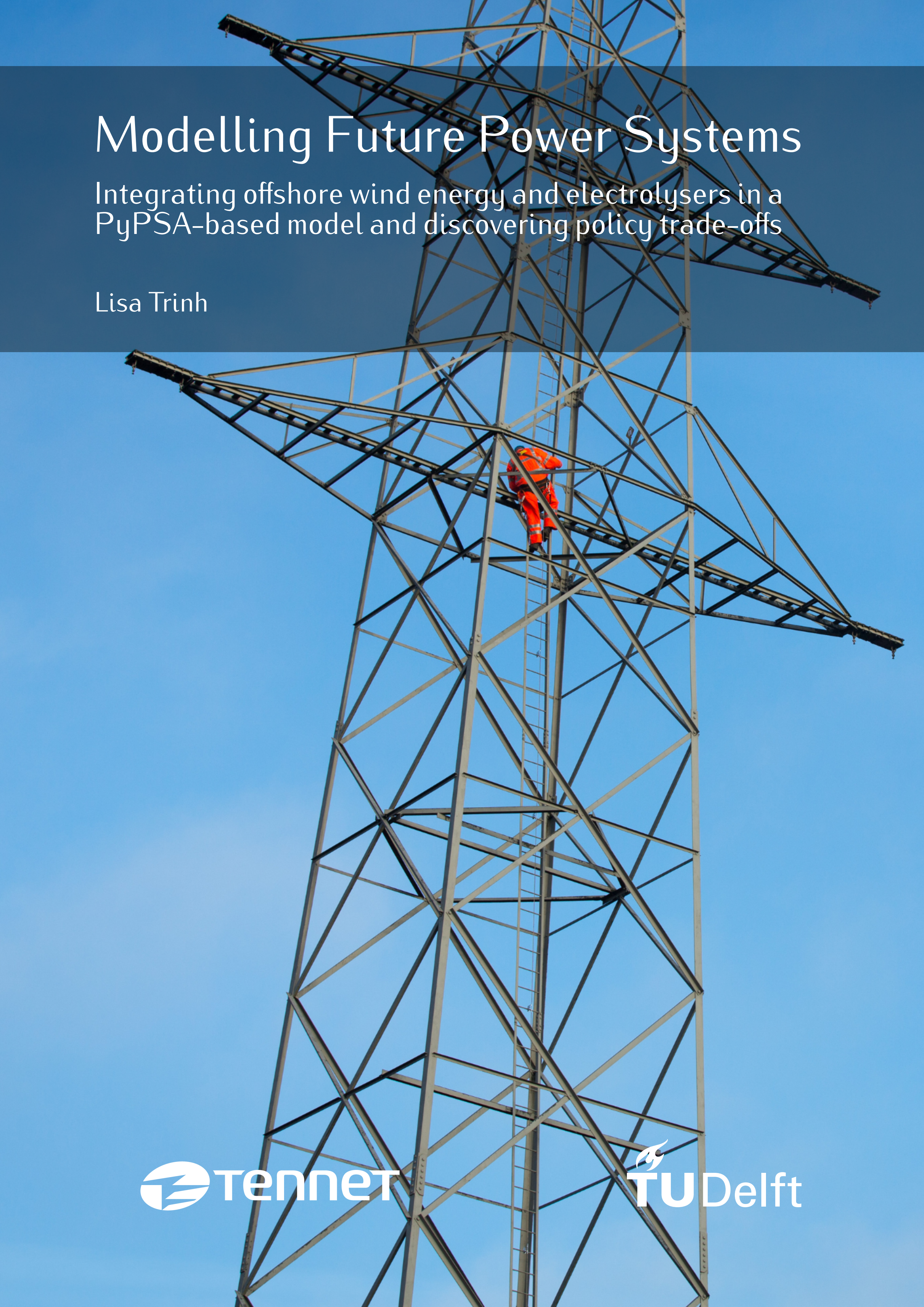
Modelling Future Power Systems

Integrating offshore wind energy and electrolysers in a PyPSA-based model and discovering policy trade-offs

Lisa Trinh



MODELLING FUTURE POWER SYSTEMS: INTEGRATING OFFSHORE WIND ENERGY AND ELECTROLYSERS IN A PYPSA-BASED MODEL AND DISCOVERING POLICY TRADE-OFFS

A master thesis submitted at the Technical University of Delft to partial completion of the requirements for the degree of

Master of Science
in Engineering and Policy Analysis

To be defended publicly on the 5th of March 2026.

by

Lisa Trinh
Student ID: 5100178

The cover photo is from TenneT

Lisa Trinh: *Modelling Future Power Systems: Integrating offshore wind energy and electrolysers in a PyPSA-based model and discovering policy trade-offs* (2026)

Faculty of Technology, Policy and Management
Delft University of Technology

Project duration:	September 2025- March 2026
Thesis committee:	Dr. S.J. (Stefan) Pfenninger-Lee, TU Delft, Chair
	Dr. F. (Francesco) Lombardi, TU Delft, First supervisor
	Dr. Ö. (Özge) Okur, TU Delft, Second supervisor
	M. (Michel) Dubbelboer, TenneT, External supervisor
	V. (Viktor) Beelen, TenneT, Second external supervisor

ACKNOWLEDGEMENTS

Dear reader,

Before you lies a manuscript that I've worked on with a lot of joy, drive and excitement. It discusses a tiny part of a big, scary monster, *the energy transition*. I'm proud that with this thesis, I've been able to make my contribution to battling this monster.

My interest in the energy transition started with, as many students do, needing an income. In my second year of my studies, I applied for a job as receptionist at a random company in Delft. And while it didn't immediately work out, they asked me to work as a course developer for offshore hydrogen and wind energy instead. In those years I got an unique chance to learn from professionals outside of academia about the technical as well as the policy side of offshore energy. Safe to say my interest in the energy transition was piqued!

This is how I slowly found my way to TenneT, who in my opinion is, at the heart of the scary monster. What better way to see what's up than to work at it? Through another offshore energy related internship I met Michel Dubbelboer, a policy advisor within TenneT who kindly mentored me through the thesis. Michel, thank you for the weekly yapping sessions! And Viktor Beelen, thank you for the weekly PyPSA-puzzle sessions! Needless to say I'd also like to thank the rest of the Energy System Planning – Systems Outlook team for adopting me, the fun lunches and walks around the office.

For the longest time I thought academia and I were not a match. This was up until I worked with my first supervisor, Francesco Lombardi. I got a glimpse of the excitement and challenges that comes with being at the forefront of innovative, promising technologies. Francesco, thank you for these insights! Stefan Pfenninger-Lee and Özge Okur, thank you for your feedback on the thesis!

Lastly, thank you to my friends and family for making these last few months as a student a very enjoyable one. For the dinners when I was too tired to cook one myself, for the coffee breaks that turned into discovering new parts of campus and for the terror but also fun swimming practices that somehow always recharged my battery. I'm grateful to move on to the next chapter of my life with such an amazing group of people.

Now it's time I'm putting you, reader, to work. Whether you have a background in the energy sector or not, I hope to pique your interest in this tiny part of the energy transition I've tried to tackle. I'm wishing you an as enjoyable time reading it as I had writing it.

Lisa Trinh
Delft, March 2026

ABSTRACT

INTRODUCTION The Dutch electricity grid is becoming increasingly congested as electrification increases and the share of solar and wind energy continues to grow. This congestion is driven by a mismatch between when electricity is generated and when it is consumed, leading to overloaded transmission lines and the curtailment of renewable energy. To maintain reliable grid performance in the future, policymakers and grid operators need insight in what and where to invest in energy technologies at a regional level. These insights must account for electricity flows in the transmission network and identify when cost-based trade-offs in the optimal system configuration arise to better account for investment uncertainties and make robust policies.

TenneT currently uses the detailed PLEXOS and PowerFactory/PSSE models for long-term grid planning, but their long run-times limit the use for quick insights into the Dutch electricity grid. Fast analysis is becoming increasingly needed to handle the increasing amount of inquiries from policy makers from the Dutch state and industries. This thesis investigates whether a simplified model, built in the open source software PyPSA, can deliver quantitative results to support early-stage planning, particularly in scenarios involving offshore electricity and its integration with electrolysers, which are currently in high priority when planning future scenarios.

The research uses optimisation techniques to allocate installed capacities of electricity technologies and calculate power flows under Netbeheer Nederland's 2040 Koersvaste Middenweg scenario, and it is designed in collaboration with TenneT experts to ensure the outputs are relevant and usable. The research includes: 1) defining key model performance indicators, 2) assessing how to reliably add offshore wind energy and electrolysers to the model, 3) quantitatively testing the PyPSA model against the PLEXOS model, and 4) defining which cost-based trade-offs take place in the system configuration. These findings will help to answer the main research question:

> *How can a PyPSA model of the Dutch transmission grid be designed and validated to support optimal capacity expansion planning for the 2040 Koersvaste Middenweg scenario, including offshore wind, electrolysers, and cost-driven trade-offs between technologies?*

METHOD For this thesis, a PyPSA model was developed and tested in an iterative way. Their results are validated against de results from TenneT's detailed models, specifically from the Energy System Planning teams. The PyPSA model is used to optimise and allocate electricity generation capacities and calculate power flows through the grid. After that, the EMA Workbench was used as a toolbox to systematically explore uncertainty in investment costs and to analyse when important cost-based trade-offs in the system configuration appear.

The PyPSA model is developed in three phases. First, the Netherlands is modelled as a single node to replicate TenneT's current dispatch single node model, PLEXOS. The results of this phase are 'validated' by comparing the optimised capacities results from PyPSA with those from PLEXOS. Nor-

mally, validation is done by comparing model results with historical data rather than future projections. However, in this thesis it is referred to as validation because TenneT considers the PLEXOS model to be a reliable and trusted reference model. Second, the model is expanded to include the transmission network, so that power flows and grid constraints are included. Load flow results from this phase are then compared with TenneT's PowerFactory/PSSE models. In the final phase, offshore wind farms and electrolysers are added to the model to study their impact on the electricity system and the grid. Also these model results gets compared to the previous model results.

After validation, all model phases are analysed using the EMA workbench. In this step, uncertainties in the investment costs of the most important technologies are explored by varying the costs between -20% and $+40\%$. This makes it possible to identify how sensitive the results are to cost assumptions and to understand when cost-based trade-offs happen between different technologies.

RESULTS Across all modelling phases, one conclusion is very stable: large amounts of solar capacity are cost-effective. Because solar output is variable, the system then needs supporting technologies to keep supply and demand balanced. In the single-node model, the PyPSA-based results show a higher deployment of batteries and gas to provide flexibility, while the PLEXOS-based results show a relatively stronger role for offshore wind. A likely reason for this difference is that cost assumptions are represented in a more detailed way in PLEXOS. When transmission grid constraints are included, the model results show a higher cost-optimal use of offshore wind. This is explained by its coastal location near large industrial demand centres, which reduces long-distance electricity transport and congestion within the network. This shows an important modelling lesson: including the network changes investment decisions, not only because of costs, but because location and congestion start to matter.

The comparison of load flow results indicates that the PyPSA-based results represent the overall flow patterns reasonably well. It does not match every individual line perfectly, which is expected because of simplifications like clustered substations and because PyPSA is built for long-term optimisation rather than operational grid security. Still, about 70–75% of the flow directions match the results of the TenneT models, and the overload results point to the same corridors that appear in TenneT's reinforcement plans. This increases confidence that the PyPSA model can identify the same patterns as TenneT's models.

Adding electrolysers and a hydrogen layer leads to another key insight: higher electricity use does not automatically mean more generation capacity is built. In the model, electrolysers mainly run in hours with surplus electricity and low prices, so they increase yearly electricity consumption but do not increase the system peak hour that drives capacity expansion. Because of this, the total optimal installed capacities results stay the same when electrolysers are added. Sensitivity analyses, especially halving the investment costs of nuclear energy, show that PyPSA-based results responds to cost changes in a similar way to PLEXOS-based results, confirming that the models behave in the same way.

Results from the EMA Workbench analysis shows that the future electricity system depends strongly on uncertain investment costs. Instead of pointing to one single 'optimal' solution, the results show that the system can take several typical configurations. A key result is that the system balances costs, emissions, and flexibility. Different technologies play different roles in this balance. Gas CCGT helps keep investment costs and peak electricity prices low and makes the system easier to operate, but it strongly increases CO_2 emissions and operating costs. Offshore wind does the opposite: it raises investment costs but lowers emissions and operating costs. In some cases, it can still lead to higher peak prices, because expensive OCGT backup plants are needed when wind production is low. Solar energy combined with batteries provides clean electricity and helps reduce peak prices, but it requires very large installed capacities and depends heavily on affordable storage. Gas OCGT mainly acts as flexible backup power and often works together with offshore wind, but when it runs during peak moments it can increase electricity prices. Nuclear power hardly appears in the results and only becomes relevant when its investment costs are strongly reduced, showing that it is not competitive under the current assumptions.

An effort was also made to improve the computation time of the PyPSA model, so that it would run significantly faster than the PLEXOS model. The most detailed version of the PyPSA model, which includes the transmission network and electrolysers, currently requires about one and a half hours to simulate a full year with hourly resolution. Previously, this calculation took more than a day. By applying time-step aggregation methods, the runtime can be shortened even further, in some cases to around five minutes. This shows that the PyPSA model can be used for fast analysis of the Dutch transmission grid, especially when a slightly lower temporal resolution is acceptable.

CONCLUSION Important insights of the analysis in this thesis are that PyPSA can be used to support early-stage planning, particularly in scenarios involving offshore wind energy and electrolysers, and that even small changes in investment costs can shift the system from one configuration to another. At the same time, the system is still able to meet electricity demand in all cases. This aligns with recent energy modelling research, especially in the field of structural uncertainty analyses, or Modelling to Generate Alternatives, that looks beyond a single 'optimal' solution and instead studies a range of possible outcomes. These results suggest that the future electricity system is not fragile, but its structure strongly depends on which technologies become cheaper or more expensive over time. Different technology combinations can all work, but they lead to different trade-offs in flexibility, emissions or costs. For policymakers, this shows that robust planning, so being prepared for future changes in investment costs, is more important than focusing on one single 'optimal' future pathway. Multiple pathways should be ready to be deployed as time goes.

This thesis contributes to faster analysis of the Dutch electricity grid compared to existing methods. By testing different model setups, this thesis demonstrates how optimisation runtimes can be reduced while maintaining results that are similar to existing models. Finally, the thesis highlights that optimal solutions are sensitive to cost uncertainty and that there are multiple system configurations that can fulfil system needs. For policy analysts, it is therefore important to take these cost uncertainties into account and to look for flexible solutions that perform well under different assumptions. A practical first step would be to model different system configurations under a range of cost assumptions, then test these outcomes using structural uncertainty analysis. Based on this, a roadmap can be developed that does not rely on one single set of assumptions but remains robust under different conditions.

CONTENTS

1 INTRODUCTION 2
- 1.1 The need for simplified electricity grid modelling . . . 2
- 1.2 Information gap on offshore wind energy and electrolysers . . . 3
- 1.3 Trade-offs in energy system optimisations . . . 3
- 1.4 Thesis scope and outline . . . 4

2 RESEARCH FORMULATION 6
- 2.1 State of the art literature review . . . 6
  - 2.1.1 Search strategy . . . 6
  - 2.1.2 Literature synthesis . . . 7
  - 2.1.3 Articulation of knowledge gaps or novelty . . . 11
- 2.2 Research objective . . . 12
- 2.3 Relevance . . . 13
  - 2.3.1 MSc Engineering and Policy Analysis relevance . . . 13
  - 2.3.2 Scientific relevance . . . 14
  - 2.3.3 TenneT relevance . . . 14
  - 2.3.4 Societal relevance . . . 14

3 RESEARCH DESIGN 16
- 3.1 Methods . . . 16
- 3.2 Overall design of phases . . . 18
- 3.3 Data sources . . . 20

4 THEORETICAL BACKGROUND 22
- 4.1 Electricity grid key performance indicators . . . 22
- 4.2 Role of hydrogen in the energy mix . . . 24
- 4.3 Trade-offs in the optimisation . . . 24
  - 4.3.1 Modelling to Generate Alternatives . . . 25
  - 4.3.2 EMA Workbench . . . 25
  - 4.3.3 Perspective of TenneT . . . 26

5 MODEL CONCEPTUALISATION 27
- 5.1 Key Performance Indicators . . . 27
- 5.2 PyPSA . . . 28
  - 5.2.1 Input variables . . . 28
  - 5.2.2 Core design . . . 29
  - 5.2.3 Optimisation . . . 31
- 5.3 EMA workbench . . . 32
  - 5.3.1 Framework . . . 32
  - 5.3.2 Open exploration . . . 34

6 MODEL FORMALISATION 36
- 6.1 PyPSA set-up . . . 36
  - 6.1.1 Phase 1: Single node . . . 38
  - 6.1.2 Phase 2: Transmission network . . . 39
  - 6.1.3 Phase 3: Transmission network + electrolysers . . . 44

6.2 EMA workbench set-up . . . 45
6.2.1 Uncertainty assumptions . . . 46
6.2.2 DelftBlue . . . 46
6.2.3 EMA Workbench code . . . 47
6.3 Data assumptions . . . 47
7 RESULTS 50
7.1 PyPSA results . . . 50
7.1.1 Improvement runtime PyPSA . . . 50
7.1.2 Phase 1: Single node . . . 51
7.1.3 Phase 2: Transmission network . . . 52
7.1.4 Phase 3: Transmission network + electrolysers . . . 54
7.1.5 Sensitivity analysis: halved CAPEX for nuclear . . . 59
7.1.6 Sensitivity analysis: uncapped solar expansion . . . 60
7.1.7 Synthesis of PyPSA results . . . 61
7.2 EMA workbench results . . . 62
7.2.1 Improvement runtime EMA workbench . . . 62
7.2.2 Phase 1: Single node . . . 62
7.2.3 Phase 2: Transmission network . . . 65
7.2.4 Phase 3: Transmission + electrolysers . . . 65
7.2.5 Synthesis of EMA workbench results . . . 70
8 DISCUSSION 72
8.1 Reflecting on this study . . . 72
8.2 Limitations . . . 74
8.2.1 PyPSA model structure . . . 74
8.2.2 EMA workbench structure . . . 76
8.3 Personal reflections . . . 76
8.4 Recommendation for policy developers . . . 77
8.5 Recommendation for future research . . . 78
9 CONCLUSION 79
REFERENCES 83
A APPENDIX A: PYPSA INPUT VARIABLES 89
A.1 Data files . . . 89
A.2 Coordinates of offshore wind farms . . . 90
A.3 Cost calculations . . . 91
B APPENDIX B: SOLVER OPTIONS 94
C APPENDIX C: RESULTS 96
C.1 PyPSA results . . . 96
C.1.1 Phase 1: Single node . . . 96
C.1.2 Phase 2: Transmission network . . . 97
C.1.3 Phase 3: Transmission network + electrolysers . . . 99
C.2 EMA workbench results . . . 101
C.2.1 Phase 1: Single node . . . 101
C.2.2 Phase 2: Transmission network . . . 103
C.2.3 Phase 3: Transmission + electrolysers . . . 104

LIST OF FIGURES

Figure 3.1 Flow diagram of research questions . . . 18
Figure 3.2 Grid models: (a) single node, (b) transmission network and (c) transmission + hydrogen network. . . . 19
Figure 3.3 Flow diagram of methods . . . 20
Figure 4.1 Illustration of uncertainties in a deterministic model (Tennøe et al., 2018) . . . 26
Figure 5.1 Schematic drawing of an electric bus . . . 30
Figure 5.2 Conceptual model of input and output of PyPSA . . . 31
Figure 5.3 Latin Hypercube Sampling (Altair, 2026) . . . 33
Figure 6.1 Map of the Netherlands as a single node . . . 39
Figure 6.2 380 and 220 kV substations in Eemshaven . . . 40
Figure 6.3 Map of the Netherlands with transmission network . . . 41
Figure 6.4 Comparison of optimisation solvers . . . 43
Figure 6.5 Differences in time series aggregation techniques (PyPSA Docs, 2025b) . . . 44
Figure 6.6 Map of the Netherlands for transmission network + electrolysers . . . 45
Figure 7.1 Deviation in installed capacities for single node model . . . 52
Figure 7.2 Deviation in installed capacities for transmission model . . . 53
Figure 7.3 Deviation in installed capacities for transmission model + electrolysers . . . 55
Figure 7.4 Optimal installed capacities by PyPSA and Guidehouse for transmission + electrolysers and single node . . . 55
Figure 7.5 Differences in load flows PyPSA and Grid Planning models for transmission + electrolysers model . . . 57
Figure 7.6 Planned TenneT projects for line upgrading (TenneT, 2026a) . . . 58
Figure 7.7 Deviation in installed capacities for PyPSA and PLEXOS when CAPEX of nuclear is halved . . . 59
Figure 7.8 Deviation in installed capacities for PyPSA and PLEXOS when solar is uncapped 60
Figure 7.9 Pair plot for single node model . . . 63
Figure 7.10 Clustered plot for transmission network + electrolyser model . . . 66
Figure 7.11 Theil-Sen energy plot for clustered plot with electrolysers . . . 68
Figure B.1 Installed capacities test Highs and Gurobi: 1 week per month . . . 95
Figure B.2 Installed capacities test Highs and Gurobi: 1 year . . . 95
Figure C.1 Optimal installed capacities by PyPSA and Guidehouse for single node . . . 96
Figure C.2 Optimal installed capacities by PyPSA and Guidehouse for transmission network . . . 97
Figure C.3 Explore plot of the transmission network in PyPSA . . . 98
Figure C.4 Optimal installed capacities by PyPSA and Guidehouse . . . 99
Figure C.5 Explore plot of the transmission + electrolyser network in PyPSA . . . 100
Figure C.6 Archetypes for single node . . . 102
Figure C.7 Violin plot for single node . . . 103
Figure C.8 Archetypes for transmission network . . . 103
Figure C.9 Dispatch for transmission + electrolyser network . . . 104

LIST OF TABLES

Table 5.1	PyPSA network components (PyPSA Docs, 2025a)	29
Table 6.1	Overview of clustered substations in PyPSA model	40
Table 6.2	Uncertainty ranges for non-mature energy technologies Guidehouse and TenneT (2025).	46
Table 6.3	Differences input data in PyPSA and PLEXOS	49
Table 7.1	Ranking of PLEXOS and PyPSA model from largest to smallest installed capacity for single node	51
Table 7.2	Ranking of PLEXOS and PyPSA model from largest to smallest installed capacity for transmission network	53
Table 7.3	Installed electrolyser capacities for hydrogen output in PyPSA per substation	56
Table 7.4	Top 10 most overloaded lines in the Netherlands in PyPSA	58
Table 7.5	Capital cost factors for transmission + electrolysers cluster plot	66
Table 8.1	Past, current and optimised installed capacities for the Netherlands (Netbeheer Nederland, 2025; TenneT, 2015)	73
Table A.1	Excel sheets and their input variables for the PyPSA model	89
Table A.2	Offshore wind parks, their (x,y) coordinates, landing location, and the distance between them (Ministry of Economic Affairs and Climate, 2025; Ministry of Infrastructure and Water Management, 2025; Rijkswaterstaat, 2025)	90
Table B.1	Time to run the model with Gurobi	94
Table C.1	Colors of explore plot	98
Table C.2	Capital cost factors for single node archetype plot	102
Table C.3	Capital cost factors for transmission network archetype plot	104

ACRONYMS

CAPEX Capital Expenditures 22

EPA Engineering and Policy Analysis 13

EPN-SO Energy Systems Planning - Systems Outlook 3

ETM Energy Transition Model 4

ETS Emissions Trading System 48

KM Koersvaste Middenweg 4

KPI Key Performance Indicator 22

LHS Latin Hypercube Sampling 32

MGA Modelling to Generate Alternatives 25

OPEX Operating Expenditures 22

RES Renewable Energy Sources 22

TSO Transmission System Operator 2

1 INTRODUCTION

In recent years, the Dutch electricity grid has come under growing pressure due to the mismatch between when electricity is generated and when it is consumed (Netherlands Enterprice Agency, 2023). During the day, the increasing supply of solar energy leads to temporary surpluses, while electricity demand peaks in the evening (Van Someren et al., 2025). This growing imbalance between supply and demand causes grid congestion, leading to curtailment of excess electricity, and shortages later in the day (Schermeyer et al., 2017; Van Someren et al., 2025). If unmanaged, these imbalances could lead to reliability issues such as the 2025 Spain-Portugal blackout (ENTSOE, 2025; Fotis et al., 2023).

1.1 THE NEED FOR SIMPLIFIED ELECTRICITY GRID MODELLING

To address this, TenneT, the Dutch Transmission System Operator (TSO) and the government are exploring different solutions, including grid reinforcement, strategic planning for new developments such as wind parks, demand side response, and industrial expansions (Netherlands Enterprice Agency, 2025). A key challenge in this planning process is identifying where and when congestion will arise, how severe it will be, and which policies offer the most impact.

Currently, TenneT uses a energy system model called PLEXOS to calculate long-term dispatch, market simulations and investment optimisations (Energy Examplar, 2025). The model uses detailed operational data from all generators, demand centres and cross-border interactions across Europe to perform an optimisation of the electricity system. At the moment, PLEXOS is applied as a single-node model, where each country is represented as one aggregated node. However, steps are currently being taken to regionalise the model, so that internal grid structure and spatial differences within countries can also be represented in future analyses.

Even though the results are presented at a single-node level, PLEXOS produces highly detailed outputs. Because of this level of detail and the size of the model, the computation time is relatively long (Nweke et al., 2012); internal sources have reported that one run can take up to six days on a fast computer to run an investment optimisation model. Afterwards, a power flow simulation has to be done in PowerFactory/PSSE, which increases the waiting time even more (DIgSILENT, 2026). As questions from policymakers become more frequent and time-sensitive, there is a growing need for a faster tool that can produce quick, reliable estimates that is within a certain bandwidth of the PLEXOS and PowerFactory/PSSE results. Such a tool would support early-stage planning by offering timely insights into future grid bottlenecks, expansion costs, and development options of generation technologies or batteries.

This thesis responds to that need by developing a simplified Dutch transmission grid model, built with the open source toolbox PyPSA (Brown et al., 2018). A tool that TenneT, specifically their Energy Systems Planning - Systems Outlook (EPN-SO) team, has recently begun investing in for internal use. The aim is to test whether such a model can provide quantitative, policy-relevant results for early-stage investment planning, particularly in scenarios involving offshore electricity and electrolysers, since this is high priority knowledge for TenneT in the upcoming years.

1.2 INFORMATION GAP ON OFFSHORE WIND ENERGY AND ELECTROLYSERS

At the moment, the Netherlands has ten active offshore wind parks, which have a cumulative installed capacity of 4.7 GW (RVO, 2024). Offshore wind energy is expected to grow in the upcoming years: the next goal is 21 GW in 2031, and 30-40 GW before 2040 (Rijksoverheid, 2025). Even though the timeline and exact capacity of new installations are still being decided, they will clearly have a major impact on the electricity system since it will strongly influence how the rest of the electricity system needs to be designed (Hermans, 2025). A system with a high share of offshore wind requires enough supporting technologies, such as storage, flexible generation, and hydrogen production to remain reliable at all times (Durakovic et al., 2023; North Sea Wind Power Hub, 2021). In addition, electricity must be transported from offshore landing points to areas with high demand without causing congestion in the grid (Schermeyer et al., 2017). Energy system modelling provides a way to study these challenges and to explore how different technology choices and locations can support a reliable and cost-effective electricity system in the future.

In scenarios from Netbeheer Nederland (2025), electrolysers and hydrogen are going to play a significant role in the Dutch energy supply in the upcoming years. This is partly due to the European Commission's REPowerEU initiative that sets out to cover 10% of the EU's energy needs with hydrogen (European Commission, 2026). By producing or consuming hydrogen alongside the variable offshore wind supply, the ups and downs of wind power are expected to be balanced (Martínez-Gordón et al., 2022). However, it is still unclear how much hydrogen supply will be needed in the future, where hydrogen production should be located, and how this will affect the electricity grid. These questions are important for TenneT's future investment strategy and should therefore be studied in more detail.

1.3 TRADE-OFFS IN ENERGY SYSTEM OPTIMISATIONS

Even though energy system modelling is important for providing quantitative insights for policy-makers, researchers increasingly agree that purely cost-optimisation models do not fully represent the real-world energy transition (Neumann & Brown, 2023; Trutnevyte, 2016). These models usually point to one 'optimal' solution based on a fixed set of cost assumptions. However in reality many of these costs, such as future technology prices, are uncertain and can change over time, which can strongly influence model outcomes (Neumann & Brown, 2021).

Because of these uncertainties, policy analysts should not focus only on the single lowest-cost solution calculated by a model. Instead, they should also look at the alternative solutions when key assumptions change. Small variations in costs, additions of factors such as societal acceptance, or changes in other inputs can lead to different system configurations. Understanding the different types of uncertainty in energy modelling, and how to respond when alternative solutions occur is essential for developing energy policies that remain robust under a range of possible conditions, rather than relying on one fixed outcome.

1.4 THESIS SCOPE AND OUTLINE

Netbeheer Nederland (2025) created three future scenarios that describe possible developments in the Dutch energy system towards the year 2040. These scenarios are widely used by Dutch System Operators, including TenneT. Each scenario reflects different assumptions about factors such as technological progress, government policy, and the pace of the energy transition. They provide a way to explore how different futures might affect the design and performance of the power system, and its data can be found in the Energy Transition Model (ETM) (Energy Transition Model, 2025):

- Koersvaste Middenweg (KM) or *Steady Middle Course* scenario represents the expected pathway of the Dutch energy transition based on current trends and existing policy ambitions. A key feature of this scenario is a strong and rapid electrification of energy use. This means that many sectors, such as transport, heating, and industry, increasingly switch from fossil fuels to electricity. To keep the total energy demand balanced, electricity is complemented by other energy carriers, for example, hydrogen, biogas, or heat networks in areas where full electrification is not yet practical. In the Koersvaste Middenweg scenario, the total electricity demand in 2040 is projected to be 373 TWh per year, which is in the middle of all the scenarios.
- Eigen Vermogen (EV) or *Own Power* scenario describes a future in which the Netherlands focuses strongly on energy independence and self-sufficiency. In this scenario, national policy, politics, and market developments are all aimed at producing as much energy domestically as possible, reducing reliance on imports. The country invests heavily in flexibility, large-scale heat networks, and the use of green hydrogen to balance energy supply and demand across different sectors. The scenario also assumes a rapid electrification of industry, transport, and buildings. In the Eigen Vermogen scenario, total electricity demand in 2040 is projected to reach 442 TWh per year, the highest among the three Netbeheer Nederland scenarios.
- Gezamenlijke Balans (GB) or *Shared Balance* scenario describes a future in which European cooperation and coordination are at the heart of the energy transition. In this scenario, the Netherlands works closely with neighbouring countries to develop a interconnected energy system. The shift toward sustainability in different sectors follows a hybrid approach, where both electrification and the use of gases play important roles. The existing gas infrastructure continues to be used for a mix of energy carriers such as natural gas, green gas, biofuels, and blue hydrogen. In this scenario, the total electricity demand in 2040 is projected to be 323 TWh per year, which is lower than in the other two scenarios.

TenneT recently carried out a study similar to this thesis, in which a capacity expansion optimisation was performed using the PLEXOS model. This study was conducted together with the external consultancy firm Guidehouse. It was based on the KM scenario, which represents the middle pathway of the three future scenarios (Guidehouse & TenneT, 2025). For this reason, the KM scenario is also chosen as the main focus of this thesis. While the Guidehouse report analyses the years 2030, 2035, and 2040, this thesis only analyses for the year 2040. And instead of looking at incremental steps for future investment, this thesis focuses on what the system design should look like in 2040 to ensure cost-effectiveness.

This thesis proposal is structured to provide a logical path toward answering the research questions. Chapter 2 presents the research formulation, starting with a state-of-the-art literature review, followed by the research questions and the relevance of this thesis. Chapter 3 describes the research design, outlining the methods used to answer each research question, the overall structure of the study, and the data sources. Chapter 4 introduces the theoretical framework and explains the key concepts, theories, and definitions that form the basis of this research. Chapter 5 focuses on the conceptualisation of the model and describes its design and structure. Chapter 6 details the data preparation, and assumptions used in the modelling process. The results of the analysis are presented in Chapter 7, followed by Chapter 8, which discusses the findings and outlines the limitations of the study. Finally, Chapter 9 presents the conclusions. Together, these chapters form the foundation of the thesis and show how the research is carried out in a structured way.

2 RESEARCH FORMULATION

This chapter provides an overview of existing research and tools related to electricity grid modelling. The goal is to understand what tools have already been developed, what the considerations are in electricity grid modelling, and it talks about some European case studies. It then discusses where current limitations or gaps exist, after which de research questions are formulated. The chapter ends discussing the relevance of the thesis.

2.1 STATE OF THE ART LITERATURE REVIEW

A literature review will be conducted with the aim to understand the scientific background of electricity grid modelling, the challenges associated with lightweight optimisation energy modelling tools, the implications of integrating offshore wind energy and electrolysers, uncertainties in energy modelling, and lastly a review about European case studies.

2.1.1 Search strategy

The search strategy will be designed to align closely with the research problem: the need for simplified, yet reliable, electricity system models for early-stage planning under conditions of rising electrification and growing system complexity.

Searches will be conducted using academic databases such as Research Gate, Science Direct, SpringerLink, and Google Scholar. Key search queries will include combinations of the following terms:

- “simplified electricity grid model” OR “reduced complexity power system model”.
- “PyPSA applications” OR “open-source power system modelling”.
- "offshore wind energy integration modelling" OR "offshore energy hub modelling".
- “electrolyser integration modelling” OR “sector coupling in electricity systems”.
- “electricity grid congestion” AND “Netherlands” OR “Europe”.
- “validation of energy models” OR “model comparison energy planning”.

These terms will be iteratively refined based on the initial results. The queries will be selected to help understand the practical problem (grid congestion and planning challenges in the Netherlands) into broader scientific themes: model complexity, quality, speed trade-offs, and integration of emerging technologies into system models.

Inclusion criteria will consist of:

- Peer-reviewed academic papers and technical reports published from 2015 onward.
- Studies modelling national or regional electricity grids, preferably using open-source tools.
- Research that develops, compares, or validates simplified models.
- Studies involving integration of offshore wind energy and/or electrolyser integration.

Exclusion criteria will include:

- Articles focused only on market simulation without network modelling.
- Country-specific case studies lacking methodological depth or generalisability.
- Commercial model documentation that does not explain model logic.

The results of the search will be synthesised thematically to support the development of the conceptual framework and sub-questions. This approach will ensure that the review is methods-driven, academically grounded, and methodologically robust.

2.1.2 Literature synthesis

The literature on electricity system modelling is diverse, and ranges from very detailed optimisation tools used by system operators to lightweight models designed for flexibility and speed. This review is organized around four themes: model complexity, offshore wind energy and electrolyser integration, uncertainties in energy modelling, and existing opens-source energy modelling frameworks, specifically in European studies.

Model complexity

Energy system optimisation models are essential for understanding energy systems and supporting policy decisions. In this thesis, the focus will be on electricity and a bit on hydrogen, but other forms of energy such as heating and gas are out of scope. As electricity systems become more detailed due to rising flexible demand, electrification and varying renewable resource potential, models often become more complex (Hoffmann et al., 2021). Characteristics of complex systems are path dependency, links between time steps, non-linear relationships between components, discrete decisions, and system size (Priesmann et al., 2019). This increase in model complexity can bring both benefits and challenges (Nolting & Praktiknjo, 2022).

The main benefit of model complexity is that it allows more detail to be included, such as intermittent renewable energy sources, storage technologies, and sector interconnections (Hoffmann et al., 2021). Greater complexity also enables higher temporal and spatial resolution, meaning that models can capture variation across shorter time steps and smaller regions (Hoffmann et al., 2021). It is often assumed that added detail provide more nuanced results that better reflect real system behaviour (Priesmann et al., 2019).

However, the challenges of model complexity are higher computational costs, longer run-times, and higher data quality needs (Hoffmann et al., 2021). Complex models are harder to interpret and more difficult to explain to non-experts. Also, more complex models don't always give more qualitative results, especially when input data is uncertain or poor (Nolting & Praktiknjo, 2022). This is a common issue in long-term scenario analysis, where lots of future uncertain data is needed.

So the appropriate level between model complexity and quality needs to be deliberated. There is an insightful part of Box (1976) his research paper that underlines this:

> Since all models are wrong, the scientist cannot obtain a 'correct' one by excessive elaboration. ... Just as the ability to devise simple but evocative models is the signature of the great scientist, so overelaboration and overparametrization is often the mark of mediocrity.

To manage the line between complexity and quality, the literature mentions several recommendations (Priesmann et al., 2019). The first is using the right type of electricity grid modelling. Priesmann et al. (2019) compared two types of models, one without transmission limits and one with transmission constraints, and it showed that the second showed more similar results to historical electricity prices and dispatch schedules. The study recommends that in cases of small increase of computational costs, especially in large models, to use the more complex model. The second is time-coupling constraints, for example for dispatched or dynamic storage units. If the number of time steps is large and multiple time-coupling variables are present, solver time will increase significantly. Thirdly and lastly, temporal, spatial and technological resolution. Increasing temporal and spatial resolution results in an exponential growth in solving time and model size, while its impact on accuracy is low. For example, using representative days through clustering can better preserve the impact of solar and wind fluctuations than simply adding more time steps (Hoffmann et al., 2021).

To summarize the complexity dilemma: the more detailed the system, the more complex the model needs to be. However more complexity also increases uncertainty, reduces transparency, and demands more resources (Nolting & Praktiknjo, 2022). In many cases, a simpler model may be more useful, especially when data is limited, when decisions need to be made quickly, or when results must be communicated to non-technical audiences. Most importantly is a close cooperation between modellers and users to truly unlock the potential of models for its purposes (Süsser et al., 2022).

Integration of offshore wind energy and electrolysers

On an European level, offshore renewable energy sources will increase significantly in the upcoming years to support the goal of net zero greenhouse gas emissions by 2050 (North Sea Wind Power Hub, 2021). This is reflected in the Dutch government's ambition to install 30–40 GW of offshore wind capacity by 2040 (Hermans, 2025). Depending on the capacity factor used, this could roughly match the Netherlands' total electricity consumption in 2024, which was 374 PJ (or, 103.611 GWh) (Energie Beheer Nederland, 2024). However, the variable nature of wind energy means that production does not always align with demand. To manage this mismatch, forms of energy storage are essential (North Sea Wind Power Hub, 2024).

Two main storage options are discussed: batteries and hydrogen storage (Feijoo et al., 2022). While small-scale battery storage is already common, large-scale battery deployment faces several challenges (Van Someren et al., 2025). These include high costs, the use of scarce raw materials, and the need for grid connections, which are difficult given the current congestion issues on the Dutch electricity grid (Ronde, 2025; Vandebron, 2023). As a result, hydrogen is seen as a promising alternative. It can function both as an energy source and an energy carrier, which can be used as a long-term storage medium, offering flexibility and transportability (Xing et al., 2018).

Hydrogen is especially interesting when it is combined with offshore wind energy (North Sea Wind Power Hub, 2024). Offshore wind farms often produce large amounts of electricity, sometimes more than what is needed on land at that moment. When this happens, the extra electricity cannot always be transported or used and is therefore curtailed. By using this surplus offshore wind electricity to produce hydrogen, curtailment can be reduced and the energy can be stored in another form, which can be used in other sectors, such as industry or transport. Many studies agree that producing hydrogen from offshore wind can lower curtailment and reduce total system costs (Durakovic et al., 2023; Gea-Bermúdez et al., 2023; Lüth et al., 2023; Martínez-Gordón et al., 2022). In theory, this makes the electricity system more efficient and better suited for a high share of renewable energy.

However, recent developments show that hydrogen ambitions have slowed down (Hermans, 2025). Expected cost reductions have not been realised as anticipated, and hydrogen demand hasn't grown as fast as planned (Hermans, 2025). As a result, several hydrogen projects have been delayed or reconsidered, which raises questions about the short- and medium-term role of hydrogen in the energy system.

Most existing research focuses on offshore wind and hydrogen integration at the level of the North Sea region or multiple countries combined (Durakovic et al., 2023; Glaum et al., 2024; Martínez-Gordón et al., 2022). Much less attention has been given to the specific effects of connecting offshore wind farms and electrolysers to only the Dutch electricity grid. This local perspective is important, because grid congestion, industrial demand locations, and infrastructure constraints strongly influence how these technologies perform in practice. More research is therefore needed to better represent offshore wind and electrolysers in Dutch electricity system models and to understand their role in supporting the energy transition.

Modelling practices in European case studies

PLEXOS is a tool that is used by TenneT, and a few other European TSO's for energy system modelling (Spasić, 2023). It simulates electricity dispatch in great detail, but usually represents an entire country as a single node. In recent years, there has been a shift to modelling energy systems from a GIS perspective that are open-source and also includes geographical dimensions (Valdes et al., 2020). The most used model is PyPSA, which is developed at TU Berlin (Brown et al., 2018; Henke et al., 2022; Limpens et al., 2024; Valdes et al., 2020). Another noteworthy open-source energy system model is Calliope, which is developed at TU Delft (Pfenninger & Pickering, 2018). While the two models are very similar, there are some properties that differentiate the two.

According to Pfenninger and Pickering (2018), Calliope is a framework to build energy system models, designed to analyse systems with arbitrarily high spatial and temporal resolution. Its formulation of energy system components was influenced by the power nodes modelling framework by Heussen et al. (2010) and cleanly separates the general framework (code) from the problem-specific model (data) (Pfenninger & Pickering, 2018). One of the biggest advantages of Calliope is its intuitive interface which makes it easy for non-programmers to use the model. However, Calliope does not account for in-depth power flow simulations (Valdes et al., 2020).

PyPSA is an open-source tool for simulating and optimising modern power systems across multiple time periods (Brown et al., 2018). It bridges the gap between traditional power system analysis software and broader energy system modelling tools. PyPSA supports operational and investment optimisation, full and linearised load flow calculations, and is designed to handle variable renewables, storage, sector coupling, large networks, and long time series efficiently (Brown et al., 2018).

Both Calliope and PyPSA have been used in research to support the planning of energy grids. Valdes et al. (2020) have used both to make a framework for regional smart energy planning for the optimal location and sizing of small hybrid systems. Henke et al. (2022) compared the two, but eventually built their analysis of dynamic power sector expansion in OSeMOSYS. Limpens et al. (2024) also compared the two, after which they made their own model, EnergyScope Pathway, to better capture time resolution. While many more research has been done with these models, these were most interesting ones because they contained comparisons between them.

In this thesis, PyPSA is selected over other open-source tools like Calliope because PyPSA is as of time of writing the thesis the most widely used open-source energy modelling tool, and because of path dependency of internal TenneT processes. This is because it offers a good balance between model transparency, computational efficiency and in-depth power-flow calculations. It is well-documented, actively maintained, and widely used and validated in European contexts, which makes it a trusted choice for system-level analysis (Unnewehr et al., 2022). PyPSA has also already been used to research the integration of offshore wind energy and to what extend hydrogen should play a role (Glaum et al., 2024). However, not much is known about energy system modelling with PyPSA for TSO's.

Parametric and structural uncertainty of energy modelling

A recent study by Bock and Pfenninger-Lee (2025) discusses that there are different views within the modelling community on truth, objectivity and that the way that results are presented are influenced by explanatory power. Seeing models as tentative emphasizes flexibility and supports informed decision-making, while treating them as objective truths may restrict perceived options and transfer decision-making power to the model. To address these tensions, the authors call for diverse modelling teams and greater discussion of epistemic beliefs. For this thesis, it is important to remember that optimisation models often generate one 'optimal' outcome even though there are many near-optimal outcomes, therefore it should not be discussed as the absolute truth.

In recent years, there is the growing consensus that cost-optimisation energy models on their own are not sufficient for energy policy analysis, because they do not fully reflect how the real world works (Trutnevyte, 2016). These models usually identify one 'optimal' solution based on a fixed set of assumptions. However, focusing on a single optimal outcome can be short-sighted, especially in a complex system like the energy transition, where many different actors are involved, interests often conflict, and future developments are highly uncertain (Lombardi et al., 2025; Yue et al., 2018).

Researchers distinguish between two main types of uncertainty in energy system models: parametric and structural uncertainty (DeCarolis et al., 2017). Parametric uncertainty refers to uncertainty in the input values of the model, such as future technology costs or fuel prices (Trutnevyte, 2016). Structural uncertainty refers to uncertainty in how the model itself represents the energy system, including the mathematical relationships that describe how the system develops and operates (Neumann & Brown, 2021). In recent years, there's a growth of studies on structural uncertainty, such as by a technique called Modelling to Generate Alternatives (Lombardi et al., 2025; Neumann & Brown, 2021). These models are well suited to explore system designs that prefer other factors beyond costs, such as social acceptance, asset values or compromises between stakeholders (Lombardi et al., 2020; Lombardi et al., 2025; Neumann & Brown, 2023).

Parametric uncertainty focuses on the fact that many input values used in models are not known with certainty, such as investment costs, fuel prices, or technology performance (DeCarolis et al., 2017). Instead of treating these numbers as fixed, parametric uncertainty analysis explores how changes in these values affect the outcomes of a model. Traditionally, this kind of analysis has been done by varying one or a few parameters at a time to see how sensitive the results are to each assumption. In recent years, there has been increasing interest in approaches that look more broadly at this kind of uncertainty, known as decision making under deep uncertainty (Kwakkel, 2017). Deep uncertainty refers to situations where experts and decision-makers do not fully agree on the system itself, the most important outcomes to consider, or even the probability distributions of uncertain inputs, making it difficult to agree on future conditions (Haasnoot et al., 2013; Walker et al., 2013).

The EMA Workbench, developed by researchers at TU Delft, is a toolbox designed to support this type of exploration by allowing many uncertain input parameters to be varied systematically and repeatedly across a wide range of values (Kwakkel, 2017). Instead of manually testing a few combinations, the toolbox runs many model experiments automatically and helps analyse how outcomes change across all of them. The emphasis is not on predicting one future, but on exploring many plausible combinations of assumptions and identifying patterns, vulnerabilities, and trade-offs in the results. This approach fits well with the PyPSA model in this thesis, because the goal is exploratory rather than predictive. To keep the analysis manageable within the time frame of the thesis, only investment costs of technologies will be tested, as these are expected to have the most impact. Applying parametric uncertainty analysis with the EMA Workbench to a PyPSA model has not been done before, which makes this thesis academically relevant.

2.1.3 Articulation of knowledge gaps or novelty

The literature shows a knowledge gap in model complexity when building electricity models. Greater complexity enables higher temporal and spatial resolution, however it also results in higher computational costs, longer run-times and higher data needs. They are also harder to interpret and explain to non-experts. Faster, simplified models are needed, but these need to be of enough detail for policymakers. A middle ground needs to be found for good tooling.

Another gap is about the integration of offshore wind energy and electrolysers in simplified models. Hydrogen is expected to play a role in balancing supply and demand in future energy systems. Most studies look at the inclusion of offshore wind energy and electrolyser on a North Sea scale, while there's little research on the Netherlands on its own, and the effects on its transmission network.

A third gap in the literature is the limited use of simplified, open-source energy system models by TSO's, especially models based on PyPSA. However, there is still limited research on how well these simplified models perform in practice. In particular, it is not always clear whether they produce results that are broadly consistent with more detailed and resource-intensive tools such as PLEXOS. This raises an important question about the reliability and usefulness of simplified open-source models for real-world capacity expansion planning.

Finally, there has been little to no research that combines deep uncertainty analysis with a PyPSA model using the EMA Workbench. In most parametric uncertainty studies, only a few input values are changed one at a time. While this can show sensitivity to individual assumptions, it does not explore the full range of ways a system might behave when many uncertain factors vary together. A more systematic and exploratory approach makes it possible to see how combinations of uncertain investment assumptions affect model outcomes and to identify cost-based trade-offs between different technologies. It also helps show which system configurations remain relatively stable when investment assumptions change. These kinds of insights are important for designing energy policies that are strong and useful across a wide range of plausible conditions, rather than relying on a single set of assumptions.

This thesis addresses these gaps by testing whether offshore wind energy and electrolysers can be added to a simplified PyPSA-based model, if this can give the same optimisation outputs as more complex models, and to test what trade-offs occur between technologies. The goal of this thesis is to make a model for TenneT that can make quick data-backed optimisations without needing large, time-consuming models. By focusing on a practical use case with real data and stakeholder input, the research helps bridge the gap between academic modelling and energy planning.

2.2 RESEARCH OBJECTIVE

This section presents the main research question and sub-questions that guide this thesis. The aim of the research is to add offshore wind energy and electrolysers to a simplified electricity grid model using PyPSA, to support early-stage investment planning in the context of the Dutch energy transition, to validate the results with data from TenneT experts, and to find cost-based trade-offs in the system configuration. The research takes an optimisation based approach and explores whether a simplified model can deliver results that are of sufficient quality to inform policy decisions while remaining computationally efficient. To address this, the following research questions are formulated:

Main question:

> *How can a PyPSA model of the Dutch transmission grid be designed and validated to support optimal capacity expansion planning for the 2040 Koersvaste Middenweg scenario, including offshore wind, electrolysers, and cost-based trade-offs between technologies?*

To answer the main research question, four sub-questions are formulated to ensure a systematic analysis and to address all relevant aspects of the problem.

Sub-questions:

1. What are key indicators to assess the results of electricity capacity expansion planning?
2. How can offshore wind energy and electrolysers be reliably added to the PyPSA model?
3. How can the quality of a simplified PyPSA model be tested against a PLEXOS model?
4. Which trade-offs emerge in the electricity mix under different cost assumptions in the 2040 Koersvaste Middenweg scenario?

The sub-questions break the main research question into four parts, each addressing a key element needed to evaluate the usefulness of a PyPSA model for TenneT. The first identifies which performance indicators are most relevant for addressing electricity grid expansion and investments, providing a clear target for what the model should optimise. The second assesses how offshore wind energy and electrolysers can be added to the PyPSA model while giving reliable results. The third explores if the PyPSA model can be meaningfully compared against the PLEXOS model. For this comparison to be valid, PyPSA should not only produce results that are similar to those of PLEXOS, but it should also respond in a similar way when input parameters are changed. The fourth discusses the policy implications and cost-based trade-offs around system configuration results.

2.3 RELEVANCE

This research can be valuable for different types of organisations if carried out successfully. It addresses a complex societal challenge, applies open-source energy modelling methods in a practical context, and is designed for practical use by a TSO. This section outlines its relevance for these organisations.

2.3.1 MSc Engineering and Policy Analysis relevance

The congestion of the electricity grid is a highly urgent, costly, and complex challenge, a key element of the Engineering and Policy Analysis (EPA) programme. If not addressed properly, it may lead to blackouts, as the grid struggles to function when it's lines are overloaded (Netherlands Enterprice Agency, 2023). Expanding the grid seems like an obvious solution, however it is capital-intensive and time-consuming, making it essential for policymakers to carefully prioritise where reinforcements are most needed and what the trade-offs are in each location. The addition of two new components, offshore electricity and electrolyser integration, make this case extra complex.

An additional source of complexity is the interconnected nature of the electricity grid across European countries. For example, in Spring of 2025, the power systems of Spain and Portugal experienced a black out that also affected parts of France (ENTSOE, 2025). Today, the situation is made more complex by the growing share of intermittent solar and wind energy, rising electrification from electric vehicles and heat pumps, and uncertainty about how the energy system will evolve in the coming decades (Netherlands Enterprice Agency, 2023). Small changes in demand or generation can have large effects on grid stability, which often fluctuates on short timescales. This makes it essential to analyse the system as a whole, considering how different factors influence each other over time.

This thesis is relevant to the EPA programme because it connects a pressing societal challenge, electricity system congestion, to the tools and methods of EPA. By developing a simplified optimisation model of the Dutch electricity grid, the thesis aims to support policy decision making using the systemic and analytic ways of thinking that is taught at the EPA programme. The research applies modelling and stakeholder analysis which are central to the EPA curriculum, to support credible, relevant, and useful policy advice.

2.3.2 Scientific relevance

The main scientific contribution of this thesis is the development and extensive validation of a open-source optimisation model of the Dutch electricity grid using PyPSA. Existing models primarily fixate on 1) the whole European continent and 2) one optimisation, while this thesis focuses on 1) the Netherlands and 2) a range of uncertainties and possible outcomes. The focus on the Netherlands is because of the importance of regionalised data and speed of the model, to support Dutch energy policy decision-making.

In doing so, this thesis supports the broader scientific understanding of how to balance complexity and quality in electricity grid modelling. It aims not only to produce a fast and usable tool for grid and investment planning but also to inform future research on the design of energy models for decision-making under uncertainty and computational feasibility.

2.3.3 TenneT relevance

For TenneT and the government to take effective action against grid congestion, it is important to simulate possible future scenarios in a data-driven way. There is a clear need for fast optimisation models that can evaluate investments in electricity generation technologies and takes into account the power flows of the grid. The PyPSA model developed in this thesis meets this need as it has been carefully built, tested, and validated against existing TenneT models. This, together with the EMA Workbench analysis, allows TenneT to better understand how different technology choices affect system configurations and congestion. As a result, this thesis is not only academically relevant, but also directly useful for real-world planning and decision-making for policymakers.

2.3.4 Societal relevance

This research helps policymakers and others interested in the energy transition better understand what is needed to build a robust future energy system. A key goal of this work is to make complex model results easier to understand. Clear and intuitive figures are used to translate technical modelling outcomes into plain language that can also be understood by people without a technical background. By explaining the results in an accessible way, this research increases trust in the model outcomes and the conclusions drawn from them.

3 RESEARCH DESIGN

This chapter outlines how the research will be conducted. It explains the the methods used to answer each sub-question, the overall design of the study and lastly the sources of data.

3.1 METHODS

Several methods will be used to answer the research questions. These include literature review, expert interviews, and modelling. This section will provide more explanation on the methods for data analysis, data interpretation and data validation of the sub-questions.

- **[SQ1]** What are key indicators to assess the results of electricity capacity expansion planning?
  - A combination of literature review, theoretical frameworks and semi-structured interviews with TenneT experts will be used to identify and assess relevant indicators (Amos et al., 2021). The interviews are especially important because these will tell if the metrics are operationally meaningful and used in TenneT's current models.
  - The indicators from the literature and experts will be thematically categorised based on their purpose (e.g. reliability, capacity, price). Differences between literature-based indicators and TenneT preferences will be explicitly noted.
  - These differences will serve as a cross-validation step. If they show, further interviews will give rationale and prioritisation. The selected indicators will be reviewed with the modelling experts to ensure practical relevance.
- **[SQ2]** How can offshore wind energy and electrolysers be reliably added to the PyPSA model?
  - The offshore wind parks and their corresponding lines will be added to the PyPSA model. Then, electrolysers and hydrogen pipelines will be added to the model. The key performance indicator results that were defined in the previous sub-question will be checked with data from TenneT and Netbeheer Nederland.
  - Results from PLEXOS/Guidehouse and PyPSA will be compared, and if they show both similar outcomes and behaviour, it can be concluded that the models show similar results. Since both the PyPSA and PLEXOS model will then contain offshore wind and electrolysers, a fair comparison can be made.
  - Behaviour from the PyPSA model will be discussed with TenneT experts to see if they match expert judgement.

- **[SQ3]** How can the quality of a simplified PyPSA model be tested against a PLEXOS model?
  - The key indicators from the previous sub-question from the PyPSA model will be compared to the ones from the PLEXOS model from the Guidehouse and TenneT (2025) report. The goal is for the indicators to stay in 10% difference from each other. Furthermore, the Guidehouse and TenneT (2025) does a sensitivity analysis. To test if the PyPSA and PLEXOS models behave the same way, the same sensitivity analysis will be performed to test if the behaviour is the same.
  - Results from PLEXOS/Guidehouse en PyPSA will be compared, and if they show both similar outcomes and behaviour, it can be concluded that the models show similar results. However, the PLEXOS model is built as a single node model, while the later discussed phase 2 and 3 models are on a transmission level. So there is a difference in granularity of the data, which have to be taken into account when comparing the PLEXOS and PyPSA model.
  - Behaviour from the PyPSA model will be discussed with TenneT experts to see if they match expert judgement.
- **[SQ4]** Which trade-offs emerge in the electricity mix under different cost assumptions in the 2040 Koersvaste Middenweg scenario?
  - To test which energy policies cause trade-offs in the optimisation, a tool is needed that can systematically sample different input variables with each other. EMA workbench is a tool developed by Kwakkel (2025) to test decision-making under deep uncertainty. This tool is fitting because it can mix a range of variables and their uncertainty bands for a large number of times and has an explorative nature (Kwakkel, 2025).
  - The EMA workbench shows a range of possible outcomes under uncertainty of input variables. These outcomes can clearly show when cost-based trade-offs in energy technologies take place. These trade-offs can be used to conclude which energy policies are robust.
  - Results from the EMA model will be discussed with TenneT experts to see if they match expert judgement.

An overview of how these research questions are linked together can be found in Figure 3.1.

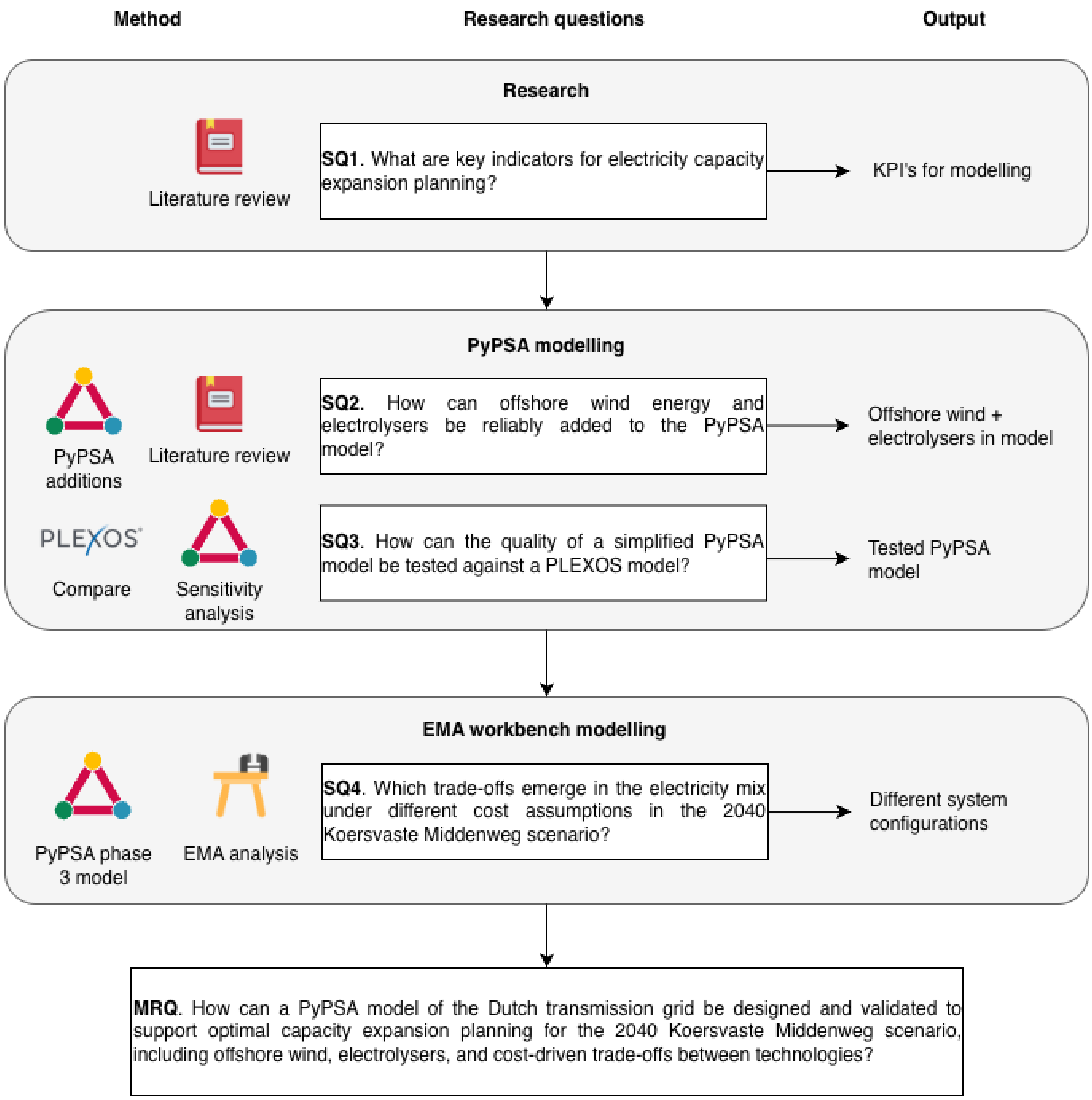


Figure 3.1: Flow diagram of research questions

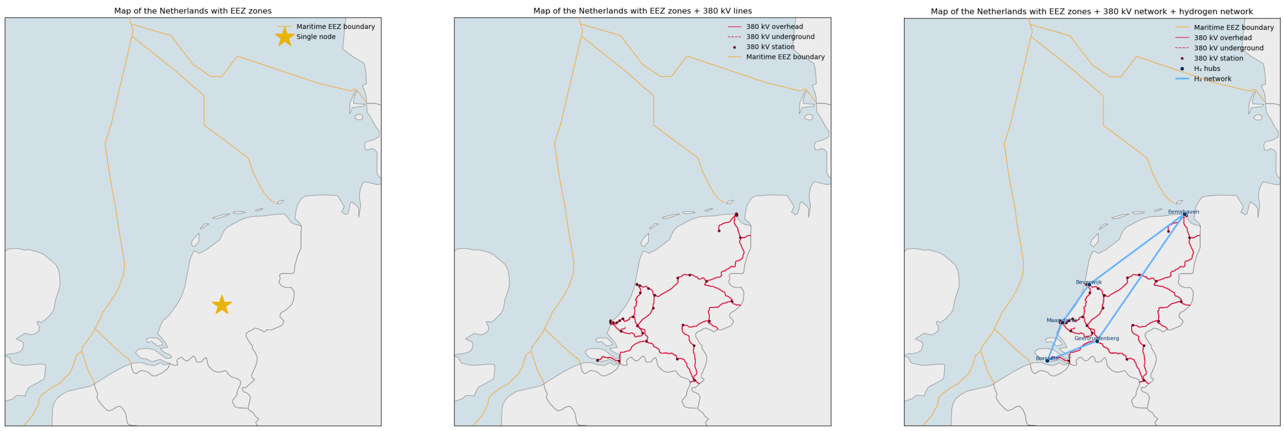
3.2 OVERALL DESIGN OF PHASES

Before starting with the modelling component of the research, short, informal interviews will be taken with TenneT experts to find out what is important when modelling parts of the Dutch electricity grid, and how their current modelling tool, PLEXOS, is used. In parallel, a theoretical background review will analyse other modelling approaches and see which performance indicators are commonly used in research. This ensures the model is both practically relevant and scientifically grounded. Then the modelling part will begin, which will be conducted in three phases.

The **first modelling phase** involves improving and validating the single node PyPSA model which TenneT has already built. A conceptualisation of this phase can be seen in Figure 3.2a. The results from this model are compared with a recent study by Guidehouse and TenneT (2025), which also used PLEXOS as a single-node model and applied almost the same input data to optimise installed capacities for different generation technologies in the Netherlands under the 2040 KM scenario. The goal is to assess whether similar capacity outcomes can be obtained.

The **second modelling phase** follows a similar approach as the first phase, improving and validating, but applied to a transmission model of the Netherlands (380 and 220 kV), illustrated in Figure 3.2b. In this phase, offshore wind will be correctly placed on the map as well. Installed capacity results will be compared with the results from the PLEXOS model, and power flows will be compared to the PowerFactory/PSSE models. Initially, the transmission lines will be modelled with unlimited capacity, effectively functioning as a single node. If this reproduces the results of the first phase, the line capacities will then be adjusted to realistic values. The objective is for the PyPSA model to give installed capacities and load flows consistent with TenneT's results.

In the **third modelling phase**, electrolysers will be added to the transmission network. These electrolysers will be connected to each other in a simple circle topology to create a simplified hydrogen network. A conceptual model of this can be seen in Figure 3.2c. The PyPSA outcomes will again be validated against internal data from TenneT's teams, focusing on both installed capacities and load flows. A sensitivity analysis will be conducted to assess if the PyPSA model has the same behaviour as the PLEXOS/PowerFactory/PSSE models. Together, the three phases aim to deliver a model of the Dutch transmission grid plus electrolysers that can timely help in energy policy decision-making processes.



(a) Map of the Netherlands with single node

(b) Map of the Netherlands with transmission lines

(c) Map of the Netherlands with transmission lines and a simplified hydrogen network

Figure 3.2: Grid models: (a) single node, (b) transmission network and (c) transmission + hydrogen network.

A visualisation of the phases can be seen in Figure 3.3. By involving TenneT in the design process, testing the model across a range of uncertainties, and refining it based on performance, the research aims to develop a tool that produces reliable results, reflects real planning needs, and can be used in practice. This approach supports key research quality standards: validity (the model measures what it intends to), trustworthiness (results are consistent and transparent), and CRELE (credibility, relevance, and legitimacy in the eyes of users and stakeholders).

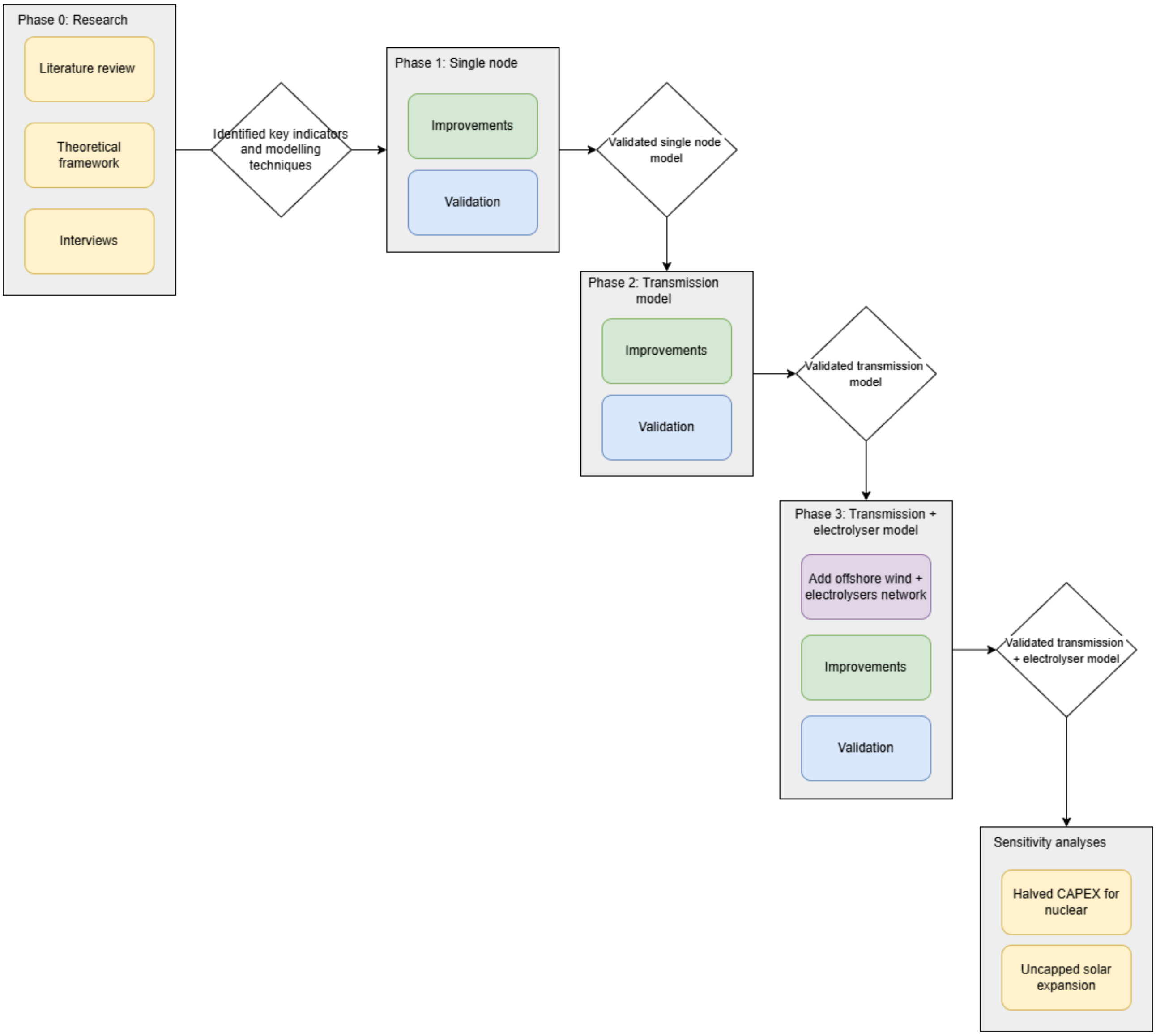


Figure 3.3: Flow diagram of methods

3.3 DATA SOURCES

The main sources of data and knowledge for this thesis will come from TenneT, different groups of people with relevant experience and academic literature.

First, the Energy System Planning teams at TenneT, will play an important role. These teams includes people with different backgrounds. Some focus more on the policy and planning side of electricity systems, while others work more on the technical modelling side. The Market Analysis team works with PLEXOS, while the Grid Planning team works with PowerFactory/PSSE. By interviewing and working with both types of experts, the thesis will gain insights into what the model should be able to do, how it should connect to real-world decisions, and how accurate and useful it needs to be in practice. This will especially be useful for sub-questions one and three.

Second, the developers of PyPSA will also be an important source (Brown et al., 2018). They understand how the model works on a deep technical level and can give guidance on what is possible within the tool, what its strengths are, and where its limits lie. Their documentation and direct communication (if possible) can help clarify how to best set up the model. This will be useful for sub-questions two and three.

Third, people who have already used PyPSA to optimise national electricity systems, for example, in Germany or other EU countries, provide useful examples. Their papers, reports, or shared models can show how PyPSA has been applied in real cases, which design choices were made, and how accurate or useful their results were. Learning from these projects will help improve the quality and design of this thesis' model. Also people who have not used PyPSA, but who have made a optimisations in an open-source model of a national energy system will be able to provide valuable input. This will be useful for all sub-questions.

While broader academic literature on energy system modelling will be reviewed, this thesis prioritises sources directly involved in using or developing PyPSA, as they provide more tool-specific insights.

4 THEORETICAL BACKGROUND

This chapter explains the concepts, theories and definitions that are used for the set up of the thesis. It shows through which lens the models are built. The chapter starts with the identification of the key performance indicators in the literature. Then the role of hydrogen in the electricity mix. Lastly the method with which the trade-offs in the energy system are found.

4.1 ELECTRICITY GRID KEY PERFORMANCE INDICATORS

Many studies have examined which Key Performance Indicator (KPI)'s are most important when analysing and planning electricity grids (Bissell et al., 2014; ENTSOE, 2018; Lehtveer et al., 2021; Mediterranean Transmission System Operators, 2018). These KPI's are usually used to measure how well a power system performs in terms of reliability, cost, and efficiency, and they help grid operators and policymakers make informed decisions.

TenneT, the Dutch transmission system operator, published a study in which they identified grid availability as their most important performance indicator (TenneT, 2019). Grid availability describes how often the electricity network is able to operate without failures or outages. A high level of availability is essential, because it ensures that electricity can be delivered reliably to consumers at all times.

Other studies focus more strongly on cost-related indicators. For example, SUMICSID and Council of European Energy Regulators (2024) highlighted the importance of Capital Expenditures (CAPEX) and Operating Expenditures (OPEX) for different asset types in the electricity grid. These costs are analysed both in a normalised way, to allow fair comparison between technologies, and at the asset-specific level, to understand which components contribute most to total system costs.

A broader perspective is provided by Coordinet (2019), who studied relevant KPI's for TSO's in Spain, Sweden, and Greece. Their work identified several indicators as particularly important, including CAPEX, OPEX, Renewable Energy Sources (RES) integration, reduction of RES curtailment, average cost per service, transaction volumes, and computational runtime. This shows that both technical performance and modelling efficiency are relevant when evaluating grid development options.

In addition, a study by the Mediterranean Transmission System Operators (2018) defined KPI's grouped into several categories. These include transmission line unavailability duration per year, power quality indicators, security of supply, transmission asset utilisation, maintenance cost indicators, and interconnection indicators. Together, these categories provide a structured framework for assessing how well a transmission system performs from both a technical and economic perspective.

These studies show that there is no single KPI that alone fully describes grid performance. Instead, a combination of reliability, cost, renewable integration, and operational efficiency indicators is needed to properly evaluate electricity grid models and support decision-making.

Based on responses from several TSO's, a set of nine KPI's was identified as being widely used in electricity grid operation and planning (Mediterranean Transmission System Operators, 2018):

- Frequency Deviation Index, which measures how stable the grid frequency is.
- Transmission Losses, which indicate how much electricity is lost while being transported through the network.
- Energy Not Supplied, which reflects the amount of electricity demand that could not be met.
- Average Interruption Time, which measures how long outages last on average.
- Interconnection Availability, which shows how often cross-border connections are available for power exchange.
- Transmission Line Availability, which indicates how often transmission lines are operational and able to carry electricity.
- Transmission Transformer Availability, which measures how reliably transformers are available to convert and transport electricity.
- System Average Interruption Frequency Index, which shows how often the average customer experiences a power interruption, and
- System Average Interruption Duration Index, which measures the average total duration of power interruptions experienced by customers.

These indicators mainly focus on system reliability and outage behaviour, which are critical in real-world grid operation. This is similar to a study from ENTSOE (2018), who focused on the costs, residual impacts and benefits, primarily regarding system adequacy and security. However, within the context of this thesis, not all of these KPI's are applicable. The PyPSA model optimises the electricity system for cost efficiency, assuming perfect system operation (PyPSA Docs, 2025b). This means that outages, blackouts, or equipment failures do not occur in the model. If electricity demand cannot be met, PyPSA will always add generation capacity to avoid shortages. As a result, indicators related to security of supply, energy not supplied, and transmission interruptions are not meaningful in this modelling framework. To do this, other specific tools are needed, like PowerFactory/PSSE.

Instead, the KPI's that are most relevant for this study are those related to cost and grid availability. In particular, CAPEX and OPEX are important indicators, as they directly reflect investment and operational decisions. In addition, load flow remains a key indicator, as it shows how effectively the transmission network can be used without congestion. These KPI's align well with the optimisation goals of PyPSA and the research questions addressed in this thesis.

4.2 ROLE OF HYDROGEN IN THE ENERGY MIX

Hydrogen has become an increasingly important topic in the energy transition over the past few years. At one point, it was widely believed that offshore hydrogen production would be necessary to transport energy from offshore wind farms to shore, due to limited electricity grid capacity (Durakovic et al., 2023; Martínez-Gordón et al., 2022). Producing hydrogen at sea would then reduce the need for large offshore electricity cables. However, in recent years this idea has become less dominant, mainly because offshore hydrogen production has not achieved the expected cost reductions, and because the Dutch hydrogen demand hasn't made the growth that was expected (Hermans, 2025). Therefore the government's hydrogen demonstration projects have been put to a halt (Hermans, 2025). Despite this, hydrogen remains an important part of the future energy mix (Ministry of Climate and Green Growth, 2025).

Currently in the Netherlands, around 50 TWh of hydrogen was produced in 2025, mainly using natural gas through steam methane reforming (Netbeheer Nederland, 2025). This hydrogen is currently used mostly in fertiliser production and oil refining. Looking ahead to 2040, hydrogen is expected to play a larger role in additional sectors, including the transport sector, where it may be used as fuel (Netbeheer Nederland, 2025). At the same time, electricity is increasingly expected to be used to produce hydrogen through electrolysis, especially in large industrial areas.

According to the scenarios developed by Netbeheer Nederland (2025), the production of green hydrogen will change significantly over time. Between 2030 and 2035, green hydrogen is assumed to be produced only onshore, using electricity generated by offshore wind, onshore wind, and land-based solar energy. From 2035 onwards, green hydrogen production is expected to gradually move offshore, connected directly to offshore wind turbines or energy hubs. In the long term, by 2050, all green hydrogen production is assumed to take place offshore. Since the hydrogen demonstration projects have been put to a indefinite halt, it is assumed in this thesis that no offshore hydrogen production will take place in 2040 (Hermans, 2025). Instead, hydrogen production is assumed to occur onshore at the landing locations of offshore wind farms. These locations are already well connected to the electricity grid and are therefore logical places to install electrolysers.

The data used in this thesis, based on the ETM KM scenario for 2040, includes hydrogen demand for a wide range of applications (Netbeheer Nederland, 2025). Although most hydrogen today is still produced using fossil fuels, the scenario assumes that all hydrogen demand in 2040 is produced by renewable energies or natural gas. For simplicity, to only model the electricity and hydrogen network in this thesis, it is assumed that all hydrogen is produced using electricity. There is no explicit preference for fossil or renewable energy for hydrogen production.

4.3 TRADE-OFFS IN THE OPTIMISATION

When optimising an energy system, it is unavoidable to make assumptions and simplifications. These choices are necessary to keep the model solvable, but they also influence the results. The optimisation outcome itself does not show how sensitive it is to these modelling choices. As a result, even small changes in assumptions or input values can lead to trade-offs that significantly change the optimal solution (DeCarolis et al., 2016). To better understand these effects and to support well-informed decision-making, several modelling approaches for trade-offs have been developed. Although these approaches work in different ways, they share the same goal: providing more insight into the range of possible solutions rather than focusing on a single 'optimal' outcome. These approaches can broadly be divided into methods that address structural uncertainty and those that address parametric uncertainty.

4.3.1 Modelling to Generate Alternatives

One method used to deal with structural uncertainty is Modelling to Generate Alternatives (MGA). This approach has gained increasing attention in energy system modelling as an extension of traditional multi-objective optimisation and as a way to bridge the gap between simulation and optimisation models (Lombardi et al., 2025; Neumann & Brown, 2023). Instead of focusing only on the single most cost-optimal solution, MGA searches for a set of near-optimal solutions that have almost the same total cost but differ in structure or technology choices (Neumann & Brown, 2021).

These near-optimal solutions allow researchers and policymakers to explore a wider range of system designs. Each alternative can be evaluated using different criteria, such as environmental impact, social acceptance, or system resilience (Prina et al., 2023). Because these solutions are still 'close to cost-optimal', they remain realistic while leaving room for other important objectives beyond cost minimisation (Grochowicz et al., 2023). This makes MGA especially useful for identifying solutions that may be slightly more expensive but offer clear benefits in terms of sustainability or robustness.

4.3.2 EMA Workbench

Another approach focuses on parametric uncertainty, which refers to uncertainty in input values such as technology costs, demand levels, or fuel prices (Loucks et al., 2005). In recent years, there has been growing interest in model based decision support under deep uncertainty, where it is not possible to reliably predict future conditions. The EMA Workbench is designed for this purpose and uses models in an exploratory rather than predictive way (Kwakkel, 2025).

Exploratory modelling also does not aim to predict a single future, but instead examines how different uncertain inputs affect model outcomes (Kwakkel, 2025). The illustration in Figure 4.1 shows the difference between a traditional deterministic model, and one where uncertainty in the input parameters are taken into account (Tennøe et al., 2018). Model A shows that parameters with a fixed input result in a single output of the model (gray). While model B takes into account the distributions of the parameters, which results in a range of outcomes (light grey). The EMA workbench follows the same principle as model B, as it has as output a range of possible values.

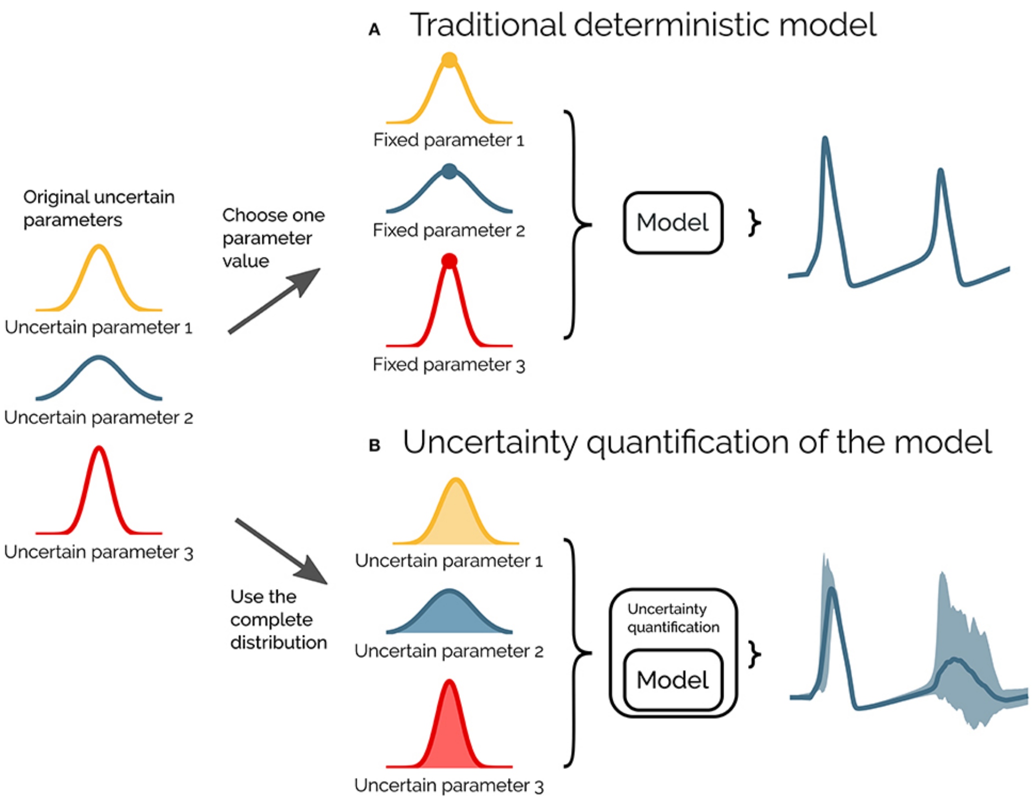


Figure 4.1: Illustration of uncertainties in a deterministic model (Tennøe et al., 2018)

The key difference between deep uncertainty modelling and MGA is the focus of the analysis. In deep uncertainty modelling, the input parameters are systematically varied, while the optimisation structure remains the same. In contrast, MGA keeps the input parameters fixed and explores multiple near-optimal solutions within the same parameter setting. Both approaches provide valuable insights, but they answer different types of questions.

4.3.3 Perspective of TenneT

From the perspective of EPN-SO, it is just as important to understand the results of parametric uncertainty analysis as well as structural uncertainty analysis. As a system operator involved in long-term policy and infrastructure planning, this team places a strong emphasis on robust solutions that perform well under many different future conditions, so both MGA as EMA workbench are valuable tools. However, because the EMA Workbench supports an exploratory approach that aligns well with the purpose of the PyPSA model, and due to time constraints, only the EMA Workbench was applied in this thesis. Future research should also include MGA to capture a wider range of near-optimal possibilities.

5 MODEL CONCEPTUALISATION

This chapter describes the design of the model and explains how its components function within the system. It begins by presenting the key performance indicators, and therefore answering the first sub-question. Next, it describes the structure of the PyPSA model. Finally, the chapter explains the structure of the EMA Workbench. In doing this, this chapter provides an overview of the model's foundation and technical setup.

5.1 KEY PERFORMANCE INDICATORS

To identify the most relevant factors for optimising an electricity system model, three informal interviews were conducted with modelling experts from TenneT. Each expert represented a different sub-team within the overarching Energy Systems Planning department, ensuring that perspectives from across the modelling chain were included. By looking into these KPI's in this section, the first subquestion *"What are key indicators for electricity capacity expansion planning?"* will be answered.

The first team, Systems Outlook, focuses on translating societal needs into a vision for a future-proof electricity system. They identify long-term developments and define the processes required to achieve a robust and resilient grid. The second team, Market Analysis, develops national-scale energy scenarios for the Netherlands and performs optimised capacity assessments at the system (single-node) level. The third team, Grid Planning, builds on these scenarios by introducing spatial detail. They determine where new generation capacity and grid reinforcements are needed and perform corresponding power flow analyses. Across these interviews, the experts consistently highlighted the following KPI's as the most important for evaluating model performance:

- **Installed capacity per generator type**
  Reflects the model's investment decisions in certain technologies. This should be done in a cost-effective and technically feasible manner.
- **Dispatch**
  Indicates how the installed generation assets are operated over time to meet demand. Dispatch results reveal whether the model can reproduce realistic system dynamics, such as the utilisation of base-load and peaking units, curtailment of renewables, and the balancing of variable generation.
- **Electricity price**
  Serves as an aggregated outcome of the system's balance between supply, demand, and operational costs. Price signals are critical for evaluating the economic plausibility of the model

results and their alignment with market behaviour. Unrealistic or highly volatile price patterns may indicate shortcomings in cost assumptions, flexibility modelling, or market representation.

- **Load flows**
  Represent the physical movement of electricity across the grid and is essential for assessing spatial feasibility. Load flows allows analysis of congestion and bottlenecks. This ensures that the optimised system design is not only cost-optimal but also technically viable from a grid perspective.

- **Interconnector capacity**
  Determines the level of cross-border exchange with neighbouring countries. This is crucial for capturing the role of imports and exports in balancing supply and demand in the EU picture. Also has a large influence on the Dutch load flows.

In the previous chapter, the most important KPI's identified in the literature were discussed. Based on this review, it was concluded that cost-related and grid-usage indicators are the most relevant for this thesis. Because of time constraints, the analysis mainly focuses on installed capacities and load flows. These indicators are consistent with those commonly used in the literature and are therefore selected as the main KPI's in this study. They are used to compare, or 'validate' the PyPSA-based results against the results from PLEXOS and PowerFactory/PSSE. Normally, validation is done by comparing model results with historical data to see how well the model can reproduce past developments. In this thesis, however, the results are compared with projections from another model rather than with historical data. It is still referred to as validation because TenneT considers the PLEXOS and PowerFactory/PSSE models to be reliable and well-established references for long-term system analysis.

5.2 PYPSA

PyPSA stands for Python for Power System Analysis and is an open-source Python software used for simulating and optimising power and energy systems. It was developed by a research team at TU Berlin and is designed to handle large electricity networks and long time series efficiently (PyPSA Docs, 2025c). The objective of PyPSA is to minimise total system costs while complying to all constraints of the optimisation problem, such as energy balance and dispatch limits (PyPSA Docs, 2025d). As explained earlier in Chapter 1, PyPSA was chosen for this thesis because the model is used the most in the energy modelling community and because of path dependency: PyPSA is already in development by TenneT.

5.2.1 Input variables

To ensure meaningful analysis, TenneT interviewees emphasised that high-quality and reliable input indicators are required. According to them, the most important input variables are:

- **Annualised capital expenditure**
  Represents the investment costs for building generation assets, recalculated to include lifetime.

- **Marginal costs**
  Covers the costs to produce one extra product.

- **Electricity demand**
  Defines the overall system load that must be met. This data needs to be known for every 380 and 220 kV substation.

- **Current and maximum installed power**
  The capacity that's already in the Netherlands, and up until what it can be expanded to.

- **Weather data**
  Determines the availability and variability of renewable generation such as wind and solar.

The sources for these variables will come from the ETM and the regional system operators, which are both sources that TenneT and Guidehouse also used.

5.2.2 Core design

All functions and components in the PyPSA network are linked to the object: Network. In this thesis, a network has the following components, see Table 5.1. Each component has a set of attributes which contain data types, values, and descriptions for each attribute, which can be found in Appendix A. Since in this thesis, there will be three different modelling phases (single node, transmission and transmission with electrolysers), there will also be 3 different datasets with each of these components.

Components	**Description**
Bus	Node from where components attach
Carrier	Energy carriers of buses (AC/DC, hydrogen or heat), or technologies of components (wind, gas, electrolyser, heat pump)
Generator (p_max_pu)	Power generator for the bus carrier it is attached to
Load (p_set)	Represents demand at the bus they are attached to
Link	Used for controllable directed flows between two or more buses with energy carriers (HVDC, converters, conversions between carriers)
StorageUnit	Enable inter-temporal energy shifting
Line	Include distribution and transmission lines, overhead lines and cables
Snapshots	Time series

Table 5.1: PyPSA network components (PyPSA Docs, 2025a)

Buses
A bus is the basic building block of the network. It is the point where all other components, such as generators, loads, and lines, are connected. At each time step, the bus makes sure that the amount of energy flowing into it is equal to the amount flowing out. A bus can represent an electricity substation, but it can also be used for other energy carriers, such as hydrogen or heat, or even for non-energy carriers like CO_2 or steel (PyPSA Docs, 2025b). In Figure 5.1, a schematic electrical drawing, it's represented as the thick, black horizontal line. In this thesis, every bus has a name, nominal voltage, coordinates and a carrier type.

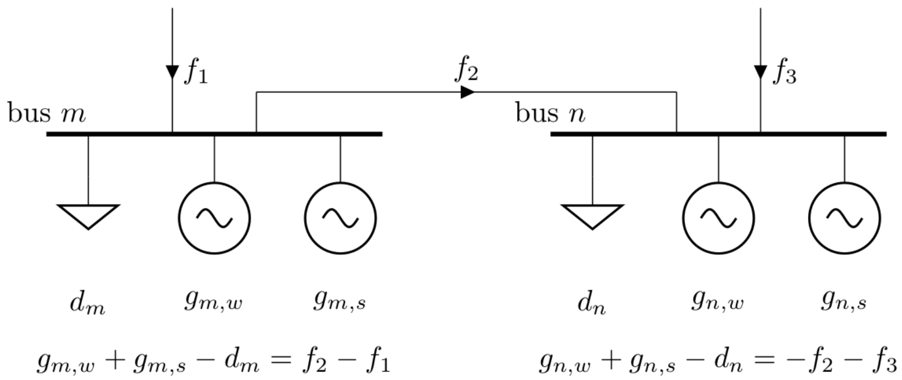


Figure 5.1: Schematic drawing of an electric bus

Carrier

A carrier describes the type of energy or technology used in the model. For buses, this can be different forms of energy such as alternating current (AC), direct current (DC), hydrogen, or heat. In this thesis, each carrier is defined by its name, colour, and CO_2 emissions.

Generators

Generator components are connected to a single bus and are used to supply power to the system. They convert energy from their own carrier, such as wind or gas, into the type of energy used by the bus, in this case electricity. In this thesis, generators include many parameters, of which the most important are the generator name, nominal and maximum power, carrier type, annualised capital costs, and marginal costs.

Loads

Load components are connected to a single bus and represent the demand for the type of energy that the bus carries. In this thesis, each load is defined by its name and the bus it is connected to for each time step.

Storage Units

StorageUnit components are connected to a single bus and allow energy to be stored and used at a later time. This makes it possible to shift energy between different moments in time. This component is suitable for technologies such as batteries. In this thesis, the most important attributes of a StorageUnit include the name, the connected bus, nominal and maximum power, capital costs, and marginal costs.

Conceptually, the system input and output of PyPSA can be seen in Figure 5.2.

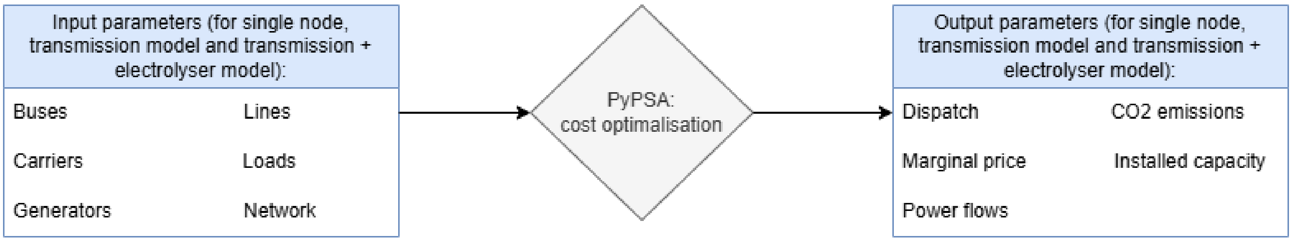


Figure 5.2: Conceptual model of input and output of PyPSA

Using data files
PyPSA can import and process several different data formats, including Excel, CSV, netCDF, and HDF5. In this thesis, CSV files are used as input. After the optimisation is completed, PyPSA produces results, such as statistics on costs, installed capacities, energy supply, curtailment, revenues, market values, energy balances, electricity prices, and transmission flows. It can also generate detailed plots, including energy balance charts, installed capacity overviews, network maps, system balances, and power flow visualisations.

5.2.3 Optimisation

PyPSA can optimise the following types of problems (PyPSA Docs, 2025d):

1. Economic Dispatch market model with unit commitment and storage operation with perfect foresight or rolling horizon, coupled across different energy carriers (electricity, heat, hydrogen, etc.) and with conversion between them.
2. Linear Optimal Power Flow with network constraints for Kirchhoff's Voltage Law (KVL) and Kirchhoff's Current Law (KCL).
3. Security-Constrained Linear Optimal Power Flow for network contingency analysis.
4. Capacity Expansion Planning with single or multiple investment periods and system-wide constraints.
5. Stochastic Optimisation in form of a two-stage stochastic program with investments as first-stage decisions and dispatch as recourse decisions across weighted scenarios of uncertain input parameters.
6. Modelling-to-Generate-Alternatives for near-optimal space exploration.

In this thesis, only the first two types of analysis are considered, because the goal is to compare the PyPSA model output to the PLEXOS one in the 2040 KM scenario. Due to time limitations, the other types of analysis are not included. To do the analysis, costs calculations are made, which are explained in Appendix A.

5.3 EMA WORKBENCH

The EMA (Exploratory Modelling and Analysis) Workbench is an open-source toolkit developed by researchers at TU Delft to support decision-making under deep uncertainty (Kwakkel, 2025). Instead of trying to predict one specific future, exploratory modelling focuses on understanding how different assumptions and decisions lead to different outcomes.

In exploratory modelling, the goal is to see how changes in uncertain inputs or decisions affect the range of possible results. There are two main ways to explore this relationship. The first approach is systematic sampling of the uncertainty or decision space, where many different combinations of input values are tested (Kwakkel, 2025). This is also known as open exploration. The second approach is a directed search, where optimisation methods are used to explore the space in a more targeted way (Kwakkel, 2025). In this thesis, the first approach will be used for the trade-off analysis because of its explorative nature. The aim is to test the robustness of the optimisation, not to find answers for policies.

5.3.1 Framework

The EMA workbench needs various elements to set up, run and analyse the experiments. They will be discussed below:

- **Model**
  A base model is needed in the workbench to be tested a number amount of times.
- **Samplers**
  Different sampling methods are available to generate combinations of uncertain inputs and decisions.
- **Uncertainties**
  Parameter classes that are used to define uncertain inputs or decision levers in the model.
- **Outcomes**
  Classes that describe which model results are recorded and analysed.
- **Evaluators**
  Tools for running experiments either one after another or in parallel to speed up computation.

Model
EMA can be connected to various simulation modelling environments, for example Vensim, Excel, Netlogo, Simio, but for this thesis, Python, specifically the PyPSA model is used.

Samplers
By default, the EMA Workbench uses Latin Hypercube Sampling (LHS) to sample both uncertain parameters and decision levers. In addition to LHS, the workbench also supports other sampling methods, such as full factorial, partial factorial, and Monte Carlo sampling, as well as several sampling techniques provided through SALib.

LHS is a statistical method designed to select samples in a way that evenly covers the entire range of possible input values (Altair, 2026). When sampling a model with N uncertain parameters, the range of each parameter is divided into M equally probable intervals. One sample is then taken from each interval, ensuring that all parts of the input space are represented. The number of intervals M is the same for every parameter, which helps create a balanced and well-distributed set of samples. This can be schematically seen in Figure 5.3.

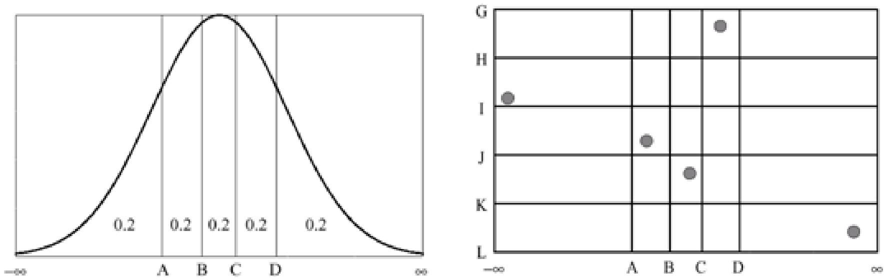


Figure 5.3: Latin Hypercube Sampling (Altair, 2026)

Because of this structured approach, LHS is often used in Monte Carlo analyses, as it can achieve reliable results with far fewer model runs than simple random sampling. The samples are still random, but in a controlled way that avoids clustering and unnecessary repetition.

In this thesis, LHS is used because it provides a good representation of the uncertainty space without redundant simulations. Its main advantage is that it reflects the underlying input distributions well while requiring a relatively small number of model runs compared to basic random sampling.

Uncertainties

To test the parametric uncertainties in the model, uncertainty ranges must be defined for both input variables and policy levers. This means setting a lower and upper bound for each parameter. This can be done with real, integer and categorical parameters. The chosen sampling method then ensures that values within these ranges are selected in a distributed way. Uncertainties can be defined in different forms, such as fractions of a base value, a set of discrete values, or as a continuous range. In this thesis, the uncertainties surrounding investment costs of important technologies are tested.

Outcomes

The outcomes are the variables of interest in the analysis, the KPI's. These outcomes can be single values, such as total system cost, or time series, such as electricity production over time.

Evaluators

Different types of evaluators can be used to run the EMA workbench model. The sequential evaluator runs the EMA workbench in sequence, while the multiprocessing one can work in parallel.

5.3.2 Open exploration

As discussed before, two types of analysis can be done in EMA workbench: open exploration and directed search. For this thesis, open exploration was used because of its explorative nature and ability to test the robustness of the model.

In open exploration, a scenario represents a specific combination of uncertain input values, while a policy represents a specific set of decisions (Kwakkel, 2025). Together, a scenario and a policy form an experiment. LHS is used to generate experiments by sampling across all uncertain variables and decision levers. After the experiments are completed, different analyses can be performed on the results.

Pair plot
A pair plot is a useful tool for visualising relationships between many variables at the same time (Ahluwalia, 2024). It consists of a grid of scatter plots, where each plot shows how two variables relate to each other. In the context of this thesis, the pair plot helps to quickly identify patterns, trends, and trade-offs between technologies and system indicators such as emissions, costs, energy share, and installed capacity. It also makes it easier to spot clusters, correlations, and extreme values. This overview is especially useful when analysing results from many model runs, because it allows complex relationships to be understood at a glance rather than by looking at individual runs one by one.

Feature scoring and Theil-Sen analysis
The EMA Workbench includes a built-in method called feature scoring analysis. This method identifies which model inputs are most important for explaining changes in the results (Kwakkel, 2025). Feature scoring is useful because it shows which technologies or assumptions have the strongest influence on system outcomes. However, for this thesis, knowing only how important a variable is is not enough. It is also important to understand the direction of the effect, whether a change in a variable makes an indicator increase or decrease.

To capture this, the Theil–Sen analysis is used. It is a type of linear regression that estimates trends in the data, but is less sensitive to extreme values than normal linear regression analysis (Satpati & Siddarth, 2023). It does this by looking at the median of many possible slopes instead of fitting one line through all data points. As a result, it not only shows whether a variable has a strong influence, but also whether that influence is positive or negative.

Archetype plot
Looking at all the EMA Workbench results, specifically the total installed capacities, one by one would be confusing and hard to interpret. To make the results easier to understand, the many system configurations need to be grouped into a smaller number of representative cases. This is why clustering with K-means is used. It is an unsupervised machine learning method that divides a dataset into a chosen number of groups, called clusters, or in this thesis, archetypes (Abhay, 2025). The idea is to group together system configurations that look alike, while keeping different types of systems in separate clusters. In this way, the main patterns in the results become visible.

The algorithm works by assigning each data point to the cluster with the nearest centre, called the centroid. It then repeatedly adjusts these centroids so that the total distance between data points and their assigned centroid is as small as possible. As a result, data points within the same cluster are very similar, while data points in different clusters are clearly different. These archetypes capture the main ways in which the electricity system can be configured under uncertainty, making the EMA results much clearer and easier to interpret.

6 MODEL FORMALISATION

This chapter explains how the PyPSA and EMA Workbench models are set up. It describes the steps taken to prepare the data, build the models, and define the main assumptions. The goal is to show how both models are designed and what choices were made to make them work. Lastly, it discusses the data assumptions that needed to be made in order to make the models.

6.1 PYPSA SET-UP

The PyPSA model is analysed in three phases. Although each phase uses different datasets for elements such as buses, loads, and lines, several inputs and assumptions remain the same across all phases. These common aspects are discussed below.

Annualised costs
The PyPSA model uses annualised costs of generators as input for its generators. These values were not directly available in the Guidehouse and TenneT (2025) report, so they were calculated manually. Although newer versions of PyPSA now include a function for this, at the time of coding, the calculations had to be done beforehand using the information provided in the report, such as capital costs, lifetime, discount rate, and OPEX.

Normally, building an energy asset such as a generator, line, or storage unit involves a large one-time investment at an operational lifetime of many years. If this full investment were added to a single model year, that year would appear unrealistically expensive. To avoid this, PyPSA annualises the capital cost by spreading the total investment evenly across its lifetime and adjusting it for the discount rate, which represents the value of money over time. This process is done using the annuity factor, which converts the total capital cost into an equivalent yearly cost. The result is then added to the annual OPEX costs, giving the total annual cost of owning and operating the asset.

By converting all investment and operational costs into the same per year unit (euros per year), PyPSA can fairly compare technologies and optimise between them within its objective function. This ensures that investment decisions in the model reflect both the upfront costs and the long-term financial implications of each asset, see Equation 6.1.

$$\text{Annualised costs} = \text{capital costs} \times \left(\frac{\text{discount rate}}{1 - (1 + \text{discount rate})^{-\text{lifetime}}} \right) + \text{OPEX}_{\text{annualised}} \tag{6.1}$$

Marginal costs
The Guidehouse and TenneT (2025) report did also not directly provide marginal costs, but it did include all the information needed to calculate them. Marginal costs describe the extra costs of producing one more unit of electricity, one extra MWh.

In the PyPSA model, marginal costs represent the operational costs that depend on how much a generator is used. These include several elements:

- Fuel costs, which apply to generators such as gas, coal, or biomass plants. These costs depend on the fuel price and the efficiency of the generator.
- CO_2 emission costs, which reflect the price for emitting CO_2.
- Variable operation and maintenance (O&M) costs, which account for wear and tear that increase with the amount of energy produced.

This means that every time a generator produces electricity, it adds to the total cost through these components. In contrast, capital costs are paid only once when the asset is built and are then spread out over its lifetime through annualisation.

By including marginal costs, PyPSA can use this to decide which generators to use and when. Generators with low marginal costs, such as wind and solar, are used first because they are cheaper to operate, while plants with high marginal costs, like gas or coal, are only used when necessary to meet demand. This makes marginal costs an important factor in how PyPSA models the energy system and determines the most cost-efficient way to supply electricity. The formula for marginal costs can be seen in Equation 6.2.

$$\text{Marginal costs} = \frac{\text{fuel price}}{\text{efficiency}} + \frac{\text{CO}_2\text{ intensity}}{\text{efficiency}} \times \text{CO}_2\text{ price} \tag{6.2}$$

Maximum per-unit availability
Renewable energy sources such as wind and solar depend directly on weather conditions. Unlike gas or coal plants, which can produce electricity whenever needed, wind turbines and solar panels can only generate power when the wind blows or the sun shines. This means their electricity output is very dependent on the weather. To capture this behaviour in PyPSA, the model needs information about how much energy these technologies can produce over time. This is where maximum per-unit availability data comes in.

The maximum per-unit availability shows the share of the installed capacity that is technically available at a given time. For example, if a wind turbine has an installed capacity of 10 MW and the maximum per-unit availability in a certain hour is 0.5, then at most 5 MW can be produced in that hour. These hourly availability values are based on weather data and reflect how wind speeds and solar radiation change throughout the year. By using these time series, PyPSA determines the maximum possible output of each renewable generator in every hour of the model.

In this thesis for solar energy, maximum per-unit availability data from the regional system operators are used. This data is the only not publicly available data of this thesis, because this data is more detailed than the ETM data that was used for the rest of the model. For wind energy, the ETM data is used. For simplicity, the same standardised maximum per-unit availability data profiles are applied to all wind and solar parks in the model. This means that every park experiences the same hourly pattern of generation, regardless of its location. In reality, conditions vary between regions. Wind speeds near the sea are generally higher and more consistent than inland, and solar radiation can differ between the north and south. Ideally, separate profiles would be used for each region to capture these differences and make the model more accurate. However, due to time limitations, this was not implemented in this thesis. Future work could include region specific maximum per-unit availability data to better reflect local weather conditions and their effect on renewable generation, as well as wake effects for wind farms.

Maximum capacity for technologies

In the ETM, maximum limits, or caps, are set for several technologies based on practical and policy assumptions. For example, it is assumed that onshore wind capacity will not expand further than 11 GW. Solar capacity is limited to 131 GW, mainly because of the estimated amount of available rooftop space. Nuclear energy is capped at 3.3 GW, which corresponds to the construction of two new nuclear power plants (Antea Group, 2025). The same caps were also applied in this thesis to ensure consistency with the ETM and Guidehouse and TenneT (2025) assumptions.

6.1.1 Phase 1: Single node

In the first phase, the model optimises the installed electricity generation capacities for the Netherlands represented as one single node, meaning that no regional differences are considered, see Figure 6.1. This setup follows the same approach used by Guidehouse and TenneT (2025) in their optimisation study in PLEXOS. To ensure a fair comparison, almost all the input data is the same. The ETM KM scenario from Netbeheer Nederland (2025) was used. This includes both energy supply and energy demand data. One difference is the data for solar energy, as previously explained in Section 6.1. However, an important note is that the PLEXOS model from the Guidehouse and TenneT (2025) research does include electrolysers, while the PyPSA single node model does not, as this will only be added in the third phase. This is discussed further in the discussions chapter.

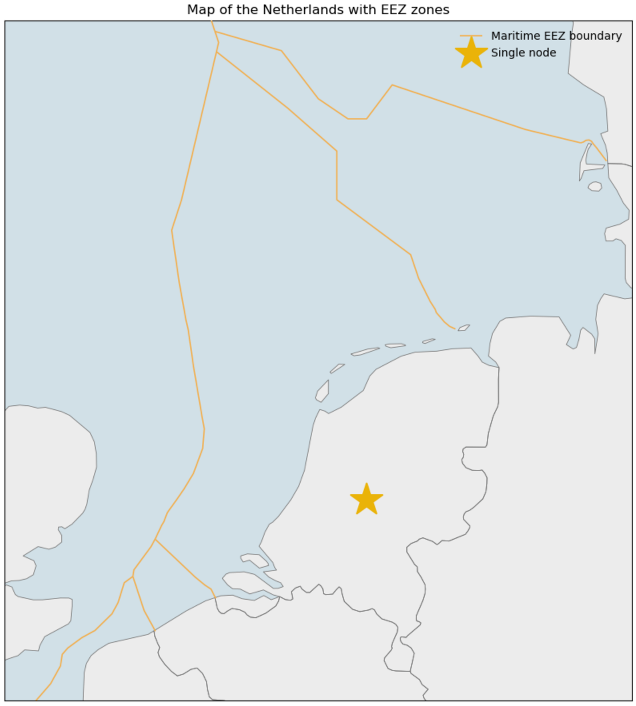


Figure 6.1: Map of the Netherlands as a single node

6.1.2 Phase 2: Transmission network

In phase 2, the Netherlands is represented in the model with its transmission network. This includes all network expansions that are considered almost ready (status G2) as of January 2026 (TenneT, 2026b). By modelling the transmission grid in this way, the model can show how electricity flows between different regions in the country and how the grid balances electricity generation and demand across locations.

The data for the substations, their connections, and their geographical coordinates comes from researchers from TU Delft, Zomerdijk et al. (2022), who created a detailed model of the Dutch transmission grid. Their work includes the full grid, covering voltage levels of 110 kV, 150 kV, 220 kV, and 380 kV. However, for this thesis only the 220 kV and 380 kV networks are included. Lower voltage levels were left out to reduce computing time and keep the model manageable. This TU Delft dataset was used because TenneT does not currently have this type of grid data readily available in a format that can be directly used in PyPSA. Using this existing research made it possible to build a realistic transmission model within the time limits. A few additions have been made manually to reflect the current status of the grid. These are substations Veenoord Boerdijk (VOB), Musselkanaal (MSK), Halsteren (HST), Tilburg (TLB) and Boxmeer (BMR).

To make the optimisation process manageable and reduce computing time, several assumptions and simplifications have to be applied to the network model. These include clustering the substations, finding coordinates for offshore wind parks, aggregating time series and using an academic solver, Gurobi.

Clustering substations

To simplify the model and reduce computing time, some 380 and 220 kV substations have been clustered together. This means combining several nearby substations into a single representative node in the model. The substations that were clustered are those located close to each other, such as the four substations in the Eemshaven industrial area that are clustered to one (Eemshaven (EEM)), see Figure 6.2.

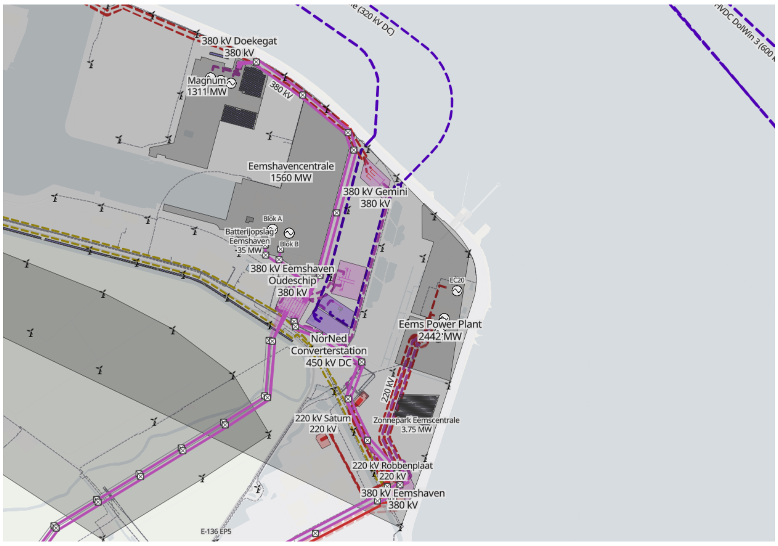


Figure 6.2: 380 and 220 kV substations in Eemshaven

The three substations that were removed are Gemini Onshore (WGO), Eemshaven Filter (ECF) and Doekegat (DKG). Because these substations are part of the same industrial zone and serve similar functions, merging them into one node does not significantly affect the overall electricity flows in the Netherlands. There are other industrial area's where substations have been clustered. These can be found in Table 6.1. The first column describes the area, the second column describes the substation that the removed substations (third column) are clustered to.

Industrial area	Main substation	Removed substations
Eemshaven	Eemshaven (EEM)	Gemini Onshore (WGO) Eenshaven Filter (ECF) Doekegat (DKG)
Rotterdam	Maasvlakte (MVL)	Hollandse Kust Zuid Landstation (HZL) Shell Holland Hydrogen 1 (SHH*)
Amsterdam	Beverwijk (BVW)	Wijk aan Zee (WAZ)
Borssele	Borssele (BSL)	Borssele Ofsshore (BSL*)

Table 6.1: Overview of clustered substations in PyPSA model

This clustering approach is preferred over typical techniques like k-means, because it is based on load flow relevance and network redundancy rather than purely geographic distance. The simplifications were made with an understanding of how the grid functions electrically, ensuring that the key load flows and potential bottlenecks in the transmission system are accurately represented in the model. A map of the clustered substations and the transmission network can be found in Figure 6.3.

Figure 6.3: Map of the Netherlands with transmission network

Coordinates for offshore wind parks

In the second phase of the model, and to partly answer the second research question: *'How can offshore wind energy and electrolysers be reliably added to the PyPSA model?'*, it was important to correctly assign the landing locations for all offshore wind parks. These landing points determine where the power from the sea enters the onshore transmission network. Assigning them correctly is important because it directly affects how electricity flows through the grid and where bottlenecks or congestion might occur.

In May 2025, the Dutch government published a report defining the official landing locations for all new offshore wind parks (Ministry of Economic Affairs and Climate, 2025). In some cases, such as for the upcoming Hollandse Kust West, the new transmission connections go to another substation instead of Hollandse Kust. However, in order to reduce computational time, these substations were clustered, like explained in the previous paragraph. In this case, the substations for Hollandse Kust Noord and West were merged with the Beverwijk cluster, since they are located in the same area and this substitution does not significantly affect the accuracy of the load flow results.

Next, each offshore wind park had to be placed in the correct geographical location within the model. This means that the wind parks were assigned to its actual position on the map. While many wind park areas are publicly available, their exact x- and y-coordinates are not directly accessible online. Therefore, several sources were combined to estimate their positions as accurately as possible. The process began by consulting the Open Infrastructure Map, which shows existing and planned energy infrastructure. When offshore wind parks or substations were visible there, their coordinates were noted down directly. The distance between each offshore wind park and its corresponding onshore substation (landing point) was then measured using the measure tool in TenneT's GeoWeb platform. To maintain a margin of error, estimates have been rounded up to the nearest multiple of one hundred. These distances were entered into the model's lines or links data files to correctly represent cable lengths.

If newer offshore wind parks or substations were not yet shown on the Open Infrastructure Map, additional sources were used, such as Rijkswaterstaat (2025) and Ministry of Infrastructure and Water Management (2025). The maps from these reports were overlayerd with Google Maps, and a location point was manually selected to estimate the coordinates. The distance between these estimated offshore points and their landing locations was again measured in GeoWeb and added to the lines or links file.

A complete overview of the offshore wind park's coordinates, landing location, and cable distance is presented in Appendix A. Wind parks marked with an asterisk (*) are those for which coordinates were estimated using Google Maps rather than the Open Infrastructure Map. These coordinates are therefore less precise, but still close enough to give a realistic representation of their location and distance to shore.

Optimisation solver Gurobi

The PyPSA model was initially set up to perform its optimisation using the HiGHS solver, an open-source optimisation tool. However, after several test runs, it became clear that Gurobi could complete the same optimisation tasks much faster and more efficiently. Research by Parzen et al. (2022) compared HiGHS to other solvers and found that while HiGHS performs well on small or medium-sized problems and is even competitive with Gurobi in those cases, its performance drops significantly for larger, real-world optimisation problems, see Figure 6.4. According to their study, HiGHS can be 60 to 100 times slower than Gurobi when solving large-scale models. Since the PyPSA model used in this thesis is of bigger size and complexity, HiGHS was found to be too slow for efficient use. Therefore the choice was made to use Gurobi. Tests have been done to compare HiGHS and Gurobi results, these can be found in Appendix B.

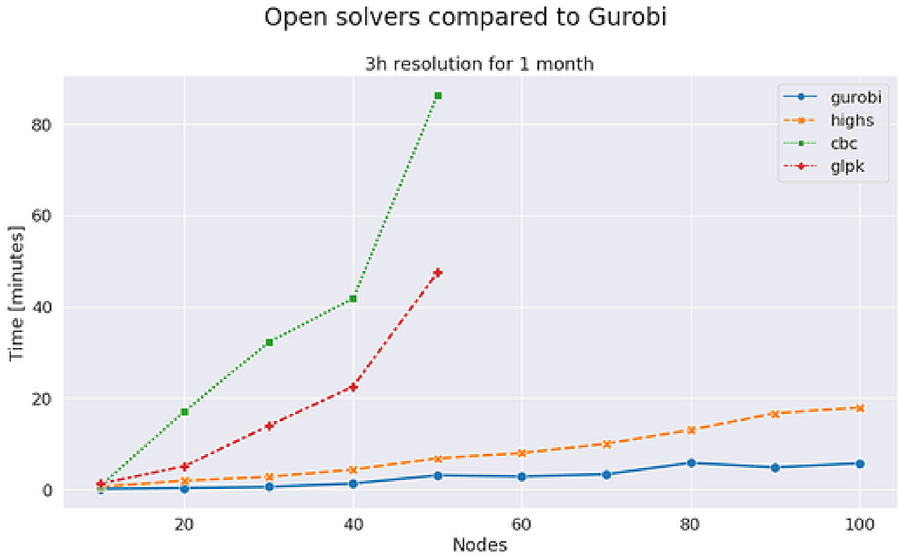


Figure 6.4: Comparison of optimisation solvers

Gurobi Optimizer is a commercial software tool that specialises in solving mathematical optimisation problems such as linear programming (LP), mixed-integer programming (MIP), and quadratic programming (QP) (Gurobi, 2025). With its developed Python interface, Gurobi integrates well with modelling frameworks like PyPSA. The solver is known for its speed, scalability, and robustness, which makes it a common choice in both academic and industrial energy system modelling.

Several studies have demonstrated Gurobi's effectiveness in energy system modelling. For example, Kwon and Hofmann (2023) applied Gurobi to solve a mixed-integer linear programming (MILP) model for optimising a district heating network. The solver successfully handled a complex system with multiple technologies, including combined heat and power plants, heat storage, and diverse operational constraints. Similarly, Parzen (2021) highlighted Gurobi's ability to process large-scale optimisation scenarios efficiently. They showed that combining Gurobi with time series aggregation techniques can significantly reduce computation time while maintaining model accuracy which is a key advantage for large energy system models like PyPSA.

To ensure consistency, the solver settings used for Gurobi were matched to those previously applied with HiGHS. The primal feasibility tolerance was set to 1×10^4, and the dual feasibility tolerance was set to 1×10^5. The solver was configured to use seven threads. These parameters provided a balance between solution accuracy and computational efficiency, allowing the PyPSA optimisation to run effectively with Gurobi.

Aggregating time steps

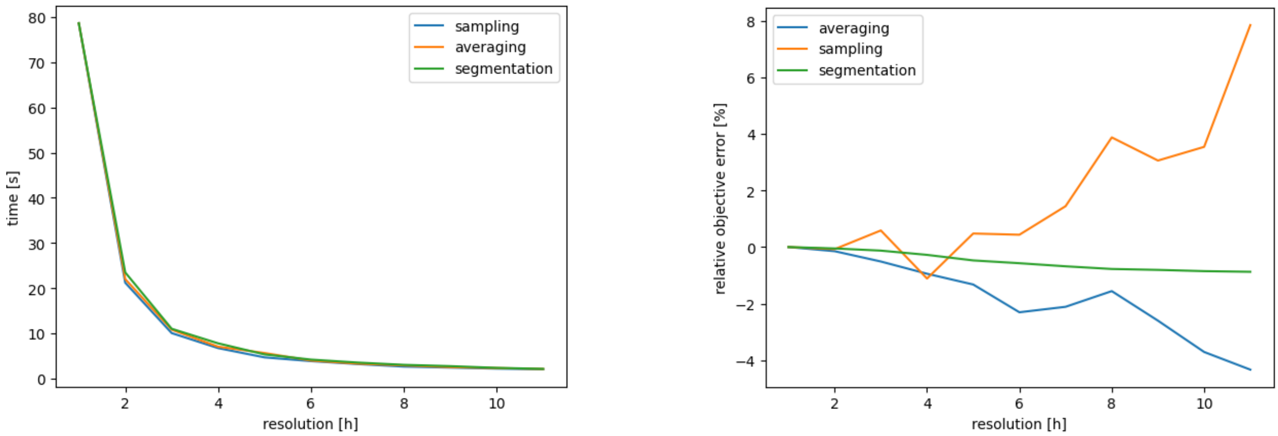
Because the optimisation model requires a large number of hourly input values, running full-year simulations can take a very long time. This makes it difficult to test different scenarios or perform multiple runs. To reduce the run time while still keeping the results realistic, it was necessary to use time series aggregation. Time aggregation reduces the number of hours the model has to process by selecting or combining representative time periods.

At first, the time aggregation method developed by TenneT was used. Their approach selects only a few representative periods from each year. The initial version of the model used one full week per month, resulting in 12 representative weeks. Later, this was changed to two days per month, which further reduced the number of snapshots. Both methods significantly speed up the optimisation, because the solver only has to work with a small subset of the original 8760 hours.

As work on this thesis progressed, PyPSA released three new official methods for time series aggregation:

- Sampling: selecting snapshots based on a fixed frequency (e.g. every 10th hour).
- Averaging: combining snapshots by averaging the values within a given time interval.
- Segmentation: clustering snapshots into groups of similar conditions and selecting a representative snapshot for each group.

PyPSA evaluated all three techniques and found that, although they all reduce run time by a similar amount, they differ in accuracy, see Figure 6.5 (PyPSA Docs, 2025b). Their tests showed that Segmentation consistently produced the lowest modelling error, meaning it best preserves the original time patterns of demand and renewable generation. Because of this, Segmentation is chosen as the time aggregation method in this thesis.



(a) Differences in time when aggregating time steps

(b) Differences in error when aggregating time steps

Figure 6.5: Differences in time series aggregation techniques (PyPSA Docs, 2025b)

6.1.3 Phase 3: Transmission network + electrolysers

This part discusses the final part of the second research question: *'How can offshore wind energy and electrolysers be reliably added to the PyPSA model?'*. Discussions with experts at TenneT made it clear that the most likely locations for future electrolysers are the landing points of offshore wind parks. These sites already have strong grid connections and large amounts of incoming renewable electricity. In the Netherlands, the landing locations are Eemshaven, Geertruidenberg, Maasvlakte, Borssele, and Beverwijk. In this thesis, these five locations are therefore modelled as the sites where electrolysers can potentially be built, see Figure 6.6.

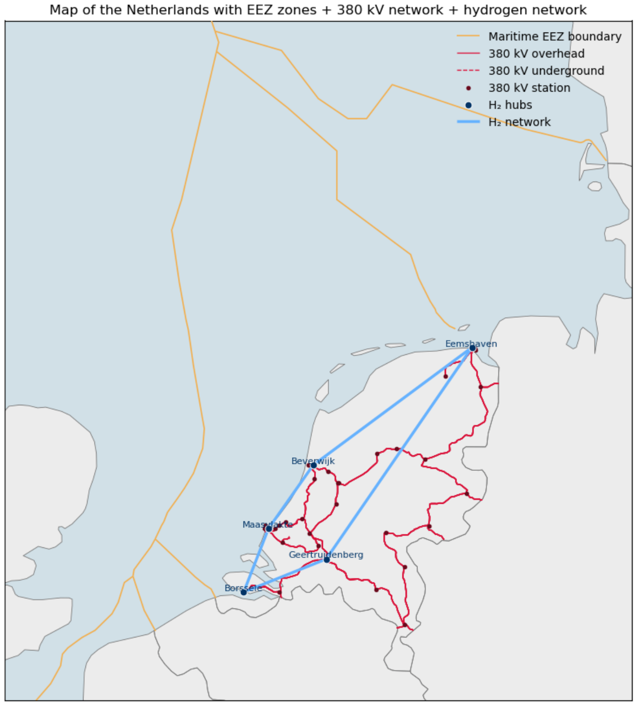


Figure 6.6: Map of the Netherlands for transmission network + electrolysers

Since the focus of this thesis is on where electrolyser capacity should be placed to reduce electricity congestion and not on modelling the full hydrogen system, the electrolysers are connected to each other using a simple circle topology. This allows hydrogen to move freely between the electrolyser locations without modelling a detailed hydrogen pipeline network. With this setup, PyPSA calculates which locations for electrolysers are most cost-effective, based on the availability of electricity at each site.

The hydrogen demand of the electrolysers is placed on the same node where each electrolyser is located. A fully realistic hydrogen network with separate producers, consumers, and regional hydrogen flows is not created, because building such a network would require much more time and is not the main focus of the thesis.

The time-series data for electrolyser loads comes from the ETM (Energy Transition Model, 2025). This dataset provides hourly supply and demand profiles for the 2040 KM scenario. These profiles are used as the expected hydrogen demand for the electrolysers in the model. For simplicity, the total hydrogen load profile will be divided into five, and evenly distributed over all possible electrolyser nodes. Just like for electricity, there's also hydrogen storage built into the model, so that supply and demand do not have to take place at exactly the same time.

6.2 EMA WORKBENCH SET-UP

When building the EMA Workbench model, several assumptions had to be made to keep the analysis focused and manageable.

6.2.1 Uncertainty assumptions

One important decision was to not include electricity demand (loads) as an uncertainty. If demand were varied, the effects of changing technology investment prices would become harder to see. For this thesis, the main interest is to understand how investment cost changes lead to trade-offs in the energy system, because these costs are something the government can directly influence through subsidies, taxes, or policy choices. In a more holistic or long-term study, it would be useful to also include load uncertainty, but for this work it would distract from the core objective.

In the EMA Workbench setup, the uncertainties applied are CAPEX variations. CAPEX is expected to have a large influence on system decisions and trade-offs, especially when combining technologies. Guidehouse and TenneT (2025) has researched the uncertainty ranges for the CAPEX for non-mature technologies, see Table 6.2. However, for a few technologies these are non-existent even though their costs are still very variable. With these input numbers, the model cannot show changes in system behaviour, making it difficult to identify trade-offs between technologies. In order to see more distinct results, the choice has been made to do for every extendable technology +40% and -20%/ on the CAPEX. That is for solar, offshore wind, gas CCGT and OCGT, batteries and nuclear energy. The lower bound of –20% is based on the lower bound estimates reported by Guidehouse and TenneT (2025). The higher bound of +40% is chosen to reflect the expectation that future costs are more likely to increase than decrease, for example due to delays in large-scale R&D or deployment. By setting the upper bound twice as large as the lower bound, the analysis allows for more visible shifts in system configurations and trade-offs.

Carrier	Used CAPEX by Guidehouse (€/kW)	Upper bound	Lower bound
Solar	275	590 (+114.5%)	200 (-27.3%)
Wind offshore	2900	2900 (0%)	2200 (-24.1%)
Battery (4h)	70	160 (+128.6%)	70 (0%)
Nuclear	12000	12000 (0%)	12000 (0%)

Table 6.2: Uncertainty ranges for non-mature energy technologies Guidehouse and TenneT (2025).

Using these uncertainty ranges allows the EMA Workbench to explore how different cost developments affect the optimal design of the energy system and to identify the conditions under which certain technologies become more or less attractive.

6.2.2 DelftBlue

Running the EMA Workbench model with the PyPSA model can take a very long time. Depending on the model setup, a run can take several weeks. Because of these long runtimes, the TU Delft supercomputer DelftBlue was used to carry out the EMA simulations (DelftBlue, 2026). As a student, access to DelftBlue came with certain limits. Each job could run for a maximum of 24 consecutive hours, use up to 32 CPU cores, and was allowed 3900 MB of memory per CPU. These limits influenced how the simulations were set up and how many runs could be performed at the same time. The single node model could be ran for 1000 times, and the transmission models could be ran for 25 times.

To speed up the modelling process, an attempt was made to use the parallel evaluator instead of the sequential evaluator, as discussed in Section 5.3. The idea was that multiple model runs could be executed at the same time, instead of one after another, which would significantly reduce computation time. However, this approach did not work as expected. Further research and debugging are needed to make the parallel evaluator function properly and to benefit from faster computation. For now, the use of DelftBlue and the sequential evaluator provided a solution to speed up the simulations. It reduced the total runtime and made the large number of model runs manageable.

6.2.3 EMA Workbench code

The EMA workbench code can be found on Github: https://github.com/TLisaaa/EMA-TenneT.

6.3 DATA ASSUMPTIONS

Several data assumptions were made in this thesis, which influence the results and should therefore be discussed.

First, different data sources were used for maximum solar output in the PyPSA and PLEXOS models. The PyPSA model uses maximum solar output data from regional system operators, while the Guidehouse study relies on data from the ETM (Energy Transition Model, 2025; Guidehouse & TenneT, 2025). The reason for using regional operator data in PyPSA is that this research focuses on regional grid planning, while the ETM only provides data for the Netherlands as a whole and does not include regional detail. However, using different data sources makes it more difficult to directly compare the PyPSA and PLEXOS results. Ideally, both models should have used the same input data, but in that case, the ETM data would have needed to be regionalised before being used in PyPSA. Part of the difference in (solar) capacity outcomes may be caused by the input data itself, rather than by differences in how the models optimise the system.

Second, the exact locations of offshore wind farms were not always known. For some wind farms, their location could be identified clearly, but for others this required assumptions and estimations. The exact length and therefore costs of the offshore wind cables have an uncertainty bandwidth, and are likely more expensive than assumed for this thesis.

Third, there is limited publicly available data on the geographical distribution of hydrogen demand in the Netherlands. Because of this, hydrogen loads were assumed to be evenly distributed across the landing points of offshore wind farms, also the points where hydrogen production is. The total hydrogen demand is taken from the ETM, but the spatial distribution is simplified. In addition, the model allows unlimited hydrogen transport between electrolysers, which does not reflect real-world constraints in hydrogen infrastructure.

Fourth, the same weather data is used for the entire Netherlands. In reality, wind speeds and solar irradiation differ by location. Also, only one weather year is used in the model. Weather conditions can vary significantly from year to year, so using a single year does not capture long-term variability.

Fifth, all hydrogen production in the model is assumed to be fully electrified. Hydrogen is therefore produced only through electrolysis and not through natural gas–based processes such as steam methane reforming. This assumption leads to higher electricity demand than is currently realistic, especially in the short and medium term.

Sixth, the model uses the current electricity grid layout. Planned and ongoing grid expansions are not included. As a result, the model likely overestimates grid congestion and line overloads in 2040, because in reality the grid will be reinforced before then.

Seventh, phases 1 and 2 of the PyPSA model do not include electrolysers, as these are only added in phase 3. However, later in the research it became clear that the single-node PLEXOS model already includes electrolysers. For a fair comparison between the two models, electrolysers should therefore also be included in phase 1 of the PyPSA model. Because they are currently not included in the earlier phases, part of the electricity demand may be missing in phase 1 of the PyPSA model. This could affect the calculated capacity mix and system costs, since electrolysers are expected to increase electricity demand and influence how generation and flexibility technologies are deployed.

Eighth, in the PyPSA model the interconnector capacity and the installed capacities of neighbouring countries are treated as fixed input values. This means they are predefined and do not change during the optimisation. In contrast, the PLEXOS model optimises these capacities at the same time as the Dutch generation mix, allowing them to influence each other. This approach is more realistic, because cross-border capacity and neighbouring systems can affect national investment decisions. However, optimising everything simultaneously also increases the size and complexity of the model, which leads to much longer computation times. Since one of the goals of this thesis is to reduce runtime and create a faster exploratory tool, these elements were simplified in the PyPSA setup.

Ninth, the PyPSA model optimises the system only for the target year 2040. In contrast, the PLEXOS model performs the optimisation for multiple years, 2030, 2035, and 2040, and does this in an incremental way. This means that investments can be added step by step over time, which reflects how infrastructure is typically developed in reality. By spreading investments across several years, costs are distributed over a longer period. This can influence which technologies are cost-effective at each stage. Since the PyPSA model focuses only on 2040 and does not represent this gradual build-up, the calculated investment mix may differ from the PLEXOS results, where earlier investment decisions affect later outcomes.

Finally, the model does not include CO_2 constraints such as emission caps or the Emissions Trading System (ETS). This means that emissions are only influenced indirectly through technology costs, and not through explicit climate policy instruments. Furthermore, in the PyPSA model, the curtailment of solar and wind energy are free. While in the PLEXOS model, complex calculations are used to determine its curtailment price.

Together, these assumptions simplify the modelling process and make the analysis feasible, but they also limit how closely the results reflect real-world conditions. They should therefore be taken into account when interpreting the outcomes of this study. An overview of the most important differences between the model input in PyPSA and PLEXOS can be found in Table 6.3.

Subject	PyPSA	PLEXOS
Solar data	Data from regional system operators.	Data from ETM.
Electrolysers	Phase 1 & 2: no electrolysers. Phase 3: does include electrolysers.	Single node: includes electrolysers.
Interconnector capacity	The capacity is taken out of PLEXOS and used as input. Does not increase or decrease.	The capacity is the output. Capacities from all countries get optimised at the same time.
Installed capacities neighbouring countries	The capacity is taken out of PLEXOS and used as input. Does not increase or decrease.	The capacity is the output. Capacities from all countries get optimised at the same time.
Calculated years	Only 2040. Investments are made once.	Calculated for 2030, 2035 and 2040. Investments are made incrementally.
Weather year	One weather year.	Multiple and expected weather years.
Curtailment of RES	Is free.	Costs money.

Table 6.3: Differences input data in PyPSA and PLEXOS

7 RESULTS

This chapter presents and discusses the results from the PyPSA and EMA workbench models. The first part focuses on the results from the PyPSA model and discusses them step by step for each modelling phase. The focus will be on the most complete phase, phase 3. After this, the sensitivity analysis of the PyPSA model is discussed. This analysis tests whether PyPSA responds to changes in input assumptions in the same way as the PLEXOS model, by applying the same tests to both models. The last part of the chapter presents the results from the EMA workbench, again per phase. Together, these results provide the basis for the discussion and conclusions presented in the following chapters.

7.1 PYPSA RESULTS

In this section, the results from all phases of the PyPSA model are discussed and compared with data from TenneT. In the first phase, the model output in terms of optimised installed capacities is compared with the results from the Guidehouse and TenneT (2025) report, that was made with the Market Analysis team and the PLEXOS model. In the second phase, the optimised capacities are again compared with the Guidehouse and TenneT (2025) results, while the load flows are compared with data from TenneT's Grid Planning team and the PowerFactory/PSSE model. The same type of comparison is also made for the third and final phase. This last phase is discussed in more detail in this section, because it represents the most complete version of the model.

7.1.1 Improvement runtime PyPSA

An effort was made to reduce the runtime of the PyPSA model, as computation time increased with each modelling phase. Phase 3, which includes the transmission network and electrolysers and is the most detailed setup, initially needed several days to complete a single run with an hourly resolution. The exact runtime was not determined, but it went over seven days, after which the run was stopped because it was considered too long.

Introducing the Gurobi solver significantly reduced the computation time. With Gurobi, a full year simulation at hourly resolution in phase 3 takes about one and a half hours. Using time aggregation through segmentation, where four hours are clustered together, the runtime can be reduced further to around five minutes. However, additional validation is needed to assess whether this aggregated approach still provides results of sufficient quality for TenneT's purposes. In this thesis, all reported results are based on the full hourly resolution runs.

7.1.2 Phase 1: Single node

When comparing the PyPSA single-node results with the PLEXOS single-node results from Guidehouse and TenneT (2025), both models identify solar energy as the most cost-effective technology in the 2040 KM electricity system. However, beyond this shared result, the ranking of technologies by installed capacity between the two models is slightly different, see Table 7.1. The asterisk (*) means that the technology is capped and cannot increase in capacity.

PLEXOS ranking	PLEXOS capacity (GW)	PyPSA ranking	PyPSA capacity (GW)
Solar	131	Solar	111
Offshore wind	37	Gas	28
Battery	24	Battery	27
Gas	13	Offshore wind	21
Onshore wind*	11	Onshore wind*	11
Nuclear	0.4	Nuclear	0.4

Table 7.1: Ranking of PLEXOS and PyPSA model from largest to smallest installed capacity for single node

This difference shows that the two models treat system flexibility and controllable power in different ways. PyPSA places more value on gas and batteries to provide flexibility and reliability, while the PLEXOS model favours offshore wind more strongly. This difference may be caused by hidden CO_2 related penalties or constraints in the PLEXOS model, or by an implicit preference for renewable technologies in its optimisation setup.

Another explanation for the differences between the PyPSA and PLEXOS results relates to how batteries are represented and used in the two models. In the PyPSA-based results, the installed battery capacity is larger than in the PLEXOS results, even though the PLEXOS setup shows a higher total amount of renewable generation capacity, see Figure 7.1. At first glance this might seem surprising, because people often assume that more renewables automatically means more battery storage.

However, batteries do not produce electricity. They store and release energy when needed. In the PyPSA setup, batteries are calculated to be cost-effective for covering peak demand and balancing short-term fluctuations as part of the optimisation across the whole network. In contrast, PLEXOS uses detailed dispatch and additional market logic that takes into account operational constraints and market conditions in greater depth, which can result in different patterns of battery use. These differences come from how each modelling tool represents system operation and optimisation priorities. In the PyPSA results, this leads to a cost-effective configuration where solar energy provides bulk electricity production, batteries provide peak and short-term flexibility, and gas CCGT provides reliable backup power.

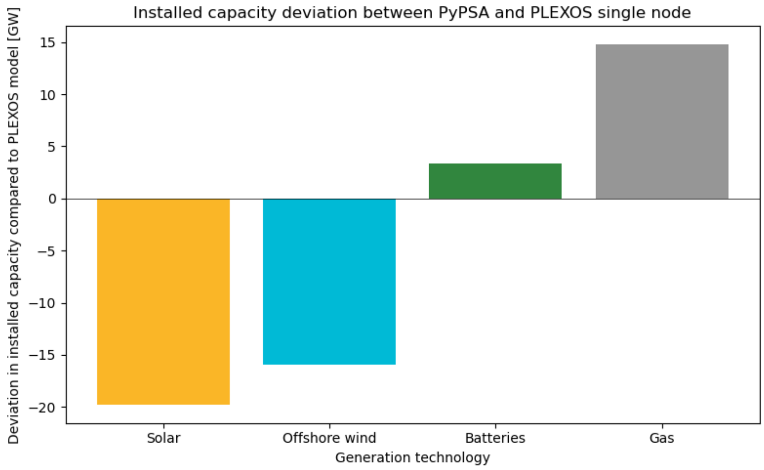


Figure 7.1: Deviation in installed capacities for single node model

Another noticeable difference in the PyPSA and PLEXOS results, shown in Figure 7.1, is the amount of optimised offshore wind capacity. The PyPSA-based results show about 16 GW less offshore wind capacity than the PLEXOS results. Both models optimise generation capacity while taking into account dispatch over time, but PLEXOS includes additional operational details such as ramping limits, minimum load constraints, and reserve requirements in greater depth. Because of this, the PLEXOS formulation can place a higher relative value on steady and predictable generation like offshore wind, while the PyPSA setup may find that combinations such as solar plus batteries appear more cost-effective under the given assumptions. To understand whether these differences are caused by specific modelling assumptions, further analysis of the operational constraints and detailed cost interactions in both models would be needed to see if changes in those assumptions affect the calculated offshore wind capacity in the PyPSA results.

7.1.3 Phase 2: Transmission network

For the transmission network, two types of indicators are analysed, the optimal installed capacities and the load flows.

Optimal installed capacities

When analysing the PyPSA transmission model, the results look very similar to those of the single-node PyPSA model, see Figure 7.2.

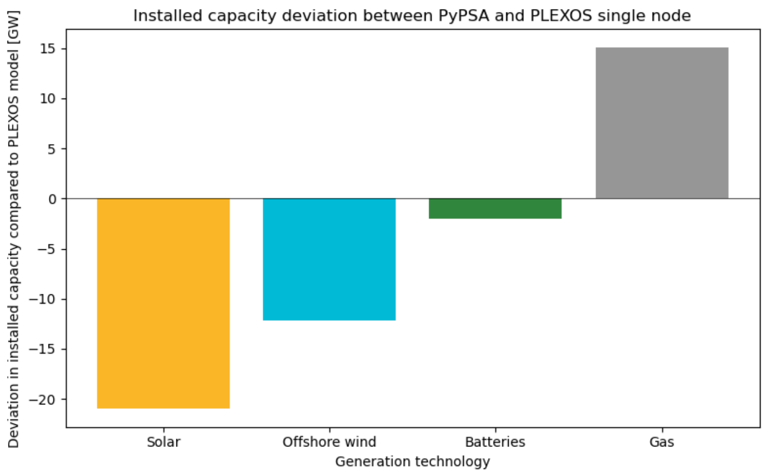


Figure 7.2: Deviation in installed capacities for transmission model

However, one important difference is that the installed capacities for offshore wind are higher, and therefore more similar to the PLEXOS results. The ranking of the technologies can be seen in Table 7.2. Effectively, since the previous phase, batteries and offshore wind have swapped places.

PLEXOS ranking	PLEXOS capacity (GW)	PyPSA ranking	PyPSA capacity (GW)
Solar	131	Solar	111
Offshore wind	37	Gas	28
Battery	24	Offshore wind	25
Gas	13	Battery	22
Onshore wind*	11	Onshore wind*	11
Nuclear	0.4	Nuclear	0.4

Table 7.2: Ranking of PLEXOS and PyPSA model from largest to smallest installed capacity for transmission network

The inclusion of grid constraints and congestion in the transmission-level PyPSA results increases the relative value of offshore wind compared to the single-node case because locational effects start to matter. In the transmission representation, offshore wind is connected at coastal landing points that are closer to large industrial demand centres such as Maasvlakte, Eemshaven, Geertruidenberg, Beverwijk, and Borssele. Even though offshore wind has high investment costs, it also has low operating costs, relatively high capacity factors, and a generation profile that is more evenly distributed over time than solar. With the network included, electricity produced close to demand can reduce the need to send large amounts of power from inland areas, which helps to avoid congestion and lower overall system costs. Because congestion and transport limits are considered at the transmission level, placing more offshore wind capacity near high demand areas can be cost-effective in the calculated outcomes. This is in line with recent research that came to the same conclusions (Frysztacki et al., 2021).

This network effect also helps explain why the transmission results show a larger deployment of gas OCGT and less gas CCGT compared to a single-node configuration. Gas OCGT plants typically have lower upfront costs than CCGT plants and provide short-term flexibility, which can support the system during hours when offshore wind production is limited. Because OCGT operates mainly in limited time periods, its higher fuel costs are less influential on total system cost in the optimisation. Maps in Appendix C confirm that in the transmission model, gas OCGT capacity is mainly associated with buses near high industrial demand, highlighting how network constraints and locational demand patterns influence the deployment patterns in the calculated results.

These results show that moving from a single-node model to a transmission network model has a clear impact on the optimal capacity mix. Once spatial detail and grid constraints are included, the model shifts towards technologies that better fit the geographical structure of the electricity system, leading to different and more realistic investment decisions.

Load flows
The direction of power flows and the deviations in minimum and maximum line loadings are compared between the PyPSA results and the results from Grid Planning. The direction of the power flow matches for 75% of the lines, which are 31 transmission lines. It does not match for 10 lines, which are defined in Appendix C. These mismatches can partly be explained by the simplifications made in the PyPSA model, such as merging buses. By combining nearby substations into one node, the network structure is slightly altered, which can change how electricity chooses its path through the grid and cause the flow direction on individual lines to reverse. Another important reason is that PyPSA and Grid Planning have different starting points for installed capacities, as discussed in the previous section. For example, Grid Planning works with more offshore wind capacity and less gas capacity than PyPSA. Even small differences in where power is generated or consumed can shift local power balances and change the direction of flow on a line. Almost all mismatches take place where substations are clustered because of industrial areas: in Amsterdam (4) and Eemshaven (3). The before mentioned theories could apply to these areas.

In addition, Grid Planning's models already include electrolysers, while in this phase of the PyPSA model they are not yet included. Electrolysers have a high electricity demand and can significantly influence local power flows. The absence of electrolysers in PyPSA could therefore contribute to differences in both flow direction and loading levels.

Overall, these results indicate that the PyPSA results provide a reasonable approximation of power flow patterns, but that differences at the individual line level are expected due to modelling choices, simplifications, and differences in purpose between the two models. More research is needed to test the assumptions PyPSA makes regarding the power flows of lines.

7.1.4 Phase 3: Transmission network + electrolysers

In this phase, the analysis is again divided into two parts: the optimal installed capacities and the load flows. It should be noted again that the PyPSA results, which include the transmission network and electrolysers, are compared with the PLEXOS results, which are at a single node level. However, the PLEXOS results also include electrolysers, even though they are represented without spatial detail.

Optimal installed capacities
The transmission network with electrolysers results in exactly the same optimised installed capacities

as the transmission network without electrolysers, both per technology and in total, see Figure 7.3 and Figure 7.4. At first, this is surprising, because adding one electrolyser per landing location increases electricity demand and one would expect that more generation capacity is needed. However, this result can be explained by looking at what actually drives capacity expansion in the PyPSA model.

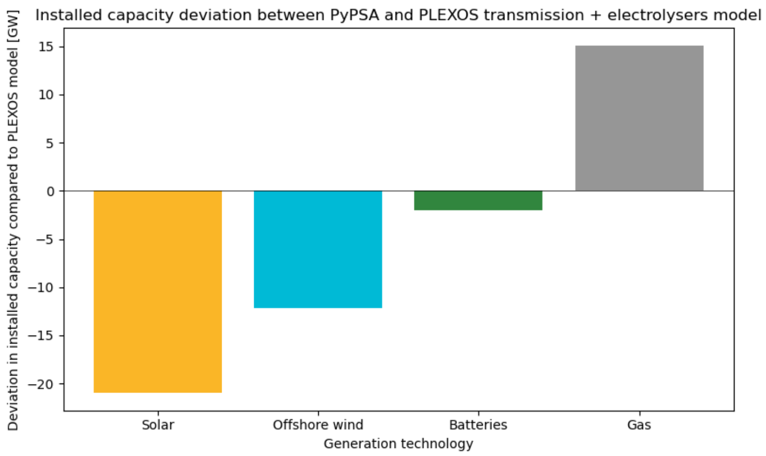
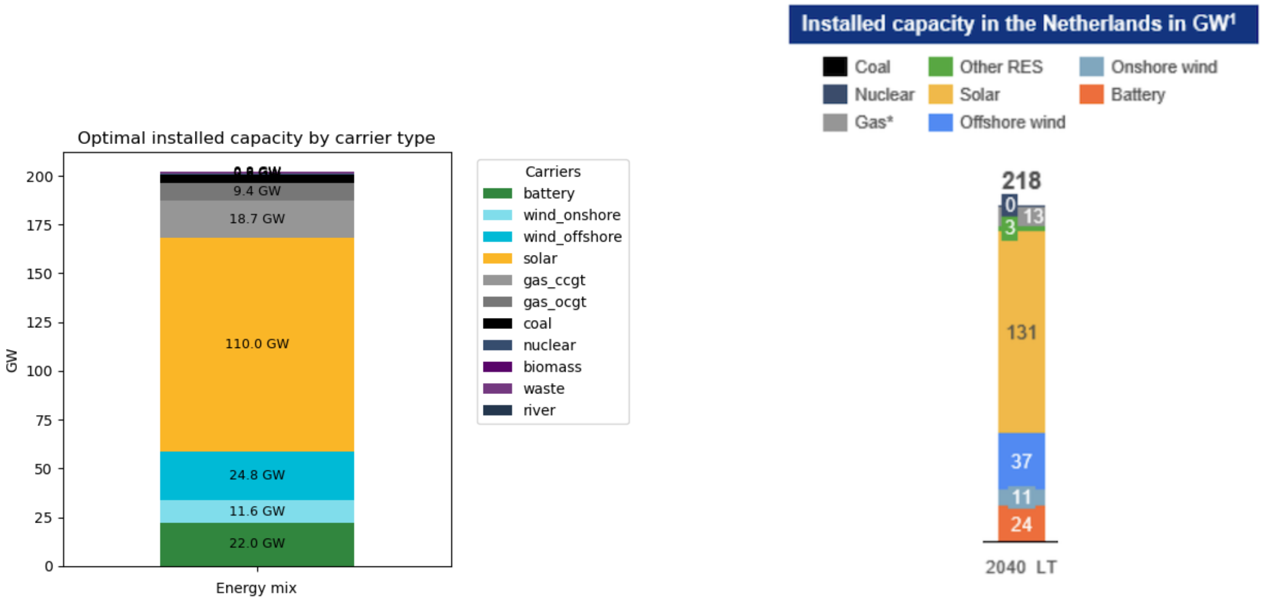


Figure 7.3: Deviation in installed capacities for transmission model + electrolysers



(a) Optimised installed capacities PyPSA transmission + electrolysers

(b) Optimised installed capacities Guidehouse single node

Figure 7.4: Optimal installed capacities by PyPSA and Guidehouse for transmission + electrolysers and single node

In the PyPSA-based results, most new generation capacity is associated with covering the highest electricity demand hour of the year, also known as the system peak. Although electrolysers add additional electricity demand, this extra demand generally remains below the existing peak level that the model's capacity is already sized for. The system already includes a substantial amount of solar energy, offshore wind, and some gas and battery capacity, which means there are many hours when electricity supply exceeds demand and marginal prices are low. Under these conditions, hydrogen production is more cost-effective during hours with surplus electricity rather than during the highest demand hours. As a result, total electricity consumption increases over the year, but the highest hourly demand does not rise enough to trigger additional generation capacity in the calculated outcomes.

Hydrogen storage also reduces pressure on the electricity system. When hydrogen production is high during periods with excess electricity, the surplus can be stored and used later during times with lower electricity supply. This separation of hydrogen demand from immediate electricity generation means that hydrogen production does not force the system to increase generation during already constrained hours. The calculated capacities for hydrogen production at the landing sites are provided in Table 7.3.

Substation	Installed electrolyser capacities (GW)
Eemshaven	8.3
Geertruidenberg	6.4
Beverwijk	10.8
Maasvlakte	6.6
Borssele	4.4

Table 7.3: Installed electrolyser capacities for hydrogen output in PyPSA per substation

Overall, this shows that under the current model assumptions, electrolysers increase electricity use but do not increase capacity needs. As long as hydrogen demand does not raise the system peak or introduce new grid constraints, the optimal installed capacities in the electricity system remain unchanged.

Load flows

The load flow results in the transmission network with electrolysers largely follow the same main routes as in the model without electrolysers, but there are two less matches in direction with Grid Planning their model results. The flow results match in 70% of the lines, which are 29 lines. These can also be found in Appendix C. One possible reason for this difference is the way electrolysers were modelled. In this thesis, specific assumptions were made about their locations, as well as their supply and demand profiles, as explained in Subsection 6.1.3. These assumptions add an extra layer of complexity to the model. Since this could not be fully verified against the PLEXOS model, because PLEXOS only exists as a single node model, this could partly explain the mismatch between the models.

However, when looking at the deviations in minimum and maximum line loadings, the results improve compared to the transmission phase model, see Appendix C. As shown in Figure 7.5, the deviation results are generally smaller than in the model without electrolysers. Although there is still a clear concentration of deviations in the 40–50% and 70–80% ranges, more lines now fall below

20%, and there are even eight cases with deviations under 5%. At the same time, the number of lines with deviations between and above 70–80% is noticeably lower. This suggests that the inclusion of electrolysers leads to power flows that better resemble the magnitudes seen in the Grid Planning model results.

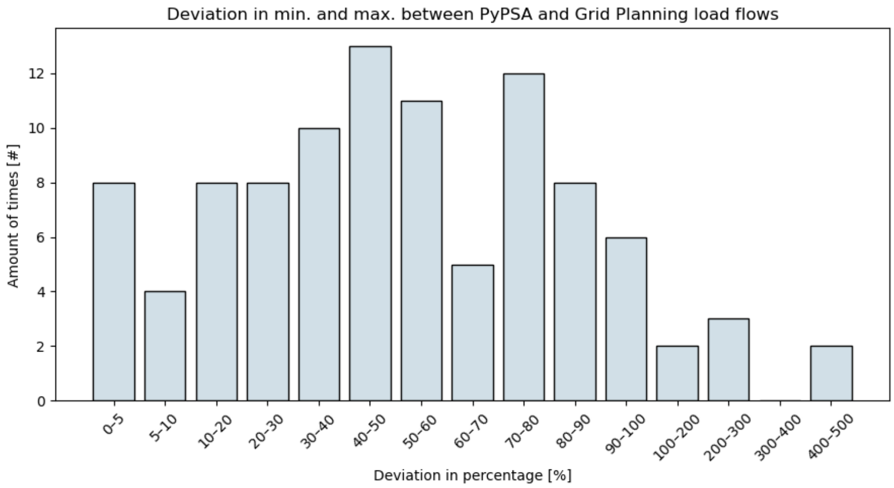


Figure 7.5: Differences in load flows PyPSA and Grid Planning models for transmission + electrolysers model

In the PyPSA-based results, electricity flows are analysed along with how often and by how much transmission lines exceed a safe operating level. A transmission line is flagged as overloaded when, in certain hours, the amount of electricity flowing through it is higher than what is considered acceptable for normal operation. In practice, TenneT applies a thermal rating of 70%, meaning that under normal conditions only up to 70% of a line's physical capacity is used. This limit is set to prevent overheating, which can damage the conductor or create safety issues. The thermal rating depends on factors such as the conductor material and the surrounding temperature. When calculated power flows exceed this threshold in the model, the line is counted as overloaded.

About half of all transmission lines experience at least one overload during the year in the phase 3 results, but many of these occur for only a few hours. For example, the Tilburg-–Rilland line is identified as overloaded for just one hour in the whole year, which is unlikely to be a big issue for system operation. In other cases, lines experience much more congestion. The Dodewaard—Doetinchem line, for instance, is calculated to be overloaded for 5623 hours over the year, with flows higher than the thermal limit by up to 1.6 GW. This pattern suggests a structural bottleneck in the transmission network rather than a short or incidental constraint. An overview of the ten most frequently overloaded lines in the phase 3 PyPSA results is shown in Table 7.4, showing where the system is under the most pressure in the future configuration results.

Name	Amount of over-load (GW)	Hours of overload (h)	Nominal voltage (kV)
Dodewaard – Doetinchem	1.6	5623	380
Diemen – Lelystad	2.0	4007	380
Boxmeer – Dodewaard	0.7	3911	380
Maasbracht – Boxmeer	1.2	2824	380
Musselkanaal – Meeden	1.7	2780	380
Doetinchem – Hengelo	2.1	2657	380
Krimpen aan de IJssel – Diemen	1.0	2507	380
Lelystad – Ens	1.2	2504	380
Vierverlaten – Zeyerveen	0.4	1652	220
Musselkanaal – Veenoord Boerdijk	1.2	1598	380

Table 7.4: Top 10 most overloaded lines in the Netherlands in PyPSA

Importantly, these results align well with TenneT's grid expansion plans, shown in Figure 7.6. For example, the Dodewaard–Doetinchem corridor is already scheduled for a major upgrade. TenneT plans to reinforce 42 km of transmission line in this area, including the replacement or reinforcement of 104 masts (TenneT, 2026a). This upgrade aims to increase the transport capacity of the transmission grid and directly addresses the congestion identified by the PyPSA model. This agreement between the model results and actual grid planning increases confidence in the model's ability to identify realistic bottlenecks.

Figure 7.6: Planned TenneT projects for line upgrading (TenneT, 2026a)

Some of the overloaded lines identified by PyPSA, such as Musselkanaal–Meeden, Musselkanaal–Veenoord Boerdijk, Vierverlaten–Zeyerveen, and Doetinchem–Hengelo, do not appear on the current TenneT upgrade map. This can be explained in several ways. Musselkanaal and Veenoord Boerdijk are relatively new substations and may not yet be included in the published planning overview. The Doetinchem–Hengelo line is located close to the Dodewaard–Doetinchem corridor, which is already planned for reinforcement, and may therefore benefit indirectly from nearby upgrades.

7.1.5 Sensitivity analysis: halved CAPEX for nuclear

To further check whether the PyPSA model behaves in a similar way to the PLEXOS model, and to answer the third research question: *'How can the quality of a simplified PyPSA model be tested against a PLEXOS model?'*, a sensitivity analysis is performed using the phase 1 single-node version of the PyPSA model, since PLEXOS is also built as a single node model. In this analysis, parameters are changed in exactly the same way as in the PLEXOS study. This makes it possible to directly compare how both models respond to the same changes. In this case, the CAPEX for building new nuclear capacity is reduced by half, from 12.000 EUR/kW to 6.000 EUR/kW.

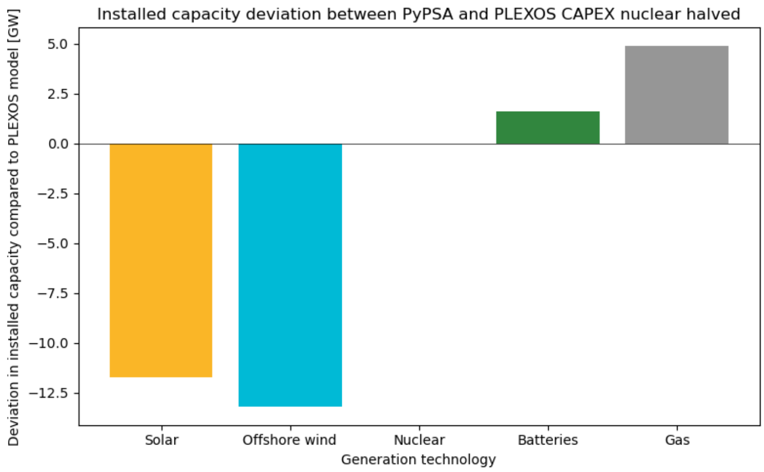


Figure 7.7: Deviation in installed capacities for PyPSA and PLEXOS when CAPEX of nuclear is halved

The results of this change are shown in Figure 7.7. Change for nuclear is barely visible, but it is there: in all previous analyses, nuclear capacity was not expanded and remained at 0.4 GW. After halving the CAPEX, both PLEXOS-based and PyPSA-based results now expand nuclear capacity to 3.3 GW, which is the maximum capacity. This shows that nuclear power becomes economically attractive when its investment costs are highly reduced. More importantly, it shows that PyPSA responds to this cost change in the same way as PLEXOS.

When comparing the other technologies, similar patterns appear in both results. In both the PyPSA-based and PLEXOS-based outcomes, the installed capacities of solar, offshore wind, and batteries decrease when nuclear investment costs are less cost-effective. This happens because larger amounts of nuclear capacity provide low-carbon electricity, which reduces the need for variable renewable generation and storage. At the same time, gas capacity appears more cost-effective in both sets of results. However, the order of this increase differs significantly between the two models: in the

PyPSA-based results, gas capacity increases by about 0.9 GW, while in the PLEXOS-based results it increases by around 9 GW. In the PyPSA-based results, a substantial amount of gas capacity, particularly CCGT, was already present before the change in nuclear costs, as discussed in the previous section. As a result, less additional gas capacity is observed after the cost change. This explains why the gas capacity increase appears much larger in the PLEXOS results than in the PyPSA results, even though both sets of outcomes move in the same general direction under the new assumptions.

7.1.6 Sensitivity analysis: uncapped solar expansion

In this analysis, the limit on how much solar capacity can be installed is removed. This leads to very different outcomes in the PLEXOS and PyPSA models. In the PLEXOS results, large amounts of solar capacity (256 GW) are cost-effective, which is 125 GW more than in the base case. At the same time, gas capacity increases slightly by 2 GW, offshore wind capacity decreases by 14 GW, and battery capacity increases strongly by 81 GW. This shows that once the solar cap is removed, larger amounts of solar capacity are calculated to be cost-effective, and more battery storage is needed to support the additional variable energy generation.

In the PyPSA-based results, the capacity levels after removing the solar cap are the same as in the normal single-node analysis. This suggests that, under the given assumptions, large amounts of solar capacity are already calculated to be cost-effective in the base case, and there is no calculated benefit to increasing it further in the single-node setup. One possible reason for the difference with the PLEXOS results is that the PyPSA single-node setup does not include electrolysers, while the PLEXOS single-node setup does. If electrolysers were included in PyPSA, higher levels of solar capacity could be associated with lower combined system costs, because relatively low-cost solar electricity could be used for hydrogen production. In that scenario, larger solar capacity and lower hydrogen storage might be compatible with cost-effective configurations. However, this explanation cannot be verified with the current model setup, and additional analysis would be needed to fully understand why the PyPSA-based results do not show an increase in solar capacity. The difference between the PyPSA and PLEXOS sensitivity outcomes is shown in Figure 7.8.

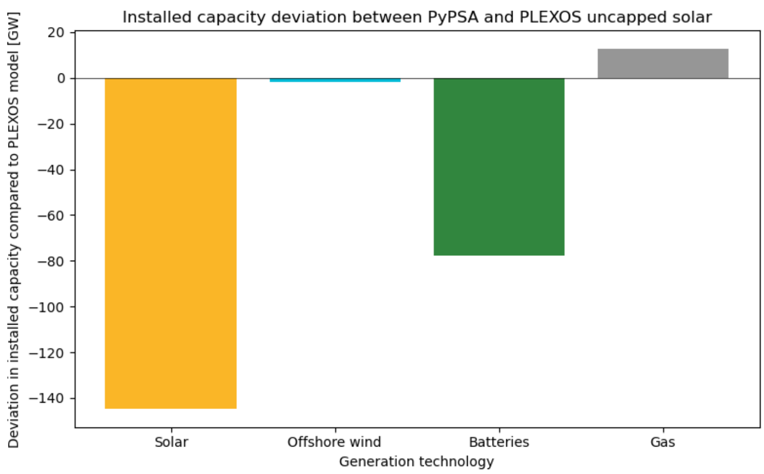


Figure 7.8: Deviation in installed capacities for PyPSA and PLEXOS when solar is uncapped

7.1.7 Synthesis of PyPSA results

Across all phases of the analysis, the PyPSA results are similar to that of PLEXOS and PowerFactory/PSSE, under the 2040 KM scenario. In every phase, large amounts of solar energy are cost-effective, and is so across different spatial levels and sensitivity tests.

In the single-node phase, the PyPSA results focuses strongly on flexibility. Besides solar, the model calculates for relatively large amounts of batteries and gas capacity to be cost-effective. Batteries are mainly used to handle peak demand and short-term fluctuations, while gas provides reliable backup power for moments when renewable generation is low. Offshore wind plays a smaller role in this phase, mainly because of less detailed data assumptions than PLEXOS.

When moving to the transmission network phase, the results become more spatially realistic. Once grid constraints and congestion are included, offshore wind becomes more cost-effective in PyPSA. Offshore wind farms connect close to large industrial demand centres, which reduces the need to transport electricity over long distances and lowers congestion in the grid. As a result, more offshore wind capacity is cost-effective than in the single-node case. At the same time, the model shifts from gas CCGT towards gas OCGT. OCGT plants are cheaper to build and are well suited to cover short periods of low wind generation near demand centres. This shows that including the electricity network has a clear impact on which technologies are preferred and where they are built as opposed to a single node network.

The load flow comparison with PowerFactory/PSSE results shows that the PyPSA results capture the main flow patterns reasonably well, but not the exact loading of individual lines. In about 70-75% of the lines, the direction of flow matches. Differences can largely be explained by model simplifications, such as merged buses, different assumptions on installed capacities, and the fact that PyPSA is designed for long-term system optimisation rather than detailed operational planning. Most deviations in minimum and maximum line loadings remain below 100%. Overall, this suggests that PyPSA provides a useful approximation of grid behaviour at system level, but is not meant to replace detailed grid planning models.

In the transmission model with electrolysers, the optimal installed capacity results remain exactly the same as in the model without electrolysers. Although this may seem surprising, it can be explained by the fact that electrolysers mainly operate during hours with surplus electricity. They increase total electricity consumption, but they do not increase the system peak, which is what drives capacity expansion in PyPSA. Simplified hydrogen storage further smooths demand by allowing production to be shifted over time. While installed capacity results remain unchanged, the inclusion of electrolysers slightly decreases the match between PyPSA and Grid Planning load flow results, but improves the deviation in minimum and maximum load, suggesting a more realistic representation of electricity use in the network. An analysis of the overloaded lines show that the lines that PyPSA highlights as overloaded are the same ones that TenneT is currently or planning to upgrade. This shows that the model is fit to pinpoint congestion in the grid.

Finally, the first sensitivity analysis confirms that the PyPSA results responds to changes in input assumptions in a similar way to PLEXOS results. When nuclear investment costs are halved, nuclear and gas capacity become more cost-effective. This shows that PyPSA correctly captures cost-driven trade-offs between technologies. Differences remain in how much gas capacity is added, but these can be explained by differences in the starting system configuration. The second sensitivity analysis, uncapping the solar expansion delivers counter-intuitive results and should be studied further.

Overall, the PyPSA results show that the model is well suited to explore long-term capacity expansion in the Dutch electricity system. The results are consistent, explainable, and broadly in line with Market Analysis and Grid Planning data.

7.2 EMA WORKBENCH RESULTS

After the final phase of the PyPSA model was validated, all three model phases were used as input for the EMA workbench. The goal of this part of the analysis was to test whether changes in CAPEX influence which technologies are chosen in the optimal system configuration, and when and why cost-based trade-offs between technologies happen. Built-in analysis tools in the EMA workbench were applied to identify patterns, sensitivities, and trade-offs in the results.

7.2.1 Improvement runtime EMA workbench

Several attempts were made to run the EMA Workbench without using the TU Delft supercomputer. However, running the phase 3 model four times required more than four hours, which shows how computationally demanding this is. One trial was carried out to run the model 100 times, but it took more than two weeks before the run was stopped because it was taking too long to complete.

When using DelftBlue, a maximum of 25 runs for phase 3 was feasible within the allowed time for the supercomputer. For the single-node model, which is much simpler, up to 1000 runs could be done within the same time. The EMA Workbench can be made faster by parallelising the model runs so that multiple runs are done at the same time instead of one after another. However, implementing this did not fit within the time frame of this thesis. Further research should explore parallelisation and other performance improvements to make EMA workbench analysis with many runs more practical.

7.2.2 Phase 1: Single node

The pair plot for the single-node model shows almost the same patterns as for the other phases. Because the results are so similar, it is only discussed for the single-node case to avoid repetition.

Pair plot

In this pair plot, see Figure 7.9, each model run contributes multiple dots: one dot for each energy carrier. It shows how each technology contributes to system outcomes such as emissions, energy share, installed capacity, CAPEX, and OPEX across all scenarios. In the plot, not the individual dots are important, but the trends that it shows.

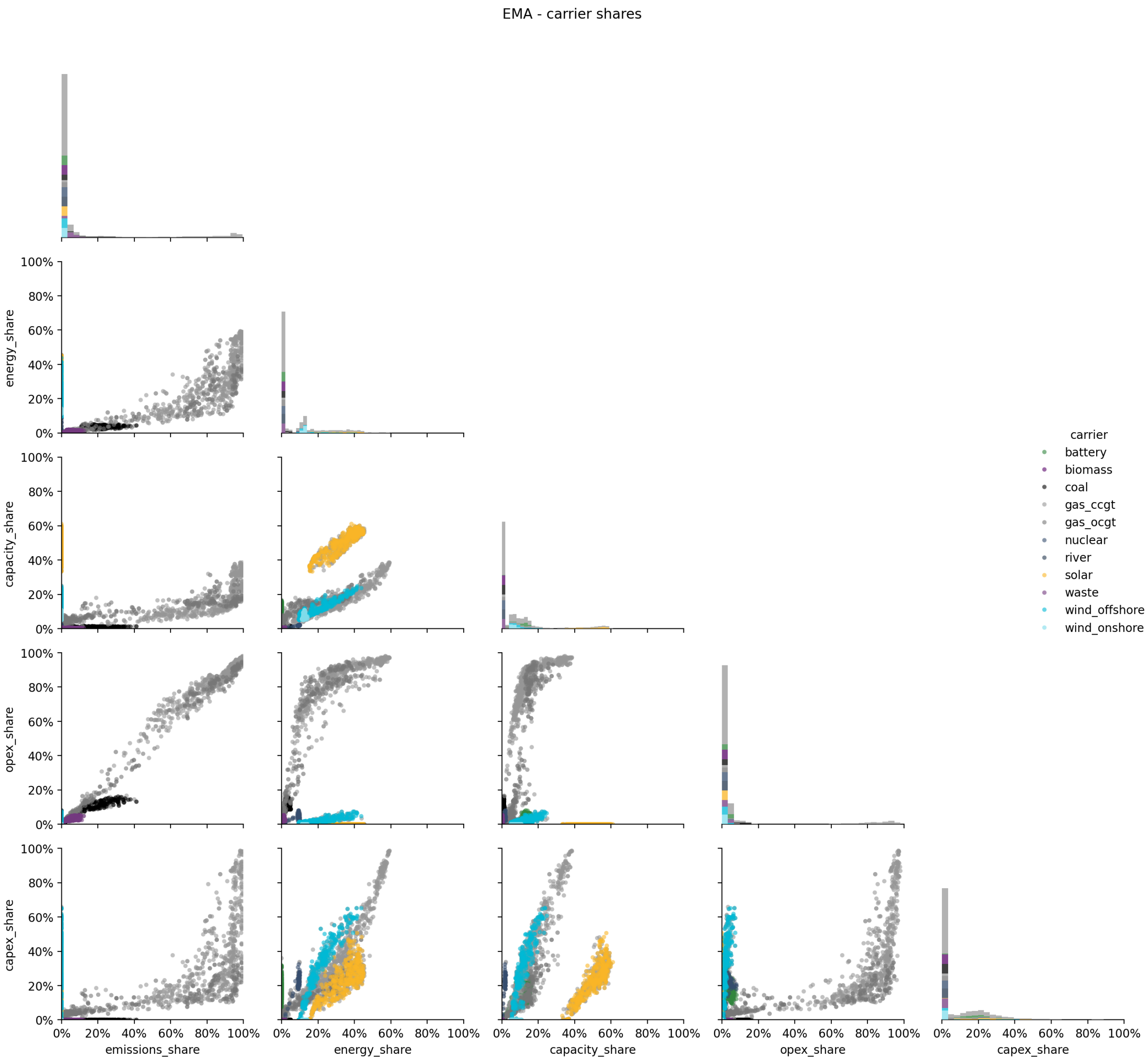


Figure 7.9: Pair plot for single node model

Looking first at emissions share on the x-axis and energy share on the y-axis, gas clearly stands out. The dots move to the far upper right of the emissions axis, showing that when gas supplies an increasing amount of energy, the total emissions in the system also increase. On a smaller level, this also happens for coal and biomass. Renewables like solar and wind technologies are clustered at zero emissions, since they do now emit emissions.

The relationship between energy share (x-axis) and capacity share (y-axis) highlights technical differences between technologies. Solar energy shows relatively high capacity shares for moderate energy shares. This happens because it strongly depends on weather and does not produce electricity all the time, so more installed capacity is needed to deliver the same amount of energy as other technologies. Offshore wind energy and gas show a similar opposite pattern: it needs less capacity share to achieve the same amount of energy.

When looking at the energy share (x-axis) and OPEX share (y-axis) plot, it can be seen that gas dominates the high OPEX region, which is because of the fuel and CO_2 costs. In the beginning, as gas use increases, operational costs rise by a lot. However, after 20% this reaches a plateau. Renewable technologies cluster near zero OPEX share, because once they are built, their operating costs are very low. This makes clear that gas mainly affects how expensive the system is to run, while renewables mainly affect how expensive it is to build.

There is also a clear pattern visible for gas in the graph that compares CAPEX share (on the y-axis) with energy share (on the x-axis). Gas lies between solar and offshore wind. This means that gas requires more investment than solar, but less than offshore wind. The graph also shows that gas can reach very high energy shares in some system configurations. In a few cases, gas provides more than 50% of total electricity production. Solar, on the other hand, has a lower energy share, but the values are more spread out across different model runs. This suggests that uncertainty in solar investment costs has a strong influence on the total investment costs of the system.

Overall, this pair plot shows how different technologies shape system outcomes in different ways. Gas mainly drives emissions and operational costs, while solar dominates installed capacity and low costs. The plot shows why uncertainty in technology costs leads to different system designs, even when total system performance remains similar.

Archetypes of the system

Five archetypes for the single node analysis have been identified and are briefly discussed in this section and in Appendix C. One of the most clear results is that solar energy is the cost-effective technology in all archetypes. This is the case even though solar is more expensive than in the base run in every archetype. It shows that solar is a very robust technology for expanding the electricity system.

Another clear pattern is that offshore wind and gas OCGT often grow together. These two technologies are only built in larger amounts when both are relatively cheap in an archetype. Offshore wind provides large amounts of fairly stable electricity, while gas OCGT is well suited to cover short periods of high demand or low wind production. In archetypes with a lot of offshore wind and gas OCGT, there is usually much less solar energy and battery capacity. This suggests that offshore wind and solar partly act as substitutes: when one becomes more attractive, the other is built less.

Nuclear energy is almost never expanded in the archetypes. This suggests that nuclear is too expensive under the assumptions used in this analysis. Even when the costs of other technologies change, nuclear does not become an attractive option for the system. In the few cases where nuclear is built, the amounts are very small, which confirms that it does not play an important role in the optimal system configurations.

This pattern is also visible in the violin plot of installed capacities across all scenarios, see Appendix C. Solar energy shows a very wide spread in installed capacity, with a clear concentration around 110 GW. This means that in many scenarios, solar capacity ends up close to this value, but it can vary strongly depending on cost assumptions. In contrast, the variation in gas, offshore wind, and battery capacities is much smaller. For offshore wind, many runs cluster around 20 GW, which is noticeably lower than earlier government ambitions of 50 GW by 2040, although these targets have recently been scaled down. Nuclear capacity shows very little variation and is only expanded slightly, by less than 4 GW, in a small number of runs. Because this happens so rarely, nuclear does not appear as a dominant technology in the archetype analysis.

7.2.3 Phase 2: Transmission network

The results from the EMA analysis in phase 2 do not provide any new insights compared to the single-node analysis or the more detailed phase 3 analysis that includes electrolysers. The same patterns, trade-offs, and system behaviours appear in this phase, without adding extra understanding of how the system responds to uncertainty. To avoid unnecessary repetition, the EMA results for phase 2 are therefore not discussed. However, for transparency, the archetypes and its cost factors can be found in Appendix C.

7.2.4 Phase 3: Transmission + electrolysers

In the EMA model that includes both the transmission network and electrolysers, several analyses were carried out to better understand when and why cost-based trade-offs occur between different energy technologies. First, the main system archetypes are discussed. These archetypes show how the electricity system can take different forms depending on the cost assumptions that are made. After this, the influence of individual technologies on key system outcomes are discussed. In this section, the fourth and last research question is answered: *'Which trade-offs emerge in the electricity mix under different cost assumptions in the 2040 Koersvaste Middenweg scenario?'*.

Archetypes of the system

The archetype plot shows five representative model runs, as shown in Figure 7.10 and Table 7.5. By comparing each archetype to the base run, it becomes clear how the electricity system changes when cost assumptions change, and which technologies become more or less important as a result.

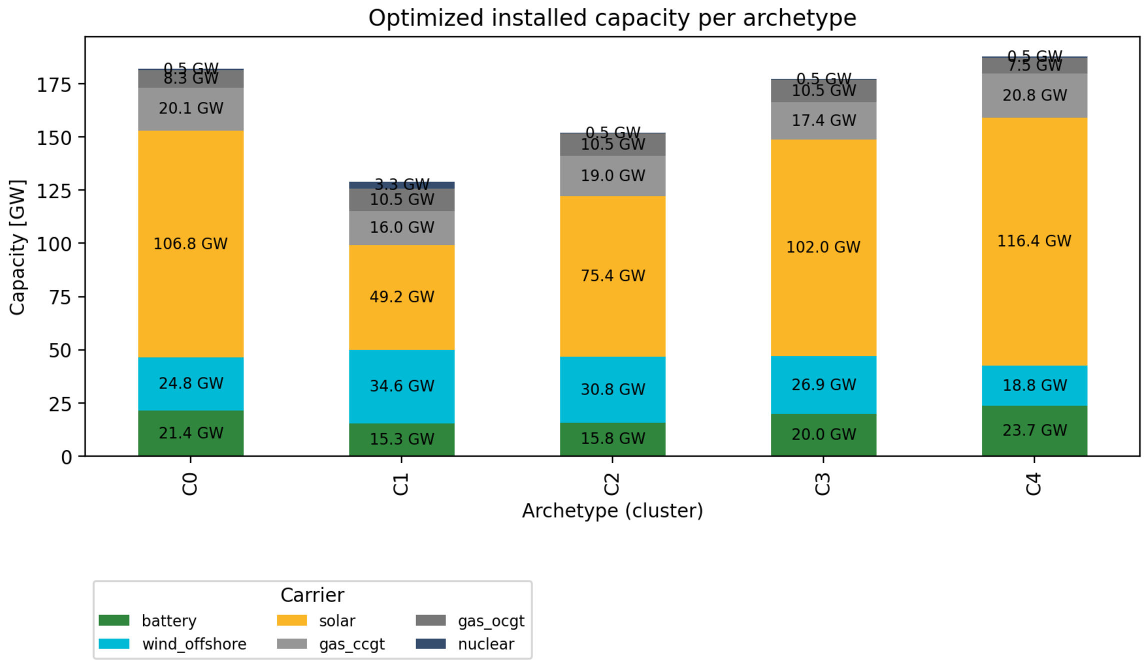


Figure 7.10: Clustered plot for transmission network + electrolyser model

Archetype	Battery	Gas CCGT	Gas OCGT	Nuclear	Solar	Offshore wind
C0	1.01	0.89	0.89	1.25	1.02	0.97
C1	1.08	1.36	1.13	1.36	1.36	0.84
C2	1.30	1.23	1.01	1.24	1.24	0.95
C3	1.14	1.35	1.25	1.05	1.05	0.97
C4	0.96	0.94	0.93	1.01	1.01	1.19

Table 7.5: Capital cost factors for transmission + electrolysers cluster plot

C0: Closest to base run

The main difference in this case is that both types of gas technologies become cheaper, by about 10% compared to the base run. Even with this cost reduction, the optimal system does not change very much. Gas CCGT capacity increases by just over 1 GW, while solar capacity decreases slightly, by around 3 GW. These are relatively small changes compared to the total installed capacities in the system. This shows that the optimal configuration is quite stable: moderate reductions in gas costs are not enough to strongly shift the balance between technologies. In other words, the system is fairly robust to small changes in gas CAPEX prices, and the overall mix of technologies remains mainly the same.

C1: High-cost system

In this archetype, almost all technologies become much more expensive, except for offshore wind. Because offshore wind is the only option that remains relatively affordable, the model relies on it much more than in the other cases. As a result, this archetype shows the highest amount of offshore wind capacity of all the runs. It mainly replaces technologies that become too expensive, especially solar energy and batteries. Solar capacity is reduced by more than half compared to the base run.

In addition, nuclear capacity increases, even though nuclear power is also more expensive in this archetype. This indicates that the model has fewer affordable options left and is therefore forced to use technologies that are normally less attractive. While the model still finds the cost-optimal solution given these assumptions, the overall system in this archetype is likely very expensive. This case clearly shows how strongly the optimal system can change when many technologies become costly at the same time.

C2: Gas OCGT and offshore wind dominated
In this archetype, the direction of the costs follow the same direction as in the previous archetype, however the order of magnitudes are slightly different. In this case, the investment costs for batteries, gas CCGT, nuclear power, and solar energy are all high. Because these technologies become less attractive, the model looks for cheaper alternatives to still meet electricity demand in a reliable way. As a result, the system shifts toward building more gas OCGT and offshore wind capacity.

Gas OCGT is relatively cheap to build and is well suited for covering short periods of high demand, even though it is expensive to operate. Offshore wind, on the other hand, can deliver large amounts of electricity at low operating costs once it is built. Together, these two technologies compensate for the reduced solar capacity and help the system stay balanced. This archetype shows how the model adapts when many commonly used technologies become expensive, by relying more on flexibility from gas OCGT and steady energy production from offshore wind.

C3: Balanced system
In this run, gas technologies and batteries become more expensive, while the costs of the other technologies stay roughly the same. Because batteries are now more costly, it is less attractive to install large amounts of solar energy. Solar power is variable and often needs batteries to store excess electricity and supply power during peak hours. When batteries are expensive, this support becomes less affordable.

As a result, the model reduces solar capacity and shifts toward offshore wind and gas OCGT instead. Again, offshore wind provides a more stable and evenly spread electricity supply over time. Gas OCGT is added because it can quickly respond to changes in demand and act as flexible backup power. Together, these changes allow the system to remain reliable, even though key technologies like gas and batteries have become more expensive.

C4: Solar dominated
In this run, almost all technologies become cheaper or stay about the same in cost, except for offshore wind, which becomes more expensive. Because offshore wind is less attractive under these assumptions, the model installs much less of it. To make up for this loss in energy production, the system adds more solar capacity and more batteries.

At the same time, the model builds more gas CCGT and less gas OCGT. Gas CCGT plants are better suited to provide steady, long-term electricity generation, while OCGT plants are mainly used for short peak moments. By shifting toward more CCGT, the system becomes more stable and less dependent on flexible but expensive peak generation. This run shows how the model adjusts the balance between technologies when offshore wind becomes relatively expensive.

Influences of technologies on system totals

Another analysis carried out in the EMA workbench is the feature scoring analysis and the Theil–Sen analysis, see Figure 7.11. The Theil–Sen method is a robust way to estimate simple linear relationships. In this case, the analysis is performed for energy output. This means that the results describe what happens in the system when the annual electricity production of a technology increases. Each row in the plot represents a technology, measured by its energy share (how much electricity it produces in a year). Each column represents an important system outcome, such as costs, emissions, or prices.

The values in the plot show both the direction and the strength of the relationship. A positive value means the outcome increases when the technology produces more energy, while a negative value means it decreases. To avoid misleading results, self-correlation is removed. For example, more gas production does not automatically lead to higher gas OPEX in the plot. Instead, the focus is on how changes in one technology affect the rest of the system. In the following sections, the effects of each technology are discussed in more detail.

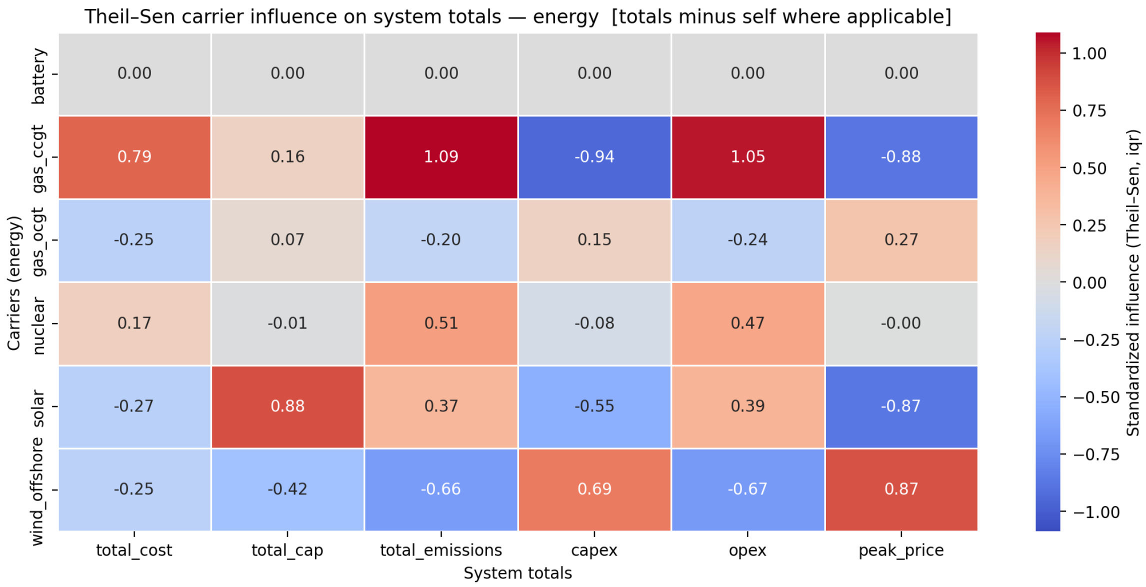


Figure 7.11: Theil-Sen energy plot for clustered plot with electrolysers

Batteries

This whole row shows values close to zero. This means that batteries do not add new electricity to the system, but only store electricity that has already been produced by other technologies. Because of this, and because of the minus self principle, increasing battery energy output does not directly change the total amount of energy produced, system costs, or emissions in this plot.

Gas CCGT

This technology has a very strong impact on the overall system. Most importantly, it leads to much higher CO_2 emissions and OPEX. This is because an increase in CCGT capacity is linked to higher battery capacity in the system. Batteries have noticeable OPEX costs and are used frequently to balance supply and demand. As a result, when both CCGT and batteries increase, total OPEX costs rise.

At the same time, gas CCGT reduces total CAPEX. Because it can deliver large amounts of reliable electricity, the system does not need to build as many other expensive technologies with high up-front costs, such as offshore wind or large battery systems. This lowers overall investment needs. Although gas CCGT often set electricity prices during peak hours, increasing this capacity appears to have a negative effect on peak prices. This is because this dispatchable capacity lowers scarcity, and high price events become less extreme.

In short, gas CCGT makes the system cheaper to build and more stable to run, with lower peak prices and lower CAPEX. However, this comes at a clear downside: it is expensive to operate and leads to high CO_2 emissions. This highlights an important trade-off between cost, reliability, and sustainability in the energy system.

Gas OCGT

Gas OCGT shows a different pattern compared to gas CCGT. First of all, the values in the Theil–Sen plot are much closer to zero. This indicates that gas OCGT does not play a very large role in most of the EMA model runs as a dominant technology. It is mainly used as a supporting or peak technology rather than as a main source of electricity.

The total OPEX decreases slightly then gas OCGT produces more energy, even though gas OCGT has a high OPEX. However, more gas OCGT means less gas CCGT, which also has a substantial OPEX. And because of the minus self principle, the total OPEX appears to decrease.

Total CO_2 emissions also decrease slightly. This is because archetypes with more gas OCGT, as shown in Figure 7.10, also tend to include more offshore wind energy. Offshore wind provides a stable and emission-free electricity supply, which reduces the need for fossil generation overall. However, peak electricity prices increase slightly in these cases. This is likely because gas OCGT is very expensive to operate, so when it is needed during peak demand hours, it pushes prices up.

Nuclear

Just like gas OCGT, the values for nuclear energy are much smaller than those for gas CCGT, solar, and offshore wind. This shows that nuclear is not a major driver in the system across the EMA runs. Even when nuclear capacity increases, it does not strongly shape the overall system outcomes. At first glance, the increase in total CO_2 emissions is surprising, because nuclear power is defined as emission-free. However, this result can be explained by looking at the full system configuration. In the archetypes where nuclear capacity increases, there is also a clear increase in gas CCGT capacity.

Since gas CCGT has high emissions and runs for many hours, this leads to higher total emissions overall. The same reasoning explains the increase in total OPEX.

One might also expect total CAPEX to increase when more nuclear capacity is built, since nuclear plants are very expensive to construct. However, this effect is likely offset by a strong reduction in solar capacity, which decreases by more than half in these archetypes, and by the minus self principle. As a result, the increase in nuclear CAPEX is partly compensated by lower investments in solar.

Solar

The indicators for solar energy are relatively strong compared to the other technologies. The most noticeable result is the strong positive effect on total installed capacity. It appears that when large capacities of solar are installed, other technologies also have to increase in capacities. This is primarily the case for batteries.

Solar energy also helps reduce peak electricity prices. This is likely because solar, especially when combined with batteries, reduces the mismatch between electricity supply and demand during peak hours. Solar power is usually available during daytime demand peaks and is used before more expensive technologies, such as gas plants, are turned on. This lowers scarcity and prevents extreme price spikes.

At first, it may seem counter-intuitive that total CO_2 emissions and total OPEX increase when solar energy production rises. However, this can be explained by looking at the full system setup. In many archetypes, higher solar capacity is paired with more gas CCGT. Because gas CCGT has high emissions and operating costs, this combination leads to higher overall emissions and OPEX, even though solar itself is clean and cheap to operate.

Offshore wind

Offshore wind is also a strong driver of the system and, in many ways, has the opposite effect of solar energy. One surprising result is that offshore wind is linked to higher peak electricity prices. This can be explained by looking at the full system configuration. In scenarios with more offshore wind, there is often more gas OCGT, which is very expensive to operate, leading to higher price spikes during those moments.

Total CAPEX increase with more offshore wind, probably because this comes with the increase of gas OCGT, which has substantial investment costs. On the other hand, total CO_2 emissions and OPEX decrease. Offshore wind does not produce CO_2 and has very low operating costs once it is built. Because gas OCGT is primarily used for backup, these plants only run for a limited number of hours. As a result, their contribution to emissions and OPEX remains small, keeping both indicators relatively low.

7.2.5 Synthesis of EMA workbench results

Across all EMA phases, one result is consistent: solar remains the most cost-effective technology in installed capacity. Even when solar is more expensive than in the base case, the model still builds a lot of it. At the same time, solar shows a wide spread in CAPEX share across runs, which suggests that uncertainty in solar investment costs can strongly influence total system investments and can help trigger trade-offs with other technologies.

A second clear finding is that gas and renewables play very different roles. Gas, especially CCGT, is the main driver of emissions and OPEX. In the pair plot, gas quickly moves to high OPEX and high emission outcomes as its energy share increases. The Theil–Sen analysis confirms this: when gas CCGT supplies more energy, total emissions and OPEX rise strongly. However, gas CCGT also has a system benefit: it can reduce total CAPEX and reduce peak prices, because dispatchable power lowers scarcity. This creates a clear trade-off: gas makes the system easier and sometimes cheaper to build, but it becomes dirtier and more expensive to operate.

The archetypes show that the model mainly switches between a few typical system designs depending on which technologies become relatively cheaper. One important pattern is that offshore wind and gas OCGT often increase together. This happens because they complement each other: offshore wind supplies a stable, low-emission energy flow over the year, while gas OCGT provides fast backup during peak moments or low-wind periods. In these offshore-wind heavy archetypes, there is usually less solar and less battery capacity. This suggests that, in the model, offshore wind and solar act partly as substitutes, especially when storage becomes more expensive or when the system prefers a more stable generation profile. Offshore wind also shows a strong trade-off profile in the Theil–Sen plot: it raises CAPEX (because it and gas OCGT is expensive to build), but it lowers emissions and OPEX (because gas OCGT does not run often). It can increase peak prices in some runs, this is because those runs are covered by the expensive OCGT peak units.

Finally, the EMA results suggest that nuclear is not a key option under the cost assumptions used here. Nuclear hardly expands in most runs and does not appear as a dominant technology in the archetypes. In the few cases where it does increase, the amounts are small. This indicates that even large changes in the other technologies do not automatically make nuclear attractive in this modelling setup.

Overall, the EMA analysis shows that the Dutch 2040 KM system is likely to stay solar-heavy, but the supporting structure around solar can change quickly. The system can shift toward more offshore wind and gas OCGT in high-cost or low-storage cases, or toward more solar and batteries when offshore wind becomes expensive. This is the main message of the EMA workbench: the optimal system is sensitive to CAPEX assumptions, so robust policy should plan for multiple credible system designs, not just one.

8 DISCUSSION

This chapter discusses the main findings of this thesis and places them in a broader context. It begins with a reflection on the overall study, specifically if the research achieved the goal it is built for. Next, the limitations of the PyPSA and EMA workbench model are discussed. This helps understanding how reliable the results are and where caution is needed when interpreting them. Then, personal reflections are discussed to evaluate where the research process could have been improved. The discussion then considers the implications of the results for policymakers, and explains how the findings can support decision-making in energy system planning. Finally, the chapter ends with recommendations for future research. These suggestions highlight how the model and analysis could be improved.

8.1 REFLECTING ON THIS STUDY

The model developed in this thesis was created in response to TenneT's need for a tool that can quickly calculate capacity expansion and load flows for the Dutch electricity system. This is needed to answer questions from governmental policymakers who need insights into the effects of new policies or additions of infrastructure. At the moment, answering these questions is usually done with detailed models that are built and operated by different teams and can take several months. However, for early-stage policy discussions, their level of precision is not always necessary. At the stage where the PyPSA model is used for, quick first insights are more important than fully detailed results.

Usefulness of the PyPSA model

The PyPSA model developed in this thesis is intended to fill that gap. It is designed as a faster and more flexible tool that policymakers within TenneT, especially from EPN-SO, can use to explore different scenarios and policy options. This thesis has shown that PyPSA has strong potential for this purpose. The results are close to those from PLEXOS and PowerFactory, and more importantly, the model shows similar behaviour when key input variables are changed. For example, when the CAPEX of nuclear power is halved, the system responds in a comparable way. The model also presents results in a clear and visual manner, which is helpful for policymakers who may not have a deep technical background. It becomes easier to understand the impact of policy choices, such as subsidies for nuclear energy on the overall electricity mix.

However, more work and further validation are still needed before the model can be fully used in real-world decision-making. There are still noticeable differences in installed capacities and load flows compared to the more detailed models. In addition, some modelling assumptions are simplified. For example, curtailment of renewable energy is currently assumed to be free in the model, while in reality it has economic and system consequences. This can influence investment decisions and system behaviour. Another important limitation is that interconnector capacity is treated as a fixed input, while in reality it could also be optimised as part of the expansion planning. These points would need to be fixed before the model can be used for real-world decision-making.

Past, present and future installed capacities

The model results indicate that the total installed capacity in the 2040 KM scenario can exceed 200 GW, compared with about 150 GW of installed capacity in 2025 (Netbeheer Nederland, 2025). Table 8.1 shows the installed capacities for 2015, 2025, and 2040, and highlights the large difference between the current system and the future system under these assumptions. To supply electricity in 2040 in a cost-effective way, a large amount of additional capacity is calculated to be necessary.

Technology	Installed capacities in 2015 (GW)	Installed capacities in 2025 (GW)	Optimised installed capacities for 2040 KM (GW)
Solar	1.4	33	110
Offshore wind	3.6	5	24.8
Onshore wind	4	7	11.6
Gas (CCGT and OCGT)	15	17	28.1
Nuclear	0.5	0.5	0.5
Flexible sources	/	12 (batteries + interconnections)	22 (batteries only)

Table 8.1: Past, current and optimised installed capacities for the Netherlands (Netbeheer Nederland, 2025; TenneT, 2015)

When looking at how installed capacities developed in the past, from 2015 to 2025, it becomes clear that, except for solar energy, growth has been limited for most technologies. This makes the large expansion required between 2025 and 2040 uncertain. An important question arises: is such rapid growth realistic within the given time frame? Achieving this expansion will require careful planning and strong policy support. In particular, solar power, offshore wind, and gas-fired generation will need to grow significantly and in a coordinated way to make the future system feasible.

Improving computational time

An important part of this thesis was improving the computation time of the PyPSA model. The goal was to make it run faster than before and significantly faster than the PLEXOS model. The most detailed version of the model, which includes the full transmission network and electrolysers, currently takes about one and a half hours to simulate one year with hourly resolution. Earlier in the project, this same calculation required more than a full day, so the runtime has already been reduced considerably.

The computation time can be reduced even further by applying time-step aggregation methods. By grouping certain hours together, the number of time steps decreases, which lowers the calculation time. In some cases, this reduces the runtime to around five minutes. Although this comes at the cost of slightly lower temporal detail, it still provides useful system insights for EPN-SO. This demonstrates that the PyPSA model is suitable for relatively fast analysis of the Dutch transmission grid, especially in situations where quick, order-of-magnitude results are more important than highly detailed hourly precision.

An optimal solution is not a robust one

The use of the EMA Workbench in this thesis also shows its value for policy support. The analysis demonstrates that the electricity system can change quickly when investment costs vary slightly. Technology choices can shift, and the ranking of cost-optimal options can change. This highlights the importance of testing parametric uncertainty not only by comparing predefined scenarios, but also by varying assumptions within those scenarios. The results show that energy policy should focus on robustness rather than aiming for one single 'optimal' solution. This should not only be tested with parametric uncertainty, but also structural uncertainty, to take into account as much factors that influence the energy system. This thesis shows that future energy policies should allow for flexibility and alternative pathways instead of relying on one fixed outcome.

8.2 LIMITATIONS

To build a functional model, this study relies on a set of assumptions, data sources, and theoretical frameworks. These choices made it possible to develop the model and carry out the analyses presented in this thesis. However, they also introduce limitations that can influence the results and how they should be interpreted. In this section, the limitations are discussed in two main categories: the PyPSA model structure and the use of the EMA Workbench.

8.2.1 PyPSA model structure

Identification of KPIs

The research started by identifying the most important KPI's for evaluating the PyPSA model. To do this, both scientific literature and input from TenneT experts were used. The main KPI's selected were installed capacities of energy technologies and load flows. These indicators focus mainly on the economic and technical performance of the electricity grid. However, recent studies show that these traditional KPI's are no longer sufficient to fully analyse electricity systems. Factors beyond costs, such as public acceptance, political feasibility, and broader techno-economic feasibility, can be just as important as economic indicators (Lehtveer et al., 2021; Pye et al., 2021). Lehtveer et al. (2021) also highlights KPI's such as import and export dependency and sectoral emission reduction rates, especially in the context of recent geopolitical tensions. This raises the question of whether additional KPI's should have been included in this thesis to provide a more complete analysis. Including these KPI's would move this thesis in the direction of a MGA approach.

Absence of gas and heat network

The current PyPSA model does not include gas or heat networks. In the real energy system, these networks supply part of the energy demand that would otherwise be met by electricity, for example because of hybrid boilers. The interaction between electricity, gas, and heat networks plays an important role in the energy transition and can significantly influence electricity demand and investment decisions. Including these interactions would improve the realism of the model and provide a more complete picture of the future energy system.

Unfair comparison due to spatial detail

The installed capacity results from the PyPSA transmission network models have been compared to those from PLEXOS' single node model. The lack of spatial resolution in the PLEXOS model does not make it a completely fair comparison. Luckily the TenneT teams and Guidehouse are working on a new model that does include the transmission network in PLEXOS. In the near future a better comparison can be made.

PyPSA vs. PyPSA-EUR

It was decided to work on a custom PyPSA model for TenneT instead of using existing PyPSA models, like PyPSA-EUR. The main reason for this choice is because of the need for speed for the model. A larger model like PyPSA-EUR, that contains multiple nodes per country in the EU takes a long time to run, while EPN-SO is primarily interested in fast results which provide a reliable order of magnitude for only the Netherlands. To do this, modelling the entire European electricity system in detail is not necessary. A focused model of the Netherlands with simplified cross-border connections is therefore better suited to answer grid-related questions at the national and regional level. However, because of this, the interconnector capacity behaviour is not included in the model.

Different sources for transmission network

The transmission network that was used in this thesis comes from researchers from TU Delft (Zomerdijk et al., 2022). They have plotted all the 2014 380 and 220 kV substations and lines in (x,y) format, information that TenneT doesn't readily have available. In the meanwhile, a few substations have been added (Tilburg, Halsteren, Veenoord Boerdijk and Musselkanaal). These have been added by hand, and have been added because these substations are included in the Grid Planning load flow calculations. To better compare these calculations with the PyPSA results, it is key that these lines/-substations match as much as possible. The substations do match now, but it must be noted that the substations are from different sources. Also there have been reinforcements made to the grid, but these are not included.

Clustering of substations

There are a few clusters of 380 and 220 kV stations that have been grouped to reduce computational time. This has been explained in Subsection 6.1.2. Substations are grouped only when they are in a near vicinity of each other and when, if they are grouped, they do not influence how the load flows go. For example, for big industrial clusters like the Eemshaven cluster. This clustering approach was chosen instead of using a clustering algorithm such as k-means because of how it affects the load flow results. If an algorithm like k-means were used, the resulting clusters would have a different spatial layout, which would change how the electricity cables are represented and therefore change the calculated load flows. Even with the clustering method used here, grouping any substations already changes the layout and means that the load flows are not exactly the same as they would be in a fully detailed network.

Costs for hydrogen network
When modelling the hydrogen network, transport was assumed to be free of costs and unlimited. This choice was made deliberately, because the focus of this thesis was on identifying where hydrogen production would take place in an optimal system without constraints. However, in reality, plans already exist for a dedicated hydrogen network with specific routes and costs (Gasunie, 2025). For a more realistic analysis, this planned hydrogen infrastructure and its costs should be included.

Curtailment of renewable energy sources
Lastly, in this model, curtailment of offshore wind energy and solar energy is free. Wind and solar parks can curtail their energy without any financial implications. In reality, curtailment leads to lost revenue for producers and network companies. To combat this, they would install less capacity, just enough, to prevent this. So in this model, more RES is installed than would probably happen in reality.

8.2.2 EMA workbench structure

To explore uncertainty in the model, several method choices were made that affect the results.

LHS was used to sample uncertainties and policy levers. Ideally, a Monte Carlo approach would have been preferred, because it samples the uncertainty space better. However, Monte Carlo sampling requires a very large number of model runs to produce reliable results. Due to limited computational resources and time constraints, this was not feasible for this study. LHS was therefore chosen as a compromise, as it provides good coverage of the uncertainty space with far fewer simulations.

While very important for modelling future energy systems, electricity demand (loads) was not included as an uncertainty. Including loads as uncertain parameters would have greatly increased the number of variables in the analysis, making the experiments computationally too heavy. As a result, demand profiles were kept fixed across all scenarios. This means that the results focus mainly on uncertainty in technology costs and do not capture uncertainty related to future demand growth or changes in consumption patterns.

These choices help keep the analysis computationally manageable, but they also limit the scope of uncertainty that is explored. This should be kept in mind when interpreting the robustness of the results.

8.3 PERSONAL REFLECTIONS

In hindsight, the order in which the research was carried out could have been improved. The study started by validating the single-node PyPSA model, after which the model was expanded to a transmission network, and later offshore wind and electrolysers were added. However, it was discovered late in the research that the PLEXOS single-node model used for comparison already includes electrolysers. This means that, ideally, electrolysers should also have been included in the PyPSA single-node model. This modelling choice likely influenced the installed capacities, especially in the sensitivity analysis where the solar capacity limit was removed.

While interviews with TenneT staff provided valuable input, the number of people involved and the time available for interviews is limited. As a result, the findings reflect the views of a few individuals rather than the full range of experts within the organisation. To mitigate this as much as possible, interviewees were selected to reflect both policy as technical roles within TenneT.

As this project is a master thesis, there was limited time to fully explore a large range of possible scenarios or test all model variations. Some research steps, like market analysis, or broader stakeholder involvement, were not done due to time limits.

8.4 RECOMMENDATION FOR POLICY DEVELOPERS

Five key recommendations can be made to policy developers reading this thesis:

One of the strongest and most consistent findings is that solar energy appears to be most cost-effective in the PyPSA and EMA workbench results in all model phases. From an energy policy perspective, especially for the government, this suggests that continued large-scale deployment of solar energy is a low-regret option. Even if future costs turn out higher than expected, solar remains an attractive backbone technology. However, the results also show that solar mainly dominates in terms of installed capacity, not necessarily in annual energy production. This means policymakers from TenneT and the government should not only focus on installed capacity numbers, but also on system integration: storage, flexibility, and grid infrastructure are essential to make large solar volumes work reliably.

The EMA workbench results show that depending on cost assumptions, the system shifts between: solar + batteries, offshore wind + gas OCGT, or combinations of both. This means that flexibility technologies compete with each other, and investing in all could improve system resilience. This is especially important for TenneT, because as the operator of the electricity system, they play a key role in deciding which flexibility technologies are developed and integrated into the grid.

Offshore wind becomes much more attractive once the transmission network is included. The transmission results clearly show that offshore wind helps reduce congestion by supplying electricity close to large industrial demand centres. At the same time, offshore wind has very high investment costs and increases total CAPEX. The EMA archetypes show that when offshore wind becomes expensive, the system quickly shifts back toward solar and batteries. This means offshore wind is not automatically the best solution in all futures, especially under high cost uncertainty. So, offshore wind expansion should be closely coordinated with grid planning and industrial demand locations. It should not only be evaluated on energy output, but also on its impact on congestion. This is important for TenneT and the government, because they are responsible for planning how offshore wind is rolled out and for making sure that this expansion is well aligned with grid development, industrial demand, and other energy technologies.

The overload analysis shows that several transmission lines experience thousands of hours of congestion, and these results strongly match TenneT's real-world upgrade plans. This confirms that grid expansion is not optional, regardless of which generation mix is chosen. Importantly, congestion appears even in cost-optimal systems. This means that market-based optimisation alone will not solve grid bottlenecks. So, system models like PyPSA can be used to identify future congestion early. These insights can be valuable for TenneT when planning and preparing for future grid conditions.

Lastly, the EMA workbench results show that small changes in CAPEX assumptions can lead to different optimal system designs, even when overall system performance remains similar. This means that choosing one 'optimal' pathway is risky. The main lesson is that robust policy is more important than optimal policy, which is important for governmental policy makers when developing new energy policy. So next to scenario analysis, also uncertainty analyses should be performed. In this thesis, parametric uncertainty was used, but structural uncertainty should also be tested.

8.5 RECOMMENDATION FOR FUTURE RESEARCH

While this thesis provides useful insights, several areas could be improved in future research.

The load flow results from the PyPSA model could be brought closer in line with the results used by TenneT's Grid Planning team. While the current model shows the general patterns of how electricity flows through the grid, there are still large differences on individual transmission lines compared with the detailed planning models. Future research could work on improving that alignment by using more detailed grid data and logic. Improving this alignment would help increase confidence in the model's results for analyses that focus on grid behaviour and congestion.

Interactions with neighbouring countries should be modelled in more detail. Cross-border electricity flows can have a big impact on the Dutch grid, especially in a highly interconnected European power system. In addition, future studies should better integrate interactions between the electricity system and other energy networks, such as gas, hydrogen, and heat. Including these sector couplings more explicitly would give a more complete picture of system behaviour and congestion effects.

Although time aggregation methods were tested during this study, the results from those methods were not thoroughly validated. Instead, the full hourly resolution results were used throughout the thesis, because they ran fast enough for the current model setup. However, as the model is developed further and becomes more detailed, the computation time is likely to increase, which will make time aggregation more useful or even necessary. For that reason, additional validation of time-aggregated results is needed in future research to ensure that they still produce results that are meaningful and reliable.

Finally, this research could be repeated using exact input data from the ETM. Using the same data source throughout would allow for a more direct comparison with other studies and reduce uncertainties caused by differences in assumptions or data quality. This would help isolate the effects of modelling choices from those of the input data itself.

Together, these improvements would increase the accuracy and applicability of the model and support more robust decision-making for future energy system planning.

9 CONCLUSION

TenneT is in need of a modelling tool that can do quick calculations on capacity planning and power flows of the Dutch electricity grid. At the moment, answering these questions requires multiple tools and teams, which makes the process complex and time-consuming. This creates a clear need for a simple tool such as PyPSA, which first needs to be thoroughly validated before it can be used in practice. Before answering the main research question, first the sub-questions are answered.

SUB-QUESTION 1.

What are key indicators to assess the results of electricity capacity expansion planning?

The literature review shows that many studies mainly use indicators related to system reliability, outages, and costs. However, in this thesis, indicators related to reliability and outages are not useful, because the PyPSA model assumes perfect system operation and focuses on minimising costs. This means that outages and blackouts do not occur in the model. As a result, the most relevant indicators for this study are those related to costs and how available and usable the grid is. Based on this reasoning, and supported by interviews with experts from TenneT, the selected KPIs are installed capacity per generator type, dispatch, electricity prices, load flows, and interconnector capacity.

SUB-QUESTION 2.

How can offshore wind energy and electrolysers be reliably added to the PyPSA model?

For the transmission-level PyPSA model to work properly, the locations of offshore wind farms, their landing points, and the lengths of the sea cables needed to be included. This information was created by combining and overlaying maps from the Open Infrastructure Map, Rijkswaterstaat, TenneT's Geoweb, the Ministry of Infrastructure and Water Management, and Google Maps. While this approach provides a good estimate of the locations and distances, the exact coordinates could not be determined. For this, more specific data from TenneT was needed, which was not available at the time of writing.

Electrolysers were added to the model using data from the Energy Transition Model, with the assumption that they are only built at offshore wind landing locations. Because the installed capacity results of this regionalised PyPSA model are compared with results from a single-node PLEXOS model, a fully fair comparison is not yet possible. At this stage, the results of the two models were compared, but a more fair and comparison will only be possible in the future, once PLEXOS is also extended to include regionalised network data.

SUB-QUESTION 3.

> *How can the quality of a simplified PyPSA model be tested against a PLEXOS model?*

The KPIs identified in sub-question 1 were taken from the outputs of the PyPSA and PLEXOS models. For each modelling phase in PyPSA, these indicators were compared with the results from PLEXOS. Where differences appear, an effort is made to explain these as clearly as possible, with many of them being caused by differences in data assumptions between the two models.

To check whether PyPSA and PLEXOS show similar system behaviour, the sensitivity analyses that were performed in PLEXOS are repeated in the PyPSA model. One of these analyses, in which the CAPEX of nuclear power is halved, shows that both models respond in the same way. However, this is not the case for the analysis where solar capacity is uncapped. This difference can be explained by the fact that electrolysers are not included in the single-node PLEXOS model, which affects how excess solar energy can be used.

SUB-QUESTION 4.

> *Which trade-offs emerge in the electricity mix under different cost assumptions in the 2040 Koersvaste Middenweg scenario?*

In almost all archetypes, solar emerges as the most cost-efficient technology. Even when solar becomes more expensive, it stays that way. However, changes in solar CAPEX strongly affect total system investments and can trigger trade-offs with other technologies. For example, when offshore wind becomes relatively cheaper, the system shifts toward more offshore wind combined with gas OCGT for backup, while solar and batteries decrease. This shows that solar and offshore wind partly substitute each other, depending on their relative costs and the need for storage.

Gas creates one of the clearest trade-offs in the system. When gas CCGT produces more electricity, total emissions and OPEX increase significantly. At the same time, gas CCGT reduces total CAPEX and can lower peak electricity prices because it provides reliable, dispatchable power during scarcity hours. Offshore wind shows the opposite pattern: it increases CAPEX due to high investment costs, but lowers emissions and OPEX because it is clean and cheap to operate. Nuclear does not play a major role in any of the cost scenarios. Overall, the main trade-offs emerge from changes in the relative costs of solar, offshore wind, gas, and storage, which cause shifts in the cost-optimal ordering of technologies and lead to different but internally consistent system designs.

Using the information from the sub-questions, finally, the main research question can be answered:

> *How can a PyPSA model of the Dutch transmission grid be designed and validated to support optimal capacity expansion planning for the 2040 Koersvaste Middenweg scenario, including offshore wind, electrolysers, and cost-driven trade-offs between technologies?*

Supporting optimal capacity expansion planning The results show that PyPSA is a strong tool for energy policy decision-making. Across all model phases, PyPSA finds cost-optimal system designs that show similar results as existing models. One of the most robust findings is that solar energy is a cost-efficient solution for the Dutch electricity system in 2040, regardless of spatial detail or uncertainty in costs. Even when solar becomes more expensive, it stays cost-efficient. This shows that solar is a low-regret technology for future planning. However, the model also makes clear that solar dominance is mainly in installed capacity, not necessarily in energy production, which highlights the importance of flexibility and system integration.

Including the transmission network strongly improves realism. Offshore wind becomes more attractive once grid constraints are included, because it can deliver electricity near large industrial demand centres at the coastal landing sites. This reduces the need to transport power from inland and therefore lowers congestion on the grid. At the same time, gas technologies shift from CCGT to OCGT, reflecting the need for less expensive investment costs and flexible peak capacity near demand locations. This shows that PyPSA does not just optimise technologies in isolation, but also accounts for where electricity is produced and consumed.

The overload analysis further demonstrates PyPSA's usefulness for planning. The model identifies structural bottlenecks that closely match TenneT's real grid reinforcement plans, which increases confidence that PyPSA can be used to signal future grid problems early.

Including offshore wind farms and electrolysers. This research was designed in multiple phases, starting from a simple single-node model and gradually adding more realism by including the transmission network, offshore wind farms, and electrolysers. This step-by-step approach makes it possible to understand how each modelling choice affects the results and allows for structured validation.

Validation against both Market Analysis (PLEXOS) and Grid Planning (PowerFactory/PSSE) models shows that PyPSA produces credible and explainable results. In terms of installed capacities, PyPSA shows results close to PLEXOS, and also responds to cost changes in the same direction, as shown clearly in the nuclear CAPEX sensitivity analysis. In terms of load flows, PyPSA captures the main flow directions and congestion patterns in the Dutch grid, even though exact line loadings differ due to simplifications and different modelling purposes.

When electrolysers are included, PyPSA does not artificially inflate generation capacity. It is assumed this is because hydrogen production mainly takes place during hours with surplus electricity. This shows that PyPSA can realistically integrate electrolysers in the model. Overall, the model is well suited to represent offshore wind and electrolysers in a way similar to existing models.

Trade-offs in the 2040 Koersvaste Middenweg scenario The EMA Workbench analysis reveals the key trade-offs that emerge in the 2040 KM scenario when technology costs are uncertain. Rather than one single optimal future, the system moves between a small number of typical system designs (archetypes). A central trade-off is between cost, emissions, and flexibility:

- Gas CCGT lowers investment costs and peak prices and makes the system easier to operate, but strongly increases CO_2 emissions and operating costs.

- Offshore wind increases investment costs but lowers emissions and operating costs, while sometimes leading to higher peak prices due to reliance on expensive backup during low-wind periods.
- Solar + batteries reduce peak prices and provide clean energy, but require very large installed capacities and depend heavily on storage costs.
- Gas OCGT works as flexible backup and complements offshore wind, but increases peak prices when used.
- Nuclear barely appears in the results and only becomes relevant when investment costs are halved, showing it is not competitive under current assumptions.

A key insight from the EMA results is that small changes in CAPEX assumptions can shift the system between different archetypes, even though total system performance (meeting demand reliably) remains similar. This means the future electricity system is not fragile, but it is path-dependent: different combinations of technologies can work, depending on costs and investments.

The most important takeaway is that there is no single 'optimal' future system. Instead, the Dutch electricity system can be built in several robust ways, each with its own strengths and weaknesses. This means that robust policy, which prepares for multiple credible system designs instead of one optimal pathway, is more important than aiming for one exact solution. A PyPSA-based model, combined with uncertainty analysis, provides a strong foundation to support such policy decisions.

REFERENCES

Abhay, S. (2025, July). K-Means Clustering: A Deep Dive into Unsupervised Learning. Retrieved January 21, 2026, from https://medium.com/@abhaysingh71711/k-means-clustering-a-deep-dive-into-unsupervised-learning-81213f56cfc9

Ahluwalia, H. (2024, February). Pair Plots in Machine Learning. Retrieved January 21, 2026, from https://www.analyticsvidhya.com/blog/2024/02/pair-plots-in-machine-learning/

Altair. (2026). Latin HyperCube. Retrieved January 23, 2026, from https://help.altair.com/hwdesktop/hst/topics/design_exploration/method_latin_hypercube_sto_r.htm

Amos, D., Au-Yong, C. P., & Musa, Z. N. (2021). A Literature Review on Performance Measurement. In D. Amos, C. P. Au-Yong, & Z. N. Musa (Eds.), *Measurement of Facilities Management Performance in Ghana's Public Hospitals* (pp. 21–34). Springer. https://doi.org/10.1007/978-981-33-4332-0_3

Antea Group. (2025, May). *Notitie Reikwijdte en Detailniveau Plan-MER locatiestudie twee nieuwe kerncentrales* (tech. rep.). https://www.rvo.nl/files/file/2025-05/Concept-NRD-nieuwbouw-kerncentrales.pdf

Bissell, G., Costa, C., Nigris, M. D., Losa, I., Margarone, M., Distribuzione, E., Galant, S., Varela, J., & Invernizzi, M. (2014). DEFINITION AND PRACTICAL APPLICATION OF KEY PERFORMANCE INDICATORS TO SUPPORT EUROPEAN GRID OPERATORS TO ENABLE THE ENERGY POLICY GOALS. Retrieved June 5, 2025, from https://www.semanticscholar.org/paper/DEFINITION-AND-PRACTICAL-APPLICATION-OF-KEY-TO-GRID-Bissell-Costa/4cd11aa0544e69510a2f0b5737294919b299cb79

Bock, F., & Pfenninger-Lee, S. (2025). Rarely pure and never simple: Exploring perceptions of truth and objectivity in energy modelling and scenarios. *Energy Research & Social Science*, *127*, 104229. https://doi.org/10.1016/j.erss.2025.104229

Box, G. (1976). Science and Statistics. *Journal of the American Statistical Association*, *71*(356), 791–799. Retrieved September 11, 2025, from https://www-sop.inria.fr/members/Ian.Jermyn/philosophy/writings/Boxonmaths.pdf

Brown, T., Hörsch, J., & Schlachtberger, D. (2018). PyPSA: Python for Power System Analysis. *Journal of Open Research Software*, *6*(1), 4. https://doi.org/10.5334/jors.188

Coordinet. (2019). *D1.6 – List of KPIs: KPI and process of measures* (tech. rep.).

DeCarolis, J., Babaee, S., Li, B., & Kanungo, S. (2016). Modelling to generate alternatives with an energy system optimization model. *Environmental Modelling & Software*, *79*, 300–310. https://doi.org/10.1016/j.envsoft.2015.11.019

DeCarolis, J., Daly, H., Dodds, P., Keppo, I., Li, F., McDowall, W., Pye, S., Strachan, N., Trutnevyte, E., Usher, W., Winning, M., Yeh, S., & Zeyringer, M. (2017). Formalizing best practice for energy system optimization modelling. *Applied Energy*, *194*, 184–198. https://doi.org/10.1016/j.apenergy.2017.03.001

DelftBlue. (2026). DelftBlue Documentation. Retrieved January 26, 2026, from https://doc.dhpc.tudelft.nl/delftblue/Accounting-and-shares/#msc-students

DIgSILENT. (2026). PowerFactory. Retrieved January 23, 2026, from https://www.digsilent.de/en/powerfactory.html

Durakovic, G., del Granado, P. C., & Tomasgard, A. (2023). Powering Europe with North Sea offshore wind: The impact of hydrogen investments on grid infrastructure and power prices. *Energy*, *263*, 125654. https://doi.org/10.1016/j.energy.2022.125654

Energie Beheer Nederland. (2024). *EBN-Infographic-2024* (tech. rep.). Retrieved September 5, 2025, from https://www.ebn.nl/wp-content/uploads/2024/01/EBN-Infographic-2024.pdf

Energy Examplar. (2025). PLEXOS Energy Modeling Software. Retrieved June 4, 2025, from https://www.energyexemplar.com/plexos

Energy Transition Model. (2025). Hourly curves for hydrogen - Data export - Energy Transition Model. Retrieved November 17, 2025, from https://2025-01.energytransitionmodel.com/scenario/data/data_export/hourly-curves-for-hydrogen

ENTSOE. (2018, October). *4th ENTSO-E Guideline for cost-benefit analysis of grid development projects* (tech. rep.).

ENTSOE. (2025, March). 28 April 2025 Blackout. Retrieved December 30, 2025, from https://www.entsoe.eu/publications/blackout/28-april-2025-iberian-blackout/

European Commission. (2026). Hydrogen. Retrieved January 23, 2026, from https://energy.ec.europa.eu/topics/eus-energy-system/hydrogen_en

Feijoo, F., Pfeifer, A., Herc, L., Groppi, D., & Duić, N. (2022). A long-term capacity investment and operational energy planning model with power-to-X and flexibility technologies. *Renewable and Sustainable Energy Reviews*, *167*, 112781. https://doi.org/10.1016/j.rser.2022.112781

Fotis, G., Vita, V., & Maris, T. I. (2023). Risks in the European Transmission System and a Novel Restoration Strategy for a Power System after a Major Blackout [Publisher: Multidisciplinary Digital Publishing Institute]. *Applied Sciences*, *13*(1), 83. https://doi.org/10.3390/app13010083

Frysztacki, M. M., Hörsch, J., Hagenmeyer, V., & Brown, T. (2021). The strong effect of network resolution on electricity system models with high shares of wind and solar. *Applied Energy*, *291*, 116726. https://doi.org/10.1016/j.apenergy.2021.116726

Gasunie. (2025). Waterstofnetwerk. Retrieved January 15, 2026, from https://www.gasunie.nl/expertise/waterstof/waterstofnetwerk

Gea-Bermúdez, J., Bramstoft, R., Koivisto, M., Kitzing, L., & Ramos, A. (2023). Going offshore or not: Where to generate hydrogen in future integrated energy systems? *Energy Policy*, *174*, 113382. https://doi.org/10.1016/j.enpol.2022.113382

Glaum, P., Neumann, F., & Brown, T. (2024, April). Offshore power and hydrogen networks for Europe's North Sea [arXiv:2404.09721 [physics]]. https://doi.org/10.48550/arXiv.2404.09721

Grochowicz, A., Van Greevenbroek, K., Benth, F. E., & Zeyringer, M. (2023). Intersecting near-optimal spaces: European power systems with more resilience to weather variability. *Energy Economics*, *118*, 106496. https://doi.org/10.1016/j.eneco.2022.106496

Guidehouse & TenneT. (2025, July). *Energy system optimization modelling* (tech. rep.).

Gurobi. (2025). Beyond Speed: Understanding the True Meaning of Solver Performance. Retrieved November 17, 2025, from https://www.gurobi.com/resources/beyond-speed-understanding-the-true-meaning-of-solver-performance/

Haasnoot, M., Kwakkel, J. H., Walker, W. E., & ter Maat, J. (2013). Dynamic adaptive policy pathways: A method for crafting robust decisions for a deeply uncertain world. *Global Environmental Change*, *23*(2), 485–498. https://doi.org/10.1016/j.gloenvcha.2012.12.006

Henke, H. T. J., Gardumi, F., & Howells, M. (2022). The open source electricity Model Base for Europe - An engagement framework for open and transparent European energy modelling. *Energy*, *239*, 121973. https://doi.org/10.1016/j.energy.2021.121973

Hermans, S. (2025, July). Het Windenergie Infrastructuurplan Noordzee. Retrieved September 5, 2025, from https://open.overheid.nl/documenten/c8ece805-8016-435a-bcb6-ba72fdf83ab8/file

Heussen, K., Koch, S., Ulbig, A., & Andersson, G. (2010). Energy storage in power system operation: The power nodes modeling framework [ISSN: 2165-4824]. *2010 IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT Europe)*, 1–8. https://doi.org/10.1109/ISGTEUROPE.2010.5638865

Hoffmann, M., Priesmann, J., Nolting, L., Praktiknjo, A., Kotzur, L., & Stolten, D. (2021). Typical periods or typical time steps? A multi-model analysis to determine the optimal temporal aggregation for energy system models. *Applied Energy*, *304*, 117825. https://doi.org/10.1016/j.apenergy.2021.117825

Kwakkel, J. H. (2017). The Exploratory Modeling Workbench: An open source toolkit for exploratory modeling, scenario discovery, and (multi-objective) robust decision making. *Environmental Modelling & Software*, *96*, 239–250. https://doi.org/10.1016/j.envsoft.2017.06.054

Kwakkel, J. H. (2025). Open exploration - EMA workbench. https://emaworkbench.readthedocs.io/en/latest/indepth_tutorial/open-exploration.html

Kwon, D. Y., & Hofmann, M. (2023). Energy System Optimization Towards a Fossil-Free Power Plant Portfolio. *36th International Conference on Efficiency, Cost, Optimization, Simulation and Environmental Impact of Energy Systems (ECOS 2023)*, 635–645. https://doi.org/10.52202/069564-0058

Lehtveer, M., Goransson, L., Heinisch, V., Johnsson, F., Karlsson, I., Nyholm, E., Odenberger, M., Romanchenko, D., Rootzen, J., Savvidou, G., Taljegard, M., Toktarova, A., Ullmark, J., Vilen, K., & Walter, V. (2021). Actuating the European Energy System Transition: Indicators for Translating Energy Systems Modelling Results into Policy-Making. *Frontiers in Energy Research*, *9*(677208). https://doi.org/https://doi.org/10.3389/fenrg.2021.677208

Limpens, G., Rixhon, X., Contino, F., & Jeanmart, H. (2024). EnergyScope Pathway: An open-source model to optimise the energy transition pathways of a regional whole-energy system. *Applied Energy*, *358*, 122501. https://doi.org/10.1016/j.apenergy.2023.122501

Lombardi, F., Pickering, B., Colombo, E., & Pfenninger, S. (2020). Policy Decision Support for Renewables Deployment through Spatially Explicit Practically Optimal Alternatives. *Joule*, *4*(10), 2185–2207. https://doi.org/10.1016/j.joule.2020.08.002

Lombardi, F., van Greevenbroek, K., Grochowicz, A., Lau, M., Neumann, F., Patankar, N., & Vågerö, O. (2025). Near-optimal energy planning strategies with modeling to generate alternatives to flexibly explore practically desirable options. *Joule*, 102144. https://doi.org/10.1016/j.joule.2025.102144

Loucks, D., van Beek, E., Stedinger, J., Dijkman, J., & Villars, M. (2005). *Model sensitivity and uncertainty analysis.* Water Resources Systems Planning; Management.

Lüth, A., Seifert, P. E., Egging-Bratseth, R., & Weibezahn, J. (2023). How to connect energy islands: Trade-offs between hydrogen and electricity infrastructure. *Applied Energy*, *341*, 121045. https://doi.org/10.1016/j.apenergy.2023.121045

Martínez-Gordón, R., Gusatu, L., Morales-España, G., Sijm, J., & Faaij, A. (2022). Benefits of an integrated power and hydrogen offshore grid in a net-zero North Sea energy system. *Advances in Applied Energy*, *7*, 100097. https://doi.org/10.1016/j.adapen.2022.100097

Mediterranean Transmission System Operators. (2018). *Deliverable 5.2 "Key Performance Indicators of the regional electricity system"* (tech. rep.).

Ministry of Climate and Green Growth. (2025). *Klimaat- en Energienota 2025* (tech. rep.).

Ministry of Economic Affairs and Climate. (2025, May). *Ontwikkelkader windenergie op zee* (tech. rep.). Retrieved October 23, 2025, from https://open.overheid.nl/documenten/267eb9c7-793a-4528-8290-61d383b83488/file

Ministry of Infrastructure and Water Management. (2025). Concept NRD en Participatieplan Programma Noordzee 2022 - 2027. Retrieved November 12, 2025, from https://www.platformparticipatie.nl/programmanoordzee/nrd-participatieplan-programmanoordzee/default.aspx

Netbeheer Nederland. (2025, May). *Netbeheer Nederland Scenario's Editie 2025* (tech. rep.).

Netherlands Enterprice Agency. (2023, October). Maatregelen tegen vol elektriciteitsnet (netcongestie) [Last Modified: 2025-05-27T09:20 Publisher: Ministerie van Algemene Zaken]. Retrieved June 4, 2025, from https://www.rijksoverheid.nl/onderwerpen/duurzame-energie/kabinet-neemt-maatregelen-tegen-vol-elektriciteitsnet-netcongestie

Netherlands Enterprice Agency. (2025, April). Maatregelen om het elektriciteitsnet sneller uit te breiden. Retrieved June 4, 2025, from https://www.rvo.nl/onderwerpen/netcongestie/maatregelen-om-het-elektriciteitsnet-sneller-uit-te-breiden

Neumann, F., & Brown, T. (2021). The near-optimal feasible space of a renewable power system model. *Electric Power Systems Research*, *190*, 106690. https://doi.org/10.1016/j.epsr.2020.106690

Neumann, F., & Brown, T. (2023). Broad ranges of investment configurations for renewable power systems, robust to cost uncertainty and near-optimality. *iScience*, *26*(5), 106702. https://doi.org/10.1016/j.isci.2023.106702

Nolting, L., & Praktiknjo, A. (2022). The complexity dilemma – Insights from security of electricity supply assessments. *Energy*, *241*, 122522. https://doi.org/10.1016/j.energy.2021.122522

North Sea Wind Power Hub. (2021). *NSWPH Integration of offshore wind* (tech. rep.). Retrieved September 9, 2025, from https://northseawindpowerhub.eu/files/media/document/NSWPH%20Integration%20of%20offshore%20wind.pdf

North Sea Wind Power Hub. (2024). NSWPH - Pathway 2.0. Retrieved September 12, 2025, from https://northseawindpowerhub.eu/files/media/document/2024.06.24_NSWPH%20Pathway%20Study%202.0.pdf

Nweke, C. I., Leanez, F., Drayton, G. R., & Kolhe, M. (2012). Benefits of chronological optimization in capacity planning for electricity markets. *2012 IEEE International Conference on Power System Technology (POWERCON)*, 1–6. https://doi.org/10.1109/PowerCon.2012.6401421

Parzen, M. (2021, November). Open source HiGHS solver & performance boost for energy system models - Modelling / Open modelling [Section: Modelling]. Retrieved November 17, 2025, from https://forum.openmod.org/t/open-source-highs-solver-performance-boost-for-energy-system-models/2922

Parzen, M., Hall, J., Jenkins, J., & Brown, T. (2022). Optimization solvers: The missing link for a fully open- source energy system modelling ecosystem. https://doi.org/10.5281/zenodo.6409432

Pfenninger, S., & Pickering, B. (2018). Calliope: A multi-scale energy systems modelling framework. *Journal of Open Source Software*, *3*(29), 825. https://doi.org/10.21105/joss.00825

Priesmann, J., Nolting, L., & Praktiknjo, A. (2019). Are complex energy system models more accurate? An intra-model comparison of power system optimization models. *Applied Energy*, *255*, 113783. https://doi.org/10.1016/j.apenergy.2019.113783

Prina, M. G., Johannsen, R. M., Sparber, W., & Østergaard, P. A. (2023). Evaluating near-optimal scenarios with EnergyPLAN to support policy makers. *Smart Energy*, *10*, 100100. https://doi.org/10.1016/j.segy.2023.100100

Pye, S., Broad, O., Bataille, C., Brockway, P., Daly, H. E., Freeman, R., Gambhir, A., Geden, O., Rogan, F., Sanghvi, S., Tomei, J., Vorushylo, I., & Watson, J. (2021). Modelling net-zero emissions energy systems requires a change in approach [Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/14693062.2020.1824891]. *Climate Policy*, *21*(2), 222–231. https://doi.org/10.1080/14693062.2020.1824891

PyPSA Docs. (2025a). PyPSA Design. Retrieved October 27, 2025, from https://docs.pypsa.org/latest/user-guide/design/

PyPSA Docs. (2025b). PyPSA Overview. Retrieved October 27, 2025, from https://docs.pypsa.org/latest/user-guide/optimization/overview/

PyPSA Docs. (2025c, October). PyPSA Home. Retrieved October 27, 2025, from https://docs.pypsa.org/latest/

PyPSA Docs. (2025d, October). PyPSA Objective. Retrieved October 27, 2025, from https://docs.pypsa.org/latest/user-guide/optimization/objective/

Rijksoverheid. (2025). Doelstellingen [Last Modified: 2025-09-17 Publisher: Rijkswaterstaat]. Retrieved September 23, 2025, from https://windopzee.nl/onderwerpen/doelstellingen/

Rijkswaterstaat. (2025). Net op Zee [Last Modified: 2025-10-21 Publisher: Rijkswaterstaat]. Retrieved November 12, 2025, from https://noordzeeloket.nl/functies-gebruik/energietransitie-zee/net-zee/

Ronde, K. d. (2025, September). Europese stroomnetten verdrinken in aanvragen. Retrieved September 5, 2025, from https://fd.nl/bedrijfsleven/1568416/europese-stroomnetten-verdrinken-in-aanvragen-nederland-koploper-wachttijden

RVO. (2024, December). Windparken op de Noordzee. Retrieved September 23, 2025, from https://www.rvo.nl/onderwerpen/windenergie-op-zee/windparken-noordzee

Satpati & Siddarth. (2023). TheilSen Regression and Estimator. Retrieved January 21, 2026, from https://www.prepvector.com/blog/theilsen-regression-and-estimator

Schermeyer, H., Studer, M., Ruppert, M., & Fichtner, W. (2017). Understanding Distribution Grid Congestion Caused by Electricity Generation from Renewables. In C. Derksen & C. Weber (Eds.), *Smart Energy Research. At the Crossroads of Engineering, Economics, and Computer Science* (pp. 78–89). Springer International Publishing. https://doi.org/10.1007/978-3-319-66553-5_6

Spasić, V. (2023, May). Serbia's TSO EMS to use Plexos to model grid investments. Retrieved February 16, 2026, from https://balkangreenenergynews.com/serbias-tso-ems-to-use-plexos-to-model-grid-investments/

SUMICSID & Council of European Energy Regulators. (2024). *TCB21 Transmission System Cost E ciency Benchmarking TenneT* (tech. rep.).

Süsser, D., Gaschnig, H., Ceglarz, A., Stavrakas, V., Flamos, A., & Lilliestam, J. (2022). Better suited or just more complex? On the fit between user needs and modeller-driven improvements of energy system models. *Energy*, *239*, 121909. https://doi.org/10.1016/j.energy.2021.121909

TenneT. (2015). *Market Review 2015* (tech. rep.).

TenneT. (2019). *TenneT Holding BV Additional CSR data: CSR data linked to Integrated Annual Report 2019* (tech. rep.).

TenneT. (2026a). Dodewaard - Doetinchem. Retrieved January 20, 2026, from https://www.tennet.eu/nl/projecten/provincies/gelderland/opwaarderen-380-kv-net/dodewaard-doetinchem

TenneT. (2026b, January). *Ontwerp investeringsplan 2026* (tech. rep.).

Tennøe, S., Halnes, G., & Einevoll, G. T. (2018). Uncertainpy: A Python Toolbox for Uncertainty Quantification and Sensitivity Analysis in Computational Neuroscience [Publisher: Frontiers]. *Frontiers in Neuroinformatics*, *12*. https://doi.org/10.3389/fninf.2018.00049

Trutnevyte, E. (2016). Does cost optimization approximate the real-world energy transition? *Energy*, *106*, 182–193. https://doi.org/10.1016/j.energy.2016.03.038

Unnewehr, J. F., Schäfer, M., & Weidlich, A. (2022). The value of network resolution – A validation study of the European energy system model PyPSA-Eur. *2022 Open Source Modelling and Simulation of Energy Systems (OSMSES)*, 1–7. https://doi.org/10.1109/OSMSES54027.2022.9769123

Valdes, J., Wöllmann, S., Weber, A., Klaus, G., Sigl, C., Prem, M., Bauer, R., & Zink, R. (2020). A framework for regional smart energy planning using volunteered geographic information [ISSN: 1680-7340]. *Advances in Geosciences*, *54*, 179–193. https://doi.org/10.5194/adgeo-54-179-2020

Van Someren, C., Visser, M., & Slootweg, H. (2025). How battery energy storage impacts grid congestion – An evaluation of the trade-offs between different battery control strategies. *Journal of Energy Storage*, *134*, 118100. https://doi.org/10.1016/j.est.2025.118100

Vandebron. (2023, November). Superbatterijen: 4 redenen waarom ze nog niet overal zijn. Retrieved September 5, 2025, from https://vandebron.nl/blog/over-superbatterijen

Walker, W. E., Lempert, R. J., & Kwakkel, J. H. (2013). Deep Uncertainty. In *Encyclopedia of Operations Research and Management Science* (pp. 395–402). Springer, Boston, MA. https://doi.org/10.1007/978-1-4419-1153-7_1140

Xing, X., Lin, J., Song, Y., Zhou, Y., Mu, S., & Hu, Q. (2018). Modeling and operation of the power-to-gas system for renewables integration: A review. *CSEE Journal of Power and Energy Systems*, *4*(2), 168–178. https://doi.org/10.17775/CSEEJPES.2018.00260

Yue, X., Pye, S., DeCarolis, J., Li, F. G. N., Rogan, F., & Gallachóir, B. Ó. (2018). A review of approaches to uncertainty assessment in energy system optimization models. *Energy Strategy Reviews*, *21*, 204–217. https://doi.org/10.1016/j.esr.2018.06.003

Zomerdijk, W., Digvijay, G., Palensky, P., & Cvetkovic, M. (2022, October). Open Data Based Model of the Dutch High-Voltage Power System [original-date: 2024-04-05T13:16:49Z]. Retrieved January 23, 2026, from https://github.com/WZomerdijk/Dutch-HV-Power-System

A APPENDIX A: PYPSA INPUT VARIABLES

This appendix describes the data files that are required to run the PyPSA model and explains which variables are included in each file. It also provides the coordinates of the offshore wind farms used in the model. Finally, it explains how PyPSA calculates investment and operating costs within the model.

A.1 DATA FILES

The PyPSA model needs data files for the model to work. These datafiles contain specific attributes of the components they belong to. For these components and attributes, see Table A.1. How these components work together, see Chapter 5.

Excel sheet	Input variables
Buses	Name, Nice_name, Network, V_nom, X_coordinate, Y_coordinate
Carriers	Name, Color, CO_2_emissions
Generators	Bus, Control, P_nom (extendable/min/max/min_pu), Q_set, Carrier, Capital_costs, Marginal_costs, Build_year, Efficiency, Committable, Start_up_cost, min_up_time, min_down_time, ramp_limit_up, ramp_limit_down
Generators-p_max_pu	Snapshot, P_max_pu (per wind/solar generator)
Global_constraints	Name, Carrier_attribute, Sense, Constant
Lines	Name, Bus_0/Bus_1, X, R, B, S_nom(_extendable/_min/_max_pu), I_nom, V_nom, Build_year, Length, Capital_cost
Loads	Name, Bus
Loads-p_set	Snapshot, Netherlands
Network	Name, PyPSA_version
Snapshots	Snapshot
Storage_units	Name, Bus, Control, P_nom(_extendable/_nom_min/_nom_max/_min_pu), Q_set, Carrier, Capital_costs, Marginal_costs, Max_hours, State_of_charge, Cyclic_state, standing_loss, build_year, efficiency_store, efficiency_dispatch
Transformers	Name, Bus_0/Bus_1, Model, X, R, G, B, S_nom(_extendable), Build_year

Table A.1: Excel sheets and their input variables for the PyPSA model

A.2 COORDINATES OF OFFSHORE WIND FARMS

Offshore wind parks	Coordinates	Landing locations	Coordinates	Distance (km)
Borssele, kavels I en II	51.7000000, 3.0566670	Borssele	51.433252, 3.728533	59.35
Borssele, kavels III, IV en V	51.7262240, 2.9655000	Borssele	51.433252, 3.728533	66.72
Hollandse Kust (zuid), kavels I en II	52.3193300, 4.0430000	Maasvlakte	51.98369, 4.02685	40.80
Hollandse Kust (zuid), kavels III en IV	52.2576722, 4.0846915	Maasvlakte	51.98369, 4.02685	32.40
Hollandse Kust (noord), kavel V	52.6977538, 4.2936252	Beverwijk	52.487194, 4.616822	38.38
Hollandse Kust (west), kavel VI	52.6797576, 3.8046376	Beverwijk	52.487194, 4.616822	74.1
Hollandse Kust (west), kavel VII	52.6138086, 3.7373507	Beverwijk	52.487194, 4.616822	71.44
Ten noorden van de Waddeneilanden, kavel I	5.491714, 54.119494*	Eemshaven	53.443919, 6.871198	100
IJmuiden Ver, kavel Alpha	52.8294, 3.4965*	Borssele	51.433252, 3.728533	112
IJmuiden Ver, kavel Beta	52.9210, 3.5884*	Maasvlakte	51.98369, 4.02685	132
IJmuiden Ver Gamma-A en Ijmuiden Ver Gamma-B	53.070054, 3.735274*	Maasvlakte	51.98369, 4.02685	176
Nederwiek kavel I-A en I-B	52.943770, 3.306198*	Borssele	51.433252, 3.728533	175
Nederwiek, kavel II	53.204351, 3.520736*	Maasvlakte	51.98369, 4.02685	162
Nederwiek, kavel III	53.204351, 3.520736*	Geertruidenberg	51.700908, 4.833338	250
Hollandse Kust (west), kavel VIII	52.574034, 3.711955	Beverwijk	52.487194, 4.616822	67
Doordewind (kavel 1)	54.368772, 5.380426*	Eemshaven	53.443919, 6.871198	150
Doordewind (kavel 2)	54.368772, 5.380426*	Eemshaven	53.443919, 6.871198	153
Gemini I	54.0368684, 6.0418551	Eemshaven	53.443919, 6.871198	105.7
Gemini II	54.0326264, 5.8883143	Eemshaven	53.443919, 6.871198	115.7
Windpark Egmond aan Zee	52.60536, 4.41328	Beverwijk	52.487194, 4.616822	16.3
Princes Amalia Windpark	52.5855137, 4.2353666	Beverwijk	52.487194, 4.616822	27.96
Gebied 6/7	54.179241, 3.859489*	Eemshaven	240-450	

Table A.2: Offshore wind parks, their (x,y) coordinates, landing location, and the distance between them (Ministry of Economic Affairs and Climate, 2025; Ministry of Infrastructure and Water Management, 2025; Rijkswaterstaat, 2025)

A.3 COST CALCULATIONS

Marginal costs
The formula for the marginal costs of dispatch is in Equation A.1.

$$+\sum_{t} w_t^0 \left(\sum_{n,s} o_{n,s,t} g_{n,s,t} + \sum_{n,s} o_{n,s,t} h_{\bar{n},s,t} + \sum_{n,s} o_{n,s,t} h_{n,s,t} + \sum_{l} o_{l,t} f_{l,t} \right) \tag{A.1}$$

In the formula, g stands for the generators, *barh* for the storage units, h for the stores, f for the links, o for the marginal costs, and w for the snapshots.

By running the energy system over all time steps, this formula calculates the total marginal costs. It adds up the costs of electricity generation, storage charging and discharging, and power flows through network links, weighted over the full year.

Investment costs
When p, s, or e compontents are extendable, the investment costs of expanding their capacity are Equation A.2.

$$\sum_{n,s} c_{n,s} G_{n,s} + \sum_{n,s} c_{n,s} H_{n,s} + \sum_{n,s} c_{n,s} E_{n,s} + \sum_{l} c_l F_l + \sum_{l} c_l P_l \tag{A.2}$$

In the formula, n represents the bus, and s represents the specific generator or storage type connected to that bus. l refers to a branch in the network, such as a transmission line. The symbols G, H, F, and P represent the nominal power capacities, while E represents the nominal energy capacity of a storage unit.

To minimise long-term annual system costs (expressed in currency per year), investment costs for system components should be converted into annualised capital costs (currency per MW per year or currency per MWh per year). Time-step weightings (hours per year) should be chosen so that it equals to 8760 hours per year. These cost conversions and weightings are applied at a later stage of the thesis.

Energy balance The energy balance equations are the most important constraints in the model. They ensure that, at every bus and for each time step, the total incoming energy is equal to the total outgoing energy. This is similar to Kirchhoff's current law for electrical circuits. When all system components are included, this balance in the following formula is used Equation A.3.

$$\sum_{s} g_{n,s,t} + \sum_{s} \left(h^-_{n,s,t} - h^+_{n,s,t} \right) + \sum_{s} h_{n,s,t} + \sum_{l} L_{n,l,t} f_{l,t} + \sum_{l} K_{n,l} p_{l,t} = \sum_{s} d_{n,s,t} \quad \leftrightarrow \quad w_t^0 \lambda_{n,t} \tag{A.3}$$

In the formula, d represents the electricity demand of load components, g represents the power produced by generators, h- represents the discharge of storage units, h+ represents the charging of storage units, f represents the power flow through links, and p represents the power flow through transmission lines.

The incidence matrices K and L describe how electricity flows between buses through lines, transformers, and links, including the direction of flow and any efficiency losses. The dual variable lambda represents the shadow price of the constraint, the value of electricity at each bus and time step.

Storage For extendable storage units, which is what's used in this thesis, PyPSA does not fix the power capacity in advance. Instead, the model optimises the storage size by deciding how much power capacity should be installed. The charging and discharging at each time step are then limited by this optimised capacity. This allows PyPSA to choose both when the storage is used and how large it should be, so that the total system cost is minimised while still meeting demand, see Equation A.4, Equation A.5, Equation A.6 and Equation A.7.

$$h_{\bar{n},s,t} \geq 0 \tag{A.4}$$

$$h_{\bar{n},s,t} \geq \bar{h}_{n,s,t} H_{n,s} \tag{A.5}$$

$$h^{+}_{n,s,t} \geq 0 \tag{A.6}$$

$$h^{+}_{n,s,t} \geq h^{+}_{n,s,t} H_{n,s} \tag{A.7}$$

In the formulas, h- represents the discharge of storage units, h+ represents the charging of storage units, and H is the power capacity that needs to be optimised.

Linear power flow First, the optimisation step is done, afterwards the power flow calculations can be done. They work independently from each other. In the power flow step, the power imbalances from during the optimisation are used as input values. For AC networks, the linear power flow model is based on several simplifying assumptions. Reactive power is assumed to be decoupled from active power, voltage magnitudes are kept constant, voltage angle differences between connected buses are small, and line resistances are much smaller than line reactances. Under these assumptions, power flows depend mainly on the series reactances of the lines.

In this approach, the active power injections are known for all buses except one, called the slack bus. The voltage angle of the slack bus is fixed at zero and serves as a reference point. The task of the linear power flow calculation is then to determine the voltage angles at all other buses. These voltage angles are found by solving the following equation Equation A.8.

$$P_i = \sum_j (KBK^T)_{ij}\theta_j - \sum_l K_{il} b_l \theta_l^{\text{shift}} \tag{A.8}$$

In the formula, K and B are incidence and diagonal matrixes of the network. The rest of the formula is for the phase shift of a transformer.

For DC networks, the linear power flow model also uses simplifying assumptions. It is assumed that the voltage magnitude differences between connected buses are small. Under this assumption, the power flow can be calculated using only the series resistances of the lines. The linear power flow calculation for DC networks is similar to that of AC networks. The main difference is that voltage angle differences are replaced by voltage magnitude differences, and series reactances are replaced by series resistances.

B APPENDIX B: SOLVER OPTIONS

PyPSA normally uses HiGHS as its standard solver, but in practice this solver turned out to be very slow for the models used in this thesis. For that reason, Gurobi was chosen instead. To check whether this choice would affect the results, both solvers were tested and their outputs were compared.

Different solvers can produce different results, even when they solve the same optimisation problem. This is because each solver uses its own mathematical techniques, numerical tolerances, and decision rules to find the optimal solution. In this thesis, the comparison between HiGHS and Gurobi was done using the single-node model, because the transmission network was already too large. Running the transmission network with HiGHS took more than a week and often failed to finish. The version with electrolysers would have taken even longer, so no solver comparison was done for those cases.

The runtime differences between the two solvers were very large. In all tested cases, Gurobi was at least twice as fast as HiGHS. The runtimes for Gurobi are in Table B.1. These solver tests were carried out using an older time-simplification method rather than the final segmentation method. However, it is expected that the general conclusions would remain the same.

Phase	Optimised timeframe	Time (minutes)
Phase 1: Single node	Whole year	1
Phase 2: Transmission network	2 days per month per year	1
	1 week per month per year	7
	Whole year	90
Phase 3: Transmission + electrolyser network	Whole year	150

Table B.1: Time to run the model with Gurobi

The results showed that HiGHS and Gurobi do not always produce exactly the same optimal system configuration. For the single-node model with a full-year simulation, Gurobi builds almost 6 GW more solar capacity, around 4 GW more offshore wind, and about 1 GW more batteries, while building roughly 1 GW less gas capacity than HiGHS (Figure B.2). For the single-node model using a 'one week per month' time simplification, the differences are also noticeable: Gurobi builds about 2 GW less gas, 8 GW less solar, 7 GW more offshore wind, and 1 GW more batteries compared to HiGHS (Figure B.1).

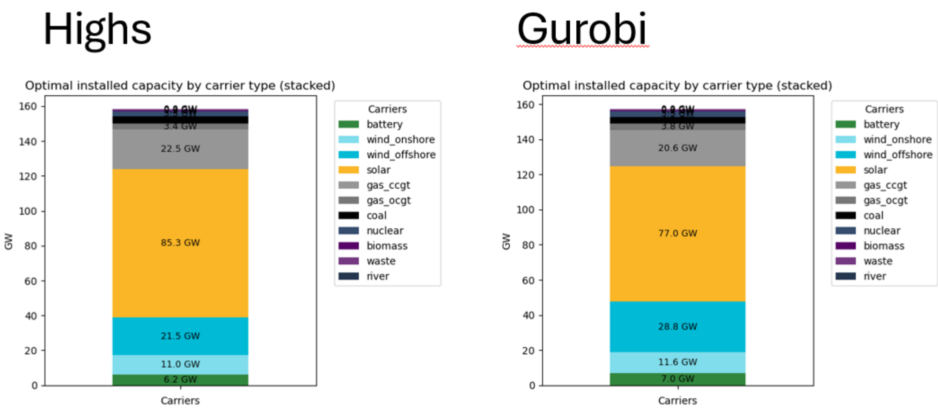


Figure B.1: Installed capacities test Highs and Gurobi: 1 week per month

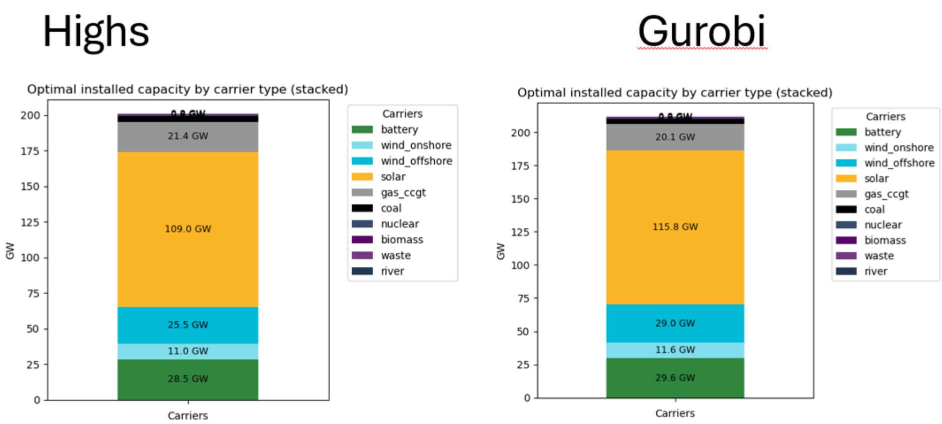


Figure B.2: Installed capacities test Highs and Gurobi: 1 year

These differences can be explained by how the solvers work internally. Gurobi uses highly optimised, proprietary algorithms for linear and mixed-integer optimisation, including advanced presolve techniques and heuristics. HiGHS, on the other hand, is open-source and mainly relies on simplex and interior-point methods, with different numerical settings and tolerances. Because the optimisation problem often has many solutions that are almost equally optimal, both solvers can choose different points within this feasible solution space. As a result, small differences in installed capacities can appear, even though the total system cost and overall system performance remain very similar.

Because of its much shorter runtimes and better ability to handle large and complex models, Gurobi was chosen as the solver for the main analyses in this thesis.

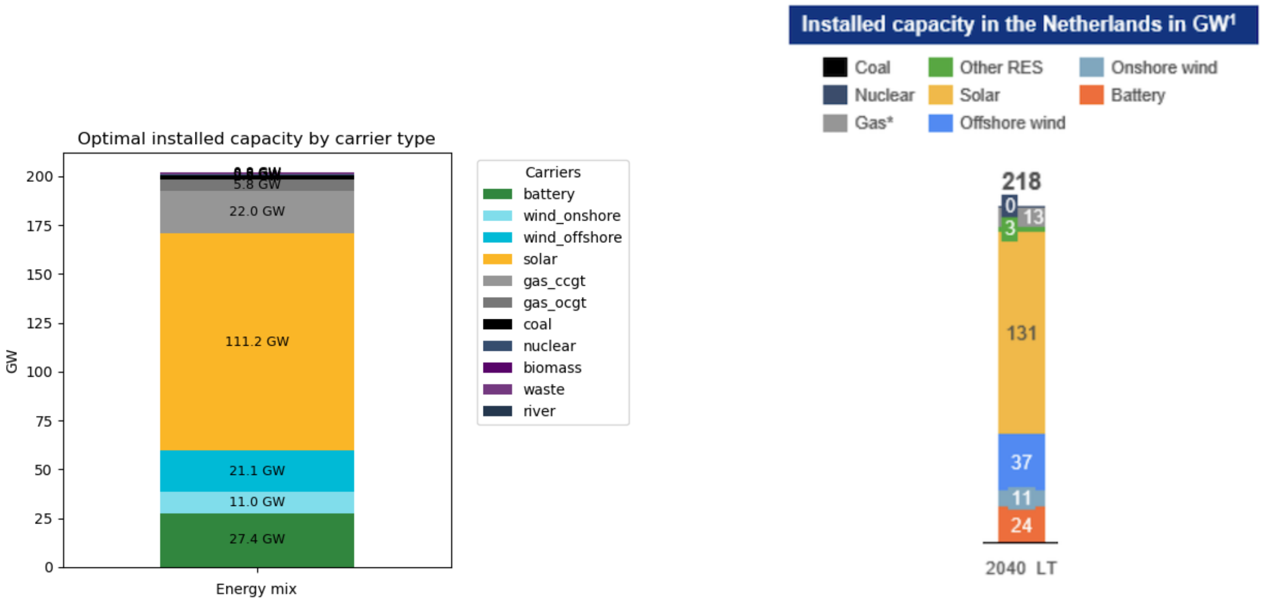
C APPENDIX C: RESULTS

This appendix contains additional results that could not be discussed in detail in the main chapters because of their size and level of detail. These results support the findings in the results and discussion chapters, but are not essential for understanding the main conclusions of the thesis. It includes both PyPSA and EMA Workbench outputs. For PyPSA, this mainly involves extra figures and comparisons related to installed capacities, load flows, and congestion patterns. For the EMA Workbench, it shows additional plots on archetypes and dispatch.

C.1 PYPSA RESULTS

C.1.1 Phase 1: Single node

The main text already describes the reason for the difference in installed capacities for the PyPSA and PLEXOS model, Chapter 7. However for full transparency, the full installed capacity plots from can be seen for the single node analysis Figure C.1.



(a) Optimised installed capacities PyPSA single node

(b) Optimised installed capacities Guidehouse single node

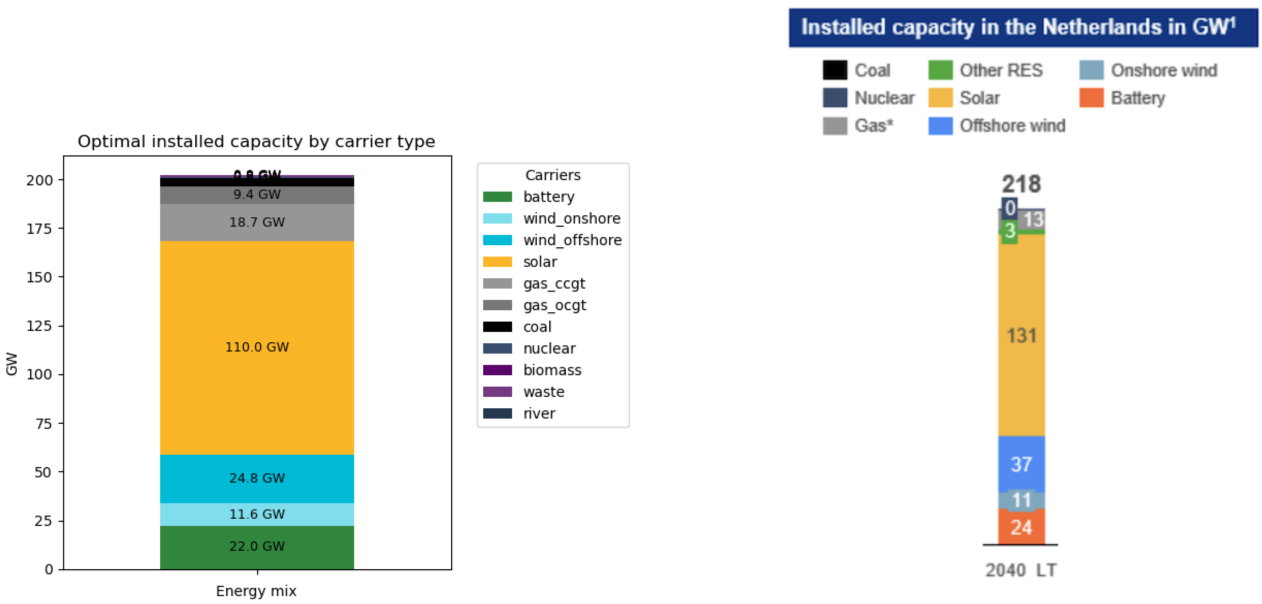
Figure C.1: Optimal installed capacities by PyPSA and Guidehouse for single node

C.1.2 Phase 2: Transmission network

Optimal installed capacities

When comparing the optimised installed capacities from Guidehouse and TenneT (2025) with the results from the PyPSA model, both approaches lead to the same overall conclusion, as shown in Figure C.2. In both models, solar energy clearly forms the largest share of the total installed capacity. Offshore wind comes second and also plays a major role in supplying electricity. Batteries and gas-fired power plants are important supporting technologies, mainly used to balance the system and provide flexibility. Onshore wind and other technologies only make up a relatively small part of the total installed capacity.

The main text already explains why there are differences between the PyPSA and PLEXOS results. What Figure C.2 adds is a clearer picture of these differences in absolute terms, rather than relative changes between technologies. When looking at the total system size, the differences between the two models are actually quite small. More importantly, the remaining differences are logical and can be explained by the modelling choices, such as how grid constraints and flexibility options are treated. This means that both models broadly agree on the overall system design, even if the exact numbers are not identical.



(a) Optimised installed capacities PyPSA transmission network

(b) Optimised installed capacities Guidehouse single node

Figure C.2: Optimal installed capacities by PyPSA and Guidehouse for transmission network

The plot in Figure C.3 adds more depth by showing where PyPSA places generation capacity and how electricity is expected to flow through the grid. Each circle represents a substation (or bus) in the model, and the colours show which technologies are installed at each bus Figure C.1.

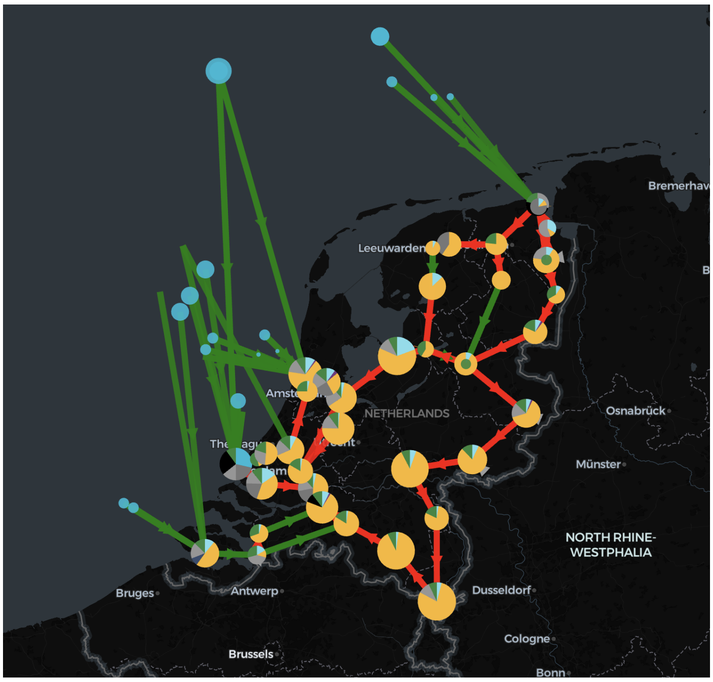


Figure C.3: Explore plot of the transmission network in PyPSA

Color	Technology
Yellow	Solar
TU Delft blue	Offshore wind
Light blue	Onshore wind
Green	Batteries
Dark grey	Gas OCGT
Light grey	Gas CCGT
Black	Coal
Dark blue	Nuclear
Purple	Biomass

Table C.1: Colors of explore plot

From this figure, a clear pattern can be seen. In large industrial clusters, a lot of offshore wind energy is connected directly to the grid. These areas show relatively little installed solar capacity, but instead rely more on gas CCGT and gas OCGT plants (and offshore wind). This supports the explanation given earlier: when transmission constraints are included, it becomes more cost-efficient to use offshore wind that is located close to large electricity consumers. Supplying these industrial clusters with offshore wind reduces the need to transport large amounts of solar electricity from inland areas, which would otherwise increase congestion on the grid.

In contrast, in non-industrial and inland regions, the combination of solar energy, batteries, and onshore wind appears sufficient to meet local demand. These areas do not need large amounts of offshore wind or gas capacity, because demand is lower and can be balanced more easily with local renewable generation and storage.

Load flows

When comparing the direction of power flows between the PyPSA results and the Grid Planning results, 31 transmission lines show the same flow direction, while 10 lines show a mismatch. This means that for most of the grid, PyPSA correctly shows the main direction in which electricity flows through the network. The mismatches are limited to a relatively small number of lines, which are summed up below:

- Breukelen Kortrijk - Diemen
- Beverwijk - Vijfhuizen
- Eemshaven - Robbenplaat

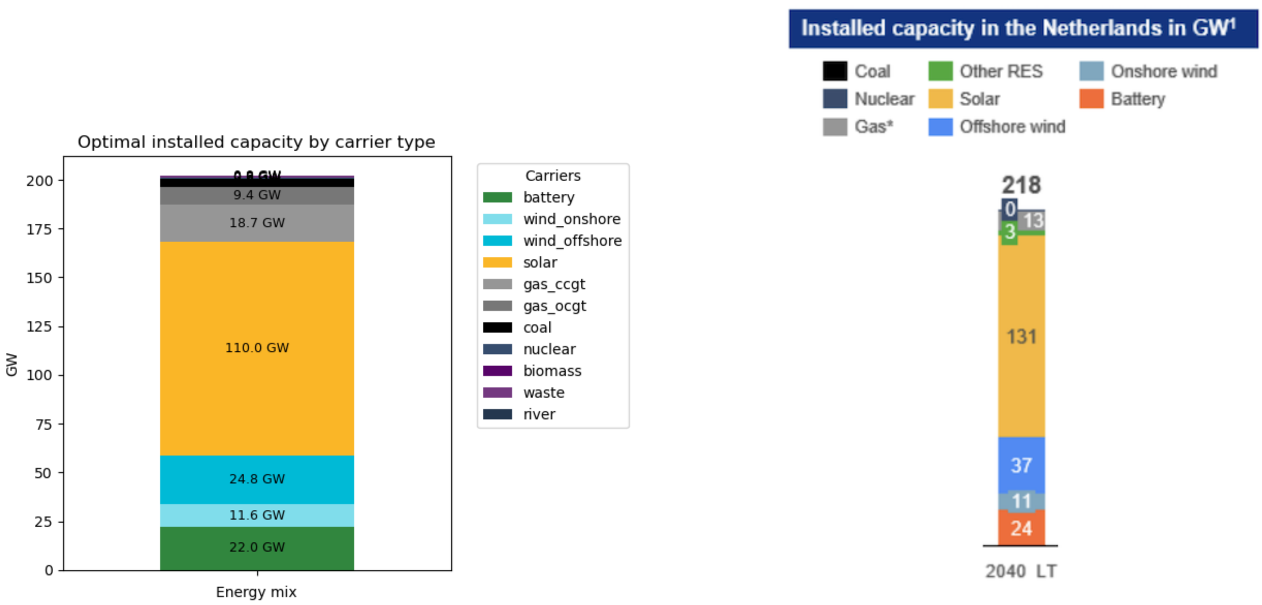
- Krimpen aan de Ijssel - Diemen
- Maasbracht - Boxmeer
- Maasbracht - Eindhoven
- Oostzaan - Diemen
- Robbenplaat - Weiwerd
- Tilburg - Eindhoven
- Weiwerd - Meeden

As discussed in the main text, there are clear reasons why these mismatches occur. Most of them can be explained by simplifications in the PyPSA model, such as the clustering of nearby substations, which slightly changes the network layout. In addition, differences in assumed installed capacities can shift local supply–demand balances and therefore reverse flow directions on specific lines.

C.1.3 Phase 3: Transmission network + electrolysers

Optimal installed capacities

As for the previous phase, the main text already explains the reasons for the differences between the PyPSA and PLEXOS models. Figure C.4 helps to make these differences clearer by showing the total installed capacities in absolute terms, rather than only relative shares. When looking at the figure this way, it becomes clear that the differences between the two models are relatively small.



(a) Optimised installed capacities PyPSA transmission model + electrolysers

(b) Optimised installed capacities Guidehouse single node

Figure C.4: Optimal installed capacities by PyPSA and Guidehouse

Additional insight is provided by the explore plot in Figure C.5, which shows where PyPSA places electricity generation, electricity transport, and hydrogen infrastructure in the system. Compared to the explore plot without electrolysers, this figure also includes the hydrogen network, shown by the purple lines. This makes it possible to see how hydrogen production and transport interact with the electricity grid.

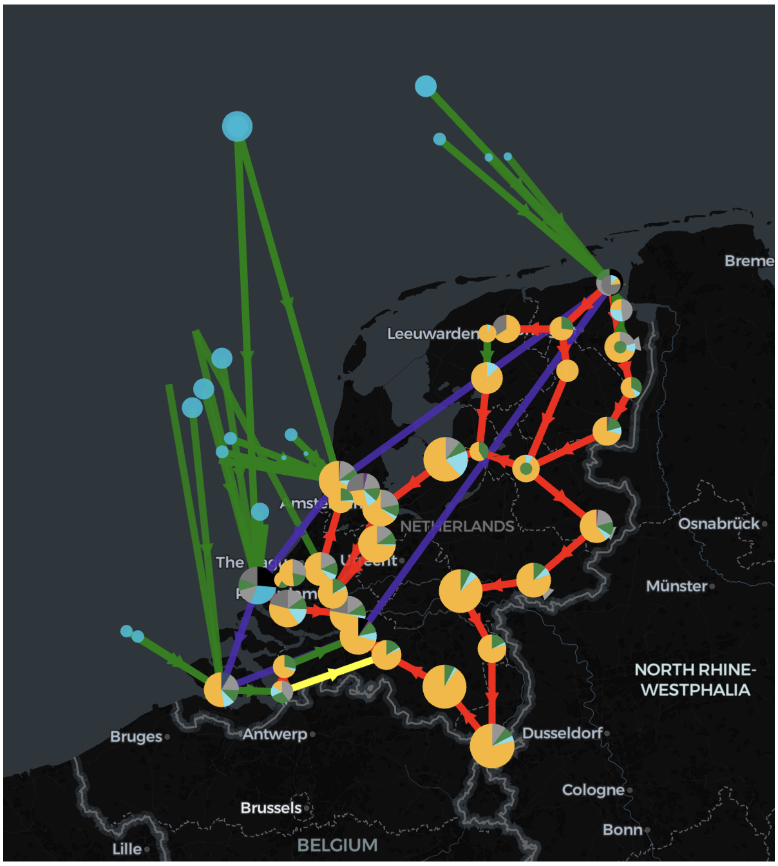


Figure C.5: Explore plot of the transmission + electrolyser network in PyPSA

One noticeable difference compared to the transmission-only phase in Figure C.3 is that the line between Rilland and Tilburg becomes slightly overloaded. This is likely caused by the additional electricity demand and supply introduced by electrolysers at Geertruidenberg. Hydrogen production increases power flows in certain parts of the grid, which adds extra pressure to some transmission corridors and can lead to new or stronger congestion.

Load flows

In this phase, several transmission lines show a different flow direction compared to the Grid Planning results. These lines are:

- Breukelen Kortrijk - Diemen*
- Boxmeer - Dodewaard

- Beverwijk - Vijfhuizen*
- Crayestein - Krimpen aan den Ijssel
- Hessenweg - Ens
- Krimpen aan den Ijssel - Breukelen Kortrijk
- Krimpen aan de Ijssel - Diemen*
- Maasbracht - Boxmeer*
- Maasbracht - Eindhoven*
- Robbenplaat - Weiwerd*
- Tilburg - Eindhoven*
- Weiwerd - Meeden*

The lines marked with an asterisk (*) already showed a mismatch in flow direction in Phase 2, meaning that their behaviour has not changed with the addition of electrolysers. The newly added mismatches are mainly located in the Maasvlakte area (two lines) and the Eemshaven area (two lines). These are exactly the regions where electrolysers and additional electricity demand have been added. This suggests that adding hydrogen production changes local supply–demand balances and can alter the direction of power flows on nearby transmission lines.

Overall, the results indicate that electrolysers have a clear and localised impact on grid behaviour, while the broader flow patterns in the network remain largely consistent with earlier phases.

C.2 EMA WORKBENCH RESULTS

For the EMA results, the focus is mainly on the archetypes that come out of the analysis. The detailed relationships between technologies and system indicators have already been explained in the main text. This section shows mainly the archetypes, as these best summarise the key outcomes of the EMA analysis and make the results easier to interpret.

C.2.1 Phase 1: Single node

A visual overview of the archetypes discussed in Section 7.2 is shown in Figure C.6. This figure presents the different representative system configurations that result from the EMA analysis for the single-node model. The corresponding cost factors that define each archetype are listed in Table C.2. Together, these figures show how changes in technology costs lead to different typical system designs.

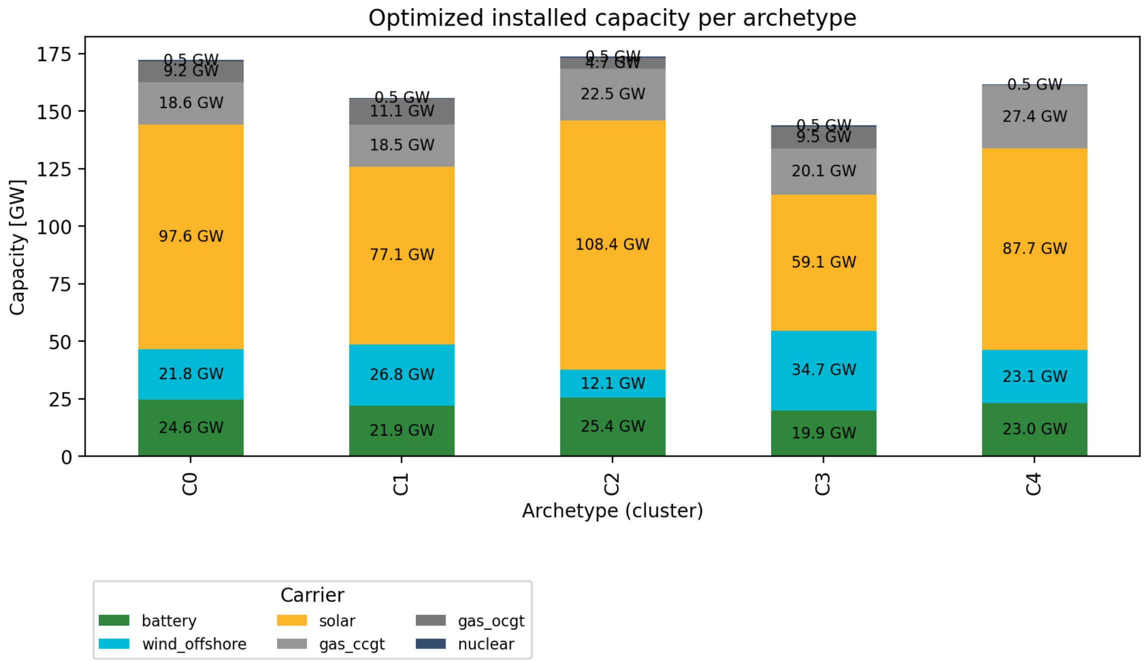


Figure C.6: Archetypes for single node

Archetype	Battery	Gas CCGT	Gas OCGT	Nuclear	Solar	Offshore wind
C0	1.11	1.34	1.21	1.38	1.13	1.06
C1	1.09	1.17	0.98	1.11	1.27	0.97
C2	1.29	1.28	1.29	1.01	1.23	1.36
C3	0.95	0.90	0.82	1.37	1.35	0.84
C4	1.17	0.93	1.29	1.13	1.19	0.96

Table C.2: Capital cost factors for single node archetype plot

Figure C.7 shows the violin plot that was discussed earlier in the main text. This plot gives additional insight into how widely the installed capacities of the different technologies vary across all scenarios in the single-node model. It helps illustrate which technologies show a lot of variation between runs and which ones remain relatively stable.

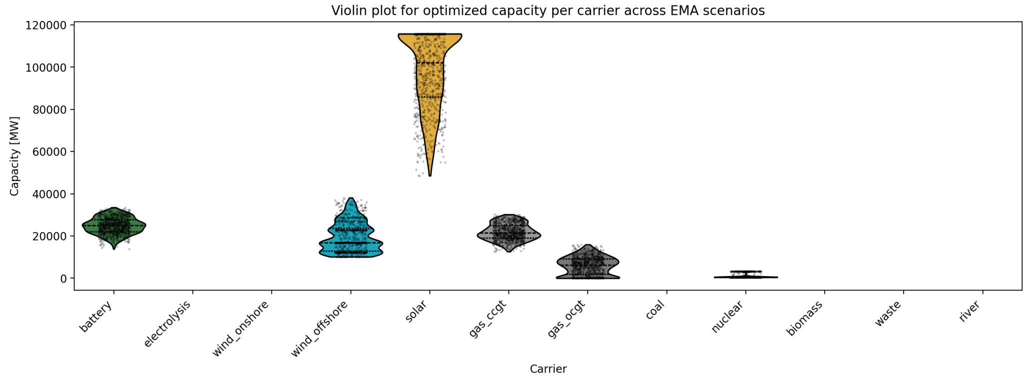


Figure C.7: Violin plot for single node

C.2.2 Phase 2: Transmission network

Although Phase 2 was not discussed in the main text, in this part of the appendix its archetypes and the corresponding cost factors are shown for transparency: Figure C.8 and Table C.3.

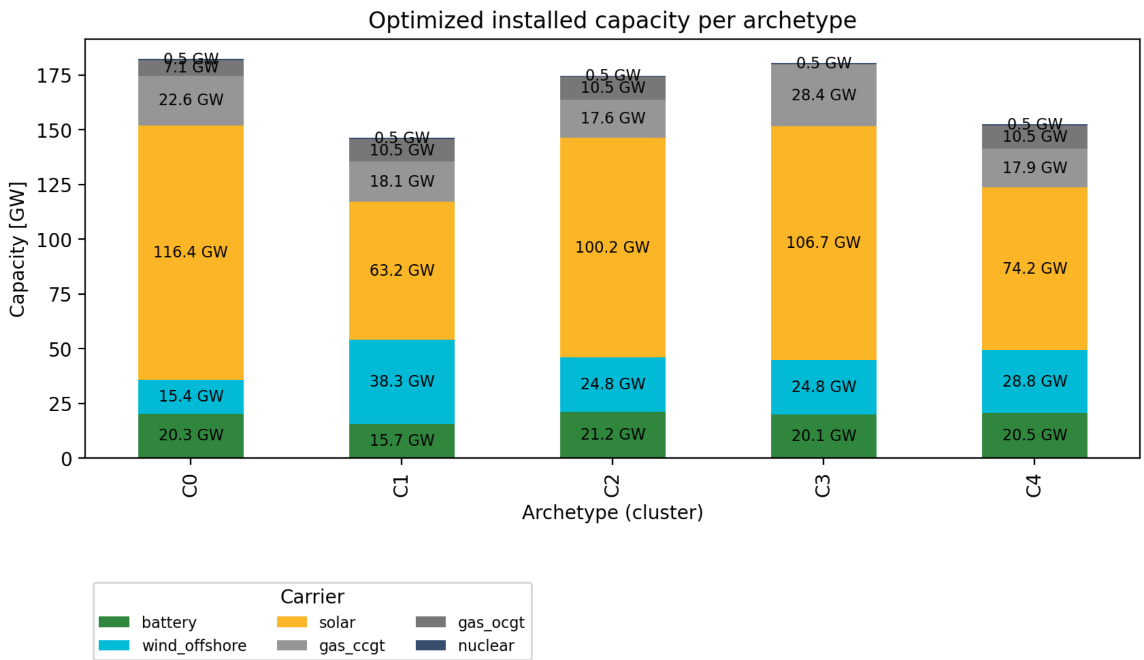


Figure C.8: Archetypes for transmission network

Archetype	Battery	Gas CCGT	Gas OCGT	Nuclear	Solar	Offshore wind
C0	1.28	0.90	0.96	0.98	0.83	1.27
C1	1.27	1.23	0.98	1.08	1.22	0.82
C2	1.00	1.23	0.81	1.09	1.12	1.04
C3	1.15	0.83	1.38	1.27	1.00	0.97
C4	0.82	1.08	1.09	1.00	1.39	1.00

Table C.3: Capital cost factors for transmission network archetype plot

C.2.3 Phase 3: Transmission + electrolysers

The dispatch plot that is discussed in the main text is shown in Figure C.9. This figure helps explain how different technologies are actually used over the year, rather than only how much capacity is installed. The dispatch plot clearly shows the role of gas CCGT in the system in the first month of the year. It runs so frequently, that a large amount of gas is burned, which increases both emissions and fuel costs, just as shown in the Theil-Sen plot.

At the same time, the dispatch plot also helps explain why gas CCGT lowers total CAPEX. Because CCGT provides large amounts of reliable electricity whenever needed, the system does not have to invest as much in expensive technologies with higher investment costs, such as offshore wind or large battery systems. This makes the overall system cheaper to build.

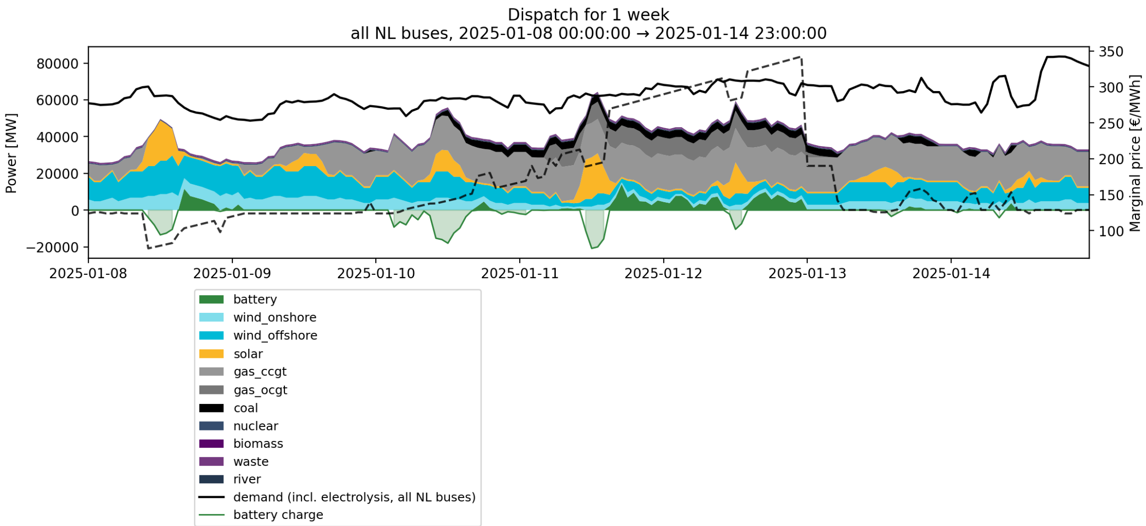


Figure C.9: Dispatch for transmission + electrolyser network