

Reactive Power Planning and Forecasting in Renewable-Integrated Power Systems

by

Arnaud Rosseel

A thesis presented for the degree of
Philosophiae Doctor in Engineering Sciences & Technology
pursued in the Electrical Power Engineering Unit of the
University of Mons

Members of the Jury

Prof. Jacques Lobry, Université de Mons Chairman

Prof. Florin Capitanescu, Luxembourg Institute of Science and
Technology

Prof. Pierre Henneaux, Université Libre de Bruxelles

Dr Jonathan Sprooten, Elia

Prof. François Vallée, Université de Mons

Prof. Jean-François Toubéau, Université de Mons Secretary

Prof. Zacharie De Grève, Université de Mons Supervisor

Prof. Bashir Bakhshideh Zad, Université de Mons Co-supervisor

September 2026

The most important step a man can take.
It's not the first one, is it?
It's the next one. Always the next step,
Dalinar.

Brandon Sanderson, Oathbringer

Acknowledgements

Je n'aurais jamais commencé cette thèse sans les relances incessantes de François Vallée pendant toute ma deuxième année de master. À l'époque j'étais un étudiant plus intéressé par les amis et les guindailles que par les cours et les projets. Ensuite, lorsque le Covid-19 est arrivé, je me suis pleinement concentré sur mon travail de fin d'études et c'est à ce moment-là que j'ai compris que je voulais faire une thèse. Merci François de toujours être ce moteur pour motiver les étudiants à se lancer dans la recherche.

Ensuite, je tiens à remercier mes promoteurs Zacharie De Grève et Bashir Bakhshideh Zad pour leur encadrement, leur soutien et leurs conseils tout au long de mon parcours. Merci Zacharie d'avoir toujours été à l'écoute et de m'avoir aidé à résoudre mes problèmes lorsque j'exprimais mes doutes. Même si parfois ça ne durait que 5 minutes autour d'un café, ces moments ont été très précieux pour moi. Merci Bashir pour m'avoir suivi au plus proche pendant ces 6 dernières années, pouvoir venir toquer à ton bureau à n'importe quel moment m'a beaucoup aidé. On est parfois très têtus tous les deux, mais je pense que c'est ce qui m'a permis de progresser et d'argumenter au mieux mes idées. Merci à tous les deux de m'avoir toujours poussé à aller plus loin dans la recherche.

Merci à l'ensemble de mes collègues pour ce magnifique environnement de travail. Je pense sincèrement que je n'aurais pas pu avoir une meilleure ambiance au travail. Tout au long de cette thèse, j'ai côtoyé des personnes incroyables. Un grand merci à Marie, Louise, Julien, Apolline, Chloé, Clélia, Cyril, Thuy-Hai, Martin, Maxime, Noé, Manu, Jenö, Jean-François, Victor, Pietro, Hooman, Hossein, Behzad, Ahmad, Jamal, Olivier, Thomas, Camilo, William, Loïc. Pour toutes ces discussions enrichissantes, ces parties de ping-pong endiablées, ces activités en dehors du travail, MERCI. Un merci spécial à Marie et Louise qui m'ont supporté directement dans le même bureau, j'espère que ma bonne humeur (et mes râleries) ne vous manquera pas trop !

Je ne peux pas ne pas remercier mon Big Bro, Martin Scohier, aka Martin Sempaï (selon ses propres mots). Depuis notre rencontre lors de cette bleusaille en 2015, on a plus su se quitter. Les premiers TPs ensemble, les groupes de travail, les veilles d'examen à s'expliquer la matière l'un à l'autre, les projets, les voyages, la coloc. Le doctorat et l'assistanat ensemble, les activités sportives, les débats philosophiques et politiques. C'est la fin d'une sacrée époque mais sûrement pas la fin de notre amitié. Merci pour tout Martin.

Je pense que personne ne peut se construire sans l'aide des autres. Bien sûr, j'ai travaillé pour arriver où j'en suis aujourd'hui. Mais je n'en serais jamais arrivé là sans tous mes proches. Tout d'abord, merci à ma famille de m'avoir soutenu sans jamais douter de moi. Merci papa, merci maman de m'avoir élevé avec toutes ces belles valeurs et surtout de m'avoir toujours poussé à donner le meilleur de moi-même dans la bienveillance. Merci mon frère Eliès, on est peut-être différents mais on a toujours su se comprendre et s'entraider et c'est aussi cette différence qui m'a permis de grandir. Merci à tous mes amis pour tous les moments partagés.

Les remerciements les plus importants vont à ma merveilleuse épouse Charlotte. Merci de me soutenir et de m'encourager quelque soit l'aventure dans laquelle je me lance. Merci d'être un pilier dans ma vie et de toujours croire en moi. Merci pour les 10 années déjà passées ensemble et vivement les 50 prochaines.

Merci tout le monde, vous avez contribué d'une manière ou d'une autre à faire de moi une meilleure personne.

Declaration on the Use of Artificial Intelligence Tools: During the preparation of this thesis, the author used Claude (Anthropic) language models for three purposes: improving the phrasing and clarity of the written English text, supporting bibliographic research, in particular the identification and organization of relevant literature, and assisting with the writing of code. All content generated or suggested with the assistance of these tools was reviewed, verified, and edited by the author, who takes full responsibility for the content of this thesis, including all bibliographic references and the correctness of the underlying code.

Summary

The energy transition is reshaping the reactive power balance of high-voltage transmission networks. The large-scale integration of renewable generation, connected predominantly at the distribution level, and the progressive replacement of overhead lines by underground cables are shifting these networks from historically inductive operating conditions toward increasingly capacitive ones. Because reactive power does not travel efficiently over long distances, it must be managed close to where it is injected or absorbed. How reactive power management should adapt to this shift remains an open question.

This thesis addresses this question through two complementary contributions, developed in turn at two levels of the network: the transmission system level, where the system operator decides where and how much reactive compensation capacity to install, and the transmission-distribution interface level, where the actual reactive behaviour that these decisions must anticipate can be characterised.

The first contribution, VARPlan, is a security-constrained, scenario-based methodology for reactive power planning at the transmission system level. It jointly optimises investment in compensation assets and the operational cost of dispatching them, evaluated across representative operating conditions and contingencies through a second-order cone relaxation of the power flow equations. Distinctively, this operational cost reflects the reactive dispatch cost actually borne by a system operator, rather than a proxy quantity such as active losses or voltage deviation, and a weak bus analysis restricts the investment search to the network locations where compensation is most needed. On the IEEE 14-bus test case, the relaxation was found to lose exactness under capacitive loading, a regularisation term was introduced to restore it, and a reactive scarcity indicator derived from the problem itself was shown to identify effective investment locations with far fewer candidate buses than conventional voltage stability indicators.

The second contribution, VARLearn, forecasts active and reactive power exchanges at transmission-distribution interfaces, quantities that the transmission

system operator cannot characterise through direct network modelling, since it has no access to distribution network topology. The approach combines an inverse load flow formulation, which identifies the parameters of an equivalent two-bus representation of the distribution network from historical interface measurements, with a machine learning regression that predicts these parameters from aggregate load conditions. Benchmarked against a purely data-driven predictor on a representative distribution network, this hybrid approach achieved a normalised root-mean-square error of about one percent, against more than four percent for the data-driven benchmark, and reduced the cost of curative redispatch attributable to forecasting errors by a factor of 3.5, an advantage preserved under realistic levels of input uncertainty.

Taken together, these two contributions respond to the same underlying requirement. Reactive power management must rely on a representation of the network that stays accurate as it moves toward increasingly capacitive operation. VARPlan and VARLearn are connected thematically, since both stem from the same physical need to manage reactive power close to where it is injected or absorbed, but they operate at two complementary decision timescales, VARPlan over a multi-year investment horizon and VARLearn at the quarter-hourly horizon of interface forecasting. Using VARLearn’s interface forecasts as inputs to VARPlan’s scenario generation is identified as the most direct perspective opened by this work.

Contents

Acronyms	XIII
1. General Introduction	3
1.1. Fundamentals of Reactive Power	4
1.2. Reactive Power Behaviour in Transmission Networks under the Energy Transition	9
1.2.1. The Q-V Relationship in HV Networks	9
1.2.2. Reactive Behaviour of Overhead Lines and Underground Cables	10
1.2.3. Reactive Balance and Its Evolution under the Energy Tran- sition	12
1.3. Open Questions and Research Motivation	14
1.4. Research Objectives and Thesis Contributions	15
1.5. List of Publications	16
2. VARPlan: Reactive Power Planning in HV Transmission Net- works	19
2.1. Problem Setting and State of the Art	19
2.1.1. European Regulatory and Economic Framework	20
2.1.2. From Compensation Resources to Behavioural Classes . . .	23
2.1.3. From Optimal Power Flow to Optimal Reactive Power Dis- patch	24
2.1.4. Convex Relaxations of the AC Optimal Power Flow	26
2.1.5. Conic Relaxation under Capacitive Loads	30
2.1.6. Reactive Power Planning	33
2.1.7. Security in Planning	34
2.1.8. Decision-Space Reduction and Weak Bus Identification . .	36
2.1.9. Synthesis, Research Gaps and Contributions	37
2.2. Methodology: the VARPlan Pipeline	39
2.2.1. Market Clearing Imitation (MCI / SC-OPF)	40

2.2.2.	Security-Constrained ORPD (SC-SOCP)	45
2.2.3.	Weak Bus Analysis	49
2.2.4.	Reactive Power Planning (SC-SOCP RPP)	51
2.2.5.	Summary of Modelling Assumptions	52
2.3.	Scenario Generation and Operating Point Construction	53
2.3.1.	Source Data	54
2.3.2.	Nodal Load Construction	54
2.3.3.	Wind Capacity Scaling and Injection	55
2.3.4.	Two-Stage Clustering of Operating Points	55
2.3.5.	Policy Trajectories	56
2.4.	Simulation Framework and Test Case	58
2.4.1.	Test network	58
2.4.2.	Cost parameters	60
2.4.3.	Computational setup	61
2.5.	Results	62
2.5.1.	Validation of the SOCP formulation	62
2.5.2.	Weak Bus Analysis	67
2.5.3.	Trajectory Comparison	69
2.5.4.	Sensitivity to the Number of Operating Conditions	71
2.5.5.	Sensitivity to the Number of Contingencies	72
2.6.	Discussion and Limitations	73
2.6.1.	Technical interpretation	73
2.6.2.	Limitations	74
2.7.	Chapter Conclusion	76
3.	VARLearn: Reactive Power Forecasting at TSO/DSO Interfaces	77
3.1.	Introduction	77
3.1.1.	Power exchanges at T-D interfaces: context and challenges	77
3.1.2.	Existing approaches and their limitations	79
3.1.3.	Research gaps and contributions	81
3.2.	Methodology	82
3.2.1.	Step 1 — Database creation via Inverse Load Flow	83
3.2.2.	Step 2 — Machine learning algorithms for equivalent impedance prediction	86
3.2.3.	Step 3 — Performance evaluation framework	89
3.3.	Test Case and Simulation Parameters	92
3.3.1.	Distribution network	93
3.3.2.	Transmission network	94
3.3.3.	Data parameters and benchmarks	95
3.4.	Results	97
3.4.1.	Statistical evaluation	97
3.4.2.	Application-based evaluation	98

3.4.3. Uncertainty sensitivity analysis	99
3.5. Discussion and Limitations	100
3.5.1. Interpretation of the results	101
3.5.2. Limitations of the proposed approach	101
3.6. Chapter Conclusion	102
4. General Conclusion and Perspectives	105
4.1. Summary of Contributions	105
4.1.1. VARPlan: Reactive Power Planning at the System Scale .	105
4.1.2. VARLearn: Reactive Power Forecasting at the Local Scale	106
4.2. Limitations and Open Questions	106
4.3. Future Research Directions	107
4.3.1. VARPlan: Model Refinements and Scalability	107
4.3.2. VARPlan: Towards Operational Deployment	109
4.3.3. VARLearn: Model Extensions and Physical Interpretation	109
4.3.4. Towards a Joint VARLearn-VARPlan Framework	110
4.4. Closing Remarks	111
A. Derivations for the Pi-Equivalent Line Model	127
A.1. Linearised Voltage Drop Relationship	127
A.2. Exact Power Flow Equations	128

Acronyms

DP	Direct Prediction.
DSO	Distribution System Operator.
GB	Gradient Boosting.
HV	High Voltage.
ILF	Inverse Load Flow.
KNN	K-Nearest Neighbours.
LV	Low Voltage.
MCI	Market Clearing Imitation.
MINLP	Mixed-Integer Nonlinear Program.
MISOCP	Mixed-Integer Second-Order Cone Programming.
ML	Machine Learning.
MLP	Multilayer Perceptron.
MV	Medium Voltage.
NRMSE	Normalised Root Mean Square Error.
OL	Overhead Line.
OPF	Optimal Power Flow.
ORPD	Optimal Reactive Power Dispatch.
PV	Photovoltaic.
QC	Quadratic Convex.
RES	Renewable Energy Sources.

RF Random Forest.

RPP Reactive Power Planning.

SC-OPF Security-Constrained Optimal Power Flow.

SIL Surge Impedance Loading.

SLPs Synthetic Load Profiles.

SOCP Second-Order Cone Programming.

SPPs Synthetic Production Profiles.

STATCOM Static Synchronous Compensator.

SVC Static VAR Compensator.

T-D Transmission-Distribution.

TSO Transmission System Operator.

UC Underground Cable.

VSI Voltage Sensitivity Index.

WBA Weak Bus Analysis.

CHAPTER 1.

General Introduction

The energy transition is imposing a structural transformation on European power systems. Historically, power flowed from centralized generators to passive consumers in a stable, top-down configuration. In High Voltage (HV) transmission networks, bus voltage must be kept within the bounds required for the safe operation of equipment, a task that falls to the Transmission System Operator (TSO). This task depends critically on reactive power, since its local availability determines whether bus voltage remains within these bounds. This arrangement is now being disrupted by the large-scale deployment of Renewable Energy Sources (RES), primarily Photovoltaics (PVs) and wind power. Being inherently decentralized and intermittent, and connected predominantly at the distribution level, these sources introduce new dynamics that propagate up to the HV transmission level, altering the reactive power balance that voltage regulation depends on. As variable and distributed resources continue to expand, this balance becomes increasingly difficult to ensure. A fragmented regulatory framework, in which TSOs and Distribution System Operators (DSOs) share, yet do not always coordinate, responsibility over reactive power, adds a further layer of complexity. This chapter addresses precisely these challenges.

Section 1.1 establishes the physical foundations of reactive power, from the elementary behaviour of AC circuits to the power triangle and the power factor. Section 1.2 grounds these concepts in the reality of HV networks, examining the relationship between reactive power and voltage, the contrasting behaviour of Overhead Lines (OLs) and Underground Cables (UCs), and the shift in reactive power balance driven by the energy transition. Section 1.3 identifies the open questions that emerge from this context and that motivate the present investigation. Section 1.4 presents the two main contributions of the thesis and explicitly delimits its scope.

1.1. Fundamentals of Reactive Power

Electrical power can be distributed either as direct current (DC), whose voltage and current do not vary in time, or as alternating current (AC), whose voltage and current oscillate periodically. The European transmission grid operates in AC at a fixed frequency of 50 Hz, so that one complete cycle spans an angular interval of 2π radians, equivalent to a period of 20 ms.

Typical power system analysis represents an ideal voltage source as a sinusoidal signal $v(t)$, characterised by an amplitude, a frequency, and a phase, as shown in (1.1). A load connected to such a source carries a current $i(t)$; assuming a linear voltage-to-current relationship, typical of power system loads, this current is likewise sinusoidal, at the same frequency as the voltage though not necessarily in phase with it, as shown in (1.2):

$$v(t) = \sqrt{2} V \cos(\omega t + \theta_u), \quad (1.1)$$

$$i(t) = \sqrt{2} I \cos(\omega t + \theta_i) \quad (1.2)$$

where V and I denote root-mean-square values, $\omega = 2\pi f$ the angular frequency, and θ_u , θ_i the respective phases of the voltage and the current.

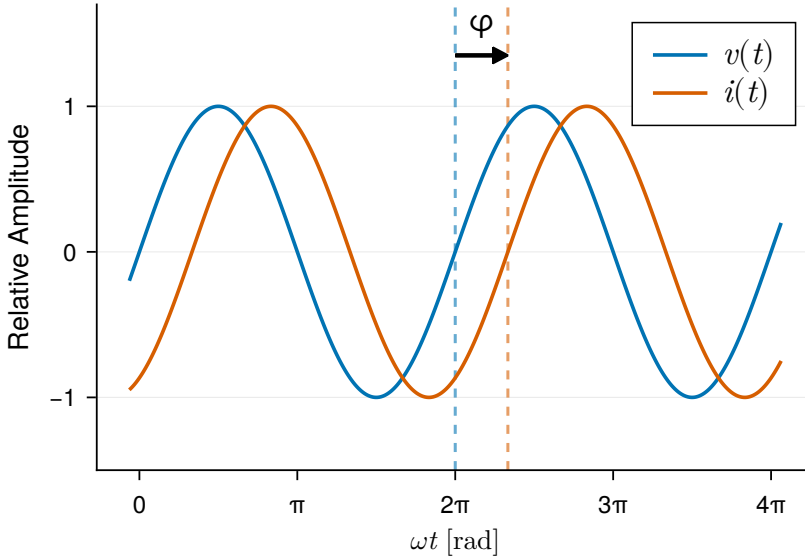


Figure 1.1.: Phase shift φ between $v(t)$ and $i(t)$.

Fig. 1.1 shows $v(t)$ and $i(t)$ plotted on a common time axis. As already noted, the two waveforms share the same frequency but are shifted relative to one another, and the instants at which each signal crosses zero do not coincide. This shift,

indicated by the horizontal arrow between the two dashed vertical lines marking the respective zero-crossings, corresponds to the phase angle $\varphi = \theta_u - \theta_i$. Depending on the nature of the load, the current may either lag behind the voltage ($\varphi > 0$) or lead it ($\varphi < 0$).

The phase shift observed in Fig. 1.1 has a precise physical origin in the nature of the elements composing the load. AC systems are fundamentally governed by the variation of electromagnetic quantities. In addition to Ohm's law,

$$v(t) = R i_R(t) \quad (1.3)$$

the voltage across the inductor and the capacitor is related to their respective currents through linear differential equations that capture the rate of change of these electromagnetic quantities:

$$v(t) = L \frac{di_L(t)}{dt}, \quad (1.4)$$

$$v(t) = \frac{1}{C} \int i_C(t) dt \quad (1.5)$$

These relations follow from Maxwell's equations [1], and highlight the dissipative nature of the resistive element as opposed to the energy transfer mechanisms of the inductive and capacitive elements.

The circuit in Fig. 1.2 provides the reference framework. It consists of an AC voltage source connected to a parallel combination of a resistor R , an inductor L , and a capacitor C .

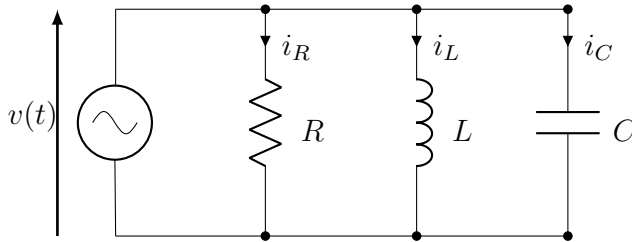


Figure 1.2.: Parallel RLC circuit fed by an AC voltage source.

This represents the simplest possible power system, in which R represents the energy consumed by the connected equipment, while L and C represent its inherent inductive and capacitive transfer behaviour.

- For a purely resistive load, (1.3) shows that the resistive current i_R remains in phase with the voltage at all times ($\varphi = 0$), since $i_R(t) = v(t)/R$ is directly proportional to $v(t)$ at every instant. The energy delivered by the

Chapter 1. General Introduction

source is entirely and irreversibly dissipated. No energy is returned to the source, the power transfer is strictly unidirectional.

- The behaviour of an inductive element, governed by (1.4), is fundamentally different. When a sinusoidal voltage is applied across an ideal inductor (i.e. linear and lossless), the inductive current i_L lags behind the voltage by exactly 90° ($\varphi = +90^\circ$). During the first half of each cycle, energy is drawn from the source and stored in the magnetic field surrounding the conductor; during the second half, this energy is returned to the network. Over a complete cycle, the net energy consumed is strictly zero.
- A capacitive element, governed by (1.5), exhibits the symmetric behaviour: the capacitive current i_C leads the voltage by 90° ($\varphi = -90^\circ$). Energy is alternately stored in the electric field surrounding the conductor and returned to the network. Here again, no net energy is consumed over a full cycle.

This establishes a key physical intuition: certain circuit elements continuously borrow energy from the source and return it within the same cycle, without ever permanently consuming it. It is precisely this oscillatory, non-dissipative exchange that underlies the concept of reactive power.

The instantaneous power exchanged between the source and the load is the product $p(t) = v(t) \cdot i(t)$. Without loss of generality, the time origin can be chosen such that the voltage is the phase reference, i.e. $\theta_u = 0$, so that $\theta_i = -\varphi$; (1.1) and (1.2) then reduce to $v(t) = \sqrt{2} V \cos(\omega t)$ and $i(t) = \sqrt{2} I \cos(\omega t - \varphi)$. Applying the standard product-to-sum identity gives:

$$p(t) = VI [\cos \varphi + \cos(2\omega t - \varphi)] \quad (1.6)$$

Expanding the oscillating term, this expression separates naturally into two contributions:

$$p(t) = \underbrace{VI \cos \varphi [1 + \cos(2\omega t)]}_{\text{active term}} + \underbrace{VI \sin \varphi \sin(2\omega t)}_{\text{reactive term}} \quad (1.7)$$

The active power P is defined as the average value of $p(t)$ over one period $T = 2\pi/\omega$:

$$P = \frac{1}{T} \int_0^T p(t) dt \quad (1.8)$$

Since $\cos(2\omega t)$ and $\sin(2\omega t)$ both average to zero over one period, applying (1.8)

to (1.7) leaves only the constant part of the active term:

$$P = VI \cos \varphi \quad (1.9)$$

while the reactive term averages to zero. The first term thus oscillates at twice the grid frequency around a strictly positive average value P : it represents power flowing unidirectionally from the source to the load, producing useful work. The second term oscillates at 2ω about a zero average, and corresponds to the reversible energy exchange established above.

The case $\varphi = \pi/2$ makes this distinction particularly transparent:

$$p(t)\Big|_{\varphi=\pi/2} = VI \sin(2\omega t) \quad (1.10)$$

The average power is strictly zero, but the instantaneous power oscillates with amplitude VI . A non-negligible current flows through the conductors and a non-zero voltage is present at the terminals, while no useful energy is transferred over a complete cycle. From the perspective of the source and the transmission infrastructure, this current has a physical reality: it produces resistive losses, loads the conductors, and stresses the insulation. A dedicated quantity is therefore needed to account for this current that circulates without producing work; this is the role of reactive power Q .

The three power quantities are formally defined as follows. The active power P [W], derived above in (1.9), is the average of $p(t)$ over one period. The reactive power Q [VAR] captures the amplitude of the oscillatory exchange:

$$Q = VI \sin \varphi \quad (1.11)$$

The powers are defined here with the receiver sign convention, as in the circuit of Fig. 1.2: the power absorbed by the load is counted as positive. With this convention, the sign of Q follows directly from the nature of the load, taking a positive value for an inductive load, whose current lags the voltage, and a negative value for a capacitive load, whose current leads it. Unlike active power, Q is not dissipated but exchanged, and it nonetheless increases the apparent power S [VA], defined as:

$$S = VI \quad (1.12)$$

accounts for the total current magnitude that the network must carry, regardless of whether that current produces useful work. These three quantities satisfy:

$$S^2 = P^2 + Q^2 \quad (1.13)$$

Chapter 1. General Introduction

This relationship is represented geometrically by the power triangle shown in Fig. 1.3, with P , Q , and S forming its three sides.

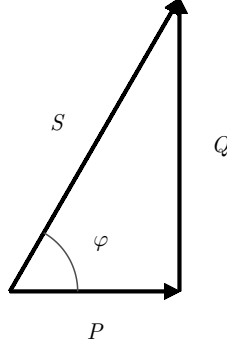


Figure 1.3.: The power triangle.

The power factor $\cos \varphi = P/S$ measures how effectively the current flowing in the network produces useful work. A power factor close to unity indicates that nearly all current drawn from the source is converted into active power; a low power factor implies that a significant fraction of the current serves only the oscillatory reactive exchange, without contributing to useful energy transfer.

This is directly visible in the circuit of Fig. 1.2, where the resistive branch dissipates active power P , the inductive branch absorbs reactive power ($Q > 0$), and the capacitive branch injects reactive power ($Q < 0$), so that the total power delivered by the source combines all three contributions. Since $S = \sqrt{P^2 + Q^2}$ from (1.13), the power factor can equivalently be written $\cos \varphi = P/\sqrt{P^2 + Q^2}$, making explicit that balancing Q locally between the inductive and capacitive branches reduces the net reactive power drawn from the source and therefore increases the power factor. A low power factor has direct and immediate practical consequences. For a given active power demand, a reduced value of $\cos \varphi$ requires a proportionally higher current to be drawn from the source, since $I = P/(V \cos \varphi)$. This excess current flows through the entire upstream network, increasing resistive losses in proportion to $I^2 R$. At the scale of a single installation, such effects may appear marginal; at the scale of a HV transmission network, they become a central operational concern.

Beyond losses, reactive power has a direct influence on bus voltage magnitudes throughout the network. This Q-V relationship, central to reactive power management in transmission systems, is examined in Section 1.2.

1.2. Reactive Power Behaviour in Transmission Networks under the Energy Transition

The physical foundations established in Section 1.1 provide the tools needed to examine how reactive power behaves in actual HV transmission networks. Three aspects are addressed in turn. Section 1.2.1 derives the fundamental relationship between reactive power and bus voltage, the cornerstone of reactive power management in HV systems. Section 1.2.2 examines the contrasting reactive behaviour of OLs and UCs, and how this behaviour is amplified by line length. Section 1.2.3 situates these elements within a broader dynamic: the energy transition is fundamentally altering the reactive power balance of HV networks, exposing the limits of approaches conceived for a more predictable system.

1.2.1. The Q-V Relationship in HV Networks

The circuit of reference for this section is the single-line diagram shown in Fig. 1.4. It consists of a voltage source connected to a load through a transmission line represented by its pi-equivalent model. This model has a series branch of resistance R and inductive reactance X , and two shunt susceptances $B/2$ placed at each end of the line to account for the distributed capacitive coupling to ground. This representation, standard in power systems analysis, will serve as the working basis throughout this section.

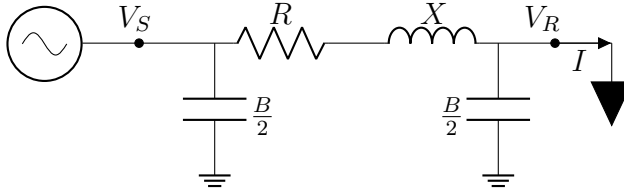


Figure 1.4.: Pi-equivalent model of a transmission line connecting a voltage source to a load. V_S , V_R and I are magnitudes.

Let $\Delta V = V_S - V_R$ denote the voltage magnitude difference between the sending and receiving ends of the line. In this section, V_S , V_R and I are magnitudes, as in Fig. 1.4. The corresponding phasors are used only in the derivation of Appendix A.1. Applying Kirchhoff's voltage law to the pi-equivalent model, and assuming the phase angle between the two ends remains small, a condition met under normal HV operating conditions, this difference reduces to:

$$\Delta V \approx \frac{PR + QX}{V_R} \quad (1.14)$$

The full derivation is provided in Appendix A.1. This result is central to the analysis of reactive power in transmission networks. It establishes that voltage drop across a line depends on both active and reactive power flows, weighted respectively by the line resistance R and inductive reactance X .

In HV transmission networks, line parameters are determined by the physical construction of the lines. The construction of OLs, with bare conductors suspended at large inter-phase spacing, favours a large series inductance and limits the shunt capacitance. From a reactive balance point of view, the reactive power absorbed by the series inductance exceeds the reactive power generated by the shunt capacitance under usual loading conditions, resulting in an inductive behaviour of the line (this contrast with UCs is developed in Section 1.2.2). For typical HV OLs operating at 150 kV, 220 kV, or 380 kV, the ratio X/R ranges from approximately 3 to 9 [2]. This ratio is not uniform across the HV range: it depends on both the voltage level and the conductor cross-section. It takes its lower values at lower HV levels equipped with reduced cross-sections, and its higher values at 220 kV and 380 kV, where large bundled conductors are used. The system remains reactance-dominated throughout, so the resistance term PR remains secondary with respect to QX , and the voltage drop approximation reduces to:

$$\Delta V \approx \frac{QX}{V_R} \quad (1.15)$$

The implication is direct: in HV transmission networks, the voltage drops between buses are governed primarily by the transit of reactive power in transmission lines. For a given source bus voltage V_S , a local reactive injection at the receiving bus reduces the reactive power transmitted through the line. The voltage drop then decreases and V_R increases. A local reactive absorption has the opposite effect. To a first approximation, controlling the reactive power flow in a line is therefore equivalent to controlling the voltage drop across it.

1.2.2. Reactive Behaviour of Overhead Lines and Underground Cables

Transmission lines are not a homogeneous class of infrastructure: OLs and UCs have fundamentally different electrical characteristics, with direct consequences on their reactive power behaviour. Table 1.1 lists representative per-unit-length parameters for each type.

In both line types, reactive power exchange results from two competing contributions: the series reactance X absorbs reactive power in proportion to XI^2 , while the shunt susceptance B injects reactive power in proportion to BV^2 . What

	R [$\Omega \text{ km}^{-1}$]	X [$\Omega \text{ km}^{-1}$]	B [$\mu\text{S km}^{-1}$]
Overhead line	0.012–0.119	0.250–0.488	2.8–5.2
Underground cable	0.013–0.059	0.131–0.339	96.5–298.8

Table 1.1.: Typical per-unit-length parameters for HV OLs and UCs [3, 2].

Table 1.1 reveals is the stark difference in the relative weight of these two terms. OLs, with bare conductors separated by air over large inter-phase distances, are inductance-dominated: their X is high and their B modest, as Table 1.1 confirms. UCs present the opposite configuration: the coaxial geometry of insulated conductors separated by a thin dielectric layer produces a shunt susceptance roughly 20 to 50 times larger than that of an OL [4, 5].

The active power level at which the net reactive exchange between the two competing contributions is exactly zero defines the Surge Impedance Loading (SIL) of the line [6]. Below the SIL, the capacitive injection from the shunt susceptance prevails; above it, the inductive absorption of the series branch dominates. For OLs, the SIL is typically reached at moderate loading levels, well within the normal operating range of a HV line. For UCs, the very large shunt susceptance results in a SIL that far exceeds the thermal current limit of the cable: no admissible operating condition places the cable above its SIL, and it therefore invariably injects reactive power. Under light loading, this sustained capacitive injection raises the voltage toward the receiving end of the line, a phenomenon known as the Ferranti effect.

In Fig. 1.5, the vertical axis represents Q_{line} , the net reactive power exchanged between the line and the network: positive values indicate an inductive absorption, negative values a capacitive injection into the network, consistent with the convention established in Section 1.1. The linearised approximation (1.15) offers direct physical insight into the Q-V coupling. However, the curves shown here are obtained without linearisation, by solving the exact power flow equations of the pi-equivalent model of Fig. 1.4 for a purely active load at the receiving end. The full derivation is provided in Appendix A.2. The line is operated at a nominal voltage of 220 kV, using representative per-unit-length parameters within the ranges of Table 1.1: $R = 0.031 \Omega \text{ km}^{-1}$, $X = 0.406 \Omega \text{ km}^{-1}$, $B = 3.8 \mu\text{S km}^{-1}$ for the OL, and $R = 0.036 \Omega \text{ km}^{-1}$, $X = 0.292 \Omega \text{ km}^{-1}$, $B = 171 \mu\text{S km}^{-1}$ for the UC. Fig. 1.5 is computed for a line length of 10 km.

Fig. 1.5 illustrates this directly. Negative values of P correspond to an active power flow reversed with respect to the nominal direction, from the receiving end toward the sending end. The zero crossing of the OL curve identifies its SIL: below this point the line is capacitive, above it inductive. Normal HV operating

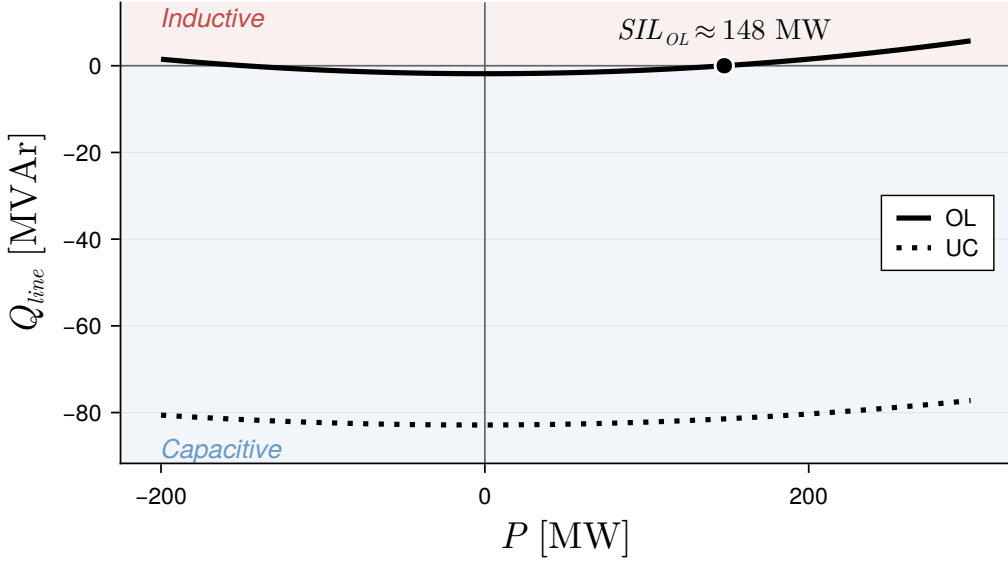


Figure 1.5.: Net reactive power balance as a function of active load: OL vs UC.

conditions place the OL comfortably above its SIL, confirming its predominantly inductive character under load. The UC curve shows no zero crossing within the plotted range: its SIL lies beyond the thermal limit of the cable, so the cable remains capacitive under all admissible operating conditions. Both curves also show that reactive behaviour becomes more capacitive as active power transit decreases, a characteristic that is increasingly relevant as periods of high PV production and low demand become more frequent under the energy transition.

1.2.3. Reactive Balance and Its Evolution under the Energy Transition

The analysis of Section 1.2.2 illustrates a key reality: an HV transmission network is a system in which sources and consumers of reactive power are distributed, heterogeneous, and spatially mismatched. Lines absorb or inject Q depending on their type and loading, transformers consume it, and loads demand it in varying proportions. Their combined imbalance translates directly into voltage deviations at the network buses, through the mechanism established in Section 1.2.1.

A direct consequence of the coupling between reactive power and voltage established above is that reactive power does not travel efficiently over long distances. Attempting to compensate a reactive deficit at one point in the network by drawing on a surplus located far away generates voltage drops along the entire con-

necting corridor, and increases resistive losses through the accompanying rise in current magnitude. Reactive power balance must therefore be ensured locally, close to where Q is consumed. This principle of local compensation structures the way HV networks are operated. Dedicated reactive compensation assets (capable of both injecting and absorbing reactive power) are deployed throughout the network to address local reactive imbalances in either direction. These imbalances are either the inductive consumption of loads, transformers, and OLs under load, or a capacitive excess under lightly loaded conditions.

The Q-V relationship established in Section 1.2.1 makes the consequence of reactive imbalance concrete: depending on its sign, bus voltages are either pushed above or depressed below admissible limits. The capacitive excess driven by UC proliferation and PV generation under light load pushes bus voltages upward: left uncontrolled, sustained overvoltages stress insulation and trigger protection systems. At the other extreme, a reactive deficit under heavy inductive loading depresses voltages below admissible limits, with cascading effects on the ability of the remaining reactive sources to support the network. These two boundaries, the upper and lower operational voltage limits, define the security envelope within which the transmission network must operate under all plausible conditions. Ensuring that this envelope is respected is therefore not merely an operational concern, but a central requirement of any reactive power management approach.

This classical picture, however, is being durably altered by the energy transition. The progressive roll-out of conventional synchronous generators [7, 8], combined with the large-scale deployment of PV generation connected predominantly at the distribution level [9], introduces a decentralized and intermittent source of active power that propagates upward through the network hierarchy. At the Transmission-Distribution (T-D) interfaces, the reactive power balance is shifting: the growing penetration of PV, combined with the progressive replacement of OLs by UCs in certain areas, is introducing a net capacitive tendency that was not present in the original design assumptions of these networks [4, 5]. Wind generation introduces a complementary mechanism, this time directly at the HV level. Wind farms, both onshore and offshore, are typically connected to the transmission network via dedicated HV UC feeders. During periods of low wind output, these cables carry little active power while remaining fully energized, keeping them well below their SIL and producing persistent capacitive injections directly at the HV substation level. Empirical evidence of this trend, measured at T-D interfaces in the Netherlands, is shown in Fig. 1.6.

Fig. 1.6 confirms this capacitive trend, while also highlighting the significant variability that surrounds it. This evolution is not uniform across the network; it depends on the local mix of OLs and UCs, the rate of PV penetration by zone,

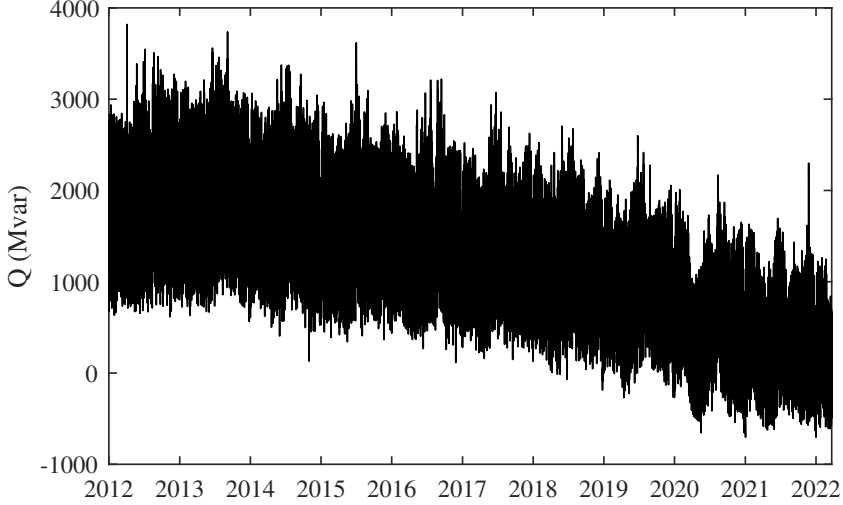


Figure 1.6.: Measured evolution of the reactive power balance at T-D interfaces in the Netherlands, illustrating the growing capacitive tendency driven by PV and UC penetration [10]. Q is positive when reactive power flows from the transmission network to the distribution network (inductive exchange) and negative in the opposite direction (capacitive exchange), consistent with the convention established in Section 1.1.

and the demand profile of the underlying distribution network. Its magnitude and spatial distribution are therefore difficult to anticipate with the tools and methods calibrated for a predictable system in which power flowed unidirectionally from centralized generators to passive loads. It is this structural shift, and the questions it raises for reactive power management, that motivates the present work.

1.3. Open Questions and Research Motivation

The structural evolution described in Section 1.2.3 has a direct consequence for reactive power management: both investment planning for compensation assets and the characterisation of reactive flows at T-D interfaces have become substantially more complex. Investment decisions must now account for a much wider range of plausible future states, while the reactive behaviour at the interfaces can no longer be reliably anticipated using the planning rules inherited from a more predictable system.

This context gives rise to two distinct but related questions, which this thesis

addresses in turn.

The first concerns planning. Determining where and how much reactive compensation capacity to install, while accounting for the range of future operating conditions and for system security, remains an open problem at the system scale. As reviewed in Chapter 2, existing formulations most often represent the operational cost of reactive management through simplified proxies, and rely on modelling tools whose underlying assumptions were established for predominantly inductive loading. The capacitive shift described above places transmission networks in a context where these assumptions no longer hold, yet where reactive investment decisions must nonetheless be made. This gap motivates the first contribution of this work: **VARPlan**.

The second concerns the T-D interface itself. The reactive power flows observed at these interfaces reflect the aggregate behaviour of the underlying distribution network: variable PV production, fluctuating loads, and a reactive signature that is difficult to anticipate from the transmission side alone. Modelling these flows accurately is a necessary condition for the TSO to make well-informed decisions. Existing approaches to this problem fall into two families: model-based methods, which require a distribution network representation the TSO does not have access to, and purely data-driven methods, which discard the underlying physical structure and offer limited reliability outside their training conditions. Bridging this gap motivates the second contribution: **VARLearn**.

These two questions are naturally connected: VARLearn addresses what happens at the interfaces, while VARPlan addresses how the TSO should invest given these evolving conditions. A full co-optimisation of both levels would be the direct continuation of this work, but lies outside the scope of this thesis.

1.4. Research Objectives and Thesis Contributions

This thesis develops two complementary approaches to improve reactive power management in HV transmission networks under the energy transition. The first operates at the system planning scale; the second addresses the modelling of power exchanges at T-D interfaces. Together, they respond to the two open questions identified in Section 1.3.

The first contribution, **VARPlan**, is a security-constrained methodology for reactive compensation planning that jointly optimises investment and operation across a range of representative operating conditions. Rather than relying on simplified cost proxies, it captures the operational cost actually borne by the

system operator; it addresses the loss of fidelity of the underlying power flow relaxation under capacitive conditions; and it restricts the investment search space through a weak-bus analysis. These specific contributions are developed in Chapter 2.

The second contribution, **VARLearn**, is a predictive model for active and reactive power exchanges at T-D interfaces that combines Inverse Load Flow (ILF)-based physical equations with Machine Learning (ML) regression. Unlike model-based approaches, it requires no knowledge of the distribution topology. Instead, it builds an equivalent model of the distribution network as seen from the transmission side, derived only from measurements available to the system operator. It includes distributed renewable generation in this equivalent and compares several learning algorithms to predict its parameters. Finally, it is assessed under realistic TSO operating scenarios and under input uncertainty. These specific contributions are developed in Chapter 3.

Several dimensions of the reactive power management problem are deliberately left outside the scope of this work. The thesis adopts an exclusively TSO perspective, leaving decisions on the DSO side unaddressed, and treats reactive power as the primary control variable while assuming active power to be given, without co-optimising the two. Neither contribution is designed for real-time operation: VARPlan is a planning tool and VARLearn a predictive model, not an online control scheme. The natural junction between them, using VARLearn predictions as inputs to the VARPlan optimisation, is recognised as the most direct perspective opened by this work but is not treated within it.

1.5. List of Publications

The following publications were produced in the context of this thesis.

Chapter 2 (VARPlan)

- [11] A. Rosseel, B. Bakhshideh Zad, and Z. De Grève, “A novel scenario-based reactive power planning method including investment and operational costs,” in *Proc. 2025 IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT Europe)*, 2025, pp. 1–5.
- [12] A. Rosseel, B. Bakhshideh Zad, and Z. De Grève, “A Cost-Reflective MISOCP Formulation for Reactive Power Planning with Weak-Bus-Guided Decision-Space Reduction,” *Sustainable Energy, Grids and Networks*, submitted August 2026.

Chapter 3 (VARLearn)

- [13] A. Rosseel, B. Bakhshideh Zad, Z. De Grève, and F. Vallée, “Reactive power forecasting at the transmission-distribution interfaces using physics-based machine learning,” in *Proc. 27th International Conference on Electricity Distribution (CIRED 2023)*, 2023, pp. 301–305.
- [14] A. Rosseel, B. Bakhshideh Zad, F. Vallée, and Z. De Grève, “Physics-informed machine learning for forecasting power exchanges at the interface between transmission and distribution systems,” *Electric Power Systems Research*, vol. 238, p. 111097, 2025.

Other publications

- [15] A. Rosseel, M. Hupez, J.-F. Toubéau, Z. De Grève, and F. Vallée, “On the optimisation of investments in distributed energy resources in a low-voltage energy community,” in *Proc. 26th International Conference on Electricity Distribution (CIRED 2021)*, 2021, pp. 1792–1796.
- [16] J. Allard, A. Rosseel, L. Sadoine, J. Faraji, T. Brihayé, F. Capizzi, B. Nteziyaremye, F. Bordes, F. Vallée, and Z. De Grève, “Technical impacts of the deployment of renewable energy communities on electricity distribution grids,” in *Proc. 27th International Conference on Electricity Distribution (CIRED 2023)*, 2023, pp. 2923–2927.

CHAPTER 2.

VARPlan: Reactive Power Planning in HV Transmission Networks

The methodology and results presented in this chapter build on work first reported in [11], and are further developed in a journal article currently under review at *Sustainable Energy, Grids and Networks* [12].

2.1. Problem Setting and State of the Art

Among its responsibilities, the TSO must ensure voltage regulation across the network in near-real time and size compensation assets within its investment cycle. These two responsibilities operate on distinct time horizons but are tightly coupled, since investment decisions determine the operational margin available to the operator at any future point in time. The ongoing transformation of transmission networks, marked by the evolving reactive power balance at T-D interfaces, driven by increasing underground cabling and a growing share of renewable generation, is intensifying reactive power needs, as established in Chapter 1 [17]. This evolution reinforces the importance of sizing compensation assets in a way that faithfully anticipates future operating conditions and system evolutions.

The quality of a reactive power planning decision is bounded by the fidelity of the operational model it incorporates. This fidelity unfolds along three dimensions: the cost structure represented in the objective, the physical representation of the network, and the extent of the decision space considered. The modelling choices underlying each of these dimensions were established against a historically inductive operating regime, one that the energy transition is progressively displacing toward periodically capacitive conditions [18]. The review that follows

examines, for each dimension in turn, the extent to which existing approaches remain representative as operating conditions evolve.

The review opens with the European regulatory and economic framework that defines the system operator's obligations and the cost structure of reactive power, followed by a characterisation of the available compensation resources. It then examines the formulation of optimal reactive power dispatch and the convex relaxations of the AC power flow on which tractable approaches rely. It subsequently addresses reactive power planning, the incorporation of security constraints, and the reduction of the decision space. A synthesis gathers the identified gaps and situates the contributions of the chapter.

2.1.1. European Regulatory and Economic Framework

The operational framework of European TSOs is established by the Commission Regulation on Electricity Transmission System Operation, which defines the mandatory security requirements governing network management [19]. This framework imposes compliance with the N-1 criterion, requiring that an admissible operating state be maintained following the outage of any single network element. It also specifies admissible voltage ranges, differentiated by voltage level and synchronous area. This regulation applies to real-time operation and operational planning, not to grid design. However, the planning problem considered here includes the operation of the network: the investment decisions must allow the TSO to operate the network in compliance with these requirements. N-1 security is therefore imposed on the operating points simulated in the planning model, as a mandatory constraint and not as an optional extension.

The regulation establishing a network code on requirements for generators specifies the reactive power capabilities that generating units must be able to provide or absorb [20]. These requirements are expressed through capability charts relating the admissible reactive power output to the terminal voltage and the apparent power rating of the unit. The minimum reactive capability is defined at maximum active power, and a P-Q profile then specifies the required capability below maximum active power. The required reactive capability therefore depends on the active operating point of the unit. In this work, the reactive capability of a generating unit is represented by constant limits, independent of its active operating point. This is a simplification of the regulatory requirement.

The transformation of the generation mix and the development of UC networks are modifying the direction and magnitude of reactive power exchanges between distribution and transmission networks, as detailed in Chapter 1 [5]. The TSO does not directly control the reactive injections of the distribution network. The

Network Code on Demand Connection gives it means to limit the reactive power exchange at the interface [21]. It defines a reactive power range for transmission-connected distribution systems, and it allows the TSO to require that no reactive power is exported to the transmission network at nominal voltage when the active power flow is below 25% of the maximum import capability. Within these limits, the reactive exchange depends on the loads, generation and cables of the distribution network, and the TSO remains responsible for voltage regulation at the interface. Recent literature characterises the reactive exchange capability at this interface in the form of active and reactive power flexibility domains [22]. The adequacy of the reactive support provided on both sides of the interface has received specific attention [23]. For a planning problem, the reactive exchange at the interface therefore constitutes an uncertain input rather than a controlled parameter, a feature that formulations treating nodal reactive load as fixed do not capture; this limitation is revisited in Section 4.2.

A comparable form of exogeneity affects the active power operating point. The European wholesale electricity market is organised into price zones within which the physical network is not represented [24]. The generation schedule resulting from this market clearing may consequently be physically infeasible on the actual network [25]. The TSO restores feasibility after the fact by activating redispatch actions that adjust the generation schedules. The cost of these actions is recovered through the network tariffs. The volumes and costs associated with these actions have grown with the penetration of renewable generation and the emergence of internal congestion within price zones. The operating point that the system operator must secure is therefore not a variable it co-optimises, but the outcome of a market process that is exogenous to it. It follows that a faithful representation of the European context requires treating this operating point as an input and associating a cost with active redispatch actions.

Several actors take part in this problem, each with its own objective. The TSO keeps the transmission network secure at least cost. The DSO operates the distribution network and ensures voltage regulation on its own side of the interface. Producers bid their units into the energy market to maximise their profit. The market operator clears the energy market without representing the network inside each bidding zone. The regulator oversees the network operators and approves their tariffs. End users pay for energy and for network use.

The formulation developed in this chapter minimises the cost incurred by the TSO. This cost is not the whole cost of electricity supply for society. The societal cost also includes the production cost of the energy, the value of that energy for consumers, and the costs of the distribution network. The cost of the TSO is only one part of it. The assumption made here is that the other parts do not change with the decisions taken in this chapter. The active power schedule comes from

the market and is treated as an input, as explained above. The production cost and the value delivered to consumers are therefore fixed for the investment and dispatch decisions considered. What these decisions change is the cost of keeping the network secure.

This cost is not the profit of a private company. A European TSO is a regulated monopoly. Its obligations are set by the regulator, and its costs are recovered from network users through regulated tariffs. A cost avoided by the TSO is therefore a cost not paid by the users. Some terms of the objective are not internal costs of the operator: redispatch costs are payments made to producers, and the penalty on unserved reactive demand represents a loss of supply quality for the users of the network. Minimising the expenditure of the TSO therefore minimises the part of the societal cost that the planning decision affects.

This holds as long as the price paid by the TSO for an action is close to the real cost of that action. If the price is higher, the excess is not a cost for society but a payment from network users to the provider of the action.

The remuneration of reactive power itself, by contrast, is far less consistently defined. It is handled across Europe under heterogeneous regulatory regimes: an obligation attached to grid connection, regulated remuneration, competitive tendering, or a combination of these [26]. The principle of market-based procurement for this ancillary service admits explicit national derogations [27]. No single reference value for reactive power therefore applies uniformly across European system operators. The cost structure adopted in a planning model consequently depends on the regulatory framework of the operator under consideration, rather than on a universal constant.

Independently of this regulatory heterogeneity, the characterisation of reactive power cost itself falls into three broad approaches in the operational literature.

- The first relates this cost to the active losses induced by reactive power production [28]. This approach needs no market data and is easy to compute. However, it can underestimate the value of reactive support when the generator's reactive capability limit, rather than the increase in losses, constrains its output.
- The second evaluates it as an opportunity cost. Beyond a certain output level, reactive power provision reduces the active generation capacity available for sale on the energy market, as analysed in the same work. This approach links the cost to the market price at the time of provision, so the resulting value changes over time and depends on how close the unit already operates to its active power limit.

- The third remunerates reserved reactive capacity independently of the reactive energy actually delivered [29]. This approach rewards the availability of the resource and gives a stable incentive for investment, but it does not reflect how scarce reactive power actually is at a given operating point, so a unit can be paid for capacity that is rarely used.

These three approaches produce values that differ by orders of magnitude for the same physical resource. The choice of an approach and its associated value is therefore an explicit modelling assumption, not a quantity that can be measured directly. The condition stated above, that the prices paid by the TSO reflect real costs, is therefore not fully met for reactive power in the European context.

This richness in the operational and pricing literature has not carried over into reactive power planning. Planning formulations most often rely instead on simplified proxy quantities, disconnected from any of the approaches reviewed above. The cost structure actually borne by a TSO is thus absent from the objective of reactive power planning problems, a gap examined in detail in the review of optimal reactive power dispatch objectives in Section 2.1.3.

2.1.2. From Compensation Resources to Behavioural Classes

Reactive power compensation resources fall into several broad technological categories. Synchronous devices include generating units, whose reactive capability is bounded by their capability chart, and synchronous condensers. Passive devices comprise capacitor banks and inductors, switched in or out in discrete steps. Power electronics-based compensators, such as Static VAR Compensators (SVCs) and Static Synchronous Compensators (STATCOMs), provide continuous control of the reactive power exchanged with the network [30]. Converter-interfaced generating units, notably wind turbines and PV installations, possess an inherent reactive capability that remains available independently of their active power output [20]. Certain power electronics devices act on branch impedance rather than on nodal injection and fall outside the scope considered here [31].

From the perspective of a planning problem, these resources are distinguished by a small number of functional attributes, primarily the continuity of control and the directionality of the reactive exchange. Control is either continuous or discrete in steps; the exchange is either bidirectional or restricted to supply only or to absorption only. Response speed constitutes a further attribute, governing the dynamic behaviour of the system without entering directly into the structure of a steady-state investment problem. This attribute nonetheless overlaps substantially with continuity: power electronics-based continuous devices offer fast,

automatic response, while discrete devices are mechanically switched and operated more slowly. The characterisation useful for static planning thus reduces to continuity and directionality, the former indirectly capturing part of the dynamic difference between device families. On this basis, compensation resources can be grouped into behavioural classes, one for continuous and bidirectional devices, and others for discrete devices differentiated by the direction of their reactive exchange. This abstraction by behaviour rather than by technology allows the formulation to remain valid as compensation technologies evolve, which is of particular relevance given the rapid development of power electronics-based devices. This distinction also has a clear economic counterpart. Continuously controllable devices are substantially more expensive per unit of installed capacity than discrete devices, an order-of-magnitude difference reflected in the cost coefficients calibrated for the test case (Table 2.3).

This behavioural representation leaves aside several characteristics of real compensation devices. The reactive power delivered by a capacitor bank varies with the square of the terminal voltage, whereas a constant-output model assumes it fixed. This model overestimates the contribution of the capacitor precisely when voltage falls and the reactive need is greatest. However, the bias remains bounded by the narrow voltage ranges within which the network is operated [32]. Power electronics-based devices are further distinguished by their dynamic response: a STATCOM maintains its reactive output more effectively than a SVC under depressed voltage conditions, a difference that is consequential in dynamic analysis but irrelevant in steady-state [33]. Switching limits for discrete devices, relating to the admissible number of commutations over a given period, are not represented. The actual granularity of capacitor bank steps, which in practice often involves multiple step sizes, is reduced to a single insertion, which simplifies the sizing representation.

2.1.3. From Optimal Power Flow to Optimal Reactive Power Dispatch

The formulation of optimal network operation as a constrained optimisation problem traces back to the foundational work on economic dispatching [34], whose numerical solution through optimal power flow methods followed [35]. The optimal reactive power dispatch constitutes a specialisation of this problem, in which active generation schedules are fixed and the voltage control resources are optimised [36]. The decision variables comprise generator voltage setpoints, transformer tap positions, and switchable compensation devices. This specialisation addresses a short-term operational need, distinct from the longer-term investment decision that reactive power planning addresses.

Optimal reactive power dispatch formulations differ along several modelling axes.

- A first axis is temporal, opposing single operating point formulations to multi-period or multi-condition ones [37]. A single operating point is cheaper to solve, but nothing then guarantees that the resulting settings, in particular the status of discrete compensation devices that cannot be adjusted at will, remain admissible or near-optimal once conditions change. Multi-period formulations instead couple several operating conditions within a single program, at the cost of a problem size that grows with the number of conditions retained, in exchange for settings that are representative across the conditions actually considered.
- A second axis concerns the treatment of uncertainty, ranging from deterministic approaches to stochastic formulations based on scenario sets [38, 39]. A deterministic approach treats a given set of conditions as certain, which keeps the problem tractable but offers no guarantee once the actual conditions depart from those assumed. A stochastic approach instead represents a distribution of conditions and optimises an expected cost. This multiplies the problem size by the number of scenarios retained, and makes the quality of the resulting solution depend on how representative that scenario set is. This trade-off is examined further for the investment problem in Section 2.1.6.
- A third axis relates to the objective function, which may be single-criterion or combine several competing objectives [30]. A single-criterion objective admits one well-defined optimum and is straightforward to interpret. In contrast, combining several objectives requires either an explicit weighting between them or the construction of a full Pareto front. The weighting is itself a modelling choice with no neutral default, and the Pareto front is considerably more demanding to compute.
- A final axis concerns the network model retained, from the full AC power flow to its linearised or relaxed approximations, a choice whose implications are examined in Section 2.1.4.

Objective functions employed in optimal reactive power dispatch fall into three main families. The first minimises active network losses, a measurable quantity independent of any market data [40]. The second minimises voltage deviation, which reflects a service quality criterion without direct economic interpretation. The third minimises a reactive dispatch cost, the only criterion directly interpretable in the operating accounts of a system operator [41]. These families are therefore not economically equivalent: the first two represent proxy quantities for the cost actually borne, rather than that cost itself.

The use of proxy quantities does not reflect a lack of rigour, but rather the absence of a directly transposable pricing framework and the practical need to compare results across systems.

- Minimising active losses constitutes a reasonable approximation of cost when the binding constraint of the network is power transit. Under capacitive conditions, however, the binding constraint becomes overvoltage and the need for reactive power absorption. Active losses scale with the square of the current magnitude on each branch, so an objective that minimises them carries little signal when lines are lightly loaded and losses are already small everywhere, precisely the conditions under which capacitive behaviour dominates. Moreover, since a higher voltage reduces the current needed to transfer a given active power, loss minimisation can push the voltage upward to reduce current, moving the operating point further from the reactive absorption that relieving overvoltage actually requires.
- The voltage deviation objective, for its part, penalises the magnitude of the difference from a reference voltage without regard to its direction, so an undervoltage and an overvoltage of equal magnitude are treated as equally costly. This symmetry misrepresents the actual economics of correction. Relieving undervoltage draws on reactive injection capability that is widely available from generators and capacitor banks, whereas relieving overvoltage requires reactive absorption, a capability that networks historically dimensioned for inductive conditions offer in far smaller quantity [18].

These two proxy quantities were thus calibrated for an operating condition that is no longer the only condition that dimensions networks undergoing the energy transition. The reactive dispatch cost family avoids both limitations, since it evaluates the objective directly from the cost structure rather than from a proxy for it.

2.1.4. Convex Relaxations of the AC Optimal Power Flow

The AC power flow equations are quadratic in the complex bus voltages. Each branch flow is expressed through products of two voltage magnitudes and through the square of the sending-end magnitude. A set defined by such quadratic equalities is not convex: if two distinct operating points both satisfy the power flow equations, an intermediate point obtained by averaging them does not satisfy them in general. The feasible domain is therefore a curved surface rather than a solid region, and it may consist of several separate parts. The associated feasibility problem is NP-hard [42]. Standard nonlinear solution methods, in particular

interior-point algorithms, produce locally optimal solutions without any certificate of global optimality [36]. The gap between the solution obtained and the true global optimum generally remains unknown, so that an AC solution produced by a local solver does not constitute an exact reference. This limitation carries more weight in planning than in operations: an imperfect operating point can be corrected in real time, whereas an investment decision, once committed, is irreversible.

A convex relaxation replaces this non-convex feasible domain by a larger convex domain that contains it. The objective function is unchanged, and the problem is solved over the enlarged domain. In the formulations considered here, the products of voltage variables become variables in their own right, and the equalities that link them to the bus voltages are replaced by inequalities. The relaxed problem is convex, so it has a single optimal value, which is reached whatever the starting point.

Two properties of this construction matter for the reading of the results. First, the enlarged domain contains the original one, so the optimal value of the relaxed problem is lower than or equal to the optimal value of the original minimisation problem. The relaxed optimum is therefore a lower bound on the true optimal cost, and the distance between the two values measures the loss due to the relaxation. Second, the point that reaches this bound does not necessarily belong to the original feasible domain. When the optimum lies strictly inside the enlarged domain, the voltage products returned by the solver correspond to no physical set of bus voltages. The operating point obtained satisfies the relaxed constraints but violates the AC power flow equations, and it cannot be applied to the network as it is.

Several families of convex relaxations of the AC Optimal Power Flow (OPF) have been developed, each reformulating the original non-convex problem in a tractable convex domain. They differ both in the tightness of the bound they provide and in their computational cost.

- A first family reformulates the network equations within the Second-Order Cone Programming (SOCP), initially for radial networks [43] and subsequently extended to meshed topologies [44]. This family is the least expensive of the families considered here to solve, and it scales favourably with network size.
- A second family is grounded in semidefinite programming, obtained by relaxing the rank-one constraint on the outer product of the voltage vector. Foundational contributions establish its application to optimal power flow [45] and characterise the conditions under which strong duality holds [46].

This relaxation is at least as tight as the SOCP relaxation, but the resulting semidefinite program scales less favourably with network size.

- The branch flow model provides an equivalent formulation, expressed directly in terms of branch power and current rather than bus voltage products. It admits a direct physical interpretation and, for radial networks, produces a relaxation equivalent to the SOCP family [47].
- The quadratic convex relaxation adds convex envelopes of the trigonometric and bilinear terms appearing in the original problem on top of the SOCP constraints. This tightens the bound at a comparable computational cost [48].
- Moment relaxations form a hierarchy in which successive orders tighten the bound and asymptotically recover the true optimum. This comes at the cost of a rapid growth in problem size that limits practical use to low orders [49, 50].

These families are not totally ordered in tightness. While the semidefinite relaxation dominates the SOCP relaxation, as noted above, the quadratic convex relaxation and the semidefinite relaxation do not dominate each other in the general case [48]. This comparison is detailed in surveys covering both the tightness and the computational cost of each family [51, 52]. A formulation combining SOCP constraints with quadratic convex envelopes occupies an intermediate position between the conic relaxation alone and higher-order relaxations. It seeks a balance between bound tightness and program size.

A convex relaxation is said to be exact when its optimal solution allows a feasible solution of the original problem to be reconstructed with the same objective value. Sufficient conditions for exactness have been established for radial networks, notably the absence of active upper voltage bounds at the optimum [53]. Alternative formulations of these conditions invoke the monotonicity of the objective function in the power injections, the absence of flow reversal [54]. These conditions are restrictive, and several of them can only be verified a posteriori on the solution obtained, rather than guaranteed in advance. No analogous guarantee is available a priori for meshed networks, where inexactness has been documented and attributed to the insufficiency of constraints bearing on the network cycles [55]. The semidefinite relaxation is at least as tight as the SOCP relaxation. Therefore, any instance for which the former is inexact is also inexact for the latter. This reasoning does not extend directly to the quadratic convex envelopes, which are not comparable to the semidefinite relaxation.

Under low transit and dominant capacitive loading, the Ferranti effect (lightly loaded branches injecting reactive power that raises receiving-end voltage, see

Section 1.2) pushes bus voltages toward their upper bounds. In these conditions, upper voltage limits become active at the optimum, and reactive power flow reversals appear on a subset of branches. These two configurations correspond precisely to the conditions whose violation, as established above, undermines the exactness of the relaxation. They appear together, but they do not have the same role. The construction developed in Section 2.1.5 shows that the reversal of the reactive flow is what triggers the loss of exactness, while the upper voltage bounds only limit how far the relaxed solution can move. It follows that the capacitive conditions simultaneously concentrate the mechanisms identified as responsible for the loss of exactness, in a context where no theoretical safeguard applies.

Several techniques have been developed to tighten the bound provided by a convex relaxation. A first approach adds valid inequalities derived from the network cycles, strengthening the conic relaxation on meshed topologies [56]. A second approach combines convex envelopes, bound tightening, and valid inequalities to reinforce the semidefinite relaxation [57]. Domain reduction through bound propagation provides an independent source of tightening that is applicable regardless of the relaxation family retained [58]. These techniques target the quality of the lower bound; recovering a feasible solution to the original problem from a relaxed solution is a separate matter that requires dedicated treatment.

A complementary approach perturbs the objective function with penalty terms designed to steer the relaxed solution toward one from which a feasible point of the original problem can be reconstructed [59]. The underlying mechanism is that the relaxation artificially raises voltage amplitudes in order to reduce active losses, producing injections that are admissible for the relaxed program but not for the original problem. Penalising reactive power production tends to reduce these voltage amplitudes and restores admissibility across a broad class of instances. In these approaches, the penalty applies to an aggregate reactive quantity rather than to a subset of branches selected by a structural criterion available prior to optimisation. This principle has been extended to the security-constrained setting [60] and combined with moment relaxations [61]. These contributions are framed within an operational context, where the objective is to reconstruct a feasible operating point for a given instance. Whether such objective perturbations preserve investment decisions in a planning setting has not yet been examined; an undersized or mislocated asset cannot be corrected in real time, which makes this gap consequential.

2.1.5. Conic Relaxation under Capacitive Loads

The mechanism behind the loss of exactness under capacitive loading is shown here on a two-bus system. The flow equations are first rewritten in the variables of the relaxation. The relaxed set is then drawn in a plane, where the reactive power absorbed by a branch can be read directly. The reactive balance of the two-bus system finally shows why the relaxation is exploited under capacitive loading, and not under inductive loading.

Lifting into the W space: The AC flow equations are quadratic in the voltage magnitudes and trigonometric in the angle differences. The change of variables

$$W_n = V_n^2, \quad W_{ij}^R = V_i V_j \cos \theta_{ij}, \quad W_{ij}^I = V_i V_j \sin \theta_{ij} \quad (2.1)$$

makes every branch flow equation affine in $(W_i, W_j, W_{ij}^R, W_{ij}^I)$. All the non-convexity is then carried by a single identity,

$$(W_{ij}^R)^2 + (W_{ij}^I)^2 = W_i W_j \quad (2.2)$$

The relaxation replaces this equality by the inequality

$$(W_{ij}^R)^2 + (W_{ij}^I)^2 \leq W_i W_j \quad (2.3)$$

The equality forces the point to lie exactly on the boundary of a second-order cone, and this boundary is not a convex set. The inequality only requires the point to lie inside the same cone, which is convex. This single change convexifies the whole formulation [44, 48], and it is the constraint used in (2.31). In a meshed network, however, the W variables do not force the angle differences to be consistent around each cycle, which is a second source of inexactness [55]. A two-bus system has no cycle, so the example below isolates the effect of the cone.

The circle and the disc: Let W_i and W_j be fixed, and draw the pair (W_{ij}^R, W_{ij}^I) in a plane, as in Figure 2.1. The equality (2.2) is a circle of radius $\sqrt{W_i W_j}$. The inequality (2.3) is the disc bounded by this circle. A point of this plane has a radius equal to $V_i V_j$ and an angle equal to θ_{ij} . On the circle, these two values agree with the voltages of the two buses. Inside the disc, they do not: the branch sees a product of voltages smaller than the product declared by its two buses. No triplet (V_i, V_j, θ_{ij}) gives such a point.

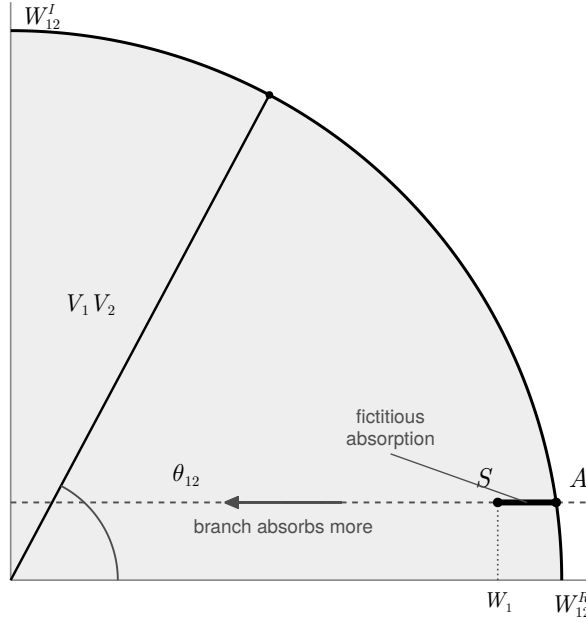


Figure 2.1.: Exact set and relaxed set of a branch in the (W_{12}^R, W_{12}^I) plane, drawn for the bus voltages of the capacitive optimum, $W_1 = 1.00$ and $W_2 = 1.28$. Point A is the only physical state for this active transfer, and the optimum settles there under inductive loading. Point S is the point where the generator is idle, and the optimum settles there under capacitive loading.

For a lossless branch, equation (2.42) reduces to $P_{ji} = B_{ij}W_{ij}^I$, so the vertical coordinate of the point is set by the active power flow and is known once the active transfer is known. The operating point then lies on a horizontal line, and the part of the disc on this line is a segment. The point reaches the circle only where W_{ij}^R takes its largest admissible value $\sqrt{W_i W_j - (W_{ij}^I)^2}$. The cone sets this maximum and no minimum, so every smaller value of W_{ij}^R places the point strictly inside the disc. The horizontal coordinate, in turn, carries the reactive power. Summing (2.41) and (2.43) cancels the terms in G_{ij} and gives the reactive power absorbed by the branch,

$$Q_{ij}^{line} = Q_{ij} + Q_{ji} = -B_{ij} (W_i + W_j - 2W_{ij}^R) \quad (2.4)$$

This identity holds for any value of the resistance.

A two-bus case: The configuration is the one already used in Section 1.2.1 (Figure 1.4), a source bus feeding a load bus through a single line. The line is reduced here to its series branch, and the capacitive behaviour comes from the load. Bus 1 is the slack bus, its voltage is imposed and W_1 is therefore a parameter, as in (2.45). A generator at bus 1 supplies the reactive power $Q_{g,1}$. Bus 2 carries the load $(P_2^{\text{load}}, Q_2^{\text{load}})$ and nothing else. The branch is lossless, so $G_{12} = 0$ and $B_{12} = -1/x_{12}$, where x_{12} is its series reactance. The objective prices the reactive output of the generator and reduces to $\gamma_1^q |Q_{g,1}|$.

Bus 2 has no generation, so the branch alone supplies the load. Adding the reactive balances of the two buses gives the reactive balance of the system,

$$Q_{g,1} = Q_2^{\text{load}} + Q_{12}^{\text{line}} \quad (2.5)$$

The generator supplies the load and the branch. Minimising $|Q_{g,1}|$ therefore amounts to bringing Q_{12}^{line} as close as possible to $-Q_2^{\text{load}}$. The reactive balance at bus 1 also gives $Q_{g,1} = B_{12}(W_{12}^R - W_1)$, so the generator is idle exactly when $W_{12}^R = W_1$.

Consider first an inductive load, $Q_2^{\text{load}} > 0$. The objective calls for a negative Q_{12}^{line} . The series branch of this example carries no shunt capacitance and cannot supply reactive power, so this target is out of reach. The best the optimisation can do is to make the absorption of the branch as small as possible, which pushes W_{12}^R to its largest admissible value. The cone is then tight, the solution lies on the circle, at point *A* of Figure 2.1, and the relaxation is exact. The generator injects the load and the branch absorption, which is the physical answer.

Consider now a capacitive load, $Q_2^{\text{load}} < 0$. The objective calls for a positive Q_{12}^{line} , an absorption that the branch can be given by lowering W_{12}^R , and nothing in the cone opposes this move. As soon as the capacitive load exceeds what the branch absorbs on its own, the optimisation reaches $Q_{g,1} = 0$ at point *S* of Figure 2.1, which lies strictly inside the disc. The cone is loose, the conic residual $\rho_{12} := W_1 W_2 - (W_{12}^R)^2 - (W_{12}^I)^2$ is strictly positive, and the branch acts as an artificial reactive power sink. It absorbs a quantity that no physical state can produce, and this absorption replaces the generator.

The two cases are therefore separated by the sign of the reactive output of the generator. The cone stays tight as long as the generator injects, and becomes loose as soon as the load asks it to absorb. The trigger is the reversal of the reactive flow, and nothing else.

This construction uses only the power balances and the conic constraint. The voltage, angle and reactive capability limits of a real network restrict the travel of the relaxed point along the segment, and can bring *S* closer to *A*. However,

they do not place it back on the circle as long as the load asks the generator to absorb reactive power. The loss of exactness therefore does not come from a saturated bound. It also shows that a small conic residual is not a negligible error, since identity (2.4) converts the horizontal gap it leaves into reactive power. The degradation of a conic relaxation under such conditions is therefore an expected consequence of the problem structure, rather than a numerical artefact specific to a particular instance. A regularisation term acting on this mechanism is introduced in the formulation of Section 2.2.2.

2.1.6. Reactive Power Planning

Reactive power planning determines the location and sizing of new compensation assets in order to ensure voltage regulation across the network over a given planning horizon [62]. The associated problem combines discrete installation variables, continuous sizing variables, and the nonlinear equations of the AC power flow. Its size grows with the product of the number of candidate sites, the number of operating conditions, and the number of contingencies considered. It therefore constitutes a mixed-integer nonlinear program whose exact solution is computationally demanding [42]. The growing scarcity of reactive power compensation resources induced by the energy transition renews the practical relevance of this problem [18].

The first integrated methodologies for reactive power source planning date back to the late 1980s [63]. From the early 1990s onward, formulations addressed reactive power planning under preventive and corrective security constraints through decomposition-based approaches [64]. The incorporation of contingency-based security into reactive power planning is therefore a long-standing objective, not a recently introduced feature.

A large share of the recent literature solves the reactive power planning problem through metaheuristic methods, such as genetic algorithms, particle swarm optimisation, or differential evolution [41, 31]. These methods handle the non-convex and mixed-integer character of the problem directly, without linearising the constraints. They provide, however, no optimality guarantee and no bound on the gap between the solution obtained and the true optimum [65]. Their reproducibility is limited, as parameter tuning and initialisation are rarely documented in sufficient detail to allow replication. Comparison across studies is further hindered by the fact that each work simultaneously modifies the objective function, the test network, and the constraint set.

By contrast, mathematical programming approaches offer optimality guarantees within the formulation they solve, at the cost of an explicit treatment of non-

convexity. A first lever for tractability is decomposition: Benders decomposition separates investment decisions from operational subproblems, containing the size of the problem [66]. Applying a convex relaxation to the operational subproblems then preserves the validity of this decomposition despite the non-convexity of the AC power flow [67]. These decomposition techniques have a long track record in transmission expansion planning such as [68, 69], whose multi-scenario security literature predates that of reactive power planning [70]. A second lever, distinct from decomposition, reduces the size of the instance to be solved by retaining only the contingencies and operating conditions that are effectively binding. This instance reduction has enabled security-constrained reactive power planning to be extended to large-scale networks [71].

A more recent direction substitutes data-driven models for the repeated solution of the optimisation problem, training them to approximate its output. Machine learning surrogates have been proposed for the security-constrained optimal power flow, replacing the solver in part [72]. Other work exploits deep learning to anticipate active constraints under contingency scenarios [73]. These approaches raise the question of the reliability of approximate solutions, which conditions their use in a decision-making context.

Reactive power planning must account for uncertainty in the future operating conditions of the network, whose modelling can draw on several paradigms [74]. The stochastic approach represents this uncertainty by a set of realisations weighted by their probability, while robust optimisation guarantees feasibility over an entire uncertainty domain, and chance-constrained formulations bound the probability of constraint violation. These paradigms differ in the tradeoff they strike between solution conservatism and the probabilistic information required on the uncertainty. The reactive power planning literature relies predominantly on the stochastic representation through a reduced set of weighted operating conditions [38]. Reducing the number of such conditions then involves a tradeoff between representational fidelity and problem size, and the degree of reduction that is sufficient has rarely been studied in a systematic manner. A second level of uncertainty, qualitatively different in nature, concerns futures that cannot be assigned probabilities, where contrasting evolutions of demand and generation mix are addressed through non-probabilistic frameworks [75]. The construction of such contrasting futures, to be explored separately rather than aggregated, corresponds to the approach adopted in European network development plans [76].

2.1.7. Security in Planning

The security-constrained optimal power flow extends the base problem by incorporating a set of contingency scenarios [77]. In a preventive approach, the base

operating point must remain feasible following each contingency without any corrective action being taken. In a corrective approach, bounded post-contingency adjustment actions are permitted, which enlarges the admissible domain [78]. A preventive-corrective approach combines both principles. A subset of controls is fixed preventively across all contingencies, while the remaining controls are allowed to adjust correctively within each contingency. Any formulation that allows corrective action assumes that these actions can be deployed within the admissible time window following the contingency. The choice between these three approaches determines both the size of the resulting problem and the degree of conservatism of the solution.

Solving the security-constrained optimal power flow on realistic networks under corrective or preventive-corrective formulations raises three difficulties specific to the incorporation of contingencies [79]. The first concerns the number of admissible corrective actions following a contingency, whose limitation reflects actual operational constraints. The second relates to the representation of voltage and stability constraints in the post-contingency state, whose inclusion increases the size and complexity of the problem. The third addresses the reduction of problem size through identification of the contingencies that are effectively binding.

Most contingencies in a given set are non-binding at the optimum, which justifies reducing their number prior to or during the solution process. Ranking contingencies by severity allows those whose effects remain admissible to be discarded [80]. Progressive identification of the binding contingencies alone during the solution process avoids the need to represent the full set explicitly [67]. These techniques reduce problem size without modifying the admissible domain they approximate. This distinguishes them from reductions applied to the investment decision space, where excluding candidate locations may remove the optimal solution from consideration.

The incorporation of security into reactive power planning problems has followed three successive approaches. Early formulations verified contingency compliance a posteriori, once investment decisions had already been made. Others deferred security treatment entirely to the operational level, without integrating it into the investment decision. More recent formulations embed contingencies directly within the investment decision, extending this integrated treatment to large-scale networks under detailed network models [71].

Beyond static voltage constraints, part of the reactive power planning literature incorporates voltage stability margins directly into investment decisions [81]. These approaches evaluate the distance of an operating point from voltage collapse, for instance through the load margin [82]. A static formulation is justified

when both the planning horizon and the regulatory criterion are themselves expressed in steady-state terms. Such a formulation does not, however, capture the dynamic phenomena of voltage instability, which require a distinct analytical framework [83, 3].

2.1.8. Decision-Space Reduction and Weak Bus Identification

A weak bus, in the sense used throughout this chapter, denotes a network location where the need for additional reactive compensation is greatest. Allowing reactive compensation to be installed at every network bus associates a discrete decision variable with each candidate site. The size of the resulting investment problem then grows with the product of the number of candidate sites, the number of operating conditions, and the number of contingencies considered. On a network of realistic scale, enumerating all candidate sites leads to a problem whose exact solution becomes computationally out of reach. A first strategy selects sites during the solution process, through iterative identification of the most constrained buses [84]. A second strategy fixes the candidate set prior to optimisation, often on the basis of expert judgement whose justification is rarely made explicit.

Weak bus identification has traditionally relied on indicators derived from voltage stability analysis. The L-index evaluates the proximity of a bus to voltage collapse from the network load relationships [85]. Modal analysis of the reduced Jacobian matrix associates the weakest modes with the buses and branches that contribute most to them [86]. A further family is based on sensitivity indices relating voltage variations to power injections at each bus [87]. Line-level indices estimate the margin before loss of a power flow solution along individual branches. Despite their differences in formulation, these indicators share a common original purpose: estimating the proximity of an operating point to voltage instability [83].

Comparative studies have shown that different indicators of proximity to voltage collapse rank the buses of a given network differently, even when applied to the same operating point [88]. Their behaviour depends on the loading level, the power factor, and the resistance-to-reactance ratio of the elements under consideration [89]. No single indicator has established itself as a universal reference, and their relative relevance varies with the dominant instability mechanism [90]. Planning studies have accordingly combined several indicators simultaneously in order to mitigate this dependence [81]. The choice of an indicator thus constitutes a modelling assumption in its own right, rather than a neutral input to the problem.

These indicators were designed and validated for estimating proximity to voltage instability, a regime characterised by heavy loading, a deficit of reactive power supply and sustained undervoltages. Nothing in their construction guarantees that they correctly localise a reactive insufficiency of capacitive nature, associated with overvoltages rather than undervoltages. Furthermore, selecting a site for the installation of compensation equipment and estimating the proximity of a bus to voltage collapse are two distinct tasks. A bus close to instability is not necessarily the bus at which adding compensation capacity produces the greatest system-level benefit. The question of which indicator, or which combination of indicators, best guides the selection of reactive compensation investment sites has not been the subject of a dedicated evaluation.

Other approaches reduce computational effort without pre-selecting sites, by identifying during the solution process the combinations of operating conditions and contingencies that are effectively binding [71]. This reduction operates on the constraint space of the operational problem, not on the investment decision space, and is therefore orthogonal to any candidate site pre-selection. Sparse formulations penalise the number of retained sites directly in the objective function, delegating selection to the optimisation itself. These two approaches are complementary to a candidate site pre-selection rather than substitutes for it.

Any pre-selection of candidate sites upstream of the optimisation restricts the decision space and may therefore exclude the optimal solution of the full problem. In the general case, no a priori bound on the optimality loss induced by such a restriction is available. The existing literature offers no criterion for assessing the fidelity of a pre-selection other than a posteriori, by comparison with the solution of the unreduced problem. The empirical evaluation of this fidelity is therefore a necessary prerequisite to the use of any pre-selection method.

2.1.9. Synthesis, Research Gaps and Contributions

The review has examined existing reactive power planning approaches along three dimensions of operational model fidelity. Cost representation relies predominantly on proxy quantities whose domain of validity narrows as operating conditions depart from the inductive regime. The physical representation of the network through convex relaxation loses exactness in capacitive configurations, where the established guarantees do not apply. The reduction of the decision space draws on indicators designed for heavily loaded conditions with a reactive power deficit, whose relevance for localising a capacitive reactive need has not been established.

The components of the reactive power planning problem have been addressed separately in the recent literature. Security incorporation and its scaling to large networks have been established under detailed network models [71, 67]. Decision space reduction has been approached through iterative site selection or through voltage stability indicators [84, 81]. Cost has been represented predominantly through proxy quantities such as active losses or investment components [41, 66]. No existing formulation, however, combines an objective representing the actual operating cost borne by the system operator, an explicit treatment of security constraints, and a reduction of the decision space, within a network relaxation that remains faithful under capacitive operating conditions.

Beyond this primary gap, the review brings to light two more specific questions that remain open. The domain of validity of convex relaxations has not been characterised for the capacitive regime, even as this regime is becoming a dimensioning condition for transmission networks. The selection of an indicator for decision space reduction has not been evaluated against the task of localising a reactive compensation need, as distinct from estimating proximity to voltage instability. These two questions, independent of the primary gap, constitute two of the contributions of the chapter.

The demonstrations pursued in this chapter concern the validation of specific properties rather than the establishment of scalability. Characterising the exactness of a relaxation or the behaviour of a weak bus indicator requires an exact reference, obtained through a global solution procedure. The richness of the operational model retained, combining a conic relaxation, an operational cost objective, and security constraints across multiple operating conditions, increases the computational cost of this reference and bounds the network size on which it remains tractable. The choice of a reduced test case thus follows from a tradeoff between model fidelity and network size, the extension to larger networks constituting a distinct perspective bearing on computational tractability rather than on the properties demonstrated here.

This chapter makes three contributions in response to the gaps identified by the review:

- A security-constrained reactive power planning formulation covering multiple operating conditions, minimising jointly an investment cost and an operational cost that represents the cost structure actually borne by the system operator.
- A regularisation term that restores the exactness of the conic relaxation in capacitive configurations, where it otherwise degrades.

- A weak bus analysis method that restricts the investment decision space prior to optimisation, evaluated with respect to its fidelity to the unreduced problem.

A sensitivity analysis complements these contributions by examining the effect of the number of operating conditions and of the demand and generation growth scenario considered for the planning horizon.

The remainder of this chapter is organised as follows. Section 2.2 presents the VARPlan pipeline, detailing the reactive power planning formulation, the security-constrained operational dispatch, the market clearing imitation step, and the weak bus analysis. Section 2.3 describes the construction of the operating scenarios used throughout the chapter. Section 2.4 introduces the simulation framework and the test network. Section 2.5 reports the results, including the validation of the SOCP formulation, the weak bus analysis, the trajectory comparison, and the sensitivity analysis. Section 2.6 discusses these results and their limitations. Section 2.7 concludes the chapter.

2.2. Methodology: the VARPlan Pipeline

The VARPlan methodology is organised as a sequential pipeline, illustrated in Figure 2.2. A pre-processing stage first simulates a secure dispatch consistent with a given operating scenario o , drawn from a set $\mathcal{O}$ of representative operating conditions describing demand levels and renewable generation, whose construction is detailed in Section 2.3. This stage, formulated as a Security-Constrained Optimal Power Flow (SC-OPF) and referred to as Market Clearing Imitation (MCI), is a dispatch problem and does not reproduce the clearing mechanism of the electricity market itself. It explicitly incorporates N-1 security constraints and produces the active power dispatch of conventional generators, the curtailment of wind farms, the transformer tap positions, and any necessary load shedding. The actual European market clearing operates on a zonal basis and does not guarantee a physically feasible outcome on the real network. The MCI instead provides a tractable nodal proxy that produces a secure and feasible operating point. This operating point is then fixed as a parameter for the reactive power formulations. A reactive slack variable, consistent with the one used in the downstream formulations, is included in the reactive power balance to guarantee convergence; any non-zero value is reported as an MCI output alongside the active dispatch. The formulation itself is standard and is detailed in Section 2.2.1.

Building on this operating point, the methodology proceeds through the security-constrained Optimal Reactive Power Dispatch (ORPD), which optimises the operational setpoints of the network for a given operating condition. A first ORPD

run, computed on the currently installed asset base, provides the performance indicators from which the Weak Bus Analysis (WBA), detailed in Section 2.2.3, ranks the nodes of the network by criticality. This ranking restricts the set of candidate buses considered for investment by the Reactive Power Planning (RPP) stage. Each candidate bus introduces additional binary investment variables, so an unrestricted candidate set would make the resulting Mixed-Integer Second-Order Cone Programming (MISOCP) intractable at realistic network sizes. The WBA confines the search to the locations where compensation is actually needed. Alternatively, the RPP can be run directly over the unrestricted candidate set, an option retained to evaluate the fidelity of the WBA restriction against the unreduced problem. The RPP formulation, presented in Section 2.2.4 as the chapter's central contribution, optimises the investment in compensation assets, embedding for each operating scenario and contingency a further minimisation of the operational cost under the same operational constraints as the ORPD. Both the ORPD and RPP are formulated as mathematical programs combining continuous and binary decision variables and explicitly embedding N-1 security constraints. Once the investment decisions are obtained, a final ORPD run, considering the newly installed capacities, quantifies their impact on system performance.

2.2.1. Market Clearing Imitation (MCI / SC-OPF)

The purpose of this stage is to determine the active power setpoints of the network under security constraints. Its output is used as a substitute for the market outcome, in accordance with the exogeneity argument of Section 2.1.1. The MCI provides the secure operating point on which the ORPD and RPP formulations of Sections 2.2.2 and 2.2.4 are built, as introduced in Section 2.2. It is formulated as a standard SC-OPF and solved as a non-convex AC nonlinear program. It can be invoked at different stages of the pipeline: prior to any investment, using only the compensation assets currently installed on the network, or after the RPP investment decision, using the resulting expanded asset base. The set of compensation assets $\mathcal{A}$ considered by the MCI is therefore a parameter of each invocation, not fixed a priori, and their installed capacity $\bar{Q}_a$ is likewise treated as a known parameter rather than a decision variable at this stage. It becomes a decision variable jointly optimised with the investment indicator x_a in the RPP formulation of Section 2.2.4.

Although the MCI solves a full AC OPF, which inherently requires balancing reactive power throughout the network to reach a feasible AC operating point, only a subset of its solution is retained and fixed as a parameter for the RPP and ORPD formulations: the active power dispatch of conventional generators

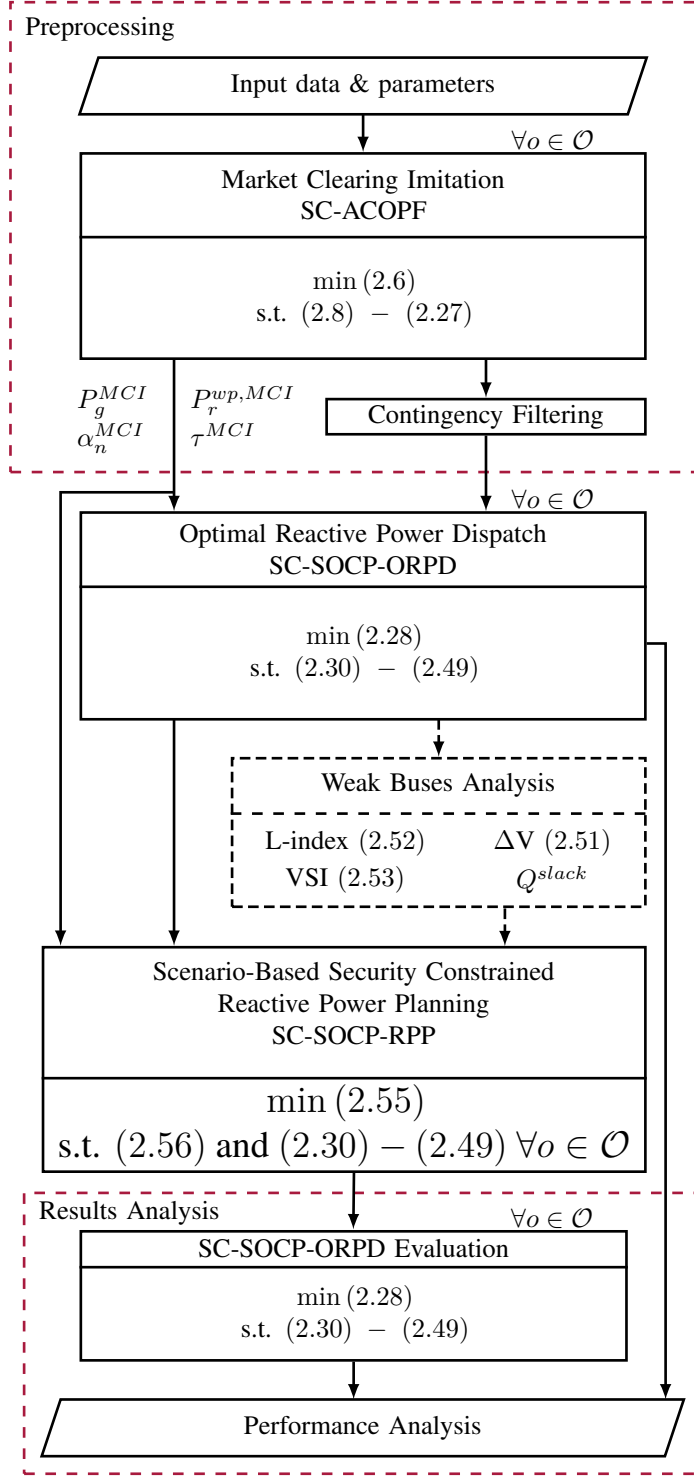


Figure 2.2.: VARPlan methodology: framework for reactive power planning in HV transmission networks.

$P_{g,k}$, the transformer tap positions τ_{ij} , the load shedding α_n , and the wind curtailment β_r . The reactive power setpoints computed internally by the MCI for generators, wind farms, and compensation assets are not retained: the ORPD formulation of Section 2.2.2 re-optimises the full reactive dispatch freely. The only reactive-related output retained is the slack variable $Q_{n,k}^{\text{slack}}$, consistent with the one used in the RPP and ORPD formulations, reported alongside the active dispatch whenever non-zero.

Objective function: The objective function, given in Equation (2.6), minimises the cost of active generation, load shedding, and wind curtailment, together with the reactive slack penalty. Only the slack term is evaluated across the full contingency set $k \in \mathcal{K}$; all other terms are evaluated at the base-case dispatch (subscript 0).

$$F_o^{\text{MCI}}(\Theta_o^{\text{MCI}}) = \sum_{g \in \mathcal{G}} \gamma_g^p P_{g,0} + \sum_{n \in \mathcal{N}_l} \gamma^{\text{shed}} \alpha_n |P_{n,0}^{\text{load}}| + \sum_{r \in \mathcal{R}} \gamma^{\text{curtail}} \beta_r P_{r,0}^{\text{wp}} + \sum_{k \in \mathcal{K}} \sum_{n \in \mathcal{N}} \gamma^{\text{slack}} |Q_{n,k}^{\text{slack}}| \quad (2.6)$$

Decision variables:

$$\Theta_o^{\text{MCI}} = \{P_{g,k}, Q_{g,k}, V_{n,k}, \theta_{n,k}, \alpha_n, \beta_r, \tau_{ij}, Q_r^{\text{wp}}, Q_{a,k}, Q_{n,k}^{\text{slack}}, b_a\} \quad (2.7)$$

Bounds: Constraints (2.8)-(2.15) define the physical and operational limits of the system. Constraints (2.8) and (2.9) bound the active and reactive power output of each conventional generator; (2.10) bounds the voltage magnitude at each bus, and (2.11) the phase angle difference across each branch. Constraint (2.12) enforces the thermal capacity limit of each transmission line, (2.13) bounds the transformer tap ratio, and (2.14) restricts the load shedding fraction to the unit

interval. Constraint (2.15) fixes the voltage angle at the reference bus.

$$\underline{P}_g \leq P_{g,k} \leq \overline{P}_g \quad \forall g \in \mathcal{G}, k \in \mathcal{K} \quad (2.8)$$

$$\underline{Q}_g \leq Q_{g,k} \leq \overline{Q}_g \quad \forall g \in \mathcal{G}, k \in \mathcal{K} \quad (2.9)$$

$$\underline{V}_n \leq V_{n,k} \leq \overline{V}_n \quad \forall n \in \mathcal{N}, k \in \mathcal{K} \quad (2.10)$$

$$\underline{\Delta\theta}_{ij} \leq \theta_{j,k} - \theta_{i,k} \leq \overline{\Delta\theta}_{ij} \quad \forall (ij) \in \mathcal{E}, k \in \mathcal{K} \quad (2.11)$$

$$P_{ij,k}^2 + Q_{ij,k}^2 \leq \overline{S}_{ij}^2 \quad \forall (ij) \in \mathcal{E}, k \in \mathcal{K} \quad (2.12)$$

$$\underline{\tau}_{ij} \leq \tau_{ij} \leq \overline{\tau}_{ij} \quad \forall (ij) \in \mathcal{E} \quad (2.13)$$

$$0 \leq \alpha_n \leq 1 \quad \forall n \in \mathcal{N} \quad (2.14)$$

$$\theta_{n,k} = 0 \quad \forall n \in \mathcal{N}^{\text{ref}}, k \in \mathcal{K} \quad (2.15)$$

Power balance and branch flows: The active and reactive power balance at each bus is enforced by (2.16)-(2.17). The reactive balance (2.17) aggregates reactive injections from conventional generators, wind farms, and compensation assets, and introduces the slack variable $Q_{n,k}^{\text{slack}}$.

$$\begin{aligned} \sum_{nj \in \mathcal{E}} P_{nj,k} &= \sum_{g \in \mathcal{G}_n} P_{g,k} + \sum_{r \in \mathcal{R}_n} (1 - \beta_r) P_r^{wp} \\ &\quad - (1 - \alpha_n) P_n^{\text{load}} \quad \forall n \in \mathcal{N}, k \in \mathcal{K} \end{aligned} \quad (2.16)$$

$$\begin{aligned} \sum_{nj \in \mathcal{E}} Q_{nj,k} &= \sum_{g \in \mathcal{G}_n} Q_{g,k} + \sum_{r \in \mathcal{R}_n} Q_r^{wp} - Q_n^{\text{load}} \\ &\quad + \sum_{a \in \mathcal{A}_n} Q_{a,k} + Q_{n,k}^{\text{slack}} \quad \forall n \in \mathcal{N}, k \in \mathcal{K} \end{aligned} \quad (2.17)$$

Branch power flows are expressed directly in terms of voltage magnitudes V , voltage angles θ , and tap ratios τ , separately for the direct (ij) and reverse (ji)

direction of each branch, through (2.18)-(2.21).

$$P_{ij,k} = \frac{V_{i,k}^2 G_{ij}}{\tau_{ij}^2} - \frac{V_{i,k} V_{j,k}}{\tau_{ij}} (G_{ij} \cos \theta_{ij,k} + B_{ij} \sin \theta_{ij,k})$$

$$\forall (ij) \in \mathcal{E}, k \in \mathcal{K} \quad (2.18)$$

$$Q_{ij,k} = \frac{-V_{i,k}^2 B_{ij}}{\tau_{ij}^2} - \frac{V_{i,k} V_{j,k}}{\tau_{ij}} (G_{ij} \sin \theta_{ij,k} - B_{ij} \cos \theta_{ij,k})$$

$$\forall (ij) \in \mathcal{E}, k \in \mathcal{K} \quad (2.19)$$

$$P_{ji,k} = V_{j,k}^2 G_{ij} - \frac{V_{i,k} V_{j,k}}{\tau_{ij}} (G_{ij} \cos \theta_{ij,k} - B_{ij} \sin \theta_{ij,k})$$

$$\forall (ij) \in \mathcal{E}, k \in \mathcal{K} \quad (2.20)$$

$$Q_{ji,k} = -V_{j,k}^2 B_{ij} + \frac{V_{i,k} V_{j,k}}{\tau_{ij}} (G_{ij} \sin \theta_{ij,k} + B_{ij} \cos \theta_{ij,k})$$

$$\forall (ij) \in \mathcal{E}, k \in \mathcal{K} \quad (2.21)$$

Security constraints: N-1 security follows a preventive-corrective scheme, as in [71]: the active power re-dispatch of each generator is bounded relative to the base case by its maximum admissible re-dispatch $\overline{\Delta P}_g$ (Equation (2.22)), and the voltage magnitude at generator buses is fixed to their pre-contingency values (Equation (2.23)).

$$|P_{g,k} - P_{g,0}| \leq \overline{\Delta P}_g \quad \forall g \in \mathcal{G}, k \in \mathcal{K}_0 \quad (2.22)$$

$$V_{n,k} = V_{n,0} \quad \forall n \in \mathcal{N}_g, k \in \mathcal{K}_0 \quad (2.23)$$

Compensation assets: Constraints (2.24)-(2.25) model the operation of compensation assets, with $\overline{Q}_a$ treated here as a known parameter. For stepped shunt devices ($a \in \mathcal{A}^d$), constraint (2.24) discretises the reactive output through the binary variable b_a , setting it to either zero or the installed capacity $\overline{Q}_a$. For continuously controllable devices ($a \in \mathcal{A}^c$), constraint (2.25) bounds the reactive output within the installed capacity.

$$Q_{a,k} = b_a \overline{Q}_a \quad \forall a \in \mathcal{A}^d, k \in \mathcal{K} \quad (2.24)$$

$$-\overline{Q}_a \leq Q_{a,k} \leq \overline{Q}_a \quad \forall a \in \mathcal{A}^c, k \in \mathcal{K} \quad (2.25)$$

Wind power plants:

$$(Q_r^{wp})^2 + ((1 - \beta_r)P_r^{wp})^2 \leq \bar{S}_r^2 \quad \forall r \in \mathcal{R} \quad (2.26)$$

$$-0.35 \cdot S_r^{wp,inst} \leq Q_r^{wp} \leq 0.4 \cdot S_r^{wp,inst} \quad \forall r \in \mathcal{R} \quad (2.27)$$

Constraint (2.27) bounds the reactive power output of each wind farm as a fraction of its installed apparent power $S_r^{wp,inst}$, with coefficients reflecting the reactive capability requirements for modern offshore wind farms specified by European grid connection regulations [20].

2.2.2. Security-Constrained ORPD (SC-SOCP)

The ORPD tries to capture the TSO's reactive power and voltage control once the active power schedule has been fixed by the market. It is formulated both as a standalone problem and as the operational subproblem embedded within the RPP of Section 2.2.4, with the objective of minimising the operational cost associated with reactive power management while satisfying network operational constraints and explicitly accounting for N-1 security requirements.

Objective function: The objective function, given in Equation (2.28), comprises seven cost components, each weighted by a dedicated cost coefficient γ . The inclusion of active power adjustment costs explicitly reflects the role of the TSO in a liberalised electricity market. Active power dispatch is determined by the MCI of Section 2.2.1, but the TSO remains responsible for ensuring secure system operation in real time. This may require redispatch actions to maintain voltage levels and satisfy security constraints, particularly under contingency conditions. As in the MCI objective of Section 2.2.1, only the reactive slack term C^{slack} is evaluated across the full set of contingencies $k \in \mathcal{K}$. All other terms are evaluated at the base-case operating point (subscript 0) only.

$$\begin{aligned} \min_{\Theta_o^{ORPD}} & \underbrace{\sum_{g \in \mathcal{G}} \gamma_g^q |Q_{g,0}|}_{C^{qg}} + \underbrace{\sum_{g \in \mathcal{G}} \gamma_g^{\delta p} |\delta P_{g,0}|}_{C^{\delta pg}} + \underbrace{\sum_{a \in \mathcal{A}} \gamma_a^c |Q_{a,0}|}_{C^{qa}} \\ & + \underbrace{\sum_{r \in \mathcal{R}} \gamma_r^{qwp} |Q_r^{wp}|}_{C^{qwp}} + \underbrace{\sum_{r \in \mathcal{R}} \gamma_r^{pwp} |\delta P_r^{wp}|}_{C^{\delta wp}} + \underbrace{\sum_{k \in \mathcal{K}} \sum_{n \in \mathcal{N}} \gamma^{slack} |Q_{n,k}^{slack}|}_{C^{slack}} \\ & + \underbrace{\sum_{ij \in \mathcal{E}^{cap}} \gamma^{line} \max(Q_{ij,0}^{line}, 0)}_{C^{line}} \end{aligned} \quad (2.28)$$

The first three terms price the reactive power output of generators (C^{qg}), the active power redispatch of generators relative to their MCI setpoint ($C^{\delta pg}$), and the reactive power output of compensation assets (C^{qa}). The next two terms price the same quantities for wind power plants: reactive power output (C^{qwp}) and active power adjustment relative to the MCI schedule ($C^{\delta wp}$). The term C^{slack} penalises the reactive slack variable introduced in the power balance (Equation (2.39)). The last term, C^{line} , regularises the base-case reactive power flow on branches $\mathcal{E}^{cap} = \{ij \in \mathcal{E} : Q_{i,o}^{load} < 0 \text{ or } Q_{j,o}^{load} < 0\}$, i.e., branches incident to a bus with a net capacitive load. Under inductive loading, the SOCP relaxation of the voltage product constraint (2.31) is essentially exact; under capacitive loading, however, the relaxation can be exploited as an artificial reactive power sink, allowing branch flows beyond what the AC equality would permit. Penalising $\max(Q_{ij,0}^{line}, 0)$ with a small coefficient γ^{line} restricts this behaviour and keeps the operating point close to AC-feasible; the resulting relaxation gap and the effect of γ^{line} are analysed quantitatively in Section 2.5 and discussed in Section 2.6. The lower bound property of Section 2.1.4 holds for the relaxation without this term. With $\gamma^{line} > 0$, the optimal value is a bound only for the penalised problem.

Decision variables: The decision variable set, given in Equation (2.29), includes voltage magnitudes, represented as squared values $W_{n,k}$ in accordance with the SOCP convexification of [48], active power adjustments of conventional generators and wind farms, reactive power outputs of generators and renewable resources, reactive power outputs of compensation assets, the reactive slack variable $Q_{n,k}^{slack}$, and the binary on/off variable b_a governing the discrete operation of stepped compensation assets.

$$\Theta_o^{ORPD} = \{W_{n,k}, W_{ij,k}^I, W_{ij,k}^R, \delta P_{g,k}, Q_{g,k}, Q_r^{wp}, Q_{a,k}, Q_{n,k}^{slack}, \delta P_r^{wp}, b_a\} \quad (2.29)$$

Bounds and SOCP relaxation: Constraints (2.30)-(2.35) define the physical and operational limits of the system. Constraint (2.30) bounds the squared voltage magnitude at each bus; (2.31) is the SOCP relaxation of the voltage product, following [48], and (2.32) linearises the phase angle difference bounds.

$$\underline{V}_n^2 \leq W_{n,k} \leq \overline{V}_n^2 \quad \forall n \in \mathcal{N}, k \in \mathcal{K} \quad (2.30)$$

$$(W_{ij,k}^R)^2 + (W_{ij,k}^I)^2 \leq W_{i,k} W_{j,k} \quad \forall (ij) \in \mathcal{E}, k \in \mathcal{K} \quad (2.31)$$

$$\tan(\underline{\Delta}\theta_{ij})W_{ij,k}^R \leq W_{ij,k}^I \leq \tan(\overline{\Delta}\theta_{ij})W_{ij,k}^R \quad \forall (ij) \in \mathcal{E}, k \in \mathcal{K} \quad (2.32)$$

This relaxation is obtained from the exact AC power flow equations of the MCI (Section 2.2.1, Equations (2.18)-(2.21)) by substituting the bilinear voltage products $W_{n,k} = V_{n,k}^2$, $W_{ij,k}^R = V_{i,k}V_{j,k} \cos \theta_{ij,k}$, and $W_{ij,k}^I = V_{i,k}V_{j,k} \sin \theta_{ij,k}$, and re-

laxing the resulting equality to the SOCP of Equation (2.31). Constraint (2.33) limits the active power re-dispatch $\delta P_{g,k}$ of each conventional generator relative to its MCI setpoint $P_{g,k}^{\text{MCI}}$; (2.34) bounds the reactive power output of each generator, and (2.35) enforces the thermal capacity limit of each transmission line.

$$\underline{P}_g - P_{g,k}^{\text{MCI}} \leq \delta P_{g,k} \leq \overline{P}_g - P_{g,k}^{\text{MCI}} \quad \forall g \in \mathcal{G}, k \in \mathcal{K} \quad (2.33)$$

$$\underline{Q}_g \leq Q_{g,k} \leq \overline{Q}_g \quad \forall g \in \mathcal{G}, k \in \mathcal{K} \quad (2.34)$$

$$P_{ij,k}^2 + Q_{ij,k}^2 \leq \overline{S}_{ij}^2 \quad \forall (ij) \in \mathcal{E}, k \in \mathcal{K} \quad (2.35)$$

Quadratic Convex (QC) envelope tightening: Constraints (2.36)-(2.37) further tighten the SOCP relaxation through the QC envelope of [48], written in terms of the angle-bound midpoint and half-range $\phi_{ij} = (\overline{\Delta\theta}_{ij} + \underline{\Delta\theta}_{ij})/2$, $d_{ij} = (\overline{\Delta\theta}_{ij} - \underline{\Delta\theta}_{ij})/2$, and the voltage-bound sums $s_{ij}^f = \underline{V}_i + \overline{V}_i$, $s_{ij}^t = \underline{V}_j + \overline{V}_j$. Constraint (2.36) is the lower envelope of the bilinear voltage product, evaluated at the lower voltage bounds $\underline{V}_i, \underline{V}_j$; constraint (2.37) is the corresponding upper envelope, evaluated at the upper voltage bounds $\overline{V}_i, \overline{V}_j$.

$$\begin{aligned} & s_{ij}^f s_{ij}^t \left(\cos \phi_{ij} W_{ij,k}^R + \sin \phi_{ij} W_{ij,k}^I \right) \\ & - \underline{V}_j \cos d_{ij} s_{ij}^t W_{i,k} - \underline{V}_i \cos d_{ij} s_{ij}^f W_{j,k} \\ & \geq -\underline{V}_i \underline{V}_j \cos d_{ij} \left(\underline{V}_i \underline{V}_j - \overline{V}_i \overline{V}_j \right) \quad \forall (ij) \in \mathcal{E}, k \in \mathcal{K} \end{aligned} \quad (2.36)$$

$$\begin{aligned} & s_{ij}^f s_{ij}^t \left(\cos \phi_{ij} W_{ij,k}^R + \sin \phi_{ij} W_{ij,k}^I \right) \\ & - \overline{V}_j \cos d_{ij} s_{ij}^t W_{i,k} - \overline{V}_i \cos d_{ij} s_{ij}^f W_{j,k} \\ & \geq \overline{V}_i \overline{V}_j \cos d_{ij} \left(\underline{V}_i \underline{V}_j - \overline{V}_i \overline{V}_j \right) \quad \forall (ij) \in \mathcal{E}, k \in \mathcal{K} \end{aligned} \quad (2.37)$$

Power balance and branch flows: The active and reactive power balance at each bus is enforced by (2.38)-(2.39). The reactive balance (2.39) is central to the formulation. It aggregates reactive injections from conventional generators, wind farms, and compensation assets, and introduces the slack variable $Q_{n,k}^{\text{slack}}$. This variable quantifies the reactive power deficit at bus n under contingency k and is a direct indicator of reactive scarcity. It is further used by the WBA of

Section 2.2.3.

$$\begin{aligned} \sum_{nj \in \mathcal{E}} P_{nj,k} &= \sum_{g \in \mathcal{G}_n} (P_{g,k}^{\text{MCI}} + \delta P_{g,k}) + \sum_{r \in \mathcal{R}_n} (P_r^{wp,\text{MCI}} + \delta P_r^{wp}) \\ &\quad - (1 - \alpha_n) P_n^{\text{load}} \quad \forall n \in \mathcal{N}, k \in \mathcal{K} \end{aligned} \quad (2.38)$$

$$\begin{aligned} \sum_{nj \in \mathcal{E}} Q_{nj,k} &= \sum_{g \in \mathcal{G}_n} Q_{g,k} + \sum_{r \in \mathcal{R}_n} Q_r^{wp} - Q_n^{\text{load}} \\ &\quad + \sum_{a \in \mathcal{A}_n} Q_{a,k} + Q_{n,k}^{\text{slack}} \quad \forall n \in \mathcal{N}, k \in \mathcal{K} \end{aligned} \quad (2.39)$$

Branch power flows are expressed as functions of the SOCP variables W , W^R , and W^I through (2.40)-(2.43), obtained by substituting the same bilinear voltage products used above directly into the exact AC power flow equations of the MCI (Equations (2.18)-(2.21)).

$$P_{ij,k} = \frac{W_{i,k} G_{ij}}{\tau_{ij}^2} - \frac{W_{ij,k}^R G_{ij} + W_{ij,k}^I B_{ij}}{\tau_{ij}} \quad \forall (ij) \in \mathcal{E}, k \in \mathcal{K} \quad (2.40)$$

$$Q_{ij,k} = \frac{-W_{i,k} B_{ij}}{\tau_{ij}^2} + \frac{W_{ij,k}^R B_{ij} - W_{ij,k}^I G_{ij}}{\tau_{ij}} \quad \forall (ij) \in \mathcal{E}, k \in \mathcal{K} \quad (2.41)$$

$$P_{ji,k} = W_{j,k} G_{ij} - \frac{W_{ij,k}^R G_{ij} - W_{ij,k}^I B_{ij}}{\tau_{ij}} \quad \forall (ij) \in \mathcal{E}, k \in \mathcal{K} \quad (2.42)$$

$$Q_{ji,k} = -W_{j,k} B_{ij} + \frac{W_{ij,k}^R B_{ij} + W_{ij,k}^I G_{ij}}{\tau_{ij}} \quad \forall (ij) \in \mathcal{E}, k \in \mathcal{K} \quad (2.43)$$

Security constraints: N-1 security follows the same preventive-corrective scheme as the MCI (Section 2.2.1), imposed by (2.44)-(2.45). Constraint (2.44) limits the active power adjustment of each generator under contingency $k \in \mathcal{K}_0$ relative to the base case, reflecting the limited re-dispatch capability between operating points. Constraint (2.45) fixes the voltage magnitude at buses connected to conventional generators and to continuously controlled compensation assets ($\mathcal{N}_g \cup \mathcal{A}^c$) to their pre-contingency values, consistent with the voltage-regulating capability of these devices.

$$|\delta P_{g,k} - \delta P_{g,0}| \leq \overline{\Delta P}_g \quad \forall g \in \mathcal{G}, k \in \mathcal{K}_0 \quad (2.44)$$

$$V_{n,k} = V_{n,0} \quad \forall n \in \{\mathcal{N}_g \cup \mathcal{A}^c\}, k \in \mathcal{K}_0 \quad (2.45)$$

Wind power plants: Constraints (2.46)-(2.47) govern the operation of wind power plants. Constraint (2.46) caps the active production of each wind farm at the available power of the considered operating point. Constraint (2.47) decouples

reactive power from active production, applying the same reactive capability bound, as a fraction of the installed apparent power $S_r^{wp,inst}$, already introduced for the MCI formulation of Section 2.2.1 (Equation (2.27)).

$$P_r^{wp,MCI} + \delta P_r^{wp} \leq P_r^{wp,prod} \quad \forall r \in \mathcal{R} \quad (2.46)$$

$$-0.35 \cdot S_r^{wp,inst} \leq Q_r^{wp} \leq 0.4 \cdot S_r^{wp,inst} \quad \forall r \in \mathcal{R} \quad (2.47)$$

Compensation assets: Finally, constraints (2.48)-(2.49) model the operation of compensation assets, using the unified capacity notation $\overline{Q}_a$ introduced in Section 2.2.1. For stepped shunt devices ($a \in \mathcal{A}^d$), constraint (2.48) discretises the reactive output through the binary variable b_a , setting it to either zero or the installed step capacity $\overline{Q}_a$. For continuously controllable devices ($a \in \mathcal{A}^c$), constraint (2.49) imposes a capacity bound whose upper limit is determined by the investment decision made at the RPP stage.

$$Q_{a,k} = b_a \overline{Q}_a \quad \forall a \in \mathcal{A}^d, k \in \mathcal{K} \quad (2.48)$$

$$-\overline{Q}_a \leq Q_{a,k} \leq \overline{Q}_a \quad \forall a \in \mathcal{A}^c, k \in \mathcal{K} \quad (2.49)$$

AC benchmark variant: For benchmarking purposes, an AC counterpart of the above formulation is considered, in which the SOCP relaxation (2.31) is replaced by its underlying equality, resulting in the exact non-convex AC power flow. To preserve tractability, the discrete on/off model (2.48) is relaxed to a continuous reactive output bounded between zero and the installed capacity, and continuously controlled devices ($\mathcal{A}^c$) are excluded so that each compensation asset operates in a single direction only. This non-convex formulation is used exclusively for the validation results of Section 2.5; all other results use the full SOCP formulation described above.

2.2.3. Weak Bus Analysis

Four performance indicators are computed from the ORPD results of Section 2.2.2, for each combination of scenario $o \in \mathcal{O}$ and contingency $k \in \mathcal{K}$, giving a distribution of values per bus across all operating conditions. As multiple scenarios and contingencies are considered, each bus is associated with an aggregate indicator value derived from the 95th percentile of its distribution, a robust near-worst-case estimate that captures stressed but realistic operating conditions. Among the four indicators considered, the reactive slack $Q_{n,k}^{slack}$ is directly available from the ORPD solution (Equation (2.39)). At optimality, a non-zero value at bus n under contingency k signals that the reactive power balance could not be achieved

without artificial injection or absorption. The reactive slack is therefore a direct measure of reactive scarcity at that bus. The three remaining indicators, voltage deviation, the Voltage Sensitivity Index (VSI), and the L-index, are computed a posteriori from the optimal voltage profile obtained from the ORPD solution, and are defined below.

For each indicator $I \in \{V^{dev}, L, \text{VSI}, Q^{\text{slack}}\}$, the aggregate value at bus n is defined as the 95th percentile over all scenario-contingency pairs:

$$\bar{I}_n = P_{95}\left(\left\{I_{n,o,k} \mid o \in \mathcal{O}, k \in \mathcal{K}\right\}\right) \quad (2.50)$$

Voltage deviation: The voltage deviation at bus n under scenario o and contingency k measures the normalised departure from the reference voltage magnitude:

$$V_{n,o,k}^{dev} = \frac{|V_{ref} - V_{n,o,k}|}{V_{ref}} \quad (2.51)$$

where V_{ref} is the nominal voltage (1 p.u.) and $V_{n,o,k}$ is the voltage magnitude obtained from the ORPD solution. Large deviations indicate buses operating far from their rated voltage, signalling insufficient reactive power support.

L-index: The L-index is a scalar indicator of voltage stability proximity, quantifying how close each load bus is to voltage collapse [87]. For each load bus n , it is defined as:

$$L_{n,o,k} = 1 - \left| \sum_{j \in \mathcal{G}} F_{nj} \frac{V_{j,o,k}}{V_{n,o,k}} \right| \quad (2.52)$$

where $V_{j,o,k}$ and $V_{n,o,k}$ are the complex voltage phasors at generator bus j and load bus n , and F_{nj} is derived from the partitioned system admittance matrix. A value close to 0 indicates a stable bus, while a value approaching 1 signals proximity to voltage instability.

Voltage sensitivity index: The VSI quantifies the sensitivity of bus n voltage to reactive power variations across the network, derived from the inverse of the Jacobian submatrix $\mathbf{J}_{\mathbf{VQ},o,k}$ evaluated at the operating point (o, k) :

$$\text{VSI}_{n,o,k} = \sum_{j \in \mathcal{N}} \left(\mathbf{J}_{\mathbf{VQ},o,k}^{-1} \right)_{nj} \quad (2.53)$$

A higher $VSI_{n,o,k}$ indicates greater voltage sensitivity to reactive power compensation at bus n under the considered operating condition.

2.2.4. Reactive Power Planning (SC-SOCP RPP)

The proposed RPP formulation deliberately remains technology-agnostic. Rather than prescribing a specific commercial device, reactive compensation assets are classified only by their operational behaviour. Continuously controllable devices ($a \in \mathcal{A}^c$, such as SVCs or STATCOMs) are distinguished from fixed-step shunt devices ($a \in \mathcal{A}^d$, such as capacitor and reactor banks). Fixed-step shunt devices provide reactive power in a single direction only: capacitor banks ($\mathcal{A}^{d,cap}$) inject reactive power, while reactor banks ($\mathcal{A}^{d,ind}$) absorb it, with $\mathcal{A}^d = \mathcal{A}^{d,cap} \cup \mathcal{A}^{d,ind}$.

Rather than selecting a particular technology, the RPP problem determines, for each candidate location and device family, whether a compensation asset should be installed and, if so, at what capacity. For every asset $a \in \mathcal{A} = \mathcal{A}^c \cup \mathcal{A}^d$, a binary investment variable x_a decides whether the device is installed, and a continuous variable $\overline{Q}_a$ sizes its installed capacity. The sizing is kept continuous for tractability: a representation by standard sizes would require additional binary variables for each combination of candidate location, device family and standard size, in a mixed-integer problem that is already expensive to solve. At the planning stage, the continuous capacity is therefore used as a sizing target, and the corresponding standard equipment is selected in a later step.

The same variables x_a and $\overline{Q}_a$ are used for both device families. The family concerned by a constraint is indicated only by the index set ($\mathcal{A}^c$ or $\mathcal{A}^d$) over which it is enforced. The long-term decision variable set is therefore

$$\Theta^{LT} = \{x_a, \overline{Q}_a : a \in \mathcal{A}\} \quad (2.54)$$

The objective, stated in Equation (2.55), is to minimise the total cost incurred over a long-term investment horizon, comprising the long-term investment cost F^{LT} , incurred once, and the probability-weighted operational cost of the ORPD formulation of Section 2.2.2, embedded for each operating scenario $o \in \mathcal{O}$. As this operational cost recurs every year of the investment horizon, it is annualised through H , the total number of hours represented by the investment horizon (i.e., the number of years considered for the investment times the number of hours per

year), and weighted by π_o , the probability associated with scenario o .

$$\min_{\Theta^{RPP}} F^{LT}(\Theta^{LT}) + H \sum_{o \in \mathcal{O}} \pi_o F_o^{ORPD}(\Theta_o^{ORPD}) \quad (2.55)$$

$$\text{s.t.} \quad \overline{Q}_a \leq \hat{Q}_a x_a \quad \forall a \in \mathcal{A} \quad (2.56)$$

$$\text{ORPD constraints (Section 2.2.2)} \quad \forall o \in \mathcal{O} \quad (2.57)$$

Constraint (2.56) links the capacity variable of each device to its investment decision, bounding it to a maximum admissible capacity $\hat{Q}_a$ whenever the device is installed, and to zero otherwise. Constraint (2.57) embeds, for every scenario, the full operational subproblem of Section 2.2.2, coupling the investment and operational decisions within a single formulation. The complete decision variable set of the RPP problem, given in Equation (2.58), therefore combines the long-term investment variables with the operational variables Θ_o^{ORPD} of every scenario's ORPD subproblem.

$$\Theta^{RPP} = \{\Theta_o^{ORPD}, \Theta^{LT}\} \quad (2.58)$$

The investment cost, given in Equation (2.59), comprises two components, over all compensation assets: a fixed installation cost $C^{fix,inv}$ (representing civil engineering and grid connection costs, linked to the binary variable x_a) and a variable capacity cost $C^{var,inv}$ (expressed per unit of installed capacity $\overline{Q}_a$).

$$F^{LT}(\Theta^{LT}) = \sum_{a \in \mathcal{A}} \left(\underbrace{\gamma_a^{i,f} x_a}_{C^{fix,inv}} + \underbrace{\gamma_a^{i,c} \overline{Q}_a}_{C^{var,inv}} \right) \quad (2.59)$$

The resulting problem is a MISOCP, combining the binary investment variables with the SOCP relaxation of the embedded ORPD subproblems, detailed in Section 2.2.2.

2.2.5. Summary of Modelling Assumptions

This subsection collects the modelling assumptions introduced in the preceding sections, organised by theme.

Network model: The ORPD and RPP use a convex SOCP relaxation of the AC power flow, tightened by the QC envelope. The MCI alone solves the exact non-convex AC formulation.

Active/reactive decoupling: The MCI fixes P_g , load shedding, and wind curtailment as parameters for the ORPD/RPP formulations. All reactive setpoints

are discarded and freely re-optimised by the ORPD. Transformer tap ratios (τ_{ij}) offer an additional voltage control lever, but their bilinear coupling with voltage magnitudes introduces non-convexities. They are therefore treated as decision variables only within the non-convex MCI, and fixed as parameters in the convex ORPD/RPP formulations.

Assets: Compensation assets are classified by operational behaviour only (continuous/bidirectional vs. discrete/single-direction), independent of the underlying commercial technology. The installation decision is binary, while the installed capacity is continuous. Wind reactive capability is bounded as a fixed fraction of installed apparent power, independent of the active power output, following the European grid connection requirements of [20].

Security: N-1 security follows a preventive-corrective scheme: post-contingency active power redispatch is bounded, and voltage is held fixed at generator buses. Continuously-controlled compensation devices (SVCs/STATCOMs) are assumed capable of the same fast, automatic voltage regulation, so their buses are held fixed as well, in both the ORPD and RPP formulations.

Market and cost: The MCI acts as a nodal proxy for the actual zonal European market clearing, which does not itself guarantee network feasibility. Reactive power costs, redispatch costs, and investment costs are calibrated to represent plausible European TSO economics.

Decision-space and horizon: The WBA aggregates each indicator's distribution across scenarios and contingencies using its 95th percentile, rather than the worst case or the mean. The RPP investment decision is static over the planning horizon: installed capacities do not vary within the horizon, and the operational cost is annualised through H .

2.3. Scenario Generation and Operating Point Construction

The optimisation formulations of Section 2.2 rely on a set of representative operating conditions, each describing a complete network operating point together with an associated weight. This section describes the construction of these conditions from Belgian historical data, and their projection along several prospective trajectories. The process proceeds in five steps: source data preparation, construction of nodal active and reactive loads, wind generation scaling, reduction through double clustering, and trajectory definition. Scenario generation is not a contribution in itself; it provides the input support on which the preceding formulations are evaluated.

2.3.1. Source Data

The IEEE 14-bus test network, like most standard benchmark cases, specifies only a single nominal load and generation snapshot, with no associated temporal profile. Representative operating conditions are therefore constructed from two distinct historical datasets covering the year 2025 at a quarter-hourly resolution, used to scale this nominal snapshot over time. The first is the aggregated Belgian electricity consumption, drawn from the open data platform of the TSO [91]. The second is the Belgian offshore wind generation over the same period, expressed as capacity factors and sourced from the same platform [91]. Each series is normalised by its respective annual peak, producing a load factor series and a generation factor series both ranging between zero and one. This normalisation renders the two series transposable to a test network of arbitrary size, independently of the absolute levels of the Belgian system.

2.3.2. Nodal Load Construction

Load factors are applied to the nominal active load of each network bus, producing a time series of active load per bus. A demand growth factor, specific to the trajectory under consideration, scales this load to the future consumption level retained for that trajectory. A net-load correction accounts for the rise of decentralized PV generation, computed from the difference between a reference PV share and a target share specific to the trajectory, then distributed across buses in proportion to their nominal load. Under conditions of high PV penetration and low demand, this correction may bring the net active load of a bus to a negative value, the bus then behaving as a net power injector.

The reactive load of each bus follows a linear relationship with its active load, whose slope is set by the power factor of that bus. This power factor is drawn, for each bus, around the nominal value of the trajectory with a small dispersion, introducing heterogeneity in reactive behaviour across the network. A capacitive offset δ_i , expressed as a fraction of the peak demand of bus i , translates this linear relationship toward negative reactive values. When δ_i is strictly positive, buses operating below δ_i times their peak demand exhibit a negative net reactive power, reflecting the structural reactive injection of UCs under light load. An additive Gaussian noise term introduces residual variability in the active-reactive pairs obtained. The construction of the reactive load at each bus is summarised by Equation (2.60), whose combination across all buses produces, at each time step, a complete network operating point.

$$Q_{i,t} = \tan(\arccos \text{PF}_i) (P_{i,t} - \delta_i P_i^{\text{peak}}) + \varepsilon_{i,t}, \quad \varepsilon_{i,t} \sim \mathcal{N}(0, \sigma^2) \quad (2.60)$$

where PF_i is the power factor drawn at bus i , δ_i the capacitive offset, P_i^{peak} the peak demand of bus i , and $\varepsilon_{i,t}$ the additive Gaussian noise term. This synthetic representation stands for the reactive behaviour of the T-D interface connected at each bus, a behaviour that Chapter 3 estimates from measurements instead of assuming it. Replacing Equation (2.60) by such estimated profiles is discussed as a perspective in Section 4.3.4.

2.3.3. Wind Capacity Scaling and Injection

Offshore wind generation is connected to a dedicated bus of the test network, whose installed capacity must be scaled from the Belgian system. This scaling preserves the ratio between wind capacity and peak demand, by multiplying the Belgian installed capacity by the ratio of the active load peaks of the two systems, according to Equation (2.61).

$$N_{\text{net}}^{\text{wpp}} = \frac{P_{\text{net}}^{\text{peak}}}{P_{\text{be}}^{\text{peak}}} \cdot N_{\text{be}}^{\text{wpp}} \quad (2.61)$$

where N^{wpp} denotes the installed wind capacity, P^{peak} the active load peak, and the subscripts net and be refer to the test network and the Belgian system respectively. The wind generation time series at the dedicated bus is then obtained by multiplying the normalised capacity factors by this scaled installed capacity. The installed capacity itself depends on the trajectory under consideration, prospective trajectories retaining a higher wind capacity than the reference case. The generation series thus constructed enters the operating conditions as an input alongside the nodal loads.

2.3.4. Two-Stage Clustering of Operating Points

The full annual set of quarter-hourly operating points is reduced to a representative set through a two-stage clustering procedure. A first k -means is applied to the full load dataset, grouping the active and reactive columns of all buses jointly, and produces K_l clusters whose centroids constitute the representative load conditions. For each load cluster, a second k -means is applied in cascade to the wind generation values recorded at the same time steps as the points belonging to that cluster, producing K_w sub-clusters. The combination of both stages produces $K_l \times K_w$ operating conditions per trajectory, each pairing a representative load condition with a representative wind generation condition. The weight

Table 2.1.: Policy trajectory parameters.

Parameters	Present-Day	High-Demand	Demand Sufficiency
Power factor PF_i	0.90	0.85	0.85
WPP capacity (MW)	45	106.1	106.1
α_{PV}^{target}	0.127	0.299	0.299
α_{wind}^{target}	0.154	0.363	0.363
Capacitive offset δ_i	0	0.6	0.6
Growth factor	1.0	1.5	1.0
Decommissioned generators	—	G4, G5	G3, G4, G5

of an operating condition is the fraction of data points it aggregates, according to Equation (2.62), so that weights sum to unity over the full condition set.

$$\pi_o = \frac{N_o}{N_{\text{tot}}}, \quad \sum_{o \in \mathcal{O}} \pi_o = 1 \quad (2.62)$$

where N_o is the number of data points aggregated into operating condition o , and N_{tot} the total number of points. Unless otherwise specified, $K_l = 20$ and $K_w = 1$, producing 20 operating conditions per trajectory; the choice of K_l is supported by the sensitivity analysis of Section 2.5.4.

2.3.5. Policy Trajectories

Three prospective trajectories are considered: Present-Day, High Demand, and Demand Sufficiency, each defined by the parameter set of Table 2.1. These parameters cover the load power factor, the installed wind capacity, the PV and wind penetration targets, the capacitive offset, the demand growth factor, and the decommissioned conventional generation fleet. The trajectories draw inspiration from the network development scenarios established by the TSO, without reproducing their quantitative assumptions, and constitute contrasting futures specific to this work. They are explored separately rather than aggregated, no probabilistic weight being assigned to any of them. This treatment distinguishes the trajectories from the operating conditions internal to each one, which are weighted by their probability of occurrence.

The Present-Day trajectory represents the current state of the network, with no capacitive offset, no demand growth, and a full conventional generation fleet. The High Demand trajectory combines strong demand growth with increased wind capacity and higher renewable penetration targets, together with a capacitive

offset reflecting a more extensively cabled network. The Demand Sufficiency trajectory retains the same renewable assumptions and capacitive offset as High Demand, but without demand growth and with a more heavily decommissioned conventional fleet. The combination of contained demand, a significant capacitive offset, and a reduced conventional fleet makes Demand Sufficiency the trajectory exhibiting the most capacitive behaviour. It is accordingly the trajectory that places the greatest stress on the network's capacity to absorb reactive power.

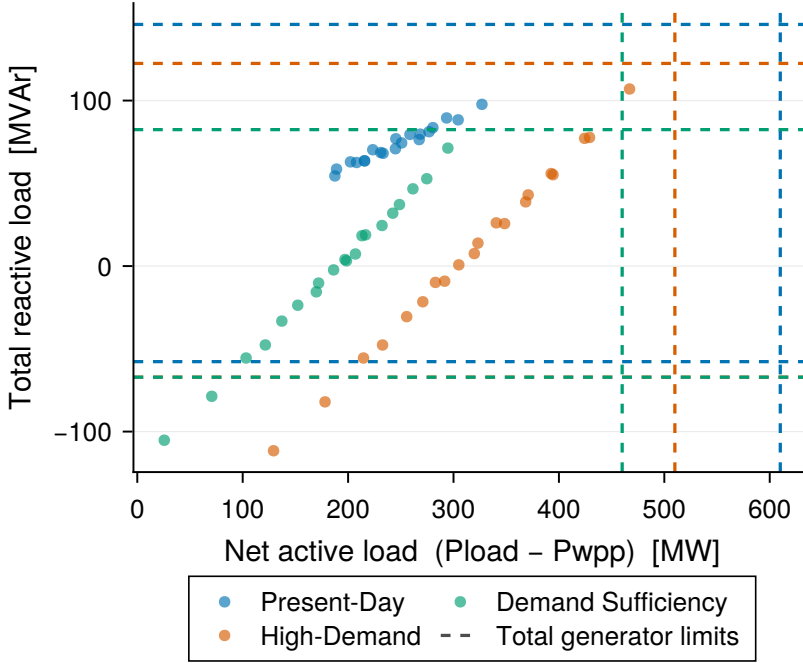


Figure 2.3.: P-Q Diagram for the three policy trajectories considered ($K_l = 20$ centroids each). For each trajectory, the dashed lines, shown in the matching colour, mark the total active power capacity of all generation sources (vertical) and the total reactive power capacity of all compensation means (horizontal).

Figure 2.3 shows the distribution of the $K_l = 20$ representative operating points obtained after clustering. Each point reports, for a given operating condition, the net active load aggregated across all buses ($P^{\text{load}} - P^{\text{wpp}}$, x-axis) against the total reactive load aggregated across all buses (y-axis), with negative values denoting a net reactive injection and positive values a net reactive absorption. For each trajectory, the dashed lines, shown in the corresponding colour, indicate the total active power capacity of all generation sources and the total reactive power capacity of all compensation means. These lines bound the domain reachable by that trajectory (without accounting for the network's own constraints, such

as line flow limits, or for the reactive power that the lines themselves inject or absorb).

The *Present-Day* trajectory displays a tighter spread in both dimensions, consistent with its absence of demand growth and capacitive offset. For the two future trajectories, certain operating points exhibit negative total reactive power at the network level, reflecting the capacitive behaviour induced by $\delta_i > 0$ under low-load conditions. Despite the noise introduced in the reactive profiles, the operating points retain a quasi-linear distribution, a direct consequence of the shared Belgian load dataset underlying all buses and the smoothing effect of centroid-based clustering. *High Demand* and *Demand Sufficiency* exhibit comparable reactive profiles but are offset in active power, with *High Demand* reaching higher P values due to its demand growth factor.

2.4. Simulation Framework and Test Case

2.4.1. Test network

All simulations are carried out on a modified version of the IEEE 14-bus transmission test network [92]. The network comprises 14 buses, 20 branches, and 11 load buses. A wind power plant is connected at bus 12, and two fixed shunt capacitors are installed at buses 3 and 9. The voltage magnitude limits are set to $\underline{V}_n = 0.95$ p.u. and $\overline{V}_n = 1.05$ p.u. at all buses. The single-line diagram is shown in Figure 2.4. Load levels and generation conditions are scenario-dependent, as detailed in Section 2.3.

N-1 security is enforced on a reduced contingency set. The screening covers the outage of each of the 20 branches and the outage of each conventional generator, evaluated at the MCI stage over all operating conditions of the trajectory. The contingencies are ranked by severity, following the principle presented in Section 2.1.7: the primary criterion is the total reactive slack Q^{slack} accumulated over the operating conditions, and the contingencies for which this quantity is zero are ranked by their worst voltage deviation. The four most severe contingencies are retained, and the screening is repeated for each trajectory, so the retained set differs from one trajectory to another, as reported in Table 2.2. The number of retained contingencies is limited by the size of the RPP. This size grows with the product of the number of candidate sites, the number of operating conditions and the number of contingencies. The limit is stronger for the AC benchmark of Section 2.5.1. The sufficiency of this reduced set is assessed in Section 2.5.5.

2.4.2. Cost parameters

All cost coefficients are calibrated to reflect the economic signals faced by a European TSO operating in a liberalised electricity market, and are summarised in Table 2.3.

- The active power price corresponds to a day-ahead market proxy based on representative European spot prices.
- Reactive power costs at generators are derived from the Belgian regulatory framework (CREG): the loss ratio of 0.032 MWh/MVArh for synchronous generating units would give a theoretical value of 3.2 €/MVArh [27]. This coefficient is recalibrated upward to preserve the cost hierarchy across reactive resources while ensuring numerical stability.
- Compensator usage costs are set lower than generator reactive costs, with a further distinction between continuously controllable devices ($\mathcal{A}^c$) and stepped shunt devices ($\mathcal{A}^d$), reflecting their relative flexibility and switching behaviour.
- The reactive slack penalty is set several orders of magnitude above operational costs to enforce voltage feasibility as a binding constraint.
- Investment costs are sourced from the MISO Transmission Cost Estimation Guide [93], with SVCs/STATCOMs approximately 13 times more expensive per MVAr than capacitor and reactor banks; fixed costs represent the civil engineering and grid connection costs, while variable costs reflect the cost of reactive power capacity.

The resulting cost hierarchy reflects the relative economic signals faced by the TSO: active power adjustments are more expensive than reactive power provision. Reactive power provision is in turn more costly at generators than through dedicated compensation assets. Investment costs are significantly higher than operational costs, to reflect the long-term nature of capacity expansion decisions.

The annualisation factor H used in the RPP objective (Equation (2.55)) is computed over an investment horizon of $T = 10$ years at 8760 h/year, scaled by a basic annualisation rate $r = 2\%$:

$$H = 8760 \cdot T \cdot (1 + r)^T \quad (2.63)$$

Table 2.3.: Cost coefficients.

Symbol	Description	Value
<i>Operational costs</i>		
$\gamma_g^{\delta p}, \gamma_r^{pwp}$	Active redispatch	100 €/MWh
γ_g^q	Generator reactive output	10.0 €/MVarh
γ_a^c	Compensator usage ($\mathcal{A}^c$ / $\mathcal{A}^d$)	1.0 / 0.1 €/MVarh
γ_r^{qwp}	WPP reactive output	10.0 €/MVarh
γ^{slack}	Reactive slack penalty	10 000 €/MVarh
<i>Investment costs</i>		
$\gamma_a^{i,f}$	Fixed CAPEX ($\mathcal{A}^c$ / $\mathcal{A}^d$)	3 / 1 M€
$\gamma_a^{i,c}$	Variable CAPEX ($\mathcal{A}^c$ / $\mathcal{A}^d$)	200 / 15 k€/MVar

2.4.3. Computational setup

All optimisation problems are implemented in Julia using the JuMP modelling framework [94]. The solver used for each problem instance follows directly from its mathematical structure. Convex formulations, whether continuous or mixed-integer, are solved with Gurobi [95]. Non-convex formulations without integer variables are solved with IPOPT [96]. Non-convex formulations with integer variables are solved with Juniper, which combines Gurobi for the integer branching with IPOPT for the continuous NLP relaxations. The MCI of Section 2.2.1, a non-convex AC OPF with the discrete asset variable b_a , is accordingly solved with Juniper. The ORPD of Section 2.2.2 is solved with Gurobi in its SOCP form, and with Juniper in its non-convex AC benchmark form. The RPP of Section 2.2.4, which always retains the discrete investment variables x_a in addition to the embedded ORPD operational variables, follows the same rule: Gurobi when the embedded ORPD is SOCP, Juniper when it is AC. For the sensitivity analysis of the γ^{line} regularisation term reported in Section 2.5, the discrete operational switching of shunt devices (b_a) is further relaxed to a continuous variable within the embedded ORPD. This keeps the AC variant of the RPP computationally tractable for that comparison. The investment binaries x_a remain discrete throughout, so both variants remain mixed-integer programs. They are solved with Gurobi (SOCP) and Juniper (AC), respectively. All simulations are run on the Consortium des Équipements de Calcul Intensif (CECI) [97], using AMD Epyc Rome 7542 CPUs (32 cores at 2.9 GHz).

The operational scope on which each experiment of Section 2.5 is run varies with the purpose of the experiment, and is summarised in Table 2.4.

Table 2.4.: Operational scope of each experiment of Section 2.5.

Experiment	Operating conditions	Contingencies	Trajectories
SOCP validation (Section 2.5.1)	5	Base case only	Three trajectories
Weak bus analysis (Section 2.5.2)	20	Base case + 4	Demand Sufficiency
Trajectory comparison (Section 2.5.3)	20	Base case + 4	Three trajectories
Sensitivity to operating conditions (Section 2.5.4)	5 to 60, evaluated on 60	Base case + 4	Demand Sufficiency
Sensitivity to contingencies (Section 2.5.5)	20	Base case to base case + 10	Demand Sufficiency

2.5. Results

This section evaluates the proposed SC-SOCP RPP methodology and its WBA-guided bus selection on the IEEE 14-bus test case introduced in Section 2.4. Section 2.5.1 validates the SOCP relaxation against the AC reference on a reduced operational scope (5 operating conditions, base case only), where the AC formulation remains tractable; it covers the calibration of the regularisation term and the impact of restricting the candidate bus space. The remaining sections apply the validated SOCP formulation to the full-scale problem (20 operating conditions, base case and 4 contingencies). Section 2.5.2 evaluates the four WBA indicators and quantifies the trade-off between the size of the candidate bus space and solution quality. Section 2.5.3 compares investment decisions across the three planning trajectories introduced in Section 2.3. Section 2.5.4 reports the sensitivity of the proposed methodology to the number of operating conditions retained in the planning stage. Section 2.5.5 reports the sensitivity to the number of contingencies retained in the planning stage. The operational scope of each experiment is summarised in Table 2.4.

2.5.1. Validation of the SOCP formulation

Both formulations are compared through a two-step procedure, applied identically to the AC and to the SOCP formulation:

1. **Planning step.** The RPP is solved in the formulation under test, either AC or SOCP. Only the investment decisions are retained, that is, the installed device and its capacity for each shunt family.
2. **Evaluation step.** These investment decisions are fixed, and the resulting network is evaluated through an AC ORPD on the same operational scope for both formulations.

All costs reported in this section come from the evaluation step, and not from the objective value of the formulation that produced the decisions. What is compared is therefore the quality of the investment decisions issued by each formulation, assessed on a common physical basis.

The RPP is solved in both formulations under a relaxed variant in which the discrete investment variables of the shunt families $\mathcal{A}^{d, cap}$ and $\mathcal{A}^{d, ind}$ (Section 2.2.4) are relaxed to continuous outputs between zero and the installed capacity, and the continuously controlled devices $\mathcal{A}^c$ are removed. This ensures that the AC and SOCP formulations share the same investment decision space, isolating the conic relaxation as the only methodological difference. The AC reference is obtained through a multi-start IPOPT procedure followed by a two-stage refinement, which mitigates the non-convexity of the AC formulation that prevents IPOPT from systematically reaching the global optimum.

Figure 2.5 reports the total cost difference with respect to the AC reference (top) and the maximum conic residual gap (bottom) as functions of the regularisation weight γ^{line} , for the three trajectories. Under the most capacitive trajectory (Demand Sufficiency), the SOCP relaxation is loose at $\gamma^{line} = 0$. The conic residual gap is large, and the investment decisions of the SOCP formulation lead to an AC-evaluated cost well above the AC reference. As γ^{line} increases, three regimes are observed:

1. For small values, the regularisation is insufficient to counter the relaxation slackness and the cost remains elevated.
2. Within an intermediate critical range, the conic gap drops sharply and the cost reaches its minimum, indicating that the regularisation successfully suppresses the fictitious reactive absorption responsible for the relaxation gap.
3. For larger values, the conic gap continues to decrease slowly but the cost increases, driven by over-investment in $\mathcal{A}^{d, ind}$ capacity that the regularisation forces to be physically installed rather than fictitiously emulated through the relaxed conic constraint.

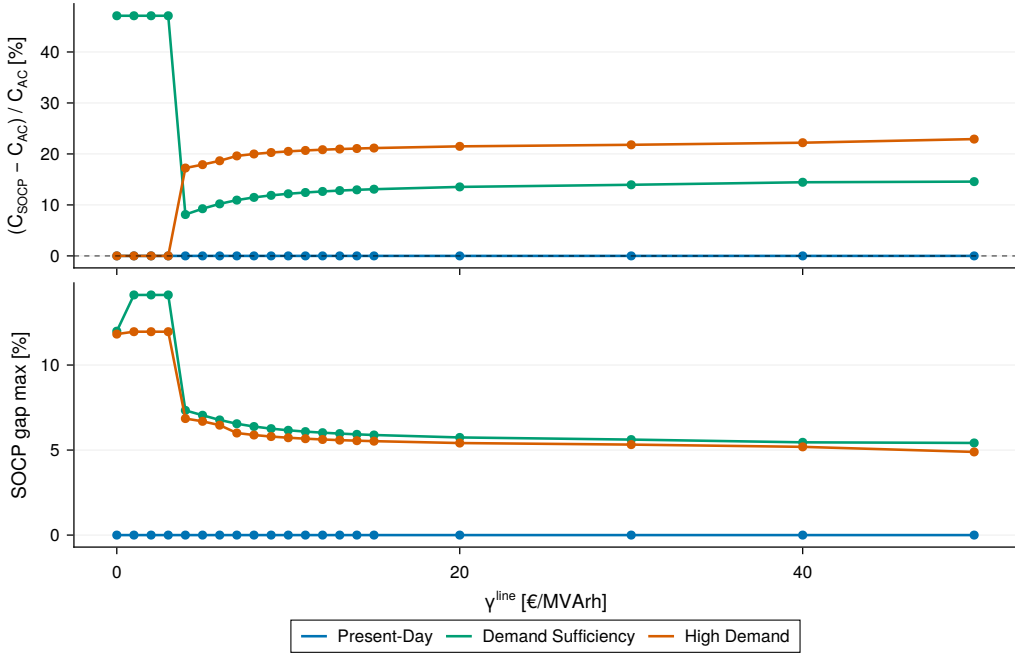


Figure 2.5.: Effect of the regularisation weight γ^{line} on the SOCP formulation across the three planning trajectories. Top: relative cost difference with respect to the AC reference. Bottom: maximum conic residual gap. Investment decisions are evaluated through an AC ORPD on the same operational scope (5 operating conditions, base case only).

The optimal γ^{line} is therefore located within the critical range, and the value $\gamma^{line} = 4 \text{ e/MVarh}$ is retained for the remainder of this chapter.

Under the High Demand trajectory, the conic residual gap is also elevated at $\gamma^{line} = 0$ (comparable in magnitude to that of Demand Sufficiency); however, the unregularised SOCP already produces investment decisions matching the AC reference. Increasing γ^{line} on this trajectory progressively reduces the conic gap but induces the same $\mathcal{A}^{d,ind}$ over-investment as in the third regime of Demand Sufficiency, monotonically degrading the total cost. The Present-Day trajectory, in contrast, exhibits a naturally tight relaxation with a negligible conic gap and a SOCP cost matching the AC reference independently of γ^{line} .

A noteworthy observation emerges from this analysis: High Demand and Demand Sufficiency exhibit comparable conic gaps at $\gamma^{line} = 0$, yet only the latter shows degraded investment decisions. The maximum conic residual gap therefore reflects local relaxation slackness but does not systematically translate into investment bias, and the cost difference evaluated through the AC ORPD remains the operational indicator of relaxation reliability. Consequently, the regularisation is activated only on trajectories where the unregularised SOCP degrades the AC-evaluated cost (exclusively Demand Sufficiency in this study). This trajectory-dependent activation is consistent with the physical interpretation of the regularisation term, which targets fictitious reactive absorption arising specifically in heavily capacitive scenarios; the design of an a priori activation criterion based on pre-solution scenario characteristics is discussed as a limitation in Section 2.6.

To further verify the reliability of the SOCP formulation, the comparison is repeated on three different bus sets restricting the candidate investment locations. The first set corresponds to the buses identified by the two-stage AC procedure as the optimal investment locations; the two additional sets explore alternative two-bus configurations. For each set, both the AC and SOCP formulations are solved on the Demand Sufficiency trajectory with $\gamma^{line} = 4 \text{ e/MVarh}$, and their resulting investment decisions are evaluated through the AC ORPD. Table 2.5 reports, for each shunt family, the number of installed devices and the installed capacity, together with the investment cost, the operational cost, and the total cost relative to the full-space AC reference.

Among the cost components of the ORPD objective (2.28), the economic terms of the operational cost reported in Table 2.5 reduce to C^{qa} . Once the compensation capacity is installed, the reactive power requirement is met by the compensation assets alone, since their usage cost is the lowest of all reactive resources (Table 2.3). The reactive power output of the generators and of the wind power plant, and the active power adjustments of both, are zero on all the runs considered here.

The regularisation term C^{line} is zero throughout, since the evaluation is carried out with the AC ORPD, which does not carry the γ^{line} penalty.

The corresponding comparison for the High Demand and Present-Day trajectories yields identical investment decisions between the AC and SOCP formulations on all three sets. This confirms that the SOCP formulation matches the AC reference exactly on these trajectories, regardless of the bus set. This comparison is therefore not reported.

Table 2.5.: Installed capacities, investment and operational costs on Demand Sufficiency for three bus sets. For each shunt family, the table reports the number of installed devices x and the corresponding installed capacity $\bar{Q}$ in MVar. Costs are in k€ and Δ is relative to the full-space AC reference. All costs are obtained from the AC ORPD evaluation of the investment decisions, whatever the formulation that produced them.

Set	Form.	$\mathcal{A}^{d,cap}$	$\mathcal{A}^{d,ind}$	C_{inv}	C_{ope}	C_{tot}	Δ (%)
		x / MVar	x / MVar				
Full	AC tot	1 / 31.0	1 / 35.6	2998.2	56.2	3054.4	—
Set 1 : {2,4}	AC	1 / 31.0	1 / 35.6	2998.2	56.2	3054.4	0.00
	SOCP	1 / 31.0	1 / 51.2	3233.1	56.2	3289.3	+7.69
Set 2 : {3,5}	AC	1 / 31.2	1 / 33.9	2976.1	55.2	3031.3	−0.76
	SOCP	1 / 31.2	1 / 54.3	3282.2	58.8	3341.0	+9.38
Set 3 : {1,6}	AC	1 / 33.3	1 / 60.2	3402.4	61.3	3463.7	+13.40
	SOCP	1 / 33.3	1 / 52.0	3279.9	61.3	3341.2	+9.39

Table 2.5 is commented set by set.

- **Set 1** {2,4}, the bus locations identified by the two-stage AC procedure. The AC formulation recovers exactly the full-space reference. The SOCP formulation on the same set retains the identical investment in $\mathcal{A}^{d,cap}$, 31.0 MVar in both cases, but over-invests in $\mathcal{A}^{d,ind}$, which rises from 35.6 MVar to 51.2 MVar. These 15.6 MVar of additional reactor capacity account for the whole 7.69% cost increase relative to the full-space reference.
- **Set 2** {3,5}, a bus composition entirely disjoint from Set 1. The AC solution reaches a slightly lower cost than the full-space reference (−0.76%). The SOCP formulation gives a 9.38% deviation, again driven by an over-investment in $\mathcal{A}^{d,ind}$: 54.3 MVar against 33.9 MVar for the AC solution, while $\mathcal{A}^{d,cap}$ stays at 31.2 MVar for both formulations.
- **Set 3** {1,6}, the configuration furthest from the active locations. The AC formulation produces a markedly degraded solution (+13.40%), with

60.2 MVar installed in $\mathcal{A}^{d,ind}$, which reflects the difficulty of the underlying non-convex problem on locations far from the active ones. The SOCP formulation remains within its usual relative deviation (+9.39%) with 52.0 MVar, and therefore outperforms the AC solution on this set.

Two observations follow. First, the AC solution of Set 2 is cheaper than the full-space reference, although Set 2 was chosen arbitrarily. Both the full-space AC solution and the two-stage refinement are therefore local optima of the non-convex problem rather than the global optimum, the latter being unreachable in practice for the considered problem size. Second, the differences with respect to the full-space reference originate almost entirely from the $\mathcal{A}^{d,ind}$ capacity, while the $\mathcal{A}^{d,cap}$ investments remain stable around the optimal value, between 31.0 MVar and 33.3 MVar for both formulations on all three sets. The $\mathcal{A}^{d,ind}$ family concentrates the whole discrepancy. The SOCP formulation sizes it in a narrow range, between 51.2 MVar and 54.3 MVar, whereas the AC formulation ranges from 33.9 MVar to 60.2 MVar. Read in installed capacity, the relaxed formulation is systematically conservative on the reactor family and much less sensitive to the bus selection than the AC formulation. This is the capacity counterpart of the consistent relative deviation of 7.7% to 9.4% observed on the total cost.

The higher costs reported for the SOCP formulation follow directly from the evaluation protocol. The SOCP solution is not feasible for the original non-convex problem (Section 2.1.4), so the investment decisions it produces are no longer optimal once they are evaluated through the AC ORPD. A higher evaluated cost is therefore the expected outcome of the comparison, and it does not affect the lower bound that the SOCP formulation provides on its own objective value. The existence of multiple AC bus configurations with comparable or significantly different costs highlights the non-uniqueness of local optima in the non-convex formulation, which is precisely the difficulty that motivates the convex SOCP formulation as a planning tool.

2.5.2. Weak Bus Analysis

The WBA procedure described in Section 2.2.3 is evaluated on the full-scale problem under the Demand Sufficiency trajectory. For each of the four indicators (V^{dev} , L-index, VSI, and Q^{slack}), the buses of the system are ranked by criticality, and the RPP is solved by restricting the candidate investment locations to the top- k buses of the ranking, for k ranging from 1 to the total number of candidate buses. The resulting solutions are compared to the AllBus reference obtained by solving the RPP on the full candidate space.

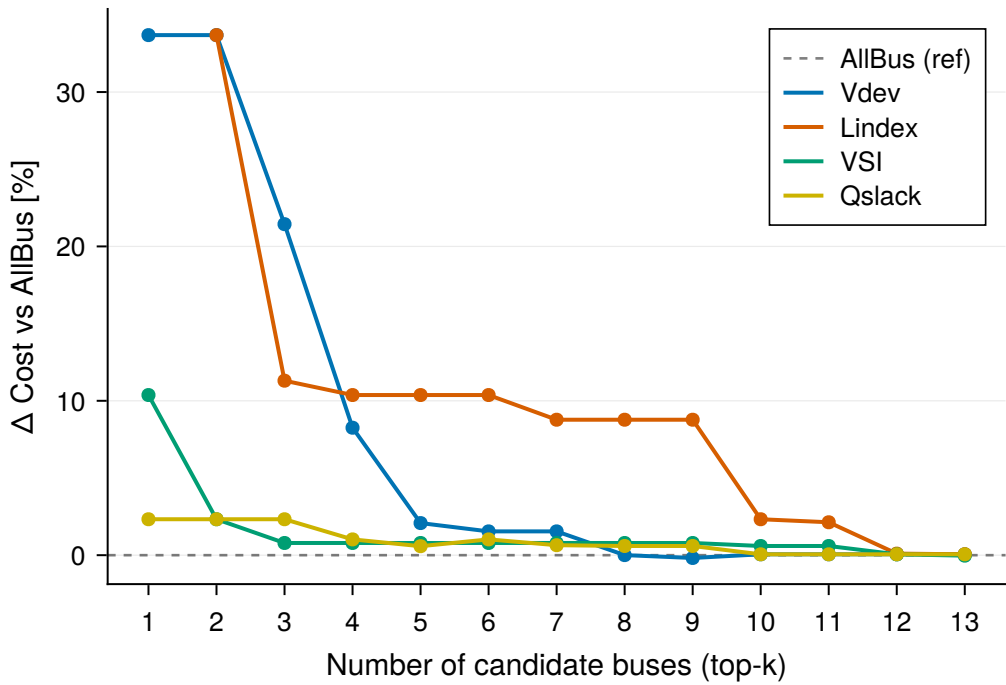


Figure 2.6.: Relative cost difference of top- k WBA configurations with respect to the AllBus reference, for the four indicators.

Figure 2.6 reports the relative cost difference of each top- k configuration with respect to the AllBus reference, for the four indicators. Two regimes clearly emerge. With Q^{slack} and VSI, the cost converges to within 1% of the AllBus reference from $k = 2$, indicating that the buses identified as most critical by these indicators capture the essential information for siting reactive investments. In contrast, V^{dev} and L-index require respectively 8 and 11 top-ranked buses to reach a comparable cost, meaning that the buses they prioritise in the ranking are not those most relevant for investment. The four indicators are therefore not equivalent, and the choice of indicator directly determines the achievable level of space reduction.

A noteworthy observation is that Q^{slack} at $k = 1$ achieves a solution within 2% of the AllBus reference while investing on buses that differ from those of the full-space solution. This confirms the non-uniqueness of the optimum already observed in Section 2.5.1: multiple bus configurations give solutions of comparable economic quality, and the WBA captures this criticality even when pointing to a different configuration than the one identified by the full-space optimisation. Based on these observations, Q^{slack} emerges as the most efficient indicator, achieving sub-1% cost deviation from the AllBus reference with as few as $k = 2$ candidate buses. For the remainder of this chapter, the candidate investment locations are fixed to the buses selected for investment by the AllBus solution; this choice does not affect the conclusions but reduces the computational burden of the subsequent experiments. The WBA, and Q^{slack} in particular, provides a scalable pre-processing strategy for larger systems where the AllBus formulation may exceed practical compute budgets.

2.5.3. Trajectory Comparison

Figure 2.7 reports the cost breakdown for the three planning trajectories introduced in Section 2.3, and Table 2.6 provides the detailed component values. The candidate investment locations are restricted to the buses selected for investment by the AllBus solution.

All three trajectories invest in capacitive compensation ($\mathcal{A}^{d,\text{cap}}$) to support voltage profiles, but Demand Sufficiency is the only one that also invests in inductive compensation ($\mathcal{A}^{d,\text{ind}}$), with a smaller capacitive capacity than the two other trajectories. This reflects the dominantly capacitive nature of Demand Sufficiency, which requires reactive absorption in addition to injection, whereas Present-Day is fully inductive and High Demand exhibits a nearly balanced profile with a slight inductive tilt. Among the latter two, Present-Day requires more capacitive capacity and reaches a higher total cost than High Demand. This is explained by the structural differences between the two trajectories: High Demand includes a

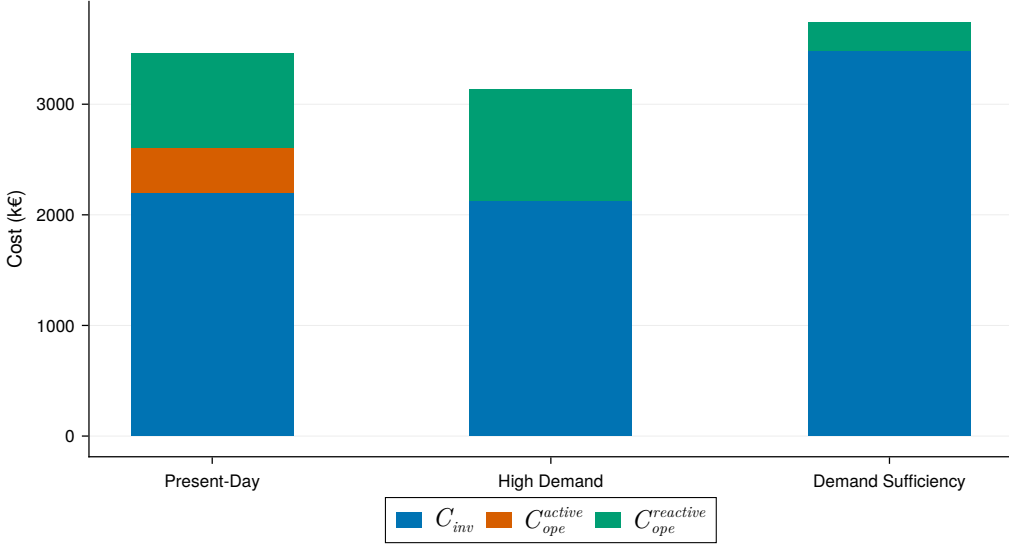


Figure 2.7.: Total cost breakdown for the three considered trajectories.

larger installed wind capacity, which provides additional reactive support, and operates with two fewer conventional generators, reducing the need for reactive adjustments.

Subscripts cc , dp , and dn denote $\mathcal{A}^c$, $\mathcal{A}^{d, cap}$, and $\mathcal{A}^{d, ind}$ respectively.

The C^{line} term reported for Demand Sufficiency is a numerical regularisation penalty rather than an economic cost borne by the operator; it is included in the reported total for completeness but does not correspond to a physical dispatch expense.

The operational costs reported in Table 2.6 clarify how each trajectory uses its installed assets. Despite having the largest investment, Demand Sufficiency exhibits the smallest operational cost: most of its operating conditions remain close to balanced configurations, so the newly installed compensators are activated only marginally on average. Present-Day, in contrast, must actuate its compensators substantially more than High Demand, again as a consequence of the operating condition distribution: at each operating condition, Present-Day uses its installed assets more intensively than High Demand, where alternative balancing resources are available. The breakdown of reactive support further reveals that High Demand draws extensively on the reactive capability of its wind farms, whereas Present-Day relies primarily on active redispatch of conventional generators combined with compensator actuation. These patterns are direct consequences of the resources available in each trajectory and illustrate how the model adapts its planning and operational decisions to the underlying physical configuration.

Table 2.6.: Investment and operational cost breakdown for the three planning trajectories (costs in k€).

Component	Present-Day	High Demand	Demand Sufficiency
<i>Investment costs</i>			
$C_{inv,cc}^{fixed}$	0.0	0.0	0.0
$C_{inv,cc}^{var}$	0.0	0.0	0.0
$C_{inv,dp}^{fixed}$	1000.0	1000.0	1000.0
$C_{inv,dp}^{var}$	1195.9	1128.4	750.1
$C_{inv,dn}^{fixed}$	0.0	0.0	1000.0
$C_{inv,dn}^{var}$	0.0	0.0	730.1
<i>Operational costs</i>			
$C_{\delta pg}$	410.5	0.0	0.0
C_{qg}	0.0	0.0	0.0
C_{qa}	851.4	370.2	261.5
C_{qwp}	0.0	639.9	0.0
$C_{\delta wp}$	0.0	0.0	0.0
C_{slack}	0.0	0.0	0.0
C_{line}	0.0	0.0	731.8
Total (C_{tot})	3457.8	3138.5	4473.5

2.5.4. Sensitivity to the Number of Operating Conditions

This subsection assesses the impact of the number of operating conditions retained through the clustering procedure of Section 2.3 on the quality of the resulting investment decisions. For each value of k ranging from 5 to 60 in steps of 5, the RPP is solved using k operating conditions, and the resulting investment decisions are evaluated through an ORPD on a fixed set of 60 operating conditions. This procedure isolates the impact of the planning sample size on the operational quality of the obtained investment decisions, independently of the optimisation process itself.

Figure 2.8 reports the total cost evaluated on the 60 operating conditions (left axis) and the CPU time of the RPP solve (right axis) as functions of k . With very few operating conditions, the planning stage converges rapidly but leads to investment decisions that are not representative of the broader operational scope, resulting in significantly higher evaluated costs. As k increases, the total cost decreases sharply and reaches a plateau at $k = 20$, while the CPU time grows exponentially. Beyond this plateau, additional operating conditions no longer improve the quality of the investment decisions but considerably increase the computational burden, which motivates the choice of 20 operating conditions retained in the previous analyses.

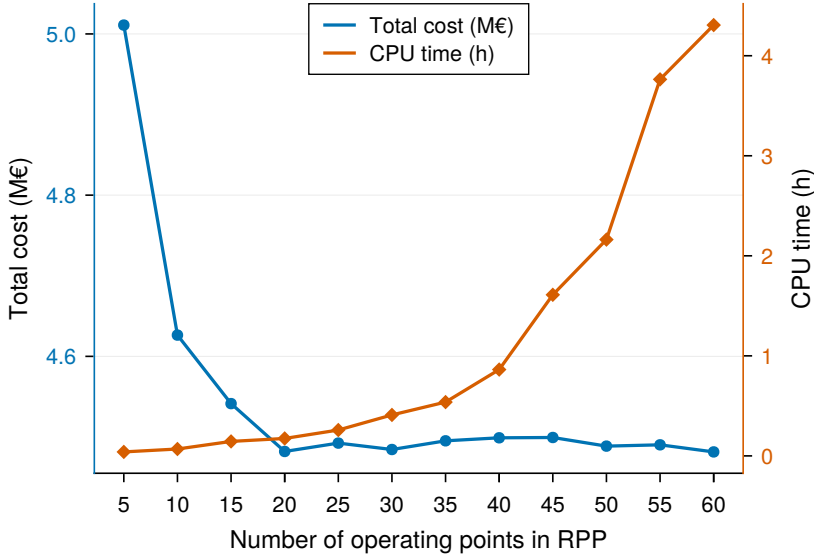


Figure 2.8.: Total cost evaluated on 60 operating conditions (left axis) and CPU time of the RPP solve (right axis) as functions of the number of operating conditions k used in the planning stage.

2.5.5. Sensitivity to the Number of Contingencies

This subsection assesses the impact of the number of contingencies retained in the planning stage. The RPP is solved on the Demand Sufficiency trajectory with 20 operating conditions, for contingency sets ranging from the base case alone to the base case and ten contingencies. The same two buses and the same two device families are selected in all cases. The installed capacities change by at most 2.3%, and the total cost by less than 0.3%. Retaining more than four contingencies therefore does not change the investment decisions on this test case. Once the compensation capacity is installed, the reactive power balance is satisfied without slack under all the contingencies considered, so the loss of a single element does not create an additional need for compensation. This behaviour is specific to the stress level of the IEEE 14-bus network, and is discussed further in Section 2.6.2.

2.6. Discussion and Limitations

2.6.1. Technical interpretation

The results of Section 2.5 converge on a consistent picture of the proposed methodology. The SOCP relaxation of the ORPD (Section 2.2.2) is exact under inductive loading but can become loose under capacitive loading, where it is exploited as an artificial reactive sink. This is not a generic property of the relaxation but a scenario-dependent one: in this test case, it appears only under the Demand Sufficiency trajectory. The γ^{line} regularisation term, introduced to counter this effect, behaves as intended: it suppresses the fictitious reactive absorption without requiring a fully non-convex reformulation, at the cost of a moderate over-investment in inductive compensation ($\mathcal{A}^{d,ind}$) once pushed beyond its critical range. That the regularisation only needs to be active on the trajectory where it actually improves the AC-evaluated cost, rather than uniformly across all trajectories, indicates that the SOCP relaxation gap is a localized, identifiable phenomenon rather than a systemic weakness of the formulation.

The choice between a nonlinear solver and a convex relaxation depends on the level of decision. In operation, the computed point is applied to the network and can be corrected in real time. Feasibility then comes first, and an interior-point solver of the AC formulation is appropriate. In planning, the operational model is used to evaluate binary investment decisions. With these variables, the AC formulation of the RPP is a non-convex Mixed-Integer Nonlinear Program (MINLP). Its branch-and-bound relies on nonlinear subproblems that are solved locally (Section 2.4.3), so it provides no valid bound and no optimality guarantee. The SOCP formulation gives a MISOCP, which is solved to the global optimum of the relaxed problem within the solver tolerance. Its optimal value is unique and does not depend on the starting point (Section 2.1.4). The AC formulation does not have this property. In Table 2.5, an arbitrarily restricted configuration gives a lower cost than the full-space AC solution. The relaxation also has limits. It loses exactness in the capacitive conditions toward which the energy transition moves the network (Section 2.1.5). With $\gamma^{line} > 0$, its optimal value is a lower bound only for the penalised problem, and not for the original AC cost. Its reliability was also established by comparison with the AC solver. The two approaches are therefore complementary and not opposed.

The WBA results reinforce this picture from a different angle. Among the four indicators considered, only Q^{slack} and VSI capture the buses that actually matter for investment, while V^{dev} and the L-index rank buses that are not the most relevant for siting reactive compensation in this test case. This suggests that indicators directly tied to the reactive power balance (Q^{slack}) or to its sensitivity

to compensation (VSI) transfer more usefully into a filtering role than indicators that capture a related but distinct physical property. This property is the voltage proximity to nominal for V^{dev} , and the proximity to voltage collapse for the L-index. Finally, the trajectory comparison shows that the model’s investment and operational choices track the underlying physical configuration of each trajectory rather than a fixed pattern. This gives confidence that the cost structure introduced in Section 2.2.4 is discriminating enough to produce differentiated, trajectory-specific decisions rather than a single generic solution.

2.6.2. Limitations

Several limitations qualify the scope of these results.

1. The AC reference itself is not fully reliable: Table 2.5 shows that Set 2, an arbitrarily restricted two-bus configuration disjoint from the buses identified by the full-space multi-start procedure, reaches a slightly lower AC cost than that procedure (-0.76%). Although the margin is small, it shows that even the multi-start AC solve does not systematically recover the global optimum, so AC-based results throughout this chapter should be read as a strong reference rather than a guaranteed ground truth.
2. The activation of the γ^{line} regularisation is currently decided a posteriori, by observing which trajectory benefits from it. An a priori criterion, based on scenario characteristics available before solving the SOCP formulation, would be needed to apply the methodology without first solving the more expensive AC benchmark.
3. The test case is limited to the IEEE 14-bus network with a single investment period: the scalability of the joint WBA-RPP pipeline to larger transmission networks, and its behaviour under a multi-period investment horizon, remain untested. Approaches that reduce the operational constraint space by identifying binding contingencies and conditions have been shown to scale to large networks [71]. The decision-space reduction pursued here is orthogonal to that constraint-space reduction. Combining the two is therefore a natural direction for extending the pipeline to realistic network sizes.
4. In practice, the continuous capacity obtained by the RPP must be rounded up to the next available standard size. On the test case, where at most two devices are installed, this rounding has a limited effect. On a larger network with many investments, however, the rounding deviations accumulate and can move the applied solution significantly away from the computed optimum. If the continuous sizing is kept as a working assumption to

preserve tractability, the pipeline should therefore be completed by a post-optimisation discretisation step. In this step, each continuous capacity is converted into an available standard size, and the resulting investment plan is evaluated again with the final ORPD run.

5. The cost coefficients of Section 2.4 are calibrated to reflect plausible economic signals for a European TSO. However, they are not validated against a specific utility's actual reactive power tariffs or investment budgets. The resulting investment decisions should be interpreted as illustrative of the methodology rather than as operational recommendations. These coefficients determine which solution the formulation returns, but they are not the only distance to a societal optimum: the problem is solved over transmission assets only, the distribution network enters it through its reactive exchange at the interface, and the active power schedule is taken as given. Supplied with coefficients consistent with the societal cost, the formulation would therefore return the least-cost solution within this scope, and not a societal optimum over the whole system.
6. Although the formulation embeds N-1 security constraints throughout, no result isolates their individual impact. On the IEEE 14-bus test case, the network is not sufficiently stressed for individual contingencies to bind differently across trajectories. As a result, the reported investment and operational costs are not visibly sensitive to which contingency is considered. A dedicated sensitivity analysis of the security dimension, on a larger or more heavily loaded network where individual contingencies bind unevenly, is left for future work. The retained contingency set is also not exhaustive, and the investment decisions obtained have not been verified against the complete N-1 set, beyond the sensitivity analysis of Section 2.5.5. The severity-based screening presented in Section 2.1.7 remains defensible, since most contingencies of a set are non-binding at the optimum, but it is not a guarantee, and an a posteriori verification of the obtained decisions on the full N-1 set is a natural extension.
7. The dispatch stage takes the commitment status of the generating units as given and does not represent start-up and shut-down decisions. Representing these decisions could change the investment results substantially, because the set of units participating in voltage control would then vary across operating conditions. A synchronous unit that is decoupled from the network provides neither reactive power nor a voltage setpoint. As a result, the compensation required at the surrounding buses would be different.

2.7. Chapter Conclusion

This chapter proposed an original scenario-based, security-constrained MISOCP formulation for reactive power planning that explicitly captures the reactive power dispatch costs incurred by a TSO across multiple operating conditions and contingencies. The formulation jointly optimises investment and operational decisions, treating shunt compensation devices ($\mathcal{A}^{d,cap}$ and $\mathcal{A}^{d,ind}$) as discrete on/off assets and SVCs/STATCOMs ($\mathcal{A}^c$) as continuously controllable. To address the inexactness of the SOCP relaxation under capacitive scenarios, a regularisation term acting selectively on branches incident to capacitive loads was introduced and calibrated against an AC reference on a reduced operational scope. A WBA procedure based on four performance indicators (V^{dev} , L-index, VSI, and Q^{slack}) was developed as a pre-processing stage to identify the most critical buses for VAR investment, reducing the combinatorial complexity of the planning problem.

The numerical evaluation on the IEEE 14-bus test case, presented in Section 2.5, validated the proposed methodology along several dimensions. The SOCP relaxation was shown to remain exact on inductive trajectories and to require regularisation only under capacitive conditions. The WBA identified Q^{slack} as the most efficient indicator, achieving sub-1% cost deviation from the full-space solution with only two candidate buses, while V^{dev} and the L-index required substantially more buses to reach comparable quality. The trajectory comparison highlighted how the model adapts its planning decisions to the physical configuration of each future scenario. The Demand Sufficiency trajectory was the only one requiring investment in inductive compensation, reflecting its dominantly capacitive nature. In contrast, the High Demand trajectory exploited the reactive capability of its wind capacity to limit compensation needs. Finally, the sensitivity analysis showed that 20 operating conditions are sufficient to capture the operational behaviour of the network within the considered test case, beyond which the computational burden grows exponentially without improving solution quality.

CHAPTER 3.

VARLearn: Reactive Power Forecasting at TSO/DSO Interfaces

The methodology and results presented in this chapter build on work first reported in [13], and further developed and validated in [14].

3.1. Introduction

This chapter addresses the short-term forecasting, at the quarter-hourly horizon, of active and reactive power exchanges at T-D interfaces from the TSO perspective. The remainder of this section proceeds as follows. Section 3.1.1 introduces the challenge of characterising active and reactive power exchanges at T-D interfaces from the TSO perspective. Section 3.1.2 reviews existing model-based and data-driven approaches and identifies the gap that neither family closes on its own. Section 3.1.3 synthesises this gap and situates the contributions of the chapter.

3.1.1. Power exchanges at T-D interfaces: context and challenges

As discussed in Chapter 1, the ongoing energy transition is profoundly altering the reactive power landscape in HV transmission networks. The growing penetration of RES and the replacement of OLs with UCs at the distribution level reshape reactive power behaviour at T-D interfaces. At the same time, the progressive phase-out of conventional synchronous generators, connected at the

transmission level, reduces the reactive resources available to the TSO to manage these exchanges. These physical phenomena are not recalled in detail here but constitute the central motivation for the present chapter.

Chapter 2 addresses reactive power management from a strategic planning perspective: given a set of operating scenarios, the VARPlan framework determines where and how much to invest in compensation assets to ensure voltage security across the transmission network. Central to this formulation is the characterisation of reactive power demand at each network bus, including the T-D interfaces, where power flows originate from distribution networks that the TSO does not model directly, a modelling boundary that motivates the remainder of this chapter. Characterising active and reactive demand at these interfaces is a problem that has been identified in the literature for some time [4]. Because reactive power is a local quantity, this system-level view is naturally complemented by an interface-level perspective, focusing on where these exchanges actually take place. This chapter therefore turns to the forecasting of active and reactive power exchanges at T-D interfaces, from the TSO perspective and under its information constraints.

In Europe, and more particularly in Belgium, TSOs and DSOs are clearly separated regulated entities [21]. This institutional separation introduces a fundamental information asymmetry: the TSO has no direct access to the distribution network topology or model parameters. This inaccessibility, which extends to line parameters and network topology, reflects the commercially sensitive and proprietary nature of distribution network data, an established constraint that motivates approaches grounded in historical measurements [98]. As a consequence, it cannot rely on explicit network modelling to forecast reactive power flows at T-D interfaces, and must instead infer these quantities from aggregate measurements accessible at the interface itself.

On the DSO side, reactive power exchanges are subject to contractual constraints imposed by the TSO. Deviations beyond prescribed reactive power bounds may result in financial penalties [21]. This regulatory framework reinforces the importance of accurate forecasting for both parties.

The prediction problem is therefore characterised by two compounding difficulties: the physical complexity of the underlying phenomena and the structural opacity imposed by the TSO/DSO information barrier. The lack of access to distribution network topology prevents the direct application of model-based forecasting approaches, while purely data-driven methods applied to aggregate interface signals may lack physical consistency. These two families will be assessed, in Section 3.1.2, against two properties, physical grounding and independence from the distribution network model. The former reflects reliance on the governing

power flow equations rather than on statistical fitting alone; the latter, independence from the network’s topology and parameters. As shown there, no existing approach satisfies both simultaneously. This dual challenge motivates the hybrid approach developed in the present chapter.

3.1.2. Existing approaches and their limitations

A substantial body of work has characterised the aggregate active and reactive flexibility that a distribution network can offer at its transmission interface, quantifying how much a DSO can modulate power exchanges in response to transmission-level signals. The object of this characterisation, the region of P/Q operating points attainable at the interface under the internal constraints of the distribution network, is known as the flexibility area or feasible operating region [99]. Estimating this region has given rise to distinct families of methods, including geometric constructions, sampling-based approaches, and optimisation-based formulations, a diversity that attests to a mature and active field [100].

These works address a question of capacity: what is the set of P/Q exchanges that a distribution network can technically sustain at its transmission interface. The present chapter addresses a distinct question: which P/Q exchange will actually occur for a given load state, that is, a single operating point rather than an envelope of attainable ones. A feasibility region delimits what is possible; it does not predict which point within that region will be realised under a specific operating condition [100]. The two perspectives are complementary rather than competing: the flexibility area and the realised operating point respond to different operational needs and inform different decisions.

Despite the importance of this topic and the scientific challenges it represents, the literature on reactive power prediction at T-D interfaces remains limited. Existing approaches can be classified into two main categories.

The first category comprises **model-based approaches**, which rely on a representation of the distribution network, calibrated from historical measurements, to project future reactive power demand trends and characterise active and reactive power behaviour at T-D interfaces [4]. While physically grounded, these methods are fundamentally constrained by the information asymmetry described in Section 3.1.1: the TSO does not have access to an up-to-date and complete model of the distribution network. In practice, the distribution topology changes continuously due to network reconfigurations, new connections, and the integration of RES. Maintaining a reliable equivalent model under these conditions is not tractable from the TSO perspective.

The second category relies on **data-driven and machine learning techniques**, such as clustering and regression methods, applied directly to aggregate interface measurements [101, 102]. These methods can produce satisfactory predictive performance when sufficient representative data is available. However, they are generally treated as black boxes: by operating solely on statistical patterns in the data, they bypass the underlying physical structure of the power system. This lack of physical interpretability raises concerns regarding reliability under operating conditions not well represented in the training data, a particularly relevant limitation given the rapid evolution of distribution network behaviour discussed above.

These two limitations, model inaccessibility and physical opacity, define the gap that motivates the hybrid approach proposed in this chapter. The two families thus occupy the two unsatisfactory positions in turn; the position that would combine independence from the distribution network model with physics as an explicit driver of the computation remains unoccupied. Whether the existing methodological building blocks, namely network equivalents, identification from measurements, and learning constrained by physical equations, would make it possible to reach that position is what the remainder of this chapter sets out to establish.

Identifying an equivalent electrical model from successive state measurements constitutes an inverse problem studied in its own right, formalised under the name of inverse power flow. This problem consists in inferring the parameters of a network, such as its admittance matrix or line parameters, from quantities measured at a set of nodes. When certain nodes are not observed, what is identified is not the full network but a reduced equivalent as seen from the accessible nodes [103]. We designate by ILF the formulation that, in this spirit, identifies the parameters of a reduced equivalent of the distribution network capable of reproducing a measured state at the interface. Unlike formulations based on voltage and current phasors, the present approach relies on pairs of aggregate load conditions and measured P/Q exchanges, a variant in terms of the input quantities rather than in terms of the underlying inverse problem principle.

Parameter and topology identification from network data constitutes an active research domain: these works aim at recovering the actual structure of the network, its topology and admittances [104]. The TSO, however, neither can nor seeks to reconstruct that structure; the present chapter therefore assumes the topology to be unknown and settles for a reduced equivalent that remains operationally exploitable. This family is thus physics-driven and model-independent in its inputs, but its objective, recovering the true network, differs from the one pursued here.

Reducing a network to an equivalent of lower dimension is a long-established technique, of which the Ward equivalent constitutes the founding formulation. These methods proceed, however, by reducing the full network model: they presuppose knowledge of the network to be reduced [105]. They are therefore inapplicable from the TSO’s standpoint, which does not hold the distribution model required for such a reduction. This distinguishes identification from measurements, introduced above as a means of inferring a reduced or full network representation from data, from model-based reduction: two routes toward an equivalent, only one of which is accessible to the operator.

Combining physical knowledge with machine learning is an active research direction, initiated by neural network formulations that embed the governing physical equations directly into the learning process [106]. In power systems, this paradigm takes several forms, including physical constraints embedded in the loss function, physics-guided initialisation, architecture design, and hybrid models [107]. In these approaches, physics intervenes as a constraint imposed on a learned model that itself produces the output. The present approach proceeds differently: learning applies only to the parameters of an explicit physical model, a network equivalent, whose resolution, rather than the learned model, delivers the sought exchange. Physical consistency of the output follows by construction, rather than being encouraged through a penalty term.

Each of these approaches provides one of the required properties, but none of them provides all of them. Identification from measurements provides a physics-grounded equivalent but targets the recovery of the actual network. Classical reduction constructs the equivalent but requires the full network model. Learning constrained by physical equations retains physics as a constraint rather than as a driver of the computation. Their combination, learning the parameters of an explicit reduced equivalent from the sole quantities accessible to the TSO, then solving that equivalent, reaches the position that remained unoccupied: independence from the distribution network model with physics as the explicit driver of the computation.

3.1.3. Research gaps and contributions

The two limitations identified above point to a common gap: no existing approach simultaneously provides physical grounding and independence from the distribution network model. Model-based methods require topology data that the TSO cannot access; purely data-driven methods bypass the physical structure of the power system and offer limited reliability guarantees outside their training distribution.

To address this gap, this chapter proposes an hybrid approach combining an ILF formulation with ML regression. In an offline phase, historical records of aggregate load conditions and measured P/Q exchanges at the interfaces are used by the ILF formulation to identify, for each observed state, the equivalent resistance and reactance of a reduced network model that reproduces that state. A regression model then learns the mapping from aggregate load conditions to these equivalent parameters. In operational use, the learned model predicts the equivalent parameters from an estimated aggregate load, and the resolution of a load flow on that equivalent delivers the active and reactive exchange at the interface. Condensing the distribution network's interface behaviour into a small number of equivalent parameters removes any dependence on its internal topology, which the TSO does not hold. Renewable resources, in particular PV generation, are accounted for through the net active load. The following contributions are presented:

- a hybrid model combining ILF-based physical equations with ML regression to forecast active and reactive power exchanges at T-D interfaces, without requiring access to the distribution network topology;
- incorporation of RES into the equivalent model and comparative assessment of four ML algorithms for equivalent parameter prediction;
- application-based and robustness evaluation, assessing the model under TSO operational scenarios and under input uncertainty.

3.2. Methodology

The proposed methodology addresses the problem of forecasting active and reactive power exchanges at T-D interfaces from the TSO perspective, without requiring access to the distribution network topology or its model parameters. The approach combines a physical equivalent model of the distribution network, derived from interface measurements and the aggregate distribution network load, with supervised ML regression. It is structured in three successive steps, illustrated in Figure 3.1.

In the first step, a synthetic dataset is constructed from a representative distribution network, and an ILF formulation is applied at each operating point to derive a physical equivalent model of that network from the interface measurements and the aggregate distribution network load. In the second step, supervised ML regression is used to learn how the parameters of this equivalent model vary with the operating conditions. In the third step, the learned parameters are used

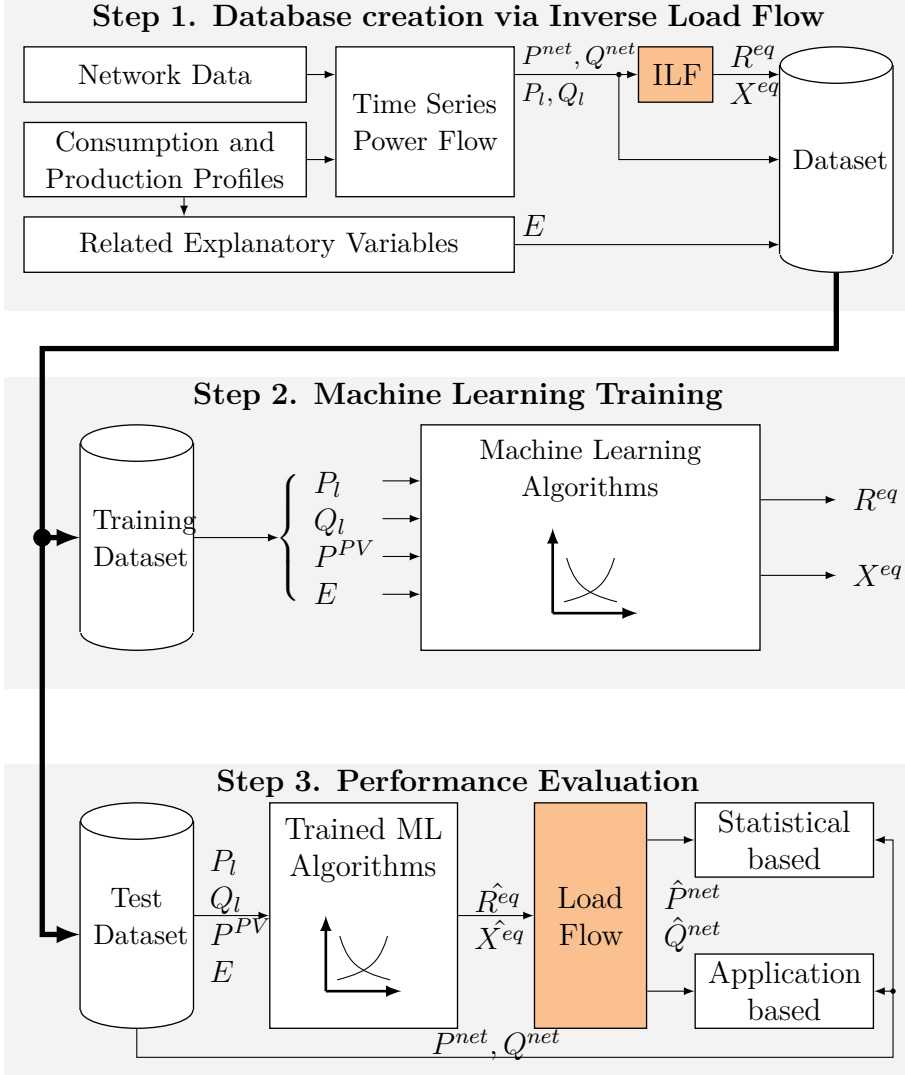


Figure 3.1.: Flowchart of the proposed three-step methodology for forecasting power exchanges at T-D interfaces.

to reconstruct the forecasted interface power exchanges, whose accuracy is then assessed from complementary perspectives.

3.2.1. Step 1 — Database creation via Inverse Load Flow

The ILF formulation aims to represent the entire distribution network connected to the transmission network at an HV/Medium Voltage (MV) substation by a

single equivalent line model. This equivalent model is parameterised by a resistance R^{eq} and a reactance X^{eq} , derived from the power flow equations applied to four quantities available to the TSO: the active and reactive power exchanges at the T-D interface (P^{net} and Q^{net} , measured directly at the substation) and the aggregate active and reactive load demand of the distribution network (P_l and Q_l , estimated from standard load profiles).

The derivation begins from the general power flow equations in hybrid form [108]:

$$\sum_{m \in \mathcal{M}_{(n)}} P_{n,m} = \sum_{g \in \mathcal{G}_{(n)}} P_g - \sum_{l \in \mathcal{L}_{(n)}} P_l \quad \forall n \in \mathcal{N} \quad (3.1)$$

$$\sum_{m \in \mathcal{M}_{(n)}} Q_{n,m} = \sum_{g \in \mathcal{G}_{(n)}} Q_g - \sum_{l \in \mathcal{L}_{(n)}} Q_l \quad \forall n \in \mathcal{N} \quad (3.2)$$

$$P_{n,m} = g_{n,m} V_n^2 - g_{n,m} V_n V_m \cos(\delta_n - \delta_m) - b_{n,m} V_n V_m \sin(\delta_n - \delta_m) \quad \forall n, m \in \mathcal{N} \quad (3.3)$$

$$Q_{n,m} = -b_{n,m} V_n^2 + b_{n,m} V_n V_m \cos(\delta_n - \delta_m) - g_{n,m} V_n V_m \sin(\delta_n - \delta_m) \quad \forall n, m \in \mathcal{N} \quad (3.4)$$

where $P_{n,m}$ and $Q_{n,m}$ are the active and reactive power flows on line $\langle n, m \rangle$; $g_{n,m}$ and $b_{n,m}$ are its conductance and susceptance; V_n and δ_n are the voltage magnitude and phase angle at bus n ; and $\mathcal{N}$, $\mathcal{M}_{(n)}$, $\mathcal{G}_{(n)}$, $\mathcal{L}_{(n)}$ denote the sets of buses, buses adjacent to n , generators at n , and loads at n , respectively.

The equivalent 1-line model reduces this general formulation to a two-bus system: bus 1 represents the HV/MV substation (the T-D interface) and bus 2 aggregates all loads of the distribution network, as illustrated in Figure 3.2.

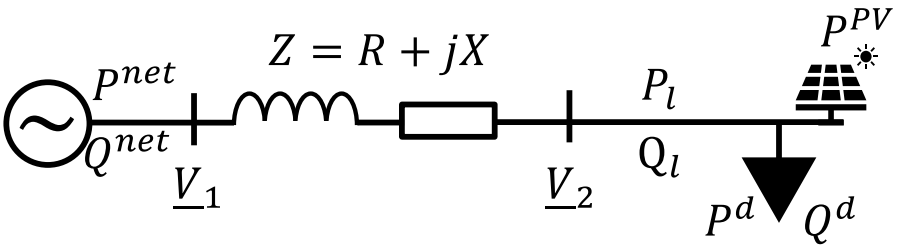


Figure 3.2.: Equivalent 1-line network used in the ILF formulation. Bus 1 represents the T-D interface (slack bus); bus 2 aggregates all distribution-level loads and PV generation.

Designating bus 1 as the slack bus with $V_1 = 1$ p.u. and $\delta_1 = 0$, the system

simplifies to four equations in four unknowns (g , b , V_2 , δ_2):

$$P^{\text{net}} = g - gV_2 \cos \delta_2 + bV_2 \sin \delta_2 \quad (3.5)$$

$$Q^{\text{net}} = -b + bV_2 \cos \delta_2 + gV_2 \sin \delta_2 \quad (3.6)$$

$$P_l = -gV_2^2 + gV_2 \cos \delta_2 + bV_2 \sin \delta_2 \quad (3.7)$$

$$Q_l = bV_2^2 - bV_2 \cos \delta_2 + gV_2 \sin \delta_2 \quad (3.8)$$

Given the four inputs P^{net} , Q^{net} , P_l , and Q_l , the nonlinear system (3.5)–(3.8) (four equations in four unknowns: g , b , V_2 , and δ_2) is solved numerically at each time step using a standard nonlinear equation solver. Since this system is nonlinear, the existence and the uniqueness of a physically meaningful solution are not guaranteed a priori, and several mathematical solutions can satisfy the four equations while only one of them corresponds to a realistic operating point. For this reason, the unknowns are initialised at values corresponding to normal operating conditions, so that the solver converges towards the solution which is physically meaningful. The obtained solution is then verified by a direct load flow computation performed with the identified parameters and the corresponding load, in order to check that the interface exchanges P^{net} and Q^{net} are correctly retrieved. The solution gives g and b , from which the equivalent impedance is obtained via:

$$R^{\text{eq}} + jX^{\text{eq}} = \frac{1}{g + jb} \quad (3.9)$$

This procedure is applied independently to each of the 30 000+ quarter-hourly time steps in the dataset, with negligible computational cost relative to the subsequent ML training.

The physical interpretation of these parameters is suggestive rather than validated. R^{eq} may be read as a reflection of aggregate active power losses within the distribution network, while X^{eq} can be interpreted as an image of its net reactive behaviour, shaped by the mix of inductive loads, UCs, and distributed generation units. These readings remain interpretive and are not empirically validated in the present work. The value of the equivalent representation is instead established on learning grounds: by providing a smoother and more structured regression target than a direct mapping from features to interface power flows, it renders the learning problem more tractable.

Both parameters vary at each time step: they are not fixed network constants but operating-point-dependent quantities that encode the instantaneous electrical state of the distribution network as seen from the T-D interface. This differs

from the classical Ward equivalent, in which the equivalent admittances are obtained by eliminating the nodes of the external subnetwork from its topology and its parameters [105]. In the Ward equivalent, these admittances are network constants, and the operating point enters only through the equivalent injections placed at the boundary nodes.

The parameters identified here are of a different nature, because they are re-identified at each time step and absorb both the network structure and the operating point into a single time-varying impedance. There is therefore no reason for them to coincide with the Ward equivalent of the same network, and they should not be read as an estimate of it. The distinction made here concerns the nature of the result. It complements the one made in Section 3.1.2, which concerns the information required to compute the equivalent.

RES, and PV generation in particular, are incorporated into the model through the aggregate net load definition:

$$P_l = P^d - P^{\text{PV}} \qquad Q_l = Q^d \qquad (3.10)$$

where P^d and Q^d are the total active and reactive demand, and P^{PV} is the aggregate PV production. The reactive contribution of PV inverters is neglected under the standard assumption of unity power factor operation.

At the end of this step, the synthetic dataset is made of, for each time step, the interface power exchanges (P^{net} , Q^{net}), the aggregate distribution network load (P_l , Q_l), and the corresponding equivalent impedance (R^{eq} , X^{eq}). The resulting dataset is then partitioned into a training set (90 % of samples), used to fit the ML models of Step 2, and a test set (10 %), held out to evaluate their performance on operating points not seen during training.

3.2.2. Step 2 — Machine learning algorithms for equivalent impedance prediction

The second step formulates the prediction of R^{eq} and X^{eq} as a supervised ML regression problem. Rather than mapping aggregate loads directly to interface power exchanges, which requires the predictor to implicitly replicate the entire power flow computation within the distribution network, the proposed approach restricts the learning task to the identification of the equivalent impedance parameters. The load flow reconstruction is then handled analytically in post-processing, providing a physically consistent route from predictions to power exchange estimates.

Feature vector and dataset structure

The model takes as input seven features, organised into three categories:

- *Load behaviour*: aggregate active load P_l , aggregate reactive load Q_l , and aggregate PV production P^{PV} ;
- *Weather data*: cloudiness and temperature, derived from Synthetic Load Profiles (SLPs) provided by the Belgian system operator Synergrid [109];
- *Calendar data*: hour of the day and day of the week.

and predicts the equivalent impedance parameters R^{eq} and X^{eq} derived in Step 1. Let $\mathbf{X} \in \mathbb{R}^7$ denote the input feature vector and $\mathbf{y} = (R^{\text{eq}}, X^{\text{eq}}) \in \mathbb{R}^2$ the prediction target.

The training dataset is structured as a set of labelled pairs:

$$\mathcal{D} = \{(\mathbf{X}_1, \mathbf{y}_1), (\mathbf{X}_2, \mathbf{y}_2), \dots, (\mathbf{X}_m, \mathbf{y}_m)\} \quad (3.11)$$

where m denotes the number of samples in the training set. The training objective is empirical risk minimisation: the learned function $\hat{\mathbf{y}} = f(\mathbf{X})$ is fitted by minimising the discrepancy between each predicted output $\hat{\mathbf{y}}_i$ and the corresponding true target $\mathbf{y}_i$ over the training set, so that it generalises to unseen samples from the test set. All features are linearly scaled prior to training to ensure comparable magnitude across input dimensions.

The term forecasting warrants clarification, as the predictive character of a model can arise from two distinct sources.

1. Explicit modelling of temporal dynamics, as in time series models, autoregressive representations, or architectures that propagate information from one instant to the next.
2. Propagation of inputs that are themselves forecasted through an instantaneous mapping: the model carries no internal representation of time but transforms a given input into a corresponding output.

The approach proposed here belongs to the second category. The feature vector $\mathbf{X}$ is available prior to the target instant. The learned mapping relates the features of a given operating point to its equivalent impedances at that same instant, without temporal offset or autoregression on past values of the target. The mapping itself is instantaneous and memoryless. Its predictive character derives from the fact that the aggregate load and meteorological quantities fed into it are, in operational use, forecasts produced upstream. A forecast for a subsequent

operating point is then obtained by evaluating this mapping on the predicted feature vector of that point.

This mechanism distinguishes the approach from the naive persistence benchmark, which belongs to the first category: it extrapolates the target from its most recent observed realisation rather than from exogenous input features. The absence of internal temporal modelling is therefore not a limitation but a deliberate choice: the physical structure of the equivalent carries the information necessary to reproduce the interface exchange, rendering an autoregressive representation of the output unnecessary.

The availability of these features at forecast time differs across categories. Calendar features are deterministic and known by construction. Meteorological inputs rely on short-range numerical weather predictions, which are available and sufficiently reliable at the quarter-hourly horizon. Aggregate load and production quantities (P_l , Q_l , P^{PV}) are, in real-world operation, among the quantities most amenable to short-term forecasting in power systems. In the present work, however, they are derived from SLPs and Synthetic Production Profiles (SPPs) in the absence of real measurements (Section 3.3). The operational forecasting argument above should not be taken to rest on these synthetic proxies.

Applied algorithms

Four supervised learning algorithms are applied, selected to cover distinct families of ML models: a non-parametric method, two tree-based ensemble methods, and a neural network. These families are commonly applied to similar regression and forecasting tasks in power systems [110, 111], which motivates their selection here. A detailed treatment of their formulation is not the focus of this chapter; the reader is referred to [112] for a comprehensive presentation of these methods, including of the regression trees underlying both Random Forest (RF) and Gradient Boosting (GB).

K-Nearest Neighbours (KNN) is a non-parametric, instance-based method with no explicit training phase. For a new input $\mathbf{X}^{\text{new}}$, it identifies the K training examples closest to it in feature space using the Euclidean distance [110, 113]:

$$d_j = \sqrt{\sum_{i=1}^7 |X_i^j - X_i^{\text{new}}|^2} \quad (3.12)$$

The predicted target is then taken as the average of the corresponding labels of the K nearest neighbours, where K is the model's main hyperparameter.

RF is a bagging ensemble that combines multiple regression trees, each trained on a bootstrap sample of the data with a randomly selected subset of features at each split [114, 111]. The final prediction is obtained by averaging the individual tree outputs. The randomization introduced at both the sample and feature levels substantially reduces variance compared to a single regression tree. The ensemble is primarily governed by the number of trees and the maximum depth allowed for each tree.

GB is a boosting ensemble that builds regression trees sequentially, each fitted to the negative gradient of the loss function with respect to the current ensemble prediction [115, 116]. Unlike RF, which constructs trees independently in parallel, GB corrects the errors of preceding models iteratively, achieving high predictive accuracy at the cost of greater sensitivity to hyperparameter choices. The learning rate, the maximum tree depth, and the number of boosting iterations are its key hyperparameters.

Multilayer Perceptron (MLP) is a feedforward neural network comprising an input layer, one or more hidden layers, and an output layer. Each neuron computes a nonlinear activation function applied to a weighted sum of its inputs, allowing the network to capture complex relationships in the data. Model parameters are learned through backpropagation [117]. The number and size of the hidden layers, together with the choice of activation function, define its architecture. While MLP offers broad representational capacity, it requires careful regularisation and architecture tuning to avoid overfitting.

For all four algorithms, hyperparameters are optimised using a Bayesian optimisation procedure with a dynamic racing mechanism [118], which efficiently explores the configuration space while discarding underperforming candidates. The resulting optimal architectures are reported in Section 3.3.3.

3.2.3. Step 3 — Performance evaluation framework

The proposed approach is evaluated from two complementary perspectives: a statistical assessment of regression accuracy and an application-based assessment of the operational consequences of prediction errors for the TSO.

Statistical evaluation

The statistical performance of the regression is quantified using the Normalised Root Mean Square Error (NRMSE), defined as:

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (3.13)$$

$$\text{NRMSE} = \frac{\text{RMSE}}{\sigma_y} \quad (3.14)$$

where n is the number of test set samples, y_i and $\hat{y}_i$ are the observed and predicted values, and σ_y denotes the standard deviation of the target variable over the test set. Because σ_y is computed separately for each load configuration and for each target variable, dividing the RMSE by σ_y removes the dependence on the natural scale of that specific configuration and target. The resulting NRMSE values can therefore be directly compared, and averaged, across the five load configurations and across the two targets R^{eq} and X^{eq} .

The NRMSE is computed for both the proposed ILF-based approach and a direct prediction benchmark (Direct Prediction (DP)), which follows the same ML framework but maps aggregate load demand directly to interface power exchanges without the ILF step. This comparison isolates the contribution of the physical equivalent model to prediction accuracy.

Application-based evaluation

The application-based evaluation places the forecasting model within a TSO short-term operational framework, where generators must be dispatched one quarter-hourly time step ahead based on predicted interface power exchanges [119]. The procedure proceeds in two steps, as illustrated in Figure 3.3.

In the first step, the predicted power exchanges $\hat{P}^{\text{net}}$ and $\hat{Q}^{\text{net}}$ at each T-D interface are treated as fixed loads in a TSO-level OPF, which minimises total generation costs:

$$f_1(\Theta) = \sum_{g \in \mathcal{G}} \gamma_{g,1} P_g \quad (3.15)$$

where $\gamma_{g,1}$ [€/MWh] is the marginal generation cost of generator g . The OPF solution determines the committed generator setpoints ($P_g^{\text{sp}}, Q_g^{\text{sp}}$) for the next time step.

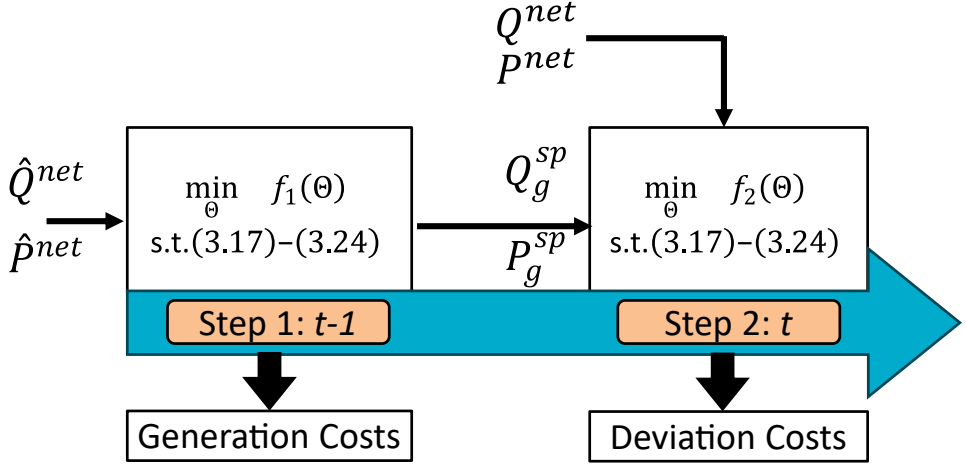


Figure 3.3.: Two-step application-based evaluation methodology from the TSO perspective. Step 1 optimises generator setpoints based on predicted interface power exchanges; Step 2 minimises re-dispatch costs against the actual realised exchanges.

In the second step, actual power exchanges become observable and differ from the predictions. A curative OPF is then run to minimise the deviation of each generator from its committed setpoint:

$$f_2(\Theta) = \sum_{g \in \mathcal{G}} \gamma_{g,2} |P_g - P_g^{sp}| \quad (3.16)$$

The curative cost coefficient $\gamma_{g,2}$ is set to five times $\gamma_{g,1}$, reflecting the higher cost of emergency re-dispatch compared to planned dispatch [120]. Both OPF

problems are subject to the same operational constraints:

$$\sum_{m \in \mathcal{M}_{(n)}} P_{n,m} = \sum_{g \in \mathcal{G}_{(n)}} P_g - \sum_{l \in \mathcal{L}_{(n)}} P_l \quad \forall n \in \mathcal{N} \quad (3.17)$$

$$\sum_{m \in \mathcal{M}_{(n)}} Q_{n,m} = \sum_{g \in \mathcal{G}_{(n)}} Q_g - \sum_{l \in \mathcal{L}_{(n)}} Q_l \quad \forall n \in \mathcal{N} \quad (3.18)$$

$$P_{n,m} = g_{n,m} V_n^2 - g_{n,m} V_n V_m \cos(\delta_n - \delta_m) - b_{n,m} V_n V_m \sin(\delta_n - \delta_m) \\ \forall n, m \in \mathcal{N} \quad (3.19)$$

$$Q_{n,m} = -b_{n,m} V_n^2 + b_{n,m} V_n V_m \cos(\delta_n - \delta_m) - g_{n,m} V_n V_m \sin(\delta_n - \delta_m) \\ \forall n, m \in \mathcal{N} \quad (3.20)$$

$$P_g^{\min} \leq P_g \leq P_g^{\max} \quad \forall g \in \mathcal{G} \quad (3.21)$$

$$Q_g^{\min} \leq Q_g \leq Q_g^{\max} \quad \forall g \in \mathcal{G} \quad (3.22)$$

$$V_n^{\min} \leq V_n \leq V_n^{\max} \quad \forall n \in \mathcal{N} \quad (3.23)$$

$$P_{n,m}^2 + Q_{n,m}^2 \leq (S_{n,m}^{\max})^2 \quad \forall n, m \in \mathcal{N} \quad (3.24)$$

with the slack bus fixed at $V_1 = 1.0 \angle 0$.

The total operational cost of each forecasting approach (generation cost plus curative re-dispatch cost) is aggregated over the entire test set. It is compared across three methods: the proposed ILF-based approach, the DP benchmark, and a naive persistence predictor (which uses the power exchange measured at $t-1$ as the forecast for t). A ground truth baseline, computed using the actual interface power exchanges in both OPF steps, serves as the reference lower bound on total cost.

3.3. Test Case and Simulation Parameters

This section describes the simulation setup used to validate the proposed methodology. Since real TSO measurements of power exchanges at T-D interfaces are not publicly available, the synthetic dataset introduced in Section 3.2.1 is built from a modelled distribution network; its concrete construction is detailed in the following subsections. The statistical evaluation is conducted on this same dataset, while the application-based evaluation is carried out on a HV transmission network benchmark. The two networks are described in the following subsections, followed by the data parameters and the benchmarks used for comparative assessment.

3.3.1. Distribution network

The proposed ILF-based approach is designed for the TSO, which can perform the ILF step using interface power measurements and aggregate load estimates to build a dataset linking operating conditions to equivalent impedance parameters. To compensate for the unavailability of real measurements, this dataset is constructed by modelling a MV distribution network, applying load profiles to it, and running time-series load flow computations to determine the power exchanges at the T-D interface for each time step.

The selected network is the Oberrhein MV network [2], a generic AC distribution network operating at 20 kV and constructed from publicly available data within the Pandapower framework. This network has been previously used as a benchmark in distribution network studies [121, 122]. The version adopted here comprises 69 nodes and 68 lines, connected to the transmission level through a step-down 110/20 kV transformer, as illustrated in Figure 3.4.

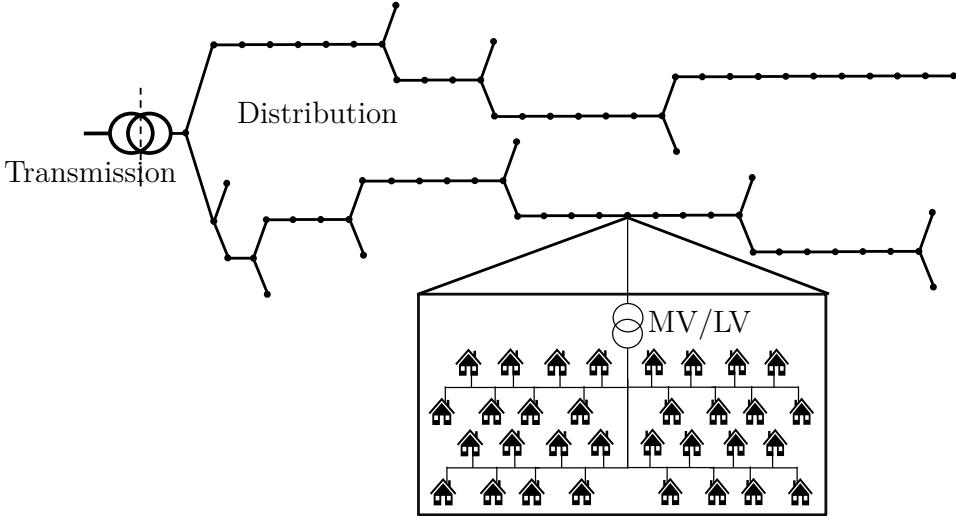


Figure 3.4.: Modified Oberrhein MV distribution network (69 nodes, 20 kV) used as the test distribution network.

To simulate realistic Low Voltage (LV) load behaviour, SLPs and SPPs provided by the Belgian grid operator Synergrid [109] are incorporated into the model. These profiles emulate aggregate LV network characteristics as loads within the MV network, parameterised by four quantities:

- C_{nom} : nominal load capacity [MVA];
- I_{PV} : total installed PV capacity [MWp];

- PF : power factor;
- NoL : nature of load (inductive or capacitive).

To cover a representative range of distribution network operating conditions, five distinct load configurations are defined, each corresponding to a different combination of the parameters above. Table 3.1 summarises the parameter values used for each configuration.

Table 3.1.: Parameters of the five distribution network load configurations.

Load	C_{nom} [MVA]	I_{PV} [MWp]	PF	NoL
L1	30	20.4	0.85	Inductive
L2	30	6.8	0.95	Capacitive
L3	15	0	0.90	Capacitive
L4	20	3.4	0.80	Inductive
L5	20	3.4	0.95	Inductive

For each configuration and each time step, a load flow computation is performed to determine P^{net} and Q^{net} at the T-D interface. The ILF formulation described in Section 3.2.1 is then applied to complete each data point with the corresponding R^{eq} and X^{eq} values. The five configurations are treated as five independent scenarios for the statistical evaluation, enabling an assessment of the model robustness across different distribution network operating conditions.

It is worth noting that the proposed methodology is not restricted to networks with a single load point at the T-D interface. When multiple distribution feeders are connected to the same transmission bus, the aggregate load powers P_l and Q_l are summed across all feeders, and the ILF is applied to the resulting aggregate quantities to obtain a single equivalent impedance for the bus. This demonstrates the generality of the approach with respect to the number of feeders served by a given T-D interface.

3.3.2. Transmission network

The application-based evaluation requires a transmission network in which the TSO OPF framework described in Section 3.2.3 can be implemented. The CIGRE HV benchmark network [123] is adopted for this purpose, with minor modifications to adapt it to the simulation context. All 220 kV buses are adjusted to 110 kV, and bus 9 is removed with its associated generator relocated to bus 7. The resulting network comprises three voltage levels: 20, 110, and 380 kV, as shown in Figure 3.5.

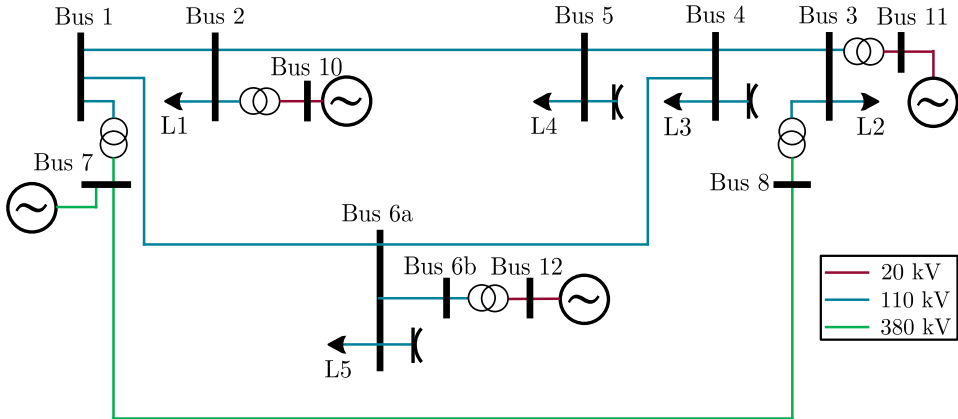


Figure 3.5.: Modified CIGRE HV benchmark network used for the application-based evaluation. The five loads (L1–L5) each represent a distinct distribution network connected at a T-D interface.

The five loads present in the transmission network (L1 to L5) each represent a distinct distribution network connected at a T-D interface. For consistency with the statistical evaluation, all five distribution networks share the Oberrhein topology but are assigned the different load configurations defined in Table 3.1. This setup allows the application-based evaluation to reflect the diversity of reactive power behaviour that a TSO would encounter across its portfolio of T-D interfaces.

The cost parameters used in the OPF are as follows. The generation cost coefficient $\gamma_{g,1}$ is set to 20 e/MWh for all conventional generators, consistent with values reported in the literature [124]. The curative re-dispatch cost $\gamma_{g,2}$ is set to 100 e/MWh, five times $\gamma_{g,1}$, to reflect the higher cost of emergency re-scheduling compared to planned dispatch [120].

3.3.3. Data parameters and benchmarks

This subsection reports the data and implementation parameters of the simulation setup: the construction and preprocessing of the dataset, the hyperparameter optimisation procedure, and the benchmarks used for comparative assessment.

Dataset construction and preprocessing

Using the SLPs and SPPs described above, the resulting dataset comprises over 30 000 data points for each load configuration, corresponding to one year of quarter-hourly measurements. It is divided into a training set comprising 90 %

of the samples and a test set comprising the remaining 10 %, used for training and evaluation respectively. All input features are linearly scaled to the $[0, 1]$ interval prior to training, ensuring comparable magnitude across dimensions with different physical units.

Hyperparameter optimisation

For all four ML algorithms and both the ILF-based approach and the DP benchmark, hyperparameters are optimised using the SMAC framework, which implements Bayesian optimisation with a dynamic racing mechanism [118]. This approach efficiently identifies high-performing configurations by concentrating evaluations on promising candidates while discarding underperforming ones. The optimal architectures obtained are reported in Table 3.2.

Table 3.2.: Optimal hyperparameter configurations for the DP and ILF-based approaches, obtained via Bayesian hyperparameter optimisation.

Algorithm	Hyperparameter	DP	ILF-based
GB	Learning rate	0.111	0.111
	Max depth	57	10
	Estimators	262	296
RF	Max depth	64	64
	Estimators	306	306
KNN	K	2	6
MLP	Layer sizes	(100, 100)	(50, 100, 50)
	Activation	ReLU	ReLU

Benchmarks

Two benchmarks are used for comparative assessment. The first is the **DP** method, which follows the same ML framework as the proposed approach but maps aggregate load demand directly to interface power exchanges, without the ILF step. This comparison isolates the specific contribution of the physical equivalent model to prediction accuracy. Both the ILF-based approach and DP are evaluated with the four ML algorithms listed above.

The second benchmark, used exclusively in the application-based evaluation, is the **Naive** predictor, which forecasts the power exchange at time step t by using

the actual measured value at time step $t - 1$. This simple persistence model serves as a lower-bound reference to establish the minimum performance level that any practical forecasting approach must exceed to be operationally relevant.

3.4. Results

The results are presented in three parts, corresponding to the three evaluation perspectives defined in Section 3.2.3. The statistical evaluation assesses regression accuracy across the five load configurations. The application-based evaluation quantifies the operational cost implications of prediction errors for the TSO. The uncertainty sensitivity analysis examines model robustness to imperfect knowledge of aggregate load inputs.

3.4.1. Statistical evaluation

Table 3.3 reports the NRMSE [%] for active and reactive power, averaged over the five load configurations of Table 3.1, for both the ILF-based approach and the DP benchmark across the four ML algorithms.

Table 3.3.: NRMSE [%] for active (P) and reactive (Q) power prediction, averaged over the five load configurations. Results shown for the DP benchmark and the ILF-based approach.

Algorithm	P		Q	
	DP	ILF-based	DP	ILF-based
KNN	9.53	0.41	9.18	2.58
RF	2.13	0.39	2.22	0.85
GB	5.07	0.40	4.68	1.85
MLP	0.88	0.40	1.46	0.98

The ILF-based approach consistently and substantially outperforms the DP benchmark across all four algorithms and for both active and reactive power. The performance gain is particularly pronounced for KNN, where the NRMSE drops from 9.53 % to 0.41 % for active power and from 9.18 % to 2.58 % for reactive power. Even for MLP, which achieves the best DP performance, the ILF-based approach reduces the active power NRMSE from 0.88 % to 0.40 % and the reactive power NRMSE from 1.46 % to 0.98 %.

This consistent pattern across all algorithms confirms that the performance gain is not attributable to the choice of ML model but to the ILF step itself. By decomposing the prediction task into the identification of equivalent impedance parameters followed by a deterministic load flow, the ILF-based approach reduces the complexity of the learning problem. The predictor does not need to implicitly replicate the full nonlinear power flow computation. It only needs to track the variation of R^{eq} and X^{eq} with the operating point, which is a smoother and more structured regression target.

Among the ILF-based results, RF achieves the lowest NRMSE for active power (0.39%) and the second lowest for reactive power (0.85%), making it the best-performing algorithm overall. It is therefore selected as the representative ILF-based predictor for the application-based evaluation in Section 3.4.2. For the DP benchmark, MLP achieves the best performance and is selected as the DP representative.

3.4.2. Application-based evaluation

Table 3.4 reports the cumulative generation and deviation costs over the entire test set for the four methods compared. These methods are the ILF-based approach with RF, the DP benchmark with MLP, the Naive persistence predictor, and the Ground Truth (GT) baseline. The GT baseline uses actual interface power exchanges in both OPF steps.

Table 3.4.: Cumulative generation and deviation costs over the test set [k€] for each forecasting method.

Cost [k€]	ILF-based (RF)	DP (MLP)	Naive	GT
Generation	8253.58	8246.44	8250.49	8253.69
Deviation	5.42	19.08	697.95	0
Total	8259.00	8265.52	8948.44	8253.69

The GT baseline gives zero deviation costs by definition, since the same power exchanges are used in both OPF steps. Among the three forecasting methods, the ILF-based approach achieves the lowest total cost (8259 k€), closest to the GT value of 8253.69 k€. The deviation costs are the most discriminating component: the ILF-based approach incurs 5.42 k€, compared to 19.08 k€ for DP (a factor of 3.5) and 697.95 k€ for the Naive predictor (a factor of 129). Generation costs, being driven primarily by the actual load level rather than prediction accuracy, are comparable across all methods.

These results confirm that the statistical superiority of the ILF-based approach reported in Section 3.4.1 translates into concrete operational gains. More accurate interface power forecasts allow the TSO to commit generator setpoints that are closer to the optimal dispatch under real conditions, thereby reducing the need for costly curative re-dispatch. The disparity with the Naive method also demonstrates that the prediction horizon (one quarter-hourly step ahead, obtained by evaluating the learned mapping on the predicted feature vector of the subsequent operating point) is not trivially predictable from the previous measurement. This underlines the value of a model-based forecasting approach.

3.4.3. Uncertainty sensitivity analysis

The previous evaluations assumed perfect knowledge of the aggregate load inputs P_l and Q_l . In practice, P_l and Q_l are not known exactly but forecast, and any forecast is subject to error; it is this input forecast error that the addition of noise is intended to simulate. The characterisation of such input uncertainty is studied in the context of distribution state estimation, where unmetered loads are represented by pseudo-measurements [125]; real load profiles deviate from the standard profiles used to construct these pseudo-measurements, and the propagation of this deviation has been quantified [126]. The present work does not perform state estimation; it borrows from this literature the characterisation of uncertainty over estimated aggregate load.

Whereas the statistical and application-based evaluations reported above relied on realised feature values, this analysis measures the degradation incurred when inputs are affected by input uncertainty, that is, the performance of the model under realistic forecasting conditions. Only P_l and Q_l are subject to noise injection in this experiment. Meteorological inputs and PV production, which are likewise forecast in practice, are held at their realised values. The analysis therefore constitutes a partial proxy for the total input forecast error. This subsection investigates the sensitivity of the total operational cost to increasing levels of input uncertainty.

Gaussian noise of progressively increasing standard deviation is added to both P_l and Q_l in the test set, and the resulting total costs for both the ILF-based approach and the DP benchmark are recorded. The results are illustrated in Figure 3.6.

Both methods exhibit a monotonically increasing total cost with growing noise standard deviation, as expected: larger input errors propagate into less accurate interface power forecasts, increasing curative re-dispatch costs. The ILF-based approach maintains a consistently lower total cost than DP across the entire

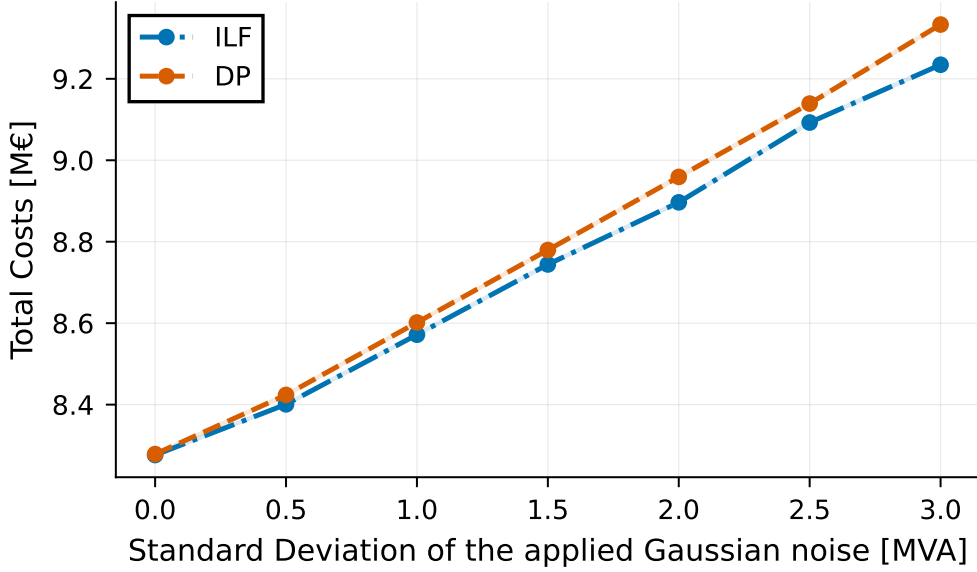


Figure 3.6.: Total operational cost as a function of the standard deviation of Gaussian noise introduced into the aggregate load inputs P_l and Q_l . The ILF-based approach (RF) is compared to the DP benchmark (MLP).

noise range tested. The two curves exhibit a similar rate of increase with noise, indicating that neither method degrades disproportionately faster than the other as uncertainty grows. The absolute cost advantage of the ILF-based approach is therefore preserved under realistic levels of load estimation error.

This result is practically relevant: it demonstrates that the superiority of the ILF-based approach does not depend on access to highly accurate load estimates. Even when the aggregate load inputs are corrupted by realistic levels of uncertainty, the physical structure embedded in the equivalent model provides a more reliable mapping to interface power exchanges than a purely data-driven predictor operating on the same noisy inputs.

3.5. Discussion and Limitations

This section interprets the results reported in Section 3.4 and discusses the limitations of the proposed approach. Section 3.5.1 provides this interpretation, while Section 3.5.2 examines the limitations.

3.5.1. Interpretation of the results

The results presented in Section 3.4 demonstrate that the ILF-based approach consistently outperforms the DP benchmark, and that this advantage holds both statistically and operationally. The underlying reason is structural. By introducing an intermediate physical representation of the distribution network, the approach separates what is known from physics (the load flow equations governing the equivalent model) from what must be learned from data (the operating point dependence of R^{eq} and X^{eq}). The ML predictor is therefore tasked with a smoother and more tractable regression problem than in the DP case, where the full nonlinear relationship between aggregate loads and interface power exchanges must be approximated entirely from data.

The operational interpretation of the cost results is also informative. The deviation costs of 5.42 k€ for the ILF-based approach and 19.08 k€ for DP, while both small relative to total generation costs, reflect the frequency and magnitude of prediction errors that force the TSO to deviate from its planned dispatch. In a real system, such deviations carry costs beyond the financial component: they reduce the margin available for other control actions and increase the risk of constraint violations. The factor-of-3.5 advantage of the ILF-based approach therefore represents a meaningful operational improvement.

The robustness analysis further shows that this advantage is not contingent on precise load estimation. The ILF-based approach degrades gracefully with increasing input noise, maintaining lower costs than DP across the full range tested. This property is critical for practical deployment, since TSO-level load estimates are inherently subject to forecast errors, particularly at the distribution level where metering granularity is limited.

3.5.2. Limitations of the proposed approach

Several limitations of the proposed approach must be acknowledged, both to bound the validity of the results and to identify directions for future improvement.

Synthetic training data. The methodology relies on a synthetic dataset generated from a simulated distribution network, since real TSO measurements of interface power exchanges are not publicly available. While the Oberrhein network provides a realistic benchmark, the trained ML models are optimised for that specific network and its load profile statistics. Deployment on a different distribution network would require retraining the models on a new dataset built from that network’s own measurements or simulations. The degree to which the

equivalent model structure generalises across topologically dissimilar networks is an open question.

Single-interface equivalent model. The ILF formulation represents the entire distribution network as a single equivalent two-bus line. This representation captures the aggregate reactive behaviour of the network, including losses and the net effect of distributed loads and generation, but it does not distinguish between individual feeders or voltage levels within the distribution system. Operating conditions in which the reactive behaviour is driven by a localized event, such as a large industrial load switching or the sudden loss of a feeder, may not be captured accurately by the aggregate equivalent.

Unity power factor assumption for PV. The model assumes that PV inverters operate at unity power factor, so that their reactive contribution to Q_l is zero. Modern grid codes increasingly require PV inverters to provide reactive support, which would invalidate this assumption. Extending the model to account for the reactive output of PV systems would require either additional measurements or a more refined aggregate load decomposition.

Static equivalent parameters. The equivalent parameters R^{eq} and X^{eq} are computed from a fixed network topology. In practice, distribution networks undergo frequent reconfigurations, new connections, and structural changes due to the integration of RES. The trained ML models would need to be periodically retrained as the underlying network evolves, and the rate of this retraining is not addressed in the present work.

Absence of reactive power control. The application-based evaluation focuses solely on active power dispatch optimisation, without incorporating reactive power control levers such as On-Load Tap Changers, SVCs, or STATCOMs. Including reactive compensation in the OPF formulation would produce a more realistic representation of TSO operational practice, and could interact with the accuracy of the reactive power forecasts in non-trivial ways.

3.6. Chapter Conclusion

This chapter addressed the problem of forecasting reactive power exchanges at T-D interfaces from the TSO perspective, under the fundamental constraint that the distribution network topology is inaccessible. The proposed approach combines an ILF formulation with supervised ML regression to predict the equivalent impedance parameters of the distribution network at each operating point, from which interface power exchanges are reconstructed via a standard load flow computation.

The main results can be summarised as follows. On the statistical side, the ILF-based approach achieves an average NRMSE of 0.98 % against 4.38 % for the DP benchmark, with consistent gains across all four ML algorithms and both active and reactive power. On the operational side, the application-based evaluation shows that the ILF-based approach reduces curative re-dispatch costs by a factor of 3.5 compared to DP, and that this advantage is maintained in the presence of realistic input uncertainty.

These results validate the core hypothesis of the chapter: embedding the physical structure of the load flow equations into the forecasting pipeline, even in a highly simplified two-bus form, leads to a fundamentally more tractable learning problem and more reliable predictions than a purely data-driven approach. The equivalent model acts as a physically grounded intermediate representation that absorbs the nonlinear coupling between aggregate loads and interface power exchanges, leaving the ML predictor to track a smoother operating-point-dependent signal.

The limitations identified in Section 3.5 point to several directions for future work: generalisation of the equivalent model to other distribution topologies, incorporation of reactive control in the OPF evaluation, extension to inverter-based reactive support from PV systems, and the development of online learning strategies to track distribution network evolution over time. These perspectives are discussed further in Chapter 4. The estimated active and reactive exchange profiles correspond to the quantity that the scenario generation of Chapter 2 represents by a synthetic model, and their coupling is examined in Section 4.3.4.

CHAPTER 4.

General Conclusion and Perspectives

Recent transformations of power systems, driven by the energy transition, are pushing them increasingly close to the limits within which system security must be maintained. Because reactive power does not travel efficiently over long distances, it must be managed close to where it arises. This locality plays out at two distinct levels of the network. At the system level, investment decisions determine the compensation capacity available and its placement. At the level of individual interfaces, the actual behaviour that these decisions must anticipate can be characterised. This thesis has approached the problem from both ends, with VARPlan at the scale of the transmission system and VARLearn at the scale of the T-D interface. This chapter synthesises the contributions of these two approaches, sets out the limits that bound their scope, and identifies the research directions to which they open.

4.1. Summary of Contributions

Section 4.1.1 recalls the contribution of VARPlan while Section 4.1.2 recalls that of VARLearn.

4.1.1. VARPlan: Reactive Power Planning at the System Scale

VARPlan formulates reactive compensation planning as a security-constrained program, covering multiple operating conditions and jointly optimising investment and operational costs. Its primary contribution lies in bringing together,

for the first time within a single formulation, three elements that the literature had addressed separately. These are an objective representing the actual operating cost borne by the system operator, a convex relaxation that remains faithful under capacitive operating conditions, and a prior reduction of the decision space achieved through weak-bus analysis. For the system operator, this means an investment decision that reflects the cost structure it actually faces, rather than a proxy quantity. The methodology has further revealed that the convex relaxation on which tractable approaches rely loses its faithfulness precisely in the capacitive operating conditions toward which networks are evolving. It has also shown that a reactive scarcity indicator derived from the problem itself localizes investment needs more effectively than conventional voltage stability indicators.

4.1.2. VARLearn: Reactive Power Forecasting at the Local Scale

VARLearn predicts active and reactive power exchanges at T-D interfaces by combining the physical structure of the network with information extracted from measurements available to the system operator. Its contribution lies in constructing an equivalent model of the distribution network as seen from the transmission side of the interface, without requiring knowledge of its topology, which the TSO does not possess. It thereby occupies a position that neither purely model-based approaches, which require this topology, nor purely data-driven approaches, which discard physical structure, can hold. For the system operator, this provides a characterisation of local interfaces that can be derived from the data it already has access to.

4.2. Limitations and Open Questions

The contributions of this thesis are accompanied by limits that bound their scope, detailed in their respective contexts in the preceding chapters. The most structural among them are recalled here at the level of the thesis as a whole, each opening onto a research direction developed in the following section.

The first limitation concerns the convex relaxation underpinning VARPlan's tractability, which is not exact under capacitive operating conditions and must be supplemented by a regularisation term to remain faithful to the original problem. The branches on which that term acts are selected through a physical criterion known prior to solving. However, whether to activate the term for a given operating trajectory is currently decided a posteriori, by observing which

trajectory benefits from it against the AC reference. No criterion identifies this need beforehand. The regularisation therefore remains an added corrective to the formulation rather than an intrinsic property of the relaxation itself. Identifying a relaxation whose exactness would be guaranteed in capacitive operating conditions, or at least a criterion that anticipates when the corrective is needed, remains an open question, to which Section 4.3 returns.

The second limitation concerns the scope on which both contributions were established: a reduced-scale test network for VARPlan, and a single T-D interface for VARLearn. This scope reflects a deliberate validation choice, necessary to maintain an exact reference solution and to isolate the properties under investigation, but it leaves open the question of how the methods behave at larger scale. Extension to realistic network sizes and to a diversity of interfaces constitutes a natural direction, taken up in what follows.

The third limitation is that the two contributions, though addressing the same problem at two distinct scales, have not been coupled. VARPlan treats T-D interfaces as exogenous load injections built from a synthetic reactive load model, rather than as the active and reactive exchange profiles that VARLearn estimates for these same interfaces. Coupling the two goes beyond a direct substitution, however, because both contributions were established on different time horizons. Constructing a framework that unifies both contributions is the most direct perspective opened by this thesis, and constitutes the subject of Sections 4.3.3 and 4.3.4.

4.3. Future Research Directions

This section develops the research directions to which the preceding limitations point. It proceeds in four steps. The first two address VARPlan: improvements to the formulation and extension to larger networks, then conditions for its operational deployment. The third covers extensions of VARLearn to new interface configurations and learning settings. The last, and most direct perspective opened by this thesis, is the construction of a joint framework articulating both contributions.

4.3.1. VARPlan: Model Refinements and Scalability

A first, incremental step towards larger networks builds directly on the decision-space reduction developed in this thesis. The weak-bus analysis restricts the set of candidate investment buses. This reduction is orthogonal to approaches

which instead reduce the operational constraint space by identifying binding contingencies and conditions [71]. Combining the two, one acting on the investment decision, the other on the constraint set, offers a natural and direct extension of the pipeline toward realistic network sizes.

Beyond this combination, a more structural route toward scaling VARPlan relies on a Benders-type decomposition, which separates the master investment problem from the operational subproblem that evaluates its quality across operating conditions. This separation requires the subproblem to be convex and free of integer variables, which would call for relaxing discrete device switching into a continuous variable, an additional approximation beyond the convex relaxation already used in this work. A recent formulation applying this route to an analogous battery sizing problem establishes its feasibility within a conic framework [127]. Beyond scalability, whether finer relaxations, such as semidefinite [46] or moment-based approaches [50], maintain its exactness in capacitive conditions without recourse to a regularisation term remains an open question warranting dedicated investigation. Extending the investment horizon to multiple periods and incorporating battery storage, whose reactive capability enriches the available compensation options, constitute further directions for model development.

A more structurally significant direction concerns the treatment of uncertainty, which appears at two scales. The first scale is the variability of operating conditions and contingencies within a trajectory. At this scale, robustness is a parameter of the model and not a result of the method. The N-1 criterion accepts a residual risk instead of covering multiple simultaneous outages [19]. The coefficient γ^{slack} (Section 2.2.2) gives a price to reactive insufficiency. When this price is high, the solution does not use the slack variable, so the reactive power balance must be met in every operating condition and contingency of the planning set. The current formulation minimises an expected cost over these weighted operating conditions. This approach averages out extreme conditions by construction, whereas the network operator must also manage rare and severe events. The second scale is the choice between trajectories, to which no probability can be assigned. The trade-off between robustness and optimality is meaningful at this scale. In this thesis, the trajectories are solved separately and are not weighted. Only the Demand Sufficiency trajectory requires inductive compensation (Section 2.5.3). If the three trajectories were combined in a single expected-cost plan, the coverage of this need would depend on a probability that cannot be estimated, and not only on the physical need. A robust formulation would replace the expected cost by a guarantee over the most constraining futures [74]. Such a change would modify the nature of the investment decisions, which would be sized for security under critical conditions rather than for economy on average.

The choice of the futures against which the network must be protected remains a strategic decision, which lies beyond the planning tool.

4.3.2. VARPlan: Towards Operational Deployment

Beyond model refinements, the operational deployment of VARPlan within a real network operator raises questions of its own. Transposing the method to a real transmission system, such as the Belgian network with its approximately nine thousand nodes, requires defining trajectories, operating conditions, and a topology specific to that system, and thereby connects directly to the scalability question addressed in the previous section. Calibrating the cost structure against the operator's actual tariffs and budgets, rather than against plausible economic signals, is a prerequisite for the decisions to serve as effective operational recommendations. Integrating the method into the operator's existing tools and articulating it with reactive power pricing mechanisms constitute complementary steps toward this use. Adaptive generation of operating conditions, updated as the system evolves, would extend the method toward a time-revisable framework.

4.3.3. VARLearn: Model Extensions and Physical Interpretation

A first direction concerns the physical interpretation of the equivalent parameters that VARLearn estimates. The actual scope of the model depends on two open questions. The first is the physical status of the equivalent resistance and reactance obtained for a given interface, reflecting either its actual physical properties or a statistical fit to the training data. The second is the extent to which the same equivalent circuit generalises to interfaces with a different topology or voltage level. If this physical interpretation can be confirmed, the equivalent parameters could be updated directly from known physical changes to the network, rather than re-estimated from new measurements each time the network changes. The progressive replacement of OLs by UCs, for instance, would translate into a predictable decrease of the equivalent reactance. This transition is the same physical shift that drives the reactive power management challenge addressed by VARPlan at system scale, so representing it within VARLearn's equivalent parameters would connect both contributions around a common physical mechanism.

A more fundamental extension would concern the temporal dimension of the model: in its current form, VARLearn estimates interface exchanges at the

present instant from quantities available at that same instant, without anticipation. Exploiting observations from prior instants to predict upcoming exchanges would shift the model from instantaneous estimation to genuine forecasting, at the cost of accounting for the temporal dependence of the model parameters. Online learning would further allow the model to adapt as new measurements become available, while uncertainty quantification would associate confidence intervals with its outputs. Generalising to other T-D interface types and network configurations would extend its domain of application beyond the case studied in this work.

4.3.4. Towards a Joint VARLearn-VARPlan Framework

The most direct perspective opened by this thesis is to feed VARPlan's scenario generation with the active and reactive exchange profiles that VARLearn estimates at each interface. These profiles would replace the synthetic reactive load model of Section 2.3.2, in which the reactive injection of each bus is set by a power factor drawn per bus and by a capacitive offset δ_i fixed by the trajectory, as given by Equation (2.60). The investment decision made at system scale would then rest on interface behaviour estimated from measurements rather than assumed.

The two contributions were nevertheless established on different time horizons, and this difference conditions the coupling. VARLearn estimates the equivalent parameters of an interface from quarter-hourly measurements of the present system, and reconstructs the exchanges at the operating point observed. VARPlan evaluates investment decisions over prospective trajectories that describe the system several years ahead, for which no measurement exists. A direct substitution is therefore possible only for a study of the present system. Such a study would already be informative, since it would indicate whether the synthetic model of Section 2.3.2 is an acceptable proxy for the measured behaviour of the interfaces. It would not, however, reach the horizon on which the investment decision is actually made.

The coupling becomes relevant to planning only if the equivalent parameters can be generalised beyond the conditions under which they were estimated. This requirement has two components.

1. Physical status of these parameters, discussed in Section 4.3.3: if they describe properties of the interface and not only a statistical fit, their evolution can be derived from the planned evolution of the distribution network rather than from measurements that do not yet exist.
2. Availability of the input features of VARLearn for a future operating point. On the active side, the trajectories of Section 2.3.2 already provide them,

since the demand growth and the PV share are applied bus by bus. On the reactive side, an assumption remains necessary, but it bears on the aggregated reactive load of the distribution network and no longer on the exchange observed at the interface, which becomes an output of the equivalent model. These features are defined per time step, so the coupling would take place on the quarter-hourly series, before the clustering described in Section 2.3.4.

If both conditions are met, VARLearn no longer acts as a predictor of present exchanges, but as a generator of interface scenarios for each trajectory. The power factor and the capacitive offset of Equation (2.60), which are drawn parameters, are then replaced by quantities that carry a physical meaning and an established estimation procedure. If these conditions cannot be met, the coupling remains restricted to the present system, and the synthetic model retains its role for prospective trajectories.

Other coupling directions are conceivable. A growing body of literature envisions the distribution network as a provider of voltage support services, capable of modulating its reactive power in response to transmission-level requirements [128]. Recognising the interface as a flexible asset, characterised by an operating domain in active and reactive power rather than a fixed profile, would alter the nature of the investment problem itself. By reducing the need for dedicated compensation, it would defer part of the investments that VARPlan dimensions. A feedback loop in which investment decisions modify the behaviour of interfaces, calling for an update of the learned model, constitutes a further direction. Taken together, these elements would link interface characterisation to investment planning within a single framework.

4.4. Closing Remarks

The energy transition is most often described in terms of active power: installed renewable capacity, energy produced, and emissions avoided. Reactive power, largely invisible and local, occupies a far less visible place in that picture, while it conditions voltage regulation and therefore the security of the network that makes the transition possible. The gradual replacement of OLs by UCs and the retreat of synchronous machines are silently shifting the reactive balance of transmission networks toward operating conditions that their operational and planning tools were not designed to handle. This thesis addresses precisely this structurally significant transformation.

Rather than treating reactive power management as a secondary operational adjustment, this work approached it as a decision problem in its own right, encompassing investment and anticipation alike. It showed that a planning decision carries value only insofar as the operational model it incorporates remains faithful to the actual conditions of the network, including when those conditions depart from what conventional tools were designed to handle. It showed, symmetrically, that a fine characterisation of local exchanges at interfaces can be obtained from the data the operator already holds, without demanding knowledge that lies beyond reach. Both findings respond to the same requirement: grounding decisions in a faithful representation of the network, at system scale as much as at the level of individual interfaces.

Although these two contributions operate at distinct scales and draw on different methods, they proceed from a shared conviction: that reactive power, because it does not travel efficiently over long distances, must be managed where it is in surplus and where it is in deficit. This locality, a structuring principle of this work, demands that the network be conceived simultaneously at system scale, where investment decisions are made, and at the level of interfaces, where actual behaviour becomes legible. Kept separate within this thesis, these two scales call to be brought together, and it is toward that union that the perspectives opened here point. A framework in which local characterisation directly informs the system-level decision would constitute the natural culmination of the approach undertaken in this work.

As transmission networks incorporate more UCs, power converters, and decentralized generation, reactive power management will cease to be a secondary operational concern and become a structuring component of long-term planning. The tools that support this planning will need to embrace that evolution rather than lag behind it, faithfully representing a network whose reactive behaviour increasingly diverges from the one for which they were originally designed. By contributing to this adaptation, at two complementary scales, this work hopes to have offered a modest contribution to reactive power management in the electric power network that the energy transition is shaping.

Bibliography

- [1] R. P. Feynman, R. B. Leighton, and M. Sands, *The Feynman Lectures on Physics, Vol. II: Mainly Electromagnetism and Matter*. Reading, MA: Addison-Wesley, 1964.
- [2] L. T. et al., “Pandapower — an open-source python tool for convenient modeling, analysis, and optimization of electric power systems,” *IEEE Trans. Power Syst.*, vol. 33, no. 6, pp. 6510–6521, 2018.
- [3] P. Kundur, *Power System Stability and Control*. McGraw-Hill, 1994.
- [4] C. G. Kaloudas and et al., “Assessing the future trends of reactive power demand of distribution networks,” *IEEE Trans. Power Syst.*, vol. 32, no. 6, pp. 4278–4288, 2017.
- [5] M. Tazky, M. Regula, and A. Otcenasova, “Impact of changes in a distribution network nature on the capacitive reactive power flow into the transmission network in Slovakia,” *Energies*, vol. 14, no. 17, 2021.
- [6] M. Ghassemi, “High surge impedance loading (hsil) lines: A review identifying opportunities, challenges, and future research needs,” *IEEE Transactions on Power Delivery*, vol. 34, no. 5, pp. 1909–1924, 2019.
- [7] “Nuclear phase-out increases co2 emissions from electricity generation,” <https://shorturl.at/bdinW>, accessed: 2023-12-19.
- [8] “A new era: Germany quits nuclear power, closing its final three plants,” <https://edition.cnn.com/2023/04/15/europe/germany-nuclear-phase-out-climate-intl/index.html>, accessed: 2023-12-19.
- [9] European Parliament and Council of the European Union, “Directive (eu) 2023/2413 on the promotion of the use of energy from renewable sources,” <https://eur-lex.europa.eu/eli/dir/2023/2413>, 2023, accessed: 2026-04-30.

- [10] A. H. Javed, P. H. Nguyen, J. Morren, J. Han Slootweg, and S. Battacharyya, “Issues of capacitive reactive power flow in electricity networks,” in *2022 57th International Universities Power Engineering Conference (UPEC)*, 2022, pp. 1–6.
- [11] A. Rosseel, B. B. Zad, and Z. De Grève, “A novel scenario-based reactive power planning method including investment and operational costs,” in *2025 IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT Europe)*, 2025, pp. 1–5.
- [12] A. Rosseel, B. Bakhshideh Zad, and Z. De Grève, “A cost-reflective misocp formulation for reactive power planning with weak-bus-guided decision-space reduction,” *Sustainable Energy, Grids and Networks*, 2026, submitted August 2026.
- [13] A. Rosseel, B. Bakhshideh Zad, Z. De Grève, and F. Vallée, “Reactive power forecasting at the transmission-distribution interfaces using physics-based machine learning,” in *27th International Conference on Electricity Distribution (CIRED 2023)*, vol. 2023. IET, 2023, pp. 301–305.
- [14] A. Rosseel, B. Bakhshideh Zad, F. Vallée, and Z. De Grève, “Physics-informed machine learning for forecasting power exchanges at the interface between transmission and distribution systems,” *Electric Power Systems Research*, vol. 238, p. 111097, 2025.
- [15] A. Rosseel, M. Hupez, J.-F. Toubeau, Z. De Grève, and F. Vallée, “On the optimization of investments in distributed energy resources in a low-voltage energy community,” in *26th International Conference on Electricity Distribution (CIRED 2021)*, vol. 2021, 2021, pp. 1792–1796.
- [16] J. Allard, A. Rosseel, L. Sadoine, J. Faraji, T. Brihay, F. Capizzi, B. Nteziyaremye, F. Bordes, F. Vallée, and Z. De Grève, “Technical impacts of the deployment of renewable energy communities on electricity distribution grids,” in *27th International Conference on Electricity Distribution (CIRED 2023)*, vol. 2023, 2023, pp. 2923–2927.
- [17] A. Shuvarina, P. H. Nguyen, J. Morren, A. Ciupuliga, M. Brands, and J. Slootweg, “Reactive Power Management Challenges with Increasing Underground Cabling,” in *2024 59th International Universities Power Engineering Conference (UPEC)*. Cardiff, United Kingdom: IEEE, 2024, pp. 1–6.
- [18] F. Capitanescu, “Evaluating reactive power reserves scarcity during the energy transition toward 100% renewable supply,” *Electric Power Systems Research*, vol. 190, p. 106672, 2021.

- [19] European Commission and Parliament, “Commission regulation (eu) 2017/1485 of 2 august 2017 establishing a guideline on electricity transmission system operation,” <https://eur-lex.europa.eu/eli/reg/2017/1485/oj/eng>, 2017, official Journal of the European Union, L 220.
- [20] European Commission, “Commission regulation (eu) 2016/631 of 14 april 2016 establishing a network code on requirements for grid connection of generators,” Official Journal of the European Union, L 112/1, 2016. [Online]. Available: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32016R0631>
- [21] ENTSO-E, “Network code on demand connection,” https://www.entsoe.eu/network_codes/dcc/, 2023, accessed on 16-02-2023.
- [22] F. Capitanescu, “TSO–DSO interaction: Active distribution network power chart for TSO ancillary services provision,” *Electric Power Systems Research*, vol. 163, pp. 226–230, 2018.
- [23] S. Stanković, L. Söder, Z. Hagemann, and C. Rehtanz, “Reactive Power Support Adequacy at the DSO/TSO Interface,” *Electric Power Systems Research*, vol. 190, 2021.
- [24] A. Eicke and T. Schittekatte, “Fighting the wrong battle? a critical assessment of arguments against nodal electricity prices in the european debate,” *Energy Policy*, 2022, volume et pages a confirmer.
- [25] Joint Research Centre, “Nodal pricing in the european internal electricity market,” European Commission, Tech. Rep. JRC119977, 2020.
- [26] J. Zhong and K. Bhattacharya, “Reactive power as an ancillary service,” *IEEE Transactions on Power Systems*, vol. 16, no. 2, May 2001.
- [27] CREG, “Projet de décision de dérogation ex ante à la fourniture fondée sur le marché des services de puissance réactive à elia transmission belgium sa pour la période de fourniture 2027-2029, y compris,” Commission de Régulation de l’Électricité et du Gaz (CREG), Brussels, Belgium, Public consultation document PRD3120–PRD3125, 2024. [Online]. Available: <https://www.creg.be/fr/consultations-publiques/prd3120-prd3125>
- [28] J. W. Lamont and J. Fu, “Cost analysis of reactive power support,” *IEEE Transactions on Power Systems*, vol. 14, no. 3, pp. 890–898, 1999.
- [29] J. B. Gil, T. Gómez San Román, J. J. Alba Ríos, and P. Sánchez Martín, “Reactive power pricing: A conceptual framework for remuneration and charging procedures,” *IEEE Transactions on Power Systems*, vol. 15, pp. 483–489, 2000.

- [30] Y. Muhammad, R. Khan, M. A. Z. Raja, F. Ullah, N. I. Chaudhary, and Y. He, "Solution of optimal reactive power dispatch with FACTS devices: A survey," *Energy Reports*, vol. 6, 2020.
- [31] S. Shekarappa G, S. Mahapatra, and S. Raj, "A novel meta-heuristic approach for optimal RPP using series compensated FACTS controller," *Intelligent Systems with Applications*, vol. 18, 2023.
- [32] N. G. Hingorani and L. Gyugyi, *Understanding FACTS: Concepts and Technology of Flexible AC Transmission Systems*. IEEE Press, 2000.
- [33] EPRI, "Technical and operational comparison of SVC, STATCOM and synchronous condensers," Electric Power Research Institute, Tech. Rep. 3002024629, 2022.
- [34] J. Carpentier, "Contribution a l'etude du dispatching economique," *Bulletin de la Société Francaise des électriciens*, vol. 3, no. 1, pp. 431–447, 1962.
- [35] H. W. Dommel and W. F. Tinney, "Optimal power flow solutions," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-87, no. 10, pp. 1866–1876, 1968.
- [36] F. Capitanescu, "Critical review of recent advances and further developments needed in ac optimal power flow," *Electric Power Systems Research*, vol. 136, pp. 57–68, 2016.
- [37] C. Zhang, H. Chen, K. Shi, Z. Liang, W. Mo, and D. Hua, "A Multi-time Reactive Power Optimization Under Interval Uncertainty of Renewable Power Generation by an Interval Sequential Quadratic Programming Method," *IEEE Transactions on Sustainable Energy*, vol. 10, no. 3, pp. 1086–1097, 2019.
- [38] P. P. Biswas, P. Suganthan, R. Mallipeddi, and G. A. Amaratunga, "Optimal reactive power dispatch with uncertainties in load demand and renewable energy sources adopting scenario-based approach," *Applied Soft Computing*, vol. 75, pp. 616–632, 2019.
- [39] N. Habib Khan, R. Jamal, M. Ebeed, S. Kamel, H. Zeinoddini-Meymand, and H. M. Zawbaa, "Adopting Scenario-Based approach to solve optimal reactive power Dispatch problem with integration of wind and solar energy using improved Marine predator algorithm," *Ain Shams Engineering Journal*, vol. 13, no. 5, p. 101726, 2022.

- [40] M. Nasouri Gilvaei, H. Jafari, M. Jabbari Ghadi, and L. Li, “A novel hybrid optimization approach for reactive power dispatch problem considering voltage stability index,” *Engineering Applications of Artificial Intelligence*, vol. 96, 2020.
- [41] N. Karmakar and B. Bhattacharyya, “A reactive power planning model for power transmission systems using meta-heuristics algorithms,” *Decision Analytics Journal*, vol. 7, 2023.
- [42] D. Bienstock and A. Verma, “Strong np-hardness of ac power flows feasibility,” *Operations Research Letters*, vol. 47, no. 6, pp. 494–501, 2019.
- [43] R. A. Jabr, “Radial distribution load flow using conic programming,” *IEEE Transactions on Power Systems*, vol. 21, no. 3, pp. 1458–1459, Aug. 2006.
- [44] —, “A conic quadratic format for the load flow equations of meshed networks,” *IEEE Transactions on Power Systems*, vol. 22, no. 4, pp. 2285–2286, 2007.
- [45] X. Bai, H. Wei, K. Fujisawa, and Y. Wang, “Semidefinite programming for optimal power flow problems,” *International Journal of Electrical Power & Energy Systems*, vol. 30, no. 6–7, pp. 383–392, 2008.
- [46] J. Lavaei and S. H. Low, “Zero duality gap in optimal power flow problem,” *IEEE Transactions on Power Systems*, vol. 27, no. 1, pp. 92–107, Feb. 2012.
- [47] M. Farivar and S. H. Low, “Branch flow model: Relaxations and convexification (parts I and II),” *IEEE Transactions on Power Systems*, vol. 28, no. 3, pp. 2554–2564, 2013.
- [48] C. Coffrin, H. L. Hijazi, and P. Van Hentenryck, “The QC Relaxation: Theoretical and Computational Results on Optimal Power Flow,” 2015.
- [49] D. K. Molzahn and I. A. Hiskens, “Moment-based relaxation of the optimal power flow problem,” in *Proceedings of the 18th Power Systems Computation Conference (PSCC)*, 2014.
- [50] C. Jozs, J. Maeght, P. Panciatici, and J. C. Gilbert, “Application of the moment-SOS approach to global optimization of the OPF problem,” *IEEE Transactions on Power Systems*, vol. 30, no. 1, pp. 463–470, Jan. 2015.
- [51] S. H. Low, “Convex relaxation of optimal power flow—part I: Formulations and equivalence,” *IEEE Transactions on Control of Network Systems*, vol. 1, no. 1, pp. 15–27, 2014.

- [52] D. K. Molzahn and I. A. Hiskens, “A survey of relaxations and approximations of the power flow equations,” *Foundations and Trends in Electric Energy Systems*, vol. 4, no. 1–2, pp. 1–221, 2019.
- [53] L. Gan, N. Li, U. Topcu, and S. H. Low, “Exact convex relaxation of optimal power flow in radial networks,” *IEEE Transactions on Automatic Control*, 2015, volume et pages a confirmer.
- [54] S. H. Low, “Convex relaxation of optimal power flow—part II: Exactness,” *IEEE Transactions on Control of Network Systems*, vol. 1, no. 2, pp. 177–189, 2014.
- [55] B. Kocuk, S. S. Dey, and X. A. Sun, “Inexactness of SDP relaxation and valid inequalities for optimal power flow,” *IEEE Transactions on Power Systems*, vol. 31, no. 1, pp. 642–651, Jan. 2016.
- [56] —, “Strong SOCP relaxations for the optimal power flow problem,” *Operations Research*, 2016, volume et pages a confirmer.
- [57] C. Coffrin, H. L. Hijazi, and P. Van Hentenryck, “Strengthening the SDP relaxation of AC power flows with convex envelopes, bound tightening, and valid inequalities,” *IEEE Transactions on Power Systems*, vol. 32, no. 5, pp. 3549–3558, 2017.
- [58] C. Chen, A. Atamtürk, and S. S. Oren, “Bound tightening for the alternating current optimal power flow problem,” *IEEE Transactions on Power Systems*, vol. 31, no. 5, pp. 3729–3736, 2016.
- [59] R. Madani, S. Sojoudi, and J. Lavaei, “Convex relaxation for optimal power flow problem: Mesh networks,” *IEEE Transactions on Power Systems*, vol. 30, no. 1, pp. 199–211, Jan. 2015.
- [60] R. Madani, M. Ashraphijuo, and J. Lavaei, “Promises of conic relaxation for contingency-constrained optimal power flow problem,” *IEEE Transactions on Power Systems*, vol. 31, no. 2, pp. 1297–1307, Mar. 2016.
- [61] D. K. Molzahn, C. Jozs, I. A. Hiskens, and P. Panciatici, “Solution of optimal power flow problems using moment relaxations augmented with objective function penalization,” in *Proceedings of the IEEE Conference on Decision and Control (CDC)*, 2015.
- [62] W. Zhang, F. Li, and L. M. Tolbert, “Review of reactive power planning: Objectives, constraints, and algorithms,” *IEEE Trans. Power Syst.*, vol. 22, no. 4, pp. 2177–2186, 2007.

- [63] S. Granville, M. V. F. Pereira, and A. J. Monticelli, “An integrated methodology for VAR sources planning,” *IEEE Transactions on Power Systems*, 1988, metadonnees a confirmer.
- [64] T. Gómez, I. J. Pérez-Arriaga, J. Lumbreras, and V. M. Parra, “A security-constrained decomposition approach to optimal reactive power planning,” *IEEE Transactions on Power Systems*, vol. 6, no. 3, pp. 1069–1076, Aug. 1991.
- [65] N. Karmakar and B. Bhattacharyya, “Techno-economic strategy for reactive power planning using Series-shunt compensation in power transmission network,” *Sustainable Energy Technologies and Assessments*, vol. 49, 2022.
- [66] X. Fang, F. Li, Y. Wei, R. Azim, and Y. Xu, “Reactive power planning under high penetration of wind energy using Benders decomposition,” *IET Generation, Transmission and Distribution*, vol. 9, no. 14, pp. 1835–1844, 2015.
- [67] E. F. Alvarez, M. Paredes, and M. J. Rider, “Semidefinite relaxation and generalised benders decomposition to solve the transmission expansion network and reactive power planning,” *IET Generation, Transmission and Distribution*, vol. 14, no. 11, pp. 2160–2168, 2020.
- [68] C. Li, A. J. Conejo, P. Liu, B. P. Omell, J. D. Sirola, and I. E. Grossmann, “Mixed-integer linear programming models and algorithms for generation and transmission expansion planning of power systems,” *European Journal of Operational Research*, vol. 297, no. 3, pp. 1071–1082, 2022.
- [69] S. Binato, M. Pereira, and S. Granville, “A new benders decomposition approach to solve power transmission network design problems,” *IEEE Transactions on Power Systems*, vol. 16, no. 2, pp. 235–240, 2001.
- [70] S. Lumbreras and A. Ramos, “The new challenges to transmission expansion planning: Survey of recent practice and literature review,” *Electric Power Systems Research*, vol. 134, pp. 19–29, 2016.
- [71] E. Davoodi, F. Capitanescu, M. I. Alizadeh, and L. Wehenkel, “A Scalable Var Planning Methodology to Mitigate Reactive Power Scarcity During Energy Transition,” *IEEE Trans. Power Syst.*, pp. 1–11, 2025.
- [72] N. Popli, E. Davoodi, F. Capitanescu, and L. Wehenkel, “On the robustness of machine-learnt proxies for security constrained optimal power flow solvers,” *Sustainable Energy, Grids and Networks*, vol. 37, p. 101265, 2024.

- [73] B. N. Giraud, A. Rajaei, and J. L. Cremer, "Constraint-driven deep learning for N-k security constrained optimal power flow," *Electric Power Systems Research*, vol. 235, p. 110692, 2024.
- [74] L. A. Roald, D. Pozo, A. Papavasiliou, D. K. Molzahn, J. Kazempour, and A. J. Conejo, "Power systems optimization under uncertainty: A review of methods and applications," *Electric Power Systems Research*, vol. 214, no. Part A, p. 108725, 2023.
- [75] A. H. Shojaei, A. A. Ghadimi, M. R. Miveh, F. H. Gandoman, and A. Ahmadi, "Multiobjective reactive power planning considering the uncertainties of wind farms and loads using Information Gap Decision Theory," *Renewable Energy*, vol. 163, pp. 1427–1443, 2021.
- [76] ENTSO-E and ENTSO-G, "TYNDP 2024 scenarios report," European Network of Transmission System Operators, Tech. Rep., 2024. [Online]. Available: <https://2024.entsos-tyndp-scenarios.eu/>
- [77] B. Stott, O. Alsac, and A. J. Monticelli, "Security analysis and optimization," *Proceedings of the IEEE*, vol. 75, no. 12, pp. 1623–1644, 1987.
- [78] A. J. Monticelli, M. V. F. Pereira, and S. Granville, "Security-constrained optimal power flow with post-contingency corrective rescheduling," *IEEE Transactions on Power Systems*, vol. 2, no. 1, pp. 175–182, 1987.
- [79] F. Capitanescu, J. Martinez Ramos, P. Panciatici, D. Kirschen, A. Marano Marcolini, L. Platbrood, and L. Wehenkel, "State-of-the-art, challenges, and future trends in security constrained optimal power flow," *Electric Power Systems Research*, vol. 81, no. 8, pp. 1731–1741, 2011.
- [80] S. Fliscounakis, P. Panciatici, F. Capitanescu, and L. Wehenkel, "Contingency ranking with respect to overloads in very large power systems taking into account uncertainty, preventive and corrective actions," *IEEE Transactions on Power Systems*, vol. 28, no. 4, pp. 4909–4917, 2013.
- [81] Y. Chi, Y. Xu, and T. Ding, "Coordinated var planning for voltage stability enhancement of a wind-energy power system considering multiple resilience indices," *IEEE Trans. Sustainable Energy*, vol. 11, no. 4, pp. 2367–2379, 2020.
- [82] S. Greene, I. Dobson, and F. L. Alvarado, "Sensitivity of the loading margin to voltage collapse with respect to arbitrary parameters," *IEEE Transactions on Power Systems*, vol. 12, no. 1, pp. 262–272, 1997, ordre des auteurs a confirmer.

- [83] T. Van Cutsem and C. Vournas, *Voltage stability of electric power systems*. Springer, 2008, vol. 13.
- [84] A. M. Shaheen, R. A. El-Sehiemy, and S. M. Farrag, “A reactive power planning procedure considering iterative identification of VAR candidate buses,” *Neural Computing and Applications*, vol. 31, no. 3, pp. 653–674, Mar. 2019.
- [85] P. Kessel and H. Glavitsch, “Estimating the voltage stability of a power system,” *IEEE Transactions on Power Delivery*, vol. PWRD-1, no. 3, pp. 346–354, Jul. 1986.
- [86] B. Gao, G. K. Morison, and P. Kundur, “Voltage stability evaluation using modal analysis,” *IEEE Transactions on Power Systems*, vol. 7, no. 4, Nov. 1992, pages a confirmer.
- [87] X. Liao, M. Zhang, J. Le, L. Zhang, and Z. Li, “Global sensitivity analysis of static voltage stability based on extended affine model,” *Electric Power Systems Research*, vol. 208, Jul. 2022.
- [88] C. A. Cañizares, A. C. Z. De Souza, and V. H. Quintana, “Comparison of performance indices for detection of proximity to voltage collapse,” *IEEE Transactions on Power Systems*, vol. 11, no. 3, pp. 1441–1450, Aug. 1996.
- [89] A. K. Sinha and D. Hazarika, “A comparative study of voltage stability indices in a power system,” *International Journal of Electrical Power & Energy Systems*, vol. 22, no. 8, pp. 589–596, Nov. 2000.
- [90] M. Glavic and T. Van Cutsem, “A short survey of methods for voltage instability detection,” in *IEEE Power and Energy Society General Meeting*, 2011.
- [91] Elia, “Open data platform,” <https://opendata.elia.be/pages/home/>, 2025, aggregated Belgian load and offshore wind generation datasets, accessed December 2025.
- [92] S. Babaeinejadsarookolae and et al., “The Power Grid Library for Benchmarking AC Optimal Power Flow Algorithms,” 2019.
- [93] MISO, “Transmission cost estimation guide for MTEP24,” Midcontinent Independent System Operator, Planning Subcommittee Document PSC Item 04, 5 2024. [Online]. Available: <https://cdn.misoenergy.org/20240501%20PSC%20Item%2004%20MISO%20Transmission%20Cost%20Estimation%20Guide%20for%20MTEP24632680.pdf>

- [94] I. Dunning, J. Huchette, and M. Lubin, “Jump: A modeling language for mathematical optimization,” *SIAM Review*, vol. 59, no. 2, pp. 295–320, 2017.
- [95] Gurobi Optimization, LLC, “Gurobi optimizer reference manual,” 2024. [Online]. Available: <https://www.gurobi.com>
- [96] A. Wächter and L. T. Biegler, “On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming,” *Mathematical Programming*, vol. 106, no. 1, pp. 25–57, 2006.
- [97] “Consortium des Équipements de calcul intensif (ceci),” <https://www.ceci-hpc.be>, funded by the F.R.S.-FNRS under Grant No. 2.5020.11.
- [98] A. G. Givisiez, K. Petrou, and L. F. Ochoa, “A review on TSO-DSO coordination models and solution techniques,” *Electric Power Systems Research*, vol. 189, p. 106659, 2020.
- [99] J. Silva, J. Sumaili, R. J. Bessa, L. Seca, M. A. Matos, V. Miranda, M. Caujolle, B. Goncer, and M. Sebastian-Viana, “Estimating the active and reactive power flexibility area at the TSO-DSO interface,” *IEEE Trans. Power Syst.*, vol. 33, no. 5, pp. 4741–4750, 2018.
- [100] G. Papazoglou and P. Biskas, “Review of methodologies for the assessment of feasible operating regions at the TSO-DSO interface,” *Energies*, vol. 15, no. 14, p. 5147, 2022.
- [101] P. D. Dapoz and et al., “Mixed regression clustering techniques for the medium term prediction of reactive power within transmission grids,” *ISGT-Europe*, pp. 1–5, 2019.
- [102] E. F. Arruda and et al., “Disaggregated active and reactive demand forecasting using first difference measured data and neural networks,” *Proceedings ICED*, 2015.
- [103] Y. Yuan, S. H. Low, O. Ardakanian, and C. J. Tomlin, “Inverse power flow problem,” *IEEE Trans. Control Netw. Syst.*, vol. 10, no. 1, pp. 261–273, 2023.
- [104] J. Yu, Y. Weng, and R. Rajagopal, “PaToPa: A data-driven parameter and topology joint estimation framework in distribution grids,” *IEEE Trans. Power Syst.*, vol. 33, no. 4, pp. 4335–4347, 2018.
- [105] J. B. Ward, “Equivalent circuits for power-flow studies,” *Trans. Amer. Inst. Elect. Eng.*, vol. 68, pp. 373–382, 1949.

- [106] M. Raissi, P. Perdikaris, and G. E. Karniadakis, “Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations,” *J. Comput. Phys.*, vol. 378, pp. 686–707, 2019.
- [107] B. Huang and J. Wang, “Applications of physics-informed neural networks in power systems – a review,” *IEEE Trans. Power Syst.*, vol. 38, no. 1, pp. 572–588, 2023.
- [108] P. Van Hentenryck, “A linear-programming approximation of ac power flows,” *INFORMS Journal on Computing*, vol. 26, pp. 718–734, 2014.
- [109] Synergrid, “Synthetic load profiles,” <http://www.synergrid.be/index.cfm?PageID=16896>, 2023, accessed on 15-07-2022.
- [110] F. H. Al-Qahtani and S. F. Crone, “Multivariate k-nearest neighbour regression for time series data — a novel algorithm for forecasting uk electricity demand,” in *The 2013 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 2013, pp. 1–8.
- [111] B. Bakhshideh Zad, J. F. Toubreau, K. Bruninx, B. Vatandoust, Z. De Grève, and F. Vallée, “Supervised learning-assisted modeling of flow-based domains in European resource adequacy assessments,” *Applied Energy*, vol. 325, p. 119875, 2022.
- [112] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd ed., ser. Springer Series in Statistics. New York: Springer, 2009. [Online]. Available: <https://web.stanford.edu/~hastie/ElemStatLearn/>
- [113] K. Chomboon, P. Chujai, P. Teerarassamee, K. Kerdprasop, and N. Kerdprasop, “An empirical study of distance metrics for k-nearest neighbor algorithm,” in *Proceedings of the 3rd International Conference on Industrial Application Engineering*, vol. 2, 2015.
- [114] L. Breiman, “Random forests,” *Machine Learning*, vol. 45, pp. 5–32, 2001.
- [115] J. Friedman, “Greedy function approximation: A gradient boosting machine,” *Annals of Statistics*, 2000.
- [116] A. Natekin and A. Knoll, “Gradient boosting machines, a tutorial,” *Frontiers in Neurorobotics*, 2013.
- [117] H. Ramchoun, Y. Ghanou, M. Ettaouil, and M. A. Janati Idrissi, “Multilayer perceptron: Architecture optimization and training,” *International Journal of Interactive Multimedia and Artificial Intelligence*, 2016.

- [118] M. Lindauer, K. Eggensperger, M. Feurer, A. Biedenkapp, D. Deng, C. Benjamins, T. Ruhkopf, R. Sass, and F. Hutter, “Smac3: A versatile bayesian optimization package for hyperparameter optimization,” *The Journal of Machine Learning Research*, vol. 23, no. 1, pp. 2475–2483, 2022.
- [119] F. Capitanescu, “Enhanced risk-based scopf formulation balancing operation cost and expected voluntary load shedding,” *Electric Power Systems Research*, vol. 128, pp. 151–155, 2015.
- [120] D. Schönheit, M. Kenis, L. Lorenz, D. Möst, E. Delarue, and K. Bruninx, “Toward a fundamental understanding of flow-based market coupling for cross-border electricity trading,” *Advances in Applied Energy*, vol. 2, 2021.
- [121] A. Dognini, M. Ginocchi, E. De Din, F. Ponci, and A. Monti, “Service restoration of ac–dc distribution grids based on multiple-criteria decision analysis,” *IEEE Access*, vol. 11, pp. 15 725–15 749, 2023.
- [122] S. Maharjan, A. M. Khambadkone, and J. C.-H. Peng, “Enhanced z-bus method for analytical computation of voltage sensitivities in distribution networks,” *IET Generation, Transmission & Distribution*, vol. 14, no. 16, pp. 3187–3197, 2020.
- [123] K. Strunz, E. Abbasi, R. Fletcher, N. Hatziaargyriou, R. Iravani, and G. Joos, *TF C6.04.02: TB 575 – Benchmark Systems for Network Integration of Renewable and Distributed Energy Resources*. Electra, 2014.
- [124] P. L. Joskow, “Comparing the costs of intermittent and dispatchable electricity generating technologies,” *American Economic Review*, vol. 101, no. 3, pp. 238–241, 2011.
- [125] R. Singh, B. C. Pal, and R. A. Jabr, “Distribution system state estimation through Gaussian mixture model of the load as pseudo-measurement,” *IET Gener. Transm. Distrib.*, vol. 4, no. 1, pp. 50–59, 2010.
- [126] A. Angioni, T. Schlösser, F. Ponci, and A. Monti, “Impact of pseudo-measurements from new power profiles on state estimation in low-voltage grids,” *IEEE Trans. Instrum. Meas.*, vol. 65, no. 1, pp. 70–77, 2016.
- [127] G. Larroux, M. Jacobs, K. Jia, F. Sossan, and M. Paolone, “Enhancing power systems transmission adequacy via optimal BESS siting and sizing using benders decomposition with feasibility cuts,” *arXiv preprint arXiv:2603.04189*, 2026, preprint; vérifier publication finale avant dépôt.
- [128] A. O. Rousis, D. Tzelepis, Y. Pipelzadeh, G. Strbac, C. D. Booth, and T. C. Green, “Provision of voltage ancillary services through enhanced TSO-DSO

interaction and aggregated distributed energy resources,” *IEEE Transactions on Sustainable Energy*, vol. 12, no. 2, pp. 897–908, 2021.

APPENDIX A.

Derivations for the Pi-Equivalent Line Model

A.1. Linearised Voltage Drop Relationship

This section derives the linearised voltage drop approximation (1.14) used in Section 1.2.1, from the pi-equivalent model of Fig. 1.4. Voltage and current are represented by their phasors $\underline{V}$ and $\underline{I}$, complex numbers encoding the amplitude and phase of a sinusoidal quantity. The series branch of the pi-model is represented by the complex impedance $\underline{Z} = R + jX$, where R is the line resistance and X its inductive reactance. This derivation neglects the shunt susceptances $B/2$ of the pi-model and retains only the series branch: under normal loading, the shunt branches draw a comparatively small current, so their effect on the series voltage drop is secondary, though they do contribute directly to the net reactive power exchanged by the line, as discussed in Section 1.2.2.

Applying Kirchhoff's voltage law along the series branch yields the fundamental sending-end to receiving-end voltage relationship:

$$\underline{V}_S = \underline{V}_R + \underline{Z} \underline{I} \tag{A.1}$$

Taking the receiving-end voltage as the phase reference, $\underline{V}_R$ is taken as real and positive, and the line current is expressed in terms of the active and reactive power delivered to the load:

$$\underline{I} = \frac{P - jQ}{V_R} \tag{A.2}$$

Appendix A. Derivations for the Pi-Equivalent Line Model

Substituting into (A.1):

$$\underline{V}_S = V_R + \frac{(R + jX)(P - jQ)}{V_R} = V_R + \frac{RP + XQ}{V_R} + j\frac{XP - RQ}{V_R} \quad (\text{A.3})$$

Under normal operating conditions in HV networks, the phase angle δ between $\underline{V}_S$ and $\underline{V}_R$ remains small. Since $\underline{V}_R$ was taken as the phase reference, this means the imaginary component of $\underline{V}_S$ is small compared to its real component, so that $|\underline{V}_S| \approx \text{Re}(\underline{V}_S)$; equivalently, the complex voltage drop $\underline{V}_S - \underline{V}_R$ is approximately aligned with the reference axis, so that its magnitude reduces to the difference of the two voltage magnitudes, $|\underline{V}_S - \underline{V}_R| \approx |\underline{V}_S| - |\underline{V}_R|$. Defining the voltage magnitude difference as $\Delta V = |\underline{V}_S| - |\underline{V}_R|$, this small-angle approximation gives:

$$\Delta V \approx \frac{RP + XQ}{V_R} \quad (\text{A.4})$$

which is (1.14) of Section 1.2.1.

A.2. Exact Power Flow Equations

This section derives the exact power flow equations underlying the curve of Fig. 1.5 in Section 1.2.2, obtained without the linearisation of Section A.1. Taking the sending end as the slack bus, $\underline{V}_S = V_S \angle 0$, and the receiving end as a PQ bus of unknown voltage $\underline{V}_R = V_R \angle \theta_R$, the bus admittance matrix of the pi-model of Fig. 1.4 is:

$$Y_{SS} = Y_{RR} = \frac{1}{R + jX} + j\frac{B}{2}, \quad Y_{SR} = Y_{RS} = -\frac{1}{R + jX} \quad (\text{A.5})$$

The complex power injected at bus R must balance the load $P_R + jQ_R$ demanded there:

$$\underline{V}_R (Y_{RS} \underline{V}_S + Y_{RR} \underline{V}_R)^* + (P_R + jQ_R) = 0 \quad (\text{A.6})$$

a system of two nonlinear equations in (V_R, θ_R) , solved numerically by a Newton-Raphson procedure. The reactive power supplied by the slack bus into the line then follows as $Q_{\text{line}} = \text{Im}\{\underline{V}_S(Y_{SS} \underline{V}_S + Y_{SR} \underline{V}_R)^*\}$. For the purely active loads considered here ($Q_R = 0$), Q_{line} coincides with the net reactive power exchanged between the line and the network.