon emission cases. Indeed, the upper bounds for wind and solar energies (which are the most restrictive boundaries) are respectively 55% and 35% of the peak load, which are already very high level of penetration. Moreover, it is assumed that the outputs given by the surrogate model, even when one or more features reach its boundaries, are more relevant than the previous assumption from the former model.

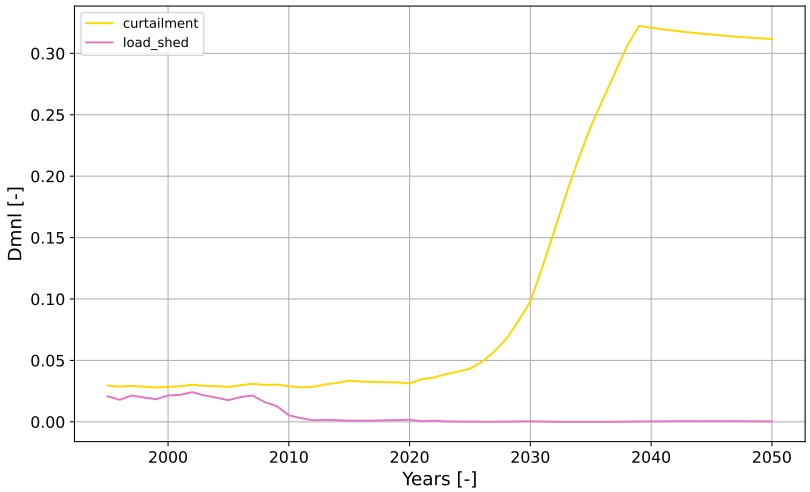


Figure 34: Prediction of the level of curtailment and load shedding for the high-RES scenario.

5.4 Random Forest or Multi-Layer Perceptron?

This subsection briefly compares the outputs of both surrogate models described in Section 3.3.2 in order to determine whether one model is more suitable to be integrated in MEDEAS. The following Figure 35 shows curtailment and load shedding computation from both surrogate models algorithm, in the context of the previous baseline scenario (Section 5.2).

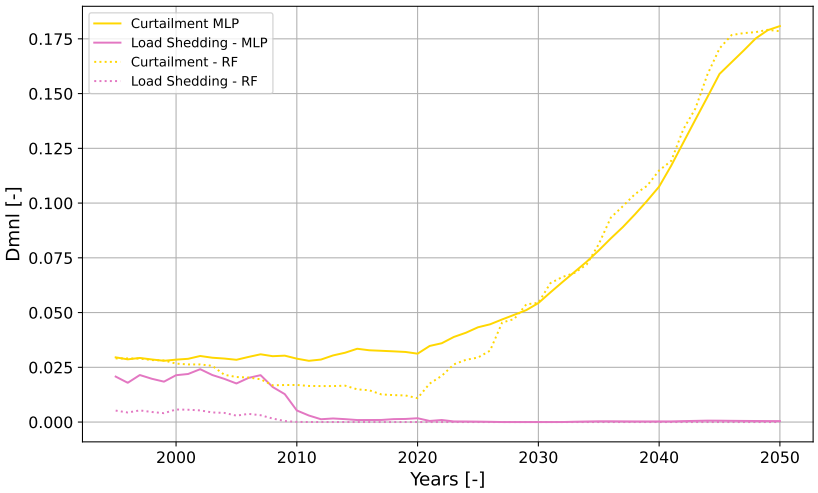


Figure 35: Prediction of the level of curtailment and load shedding for the baseline scenario, comparison between MLP and RF surrogate models.

It can be observed that both curtailment and load shedding predictions are relatively close to each other, and presents the overall same tendencies. Nevertheless, in the rest of this report, the Multi-Layer Perceptron will be preferred over the Random-Forest algorithm. This is justified by the better predictions on unseen data accorded to the MLP in Tables 2a and 2b.

5.5 Investment data

The investment predictions are not yet integrated into any feedback loop in the proposed MEDEAS version. However, to demonstrate the effectiveness in providing insights on the monetary efforts required for the energy transition, an example is presented in this section. This example focuses on a 2050 NZP scenario and is used to confront obtained results with external data to assess their accuracy, in this case a 2020 EC communication. The Figure 36 uses both the external prices from PyPSA-EUR and the endogenous installation capacity simulated by MEDEAS.

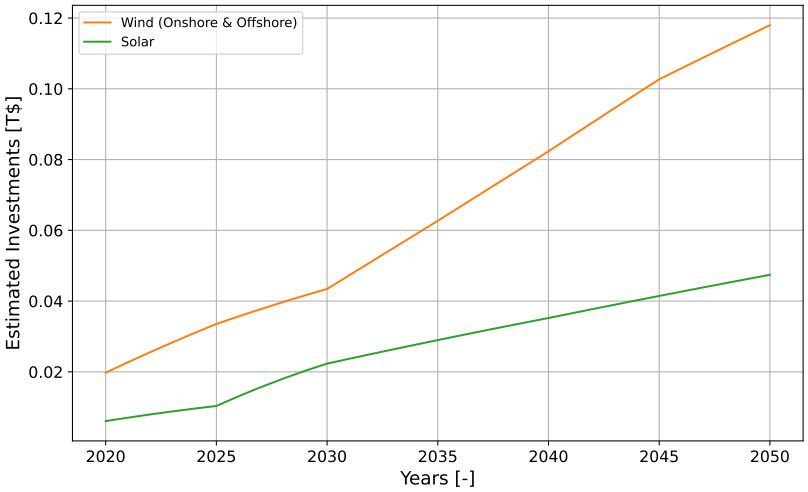


Figure 36: Investments assumption for Solar and Wind (onshore and offshore) installations in the European Grid, from 2020 to 2050.

Figure 36 is directly in line with the European Commission's communication on offshore energy strategy from November 2020 [40]: *"Annual investment in onshore and offshore grids in Europe over the decade to 2020 have amounted to around EUR 30 billion but need to increase to above EUR 60 billion in the coming decade, and then increase further after 2030."*[5]

[5]Statement expressed in *2020*-EUR, but Figure 36 in *1995*-USD.

6 Society Reaction to curtailment and load shedding

6.1 Overview

Previous Section 5 details the integration of curtailment and load shedding in MEDEAS. Once both variables are computed, one can attempt to adjust the system in order to decrease these undesired outputs, through investments in the electrical grid infrastructure. The purpose of this section is to model a potential perfect reaction from the society to reduce both outputs from the surrogate model. To do so, as it will be seen in Section 6.3, two electrical grid investment mechanisms will be implemented, which will directly impact:

1. RES installed capacity,
2. Storage installed capacity,
3. Net Transfer Capacity of the grid.

Since these capacities have a direct influence on curtailment and load shedding, an increase in grid investments should decrease both outputs of the surrogate model.

Two distinct processes are implemented in order to implement these feedback mechanisms, one for each target (Section 6.3). Before that, a brief description of the PID control is proposed in Section 6.2. This section ends with results and limitations of this modeling of society reaction (Section 6.4).

6.2 Basic PID control reminder

The following description was inspired by the book *PID Control - New Identification and Design Methods* by Michael A. Johnson and Mohammad H. Moradi [41].

The *PID* control methodology is composed by three terms associated to the three letters *P, I, D* referring to standard three parts of the controller: Proportional, Integral, and Derivative terms. This controller is well known in the field of control engineering and is widely used in the industrial sector. The aim of such a device is to minimize the error over time of some output(s) by adjusting specific system variables. Both Figures 37 and 38 respectively depict a typical industrial control loop and a PID controller representation.

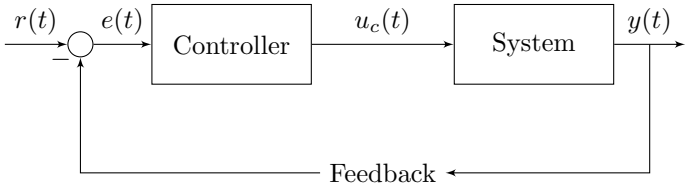


Figure 37: Components in a typical industrial control loop.

Where $r(t)$ is the reference variable (or the setpoint), $e(t)$ is the error term, $u_c(t)$ the output of the controller (or the control variable), and $y(t)$ the measured process variable.

The *System* block is the actual process or system for which a specific variable needs to be controlled. It's a set of equations that describes either a complex or simple physical phenomenon (or phenomena). A *Feedback* control loop closes the system to continuously monitor its output (and the error term) at each time step and thus adjusting the controller's reaction accordingly. The *Controller* block represents the PID control device and is illustrated in the Figure 38.

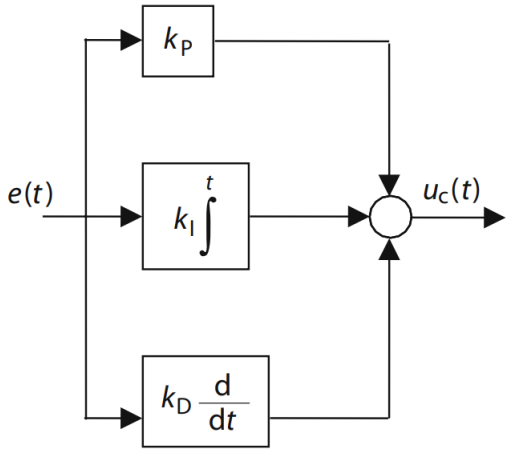


Figure 38: PID Controller Block: time domain representation, taken from [41].

Where k_P is the Proportional gain, k_I the Integral gain, and k_D the Derivative gain, all composing the *PID* controller.

Another way to interpret Figure 38, is:

$$u_c(t) = k_P \cdot e(t) + k_I \int e(\tau)d\tau + k_D \frac{\mathrm{d}}{\mathrm{d}t} e(t) \tag{31}$$

The control variable $u_c(t)$ is defined by:

1. The Proportional term, the direct action of the controller which is proportional to the error.
2. The Integral term, which considers the previous values of the error and is used for correcting steady state error. With τ the integration variable.
3. The Derivative term: which considers the rate of change of the error and adapts the control action.

The tuning procedure could be divided into two distinct parts, namely the structure of the PID controller and the scalar value of the k-coefficients.

On the one hand, the PID control structure can also be derived as P or PI (usually not PD), depending on the system requirements. Basic P controllers handle stability, PI controllers address steady-state errors, and PID controllers provide precise and fast responses. The choice depends on the system dynamics and the performance goals.

On the other hand, regarding k-coefficients, even if the implementation of a PID controller is straightforward, it heavily depends on the scalar values for the gains (k_P, k_I, and k_D in Equation 31) which can be challenging to tune. In order to obtain an optimal control function for a specific application, the k constants must be tested using a sensitive analysis. There is no universal rule for determining optimal coefficients since they depend on the system response itself, and the potential sensors or delays that could be present. However, some heuristic methods exist and give simple methodology to apprehend the tuning, as the Ziegler-Nichols methods [42]. In a nutshell, a good practice is to start with all coefficients initialized to 0, and slightly increase the proportional gain until the ultimate gain (denoted as k_U) is reached, where a stable system's output is found. Both Integral and Derivative gains could be deduced afterwards using empirical formulas given in [42].

6.3 Grid investments reaction to curtailment and load shedding

6.3.1 Link with control theory

The objective is, based on the control theory discussed in previous Section 6.2, to create a perfect reaction mechanism that will decrease both surrogate model outputs in order to assess the required efforts and investments to overcome curtailment and load shedding. These mechanisms are inspired by PID controllers, and the case of load shedding reaction is depicted in Figure 39.

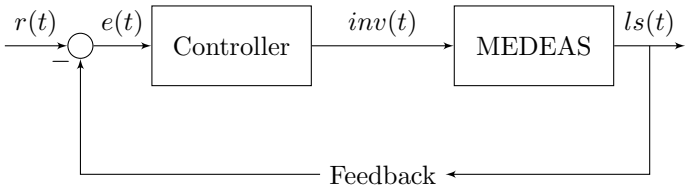


Figure 39: PID Control analogy: society reaction to load shedding.

Where $ls(t)$ represents load shedding, $r(t)$ is set to 0% being the reference for load shedding (setpoint), $e(t)$ error between setpoint and actual load shedding value, $inv(t)$ represents the investments computed by the controller and represent the reaction of the society to the increase of load shedding.

To ensure stability and avoid erratic responses in the control mechanism, only the proportional (P) gain of the PID controller will be set to non-zero. While PID controllers typically incorporate three aforementioned gains to achieve precise control, the exclusive use of the P gain ensures a more straightforward and stable response, particularly in this context of addressing load shedding and curtailment scenarios. This approach minimizes the risk of overshooting or oscillations, allowing for a controlled and reliable estimation of the investments needed to mitigate these challenges. A counter example where integral gain is used will be shown in Section 6.4.2

6.3.2 Optimal dispatch for investments

In order to assess the impact of any financial efforts on the system, directives regarding the investment allocation must be set. How will a 1000 $ investment affect the grid in terms of capacity? Since MEDEAS is an IAM and not an optimization model, it cannot compute the optimal investment dispatch at each iteration. The method proposed in this work consists in extracting the investments made in PyPSA-EUR simulations and compute the shares in percentage to track where the money is placed between all considered sectors. Once again, interpolation of the given data is required. Figure 40 represents, in percentage, how the society will react in terms of investments into new added capacities through the years for a load shedding reduction policy. For the sake of clarity, values are displayed at 5-years steps but are computed for each year.

Obviously, shares in investments are varying along the years. It can be seen that the RES integration is the field where more investments are dispatched, with significant increase in Solar and Wind shares. These shares are stored in MEDEAS as external constants that remain unchanged during the simulation. It is of course a huge assumption that investments policies cannot change regarding the simulation context, but it can give a good insight on investments allocation during the years.

6.3.3 Feedback implementation

Figure 41 presents a summary of the procedures for implementing the investment feedback mechanisms. Two distinct processes are implemented separately, one for each surrogate model output. If the model detects either the presence of load shedding or curtailment, an investment is allocated to the next time step. This investment, symbolizing a direct and instantaneous reaction of the society, depends

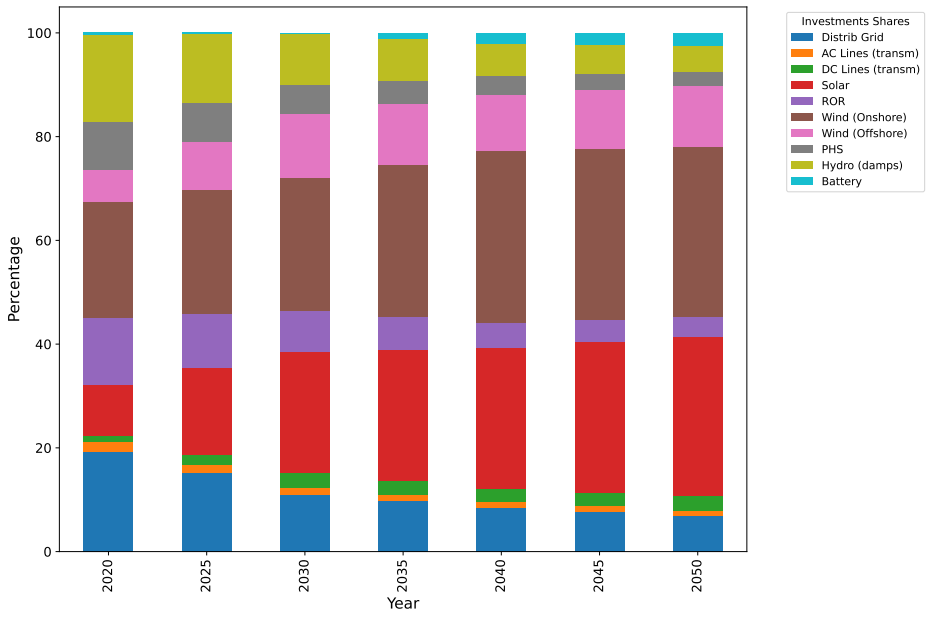


Figure 40: Technology Investments shares (load shedding) for additional capacity (every 5 years but each year available).

on the associated parameters (k-coefficients) that are defined by the user. Once this monetary value is computed, using shares depicted in Figure 40, investments are dispatched into several domains (Distribution Grid, AC/DC lines, Solar, Wind, ROR, PHS, Hydro storage, or Battery) and finally converted in additional new power installations.

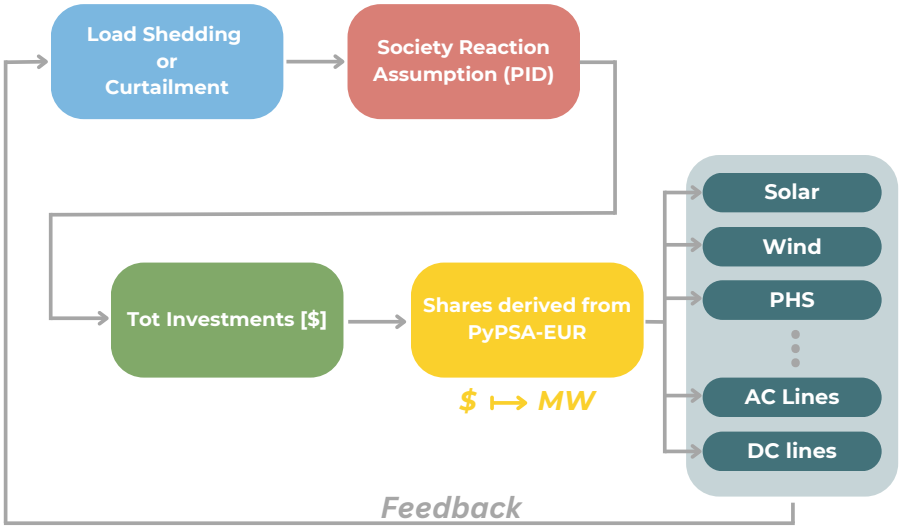


Figure 41: Diagram representation for the society feedback investments implementation.

The main difference between load shedding and curtailment related mechanisms, despite k-coefficients

parametrization, is the impact on the grid for the investments. Indeed, by definition, limiting curtailment in the grid cannot be done with simultaneous increase in VRES. The curtailment mechanism thus tends to increase only storage and transfer capacities in the electrical grid. Regarding load shedding, an increase of all parameters (RES, Storage, and Transfer capacities) is beneficial in order to balance production and consumption, thus limiting loads that might not be fulfilled. The key results and limitations are discussed in the next section.

6.3.4 MEDEAS definitions update

In order to implement this feedback mechanism, some MEDEAS variables must be updated to take into account the additional feedback term.

RES and storage capacity

The definitions of installed RES and storage capacities are updated to include the cumulated additional feedback capacity by simply adding a new variable, respectively: `cumulated_capacity_res_elec` and `cumulated_storage_feedback`. These variables are updated at each iteration and influence the associated capacity endogenous variable (*ie* are non-zero) when the feedback mechanism is activated. The following MEDEAS definitions are thus modified in the new version: `installed_capacity_res_elec` and `total_capacity_elec_storage_tw`.

rNTC

In a similar manner, $rNTC$ which was previously defined in Equation 11 using Dispa-SET variables and in Equation 24 within MEDEAS environment, is updated as follows:

$$\texttt{rNTC} = \texttt{pypsa\_rNTC} + \texttt{cumulated\_add\_rNTC\_feedback} \tag{32}$$

Where `pypsa_rNTC` is the external base load from the PyPSA-EUR model (Section 4.3.1), and the second term `cumulated_add_rNTC_feedback` represents the cumulated added increase in $rNTC$ expressed in percentage computed by the additional grid capacity from the feedback process.

6.4 Results and limitations

Different cases are shown in this section in order to present the results and limitations regarding the investments feedback mechanisms. First the baseline scenario is used with a moderate-intensity feedback, resulting in near zero load shedding and curtailment reduction. Subsequently, another implementation is analyzed where an aggressive feedback leads to undesirable and even counterproductive effects on curtailment.

The reduction mechanisms are used in this section from the beginning of the simulation (*ie* from 1995) for demonstration purposes. Nevertheless, the activation year of the mechanism can be set by the user in order to represent the society reaction to future policies, starting from 2025 for example.

6.4.1 Moderate feedback with baseline scenario

First of all, it is important to acknowledge the tradeoff that might appear with the reduction processes of the load shedding and curtailment. Indeed, to avoid load shedding, the model tends to augment VRES shares in the mix, which by definition leads to a curtailment growth. Nevertheless, by increasing

the Net Transfer Capacity ratio and the share of storage, the curtailment reduction process also helps decreasing load shedding. Since this feature is already at low levels in most cases, it will be enough to simply tune $rNTC$ feature to obtain negligible load shedding.

The following parameters are fixed for this example:

	P	**I**	**D**
Curtailment	0.01	0	0
Load Shedding	0.00001	0	0

Table 11: Parametrization for gains (k-coefficients) for a moderate feedback implementation.

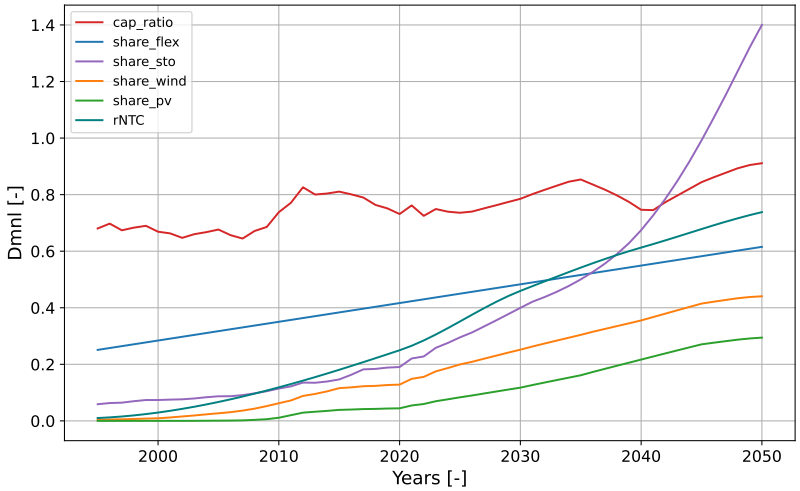


Figure 42: Evolution of the six inputs of the surrogate model in MEDEAS, for the baseline scenario with moderate curtailment reduction implemented.

Figure 42 illustrates the implications of added storage capacity in the reduction process, which is the most significantly increased feature compared to the case without feedback (Figure 29). The scenario depicted here suggests a strong policy in deploying new storage capacity, in order to counter balance the intermittency caused by the rise of VRES in the electrical mix. Investments in the Net Transfer Capacity of the grid are also present, increasing the $rNTC$ by 20% in 2050. The four other features remains the same with or without the curtailment reduction activated. Of course, the increase in grid's flexibility impacts the outputs of the surrogate model. Following Figures 43 and 44 demonstrate the effectiveness of the curtailment reduction, while the slight variations regarding curtailment.

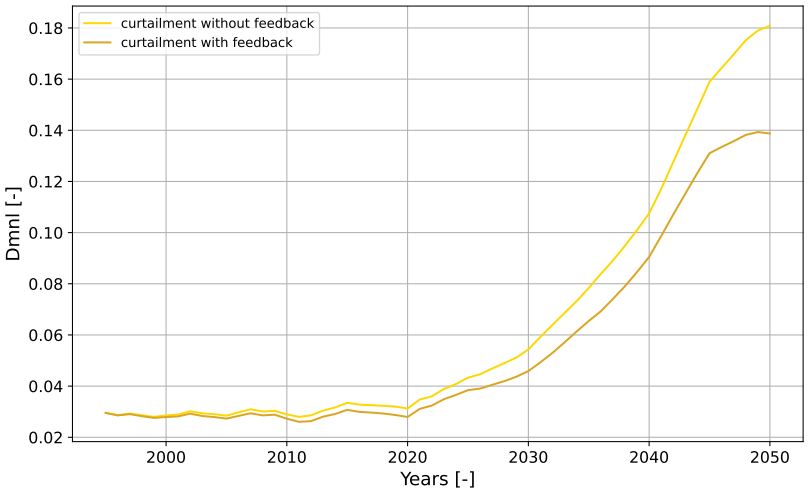


Figure 43: Prediction of level of curtailment with or without moderate curtailment reduction.

Figures 43 shows that a 6% reduction of curtailment by 2050 can be reached with the flexibility increase proposed by the feedback process. Regarding the load shedding, the Figure 44 illustrates that since the unsatisfied demand is already at a very low level (near zero), the reduction does not significantly impact this target.

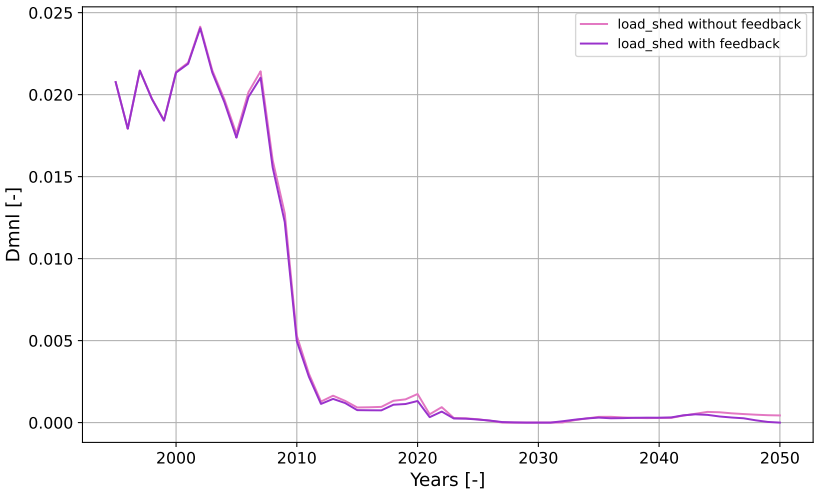


Figure 44: Prediction of level of load shedding with or without moderate curtailment reduction.

The reduction processes naturally involve higher financial commitments towards renewable energy deployment and system flexibility (through flexibility assets: storage and transfer capacity) as demonstrated in Figure 45. According to these predictions, reducing the curtailment by 6% by 2050 compared to the baseline scenario results in a cumulated amount of approximately 3 T$ additional investments in 50 years.

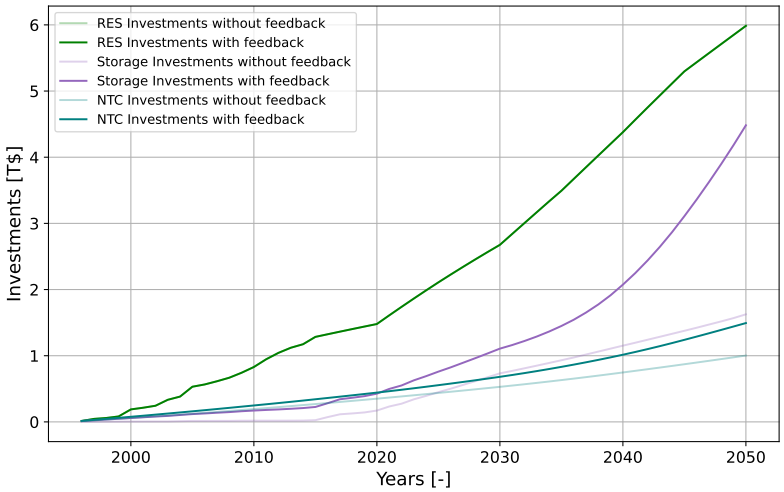


Figure 45: Investments assessment for RES deployment, new Storage installations, and transfer capacity improvement in the European grid, with or without moderate curtailment reduction.

Although the feedback mechanism indeed helps to decrease both surrogate model outputs, it is important to keep in mind that such instantaneous reductions are obviously not considered practically feasible. Indeed, the purpose of this section is to provide insights on the potential investment policies necessary to constraint curtailment and load shedding.

6.4.2 Aggressive feedback with baseline scenario

This subsection proposes a counter example where the limitations of the feedback processes are demonstrated. In this case, a too aggressive reduction process is implemented, inducing the reach of several bounds of the design space and causing undesired effects on the targets, especially for curtailment. It can also be noticed that the integral gain is set to non-zero, which will lead to unsteady results in both features and targets.

	P	**I**	**D**
Curtailment	0.1	0.1	0
Load Shedding	0.1	0.1	0

Table 12: Parametrization example for gains (k-coefficients) for a too aggressive feedback implementation.

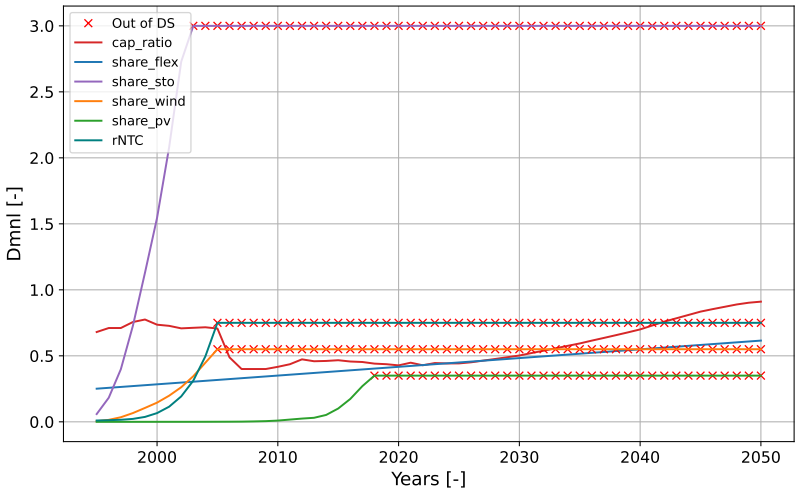


Figure 46: Evolution of the six inputs of the surrogate model in MEDEAS, for the baseline scenario with aggressive curtailment reduction implemented.

As it can be seen in Figure 46, the feature curves grow exponentially, resulting in unsteady behavior and quickly reaching the boundaries of the design space. The outputs for this pathological case are obviously not shown here.

7 Results and Comparative Analysis

7.1 Overview

This section outlines the modifications made in this master's thesis framework and compares the initial version (PyMEDEAS 2) with the updated version developed in this report.

One of the primary observations was that wind turbine installation potentials in the initial implementation of MEDEAS-EU are very low compared to the solar panels potential, which seems surprising. To address this, a literature review is conducted to evaluate RES potential deployment across Europe. Additionally, updated data from PyPSA-EUR (specifically from the *atlite* tool) is incorporated. The modifications resulting from this update are summarized in Table 13.

Following these updates, a comparative analysis is performed to evaluate the differences between the proposed MEDEAS version and the initial version. The results from this analysis are summarized in Table 16.

7.2 RES potentials

To ensure conciseness, an evaluation of some RES potentials in Europe is conducted to check the predefined values in the initial PyMEDEAS 2 version. In order to evaluate these potentials, the results from PyPSA-EUR simulations [22] (Section 3.4.) which embedded *atlite* Python package [32] are summarized in Table 13. The analysis conducts to an update for three predefined values, namely: Solar PV, Wind onshore and offshore potentials. Additionally, oceanic potential is also revised based on a European Commission communication, as detailed below.

atlite [32] is a Python-based tool designed to evaluate renewable energy potentials. It uses land-use data (*e.g.* CORINE [43], Natura2000 [44]) and environmental constraints (*e.g.* shipping lanes, GEBCO bathymetry data [45]) to compute installable capacities for renewable technologies for each node of the network. Additionally, by integrating weather datasets, the tool determines the maximum capacity potential and provides available generation time series, assuming distribution proportional to capacity factors. This ensures accurate modeling of geographically distributed energy systems, making *atlite* a decent and recent reference to compare RES potentials.

RES	Former Potentials [GW]	Update [GW]
Solar PV	13,000[6]	10,310.4
Wind Onshore	250	8,049.78
Wind Offshore	100	581.63
Ocean Energy	1	40[7]

Table 13: RES European potentials installation, former (PyMEDEAS 2) and new values from *atlite*.

Table 13 presents the update for solar, wind, and oceanic potentials. At a first sight it could be noticed that a large difference between the former and the new value for onshore potential is made. It seems that the new estimation is more in link with current literature, even being more conservative, as suggest the following articles values: 13.4 TW [46], 52 TW [47]. Globally, it seems that MEDEAS highly

[6] The former PV potential in MEDEAS is implemented in a way that it is constant between 1995 and 2020 around 8.000 GW. After that, it linearly increases up to 13.000 GW GW to 2050. Despite the fact that potentials are time-dependent since technological constraints tend to be relaxed while technological development occurs, this work suggests to be more conservative and set a constant potential limit to 10.310 GW, which is derived from *atlite*.

[7] This value does not come from *atlite* as the others but from a personal assumption, in line with the citation from [40]. This assumption is judged conservative since 40 GW is considered as the *"achievable installed capacity for 2050"*.

prioritizes solar by limiting wind integration to a maximum policy constraint of 350 GW. The proposed version of MEDEAS suggests instead to relax these constraints based on the following references which predict a more balanced energy mix (even promoting wind over solar) between respectively wind and solar PV: ∼35% - ∼25% share in power generation in 2030 [48] (Figure 2), and ∼25% - ∼11% share in electricity generation in 2050 [49] (Figure 38).

Regarding offshore power plants, a 2020 communication from the European Commission about offshore renewable strategy [40] clearly states that:

"Offshore renewable energy is among the renewable technologies with the greatest potential to scale up. Starting from today's installed offshore wind capacity of 12 GW, the Commission estimates that the objective to have an installed capacity of at least 60 GW of offshore wind and at least 1 GW of ocean energy by 2030, with a view to reach by 2050 300 GW and 40 GW of installed capacity, respectively, is realistic and achievable."

Based on the average electrical peak load reference value given in Romain master's thesis [8] (Table 4.2), the former wind potentials from PyMEDEAS 2, the initial capacity factors considered for wind (onshore and offshore), and the definition of the associated feature (Equation 6), the share of wind would have been constrained to:

$$UB_{\text{wind}} = \frac{250 \cdot 0.32331 + 100 \cdot 0.22092}{441} = 0.23\,[\text{-}] \tag{33}$$

Which seems too conservative on the basis of the aforementioned references [48], [49][8].

7.3 Comparative analysis: MEDEAS before and after (low-RES scenario)

This subsection compares two MEDEAS models: the initial version of PyMEDEAS 2 and the new version proposed in this master's thesis. In order to make a proper analysis, both models must simulate the exact same scenario. However, since some potential constraints are relaxed in the proposed version (Section 7.2), the baseline scenario cannot be selected since the integration of wind and oceanic installations exceeds previous limitations.

A new framework is thus defined: low-RES scenario, a derivative of the baseline scenario defined in Section 5.2 where everything remains the same except for the RES capacity installation, which is decrease to remains within the potentials limits of the first version. It has been decided not to activate the feedback mechanism discussed in Section 6 in the proposed version implementation during this comparative analysis.

[8] The shares presented in the Equation 33 and in the percentages extracted from [48] and [49] are not representing the same physical variable. Indeed, while the first one is defined in Equation 22 and represent a proportion of the peak load, the others represent a share in an energy mix. However, the comparison is still relevant since the upper bound computed is considered very low.

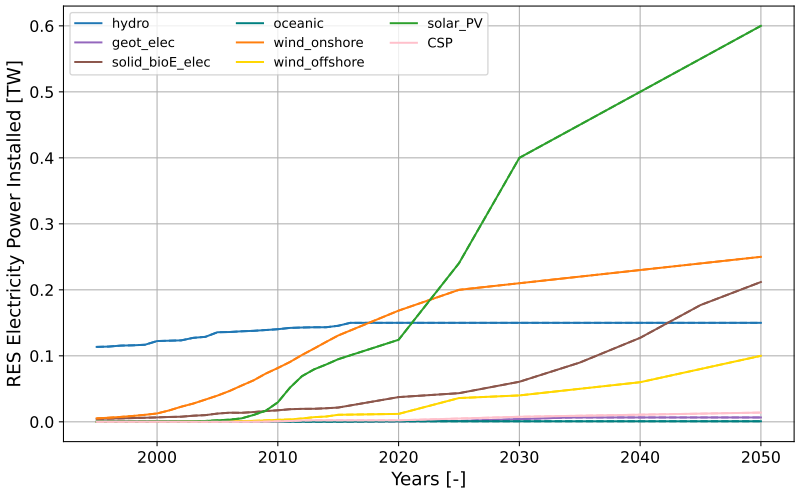


Figure 47: RES power installed evolution used in the comparative analysis ().

The power installed differences between the baseline scenario and the low-RES scenario are presented in the Table 14 for 2050.

RES	Baseline	Low-RES
Oceanic	0.0056	0.001
Wind onshore	0.336	0.250
Wind offshore	0.3	0.1
Solar PV	1	0.6

Table 14: Comparison between RES installations (TW) in 2050 between baseline and low-RES scenario.

The next paragraphs present the main differences observed when running this low-RES scenario on both MEDEAS models. At the end of this section, Table 16 summarizes key differences for a sake of clarity.

7.3.1 Computation times

Both the initial and updated implementations were simulated on a standard laptop equipped with 8GB of RAM and an Intel Core i5 processor. Simulation times ranged from approximately 14 to 20 minutes, depending on the laptop's performances, with no significant differences observed between the two models. This suggests that the integration of the surrogate model and the introduction of new variables have minimal impact on the model's efficiency. However, conducting a more extensive analysis involving a larger number of simulations would be necessary to accurately assess the impact of these modifications. Such an analysis would be computationally intensive and challenging to perform on a standard laptop so it is considered out of the scope of this master's thesis and thus will not be conducted in this report.

7.3.2 Curtailment

The first main difference is obviously the curtailment implementation made in this report's framework. As it was previously mentioned in Section 4.2.3, the former MEDEAS version proposed two variables describing this phenomenon, namely: `curtailment_res` and `res_elec_tot_overcapacity`. While the first one is an external parameter which influences the potential generation, the second is an endogenous variable which was updated (Equations 26, 27) and influences the real RES electricity generation.

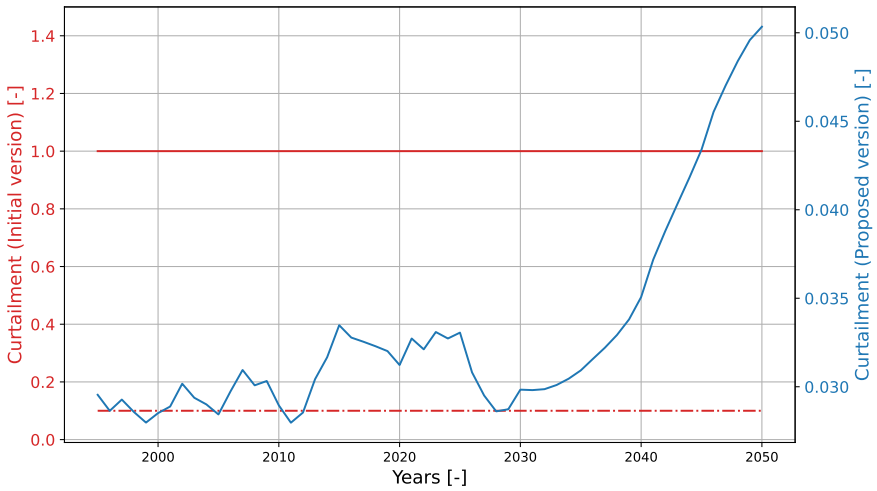


Figure 48: Comparison between initial version of PyMEDEAS 2 and proposed version: curtailment variables.

Both variables embedded within the former implementation are constant and fail to accurately represent system dynamics compared to the proposed endogenous variable produced by the surrogate model. Specifically, in the previous version, the overcapacity is incorrectly fixed at 1 (Equation 26), and the external variable is set at 10% for Solar PV, Wind (onshore and offshore), and CSP, while it is set to zero for other RES types (Hydro, Geothermal, Biomass, and Oceanic). Regarding the proposed MEDEAS model, it endogenously depends of the electrical power installation mix depicted in Figure 47 and grows along the introduction of VRES while remaining moderately present during the simulation.

7.3.3 Gap in electricity balance

This subsection compares the electrical generation mixes between both models and the electricity consumption. The electricity demand curves in Figure 49 and Figure 50 represent the total final energy electricity demand in TWh. This is the total amount of electricity required by the economy module, before the economy-energy feedback. It includes new electric uses (*e.g.* EV & HEV) and electrical transmission and distribution losses.

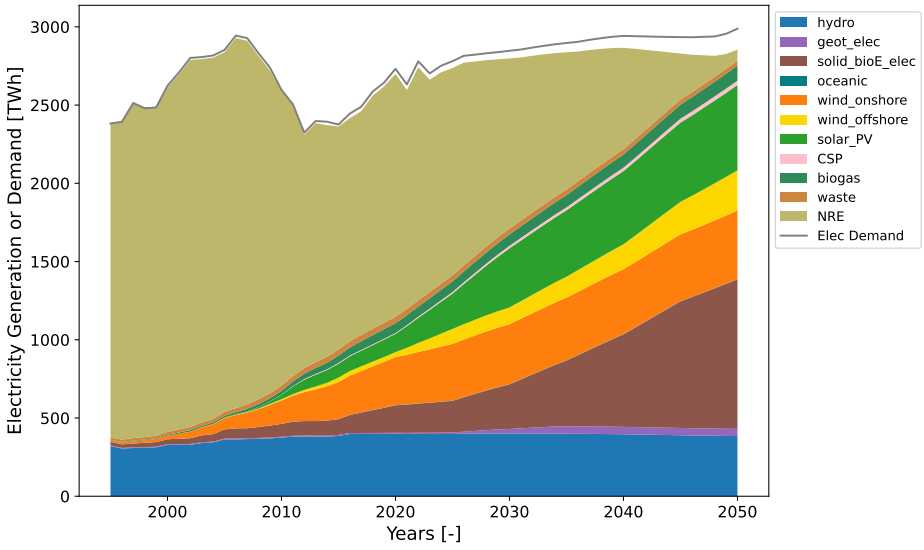


Figure 49: Electricity generation by sources and electricity demand: proposed version.

It can be noticed in Figure 49 that slight variations are present between the production and the demand, indicating that the system is not perfectly balanced. The small gaps around 2015 and 2045 represent instances of the curtailment penalty of the intermittent electricity production.

In comparison, Figure 50 shows a significant mismatch between electricity production and demand in the former MEDEAS version. The observed gap can be attributed to the RES electricity production definition of Equation 25 reminded here for the sake of simplicity.

$$\texttt{FE\_real\_tot\_generation\_res\_elec} = \frac{\texttt{potential\_tot\_generation\_res\_elec}}{1 + \texttt{res\_elec\_tot\_overcapacity}} \cdot \texttt{shortage\_bioe\_for\_elec}$$

In Equation 25 representing the former implementation, the overcapacity factor is always equal to one, resulting in the real RES generation being halved by the fractional term. Conversely, in the new implementation using Equation 28, the real generation is only penalized by a relatively small percentage of the nominal RES power installed, since it is defined by the curtailment output computed by the surrogate model. This observation was verified: the variation between electrical demand and production corresponds to the amount of electricity produced by RES, meaning that if the RES generation was doubled it would perfectly match, thus resulting in a balanced system.

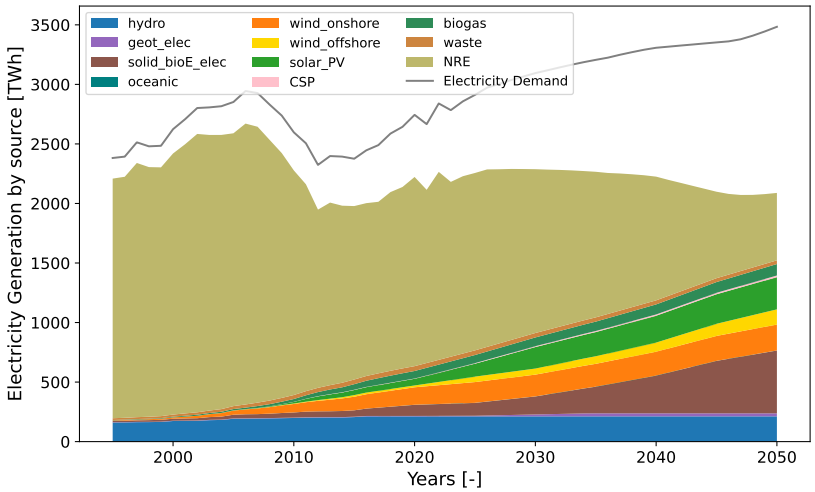


Figure 50: Electricity generation by sources and electricity demand: former version.

Regarding both Figures 49 and 50, the initial version of MEDEAS seems too conservative for the electrical power generation from RES regarding the installed capacities. With the same power installed, the new version predicts almost twice the RES electricity generated.

7.3.4 Investments

Both models evaluate the investments required for the energy transition but in a different way. This subsection details the assumptions and the computed assessments for both simulations. The cumulated investments computed by the former and proposed versions are illustrated respectively in Figure 51 and in Figure 52.

The MEDEAS framework initially identifies two main categories of costs associated with renewable energy systems: the grid reinforcement expenses and the direct costs of RES deployment. The first grid-related costs arise from the growing share of RES in the energy mix and encompass several measures: addressing grid congestion, connecting isolated RES generation units to the grid, and ensuring grid stability and reliability in response to fluctuations in renewable energy production. In the original MEDEAS model, these expenditures are classified as balancing costs and grid reinforcement costs.

For grid reinforcement costs, the original MEDEAS model assumes a fixed cost of 300 $\$/\mathrm{kW}_{\mathrm{wind}}$, based on a 2012 study by Mills *et al.* [50]. This estimate uses wind energy as a representative proxy for all RES and accounts solely for transmission costs. In contrast, the proposed MEDEAS version integrates external data from the PyPSA-EUR model, which includes additional investments required for both AC and DC transmission lines as well as the distribution network.

The proposed version of MEDEAS does not implement a new assessment methodology of balancing costs, but can obviously use same assumption as the initial version which is based on a 2011 study by Holttinen *et al* [51].

Regarding the direct costs of RES deployment, the original and updated MEDEAS models use different approaches once again. The initial version relies on external input data specified in the file `energy.xlsx`, which is derived from a 2011 study by Teske *et al.* [52], to estimate overnight investment costs. Meanwhile the updated implementation, incorporates once more investment data from PyPSA-EUR for solar, wind (both onshore and offshore), and hydro RES while retaining the earlier cost estimates for other energy sources.

Subsequently, the new version assesses storage-related investments considering both the storage capacity computed by MEDEAS and PYPSA-EUR prices.

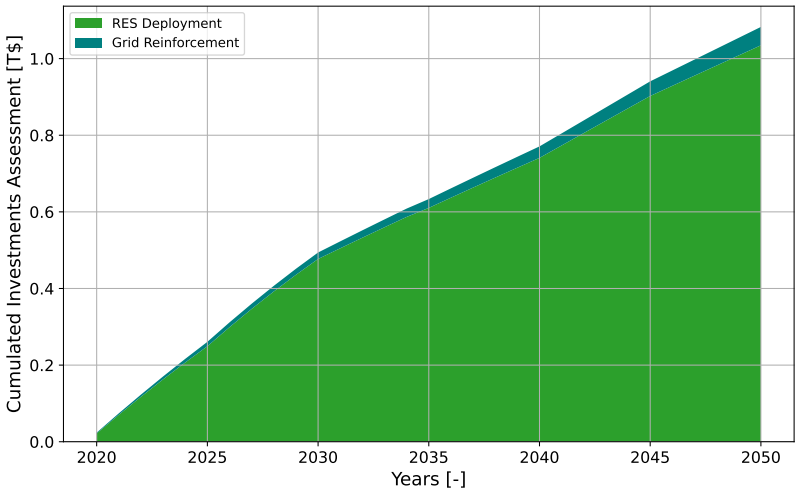


Figure 51: Energy transition investments assessment: former version.

It can be observed that the two methodologies illustrated in Figures 51 and 52 provide significantly different investment magnitudes. This discrepancy is expected, as the proposed version incorporates additional factors not considered in the original approach. Specifically, the updated methodology accounts the following additional costs: storage investments and distribution network expenses. In this way, the new approach is considered more comprehensive and provides a more conservative estimate of the resources required for the energy transition.

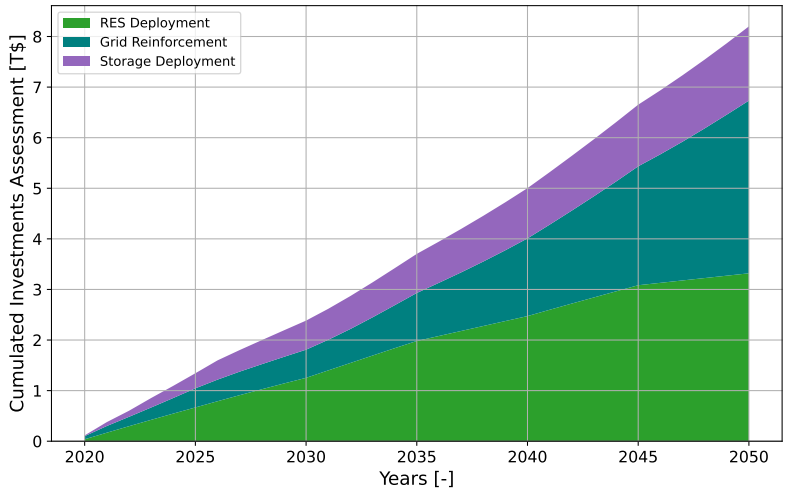


Figure 52: Energy transition investments assessment: new version.

The assessments computed by the proposed version are more consistent with the predictions given by the International Energy Agency (IEA) [53]. The following Table 15 compares the investments (for the year 2024) assessed by both IEA and the proposed MEDEAS version, in the context of a European NZP policy.

Sectors	IEA data [B$]	New Model assessment [B$]
LEE Production	121	123
Grid & Storage	81	54

Table 15: Comparison between proposed model assessment on energy transition investment and IEA data from yearly energy investment report [53]. The values presented are for the year 2024.

7.3.5 Additional Variables

Several variables are absent in the earlier MEDEAS version but are computed in the proposed model. Since these new variables have no equivalent in the former model, they are not compared here. Examples include load shedding output from the surrogate model and the rNTC feature.

7.4 Summary

The following Table 16 summarizes the main differences between the initial PyMEDEAS 2 and the MEDEAS version proposed in this master's thesis.

	Former Version	**Proposed Improvements**
Curtailment	Constant variable, either fixed as an external parameter by the user or defined by the model leading to a reduction of RES generation.	Dynamic endogenous variable which depends on the system configuration, and predicted by the surrogate model.
Load Shedding	Not supported by the model.	Dynamic endogenous variable which depends on the system configuration, and predicted by the surrogate model.
$rNTC$	Not supported by the model.	Variable derived from external data from PyPSA-EUR [21], and might be endogenously increased if the feedback mechanism is activated.
Direct RES costs	Based on a 2012 study [50] assessing the overnight capital costs.	Computed using either the same assumption as former model (Geothermal, Biomass, Oceanic, CSP) or the external data prices representing capital and operational costs (fluctuating over the time horizon) from PyPSA-EUR [21] combined with the installed RES power from the model.
Grid Reinforcement costs	Based on a 2011 study [52] and using wind power installation as a proxy for the transmission network costs.	Computed using external data prices (fluctuating over the time horizon) from PyPSA-EUR [21] considering both transmission and distribution expenses.
Storage investments	Not supported by the model.	Computed using external data prices (fluctuating over the time horizon) from PyPSA-EUR [21] but subject to strong assumptions (Section 4.3.2).

Table 16: Comparison between initial version of PyMEDEAS v2 and proposed version developed in this report.

8 Case Study: Energy Scenarios Comparison

8.1 Overview

This section precedes the conclusion of the report and aims to conduct a final analysis of the proposed MEDEAS version in a practical context. While numerous studies could have been conducted, this report focuses on the creation of three scenarios, each outlining a possible societal response to the pressing challenges humanity currently faces regarding the energy transition. These scenarios, defined in details in next Section 8.2, are named Fossil Fuel Fostering (FFF), Business As Usual (BAU), and Optimal Transition (OT).

This case study provides a practical perspective on the relevance of the work presented in this master's thesis, highlighting its potential contributions to the academic energy field through MEDEAS.

8.2 Description of the scenarios

8.2.1 Fossil Fuel Fostering (FFF)

The Fossil Fuel Fostering (FFF) scenario represents a theoretical pathway that diverges from current trends, increasing reliance on fossil fuels while slightly reducing the RES integration. This approach prioritizes traditional energy sources, while underscoring the environmental and resource depletion concerns associated with such former policies. In this work, the FFF scenario serves as a worst-case reference, used to analyze the consequences of persisting with carbon-intensive energy strategies. This scenario follows the path of SSP5 from SSPs [39].

8.2.2 Business As Usual (BAU)

The Business as Usual (BAU) scenario assumes the continuation of existing energy policies and trends without the introduction of significant new measures to improve the current trajectory. This scenario provides a critical baseline for evaluating the potential consequences of maintaining the status quo, and assesses if current trends should be revised or not. Within this master's thesis, the BAU scenario represents a middle-ground case where the energy transition begins but does not reach its full potential. This scenario follows the path of SSP2 from SSPs [39].

8.2.3 Optimal Transition (OT)

The Optimal Transition (OT) scenario represents an ambitious shift towards a renewable energy-dominated system, supported by technological innovation and sustainable energy practices. This scenario provides insights on bold climate actions, and high sustainable targets for the energy mix. In this work, the OT scenario serves as the more ambitious reference, promoting a strong progressive energy transition. This scenario follows the path of SSP1 from SSPs [39] and baseline scenario, with new predictions for the RES installations.

8.3 Implementation of the scenarios

8.3.1 RES Power installation

As mentioned in Section 2.3, the model prioritizes RES electricity production. Defining the scenarios thus involves adjusting the power installation policies as one of the primary parameters. In order to

model the BAU scenario - which simply requires the continuation of current trends - using historical data previously implemented in MEDEAS, linear regression equations for each RES type are computed using *Excel*. These equations are summarized in Table 17[9]. The historical data available in MEDEAS ranges from 1995 to 2016.

RES Variable	Linear Regression Equation	CoD R^2[-]
Hydroelectricity	$y = 0.0018x + 0.1111$	0.966
Geothermal	$y = 10^{-5}x + 0.0005$	0.9436
Biomass	$y = 0.0009x + 0.0018$	0.9817
Oceanic	$y = 10^{-6}x + 0.0002$	0.9193
Wind Onshore	$y = 0.0067x - 0.0224$	0.9484
Wind Offshore	$y = 0.0018x - 0.0016$	0.9891
Solar	$y = 0.0119x - 0.0109$	0.9704
CSP	$y = 0.0003x - 0.0006$	0.9075
PHS	$y = 0.0001x + 0.0216$	0.9454

Table 17: List of linear regression equations describing the RES power installations based on historical data.

It can be observed that each Coefficient of Determination (R^2) - which measures how well the regression equation models the data - exceeds 0.9 [-] for each equation, indicating a generally high level of predictive accuracy despite the linearity of the relation.

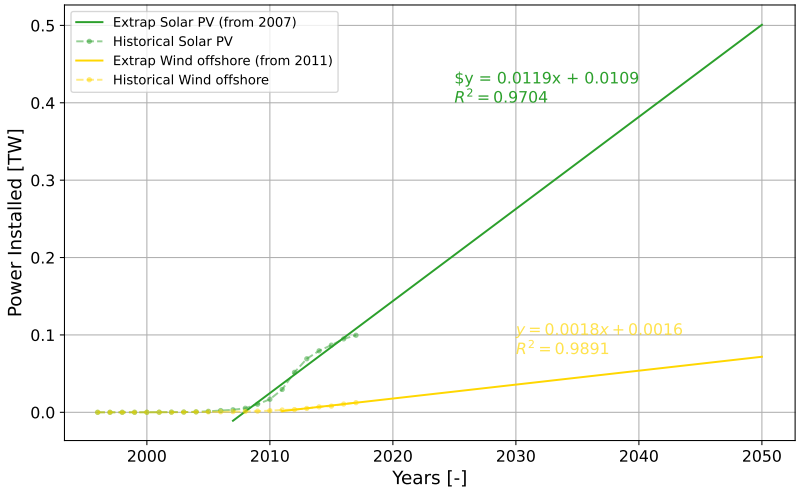


Figure 53: Linear regression examples for wind offshore and solar PV power installations.

Some variables, highlighted in blue in Table 17, are linearly approximated using data from only recent years instead of the entire regression range. Specifically, to accurately capture the recent growth in offshore wind, oceanic, and solar installations, the linear regression is performed using only the data from respectively the last six, nine, and eleven known years, rather than the entire historical dataset

[9]In this section, PHS will be considered alongside RES but is the well-established storage technology.

(example for offshore wind and solar PV illustrated in Figure 53). Additionally, since CSP is a recent technology with non-negligible installations starting from 2006 in the MEDEAS historical dataset, the linear regression is conducted using only the data from 2006.

The predictions for the BAU scenario are approximated using the aforementioned linear regressions to maintain a moderate RES integration, in contrast to the OT scenario, where polynomial regressions are preferred. This choice of regression method results in a significant difference in predictions, as illustrated by the example of wind onshore installations shown in Figure 54. However, some RES power installations are growing linearly since 1995 and are thus still approximated by the associated equation from Table 17. This is the case for: Hydroelectricity, Geothermal, Biomass, and PHS.

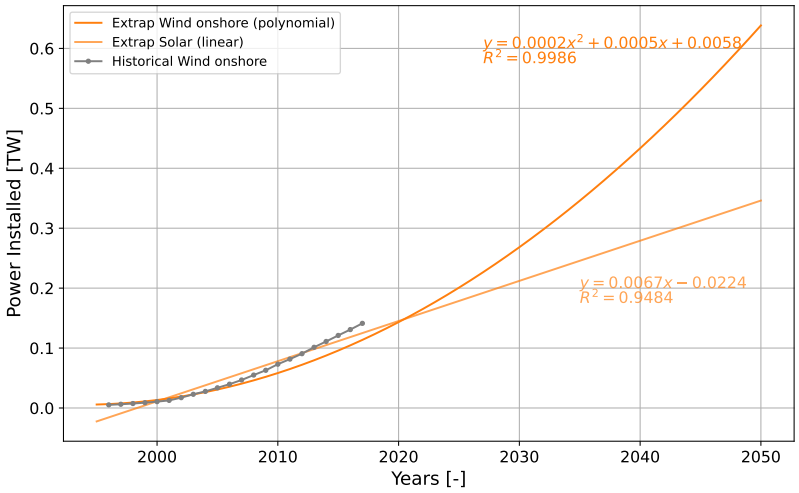


Figure 54: Onshore Wind example of RES power installations extrapolation, from historical data.

Table 18 lists the polynomial regression equations used in the case of the OT scenario.

RES Variable	**Polynomial Regression Equation**	**CoD R^2[-]**
Oceanic	$y = 4 \times 10^{-9}x^3 - 10^{-7}x^2 + 10^{-6}x + 0.0002$	0.9727
Wind Onshore	$y = 0,0002x^2 + 0,0005x + 0,0058$	0.9986
Wind Offshore	$y = 5 \times 10^{-5}x^2 + 0.0007x + 0.0017$	0.8358
Solar	$y = 0.0005x^2 - 0.0061x + 0.0138$	0.962
CSP	$y = 10^{-5}x^2 - 0.0002x + 0.0004$	0.913

Table 18: List of polynomial regression equations describing the RES power installations based on historical data.

A particular attention regarding both wind relations is necessary since polynomial regressions might result in excessive wind power in the simulation, reminding the case previously discussed in Section 5.3. To mitigate this pathological behavior, the power installation curves are slightly smoothed, thereby avoiding the upper bounds by limiting the rise of the wind installation. This adjustment is the reason why the CoD related to wind offshore in Table 18 is lower than the linear regression approximation.

In the FFF scenario, RES power installations are assumed to slightly decline, with a decrease assumed at 1% per year. The resulting power installation as well as the electricity production by source is given in Section 8.4.

8.3.2 Bio fuels annual growth rates

The assumptions listed in Table 19 have been made regarding the annual growth rates of bio energy.

	FFF	BAU	OT
Annual growth bio fuels 2nd generation [%]	0.0	4.0	7.43
Annual growth bio fuels 3rd generation [%]	0.0	4.0	7.43
Annual growth bio gas [%]	0.0	1.0	2.0

Table 19: Bio energy annual growth rates for FFF, BAU, and OT scenarios.

The assumptions in Table 19 outline different growth rates for bio energy across the FFF, BAU, and OT scenarios, highlighting varying levels of commitment to sustainable energy. The FFF scenario assumes no growth in bio energy, reflecting reliance on fossil fuels. In contrast, BAU presents moderate growth, indicating limited progress under current trends, while OT emphasizes significant expansion, demonstrating an active push towards renewable energy and sustainability.

8.3.3 Heat transition

The assumptions listed in Table 20 have been made regarding the heat commodity transition through solar panels and geothermal energy.

	FFF	BAU	OT
Annual growth solar energy for heat [%]	0.0	7.0	14.15
Annual growth geothermal for heat [%]	0.0	1.0	3.87

Table 20: Heat transition policies for FFF, BAU, and OT scenarios.

The FFF scenario assumes no growth, signifying a lack of progress in new renewable heating technologies for the benefit of traditional technologies. BAU suggests moderate increases, reflecting limited advancements under current policies. Meanwhile, OT projects substantial growth, indicating a strong commitment to transitioning towards sustainable heat solutions.

8.3.4 Oil phase-out

The assumptions listed in Table 21 have been made regarding the oil phase-out from electricity and heat.

	FFF	BAU	OT
Oil phase-out for electricity	No	Yes	Yes
Oil phase-out for heat	No	Yes	Yes

Table 21: Oil phase-out policies for FFF, BAU, and OT scenarios.

The FFF scenario retains oil usage, reflecting no commitment to reducing dependence on fossil fuels. In contrast, both BAU and OT scenarios implement a phase-out policy, indicating the wish to shift towards cleaner energy sources.

8.3.5 Nuclear policies

The assumptions listed in Table 22 have been made regarding the nuclear policies.

	FFF	BAU	OT
Nuclear Policy	Phase-out	No more nuclear installed, current capacities depreciated	Nuclear promoting but RES sufficient

Table 22: Nuclear policies for FFF, BAU, and OT scenarios.

In the FFF scenario, nuclear energy is phased out, reflecting a reliance on fossil fuels only. The BAU scenario naturally maintains current European nuclear policies, with no new installations and gradual depreciation of existing plants. In contrast, the OT scenario is implemented using a growth in nuclear energy policy, highlighting a commitment to expanding low-carbon energy sources along with the RES fostering. However, since the RES levels of power installed are already covering the demand, the installed nuclear power plants depreciate since the model prioritizes RES.

8.3.6 Sectoral final energy intensities

Regarding the evolution of the sectoral final energy intensities, no modifications have been made between the three scenarios, and the default method is used. The final energy intensities are set according to sectoral target achievements in a target year, which in this case is 2050. Despite the differences in the strategies of the energy transitions across the scenarios, the evolution of energy intensity targets is assumed to be the same in all cases. This hypothesis ensures a baseline level of energy efficiency, reflecting international commitments to improve energy infrastructures and remaining in an economic growth paradigm.

8.4 Results and analysis

This section gathers the results from the three previously discussed scenarios simulations, comparing their outcomes with an emphasis on the new outputs introduced in this master's thesis proposed implementation. The first subsections illustrate various simulation outcomes, such as the features and targets from the surrogate model, and investment assessments, but without incorporating any feedback mechanisms. In contrast, Section 8.5 proposes a brief discussion on the implementation of societal feedback for the BAU scenario.

8.4.1 Features of the surrogate model

This subsection presents, for each aforementioned scenarios, the evolution of four of the six features required by the surrogate model, namely: share of wind, solar PV, storage, and the capacity ratio. Both share of flexibility and $rNTC$ features are purposely not shown here since they are defined using assumptions discussed in Section 4.2.2 and thus exactly the same for all scenarios.

Figure 55 illustrates the different evolutions induced by the various implementations discussed in Section 8.3.1, clearly highlighting the higher RES integration rate in the OT scenario.

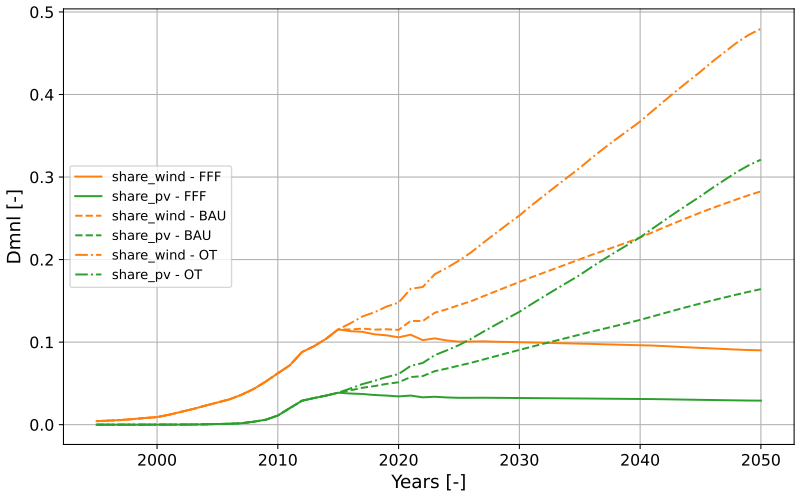


Figure 55: Evolution of wind and solar inputs of the surrogate model, comparison between the three scenarios (FFF, BAU, OT).

The share of storage curve illustrated in Figure 56 follows the same evolution than the previous two features depicted in Figure 55 while remaining moderately present in all simulation.

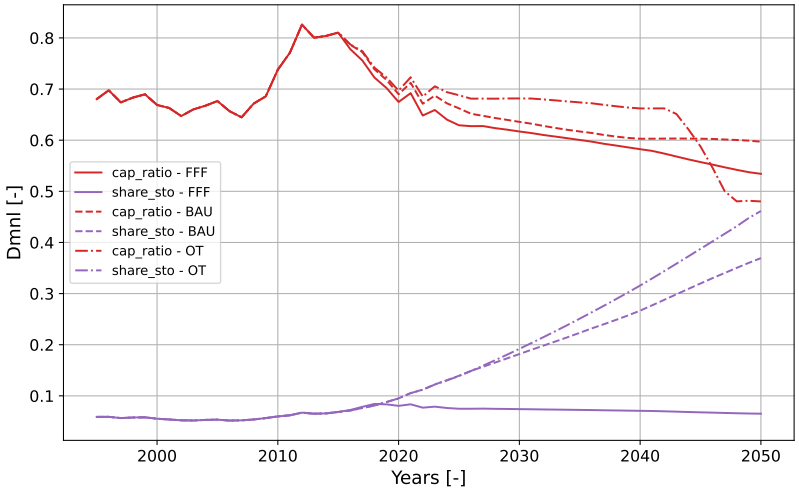


Figure 56: Evolution of storage share and capacity ratio inputs of the surrogate model, comparison between the three scenarios (FFF, BAU, OT).

It can be observed in Figure 56 that the capacity ratio suddenly drops around 2040 in the OT scenario. This decline can be explained by two phenomena. First, the higher increase in electricity demand in this scenario induced by the society electrification. Secondly, the capacity ratio does not consider both solar and wind power production in its definition. As both technologies become the dominant sources of electricity production, a decrease of the other sources is obtained thus leading to a reduction of the capacity ratio.

8.4.2 Targets of the surrogate model

This subsection presents, for each aforementioned scenarios, the evolution of both outcomes from the surrogate model, namely the curtailment (Figure 57) and the load shedding (Figure 58).

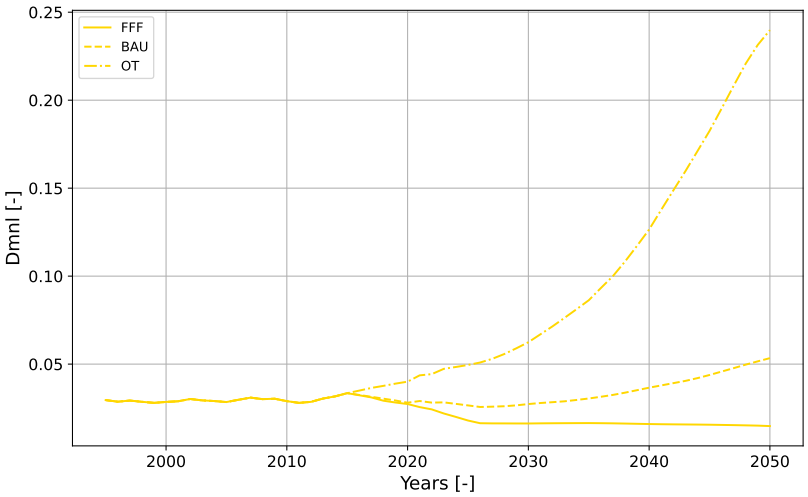


Figure 57: Prediction of the level of curtailment, comparison between the three scenarios (FFF, BAU, OT).

First, Figure 57 illustrates identical trends for each scenario before 2015, reflecting the system's curtailment when the electricity mixes are similar and the scenarios have no impact on the system. As the simulation progresses, the curves begin to diverge, reflecting expected evolutions. For the FFF scenario, a constant curtailment of 1.5% is observed, as the RES integration does not significantly impact the system, which continues to rely on fossil fuels and thus remains highly flexible. In the BAU scenario, moderate curtailment levels are reaching 5% by 2050, with a slight increase due to the growing integration of RES and the intermittency associated with these energy sources. Finally, the OT scenario, which assumes high RES penetration, exhibits a significant increase in curtailment, exceeding 20% by 2050. This high level of curtailment suggests an over-sizing of the grid, indicating that investments were not optimally allocated. The lack of flexibility occurring in the system can be counter balanced by investments in storage or transfer capacities, as it is discuss in 8.5.1.

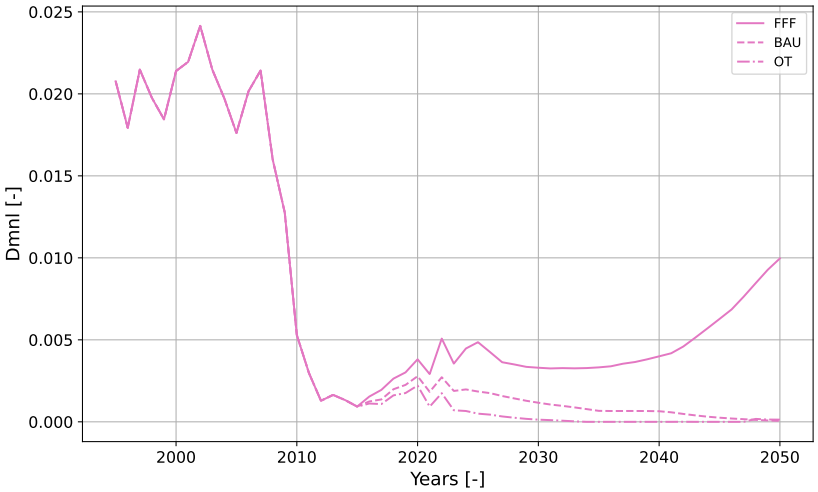


Figure 58: Prediction of the level of load shedding, comparison between the three scenarios (FFF, BAU, OT).

Regarding the load shedding outputs illustrated in Figure 58, it can be seen that it follows an opposite trend compared to the curtailment. Indeed, on the one hand, the high share of penetration occurring in bot OT and BAU scenarios induce very low levels of load shedding suggesting that the demand is always fulfilled. On the other hand, a rise of load shedding is observed in the FFF scenario suggesting that the load is harder to satisfied in this system. This is mainly induced by the depreciation policy of the RES installations which let no other choice than the reliance on fossil fuels. The latter being subject to resource depletion, an increase in load shedding is naturally observed.

8.4.3 Investments assessments

This subsection provides an analysis of the investments required under the different aforementioned scenarios. The grid reinforcement costs are deliberately excluded from this discussion, as they are derived from PyPSA-EUR as a base load and, therefore, are identical across all the considered scenarios if the feedback mechanism is not activated. These following cumulated investments are different from zero in 1995, considering the already installed capacities of the associated technology.

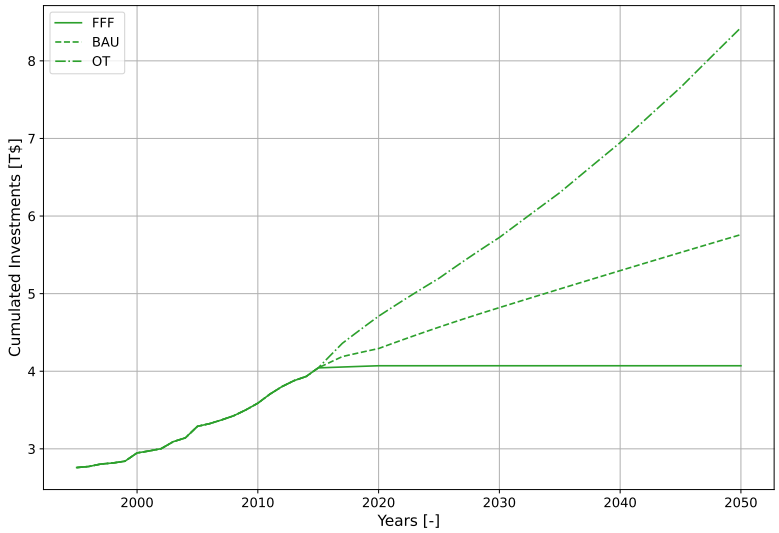


Figure 59: Investments assessment of the direct costs related to the integration of the RES, comparison between the three scenarios (FFF, BAU, OT).

Both Figures 59 and 60 exhibit similar trends in their evolution. As expected, the OT scenario requires significantly higher investments compared to the BAU scenario, reflecting its ambitious focus on renewable energy and storage capacity expansion. Conversely, the FFF scenario demonstrates constant cumulated investment levels, indicating a policy where no additional capacities are constructed. It is expected since this scenario implements a gradual depreciation of both renewable energy and storage infrastructures.

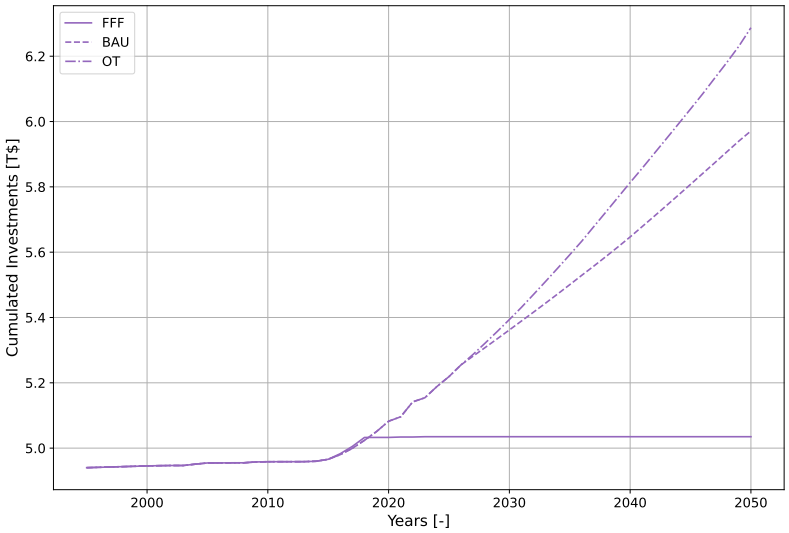


Figure 60: Investments assessment for the integration of storage, comparison between the three scenarios (FFF, BAU, OT).

8.4.4 Socio-economic and environmental aspects

Previous sections have focused on the variables directly linked to the new implementation of MEDEAS. However, these scenarios differ in numerous other aspects, which will be briefly outlined in this section.

CO_2 Emissions

The CO_2 emissions produced through electricity generation vary significantly across the considered scenarios. As shown in Figure 61, only the OT scenario achieves net zero carbon emissions, reaching neutrality around 2041, well before 2050. In contrast, both the BAU and FFF scenarios continue to emit carbon, with the FFF scenario emitting twice as much as the BAU scenario, exceeding $0.8\,\mathrm{Gt\,CO_2\,per^{-1}}$.

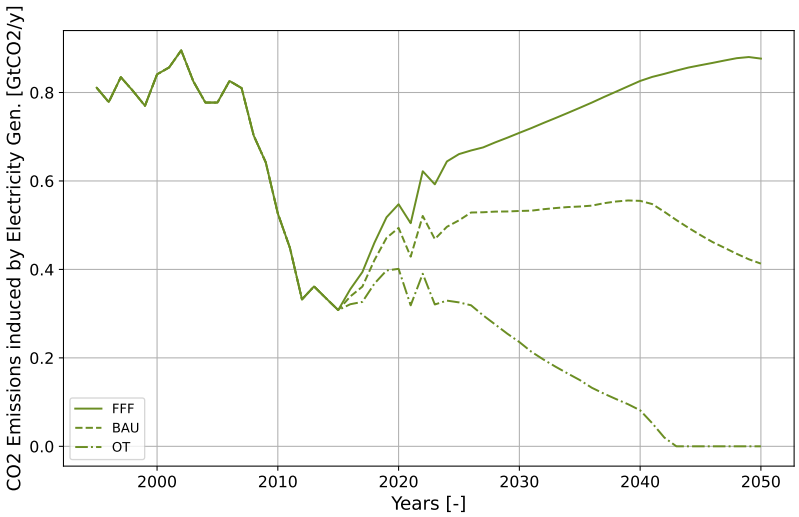


Figure 61: CO_2 emissions induced by the production of electricity, comparison between the three scenarios (FFF, BAU, OT).

Gross Domestic Product per capita

As previously mentioned, the FFF scenario experiences a reduction in satisfied electrical demand, as it relies on traditional technologies despite resource depletion challenges. Economically, this lack of innovation leads to a stagnating economy, as illustrated in Figure 62. In contrast, the other scenarios facilitate steady growth in GDP per capita by enabling higher sectoral demands and increased energy production through sustainable practices and renewable energy integration. The stagnation in GDP observed for the FFF scenario may induce several adverse effects, such as global debts increase, growth of unemployment, poverty and heightened inequalities, etc. [54]

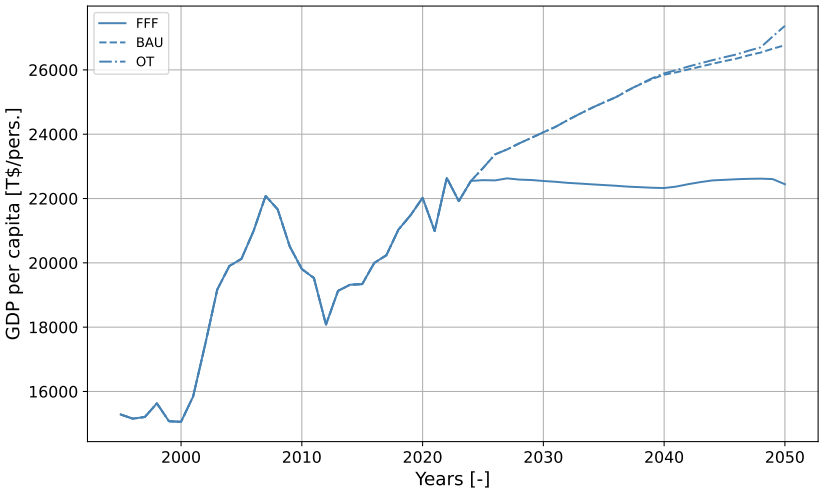


Figure 62: European Gross Domestic Product per capita, comparison between the three scenarios (FFF, BAU, OT).

8.5 Impact of society feedback mechanisms

In order to accurately depict society inertia without any external perturbation in previous analysis, the feedback mechanisms are not implemented in previous section. However, further results can be dressed with this process implemented. The dimensionless k-coefficients presented in Table 23 are set for the OT scenario:

8.5.1 Reduction of the curtailment in the OT Scenario

As it is mentioned earlier, the OT scenario invest too much in the RES integration without considering the flexibility assets such as the Net Transfer Capacity and storage. To overcome this issue, a curtailment reduction process is implemented from 2020.

P	I	D
0.01	0	0

Table 23: Feedback process parametrization for the OT scenario (dimensionless).

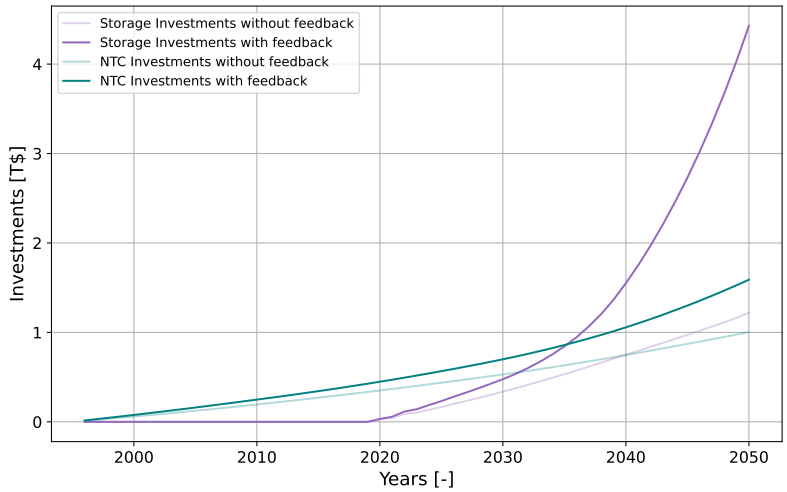


Figure 63: Grid and storage related investments with or without the feedback process, OT scenario.

As shown in both Figure 63 and Figure 64, an increase in grid-related and storage investments naturally leads to a reduction in curtailment, as electricity can flow more efficiently through the grid and additional production can be stored. However, a cumulative investment difference of 3000 B$ for storage and 500 B$ from 2020 results in only a 5% reduction in curtailment. A more detailed cost-benefit analysis is necessary to assess whether this level of investment is justified by this curtailment reduction.

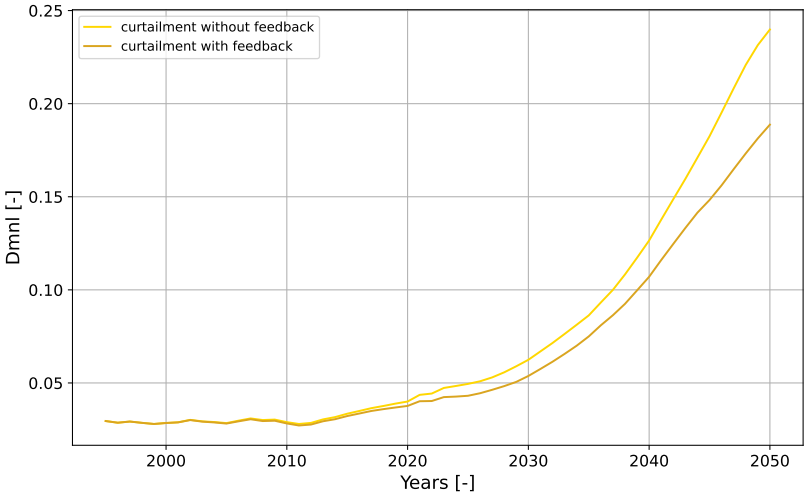


Figure 64: Prediction level of curtailment comparison with or without the feedback process, OT scenario.

8.6 Case study: conclusions

After analyzing these three scenarios, several conclusions can be drawn:

1. The surrogate model successfully captures the differences between the scenarios by integrating the simulation context through four key inputs. Both outputs are relevant and seem to accurately assess the expected electrical network situation.
2. The scenario with the largest RES share (OT) in the electricity mix naturally results in higher curtailment, reaching 23% by 2050. This level of curtailment is considerably high and suggests an over investment in RES rather than flexibility assets. The BAU scenario shows moderate curtailment, at only a few percent, while the FFF scenario experiences negligible curtailment, reflecting its reliance on fossil fuels and a more flexible electricity generation assets.
3. The FFF scenario observes an increase in unsatisfied electrical demand, leading to a non-growing economy, which might lead to significant damages. In contrast, both RES-based scenarios enable an increase in sectoral demand, particularly with the electrification of society, which drives higher energy consumption, but also leads to higher curtailment levels.
4. Only the more ambitious scenarios (OT) allows a carbon neutral electricity production, from 2041. The BAU scenario reaches $0.4\,\mathrm{Gt\,CO_2\,year^{-1}}$ by 2050, which is half of the emissions of the FFF scenario.
5. In the case of the OT scenario, a further increase in cumulative grid-flexibility investments results in a moderate reduction in curtailment of 5% in 2050, approximately 116 B$/year.

9 Conclusion and further developments

This master's thesis is concluded with a summary of key take-home messages of the important results developed throughout this work. In addition, acknowledging that this work does not claim to provide a perfect or fully comprehensive enhancement of MEDEAS, a list of proposed improvements for future works is presented.

9.1 Take-home messages

This master's thesis incorporates the previously developed surrogate model within the MEDEAS ecosystem by establishing a correlation between the inputs of the SM and the variables defined within the MEDEAS context. To achieve this linkage between both models, key assumptions were made regarding two of the six features: the share of flexibility and the Net Transfer Capacity ratio. The share of flexibility is approximated using a linear equation to represent the technological advancements in future power plants, while the NTC ratio is derived from exogenous data provided by the external model PyPSA-EUR. Despite these two primary assumptions, both proposed definitions are considered valid and enable the surrogate model to provide realistic insights regarding the expected future trends. Concerning the outputs of the surrogate model, one of them, curtailment, substitutes the predefined constant assumptions in the initial version of MEDEAS, thereby enhancing the representation of this physical phenomenon. The second output, load shedding, was not supported in its current state by the initial version and thus introduces a new endogenous evaluation of the unsatisfied electrical demand. It is worth emphasizing that two machine learning models - Multilayer Perceptron and Random Forest - were compared to determine the most suitable algorithm. The MLP algorithm was ultimately considered more suitable for this context in regards of the better predictions on unseen data.

Once the surrogate model was implemented, several new features were introduced into the proposed version. New assessments related to the required investments for the energy transition were integrated. Indeed, the model now evaluates the investments for integrating renewable energy sources (considering both capital and operational expenditures), the costs for grid reinforcement (encompassing both transmission and distribution networks), and the investments in new storage installations.

Moreover, two new processes have been integrated into the proposed MEDEAS version to simulate and assess an ideal investment reaction aimed at further reducing curtailment and load shedding. These mechanisms are user-configurable, allowing the user to tune the feedback intensity and select the activation year for the processes. The reduction of the load shedding mechanism is either not use or at very low level since - by increasing RES integration - it tends to increase too much the curtailment. However, since load shedding is generally at low levels (0 to 2%) in most of the scenarios, the curtailment reduction process provides a good insight on future grid-flexibility increase policies.

Finally, to evaluate the proposed version practically, the model was, on the one hand, compared to the initial version through a comparative analysis, and on the other hand, applied in a practical case study. First, the comparative study highlighted the modifications introduced in the proposed version, demonstrating that the updated model seems more comprehensive than the previous version by incorporating the aforementioned relevant features. Secondly, the practical case study illustrated the model's utility by comparing three different scenarios. It showed that the updated model is effective in exploring new scenarios within MEDEAS, offering new physical insights.

To conclude, the proposed work successfully integrates several new features into MEDEAS, providing modest enhancements to the model's energy module. However, several aspects of this study require further improvement, which are outlined in the following subsection.

9.2 Further proposed improvements

The list of known limitations and proposed further improvement is outlined below.

1. The definition of the share of flexibility surrogate model's input can be improved, through a literature review of the future trends and expectations for all power units characteristics. In particular regarding the following parameters evolutions: minimum nominal capacity, minimum up time, and ramping up rate.

2. The evolution of the Net Transfer Capacity ratio is currently implemented in the form of a constant base load, regardless of the scenario. Ideally, the scenario should impact this variable or at least the evolution of the grid's transfer capacity should be user-configurable. For example, several different options - from low capacity to highly efficient network - should be proposed.

3. The reduction feedback mechanisms implemented in this work only simulate a perfect society reaction but can be improved. Currently, these processes consider an instantaneous reaction of the society, but a construction-decision delay might be introduced or at least configurable by the user. Additionally, different approaches using more advanced control tools might be interesting to study. Moreover, instead of fixed investment dispatch derived from PyPSA-EUR, these shares should also depend on the scenario or at least be user-configurable to assess the impact of different investments policies. Finally, if the feedback is activated before 2020, the new capacities installed through the process are computed based on prices linearly approximated from PyPSA-EUR. However, this is not representative since the historical prices of RES technologies were well higher, leading to an undervaluation of necessary investments.

4. A literature review should also be made at a world scale regarding the RES installation potentials. Currently, at a world scale, the solar PV potential is limited to 10.6 TW based on land used while the European level of the same potential is set at 13 TW. Similar results are obtained for the Biomass and the CSP potentials.

5. A new large-scale storage technology might be introduced in MEDEAS in addition of PHS and EVs. Interesting options are new storage capacity variables representing either future grid-scale batteries or hydrogen-based technologies.

6. The capacity factors used for assessing the installed capacity of fossil fuels power plants is currently approximated using a fixed capacity factor from historical values. A future improvements might involve to integrate these CF as it is already the case for nuclear power plants. This can either be done using external predictions or by training a new ML model on dispa-set outputs.

A Added variables in MEDEAS ecosystem

Tables A.1-A.5 contain the new variables developed in this work, with a brief description. They are grouped by the file where it is defined and classified according to the implementation order.

Variables defined in `features.py`:

Name	Description	Unit
`peak_load`	Definition of the peak load variable, required by some surrogate model features and defined by Equation 16.	TW
`cap_ratio`	Feature of the surrogate model describing the capacity ratio, defined in Equation 17.	[-]
`share_flex`	Feature of the surrogate model describing the share of flexibility, defined in Equation 18.	[-]
`share_sto`	Feature of the surrogate model describing the share of storage, defined in Equation 21.	[-]
`share_wind`	Feature of the surrogate model describing the share of wind, defined in Equation 22.	[-]
`share_pv`	Feature of the surrogate model describing the share of solar PV, defined in Equation 23.	[-]
`rNTC`	Feature of the surrogate model describing the $rNTC$, defined in Equation 24.	[-]
`add_rNTC_feedback`	Instantaneous increase in $rNTC$ computed at each time step.	[-]
`cumulated_add_rNTC_feedback`	Cumulated increase in $rNTC$, in order to stores the increase through the simulation (Equation 32).	[-]

Table A.1: List of added variables in `features.py` within the proposed version of MEDEAS, names and descriptions.

Variables defined in `targets.py`:

Name	Description	Unit
`curtailment`	Target of the MLP surrogate model, defined in Equation 12.	[-]
`curtailment_delayed`	The curtailment target but delayed of one time step. The initial value is set to 0.	[-]
`energy_curtailed_twh`	It represents the curtailment target but expressed in TWh.	TW h
`load_shedding`	Target of the MLP surrogate model, defined in Equation 13.	[-]
`load_shedding_delayed`	The load shedding target but delayed of one time step. The initial value is set to 0.	[-]
`load_shed_twh`	It represents the load shedding target but expressed in TWh.	TW h
`RF_curtailment`	Target derived from the RF version of the surrogate model, for comparison purpose only.	[-]
`RF_load_shedding`	Target derived from the RF version of the surrogate model, for comparison purpose only.	[-]

Table A.2: List of added variables in `targets.py` within the proposed version of MEDEAS, names and descriptions.

Variables defined in `investments.py`:

Name	Description	Unit
`new_investments_grid_ls`	New necessary grid investments (at a specific time step) computed by load shedding reduction mechanism discussed in Section 6.3.	T$
`cumulated_new_investments_grid_ls`	Cumulated grid investments computed by load shedding reduction mechanism discussed in Section 6.3.	T$
`new_investments_grid_curt`	New necessary grid investments (at a specific time step) computed by curtailment reduction mechanism discussed in Section 6.3.	T$
`cumulated_new_investments_grid_curt`	Cumulated grid investments computed by curtailment reduction mechanism discussed in Section 6.3.	T$
`investments_shares_ls`	Load shedding investments dispatch between new capacities installed, as discussed in Section 6.3.2.	[-]
`investments_shares_curt`	Curtailment investments dispatch between new capacities installed, as discussed in Section 6.3.2.	[-]

Table A.3: List of added variables in `investments.py` within the proposed version of MEDEAS, names and descriptions.

Name	Description	Unit
sm_new_capacity_res_elec	Additional RES capacities installed (at a specific time step), computed by feedback mechanisms, investments shares, and prices.	TW
cumulated_solar_PV_feedback	Cumulated additional capacity of Solar PV installed by the feedback mechanism.	TW
cumulated_wind_offshore_feedback	Cumulated additional capacity of Wind offshore installed by the feedback mechanism.	TW
cumulated_wind_onshore_feedback	Cumulated additional capacity of Wind onshore installed by the feedback mechanism.	TW
cumulated_hydro_feedback	Cumulated additional capacity of conventional hydroelectricity installed by the feedback mechanism.	TW
sm_new_capacity_ntc	Additional transmission and distribution lines capacities installed (at a specific time step), computed by feedback mechanisms, investments shares, and prices.	TW
cumulated_ntc_feedback	Cumulated additional capacity of transmission and distribution lines, by the feedback mechanism.	TW
sm_new_capacity_storage	Additional storage capacities installed (at a specific time step), computed by feedback mechanisms, investments shares, and prices.	TW
cumulated_storage_feedback	Cumulated additional capacity of storage by the feedback mechanism.	TW
tot_investments_grid_feedback	Total of investments injected by the society feedback mechanism discussed in Section 6.	T$
annualized_investments_res_by_type	Annualized investments of the RES by type. Considering both PyPSA-EUR external data and installed capacities from MEDEAS.	T$
investments_res_by_type	Proposed RES-related investments assessment for each type, either derived from invest_cost_res_elec MEDEAS assumption, or from external model PyPSA-EUR seen in Section 3.4.	T$
tot_investments_res	Sum of the previous variable investments_res_by_type for all RES types investments, per year.	T$
annualized_investments_res	Total annualized investments res.	T$
tot_investments_ntc	Proposed grid-related investments assessment, from external model PyPSA-EUR seen in Section 3.4.	T$
tot_investments_storage	Proposed storage-related investments assessment, from external model PyPSA-EUR seen in Section 3.4.	T$

Table A.4: Table A.3 continued.

Name	Description	Unit
new_storage_installed_capacity	It represents the net new capacity of storage, considering the previous installations and the current constructions.	TW
total_capacity_elec_storage_tw_delayed	Total electrical capacity capacity installed for storage purposes, but delayed in order to compute the net installation (previous variable).	TW
TOT_investments	Sum of RES, Grid, and Storage investments.	T$
activation_year_feedback	Activation year for the feedback mechanism.	year

Table A.5: Table A.3 continued, second part.

References

[1] Hoesung Lee, Katherine Calvin, Dipak Dasgupta, Gerhard Krinner, Aditi Mukherji, Peter Thorne, Christopher Trisos, José Romero, Paulina Aldunce, and Alexander C Ruane. Climate change 2023 synthesis report summary for policymakers. 2024.

[2] United Nations. The Paris Agreement | UNFCCC, 2015. URL: https://unfccc.int/process-and-meetings/the-paris-agreement (Accessed: 1st January 2025).

[3] European Commission. The European Green Deal, July 2021. URL: https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/european-green-deal_en (Accessed: 1st January 2025).

[4] European Commission. European Climate Law, 2021. URL: https://climate.ec.europa.eu/eu-action/european-climate-law_en (Accessed: 1st January 2025).

[5] Donella H. Meadows, Dennis L. Meadows, Jorgen Randers, and William W. Behrens III. *The Limits to Growth*. Universe Books, USA, 1972.

[6] Brian Hayes. World3, the public beta, April 2012. URL: http://bit-player.org/2012/world3-the-public-beta (Accessed: on 1st January 2025).

[7] European Commission. Horizon Europe funding program, December 2024. URL: https://research-and-innovation.ec.europa.eu/funding/funding-opportunities/funding-programmes-and-open-calls/horizon-europe_en (Accessed: 1st January 2025).

[8] R. Cloux. Development of machine learning-based surrogate model in the european power system. URL: https://matheo.uliege.be/handle/2268.2/20874 (Accessed: 15th September 2024), 2024.

[9] F. Straet. Improving the simulation of variable renewable energy in the medeas integrated assessment model. URL: https://matheo.uliege.be/handle/2268.2/18379 (Accessed: 15th September 2024), 2023.

[10] Kavvadias K.; Hidalgo Gonzalez I.; Zucker A.; Quoilin S. *Integrated modelling of future EU power and heat systems: The Dispa-SET 2.2 open-source model*. EUR 29085 EN. Publications Office of the European Union, Luxembourg, 2018.

[11] Ventana Systems, Inc. *Vensim User's Guide*, 2024. URL: https://www.vensim.com/documentation/users_guide.html (Accessed: 28th November 2024).

[12] W Leontief. Input-output economics, 1986.

[13] Roger Samsó; Jordi Solé Ollé. *Pymedeas model user's manual*. CSIC, 2019. URL: https://www.medeas.eu/system/files/documentation/files/D8.11%28D35%29%20Model%20Users%20Manual.pdf (Accessed: 4th January 2024).

[14] Martin-Martinez Eneko; Samsó Roger; Houghton James; Solé Ollé Jordi. Pysd: System dynamics modeling in python. *The Journal of Open Source Software, 2022, vol. 7, num. 78, p. 4329*, 2022.

[15] Werner Zittel. Feasible futures for the common good; energy transition paths in a period of increasing resource scarcities. *Progress Report*, 1, 2012.

[16] Werner Zittel, Jan Zerhusen, Martin Zerta, and Nikolaus Arnold. Fossil and nuclear fuels–the supply outlook. *Berlin: Energy Watch Group*, 2013.

[17] Mark Whiteside and J Marvin Herndon. Disruption of earth's atmospheric flywheel: Hothouse-earth collapse of the biosphere and causation of the sixth great extinction. *European Journal of Applied Sciences–Vol*, 12(1), 2024.

[18] H.-O. Pörtner, R. J. Scholes, A. Arneth, D. K. A. Barnes, M. T. Burrows, S. E. Diamond, C. M. Duarte, W. Kiessling, P. Leadley, S. Managi, P. McElwee, G. Midgley, H. T. Ngo, D. Obura, U. Pascual, M. Sankaran, Y. J. Shin, and A. L. Val. Overcoming the coupled climate and biodiversity crises and their societal impacts. *Science*, 380(6642):eabl4881, 2023.

[19] Antonio Garcia-Olivares and Jordi Solé. End of growth and the structural instability of capitalism—from capitalism to a symbiotic economy. *Futures*, 68:31–43, 2015.

[20] Tokuta Yokohata, Katsumasa Tanaka, Kazuya Nishina, Kiyoshi Takahashi, Seita Emori, Masashi Kiguchi, Yoshihiko Iseri, Yasushi Honda, Masashi Okada, Yoshimitsu Masaki, et al. Visualizing the interconnections among climate risks. *Earth's Future*, 7(2):85–100, 2019.

[21] Tom Horsch Jonas; Hofmann, Fabian; Schlachtberger David; Brown. Pypsa-eur: An open optimisation model of the european transmission system. *Energy strategy reviews*, 22:207–215, 2018.

[22] Tareen Muhammad Umair; Quoilin Sylvain; Laterre Antoine; Meyer Sebastien; Thiran Paolo; Hernandez Aurélia. Modeling the impact of energy sufficiency on european integrated energy systems. 2024.

[23] Alireza Soroudi. *Power system optimization modeling in GAMS*, volume 78. Springer, 2017.

[24] Sylvain Quoilin, Wouter Nijs, Ignacio Hidalgo Gonzalez, Andreas Zucker, and Christian Thiel. Evaluation of simplified flexibility evaluation tools using a unit commitment model. In *2015 12th International Conference on the European Energy Market (EEM)*, pages 1–5. IEEE, 2015.

[25] Ray Antonio Rojas Candia, Sergio Luis Balderrama Subieta, Joseph Adhemar Araoz Ramos, Vicente Senosiain Miquélez, Jenny Gabriela Peña Balderrama, Hernan Jaldín Florero, and Sylvain Quoilin. Techno-economic assessment of high variable renewable energy penetration in the bolivian interconnected electric system. *International Journal of Sustainable Energy Planning and Management*, 22, 2019.

[26] Tomi J. Emanuele Taibi; Thomas Nikolakakis; Laura Gutierrez; Carlos Fernandez; Juha Kiviluoma; Simo Rissanen; Lindroos. *Power system flexibility for the energy transition: Part 1, Overview for policy makers*. VTT's Reasearch, nov 2018.

[27] European Commission. model PRIMES - Price-Induced Market Equilibrium System | Modelling Inventory and Knowledge Management System of the European Commission (MIDAS), 2024.

[28] European Commission. EU Reference Scenario 2020. URL: https://energy.ec.europa.eu/data-and-analysis/energy-modelling/eu-reference-scenario-2020_en (Accessed: 4th January).

[29] Anne Sjoerd Brouwer, Machteld Van Den Broek, Ad Seebregts, and André Faaij. Operational flexibility and economics of power plants in future low-carbon power systems. *Applied Energy*, 156:107–128, October 2015.

[30] Oliver Kramer and Oliver Kramer. Scikit-learn. *Machine learning for evolution strategies*, pages 45–53, 2016.

[31] Tom Xiong Boyuan; Fioriti Daniele; Neumann Fabian; Riepin Ilya; Brown. Prebuilt electricity network for pypsa-eur based on openstreetmap data, 2024. URL: https://doi.org/10.5281/zenodo.14144752 (Accessed: 19th December 2024).

[32] Hofmann Fabian; Hampp Johannes; Neumann Fabian; Brown Tom; Hòrsch Jonas. Atlite: a lightweight python package for calculating renewable power potentials and time series. *Journal of Open Source Software*, 6(62):3294, 2021.

[33] U.S. Energy Information Administration and International Energy Statistics. Electric generator capacity factors vary widely across the world, 2015. Available at: https://www.eia.gov/todayinenergy/detail.php?id=22832 (Accessed: 3rd December 2024).

[34] I Capellán-Pérez and C de Castro. Integration of global environmental change threat to human societies in energy–economy–environment models. In *Proceedings of the 12th Conference of the European Society for Ecological Economics, Budapest, Hungary*, pages 20–23, 2017.

[35] Carlos Medina, C Ríos M Ana, and Guadalupe González. Transmission grids to foster high penetration of large-scale variable renewable energy sources–a review of challenges, problems, and solutions. *International Journal of Renewable Energy Research (IJRER)*, 12(1):146–169, 2022.

[36] Samuel H. Williamson. Measuring Worth - GDP Deflator, 2024. URL: https://www.measuringworth.com/datasets/usgdp/result.php (Accessed: 29th December 2024).

[37] European Central Bank (ECB): Exchange rates EUR/USD from 1999 to 2024. URL: https://data.ecb.europa.eu/data/datasets/EXR/EXR.D.USD.EUR.SP00.A (Accessed: 18th December 2024).

[38] Ottmar Edenhofer, Ramón Pichs-Madruga, Youba Sokona, Kristin Seyboth, Susanne Kadner, Timm Zwickel, Patrick Eickemeier, Gerrit Hansen, Steffen Schlömer, Christoph von Stechow, et al. *Renewable energy sources and climate change mitigation: Special report of the intergovernmental panel on climate change*. Cambridge University Press, 2011.

[39] Keywan Riahi, Detlef P Van Vuuren, Elmar Kriegler, Jae Edmonds, Brian C O'neill, Shinichiro Fujimori, Nico Bauer, Katherine Calvin, Rob Dellink, Oliver Fricko, et al. The shared socioeconomic pathways and their energy, land use, and greenhouse gas emissions implications: An overview. *Global environmental change*, 42:153–168, 2017.

[40] Directorate-General for Energy European Commission. An eu strategy to harness the potential of offshore renewable energy for a climate neutral future. *Energy*, 2020. Available at: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM:2020:741:FIN&qid=1605792629666 (Accessed: 15th December 2024).

[41] Michael A Johnson and Mohammad H Moradi. *PID control*. Springer, 2005.

[42] John G Ziegler and Nathaniel B Nichols. Optimum settings for automatic controllers. *Transactions of the American society of mechanical engineers*, 64(8):759–765, 1942.

[43] CORINE Land Cover. European Union's Copernicus Land Monitoring Service information. URL: https://land.copernicus.eu/en/products/corine-land-cover (Accessed: 24th December 2024).

[44] The Natura 2000 protected areas network. URL: https://www.eea.europa.eu/themes/biodiversity/natura-2000/the-natura-2000-protected-areas-network (Accessed: 24th December 2024).

[45] General Bathymetric Chart of the Oceans. Gridded bathymetry data (General Bathymetric Chart of the Oceans). URL: https://www.gebco.net/data_and_products/gridded_bathymetry_data/ (Accessed: 24th December 2024).

[46] David Severin Ryberg; Dilara Gulcin Caglayan; Sabrina Schmitt; Jochen Linßen; Detlef Stolten; Martin Robinius. The future of european onshore wind energy potential: Detailed distribution and simulation of advanced turbine designs. *Energy*, 182:1222–1238, 2019.

[47] Peter Enevoldsen, Finn-Hendrik Permien, Ines Bakhtaoui, Anna-Katharina von Krauland, Mark Z. Jacobson, George Xydis, Benjamin K. Sovacool, Scott V. Valentine, Daniel Luecht, and Gregory Oxley. How much wind power potential does europe have? examining european wind power potential with an enhanced socio-technical atlas. *Energy Policy*, 132:1092–1100, 2019.

[48] Directorate-General for Energy (DG ENER European Commission). Non-paper on complementary economic modelling analysing the impacts of overall renewable energy targets in the context of the renewable energy directive, June 2022. Available at: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52022SC0230 (Accessed: 15th December 2024).

[49] Pantelis Capros, Alessia De Vita, Nikolaos Tasios, Pelopidas Siskos, Maria Kannavou, Apostolos Petropoulos, Stavroula Evangelopoulou, Marilena Zampara, Dimitris Papadopoulos, Ch Nakos, et al. Eu reference scenario 2016-energy, transport and ghg emissions trends to 2050., 2016.

[50] Andrew Mills, Ryan Wiser, and Kevin Porter. The cost of transmission for wind energy in the United States: A review of transmission planning studies. *Renewable and Sustainable Energy Reviews*, 16(1):1–19, January 2012.

[51] Hannele Holttinen, Peter Meibom, Antje Orths, Bernhard Lange, Mark O'Malley, John Olav Tande, Ana Estanqueiro, Emilio Gomez, Lennart Söder, Goran Strbac, et al. Impacts of large amounts of wind power on design and operation of power systems, results of iea collaboration. *Wind Energy*, 14(2):179–192, 2011.

[52] Sven Teske, Thomas Pregger, Sonja Simon, Tobias Naegler, Wina Graus, and Christine Lins. Energy [R]evolution 2010—a sustainable world energy outlook. *Energy Efficiency*, 4(3):409–433, August 2011.

[53] IEA. World Energy Investment 2024, 2024. URL: https://www.iea.org/reports/world-energy-investment-2024 (Accessed: 29th December 2024).

[54] Mikael Malmaeus and Eva Alfredsson. Potential Consequences on the Economy of Low or No Growth - Short and Long Term Perspectives. *Ecological Economics*, 134:57–64, April 2017.