U.PORTO
FEUP FACULDADE DE ENGENHARIA
UNIVERSIDADE DO PORTO

AI-based Decision Systems for Operational Planning and Real-time Power Grid Management

Ferinar Moaidi

Thesis submitted to the Faculty of Engineering of University of Porto in partial fulfilment of the requirements for the degree of Doctor of Philosophy in Electrical and Computer Engineering

Supervisor:
Ricardo Jorge Gomes de Sousa Bento Bessa
Principal Researcher at INESC TEC

July 2026

AI-based Decision Systems for Operational Planning and Real-time Power Grid Management

Thesis submitted to the Faculty of Engineering of the University of Porto in partial fulfilment of the requirements for the degree of Doctor of Philosophy in Electrical and Computer Engineering

Ferinar Moaidi

President of the Jury

Doctor Luís Miguel Pinho de Almeida, Full Professor of the Faculty of Engineering of the University of Porto.

Members

- Doctor Jochen Cremer, Associate Professor at the Faculty of Electrical Engineering, Mathematics, and Computer Science of the Delft University of Technology, Netherlands.
- Dr. Hugo Gabriel Valente Morais, Associate Professor at the Department of Electrical and Computer Engineering of the Instituto Superior Técnico of the University of Lisbon.
- Dr. José Nuno Moura Marques Fidalgo, Associate Professor with Aggregation at the Department of Electrical and Computer Engineering of the Faculty of Engineering of the University of Porto.
- Dr. Ricardo Jorge Gomes de Sousa Bento Bessa, Coordinating Researcher of the Centre for Power and Energy Systems (CPES) of the Institute for Systems and Computer Engineering, Technology and Science (INESC TEC), Supervisor.

Assinado por: **Ricardo Jorge Gomes de Sousa Bento Bessa**
Num. de Identificação: 12351548
Data: 2026.07.14 10:08:42+01'00'

Dr. Ricardo Jorge Gomes de Sousa Bento Bessa
Supervisor

You cannot force ideas. Successful ideas are the result of slow growth. Ideas do not reach perfection in a day, no matter how much study is put upon them.

Alexander Graham Bell, 1901

Acknowledgments

The completion of this thesis marks a significant milestone in both my academic and personal journey. I would like to take this moment to express my sincere appreciation to all those who have guided, supported, and encouraged me throughout this path.

I would like to express my deepest gratitude to my supervisor, Prof. Ricardo Bessa, who has been far more than a mentor. As an international student facing numerous challenges, I was fortunate to receive not only his constant support and encouragement but also the benefit of his profound knowledge and exceptional scientific expertise. His guidance, patience, and insightful advice have been invaluable in shaping this research and have deeply influenced my growth as both a scholar and a person.

I sincerely thank Prof. José Nuno Fidalgo and Prof. Hugo Morais for their valuable comments during annual evaluations, which greatly contributed to the refinement of this work.

My heartfelt thanks go to the members of INESCTEC. In particular, I am profoundly grateful to Paula Castro, whose kindness and care became like a mother's embrace to me, offering warmth and unwavering support that carried me through some of the most difficult and challenging moments of this journey. I would also like to express my appreciation to Vera Pinto and Andreia Passos from the International Relations Office. Their invaluable assistance during my arrival and administrative procedures ensured a smooth transition, allowing me to focus fully on my studies.

I would like to thank my colleagues at CPES for their support and for fostering a collaborative environment. I am grateful for the opportunity to cooperate and learn from José Ricardo Andrade, Micael Filipe Simões, Carla S. Gonçalves, João Viana, Hodjat Mariji, and Luís Brito Teixeira throughout this work.

Lastly, my deepest love and gratitude go to my parents, whose enduring love and support have carried me even across the miles that often lie between us. The physical distance never lessened the warmth of their encouragement or the strength of their belief in me. To my sister, Helia, whose messages, care, and constant presence—from far away—lifted my spirits more than she knows, I am endlessly thankful. Their love has been my anchor throughout this journey, and this achievement belongs as much to them as it does to me.

The research leading to this work was carried out as part of two European projects: Smart4RES and AI4RealNet (*AI for REAL-world NETwork operation*), European Union's Horizon Research and Innovation Programme, Grant Agreements No. 864337 and 101119527, respectively. Views and opinions expressed are, however, those of the author(s) only and do not necessarily reflect those of the European Union. Neither the European Union nor the granting authority can be held responsible for them.

This research was made possible through the primary institutional support of INESC TEC, through CPES, which provided funding, physical facilities where the majority of the work was conducted, including a dedicated workstation and meeting rooms that facilitated discussions with supervisors and colleagues, alongside essential IT services, and cooperation in EU projects, as well as FCT (*Fundação para a Ciência e a Tecnologia*), which supported this research through PhD grant 2023.04869.BD.

United Nations Sustainable Development Goals

This thesis contributes directly to three of SGDs by supporting the reliable, efficient, and transparent operation of decarbonized power systems. In particular, the proposed methods advance **SGD 7 (Affordable and Clean Energy)** by enabling higher integration of renewable energy sources through improved flexibility procurement, adaptive topology control, and optimized redispatch decisions.

In addition, the development of hybrid and interpretable artificial intelligence frameworks for real-time grid operation supports **SGD 9 (Industry, Innovation and Infrastructure)** by strengthening the resilience and digitalization of critical energy infrastructure. The emphasis on explainable and human-centered decision-support tools further aligns with **SGD 12 (Responsible Consumption and Production)** by reducing inefficiencies and supporting responsible system operation under uncertainty.

Abstract

The global transition toward decarbonized power systems introduces significant challenges for grid operation and market management, driven by increasing renewable energy penetration and widespread electrification. These developments reduce system capacity margins and amplify uncertainties in load and renewable generation forecasts, demanding novel approaches to operational flexibility, decision-making, and explainability.

This thesis addresses these challenges through a series of complementary contributions that integrate predictive algorithms, hybrid artificial intelligence, and interpretable decision-support frameworks for modern power systems. The first contribution proposes a predictive algorithm for the procurement of grid flexibility by the SO, combining SO-controlled flexible assets with active and reactive power short-term flexibility markets. By reducing the cognitive load on human operators and implementing a "time-to-decide" method, the framework allows operators to balance risk, cost, and operational stakes when deciding whether to book flexibility immediately or wait for updated forecasts. Numerical results on a public medium-voltage grid (Oberrhein) show that this approach improves a cost-loss performance indicator by up to 22% compared to immediate flexibility booking strategies.

The second contribution introduces a hybrid AI framework for power grid topology control, combining GNP, reinforcement learning, and decision trees. A novel variant of GNP is developed to evolve decision-making rules from operational data, providing transparent and traceable reasoning while managing N-1 contingencies in real-time grid operations. The method outperforms both a baseline expert system and a state-of-the-art deep reinforcement learning agent on the IEEE 118-bus system, achieving up to a 28% improvement in a key performance metric from the L2RPN competition.

The third contribution addresses the interpretability gap in AC-OPF redispatch solutions. An interpretable and adaptive framework is proposed that combines post-optimal physical analysis with a multi-agent large language model architecture to generate concise, operator-oriented explanations of AC-OPF redispatch decisions. The proposed framework uses physically grounded information, including dual variables, linearized sensitivities, and line outage distribution factors, to support counterfactual reasoning about constraint drivers, feasible redispatch alternatives, and their approximate economic and security implications. A reflective multi-agent loop composed of an Actor, Critic, Distiller, and Refiner iteratively improves the explanation process by extracting compact operational rules and integrating them into the Actor prompt through a sliding-window refinement mechanism. Evaluation on the test system shows that the evolved framework produces explanations that are physically consistent, numerically defensible, causally coherent, and highly readable, while also generalizing effectively to unseen operating snapshots. With low computational overhead, the proposed approach enhances sit-

uational awareness by translating AC-OPF outputs into transparent and operationally meaningful redispatch explanations.

Together, these contributions advance the state-of-the-art in flexible, AI-driven, and interpretable power system operation. By integrating predictive flexibility procurement, hybrid AI-based topology control, and interpretable AC-OPF reasoning, this thesis provides tools that improve operational efficiency, support human-in-the-loop decision-making, and enhance trust in automated grid management under uncertainty.

Resumo

A transição global para sistemas de energia descarbonizados introduz desafios significativos para a operação da rede e a gestão do mercado, impulsionados pelo aumento da penetração de fontes renováveis e pela ampla eletrificação da demanda. Esses desenvolvimentos reduzem as margens de capacidade do sistema e ampliam as incertezas nas previsões de carga e geração renovável, exigindo novas abordagens para flexibilidade operacional, tomada de decisão e interpretabilidade.

Esta tese aborda esses desafios por meio de uma série de contribuições complementares que integram algoritmos preditivos, inteligência artificial híbrida e estruturas de apoio à decisão interpretáveis para sistemas de energia modernos. A primeira contribuição propõe um algoritmo preditivo para a aquisição de flexibilidade da rede pelo operador do sistema (OS), combinando ativos flexíveis controlados pelo OS com mercados de flexibilidade de potência ativa e reativa de curto prazo. Ao reduzir a carga cognitiva dos operadores humanos e implementar o método "tempo-para-decidir", a estrutura permite que os operadores equilibrem risco, custo e impactos operacionais ao decidir se devem reservar flexibilidade imediatamente ou aguardar atualizações de previsão. Resultados numéricos em uma rede de média tensão pública (Oberrhein) mostram que essa abordagem melhora um indicador de desempenho custo-perda em até 22% em comparação com estratégias de reserva imediata de flexibilidade.

A segunda contribuição apresenta uma estrutura de IA híbrida para controle da topologia da rede elétrica, combinando Programação em Rede Genética (GNP), aprendizado por reforço e árvores de decisão. Uma nova variante de GNP foi desenvolvida para evoluir regras de tomada de decisão a partir de dados operacionais, proporcionando raciocínio transparente e rastreável enquanto gerencia contingências N-1 em operações de rede em tempo real. O método supera tanto um sistema especialista de referência quanto um agente de aprendizado profundo de última geração no sistema IEEE de 118 barras, alcançando até 28% de melhoria em um indicador de desempenho chave da competição Learning to Run a Power Network (L2RPN).

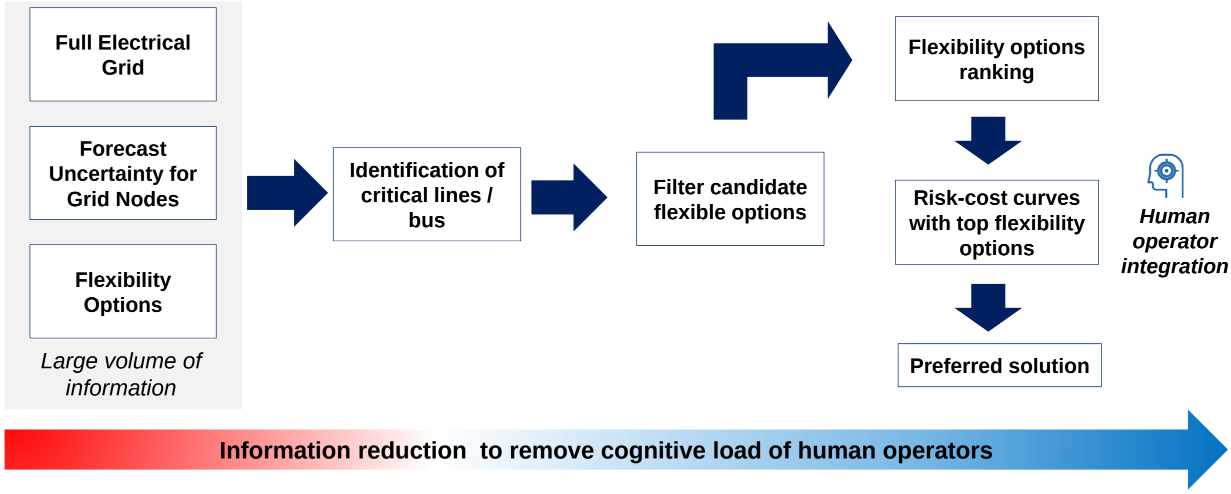
A terceira contribuição aborda a lacuna de interpretabilidade nas soluções de redespacho do fluxo de potência ótimo em corrente alternada (AC-OPF). Propõe-se um framework interpretável e adaptativo que combina análise física pós-otimização com uma arquitetura de modelo de linguagem de grande escala multiagente para gerar explicações concisas e orientadas ao operador sobre decisões de redespacho do AC-OPF. O framework proposto utiliza informações fisicamente fundamentadas, incluindo variáveis duais, sensibilidades linearizadas e fatores de distribuição de desligamento de linhas (Line Outage Distribution Factors – LODF), para apoiar raciocínio contrafactual sobre os fatores que acionam restrições, alternativas viáveis de redespacho e suas implicações econômicas e de segurança aproximadas. Um ciclo reflexivo mul-

tiagente composto por Actor, Critic, Distiller e Refiner melhora iterativamente o processo de explicação ao extrair regras operacionais compactas e integrá-las ao prompt do Actor por meio de um mecanismo de refinamento com janela deslizante. A avaliação no sistema de teste mostra que o framework evoluído produz explicações fisicamente consistentes, numericamente defensáveis, causalmente coerentes e altamente legíveis, além de generalizar de forma eficaz para estados operacionais não vistos. Com baixo custo computacional, a abordagem proposta aumenta a consciência situacional ao traduzir as saídas do AC-OPF em explicações de redespacho transparentes e operacionalmente significativas.

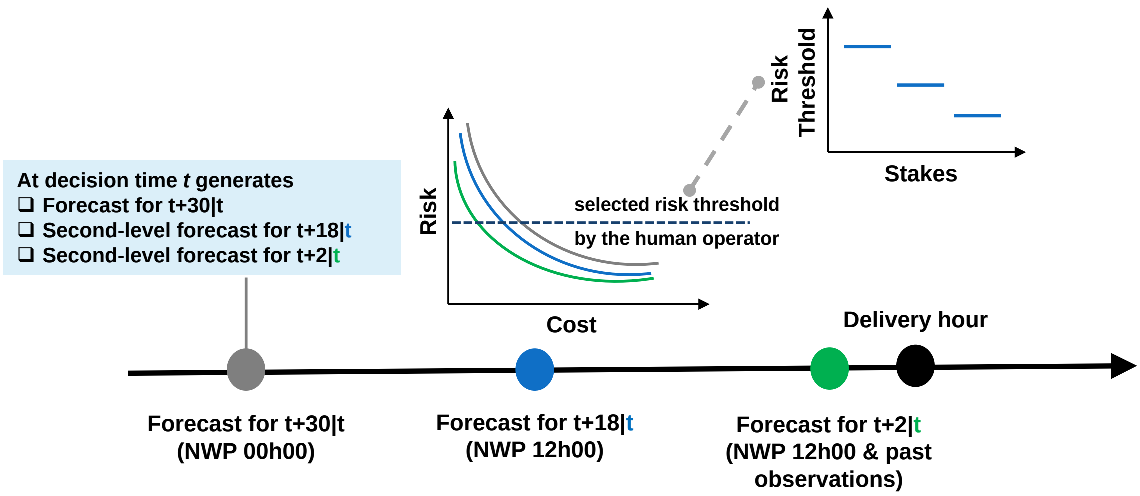
Juntas, essas contribuições avançam o estado da arte na operação de sistemas de energia flexível, baseada em IA e interpretável. Ao integrar aquisição preditiva de flexibilidade, controle de topologia baseado em IA híbrida e raciocínio interpretável para AC-OPF, esta tese fornece ferramentas que melhoram a eficiência operacional, apoiam a tomada de decisão humano-no-loop e aumentam a confiança na gestão automatizada da rede sob incerteza.

Graphical Abstract

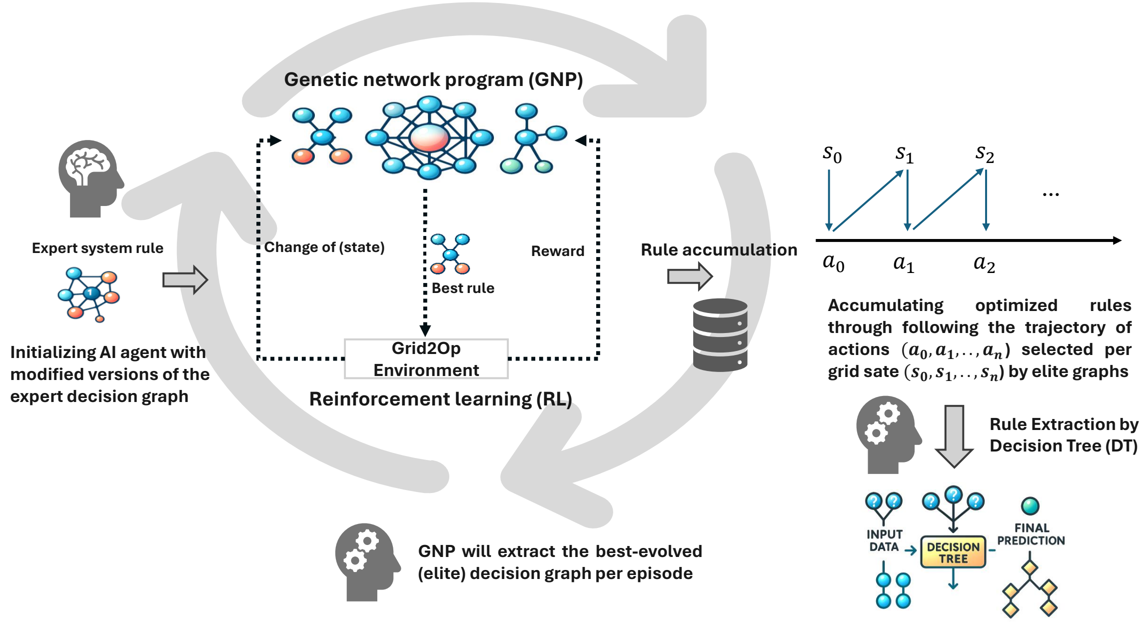
(a) Predictive Management System Architecture



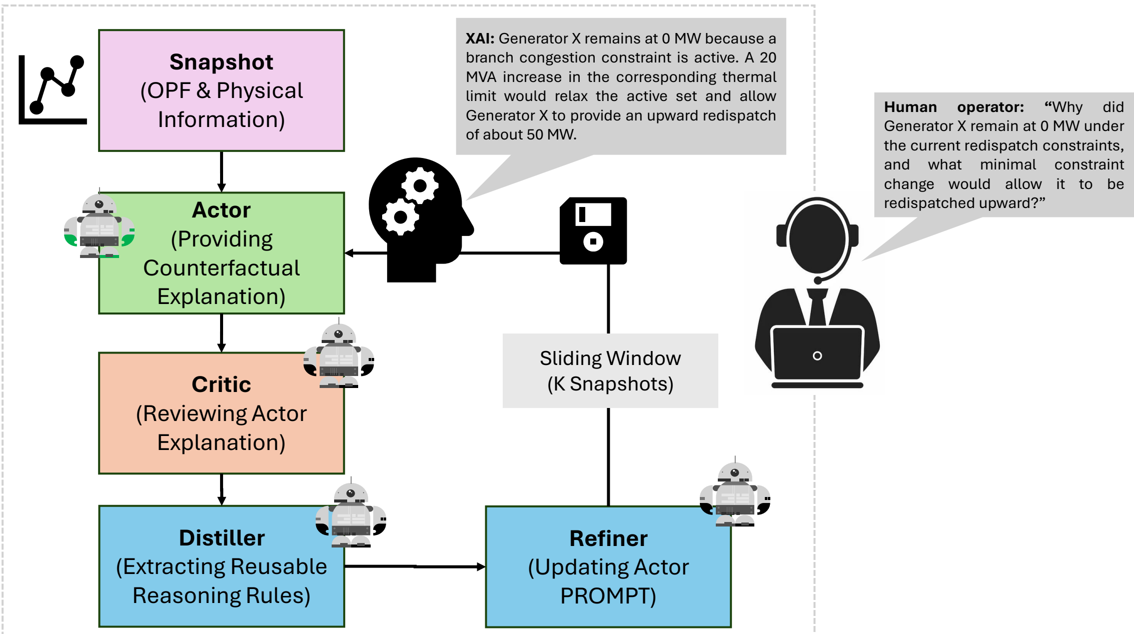
(b) Second-level Forecasting Methodology



(c) Evolving Human Operator Rules



(d) Large Language Model Optimization Explainer



Contents

List of Figures

List of Tables

List of Algorithms

List of Acronyms and Symbols

AC	Alternate Current
AC-OPF	AC Optimal Power Flow
AI	Artificial Intelligence
CAgent	CurriculumAgent
CB	Capacitor Bank
CBM	Characteristic Block Matrix
CC	Chance Constrained
CHP	Combined Heat and Power
CPES	Center for Power and Energy Systems
D1	Deterministic 1
D2	Deterministic 2
DC	Direct Current
DC-OPF	DC Optimal Power Flow
DDQN	Double Deep Q-Network
Deep-RL	Deep Reinforcement Learning
DER	Distributed Energy Resources
DF	Distribution Factor
DM	Decision-Maker
DN	Decision-Now
DR	Demand Response
DSO	Distribution System Operator
ECMWF	European Centre for Medium-Range Weather Forecasts
ED-ANN	Encoder-Decoder Artificial Neural Network

ENS	Ensemble Forecasts
ES	Expert Systems
FSP	Flexibility System Provider
FTR	Financial Transmission Rights
GBT	Gradient Boosting Trees
GNP	Genetic Network Programming
GNP-DT	Genetic Network Programming and Decision Tree
IEEE	Institute of Electrical and Electronics Engineers
IM	Incidence Matrix
INESC TEC	Instituto de Engenharia de Sistemas e Computadores, Tecnologia e Ciência
L2RPN	Learning to Run a Power Network
LLM	Large Language Model
LODF	Line Outage Distribution Factor
LSTM	Long Short-Term Memory
MILP	Mixed-Integer Linear Programming
MIP	Mixed Integer Programming
MSE	Mean Squared Error
MSTCM	Median of Survival Times Across Median
NIS	Negative Ideal Solution
NWP	Numerical Weather Predictions
OLTC	on-load tap changer
OPF	Optimal Power Flow
OTS	Optimal Transmission Switching
PIS	Positive Ideal Solution
POA	Physical Operational Analysis
PPO	Proximal Policy Optimization
PQ	Load Buses
PTDF	Power Transfer Distribution Factor
PV	Generator Buses

RES	Renewable Energy Source
RL	Reinforcement Learning
RS	Reactance Shunt
SGD	United Nations Sustainable Development Goal
SI	Sensitivity Indices
SO	System Operator
T2D	Time-to-Decide
TP	True Positives
TSO	Transmission System Operator
VaR	Value-at-Risk
VoLL	Value-of-Lost-Load
WCCI	World Congress on Computational Intelligence
WPP	Wind Power Plants
XAI	Explainable AI

$\boldsymbol{A}$	Submatrix of the compound Jacobian matrix of the network
α	Nominal proportion of the conditional quantile, with $0 < \alpha < 1$
a_t	Action taken at timestep t
A_θ	Actor agent parameterized by θ, mapping snapshots to operator-facing explanations
$\boldsymbol{B}$	Submatrix of the compound Jacobian matrix of the network
$b^s_{j\to i}$	Indicator variable for redispatch action direction, equal to $+1$ for upward and -1 for downward action
β_{th}	Threshold coefficient for determining influence path
B_{ik}	Susceptance element of the bus admittance matrix between buses i and k
$\boldsymbol{b}_{sh}$	Vector of shunt susceptances
C	Flexibility action cost (preventive cost) in cost-loss analysis
$C(\mathbf{P_g})$	Total generation cost function, often quadratic or piecewise-linear
C^{AI}	AI critic or evaluator producing intermediate assessment from Actor explanation
$C_{\text{blackout}}(t)$	Blackout or unserved load cost at timestep t

$C_{\text{dispatch}}(t)$	Dispatch cost at timestep t
$C_{\text{flex}}(t)$	Flexibility cost at timestep t
c_i	Marginal generation cost of generator i
C_j	Contribution of grid node j forecast uncertainty to the technical problem in grid element i
C_i^{load}	Load curtailment cost at bus i
$C_{\text{losses}}(t)$	Power loss cost at timestep t
$C_{\max}$	Worst-case total cost upper bound
$C_{\text{op}}(t)$	Per-timestep operational cost
cp_1	Probability for exploitation-driven crossover
cp_2	Probability for exploration-driven crossover
C_i^{res}	Renewable curtailment cost at generator i
$\boldsymbol{D}$	Submatrix of the compound Jacobian matrix of the network
$\boldsymbol{\delta}_{sh}$	Matrix whose elements are zero except at buses connected to shunts where they take values ± 1
δ	Sensitivity index associated with a flexible resource
$\delta_{i \to j,s}^{t+k}$	Sensitivity index for grid element i and flexible resource j for scenario s at lead-time $t+k$
$\delta_{\{j_1,j_2,\ldots,j_n\}\to i}^{s}$	Combined sensitivity index of multiple resources participating in redispatch for grid element i under scenario s
$\delta_{sb\to v}^{s}$	Sensitivity of slack bus sb to grid element v under scenario s
d_m^-	Geometric distance of solution m from the Negative Ideal Solution (NIS) in TOPSIS
d_m^+	Geometric distance of solution m from the Positive Ideal Solution (PIS) in TOPSIS
$\boldsymbol{E}$	Complex bus voltage vector
$e_{t+k\|t}$	Random shock term associated with the second-level forecast
$Eqcost$	Equivalent cost of a flexibility solution considering trade-off between risk and cost
e_t	Generated explanation for snapshot t
$\boldsymbol{F}$	Matrix defined as $\underline{\boldsymbol{E}}_D \cdot \frac{\partial \bar{\boldsymbol{Y}}}{\partial \boldsymbol{b}_{sh}} \cdot \bar{\boldsymbol{E}}$
f	Second-level forecasting model for each conditional quantile
$\boldsymbol{F}^d$	Real part of the matrix $\boldsymbol{F}$
$Flex_j^{down}$	Downward flexibility margin offered by flexible resource j
$Flex_j^{up}$	Upward flexibility margin offered by flexible resource j

$f_{n_t}(s_t)$	Decision rule implemented at node n_t
f_{n_t}	Decision rule implemented at node n_t
$\boldsymbol{F}^q$	Imaginary part of the matrix $\boldsymbol{F}$
$FV^s_{j\to i}$	Flexibility value required from resource j to mitigate the technical issue in grid element i under scenario s
$FV'^s_{j\to i}$	Limited flexibility value considering operational constraints of other grid elements
G_i	Individual GNP graph or agent in the population
$\mathbf{g}(\mathbf{P_g}, \mathbf{Q_g}, \mathbf{V}, \boldsymbol{\theta})$	Equality constraints enforcing AC power flow balance at each bus
$\boldsymbol{\Gamma}_t$	Set of linearized sensitivities at time t relating injections to line currents and bus voltages
γ	Performance indicator computed as weighted sum of cost-loss matrix elements by percentage of occurrence
G_i	The i-th individual graph representing a policy μ_i
G_{ik}	Conductance element of the bus admittance matrix between buses i and k
G^*	Elite graph with highest fitness in population
$\mathbf{h}(\mathbf{P_g}, \mathbf{Q_g}, \mathbf{V}, \boldsymbol{\theta})$	Inequality constraints capturing operational limits
I_{th}	Overload threshold to trigger action proposals
k	Forecast horizon index
L	Loss cost associated with emergency actions, e.g., load/generation curtailment, in cost-loss analysis
l	Number of covariates in the feature vector
λ	Operational quantity of a grid element (e.g., voltage magnitude or current loading)
λ^s_i	Forecasted operational value of grid element i under scenario s
λ^{lim}	Operational limit of a grid element
λ^{max}	Maximum operational limit of a grid element
λ^{min}	Minimum operational limit of a grid element
$\lambda^{\%}_i$	Line loading percentage of line i
λ'^s_i	Line loading or bus voltage of element i after flexibility activation under scenario s
λ^s_v	Forecasted operational value of grid element v under scenario s
$\lambda_{\max}$	Maximum mutation range as fraction of gene length

M	Total number of buses in the network
$\mathscr{M}$	Memory storing distilled rules and associated quality scores for multi-agent LLM framework
m_p	Mutation probability
μ	Trade-off parameter between risk and cost, determined by decision-maker preferences
μ_i	Policy encoded by graph G_i
N	Number of load buses
$\text{Next}(n_t)$	Deterministic next node pointer of node n_t
NIS	Negative Ideal Solution in TOPSIS multi-criteria decision making
N_{sh}	Number of shunt elements
n_t	Node visited at timestep t in a GNP graph
$\boldsymbol{O}_{sh}$	$N \times N_{sh}$ zero matrix associated with shunts
$P(s_{t+1} \mid s_t, a_t)$	State transition function
$\mathbf{P_g}$	Active power outputs of all generators
$\hat{\mathbf{P}}_s^{t+k}$	Forecasted active power nodal injection vector for scenario s at lead-time $t+k$
Φ_i	Aggregated magnitude of slack variables for power balance, voltage, and line constraints
$\phi(W_t)$	Rule aggregation function over sliding window W_t
$\pi(a_t \mid s_t)$	Policy mapping states to action probabilities
π_i	Trajectory generated by graph G_i: $\pi_i = (s_0, a_0, r_0, s_1, \dots)$
π_{trans}	Stochastic policy over node transitions
$\pi_{\text{trans}}(n_{t+1} \mid n_t, s_t)$	Stochastic node transition policy in adaptive GNP
PIS	Positive Ideal Solution in TOPSIS multi-criteria decision making
P_ℓ	Active power flow on line ℓ
$\mathscr{P}$	Space of admissible Actor prompts preserving task structure
$p^\star$	Optimal Actor prompt maximizing explanation-quality functional over horizon T
p_t	Current Actor prompt at time t
$\mathbf{Q_g}$	Reactive power outputs of all generators
$\hat{\mathbf{Q}}_s^{t+k}$	Forecasted reactive power nodal injection vector for scenario s at lead-time $t+k$

$\hat{q}^{\alpha}_{t+k\|t}$	First-level quantile forecast with nominal proportion α issued at time t for time interval $t+k$
$\hat{q}^{\alpha}_{n,t+k\|t+z}$	Quantile forecast generated at time $t+z$ for the same time interval $t+k$
Q_ℓ	Reactive power flow on line ℓ
$\frown{q}^{\alpha}_{t+k\|t}$	Second-level forecast of the conditional quantile corresponding to nominal proportion α, generated at time t for time interval $t+k$
Q^u_j	Reactive power step size of shunt element j
$R(s_t, a_t)$	Reward function
$r(s,a)$	Reward function mapping states and actions to a scalar
R_{cost}	Normalized cost-aware reward
r_t	Reward or evaluation signal fed to the Prompt Refiner at time t
S	Number of forecast scenarios
SB	Number of slack buses
Severity_i	Severity function associated with line congestion or voltage violation for element i
σ^1_{cp}	Priority indicator for highest-priority actions
σ_t	Topological score indicating quality of actions
S^{slack}_ℓ	Slack variable capturing thermal limit violations on line ℓ
$\mathscr{S}_t$	System snapshot at time t, containing pre/post-redispatch states, dual variables, sensitivities, and LODFs
s_t	System state at timestep t
stp^{max}_j	Maximum allowed step-change for shunt element j
$\boldsymbol{SV}_{a,sh}$	Voltage angle sensitivity matrix with respect to shunt variations
$\boldsymbol{SV}_{m,sh}$	Voltage magnitude sensitivity matrix with respect to shunt variations
SWS_m	Similarity to worst solution coefficient for solution m in TOPSIS, used for ranking flexibility options
t	Forecast launch time
$tap^{[\%]}$	Tap percentage representing the voltage change per tap position
τ	Trajectory under policy π, $\tau = (s_0, a_0, s_1, \ldots, s_T)$
θ_n	Heuristic parameters of node n
θ_{ik}	Voltage angle difference between buses i and k
θ_n	Heuristic parameters of node n updated from local feedback
$tpos$	Tap position of the OLTC transformer
V	Set of grid elements considered when limiting flexibility actions

$\mathbf{v}_t$	Intermediate evaluation output produced by C^{AI}
VaR	Value-at-Risk of a metric, e.g., flexibility cost or severity
$v_{M\times N}$	Weighted decision matrix in TOPSIS computed as $\hat{X}_{M\times N} \cdot W_{N\times 1}$
$V(\text{set-point})_j^s$	Voltage set-point of OLTC j under scenario s
$V(\text{set-point})_j^{ne}$	Voltage set-point of OLTC j at neutral tap position
$W_{N\times 1}$	Weighting vector for the criteria in TOPSIS
W_t	Sliding window of the last K rule sets used by the Prompt Refiner
$\boldsymbol{X}$	Partial derivative of the voltage vector with respect to shunt susceptance
$X^{\alpha}_{t+k\|t}$	Vector of covariates used by the second-level forecasting model, including quantile forecasts, engineered uncertainty features, and exogenous variables
$\acute{X}_{M\times N}$	Normalized decision matrix in TOPSIS denoted as $(\acute{x}_{mn})_{M\times N}$
$\acute{x}_{mn}$	Element of the normalized decision matrix $\acute{X}_{M\times N}$ for alternative m and criterion n
$\hat{X}_{M\times N}$	Normalized version of the decision matrix $X_{M\times N}$ in TOPSIS
$X_{M\times N}$	Decision matrix with M alternatives and N criteria
x_{mn}	Element of the decision matrix $X_{M\times N}$ for alternative m and criterion n
$\mathbf{x}_t^{\text{opf}}$	Post-redispatch optimized state at time t
$\mathbf{x}_t^{\text{pre}}$	Pre-redispatch state at time t
$\bar{\boldsymbol{y}}$	Admittance matrix of shunts in per unit
$\mathbf{Y}_{\text{bus}}$	Bus admittance matrix
Y_{bus}	Bus admittance matrix
$\hat{y}^s_{t+k\|t,j}$	Forecasted nodal injection for node j under scenario s at lead-time $t+k$
$\hat{y}^*_{t+k\|t,j}$	Point forecast of nodal injection for node j at lead-time $t+k$
z	Forecast update rate expressed in number of time intervals
Z_{bus}	Bus impedance matrix of the power network

Chapter 1

Introduction

1.1 Motivation

The global energy transition has brought unprecedented levels of RES integration into modern power systems. Driven by the urgency to reduce greenhouse gas emissions and achieve energy independence, many countries have set ambitious renewable generation targets. While these developments are crucial for sustainable energy, high RES penetration introduces significant operational challenges. The variability and uncertainty of RES, together with growing electrification of demand, increase congestion, RES curtailment, and transmission constraints. For example, wind energy dispatch-down in Ireland rose from 5.3% to 11.4%, RES curtailment in Chile increased by 30% due to inflexible thermal generation, and Germany continues to face transmission bottlenecks caused by delayed reinforcements [1]. These developments reveal that traditional deterministic planning and rule-based operational procedures are increasingly insufficient to manage the evolving complexity of modern grids [2, 3].

In response, predictive flexibility management has emerged as a key strategy to anticipate challenges before they occur. By leveraging forecasts of load and RES generation, system operators can plan the procurement of flexible resources—both active and reactive—to handle anticipated congestion and voltage issues. This anticipatory approach reduces cognitive load, supports risk-aware decision-making, and allows operators to balance cost, reliability, and resource availability [4, 5]. Predictive management shifts the focus from reactive interventions to proactive, data-driven planning, enabling the grid to accommodate high RES penetration more efficiently and reliably.

Despite these predictive measures, real-time operation still requires corrective actions. Operators must assess the system state continuously and implement remedial measures—such as generator redispatch, RES curtailment, demand response, or network topology reconfiguration—especially under contingency conditions. The combinatorial space of possible actions, coupled with strict time constraints, makes identifying optimal solutions extremely challenging. Traditional Expert Systems (ES), while structured and rule-based, struggle to scale under these high levels of uncertainty and complexity. Integrating RL into ES allows adaptive policies to be learned from experience, improving operational performance while maintaining transparency. This hybrid approach builds on Markov decision process formulations, as exemplified in the L2RPN competition [6], and provides operators with actionable guidance for timely, effective interventions.

A further challenge arises from the interpretability of optimization-based solutions. AC Optimal Power Flow (AC-OPF) is widely used for redispatch and congestion relief, but its outputs are largely opaque. While solvers provide generator setpoints and dual variables, they rarely explain why specific generators are adjusted or how constraints drive the solution. Operators often perceive AC-OPF as a “black box,” limiting situational awareness and reducing confidence in decision-making, particularly in high-stakes scenarios. Enhancing interpretability is therefore essential to bridge the gap between automated optimization and human reasoning, enabling operators to understand, validate, and act upon system recommendations effectively.

Together, these challenges underscore the need for comprehensive frameworks that integrate predictive anticipation, adaptive corrective actions, and interpretable decision-making. By anticipating future issues through forecast-informed flexibility management, enabling responsive corrective actions via hybrid Artificial Intelligence (AI)-enhanced ESs, and providing transparent explanations of optimization results, such frameworks can support operators in managing complex, uncertain, and high-stakes scenarios. Ultimately, this approach ensures that modern power systems can operate securely, efficiently, and resiliently, even under high levels of renewable energy penetration, bridging the gap between automation and human expertise. As illustrated in Figure 1.1, the modern power system operation workflow can be conceptually divided into three main stages. The first stage, **Predictive Management**, involves anticipating potential issues in the grid using forecasts of load and renewable generation, and planning the procurement of flexible resources to mitigate congestion and voltage problems. The second stage, **Real-Time Decision Making**, represents the execution of corrective actions based on the actual system state, including generator redispatch, demand response, renewable curtailment, or network topology adjustments. The final stage, **Interpretability**, ensures that the operator can understand the rationale behind decisions and optimization outputs, providing transparent and actionable insights for human-in-the-loop operation. The horizontal pipeline in Figure 1.1 emphasizes the sequential and integrated nature of these stages, showing how anticipatory planning feeds into immediate decision-making, which is then explained and analyzed to inform future operational strategies.

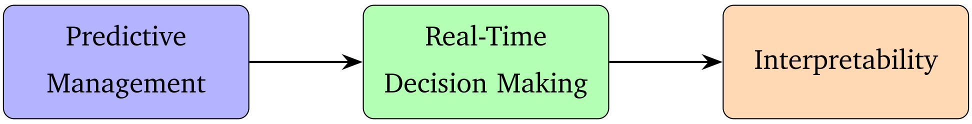


Figure 1.1: *Pipeline workflow of modern power system operation: Predictive Management → Real-Time Decision Making → Interpretability.*

1.2 Research Questions, Objectives and Contributions

The overarching objective of this thesis is to enhance the security, reliability, and human interpretability of modern power system operations under high renewable penetration. Three key research gaps were identified, motivating the following research questions, objectives, and contributions:

Objective 1. To Develop a Predictive Flexibility Management Framework for Human-in-the-Loop Decision-Making.

Research question: *How* can predictive models and risk-aware metrics be used to anticipate flexibility needs and assist human operators in multi-criteria decision-making?

Contributions:

1– A predictive framework that ranks flexibility options using sensitivity indices and risk-cost curves, enabling human operators to explore and adjust decisions according to stakes.

2– Introduction of a *time-to-decide* formulation for dynamic decision-making under frequent forecast updates.

3– Novel *second-level forecasting* model to estimate uncertainty in future forecasts using an encoder-decoder deep learning architecture.

Objective 2. To Design a Real-Time Control Framework Based on GNP and Decision Trees.

Research question: *Can* adaptive GNP-based architectures provide context-aware, interpretable real-time control decisions for uncertain power system environments?

Contributions:

1– Development of a hybrid Genetic Network Programming and Decision Tree (GNP-DT) methodology with dynamic nodes that adapt based on evolving network states.

2– Extraction of human-readable rules from GNP decision trajectories using multistage decision trees, enabling modular interpretability.

3– Demonstration of superior performance compared to baseline ESs and Deep Reinforcement Learning (Deep-RL) agents on the IEEE 118-bus system (up to 18% improvement in L2RPN score and 80% reduction in inference time).

Objective 3. To Enhance the Interpretability of Power System Redispatch Decisions via Physical Operational Analysis and LLM Reasoning.

Research question: *How* can optimization outputs, such as AC-OPF redispatch solutions, be transformed into physically meaningful, human-understandable explanations?

Contributions:

1– Introduction of counterfactual sensitivity analysis to provide local *what-if* and *why-not* insights for redispatch actions.

2– Development of an Actor–Critic–Distiller–Refiner reasoning loop integrating physical information and LLM feedback, ensuring consistency with power system physics.

3– Implementation of a verbal reasoning loop to improve interpretability and consistency across multiple redispatch analyses.

While the thesis focuses on power system operations under high renewable penetration, some contributions have broader applicability. In particular:

- **Decision Support Systems:** The predictive and interpretable frameworks can enhance human-in-the-loop decision-making in other energy systems and industrial processes.

- **Adaptive Control Systems:** GNP-based adaptive control strategies can be applied to general control problems requiring interpretable, context-aware decisions.

- **Rule-Based Reasoning and Diagnosis:** The interpretable LLM reasoning loop and post-optimization analysis can inform expert decision-makers in domains such as manufacturing, healthcare, or infrastructure management.

1.3 Structure of the Thesis

This thesis is organized into six chapters, each addressing key aspects of modern power system operation under high renewable penetration and uncertainty.

Chapter 2: Literature Review. This chapter provides a comprehensive overview of the state-of-the-art in power system predictive management, real-time control, and interpretability of optimization-based decisions. It covers methodologies for flexibility procurement, expert and data-driven control strategies, and emerging approaches for interpretable artificial intelligence in power systems. The review highlights current gaps in predictive decision support, real-time adaptive control, and human-understandable redispatch reasoning, establishing the motivation for the novel contributions presented in this thesis.

Chapter 3: Uncertainty-Aware Procurement of Flexibilities for Electrical Grid Operational Planning. This chapter introduces a predictive framework for grid flexibility procurement that integrates sensitivity analysis, risk metrics, and second-level forecasting. The methodology allows operators to anticipate potential system constraints, rank flexibility options, and make risk-informed decisions. A novel *time-to-decide* approach is presented to account for dynamic forecast updates, demonstrating improved operator support and cost-effective planning through numerical experiments on realistic power system models.

Chapter 4: Evolving Power System Operating Rules for Real-Time Congestion Management. In this chapter, a hybrid GNP-DT framework is proposed for real-time congestion management. By incorporating dynamic, context-aware nodes and rule extraction through decision trees, the framework enables interpretable, adaptive decision-making under uncertainty. Extensive simulations on benchmark systems illustrate how the approach outperforms traditional ESs and deep RL agents in both control performance and human interpretability.

Chapter 5: Bridging Optimization and Operator Insight: LLM-Assisted Counterfactual Analysis of AC-OPF. This chapter focuses on enhancing redispatch interpretability by combining Post Optimization Analysis with LLM-driven reasoning. Counterfactual sensitivity analysis is used to translate AC-OPF outputs into actionable, physically consistent explanations, while the Actor–Critic–Distiller–Refiner reasoning loop ensures interpretive validity. The chapter demonstrates how integrating LLM feedback and verbal reasoning provides operators with transparent insights into constraint drivers, the redispatch rationale, and system behavior.

Chapter 6: Conclusion and Future Work. The final chapter summarizes the main contributions of the thesis, reflecting on the improvements in predictive management, real-time control, and redispatch interpretability. It also discusses limitations, lessons learned, and potential extensions of the proposed frameworks. Directions for future research are outlined, including further integration of human-in-the-loop learning, scaling methodologies to larger power systems, and exploring additional applications in other complex, high-stakes decision-making domains.

1.4 Publications

During the course of this PhD research, several works resulting from this thesis have been published in high-impact venues. These publications reflect the contributions of the thesis in predictive grid management, real-time decision-making, and interpretable redispatch of power systems.

Journal Publications

1. F. Moaidi, R. J. Bessa, Evolving power system operator rules for real-time congestion management," Energy and AI, vol 23, pp. 100672, Jan, 2026, doi: 10.1016/j.egyai.2025.100672. (Q1, Impact Factor: 9.7)
   This publication introduces the hybrid GNP-DT framework for real-time congestion management. The paper emphasizes context-aware decision-making, interpretable policy extraction, and performance improvements over baseline expert and Deep-RL systems.

2. R. J. Bessa, F. Moaidi, J. Viana and J. R. Andrade, "Uncertainty-Aware Procurement of Flexibilities for Electrical Grid Operational Planning," in IEEE Transactions on Sustainable Energy, doi: 10.1109/TSTE.2023.3305865. (Q1, Impact Factor: 10.3)
   This paper presents the predictive flexibility management framework, integrating sensitivity indices, risk-cost curves, and second-level forecasting for human-in-the-loop operational planning. It demonstrates improved decision-making under uncertainty and highlights the novel *time-to-decide* methodology for dynamic environments.

Conference Publication

1. F. Moaidi and R. J. Bessa, "Understanding OPF Redispatch Actions through LLM-Assisted Counterfactual Analysis," 26th International Conference on Environment and Electrical Engineering & 10th Industrial and Commercial Power Systems Europe, Lisbon, Portugal, 29 June – 02 July, 2026.
   The focus of this work is on interpretable redispatch using Post Optimization Analysis and LLM-assisted reasoning. It proposes a counterfactual sensitivity approach and an

Actor–Critic–Distiller-Refiner reasoning loop to justify AC-OPF outputs through physically meaningful insights for system operators.

2. C. Martins, M. Marques, R.J. Bessa, F. Moaidi, S. Camal, "The benefits of Smart4RES predictive analytics," 27th International Conference and Exhibition on Electricity Distribution (CIRED 2023), Rome, Italy, 12-15 June 2023.
   This conference paper presents preliminary results on predictive frameworks for power systems under high renewable penetration. It highlights the practical implementation aspects, case studies, and insights gained from early experimentation.

Technical Reports

1. Ferinar Moaidi; Ricardo J. Bessa, "Evolving power system operator rules for real-time congestion management", Technical report, AI4REALNET Project, Deliverable 2.3, March 2026.
   Within the framework of Task 2.1, Knowledge-assisted AI, this section presents a hybrid AI framework for power grid topology control that integrates GNP, reinforcement learning, and decision trees. A novel variant of GNP is proposed, enabling the evolution of decision-making rules through learning from data within a reinforcement learning paradigm. The graph-based evolutionary representations of both GNP and decision trees support transparent and traceable decision-making, facilitating interpretability and knowledge extraction.

2. Ricardo J. Bessa; Ferinar Moaidi; Luis Teixeira; Hodajt Mariji; Ricardo Andrade; Joao Viana; Manuel Matos; et al, "Predictive management of voltage and congestion problems under RES uncertainty", Technical report, Smart4RED Project, Deliverable 5.3, January 2023.
   This technical report presents a detailed account of the predictive flexibility management framework developed in this thesis. It includes extensive numerical simulations, sensitivity analyses, and risk-cost evaluations that underpin the human-in-the-loop decision-aid methodology. The report provides supplementary material on the *time-to-decide* approach, second-level forecasting models, and guidelines for practical implementation in power system operational planning.

Open-Source Software and Code Repositories

- Ferinar Moaidi, *AI4REALNET*. GitHub repository, `https://github.com/AI4REALNET/XLLM`, 2026.

- Ferinar Moaidi, *AI4REALNET*. GitHub repository, `https://github.com/AI4REALNET/`

`GNPDT`, 2026.

These publications collectively demonstrate the scientific contributions of this thesis and their relevance to both the research community and practical power system operation.

Chapter 2

Background Knowledge and Literature Review

The main goal of the present chapter is to review the core concepts necessary to situate the thesis within the broader scientific and technical landscape of modern power system operation. With the increasing complexity of electrical grids—driven by the integration of renewable energy, uncertainty, and decentralized flexibility—the chapter establishes the theoretical foundations that underpin predictive, data-driven, and interpretable control methodologies.

The first section introduces the paradigm of predictive grid management, providing an overview of forecasting-based operational planning, uncertainty-aware decision tools, and the mechanisms through which flexibility resources can be procured and coordinated ahead of real-time operation. This discussion highlights the limitations of classical deterministic approaches and motivates the shift toward stochastic, robust, and multi-scenario optimization methods used for day-ahead and intraday operational planning.

The second section then transitions to real-time system operation and decision-making under uncertainty, with a particular emphasis on RL strategies. Here, the chapter examines existing methodologies for real-time congestion management, adaptive control of distribution and transmission networks, and the challenges of integrating learned policies into safety-critical operational settings. Special attention is given to the tension between the adaptability of RL agents and the reliability, guaranties, and transparency required in power system operations.

The third section focuses on interpretability and explainability, reviewing efforts towards optimization, interpretation, and recent advances in LLM-based analysis. This section positions interpretability not merely as a desirable feature but as a fundamental requirement for operator trust, regulatory acceptance, and actionable decision support. It also surveys current techniques for extracting insights, conducting counterfactual analyses, and implementing transparent decision structures from optimization and learning frameworks.

The chapter concludes with a summary of key open research gaps—namely, the integration of predictive planning with adaptive real-time control, the development of interpretable RL-driven operating rules, and the use of LLMs to bridge optimisation outputs with human reasoning—which collectively motivate the contributions of this thesis.

2.1 Predictive Management of Grid Flexibility Under Forecast Uncertainty

Predictive management in power systems refers to the planning of control actions ahead of time by using forecasts of demand, RES generation, flexibility availability, and network conditions over a future time horizon. Its main purpose is to identify suitable actions before operational limits are reached, so that system operators do not rely only on remedial actions after congestion, voltage problems, or other violations have already appeared. This need has become more important as power systems include larger shares of variable renewable generation and flexible resources, which increase both uncertainty and the need for coordinated decision-making [7, 8].

In this setting, predictive management is not a single control method, but a decision framework in which forecasts, uncertainty models, and network constraints are combined to select preventive and corrective actions over time. As new forecasts become available, these decisions can be updated to reflect improved information and the uncertainty that still remains. For this reason, Optimal Power Flow (OPF) and its variants are widely used in predictive grid management, since they provide a clear way to represent operating objectives, technical constraints, flexibility activation, and uncertainty within the same optimization problem [7, 8].

The OPF method and its variants is the standard approach for short-term management of grid flexibility, with a closed-form representation of the objective function (e.g., minimize flexibility activation cost) and technical constraints. A survey in [9] identified the major advances and challenges for the OPF problem, namely: problem decomposition (distributed optimization), management of discrete variables, contingency filtering, limiting the number of corrective actions, extend the number of uses cases for the OPF, risk-based OPF, and OPF under uncertainty. The inclusion of multi-period constraints has been also a research topic [10], in particular for smart grids with direct control of demand response and storage. With the impact of RES forecast uncertainty in the electrical grid operational planning, the risk-based and OPF under forecast uncertainty problems are presently highly relevant and, in [8], the state-of-the-art methods are categorized in: a) risk-neutral and risk-averse two-stage stochastic optimization, b) limiting risk of constraint violations (e.g., chance-constrained optimization), c) robust optimization, and d) distributionally robust optimization. For the sake of completeness, below we discuss several OPF-based approaches from the state-of-the-art, but for more details the reader is directed to comprehensive literature reviews in [8, 9, 11].

A complementary line of work formulates predictive grid management as a rolling-horizon or model predictive control problem, in which a finite-horizon optimization is repeatedly re-solved as new measurements and updated forecasts become available [12–14]. In power networks, this receding-horizon principle is attractive because it allows forecast revisions to be translated into updated control actions while preserving network and device constraints [12, 14]. At the distribution level, it has been applied to stochastic scheduling combined with predictive real-time dispatch of flexibility services, and more recent multi-time-scale formulations explicitly couple day-ahead decisions with intra-day rolling corrections [13, 14]. However, these works remain primarily centered on automated dispatch correction, schedule tracking, and cost minimization under updated forecasts, rather than on the operator-oriented question of whether flexibility should be booked immediately or deferred until the next forecast update [13, 14]. In addition, the value of forecast updates is not usually expressed through interpretable risk–cost trade-offs or multi-criteria ranking of candidate flexibility options, which limits their usefulness as decision-aid tools for human operators [12, 13].

In stochastic and Chance Constrained (CC) optimization the uncertainty is represented via a probability distribution, e.g., quantiles, moments of the distribution, parametric function, random samples or vectors from the marginal or joint probability distribution. Zhang and Li proposed a CC-OPF with a back-mapping approach and a linear approximation of the power flow equations where uncertainty (in load) was assumed to follow a parametric multivariate Gaussian representation [15]. The linearization is essential since the non-linear formulation would exponentially increase the computation time, as the dimension of the random variables increases. Roald et al. described a CC-OPF model solved using a randomized optimization technique and representing RES and load forecast uncertainty through scenarios [16]. In that work, a risk measure, inspired by the severity functions from [17], is proposed to relate the amount of line overload in post-contingency condition to the cost of remedial measures and the probability of cascading events, which is, in a second step, integrated as a constraint in DC Optimal Power Flow (DC-OPF). Roald et al. proposed corrective affine control policies that exploit the flexibility from phase shifting transformers and high voltage direct current connections to handle uncertainty (i.e., the impact of already occurred events), later integrated in a Direct Current (DC) CC-OPF [18]. These two aforementioned works were generalized to the AC-OPF problem in [19], where different formulations and solutions are discussed considering non-Gaussian representation for uncertainty, including a iterative solution scheme with deterministic AC-OPF. In [20], the McCormick and big-M relaxation methods are compared to linearize the AC-OPF problem under forecast uncertainty (represented by spatial scenarios/ensembles) and a trade-off between operating costs and system reliability is included in the objective function. To decrease the computational challenges associated to the non-linear nature of AC-OPF, in [21] is proposed a data-driven method, based on sparse regression, to reduce the number of scenarios (Monte Carlo in-the-loop validation step) required to solve the stochastic OPF.

Robust optimization does not make any specific assumptions on probability distributions and the uncertain parameters are assumed to belong to a deterministic uncertainty set, e.g., polyhedral, ellipsoidal. Soares et al. described a two-stage robust AC-OPF for a Distribution System Operator (DSO) contracting market-based Distributed Energy Resources (DER) flexibility to solve, in advance, congestion and voltage problems [22], and where uncertainty is defined as the convex hull of the deviation between the scenario set and the conditional mean forecast and the robustness level was controlled by the number of vertices of the uncertainty set. Attarha et al. proposed a tri-level adaptive robust AC-OPF where a polyhedral uncertainty set is used, robustness is controlled via a budget of uncertainty parameter and a pre-defined set of contingencies is also included [23]. Lee et al. proposed a convex inner approximation (i.e., robust convex restriction) of the non-convex feasible region of the AC-OPF, which ensures solution feasibility (in active power generation dispatch) for all possible load variations [24]. As discussed in [25], it is important to consider first and second order probabilities (i.e., probability of a probability), which leads to distributionally robust optimization methods. For instance, Guo et al. considered ambiguity sets (Wasserstein balls) for multi-period control policies robust to forecast errors and sampling errors due to limited data, and using different linearizations of the AC-OPF equations [26].

Finally, it is important to mention that RL is emerging as an alternative to traditional mathematical optimization for the OPF problem, when formulated as a sequential Markov decision process [27, 28]. However, integration of forecast uncertainty remains an open challenge for RL, at least for large-scale real electrical grids.

Research in other domains already studied the influence of information/forecast updates in decision-making. Regnier studied the impact of lead-time reduction on the preparation for tropical storm events, in terms of total cost reduction, missed detection and false alarms [29]. Wanke and Greenbaum formulated a sequential decision-making problem for airspace congestion resolution, considering a three-stage decision tree for finding the optimal time and type of action (i.e., from doing nothing to "fully" resolve the congestion) [30]. Jewson et al. proposed an extended cost–loss model to decide if a decision is based on the first forecast or wait for the second forecast, and proposed a parametric methodology with Gaussian distribution to estimate the evolution of air temperature forecast error [31]. In another research, the same authors proposed communicating to the human decision-maker information about forecast change (with information updates), introducing metrics like the mean absolute change, probability of change of size X, among others [32].

In the energy domain, some works are starting to consider forecast update in the decision process. Tankov and Tinsi formulated a stochastic sequential decision-making for energy trading, which used stochastic differential equations to describe the evolution of the forecast error as new information becomes available [33]. Crampes and Renault analyzed optimal dispatching strategies for generation units, considering a time-dependent cost function that takes into account the lag between the date of a making a decision and the instant when the decision becomes effective, e.g., generation costs increase when decisions to produce electrical energy must become effective in a short period of time [34]. Bellenbaum et al. designed a two-stage stochastic regional flexibility market mechanism, formulated as a newsvendor problem, which integrates the trade-off between early procurement at low-cost and later procurement at a higher cost but with better forecasts and knowledge about the flexibility demand [35]. Finally, in terms of predictive grid operation, Mühlpfordt et al. established a relation between the price of uncertainty and the total variational distance between the densities of the in-hindsight OPF (i.e., an optimization problem is solved for every realization of uncertainty) and CC-OPF solutions [36]. The price of uncertainty concept can be used to decide between using fast online optimization (i.e., solve technical problems in real-time) or apply CC-OPF in predictive mode.

A related recent perspective comes from decision-oriented learning, which questions the traditional forecast-then-optimize pipeline by emphasizing that a statistically more accurate forecast does not necessarily produce a better operational decision. In power systems, this literature argues that uncertainty should be evaluated through its downstream decision impact, rather than only through forecast error metrics, and reviews a growing set of methods that directly couple prediction models with optimization or control objectives [37, 38]. This perspective is highly relevant to predictive grid management because it reinforces the idea that updated information should be assessed according to its operational value. Nevertheless, the present literature is still mainly focused on data-driven forecast-to-decision pipelines and benchmark operational problems, rather than on network-constrained flexibility procurement with explicit time-to-decide choices, uncertainty-aware procurement timing, and operator-interpretable ranking of flexibility resources. Therefore, while this emerging line of work helps justify why forecast updates matter for decision-making, it does not yet provide a complete framework for interpretable and sequential flexibility booking in active grid operation.

2.2 Real-Time Grid Topology Control

Real-time decision making in power systems refers to the selection of control actions under strict time constraints, using the information available at or close to the time of operation. In this setting, operators must respond to changing grid conditions such as congestion, line overloads, voltage problems, equipment outages, and fluctuations in demand or renewable generation. Unlike predictive or multi-period decision frameworks, which optimize actions over a future horizon using forecasts, real-time decision making focuses on identifying actions that can be implemented immediately or within a short operational window. The main challenge is therefore to balance speed, feasibility, and system security while working with limited time and incomplete information.

One important class of real-time actions is topology control, where the network configuration is modified through actions such as line switching, bus-bar switching, bus splitting, or substation reconfiguration. These actions can relieve congestion, redistribute power flows, and improve the use of existing network capacity without requiring new infrastructure. For this reason, topology control has long been studied as a corrective tool for system operation, first through engineering heuristics and later through optimization-based formulations such as OPF, DC-OPF, and Optimal Transmission Switching (OTS). More recently, the increasing complexity of real-time operation has also motivated the use of data-driven and learning-based methods, especially RL, which aim to approximate fast control policies for grid management under dynamic and uncertain conditions.

The literature on real-time grid topology control and decision making can therefore be viewed along two main directions. The first direction develops optimization and heuristic methods for network reconfiguration, with emphasis on feasibility, security, and computational tractability. The second direction explores learning-based approaches that seek faster and more adaptive control policies, especially in large-scale environments where repeated optimization may be difficult to perform in real time. The following review discusses these developments and highlights their main strengths and limitations.

Early research on network topology control was primarily driven by heuristic approaches aimed at alleviating network congestion. Mazi et al. explored the switching of lines and bus bars for overload mitigation, but the method relied on heuristic rules based on engineering judgment and sensitivity analysis without a comprehensive optimization of an objective function under power flow constraints, as in standard OPF formulations, which has the potential for suboptimal or conflicting actions that fails to account for system-wide interdependencies and limit the applicability of the method to large-scale systems [39]. Bacher and Glavitsch formulated network switching actions as an OPF problem to minimize losses, highlighting the potential of topology as a controllable parameter. Although innovative for its time, the method primarily focuses on a static snapshot rather than dynamic system behavior, limiting its adaptability to real-time operation. Furthermore, its computational complexity posed challenges to scalability for large-scale systems, making it less practical for modern, data-rich grid environments [40]. Bakirtzis and Meliopoulos attempted to incorporate switching into real-time control frameworks. The main limitation lies in not fully addressing the combinatorial complexity inherent in network topology reconfiguration. The method focuses on a limited set of predefined switching options and does not explore the vast space of possible configurations, which restricts scalability and optimization performance in large or highly meshed grids [41]. Makram et al. proposed a Z-matrix-based strategy to select the transmission lines to be switched, offering improved computational efficiency. However, the approach lacked guarantees of optimality and often neglected critical constraints such as voltage limits and dynamic system behavior [42].

During the 1990s, efforts to incorporate dynamic considerations into switching strategies gained traction. Chen and Glavisch investigated stabilizing switching actions during emergencies, focusing on dynamic system response. Although innovative in its focus on transient behavior, the method was not integrated within a formal OPF framework [43]. Chen et al. extended this concept by introducing transient stability constraints in network reconfiguration decisions. Although this represented a step toward dynamical optimization, the work was tested in small systems and faced significant scalability limitations. In the early 2000s, renewed interest in optimization-based approaches emerged. Shao and Vittal developed a guided greedy search based on the Line Outage Distribution Factor (LODF) sensitivity factor, where switching actions are evaluated and selected regarding the predicted local impact, not through global optimization [44]. In a follow-up study, Shao and Vittal applied a binary integer programming formulation within an OPF context to incorporate the switching of lines and bus bars, improving the systematic identification of corrective actions [45]. Granelli et al. proposed a network reconfiguration strategy for congestion management using genetic algorithms with deterministic conditions (i.e., static load). However, a key limitation of this approach is that the search strategy lacks adaptability and contextual awareness, as it relies on predefined fitness evaluations without real-time learning or feedback. As a result, its effectiveness may degrade in dynamic grid environments or under unforeseen operating conditions, where more adaptive or learning-based control strategies would be preferable [46]. Fliscounakis et al. modeled loss minimization via topology changes as a Mixed-Integer Linear Programming (MILP) problem, further highlighting the tractability of reconfiguration within a linearized power flow context [47]. The formulation of OTS as a mathematically rigorous problem marked a turning point. Fisher et al. introduced an OTS model based on Mixed Integer Programming (MIP) within the DC-OPF framework, demonstrating substantial operational cost savings through strategic line switching [48]. However, this approach neglected the characteristics of the Alternate Current (AC) system, limiting its applicability in the real world. To improve system reliability, Hedman et al. incorporated $N-1$ contingency constraints into the OTS framework, although the computational burden increased significantly due to the presence of binary decision variables [49]. Fuller addressed this by proposing a heuristic method to speed up OTSsolution times, though such methods often lack robustness and optimality guarantees [50].Around the same time, Hedman, Oren, and O'Neill identified that OTScould create inconsistencies in the Financial Transmission Rights (FTR) markets, leading to insufficient revenue. To address this, a revenue-adequacy-constrained formulation was introduced (Hedman et al., 2011), aligning OTS with market-based settlements, although at the cost of added model complexity [51].

In more recent work, Han and Papavasiliou employed topological corrections for congestion management in Central Western Europe using a scalable MIP-based approach [52]. Although effective, reliance on zonal approximations limited the physical fidelity of network modeling. Heidarifar and Ghasemi advanced the modeling of practical grid configurations by introducing substation and node-breaker representations into the topology optimization framework [53]. This provided more realistic assessments of switching decisions, but also introduced additional computational challenges related to convergence and scalability.

Recent studies have further advanced topology optimization by integrating sophisticated methods into OPF models. Heidarifar et al. proposed a heuristic approach that incorporates AC and N-1 contingency constraints into transmission line switching and bus splitting, enhancing system reliability [54]. Despite its effectiveness, real-time management remains challenging due to the high computational burden imposed by contingency constraints. Zhou et al. addressed the optimization of the grid topology at the substation level using bus splitting actions, with the aim of reducing congestion and improving grid efficiency [55]. However, the method may face challenges in real-time applications due to the complexity of substation-level configurations and computational intensity. Zhang and Liu developed a multi-period OPF model to manage voltage stability and security through transmission switching [56]. While it provides a comprehensive solution, the method's real-time application is limited by the need for extensive computations over multiple periods. Tavakkoli and Amjady extended the AC-OPF model by incorporating bus-bar switching actions to prevent overload conditions [57]. However, its practical use in real-time is constrained by the intricacy of the overload relay dynamics and the computational time required. These advancements underscore the importance of balancing the trade-off between optimization accuracy and computational efficiency for real-time implementation in large-scale systems.

In the last decade, researchers [58, 59] have found that RL could be crucial to achieving the ambitious goal of real-time decision making without relying on predefined models, making it suitable for complex and uncertain dynamic environments [60]. Recent advances in the application of RL, particularly deep RL, have opened new avenues to address challenges traditionally addressed by OPF methods. Unlike classical OPF, which relies on mathematical programming and model-based optimization, RL offers a data-driven and adaptive approach to decision making under uncertainty. For example, [61] proposed a DRL-based framework for optimal active power dispatch, demonstrating the potential of policy learning to manage system control in real-time. Similarly, [62] and [63] developed DRL-based solutions for AC-OPF that account for system uncertainties and deliver rapid approximations, which are crucial for time-sensitive grid operations. These approaches showed competitive accuracy compared to traditional solvers while significantly reducing computation time, an essential factor for large-scale and real-time applications. Building on these foundations, several recent studies have sought to enhance the practicality and robustness of RL in power system contexts. For example, [64] integrated a physics-informed RL framework for AC-OPF with a look-ahead, incorporating system dynamics and constraints to ensure feasibility and stability. [65] extended this concept by developing a safe-RL model in which the learned policies respect operational limits during stochastic OPF. Moreover, RL has also been used to manage contingencies: [66] explored deep RL to handle generator failures, while [67] combined graph-based learning with a Deep-RL model (graph neural network) to capture the network topology in OPF problems involving renewable generation. Despite their promise, these methods often face limitations in convergence, interpretability, and the need for extensive training data. However, their flexibility and learning capability mark a significant shift toward autonomous, scalable power system control.

In 2019, RTE launched the L2RPN challenge [68], focused on congestion management. The competition introduced Grid2Op [69], an RL environment designed to train and evaluate control algorithms using realistic grid dynamics and historical data. The challenge evolved in complexity with the introduction of robustness and adaptability tracks in the NeurIPS 2020 edition. The robustness track incorporated an adversarial component that simulated $N-1$ contingency scenarios by heuristically disabling transmission lines [70], thus testing agents' resilience to sudden network disruptions. Meanwhile, the adaptability track introduced a higher share of RES, ranging from 10% to 30%, to evaluate the flexibility and responsiveness of agents under variable generation conditions [71]. The winning agent in the NeurIPS 2020 L2RPN competition [72] introduced an action-set-based policy optimization technique that directly embeds safety constraints into the learning process – crucial for bridging experimental success and real-world applicability. Unlike other top submissions that used standard Deep-RL methods like Proximal Policy Optimization (PPO) and Double Deep Q-Network (DDQN) (DDQN), this agent employed a genetic algorithm for policy optimization. However, a subsequent analysis [6] showed that no submitted solution fully met the stringent safety and reliability standards required for deployment in operational power grids. A notable advancement beyond the L2RPN NeurIPS 2020 and WCCI 2020 agents is the PowRL agent [73], a Proximal Policy Optimization (PPO)-based model designed to manage power networks under variable and adversarial conditions. While it showed improved resilience and adaptability, its effectiveness depends heavily on the simulation environment's fidelity and reward design. Similarly, CurriculumAgent (CAgent) [74] extended PPO by integrating an $N-1$ reliability strategy and systematically comparing rule-based and RL-driven topology control. Despite their promise, such RL models often lack interpretability, limiting trust in safety-critical grid operations.

These RL-based approaches represent significant progress toward intelligent data-driven grid management, demonstrating promising performance in tasks such as re-dispatch and topology optimization. However, they often suffer from two shortcomings. Firstly, the limited interpretability of the learned policies. These approaches rely heavily on deep neural networks to approximate value or policy functions, making it difficult to trace the rationale behind specific control actions, such as switching operations or dispatch decisions. For example, in [67], the agent learns to act based on graph-structured observations, but it provides no explicit explanation of the grid conditions or topological features that triggered its decisions. Similarly, in [65], while the constraints are enforced via a safe-RL framework, the internal decision pathway remains opaque. This lack of transparency is a barrier to deployment in critical grid operations, where operators require not only effectiveness but also the ability to validate and interpret automated decisions. Secondly, these methods often require large amounts of simulated data and prolonged training episodes to converge, which can be computationally expensive and hinder rapid adaptation to evolving grid conditions. This combination of opacity and high resource demands limits their practical deployment in critical real-time power system operations.

2.3 Explainable Optimal Power Flow and Interactive Interpretation

OPF is a central tool for power system operation because it determines dispatch decisions that satisfy network constraints while optimizing an operational objective, such as generation cost, losses, or security margins. However, although OPF provides mathematically optimal solutions, the reasons behind these solutions are often difficult for human operators to understand directly. In practice, dispatch outcomes are shaped by a combination of active constraints, dual variables, network sensitivities, and trade-offs between competing objectives. As a result, there is an important difference between solving an OPF problem and explaining why a given solution is obtained.

This challenge has motivated growing interest in explainable optimization for power systems. In this context, explanation means translating the internal logic of the optimization problem into a form that is understandable to operators, planners, or other decision-makers. Such explanations may clarify which constraints are binding, why some generators are redispatched while others are not, what network conditions drive congestion, or how small changes in inputs could modify the final decision. Therefore, explainable OPF is not concerned only with prediction accuracy or computational efficiency, but also with transparency, interpretability, and decision support.

More recently, this line of work has expanded from static post-solution interpretation to interactive explanation frameworks. These approaches aim not only to describe solver outputs, but also to support dialogue, counterfactual analysis, and iterative reasoning about feasible operational changes. In this setting, LLMs and related Explainable AI (XAI) methods have emerged as useful tools for connecting mathematical optimization models with natural language interaction, thereby helping operators explore why a solution was obtained, what alternatives are possible, and how system conditions influence the optimization outcome.

To bridge the gap between numerical optimization and human understanding in power system operations, recent research has increasingly focused on explainable optimization frameworks. Fritz and Bukhsh (2025) [75] proposed an explainable DC-OPF approach that generates counterfactual explanations and employs decision-tree models to provide interpretable insights into dispatch outcomes. Their framework identifies minimal changes in system inputs, such as nodal demands, that would alter generator dispatch decisions, enabling operators to understand the influence of constraints and demand variations. While this approach enhances interpretability and decision support, it is limited to linear DC formulations and relies on historical solution data for its heuristic explanations. Similarly, Pineda et al. (2025) [76] extend this effort by improving the interpretability of OPF-related computations. Pineda et al. introduce an interpretable, hybrid learning–plus–linearization AC-OPF approximation, which enhances transparency compared to neural network surrogates but does not provide explicit, human-readable explanations of why specific dispatch decisions occur. Instead, their method produces a white-box optimization model, not post-hoc interpretability. As a result, their approach still lacks semantic, operator-facing explanations—a gap especially relevant in realistic AC-OPF settings. Goerigk et al. [77] further introduced symbolic regression models using constraint decomposition and local sensitivity visualization to approximate OPF outcomes; however, these symbolic surrogates provided static explanations without feedback or dynamic introspection. Li et al. [78] extended interpretability through dual-based decomposition; however, their approach lacked natural language reasoning and counterfactual analysis. Collectively, these works demonstrate that numerical optimization can be made more interpretable by connecting solver outputs—such as dual variables, active sets, and sensitivities—to intuitive physical explanations. Nevertheless, most existing approaches remain symbolic or post hoc, offering limited interactive adaptability and little integration with real-time operator feedback.

Beyond power systems, recent work in explainable optimization has started to formalize explanation mechanisms directly for mathematical optimization problems. For example, Kurtz et al. [79] extend counterfactual explanations to linear optimization and show that some forms of relative counterfactual explanations can be computed efficiently, which suggests that optimization-specific explanations need not be limited to feature attribution alone. In parallel, Biemans et al. [80] investigate whether LLMs can transform numerical explanation outputs into clearer, more persuasive narratives for human planners in supply-chain optimization. However, these works still remain limited either to linear optimization settings with tractable convex structure or to narrative reformulations based on LIME-inspired explanation signals [79, 80]. As a result, they do not address the non-convex physics of realistic AC-OPF problems, the role of dual feasibility and sensitivities in redispatch decisions, or the need for explanations that remain explicitly grounded in power-system constraints while still being understandable to operators.

Complementary surveys (e.g., Da Ros et al., 2025 [81]) have documented the rapid emergence of such hybrid reasoning systems that unify optimization solvers, LLMs, and XAI frameworks. These developments mark a paradigm shift from post-solution interpretability to *interactive interpretability*, where models and operators co-reason about operational constraints and feasible redispatch options.

Recent advances in LLMs have opened a new frontier for interpretable optimization. LLMs are increasingly used to bridge the gap between formal mathematical solvers and human operators by providing natural language interfaces, scenario exploration, and verbal reasoning over model internals. For instance, Zhang et al. (2025) introduced *OptiChat* [82], a dialogue system driven by LLMs that enables users to interrogate Pyomo-based optimization models. OptiChat supports “what-if” and “why-not” analyses, extracting variable structures, constraints, and sensitivities to produce counterfactual explanations for model outputs. Similarly, Chen et al. (2024) [83] proposed an LLM interface for diagnosing infeasible optimization problems, allowing operators to explore relaxation options and trade-offs in natural language. These LLM-driven interpretability systems share three defining characteristics:

1. Convert optimization solvers into *dialogue-capable systems*, allowing structured queries such as “Why is bus X voltage constrained?” or “What if generator Y increases output by 10 MW?”.

2. Combine symbolic sensitivity computation with verbal reasoning to generate explanations anchored in numerical model data.

3. Support human-in-the-loop counterfactual reasoning by determining the smallest perturbations to system variables required to resolve constraint violations or enhance objective performance.

Their model demonstrated promising descriptive capabilities but limited quantitative grounding. LLMs are increasingly employed as autonomous reasoning agents capable of iterative improvement through feedback. Recent developments have introduced frameworks in which an LLM not only generates textual outputs but also learns from verbal feedback, reflection, and memory consolidation to enhance reasoning quality over time. This paradigm, referred to as *verbal RL*, provides a natural and interpretable extension to RL principles in language-based reasoning systems. Architectures such as Reflexion [84], Language Agent Tree Search (LATS) [85], and Self-Evaluation Memory [86] represent a significant step beyond traditional explainable optimization. Unlike prior LLM approaches that produce static textual interpretations, these models dynamically refine their explanations based on physical feedback.

A related stream of research couples LLMs with formal reasoning engines rather than relying on free-form language generation alone. For instance, Régin et al. [87] combine constraint-programming reasoning with LLM predictions so that the language model proposes candidate content while the formal reasoning layer ensures that structural constraints are satisfied. This idea is particularly relevant for power systems, where plausible verbal explanations are insufficient unless they remain consistent with feasibility and network structure. At the same time, recent work has shown that generic multi-agent discussion is not automatically beneficial for reasoning quality: Wang et al. [88] report that, across several reasoning benchmarks, a single-agent LLM with strong prompting can match the best multi-agent discussion approaches, and that discussion mainly helps under specific prompt conditions. Therefore, simply introducing dialogue or reflection loops does not by itself guarantee reliable interpretability. What remains missing for OPF is a domain-grounded reasoning layer in which language agents are explicitly constrained or critiqued using power-system feasibility checks, sensitivity information, and optimization structure [87, 88].

2.4 Literature Gaps and Thesis Contributions

The literature reviewed in this chapter shows substantial progress in predictive management under uncertainty, real-time decision making, and optimization explainability. However, these three lines of research have largely evolved separately, and a clear methodological gap remains at their intersection. Predictive OPF-based approaches offer rigorous mathematical tools for modeling uncertainty and supporting operational planning, but they are often too slow or too rigid to support repeated decision updates in dynamic environments with frequent forecast revisions and increasing renewable generation [8]. Real-time control approaches, especially RL and topology optimization, provide the promise of faster operational responses, but they still face important limitations in interpretability, safe adaptation, and scalability. At the same time, recent work on explainable optimization and LLM-based interfaces improves access to model outputs, but most of these methods remain post-hoc, weakly grounded in physical reasoning, and detached from actual operator decision loops [75,76,81,89]. Taken together, the literature still lacks a framework that can simultaneously provide predictive awareness of uncertainty, adaptive real-time control, and physically grounded, operator-oriented explanations.

2.4.1 Gap 1: Predictive Flexibility Management Under Forecast Uncertainty

A first gap concerns predictive management. Existing stochastic, chance-constrained, robust, and distributionally robust OPF formulations provide a rich basis for representing uncertainty and risk, yet their integration into sequential operational decision-making remains limited when forecast updates must be taken into account repeatedly and within short time windows. In practice, multi-period OPF and robust AC-OPF formulations are difficult to use as operational decision-aid tools in settings where system operators need summarized and interpretable information rather than a single optimized solution.

Existing OPF-based methods provide rigorous optimization under uncertainty, but they do not support repeated and operator-oriented decision updates when forecasts evolve over time [8]. Most approaches provide a single optimized solution, whereas system operators often require summarized and interpretable information to compare several flexibility options. Prior stochastic and robust approaches do not explicitly study how the operator should adjust the risk level according to the operational stakes [16, 19, 22, 26]. Practical flexibility and curtailment approaches either neglect forecast uncertainty or select resources only from sensitivity values, without formulating a broader multi-criteria ranking problem [90, 91]. In addition, the decision of whether to act immediately or wait for a forecast update is largely absent from the state of the art [19, 21, 22, 26, 90, 91].

This thesis addresses these limitations through a predictive grid management framework that uses sensitivity indices and risk metrics to rank flexibility options and summarize the decision problem through risk–cost curves. It also proposes a decision-aid interface that improves interaction between the operator and the model by allowing the exploration and ranking of multiple flexibility solutions. Furthermore, it introduces a multi-criteria decision formulation in which flexible resources are ranked according to several criteria, rather than selected only from the highest sensitivity value. The thesis also proposes a *time-to-decide* formulation in which the operator chooses between booking flexibility immediately or waiting for the next forecast update, depending on the trade-off between cost, risk, and information quality. Finally, it introduces a novel *second-level forecasting* concept, where a dedicated model predicts the uncertainty of future forecasts through quantile estimation; in the experiments, an encoder–decoder deep learning model achieved the best performance.

This contribution complements market-oriented flexibility mechanisms such as [35], which do not address how the system operator estimates flexibility needs. It also differs from the price-of-uncertainty perspective in [36], since the proposed framework supports operator interpretation of risk–cost relations rather than relying only on a distance metric between solutions. Finally, it differs from parametric forecast-update approaches such as [31], which are not well suited to the uncertainty structure of RES forecasts.

2.4.2 Gap 2: Interpretable and Adaptive Real-Time Topology Control

A second gap appears in the literature on real-time decision making. RL has shown strong potential for fast control in complex power system environments because learned policies can often be executed online with limited delay and can discover non-trivial strategies for network management [58–60]. However, most RL-based controllers remain difficult to interpret, since the logic of their decisions is embedded in high-dimensional neural representations. In addition, although RL policies can be fast at inference time, they are usually trained offline and then deployed as fixed controllers, making safe adaptation to new operating conditions difficult. Optimization-based topology control provides explicit treatment of system constraints, but it remains limited by combinatorial complexity and scalability barriers that restrict its real-time use.

Most deep-RL controllers behave as black boxes, which limits operator trust and makes validation difficult in safety-critical power system applications. Furthermore, RL policies are generally trained offline and deployed as fixed controllers, making adaptation to unseen operating conditions, distribution shifts, or new contingencies difficult. Optimization-based topology control methods, such as OTS and MIP-based reconfiguration, also remain constrained by combinatorial complexity and limited real-time scalability. As a result, existing methods often impose a trade-off between interpretability and performance: rule-based or heuristic methods are easier to understand but less adaptive, whereas deep-RL methods are adaptive but opaque.

This thesis addresses these limitations through a novel GNP framework with dynamic node behavior, enabling context-aware decision-making in uncertain power system environments. Unlike conventional GNP methods that rely on fixed function nodes and sequential updates [92], the proposed framework uses adaptive functional nodes whose behavior changes according to the current system state and operational context. It also introduces a hybrid GNP-DT methodology that combines Genetic Network Programming with Decision Trees to improve interpretability. In this framework, the control policy is represented explicitly as a graph, where each node corresponds to a programmed decision heuristic, making the decision process traceable and visually interpretable. In addition, a multistage decision-tree extraction process generates human-readable rules from the trajectories of top-performing decision graphs. The framework also allows the initialization of the population with expert-derived decision graphs, which can accelerate convergence in the early learning stages.

This design allows the controller to represent more sophisticated state-dependent policies suited to uncertain and non-linear power system conditions, such as fluctuating renewable injections and non-linear network responses. For example, an overloaded line does not necessarily trigger the same switching action in all situations; instead, the selected control depends on the surrounding operating context, such as nearby contingencies or evolving grid stress. On the IEEE 118-bus system, the proposed method outperformed both the baseline expert system [93] and CAgent [74], achieving up to an 18% improvement in the mean L2RPN performance score and an 80% reduction in inference time. These results show that it is possible to combine adaptability, transparency, and computational efficiency in a real-time topology control framework.

2.4.3 Gap 3: Physically Grounded and Interactive Explanation of OPF-Based Decisions

A third gap emerges from the literature on optimization explainability. Existing explainable optimization methods based on symbolic sensitivity analysis, dual-based reasoning, local surrogate modeling, or white-box approximations have improved the interpretability of OPF results, but most remain static and post-hoc [75–78]. They generally explain a solved optimization instance after the fact, without supporting iterative dialogue, counterfactual redispatch reasoning, or interaction with the operator over multiple layers of physical constraints. More recent LLM-based systems have extended interpretability by allowing users to query optimization models in natural language [81–83]. However, in power-system settings, these systems are still not sufficiently grounded in AC network physics, counterfactual OPF sensitivities, and feasibility-aware redispatch reasoning.

Existing explainable optimization approaches are mostly static and post-hoc. Most methods do not support iterative dialogue, counterfactual redispatch reasoning, or multi-step interaction with the operator. Current LLM-based systems improve accessibility but remain weakly grounded in AC power-system physics and redispatch feasibility. As a result, there is still no unified framework that combines optimization outputs, physical sensitivities, and interactive reasoning in a way that supports real operational decision-making.

This thesis addresses these limitations through a counterfactual, LLM-driven reasoning system for AC-OPF redispatch that combines physical sensitivity analysis with iterative language-based reasoning. It introduces an interpretation layer that converts AC-OPF outputs into counterfactual insights using sensitivity matrices, showing how incremental redispatch actions affect line flows and slack-variable relaxations. It also proposes a structured reasoning loop composed of an *Actor*, *Critic*, *Distiller*, and *Refiner*, designed to interpret optimization outputs while preserving consistency with power-system physics. In this loop, the Critic performs a feasibility- and dual-consistency-based validation step to evaluate whether the generated explanation remains physically plausible. The framework also includes a refinement mechanism in which the Distiller extracts compact operational rules and the Refiner integrates them into subsequent prompts. Finally, it introduces a verbal reasoning loop with short-term and long-term memory, inspired by Reflexion [84] and LATS [85], so that feedback from previous reasoning episodes improves future explanations.

This shifts explainability from a static descriptive layer to an interactive, physically grounded reasoning system for operator support.

2.4.4 Final Remarks

Taken together, the gaps identified in this chapter point to a broader methodological problem rather than three independent shortcomings. Modern power system operation requires decisions that are anticipatory, fast, and explainable at the same time, yet the literature still tends to treat these properties separately. Predictive management emphasizes robustness under uncertainty but often lacks operational flexibility, real-time control emphasizes speed and adaptability but often sacrifices interpretability or robust constraint awareness, and explainable optimization improves transparency but remains largely detached from live operational reasoning. The central contribution of this thesis is therefore not only to advance each of these areas individually, but also to place them within a common methodological perspective centered on operator-oriented decision support. Across predictive planning, real-time topology control, and redispatch explanation, the thesis develops methods that preserve physical consistency while improving interpretability, adaptability, and operational usability.

In this context, the thesis makes three main contributions. First, it proposes a predictive flexibility procurement strategy that integrates forecast updates, sensitivity-based and risk-based ranking of flexibility resources, risk–cost curves for operator support, a multi-criteria decision formulation, and a time-to-decide mechanism supported by second-level forecasting of forecast uncertainty. Second, it develops an interpretable evolutionary real-time topology control framework based on dynamic-node GNP and hybrid GNP-DT, enabling context-aware control with transparent reasoning and low inference time. Third, it introduces a counterfactual, LLM-driven reasoning system for AC-OPF redispatch, providing interactive, physically grounded explanations that support operator understanding of optimization decisions. Together, these contributions address the main literature gaps identified in this chapter and support a broader research direction in which uncertainty-aware optimization, adaptive control, and explainable reasoning are treated as complementary elements of intelligent power system operation.

Chapter 3

Uncertainty-Aware Procurement of Flexibilities for Electrical Grid Operational Planning

Based on the reviewed commercial flexibility market platforms [5] and the survey recently conducted by the Joint Research Centre (JRC) with companies like EPEX SPOT or Enedis [94], this work assumes that both long-term and short-term products will probably be requested in future flexibility markets, subject to the specific network needs in each case. The long-term flexibility products are designed to defer network investment, while the short-term flexibility products target congestion management and the strengthening of network resilience. However, a significant challenge lies in determining the appropriate price caps (or annual flexibility budget), which is comparatively more straightforward regarding long-term products associated with investment deferral [95].

According to JRC, while the present state of local flexibility markets in Europe does not provide a conclusive outlook regarding their future characteristics, a notable finding from the interviews is the anticipation of transitioning towards short-term flexibility markets [94]. The main reasons are [94]: a) by enabling smaller assets (e.g., EVs) to participate in the procurement process, there is enhanced liquidity that arises due to their ability to accurately forecast its flexibility potential only in the short-term; b) improved grid forecasts closer to real-time mitigate volume risk for network operators. Nevertheless, long-term contracts are expected to persist as a means of securing reliability and integrating flexibility into the long-term expansion of networks.

The focus of this work is on the short-term horizon, drawing inspiration from successful System

Operator pilots such as the CoordiNet project Swedish pilot [96], sthlmflex project with NODES platform [94], and the Redispatch 2.0 in Germany [97]. Additionally, insights are drawn from market operator initiatives such as Enera Flexmarkt, operated by EPEX SPOT [94], and OMIE with the IREMEL platform [5]. It is important to note that this methodology extends beyond flexibility market procurement. It can also be applied for defining curtailment signals, particularly for RES with non-firm connection contracts (like Enedis in France [90]).

For this work, the following assumptions were made. Firstly, day-ahead and intraday flexibility markets occur before each energy trading session, which aligns with the flexibility market platforms offered by NODES, GOPACS, and Enera [5]. The market operator collects the bids submitted by Flexibility System Provider (FSP) and provides this information to TSO and/or DSO. The flexibility needs can be revised in the intraday sessions with updated load and RES forecasts and changes in the flexibility band. Secondly, availability and dispatch payments are used for contracted flexibility. Thirdly, a pay-as-bid pricing method is considered for flexibility payment, widely recognized as the most common payment scheme [5].

The availability price is a mitigation measure to ensure enough flexibility volume in the market [5]. However, it is important to acknowledge that liquidity issues related to flexibility markets can still arise. For instance, the price cap established through the long-term analysis [95] may significantly limit revenue opportunities for FSPs in technical constraints management markets, particularly for those already involved in more lucrative markets such as wholesale energy trading and frequency control. While addressing liquidity problems falls beyond the scope of this work, the proposed methodology offers valuable contributions to mitigate this issue. Firstly, the implementation of *second-level forecasting* effectively reduces the overall cost associated with flexibility use, as shown in section 3.4. This approach minimizes expenses and discourages an open-book mentality in the process of flexibility procurement. Secondly, the proposed methodology also serves as a valuable tool to communicate the potential risk of flexibility scarcity to the operator, as illustrated in section 3.4.5.

Furthermore, without loss of generality, this work considers hourly market intervals and three forecasting moments: i) day-ahead forecast (for D+1) with NWP generated at 0h00, i.e., between $t+24|t$ and $t+48|t$; ii) updated NWP data at 12h00, i.e., between $t+12|t$ and $t+36|t$; iii) 2-hour forecast before delivery, which can be used for intraday participation, i.e., $t+2|t$.

Finally, the study and integration of various TSO-DSO coordination schemes are outside the scope of this work. However, it is important to note that information from both TSOs and DSOs can be effectively integrated into the proposed methodology, as elaborated in the subsequent section.

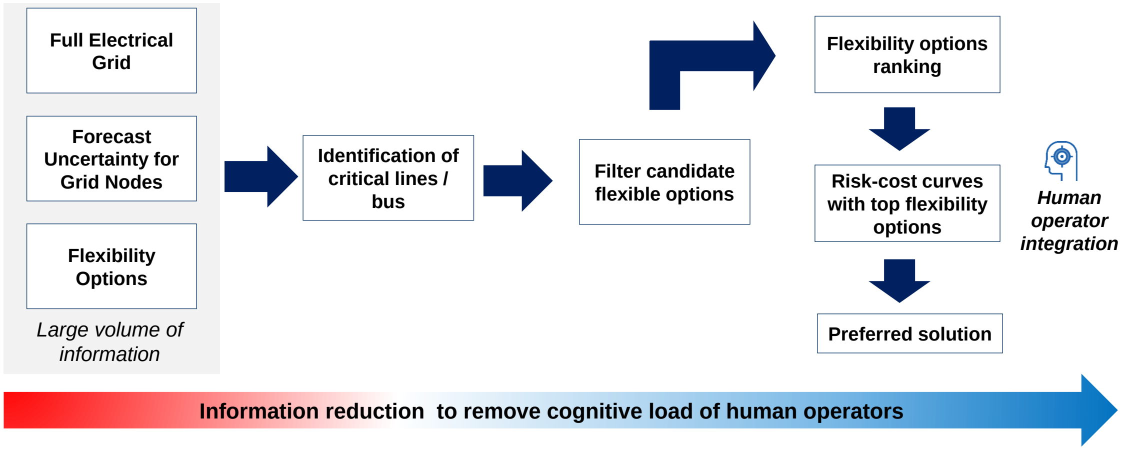


Figure 3.1: *Building blocks of the risk-aware flexibility procurement*

3.1 Methodology

Fig. 3.1 presents the building blocks of the proposed methodology. It departs from a large volume of information, namely the full electrical grid, RES/load uncertainty forecasts, and a set of flexibility options, i.e., dispatchable distributed generation, RES curtailment, Demand Response (DR), storage, network reconfiguration, on-load tap changer (OLTC) and capacitor banks/reactor shunts (Capacitor Bank (CB)/Reactance Shunt (RS)). Step by step, all this information is filtered and finally condensed into a risk-cost curve from where the human operator can select a preferred solution. This approach reduces the cognitive load of human operators and offers simplicity when analyzing the multiple flexibility options under uncertainty. Moreover, it enables simultaneous analysis of multiple lines/buses with a high probability of technical problems. The load and RES forecast uncertainty is represented by random vectors (or scenarios) that can be generated by physical or statistical approaches and capture the spatial dependency of forecast errors.

The first step focuses on running a power flow for each scenario, and computing, for each node and branch, the probability of having an over/under-voltage and congestion issue. The operator can set a minimum probability to select and analyze a reduced set of critical buses/lines. For this subset of elements, sensitivity indices (see section 3.2.1) relating active/reactive power and the line's current and node's voltage are computed. Then, in a second step, the flexibility options are filtered based on a sensitivity index minimum threshold to identify the most "relevant" options for a specific technical problem. The third step includes performing a flexibility ranking with the Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS) method [98] (section 3.3.1), considering a set of risk metrics that measure how effective a flexibility option is in solving a technical problem under uncertainty. It can also include probabilistic information (risk metrics) from the TSO or DSO associated with conflicting actions. Finally, the final step focuses on the design of a risk-cost curve (section 3.3.2) by combining a subset of flexibility options (e.g., 3 to 5 options) selected by the operator and considering only the non-dominated solutions. The flexibility capacity of these options may also be constrained in the curves due to conflicting actions arising from the activation of flexibility by both DSO and TSO. The risk can be represented by different metrics, and without loss of generality, this work uses the probability of having a congestion or under/over-voltage problem. To obtain a preferred solution, we consider two decision-making paradigms: one based on the maximum risk threshold and the other based on a trade-off value (section 3.3), where the risk threshold and trade-off values are conditioned by the stakes level (i.e., risk of cascading failure) and operator's preferences. Note that this relation between risk and stakes borrows the fundamental idea behind confidence-based decision-making theory [99].

The methodology above does not provide information about the best moment to book flexibility, considering that i) forecast uncertainty will decrease with time and ii) the flexibility availability price might increase with a shorter notification time. Thus, a *time-to-decide* problem is formulated based on the *second-level forecasting* concept (section 3.2.4), where the goal is to forecast how the RES and load uncertainty forecasts change over time, as new information is available. The term *second-level forecast* comes from the fact that we are, in essence, computing a forecast of forecast uncertainty. For the flexibility price, it is assumed that the SO has a forecast or that the price offers with different notification times are available in advance.

The *second-level forecast* of conditional quantile $\widehat{q}^{\alpha}_{t+k|t}$, Eq. 3.1, generated with information available at time instant t (forecast launch time) for time interval $t+k$, with nominal proportion α ($0 < \alpha < 1$), and a forecast update rate of z time intervals, corresponds to the conditional expected value of quantile $\hat{q}^{\alpha}_{n,t+k|t+z}$ of the next forecast (i.e., generated at $t+z$) for the same time interval $t+k$. f is the *second-level forecasting* model for each conditional quantile α, and uses as covariates i) quantile forecasts $\hat{q}^{\alpha}_{t+k|t}$ generated at t and ii) engineered features that explain the level of uncertainty in the power forecast generated at t and exogenous variables such as NWP, both denoted by $X^{\alpha}_{t+k|t}$ with l variables, plus a random shock $e_{t+k|t}$.

$$\widehat{q}^{\alpha}_{t+k|t} = \mathbb{E}[\hat{q}^{\alpha}_{t+k|t+z}|\hat{q}^{\alpha}_{t+k|t}, \mathbf{X}^{\alpha}_{t+k|t}] = f\left(\hat{q}^{\alpha}_{t+k|t}, \hat{x}^{1}_{t+k|t}, \ldots, \hat{x}^{l}_{t+k|t}\right) + e_{t+k|t} \tag{3.1}$$

Fig. 3.2 depicts the integration of *second-level forecasting* into the *time-to-decide* problem, considering: i) forecast launched at 0h00 (t) for lead-time $t+30$ (day-ahead), and ii) two *second-level forecasts* for the forecast that will be generated 12 hours later with NWP updated at 12h00 (i.e., $\widehat{q}^{\alpha}_{t+30|t} = \mathbb{E}[\hat{q}^{\alpha}_{t+30|t+12}|\hat{q}^{\alpha}_{t+30|t}, \mathbf{X}^{\alpha}_{t+30|t}]$), and another for the forecast that will be generated 2-hours before delivery time (i.e., $\widehat{q}^{\alpha}_{t+30|t} = \mathbb{E}[\hat{q}^{\alpha}_{t+30|t+28}|\hat{q}^{\alpha}_{t+30|t}, \mathbf{X}^{\alpha}_{t+30|t}]$). This leads to three risk-cost curves that inform the operator about the possible outcome of waiting for the next forecast to book flexibility, each one corresponding to three different forecasting launch times for the "delivery hour", $t+30|t$, $t+18|t$, and $t+2|t$. In this illustrative example, the preferred option would be to wait for the next forecast since the curve obtained with

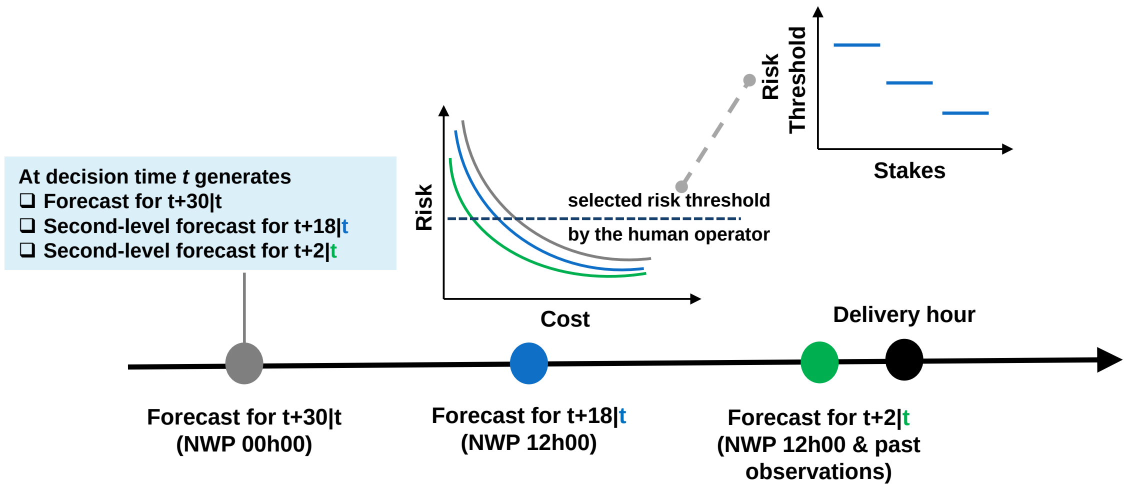


Figure 3.2: *Time-to-decide framework and risk-stakes relation. The integration of second-level forecasts into the time-to-decide problem generates risk-cost curves for a) forecast launched at 0h00 for lead-time $t+30|t$, and b) two second-level forecasts, one for the $t+18|t$ forecast (NWP updated at 12h00), and another for the $t+2|t$ forecast. The preferred solution for the human operator is selected from a risk threshold conditioned by the risk-stakes relation*

the *second-level* forecast for $t+2|t$ indicates a lower flexibility cost for the same risk level.

Finally, it is essential to mention that contingencies were not considered. However, the methodology can be easily extended to include information about the probability of contingencies in the risk modeling and combined with the scenarios.

3.2 Knowledge Construction for Flexibility Procurement

3.2.1 Sensitivity indices (SI)

Analytical method

This section presents the analytical method for deriving Sensitivity Indices (SI) for the electrical grid, and a machine learning proxy to leverage from its very fast inference time and avoid analytical calculations for each grid operating scenario. The state-of-the-art approaches [100,101] mainly follow an updated Jacobian matrix derived from the power flow algorithm. One of the critical deficiency in following these type of methods is the necessity of redefining the Jacobian matrix in case of any transition in system operation status that also may face convergence issue due to inversion procedure. In addition, these approaches are limited to Newton-Raphson (NR) powerflow methods. In this work, we follow the Y_{bus} compound matrix method [102], which offers a straightforward analytical derivation of node voltages and line currents sensitivities. This approach uses the whole admittance matrix of the network with sparsity advantage, and is applicable to a generic number of slack buses. The Y_{bus} matrix approach considers the inherent multi-phase and unbalanced nature of distribution networks.

In this work, the analytical solutions of the system of linear equations are provided by constructing a Characteristic Block Matrix (CBM), in terms of Y_{bus} matrix and voltage vector elements (before changing control variables) of the network, to extract (through a two-step algorithm) the voltage magnitude and angle sensitivity full matrices versus variations of active power/reactive power/OLTC/ shunt, and then using them to find the SIs of line currents. It is worth mentioning that /CBM of the network extracts the voltage sensitivity full matrix for flexibility without any iteration procedure, which makes it a computational efficient method for large scale networks and real-time implementations.

The Y_{bus} compound matrix approach [102] has more computational efficiency than Jacobian-based methods in computing voltage magnitude and current sensitivity coefficients, however, it still needs iteration to solve the system of linear equations through a three-step algorithm. Furthermore, this approach is applicable only for networks without Generator Buses (PV) buses since the authors assumed that by imposing perturbation in nodal power injections, the power set points of other buses are not changed while, in practice, the active and reactive power set points of PV buses are changed. Hereupon, in this work, we generalize the Y_{bus}-extended matrix approach to all electrical networks with all type of bus using the PV2Load Buses (PQ) conversion method (explained in Appendix A.2). In addition, voltage sensitivity to OLTC tap position has been derived in the rest of the Appendix A.2. We develop the same methodology to extend it for Voltage SI versus Shunt elements.

To compute the voltage SI with respect to shunt (CB/RS) deviations for a network with M buses (SB slack buses and N load buses) and N_{sh} shunts, within the framework of the Y_{bus}-extended matrix method and the PV-to-PQ conversion approach, the partial derivative of Eq. (A.4) with respect to $\boldsymbol{b}_{sh}$, the shunt susceptance vector, is first calculated. Then, by denoting $\boldsymbol{X} = \frac{\partial \boldsymbol{E}}{\partial \boldsymbol{b}_{sh}}$, recalling that only the diagonal elements of the admittance compound matrix have non-zero derivatives with respect to $\boldsymbol{b}_{sh}$, rewriting complex matrices and vectors in rectangular coordinates, and after straightforward but tedious partial-derivative, vector, and matrix manipulations, we obtain

$$-\boldsymbol{F}^d = (\boldsymbol{\alpha}_D + \boldsymbol{\gamma}) \cdot \boldsymbol{X}^d + (\boldsymbol{\beta}_D + \boldsymbol{\eta}) \cdot \boldsymbol{X}^q \tag{3.2}$$

for the real part and

$$-\boldsymbol{F}^q = (\boldsymbol{\beta}_D - \boldsymbol{\eta}) \cdot \boldsymbol{X}^d + (-\boldsymbol{\alpha}_D + \boldsymbol{\gamma}) \cdot \boldsymbol{X}^q \tag{3.3}$$

for the imaginary part. In above, the $N \times N_{sh}$ matrix $\boldsymbol{F}^d$ and $\boldsymbol{F}^q$ are the real and imaginary parts of $\bar{\boldsymbol{F}}$, respectively, with $\bar{\boldsymbol{F}} = \underline{\boldsymbol{E}}_D \cdot \frac{\partial \bar{Y}}{\partial \boldsymbol{b}_{sh}} \cdot \bar{\boldsymbol{E}}$. By denoting the admittance matrix of shunts per unit as $\bar{\boldsymbol{y}}$, the ki^{th} column of $\bar{\boldsymbol{F}}$, after some vector operations, calls:

$$[\bar{\boldsymbol{F}}]_{ki} = \underline{\boldsymbol{E}}_D \cdot \bar{\boldsymbol{E}}_D \cdot \bar{\boldsymbol{y}}_{ki}, \qquad 1 \leq ki \leq N_{sh}, \tag{3.4}$$

where $\bar{\boldsymbol{y}}_{ki} = [\frac{\partial \bar{y}}{\partial \boldsymbol{b}_{sh}}]_{ki}$, the ki^{th} column of an $N \times N_{sh}$ matrix with a non-zero element at the l^{th} row corresponding to the shunt bus number l. For an ideal shunt, that $\boldsymbol{F}^d = \boldsymbol{O}_{sh}$ and $\boldsymbol{F}^q = -|\bar{\boldsymbol{E}}|_D^2 \cdot \boldsymbol{\delta}_{sh}$, where $\boldsymbol{O}_{sh}$ and $\boldsymbol{\delta}_{sh}$ are, respectively, the $N \times N_{sh}$ zero matrix and matrix of which elements are zeros except those correspond to the buses which connected to shunts and have values ±1 with +1 (-1) for CB (RS). Now, by: defining $\boldsymbol{SV}_{m(a),sh}$ as the voltage magnitude (angle) sensitivity matrix versus variations of shunts; substituting $\boldsymbol{SV}_{m(a),sh}$ in Eqs. (A.12-A.13) in replace with $\boldsymbol{SV}_{m(a),P(Q)}$; rewriting Eq (3.4) in the rectangular coordinates and then including it in Eqs. (3.2-3.3); and after matrix and vector operations, $\boldsymbol{SV}_{m(a),sh}$ is obtained in terms of the /CBM of network as follows:

$$\begin{bmatrix} \boldsymbol{SV}_{m,sh} \\ \boldsymbol{SV}_{a,sh} \end{bmatrix} = \pm \begin{bmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{bmatrix}^{-1} \cdot \begin{bmatrix} \boldsymbol{O}_{sh} \\ |\bar{\boldsymbol{E}}|_D^2 \cdot \boldsymbol{\delta}_{sh} \end{bmatrix}, \tag{3.5}$$

where with + (−) for CB (RS) and the $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$, and $\boldsymbol{D}$ matrices are given by Eqs. (A.15-17).

For network topology reconfiguration, two approaches are used to compute the current change due to network topology reconfiguration, Distribution Factor (DF) method [103] and a modification to Incidence Matrix (IM) method explained in [103]). To model the reconfiguration of the power system network, a good candidate is the impedance (Z_{bus}) matrix method because of linearity and computational efficiency. First, we obtain Z_{bus} of the network. Although, Z_{bus} matrix of the network can be constructed by the direct step-by-step method or by using basic concepts from graph theory [104], in this work, we obtain it through inverting the Y_{bus} full matrix. The inverting Y_{bus} method is efficient since it avoids complicated procedure, such as lower and upper triangular matrix factorization and Gaussian elimination in the direct method or using the branch-path incidence matrix and primitive admittance matrix in the graph theory based method [104]. Then, having the Z_{bus} matrix (detailed in the Appendix A.3) and the branch-bus incidence matrix from graph theory, we apply the IM method to estimate the bus voltage deviations by the sequential line switch actions. Finally, the voltage deviations from the IM method has been used to find the current deviations after the foregoing switch actions.

For the sake of simplicity in notation, in the remaining of the document δ stands for the SI associated to any flexible resource.

Machine learning proxy

The SI vary with the grid operating conditions, meaning each scenario requires an analytical calculation. To decrease the computational time due to its fast inference time, the Gradient Boosting Trees (GBT) model is used as a proxy to extract functional knowledge relating SI and grid operating conditions. Firstly, a large set of grid operating scenarios is generated with the NORTA (NORmal To Anything) method [105]. Then, the SI are analytically computed for this generated set of grid operating scenarios, and the GBT is fitted over this data using active and reactive nodal power injection as input features. To keep low the number of input variables while leveraging on the localized impact of active/reactive power nodal injections, the Spearman rank correlation coefficient was applied to measure the correlation between the time series of nodal power injection and the SI. Only the nodal injections exhibiting a correlation above 0.5 were selected as input features.

The GBT model is then used to derive SI $\delta^{t+k}_{i \to j,s}$ (Eq. 3.6) for the forecasted scenario s, and for grid element i (i.e., line/bus) and flexible resource j, and using as input a matrix of forecasted active $\hat{\mathbf{P}}^{t+k}_{s}$ and reactive $\hat{\mathbf{Q}}^{t+k}_{s}$ power nodal injections for lead-time $t+k$. The mean absolute percentage error of estimated SI with GBT was less than 1%.

$$\hat{\delta}^{t+k}_{j \to i,s} = f(\hat{\mathbf{P}}^{t+k}_{s}, \hat{\mathbf{Q}}^{t+k}_{s}) \tag{3.6}$$

The analytical method is only used for generating the training set of the GBT proxy, using network operating scenarios generated by the NORTA method. Subsequently, the GBT model is used in the remaining steps of the methodology.

3.2.2 Interpretation from sensitivity indices

The SI relate the rate-of-change in nodal voltage/line current as a linear function of active/reactive power modulation, variation in CB/RS elements, or OLTC tap position. This information could assist the human operator in figuring out the causal relationship between the technical issue and the flexibility options. Fig. 3.3 depicts a causality tree divided into two levels (unrelated to the network topology), which is helpful to guide the human operator analysis to the most suitable flexibility options. In Level 1, the SI represent the influence of each flexibility option on reducing the overload in line i via downward/upward (active power) action over the current operating point of each bus j. Since the activation of flexibility can have an inverse impact on other grid elements (e.g., cause congestion in another line), the SI in Level 2 represent the negative (inverse) impact on other grid elements, which can be used to constrain the activated flexibility in each node j.

From the SI and scenarios, it is possible to compute the contribution C_j of each grid node (load or generator) j forecast uncertainty to the forecasted technical problem in grid element i, via the average deviation of a set of S scenarios around the point forecast $\hat{y}^{*}_{t+k|t,j}$ weighted by the SI:

$$C_j = \frac{1}{S} \sum_{s=1}^{S} \frac{\hat{y}^{*}_{t+k|t,j} - \hat{y}^{s}_{t+k|t,j}}{\hat{y}^{*}_{t+k|t,j}} \cdot \delta^{s}_{j \to i} \tag{3.7}$$

where $\delta_{j \to i}$ is the sensitivity of resource j to grid element i, and for scenario s.

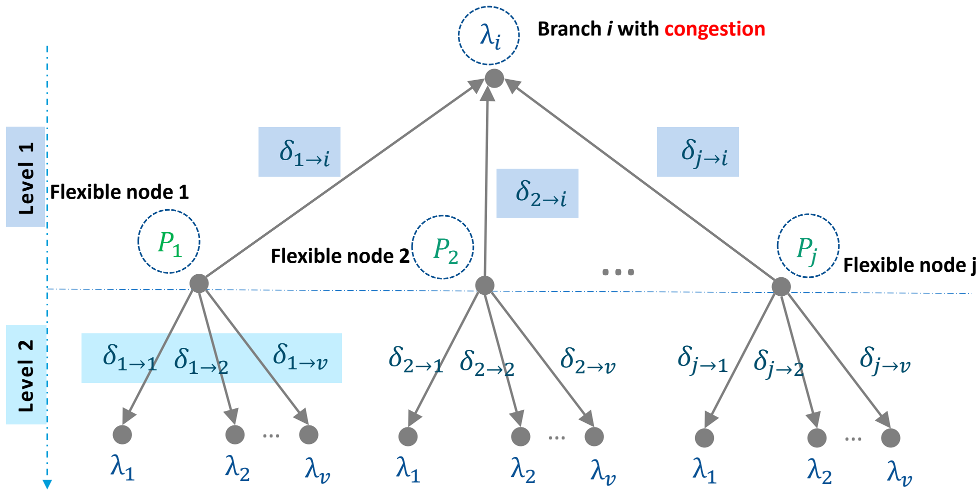


Figure 3.3: *Causality tree for a congested line i and flexibility options in nodes j (Level 1). Level 2 corresponds to the inverse impact in other grid elements by applying the Level 1 actions*

3.2.3 Flexibility modeling

Active and reactive power nodal flexibility

Comprises a FSP aggregating DER or individual RES power plants, DR, and storage system per grid node, providing flexibility via upward/downward actions around their operating point. The FSP can use advanced aggregation functions, such as [106], to effectively manage and optimize the heterogenous DER, considering technical constraints, market dynamics, operational objectives, and different asset owners.

The flexibility value (FV), Eq. 3.8, represents the amount of active/reactive power that is computed per flexible resource j, based on $\delta^s_{j\to i}$, to solve a technical problem considering the max/min operational limit λ^{lim} (e.g., voltage or current limits) and the forecasted operating conditions λ^s_i for each scenario s, where i means the element with a technical issue. The suggested upward ($Flex^{up}_j$) and downward flexibility margin ($Flex^{down}_j$) in the flexibility market is an upper limit for $FV^s_{j\to i}$.

$$FV^s_{j\to i} = \frac{\lambda^{lim} - \lambda^s_i}{\delta^s_{j\to i}} \tag{3.8}$$

As illustrated in Fig. 3.3, a specific flexibility action may create technical issues in other grid elements with opposite SI signs. Thus, Eq. 3.9 is applied to limit the $FV^s_{j\to i}$ value, conditioned by the operational limits of the other grid elements v.

$$FV'^s_{j\to i} = \text{sgn}\left(FV^s_{j\to i}\right) \cdot \min\{|FV^s_{j\to i}|, \min_{\forall v\in V, v\neq i}\{|\frac{\lambda^{max} - \lambda^s_v}{\delta^s_{j\to v}}|\}\}, \quad \text{if:} \quad \text{sgn}\left(\delta^s_{j\to i}\right) \neq \text{sgn}\left(\delta^s_{j\to v}\right) \tag{3.9}$$

In situations where conflicting actions between TSO and DSO may arise during the activation of flexibility, Eq. 3.9 can be applied to limit the $FV^s_{j\to i}$ value. For example, suppose activating a flexible resource within the distribution network leads to technical issues in the transmission network. In that case, the TSO can communicate the parameters of Eq. 3.9 to enable the DSO to compute the corresponding limitation. Alternatively, the TSO can simply communicate the limitation percentage applicable to that specific resource.

Reaching the lower limit λ^{min} for elements with the same SI sign is also possible where the elements are forced to a minimum operational limit (i.e., bus voltage should be above 0.9 p.u.). In this case, Eq. 3.10 is applied.

$$FV'^s_{j\to i} = \text{sgn}\left(FV^s_{j\to i}\right) \cdot \min\{|FV^s_{j\to i}|, \min_{\forall v\in V, v\neq i}\{|\frac{\lambda^{min} - \lambda^s_v}{\delta^s_{j\to v}}|\}\}, \quad \text{if:} \quad \text{sgn}\left(\delta^s_{j\to i}\right) = \text{sgn}\left(\delta^s_{j\to v}\right) \tag{3.10}$$

Moreover, the activation of $FV^s_{j\to i}$ also creates a change in the slack bus active and/or reactive power injection (in the opposite direction of the flexibility use), which can lead to technical problems in grid elements with the same SI sign of $\delta^s_{j\to i}$. Thus, Eq. 3.11 is applied to limit the $FV^s_{j\to i}$ avoiding a second order technical problem due to the slack bus (*sb*) operating point variation.

$$FV'^s_{j\to i} = \text{sgn}\left(FV^s_{j\to i}\right) \cdot \min\{|FV^s_{j\to i}|, \min_{\forall v\in V, sb\in SB}\{|-\frac{\lambda^{max} - \lambda^s_v}{\delta^s_{sb\to v}}|\}\}, \text{if:} \quad \text{sgn}\left(\delta^s_{j\to i}\right) = \text{sgn}\left(\delta^s_{sb\to v}\right) \tag{3.11}$$

where V is the line and bus elements set, and SB is the set of slack buses in the system. The same constraint in Eq. 3.10 is applied for the slack bus but conditioned to grid elements with the opposite SI sign of $\delta^s_{j\to i}$ due to a reverse reaction of the slack bus to the flexibility use.

Active power redispatch

Combined upward/downward action between two (or more) flexible resources/nodes, such that total active power feed-in remains virtually unchanged, but the congestion is removed. Note that, currently, redispatch actions are primarily employed by TSOs and are included in our proposed methodology for completeness. The application of such actions in distribution networks is still a subject of ongoing discussion, as highlighted in [107]. However, initiatives such as redispatch 2.0 and 3.0 in Germany [97] have the potential to extend redispatch actions to distribution networks.

In this case, the individual SI in Eq. 3.8 are replaced by the ones computed by Eq. 3.12 for several n flexibility resources (i.e., $j_1, j_2, ..., j_n$) participating in redispatch, where $b^s_{j\to i}$ is an indicator function equal to $+1$ if the action is upward and -1 if downward.

$$\delta^s_{\{j_1,j_2,...,j_n\}\to i} = \Sigma^{j=j_n}_{j=j_1} b^s_{j\to i} \cdot \delta^s_{j\to i} \tag{3.12}$$

To avoid technical issues in other elements (lines or buses) Eq. 3.9-3.11) are checked for all units participating in re-dispatch.

Reactive power flexibility from CB/RS

The FV for shunt elements follows Eq. 3.8 and moves on discrete steps determined by the steps of the CB/RS unit as formulated in Eq. 3.13. The FV is communicated to the human operator regarding step changes in shunt elements.

$$FV^s_{j\to i} \in \{\pm Q^u_j, \pm 2Q^u_j ..., \pm stp^{max}_j \cdot Q^u_j\} \tag{3.13}$$

where stp^{max}_j is an integer parameter representing the maximum allowed step-change in CB/RS and Q^u_j is the reactive power per step change in shunt element j.

OLTC

In this case, the FV in Eq. 3.8 will be equal to $tap^{[\%]} \cdot tpos$, where $tpos$ and $tap^{[\%]}$ mean tap-position and tap-percentage (i.e., percentage of voltage adjustment versus one-step change in tap-position). Thus, Eq. 3.8 is reformulated to Eq. 3.14, where λ^s_i is voltage in bus i and scenario s.

$$FV^s_{j\to i} = tap^{[\%]} \cdot tpos^s = \frac{\lambda^{lim} - \lambda^s_i}{\delta^s_{j\to i}} \tag{3.14}$$

Moreover, the goal is to determine the required voltage set-point that will be automatically translated to a tap position by the OLTC, using the relation from Eq. 3.15.

$$V\,(\text{set}-\text{point})^s_j = (1 + tap^{[\%]} \cdot tpos^s) \cdot V\,(\text{set}-\text{point})^{ne}_j \tag{3.15}$$

In Eq. 3.15, if the tap position ($tpos^s$) remains zero, the voltage will be unchanged and equal to the voltage set-point of OLTC j in neutral tap position $V\,(\text{set}-\text{point})^{ne}_j$ (i.e., zero tap position). Then, Eq. 3.15 could be revised as Eq. 3.16 using Eq. 3.14:

$$V\,(\text{set}-\text{point})^s_j = (1 + FV^s_{j\to i}) \cdot V\,(\text{set}-\text{point})^{ne}_j \tag{3.16}$$

Like other flexibilities, the FV should be limited by the operating margins of other grid buses using an equation analogous to Eq. 3.9. Similarly, the limitation of flexibility capacity can also arise due to conflicting actions between TSO and DSO, as explained earlier.

Network topology

A combination of switching actions proposed to change the network topology while ensuring it remains a traceable-connected graph. The candidate switching actions are selected based on their *area-of-effectiveness*, defined as the minimum path between two-side buses of the switched line using graph theory. This approach avoids simulating all possible switching actions. Therefore, the closing/opening actions are determined as follows: a) the closing action is applied to normally open lines whose *area-of-effectiveness* includes the congested line; b) the opening action is applied to normally closed lines in the *area-of-effectiveness* of the selected closing action. It was assumed that the opening action only applies when preceded by a closing action within the same area to prevent additional technical issues. As a result, the proposed actions can include a single closing action, sequential closing actions, and sequential closing and opening actions. This type of flexibility is considered in the decision-aid phase of section 3.3 together with the other flexibilities.

Figure 3.4 presents an illustration of finding effective switching actions using graph theory for an open-source electrical grid (MV Oberrhein). In this example, two normally open switches have been considered for comparison: switchers 144 and 61. The green minimal path associated to switch 144 includes the congested line (in red) and can be an effective switching action to solve the congestion problem. In contrast, the blue path associated with switch 61 is a less effective switching action since its area of effectiveness is far from the congested line.

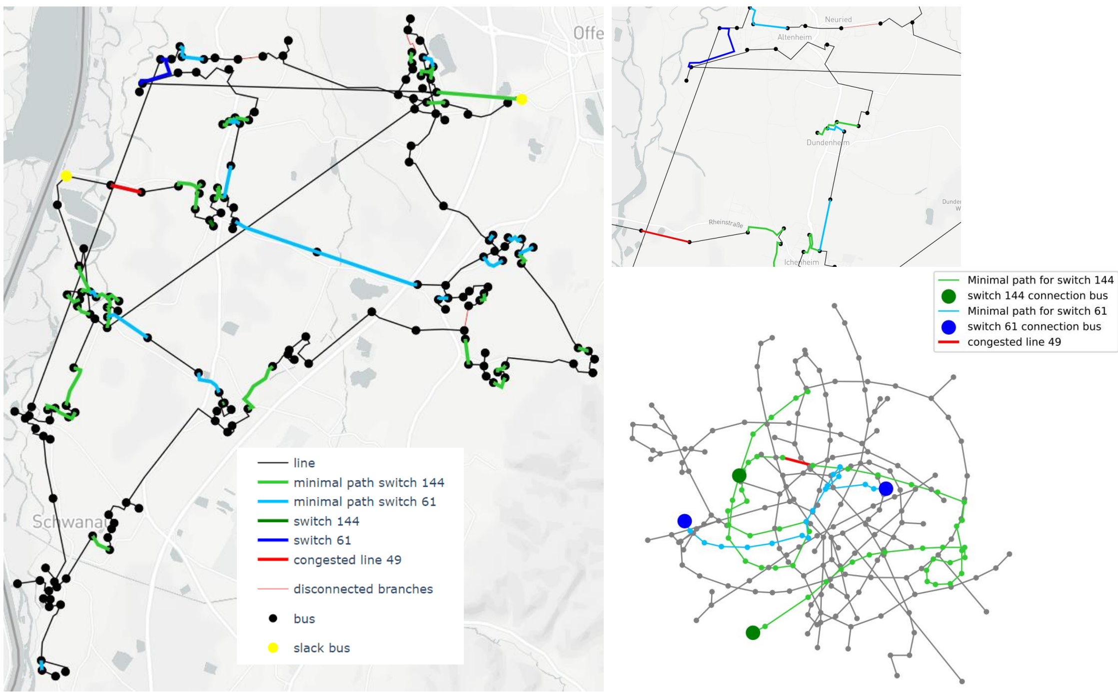


Figure 3.4: *Finding effective switching actions using graph theory for the MV Oberrhein network.*

The line loading after each switching action can be estimated based on Eq (A.28-A.31) from the Z_{bus} method.

3.2.4 Second-level forecasting

The *second-level forecasting* concept builds upon a set of computed probabilistic forecasts (set of quantiles) generated with any forecasting algorithm. In this work, we used a GBT model with feature engineering [108], with truncated generalized Pareto distribution for the distribution's tails [109]. This GBT model was extended with additional input variables for *second-level forecasting*, namely variables engineered with the quantiles from the first-level forecast, i.e., forecast for the day $D+1$ with NWP data generated at 0h00 of day D.

A second model was an Encoder-Decoder Artificial Neural Network (ED-ANN) [110]. The Encoder comprises two layers, which are Long Short-Term Memory (LSTM) network layers. The LSTM has feedback connections and can model temporal or sequential information. The Decoder comprises three internal layers: one LSTM and two fully connected layers, with linear activation function. The Decoder produces the output iteratively, and the hidden state produced at each forecast step generates the next forecast. In the training phase, teacher forcing was used, i.e., using the actual first-level forecast values as input for training the Decoder.

This model is popularly used in sequence-to-sequence (*seq2seq*) problems, where the main goal is to use one sequence of varying length as input to produce an output of not necessarily the same length as the input. One classic example for its application is machine translation. The Encoder takes the input (a set of variables, detailed in subsection 3.2.4), processes it in its internal network, and produces a context vector, a representational state of the input that will then be passed as input to the Decoder. In turn, the Decoder iteratively forms a forecast sequence based on new input variables and using the vector passed from the Encoder as an initial state. In our case, the goal is to take inputs of length L (apart from the dimension corresponding to the number of variables used) to generate a forecast sequence of the same length.

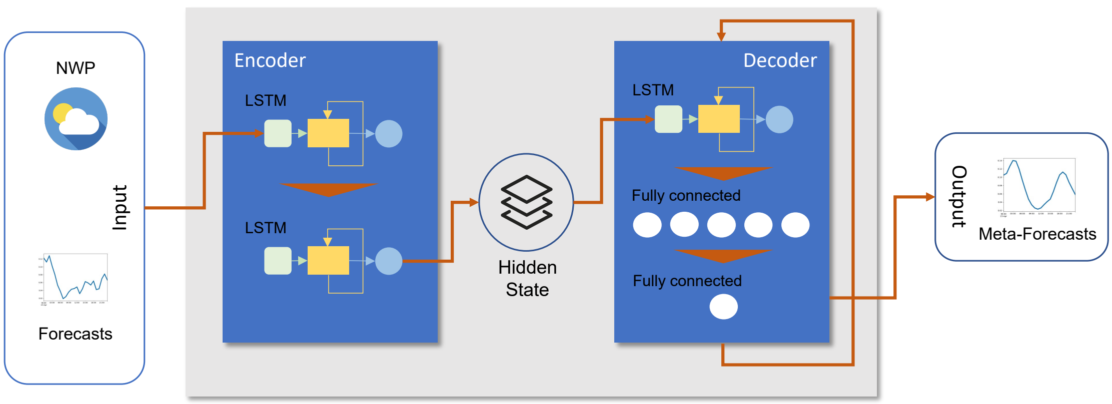


Figure 3.5: *Encoder-Decoder ANN architecture for second-level forecasting.*

Figure 3.5 depicts the architecture of the ED-ANN model. The Encoder comprises two layers, which are LSTM network layers. These structures have feedback connections and are capable of modeling temporal or sequential information, resulting in a good choice for encoding the initial input, given our adopted strategy. The output of the first layer is fed as input to the second layer, which produces a hidden state, a context vector that represents the input. This is what, as mentioned before, is fed to the Decoder. The Decoder is comprised of three internal layers: one LSTM and two fully connected layers, with linear activation function. Having only one unit, the last layer is what generates the output. The Decoder produces the output iteratively, and the hidden state produced at each forecast step is used to generate the next forecast, until a sequence of L values is formed.

Input variables for the second-level forecasting model

The input variables considered for the second-level forecasting models are summarized in Table 3.1.

The set of inputs are divided in different categories. Groups I and II refer to variables which are also input in the model used to compute the quantile forecasts at 00h00 in day D for $D+1$. In Group I, they are all averages of the NWP for the locations of the wind turbines in each wind power plant. Moreover, “Yes” in column “Lags/Leads” means that variables based on lags of the referenced inputs were also included in the model, namely for the previous 1, 2 and 3 (lags) time intervals and also for the following 1, 2 and 3 (leads) time intervals, as described in [108].

As for Group III, unless the target variables are either the 5% or the 95% quantiles, it is always composed of three variables. We use the edge quantiles to introduce more information about the first-level forecast.

Group IV pertains to statistics computed around the target quantile of the second-level forecast. The goal is to include information regarding neighbouring quantiles, such as averages and ranges between equally distant quantiles in respect to the target, as well as a simple inter-quantile range, which are proxies for the uncertainty level.

Group V includes variables computed from a poor man’s ensemble (i.e., for a given time interval, several forecasts available from different launch times in the past). The available NWP had a forecast horizon of up to 90 hours and with 12 hours update rate. Given that we are producing quantile forecasts at 00h00 of day D for day $D+1$, it means that we have 4 different available NWP runs for day $D+1$ (namely 24, 36, 48 and 60 hours before). The use of these variables did not yield any improvement in almost all quantiles. In fact, on average, the model performed was worse than the best model, i.e., -0.17% improvement in the second-level forecast for the 12h forecast and -0.03% for the 2-hours ahead forecast. Note

Table 3.1: *Listing of the input variables for the second-level forecasting model.*

Group	Type of variable	Name of variable	Lags/Leads	Description
I	NWP	mean_mod10	Yes	Average of wind module at 10 meter for all turbines of target wind farm
		mean_mod100	Yes	Average of wind module at 100 m for all turbines of target wind farm
		mean_dir10	Yes	Average of wind direction at 10m for all turbines of target wind farm
		mean_dir100	Yes	Average of wind direction at 10m for all turbines of target wind farm
		num_0	No	First component of the PCA decomposition of the NWP variables across all wind farms
		num_1	No	Second component of the PCA decomposition of the NWP variables across all wind farms
		num_2	No	Third component of the PCA decomposition of the NWP variables across all wind farms
II	Calendar variables	hour_sin	No	Hour transformation into a circular variable using sine function (for hour t, $\sin \frac{t*2\pi}{24}$)
		hour_cos	No	Hour transformation into a circular variable using sine function (for hour t, $\sin \frac{t*2\pi}{24}$)
III	Forecasted quantiles (at 00h00)	q05	No	Quantile 5% of the forecasts computed at 0h00
		q95	No	Quantile 95% of the forecasts computed at 0h00
		q{target}	No	Target quantile of the forecasts computed at 00h00
IV	Statistics from forecasted quantiles (at 00h00)	iqr_gen	No	Inter-Quantile Range (Difference between quantile 75% and quantile 25%)
		qr0_{target}	No	Quantile range (difference) between previous and following quantiles of target quantile
		qr1_{target}	No	Quantile range (difference) between second previous and second following quantiles of target quantile
		mean_q{target}	No	Average of target quantile, the two previous quantiles and the following two
		std_q{target}	No	Standard deviation of target quantile, the two previous quantiles and the following two
V	Statistical variables from consecutive NWP computations	mean_mod10_mean	No	Average of mod10_mean variable for several NWP forecast runs
		mean_mod10_std	No	Standar deviation of mean_mod10 variable for several NWP forecast runs
		mean_mod100_mean	No	Average of mean_mod100 variable for several NWP forecast runs
		mean_mod100_std	No	Standard deviation of mean_mod100 variable for several NWP forecast runs
		mean_dir10_mean	No	Average of mean_dir10 variable for several NWP forecast runs
		mean_dir10_std	No	Standard deviation of mean_dir10 variable for several NWP forecast runs
		mean_dir100_mean	No	Average of mean_dir100 variable for several NWP forecast runs
		mean_dir100_std	No	Standard deviation of mean_dir100 variable for several NWP forecast runs

that the "meteo-risk index" proposed in [111] to reflect the spread of the European Centre for Medium-Range Weather Forecasts (ECMWF) Ensemble Forecasts (ENS) was also tested as an input variable, but does not seem to have very much influence on the forecasting accuracy, i.e., the improvement was below 2.9% (when averaging per quantile), which does not justify the cost associated to the ENS data.

Finally, it is important to specify that the Decoder portion of the model requires an initial value to start the generation of the forecast sequence. As such, we decided to use the last value generated in the previous first-level forecast (for day D generated in day $D-1$), which is valid sense since this value is available at the moment of second-level forecasting computation for day $D+1$ on day D. Moreover, in the training phase, teacher forcing was used (using the actual first-level forecast values as input for training the Decoder, which improves model accuracy).

In our experiments, different models and variables were tested. Some the variations that were considered are:

Var1. Using Groups I-IV and, as a variation of Group I, including the same variables but for the previous NWP runs (not including lags and leads);

Var2. Same as the previous model, but not using lags and leads for any NWP run;

Var3. Same as the previous model, but using lags and leads for both NWP runs;

Var4. Not including Groups IV and V;

Var5. Using all groups

Var6. For Group V, experimenting with a variation, consisting in using averages and standard deviations of the differences in NWP for consecutive runs.

The variation considered the best performing was Var1, and the comparison of performances is outlined in Table 3.2. It shows the average performance per WPP (in terms of percentage improvement over the naive benchmark). It is important to note that the choice for best performing model came from a combination of the results in both subtables: While in the second-level forecast for the 12h forecast (Table a) the best performing model is actually *Var3* (a difference of 0.19 percentage points, or pp), when considering the second-level forecast for the 2-hours ahead forecast, *Var3* has a greater advantage over the first (a 1.64pp difference). In turn, when looking at *Var6*, it performs better on average in the 2-hours ahead forecast (only slightly, by 0.02pp) but is outperformed by our best model with a greater margin (a difference of 0.06pp).

Table 3.2: *Average improvement (percentage) per WPP for several model variations.*

(a) *second-level forecast for the 12h forecast*

WPP	Best Model (Var1)	Var2	Var3	Var4	Var5	Var6
		Variations				
WPP00	7,51	2,21	7,71	10,71	6,41	9,09
WPP01	15,2	9,63	13,67	12,96	16,44	14,48
WPP02	6,01	8,84	6,21	8,95	8,67	4,03
WPP03	14	11,31	15,04	15,86	15,52	12,93
WPP04	11,25	7,81	12,55	3,96	10,39	9,9
WPP05	9,12	9,82	10,94	9,91	10,45	8,48
WPP06	10,59	7,78	6,91	7,85	6,9	7,62
WPP07	6,65	4,38	6,81	5,72	6	6,7
WPP08	2,17	3,67	4,4	6,03	1,14	2,24
Average	9,17	7,27	9,36	9,11	9,10	8,39

(b) *second-level forecast for 2 hours-ahead forecast*

WPP	Best Model (Var1)	Var2	Var3	Var4	Var5	Var6
		Variations				
WPP00	17,25	15,43	12,61	16,37	17,07	14,6
WPP01	18	18,36	15,83	16,76	17,2	12,32
WPP02	12,21	11,95	13,82	10,78	12,87	11,91
WPP03	13,42	13,17	7,19	12,33	12,75	11,29
WPP04	11,85	9,38	14,91	9,71	12,42	11,73
WPP05	16,03	20,36	17,32	14,05	17,71	12,96
WPP06	11,4	10,78	9,6	10,58	10,98	9,31
WPP07	13,88	11	12,59	8,65	13,35	10,48
WPP08	12,53	12,76	7,88	8,15	12,37	8,56
Average	14,06	13,69	12,42	11,93	14,08	11,46

Finally, Figure 3.6 depicts the timeline of inputs for the ED-ANN model. The model uses the Encoder portion to process information regarding day D (with forecasts, NWP and other listed variables computed on day $D-1$) and then uses information about day $D+1$ on the Decoder (computed on day D) to produce the forecasts. Hence, except for the last variable group, all variables are doubled: one set for the Encoder (day D) and another set for the Decoder (day $D+1$, see Figure 3.6). It should be mentioned that the same methodology has been followed for predicting the tails (i.e.,quantiles - less than 5% and more than 95%).

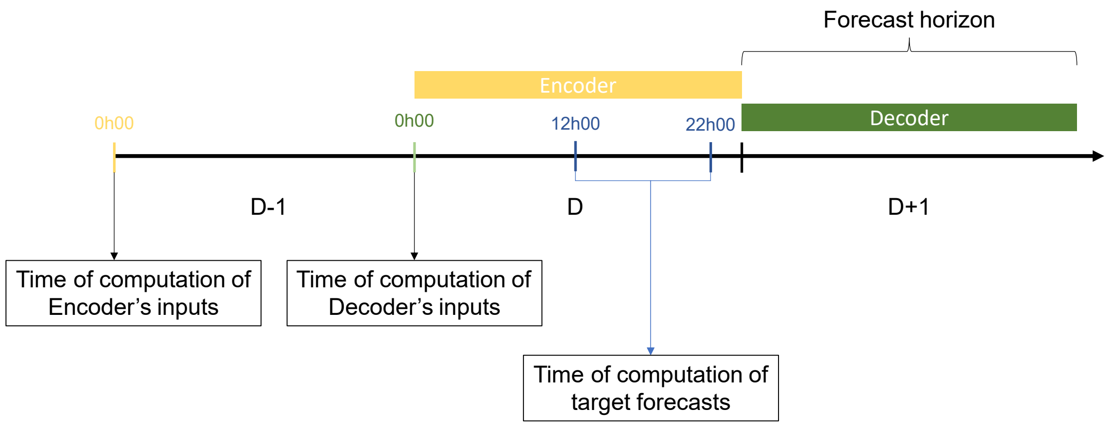


Figure 3.6: *Timeline of the inputs used for the Encoder-Decoder model.*

Results and second-level forecasting model comparison

The second-level forecasting skill is evaluated with the Mean Squared Error (MSE) considering the difference between forecasted quantile ($\hat{q}^{\alpha}_{t+k|t+z}$) and second-level forecast of the same quantile ($\widehat{q}^{\alpha}_{t+k|t}$). For assessing the performance of the GBT and ED-ANN models, a naive benchmark model where $\widehat{q}^{\alpha}_{t+k|t} = \hat{q}^{\alpha}_{t+k|t}$ was used, assuming that there is no variation in the quantile forecasts with information (NPW and/or power observations) update. In other words, the forecast generated at 00h00 in day D remains the same for the second-level forecasts. The improvement (Eq 3.17) over benchmark model was computed for model mi, and divided by i) time series, ii) quantile, and iii) hour of the day. As mentioned, two second-level forecasts are considered, one for the forecast updated with NWP generated at 12h00 and another for 2 time intervals ahead $(t+2|t)$.

$$Improvement_{mi}^{MSE} = \left(1 - \frac{MSE_m}{MSE_{benchmark}}\right) \times 100\% \tag{3.17}$$

For our experiments, we used data from 9 WPP. Data from April 2019 to March 2020 was used for model training, and from this portion, about 25% random samples were selected for validation. The testing period spanned from April to September 2020. Note that the input and target quantiles of the train set have been produced through k-fold cross validation over the training period.

Figure 3.7 depicts, for some days, illustrative examples of first-level forecast and the two second-level forecasts for different WPP and quantiles. In the first hours of the day for the first three examples, the ED-ANN seems to be closer to the actual forecast than the GBT and naive models. The results presented in this section will show that this is the general case for the models' performance. Also, for the second example (WPP03, quantile 50%), the ED-ANN seems to be the better performing model for most of the hours of the day. The bottom two examples in the figure show more comparable performances of the three models, with no clear advantage for any specific model except for the last example, where GBT appears to perform very well on the second part of the day.

Starting with the second-level forecast for the 12h00 forecast launch time in day D, Figure 3.8 shows the improvement over benchmark model per WPP, we can see that ED-ANN has results that are comparable with the GBT, although the improvement is greater for the former in 5 different WPP. The differences in improvement are no greater than 3.5pp.

Figure 3.9 depicts the average improvement per quantile. One general aspect of the results is that the improvement over benchmark is more significant on the extreme quantiles, and lowers towards the middle quantiles. The GBT outperforms the ED-ANN in the extreme quantiles (from 0.01% to 5% and also 90% and above), with improvements up to 10 pp greater (quantile 0.01%). The ED-ANN is the best performing model for the middle quantiles (10% to 85%).

In Figure 3.10, when analyzing the results per hour of day, an interesting trend is revealed: the ED-ANN model has a frankly better performance for the first hours of the forecast horizon. However, this is not unexpected, given the architecture of the model. Since the Decoder builds

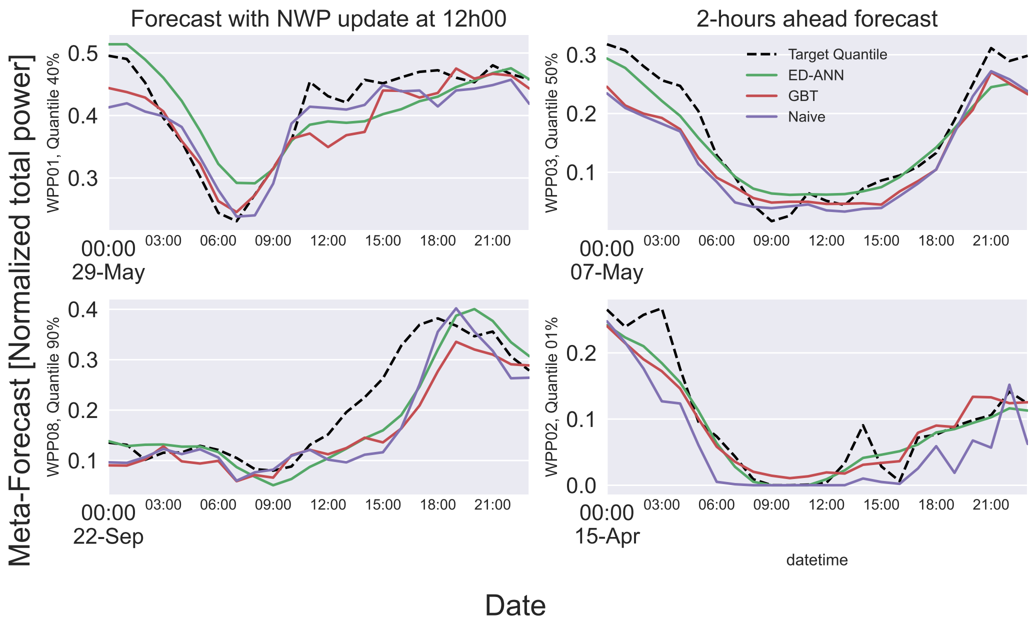


Figure 3.7: *Examples of second-level forecasts for different WPP and quantiles.*

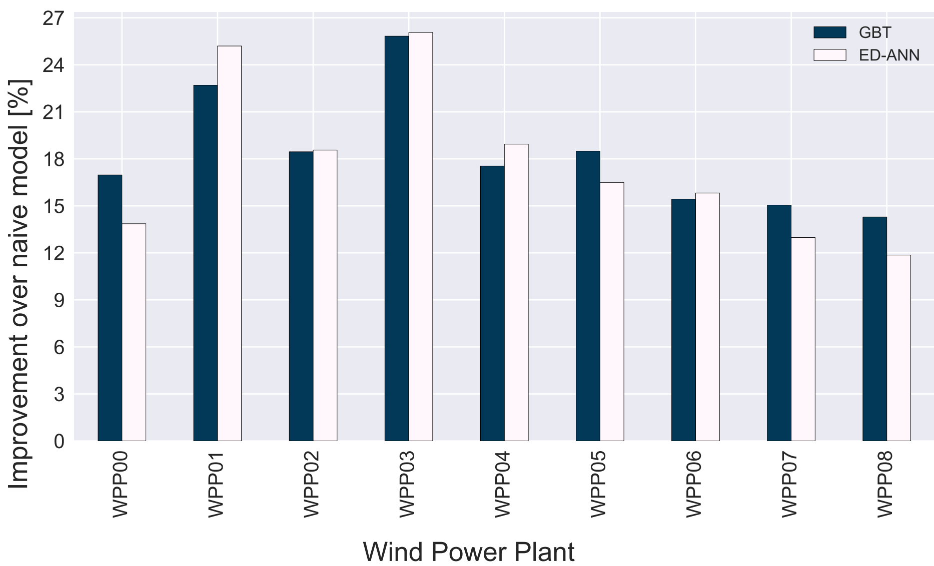


Figure 3.8: *Improvement over naive model per WPP (second-level forecast for the 12h00 forecast)*

each forecast sequentially in an iterative manner, it is somewhat common that the first steps are significantly good, but since each value generated at each step is fed into the model to compute the next one, errors in forecast gradually accumulate. Also, there is a possibility that our choice of initial value for the Decoder portion of the model helped improve the forecast performance.

For the ED-ANN, the improvement for the first three hours is 57.81%, 37.35% and 25.74%, respectively. Then it decreases until it reaches a minimum of 7.72%. However, GBT improves

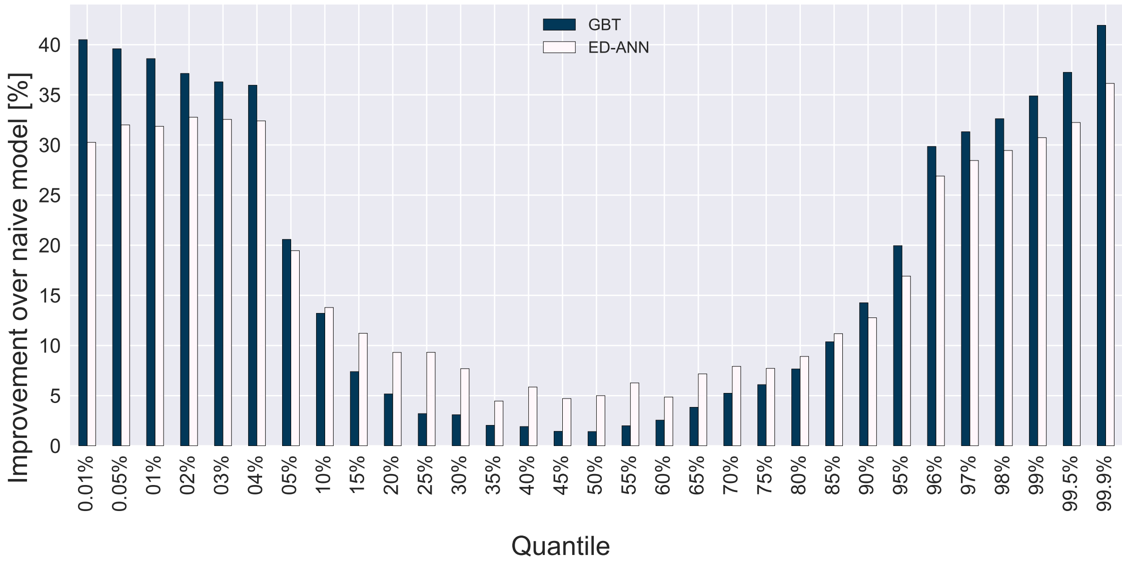


Figure 3.9: *Improvement over naive model per quantile (second-level forecast for the 12h00 forecast)*

on the naive (benchmark) model more than the ED-ANN for the rest of the hours of the day, reaching a maximum of 24,52% in the end of the day (at 21h).

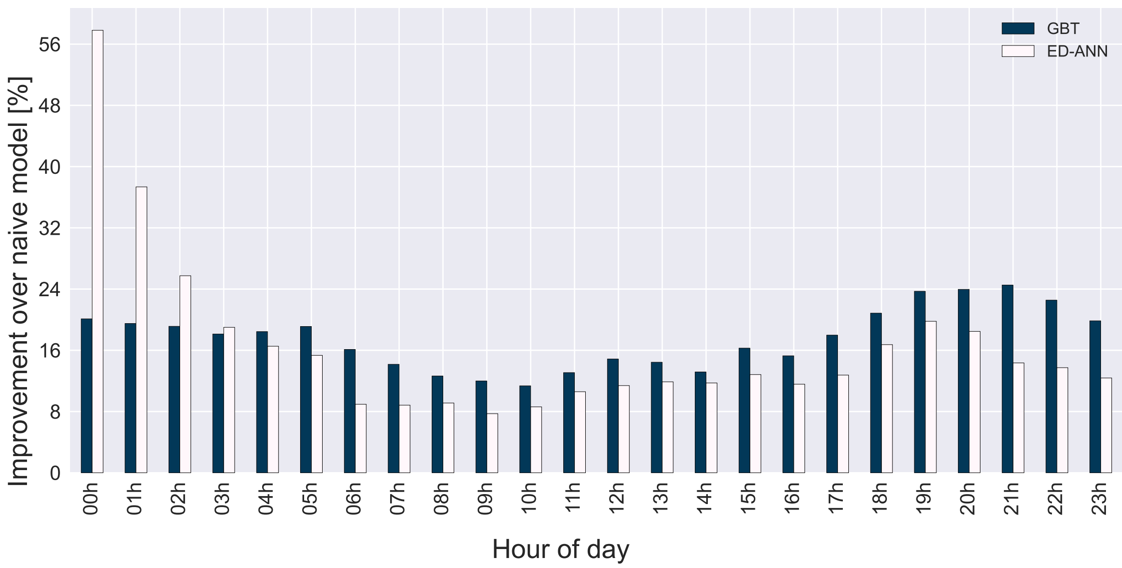


Figure 3.10: *Improvement over naive model per hour of day (second-level forecast for the 12h00 forecast)*

For 2 hours-ahead forecast, the performance is generally better for the ED-ANN model. Regarding the improvement averaged by WPP (see Figure 3.11), the ED-ANN outperforms the GBT model for 7 in 9 WPP. They both improve on the naive model and are comparable in some instances (the smallest differences in improvement are in the WPP where the GBT performs better, and that difference is of , at most, 1.50 pp).

The average improvement in the 2 hours-ahead second-level forecast per quantile is depicted

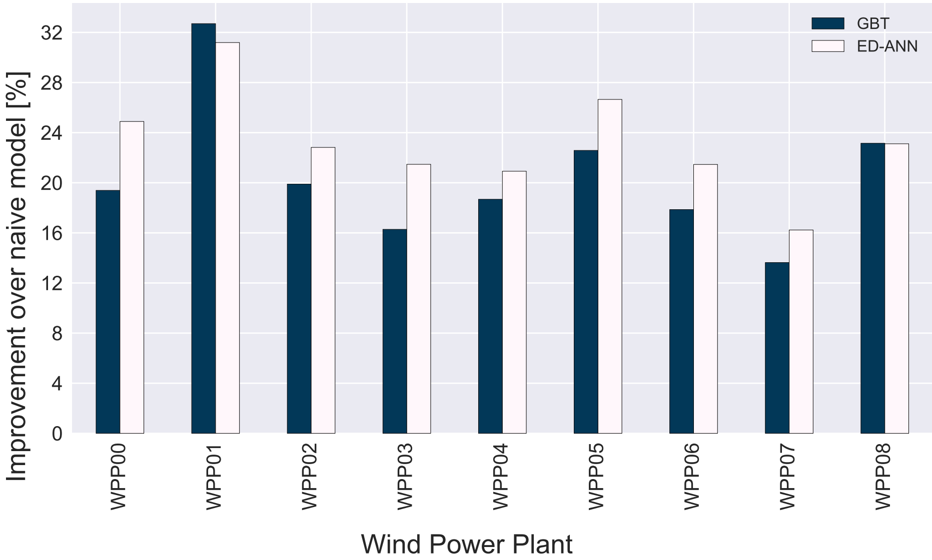


Figure 3.11: *Improvement over naive model per WPP (second-level forecast for 2 hours-ahead forecast)*

in Figure 3.12). While the GBT has a better (but comparable) improvement then the ED-ANN from quantiles 0.01% to 04%, it is outperformed for the remainder of the quantiles. The same pattern of having better performance in the extreme quantiles can be noted, just like for the forecasts with NWP update at 12h00. Again it can be noted that both models always improve on the naive model.

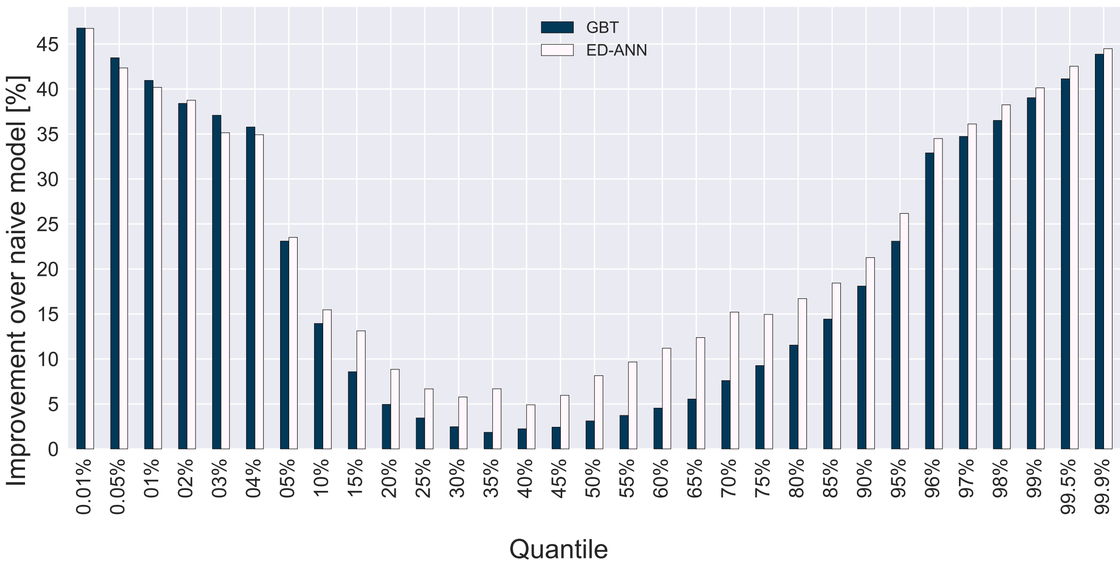


Figure 3.12: *Improvement over naive model per quantile (second-level forecast for 2 hours-ahead forecast)*

Finally, the results per hour of day (see Figure 3.13) reveal the same pattern that we commented before: the presence of a significantly higher improvement in the first hours of the forecast horizon. The improvement in the first hour is around 61%, and then decreases. Although both models outperform the benchmark for all hours, there are some instances where the GBT is the better model (from 06h to 09h and from 21h to 23h). There are even some hours where the models' results are practically the same (see hours 12h to 14h).

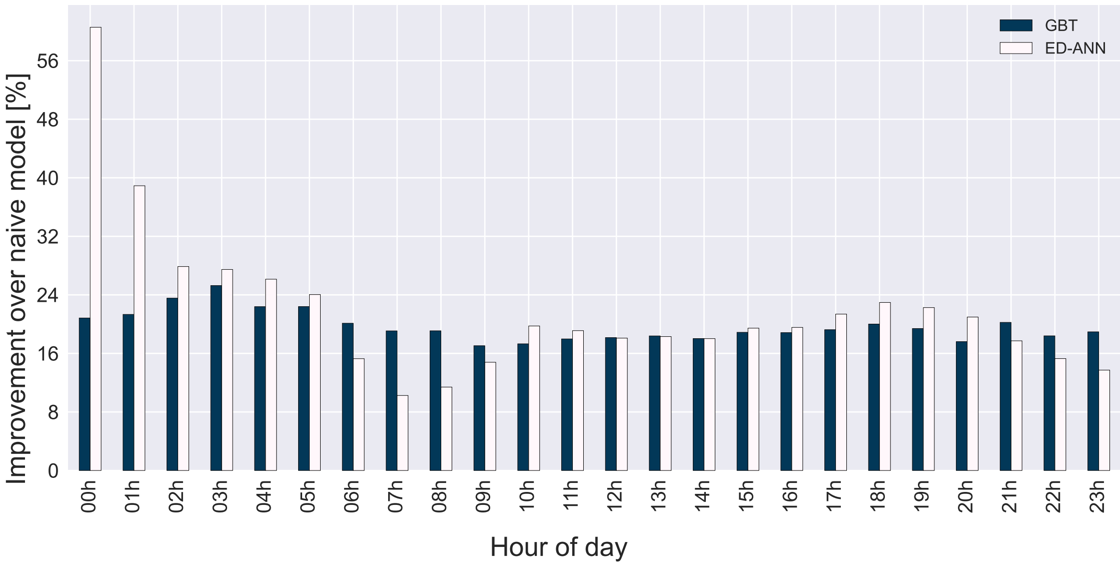


Figure 3.13: *Improvement over naive model per hour of day (second-level forecast for 2 hours-ahead forecast)*

To conclude, these forecasting results show potential for exploring the second-level forecasting concept, in particular the ED-ANN architecture to produce second-level forecasts. The addition of new inputs, such as statistics computed from the forecasted quantiles and also the introduction of poor man's ensembles improved the model performance. However, there is potential for improvement, namely in increasing the forecasting skill of lead times greater than $t+12$, which our proposed framework struggles with keeping a stable performance.

3.3 Decision-aid Phase

3.3.1 Flexibility ranking

After calculating the $FV^s_{j\to i}$ (see section 3.2), the next step is ranking the individual flexibility options according to their *effectiveness* in solving the technical problem under forecast uncertainty. Firstly, the line loading/bus voltage after flexibility use is computed for each scenario:

$$\lambda'^s_i = FV^s_{j\to i} \cdot \delta^s_{j\to i} + \lambda^s_i \tag{3.18}$$

where λ^s_i and λ'^s_i are the line loading/bus voltage of element i before and after flexibility activation.

Secondly, for each scenario, s, the severity function from [17] associated with line congestion/voltage violation (see Eq. 3.19) is applied to the value of λ'^s_i, where $\lambda^{\%}_i$ is the loading percentage of the line i (an analogous equation is applicable for voltage problems).

$$Severity_i = 0.1 \cdot \left(\lambda_i^{\%} - 90\right) \tag{3.19}$$

The flexibility cost corresponding to $FV^s_{j\to i}$ is also computed considering the flexibility bid price. The cost of the system operator assets was computed with the methodology from [112], described in Appendix A.1.

It is important to underline that the goal here is to rank each flexibility option. Thus, cost and severity are computed for each scenario and flexibility option. Then, risk metrics are computed from the severity and flexibility cost, namely: expected value and value-at-risk (VaR) of the flexibility cost, expected value and Value-at-Risk (VaR) of the severity, probability of over/under-voltage, or congestion. In situations involving conflicting actions between TSO and DSO, an alternative approach to constrain the $FV^s_{j\to i}$ in Eq. 3.9 (as explained in section 3.2.3) is to incorporate risk metrics that quantify the potential impact on the TSO or DSO networks. This can include metrics such as the probability of congestion occurring in an upstream line within the TSO network.

Note that these metrics are computed for all flexibility options described in section III-C, including network reconfiguration actions (whose impact is estimated with the Z_{bus} matrix – see section 3.2.1). Then, the TOPSIS method is applied to rank the flexibility options (i.e., set of alternatives), using the risk metrics as criteria. The TOPSIS method is then used to rank the flexibility options (i.e., set of alternatives) with the risk metrics as multi-criteria. This technique ranks the options based on the Euclidean distance to the Positive Ideal Solution (PIS) and Negative Ideal Solution (NIS) . For decision matrix $X_{M\times N}$ with M options and N criteria, $v_{M\times N}$ ($v_{M\times N} = \acute{X}_{M\times N} \cdot W_{N\times 1}$) is the weighted decision matrix considering $W_{N\times 1}$ as weighing matrix and $\hat{X}_{M\times N}$ as the normalized version of $X_{M\times N}$:

$$\acute{x}_{mn} = \frac{x_{mn}}{\sqrt{\Sigma_{k=1}^{k=M} x_{kn}^2}} \tag{3.20}$$

where $\acute{X}_{M\times N} = (\acute{x}_{mn})_{M\times N}$ and $X_{M\times N} = (x_{mn})_{M\times N}$). In this case all the mentioned metrics has negative impact, therefore:

$$\text{PIS} = \{\min_{\forall m\in M} v_{m,n} | n \in N\} \tag{3.21}$$

$$\text{NIS} = \{\max_{\forall m\in M} v_{m,n} | n \in N\} \tag{3.22}$$

Then, the geometric distance from PIS and NIS for each solution denoted by d_m^+ and d_m^-, respectively.

$$d_m^+ = \sqrt{\sum_n (v_{m,n} - \text{PIS})^2} \tag{3.23}$$

$$d_m^- = \sqrt{\sum_n (v_{m,n} - \text{NIS})^2} \tag{3.24}$$

Finally, the options will be scored by similarity to worse solution coefficient, where the option with best condition has the highest SWS value (Eq 3.25):

$$\text{SWS}_m = \frac{d_m^-}{d_m^- + d_m^+} \tag{3.25}$$

Figure 3.14 depicts an example of multiple flexibility options ranked with TOPSIS for a line congestion problem.

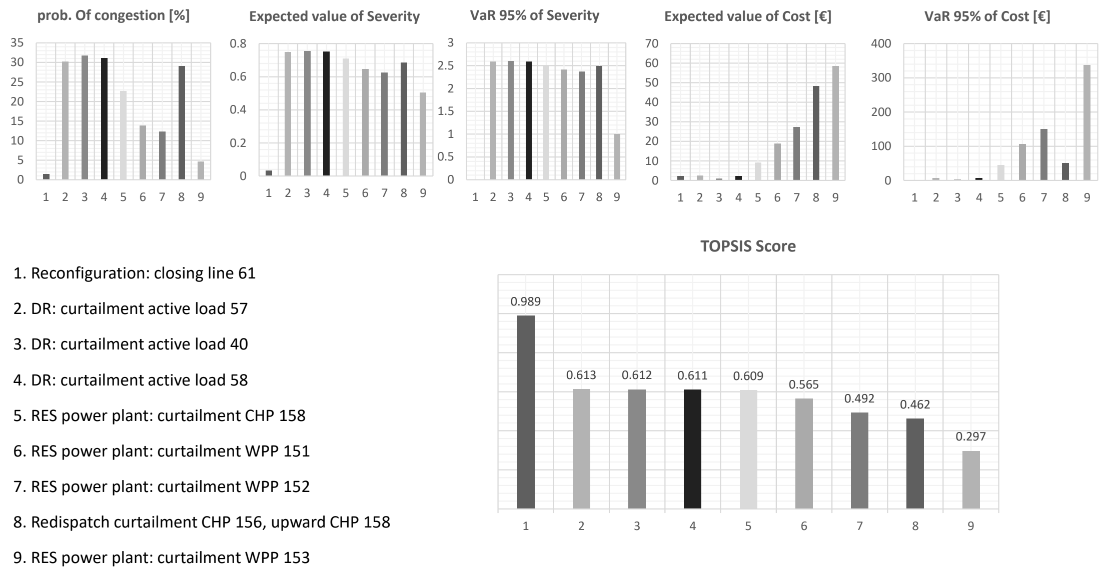


Figure 3.14: *TOPSIS score over five risk metrics considering multiple potential solutions*

There is no differentiation between flexibility types during the ranking process. For instance, if a network switching action ranks among the top solutions, it will be included in the set of actions used to construct the risk-cost curves discussed in the following subsection.

3.3.2 Risk-cost curves

The top three to five flexibility options are selected for combination following the flexibility ranking. All the possible combinations, including binary actions and actions with discrete and continuous values, are considered. Only the non-dominated ones create the risk-cost curve among all these possible combinations. Therefore, there will be several possible combinations of flexibility options, which lead to the grey points in the risk-cost mapping from Figure 3.15 for a congested line problem. Among all these possible combinations, only the non-dominated ones are used to construct the final risk-cost curve, as depicted in the blue curve of Figure 3.15, where risk is represented by the probability of congestion in line i.

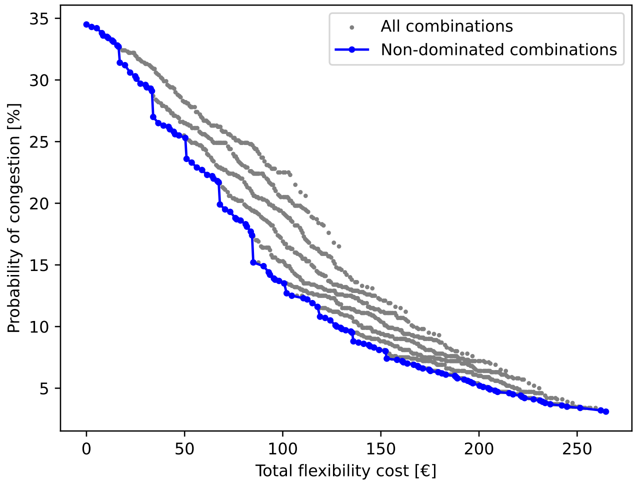


Figure 3.15: *Selection of non-dominated solutions from the combination of options*

This approach is applied to the uncertainty forecast available in the present time and to the *second-level forecasts* generated with the methodology described in section 3.2.4, which will result in multiple risk-cost curves (one for each forecast).

3.3.3 Decision-making paradigms

The final step (*decision*) requires the involvement of a human decision-maker to select a preferred solution from the risk-cost curve(s). Two alternative decision-making paradigms are considered:

Maximum risk-threshold

Sets a maximum threshold for the risk level conditioned by the decision stakes. In this work, the stakes are directly linked to the risk of cascading failures in the grid if the technical problem is a congested line. A similar concept can be derived for voltage problems, e.g., propagation effects of over-voltage in the protection system. The cascading risk is computed by simulating (using the Z_{bus} method, in each scenario s) the impact on all the grid lines of tripping the congested line i and computing the expected number of lines tripping. This expected value is normalized between 0 and a maximum pre-defined value for the cascading simulation and re-scaled into a scale of integer numbers that is easy to analyze by a human operator (e.g., between 1 and 10).

It is necessary to find, via interaction with the human decision-maker, a functional relation between stakes (risk of cascading) and maximum risk threshold so that stakes condition the risk level in each decision. A simple method to find this functional relation is a direct rating (e.g., the decision-maker is asked to state the maximum risk for stakes equal to 10). Still, other methods like direct mid-point can be used [113].

Risk-cost trade-off

The preferred solution is the one with the minimum equivalent cost, $Eqcost$ (Eq. 3.26), obtained by considering a trade-off value μ between risk and cost for each solution p in the curve. The trade-off value can be determined by interaction with the decision-maker, such as indifference judgments [113], and estimated for each stake level, creating a functional relation between trade-off value and stakes. In this case, the higher the stakes, the higher the trade-off value, i.e., the decision-maker is willing to pay more to decrease risk.

$$Eqcost_p = cost_p + \mu \cdot risk_p \tag{3.26}$$

3.4 Numerical Results

3.4.1 Case-study description

The 20 kV Oberrhein MV network supplied by two 25 MVA HV/MV substations[1] was used with some modifications to increase RES penetration and create technical problems. The network supplies 141 MV/LV (secondary) substations and 61.86 MW loads (peak power) through four MV feeders. The topology is meshed but is operated as a radial grid. It contains 147 consumers, nine WPP, four Combined Heat and Power (CHP) units, and three storage systems. The secondary substation load measurements were taken from the Iowa Distribution Test System [114], and the wind power measurements are from WPP in France (the location cannot be disclosed for confidentiality reasons). The NWP are from the ECMWF High-Resolution Forecast model. Without loss of generality, the uncertainty forecast has been applied only for WPP, and a perfect forecast was used for the substation load and CHP. The training period is from 2019-04-01 to 2020-03-30, and the testing period is six months, starting on 2020-04-01.

In terms of flexibility, the bid by generation units (i.e., WPP and CHP) was considered as 30% of the point forecast for upward/downward directions, while the upward direction for WPP was not considered. As assumed in [115], a shorter notification time can increase the flexibility price. Thus, the following was considered: 40% price increment for booking flexibility 12 hours later (assuming the present moment as the reference time) and 90% for booking two hours before the delivery time. A sensitivity analysis of these prices is presented in Section 3.4.6. Without loss of generalization, it was assumed that availability and activation prices for flexibility are equal.

For DR, we considered a specific time schedule for its availability, and the maximum flexibility

[1] `https://pandapower.readthedocs.io/en/v2.0.0/networks/mv_oberrhein.html`

Table 3.3: *Functional relation between stakes and risk threshold/trade-off value for different DM*

	Stakes (ρ) range		
Decision-making approach	$0 \leq \rho \leq a^*$	$a < \rho \leq 7$	$7 < \rho \leq 10$
DM A: Maximum risk threshold	10	6	3
DM B: Maximum risk threshold	20	15	10
DM C: Maximum risk threshold	25	20	20
DM D: Risk-cost trade-off	30	50	70
DM E: Risk-cost trade-off	70	90	110

* Parameter a for DM A is equal to 3, for DM B, DM D and DM E is equal to 4, and for DM C is equal to 2.

is 30% of the electrical energy consumption in a specific hour for upward/downward directions. In the case of storage units, flexibility was determined by the FSP as a function of battery state-of-charge and depth-of-discharge.

To evaluate the impact of different human DMs, different curves relating stakes and risk threshold/trade-off are considered in Table 3.3.

The proposed methodology can address both congestion and voltage issues. However, the main focus was on congestion problems.

3.4.2 Benchmark Strategies and Evaluation Metrics

For benchmark, the following strategies were considered:
T2D: Novel concept proposed in this work where the operator decides i) book flexibility now or ii) wait for the next forecast update and book flexibility later.
Decision-Now (DN): The operator, using the information from uncertainty forecasts, decides to book flexibility at the lowest (availability) cost based on forecasts generated with NWP generated at 0h00.
Deterministic 1 (D1): The operator decides, based on a deterministic forecast, to i) book flexibility now or ii) wait for the next forecast update and book flexibility later.
Deterministic 2 (D2): The operator always decides based on the very short-term forecast, i.e. $t+2|t$.
The *probabilistic method from [91]*.

The overall evaluation of the results considered two evaluation matrices: the classical confusion matrix and the cost-loss matrix (see Table 3.4). The confusion matrix is used to evaluate the performance of each strategy in detecting congestion problems, from which F-score accuracy metrics are computed to measure the balance between the precision (i.e., dividing the

Table 3.4: *Cost-loss matrix to evaluate the performance of the predictive grid management process*

	Event occurred	**Event did not occur**
Action taken	Cell A: Rate-of-occurrence (h%) Flex cost (C) + Loss (L)	Cell B: Rate-of-occurrence (m%) Flex cost (C)
Action not taken	Cell C: Rate-of-occurrence (f%) Loss (L)	Cell D: Rate-of-occurrence (c%) Zero cost

true positives (TP) by anything that was classified as a positive) and recall (i.e., dividing the True Positives (TP) by anything that should have been classified as positive).

The cost-loss matrix draws inspiration from the threshold-based cost-loss analysis method described in [116] for evaluating weather forecasts. For this problem, the matrix was adapted to the following: flex cost (C) corresponds to the preventive actions cost (i.e., flexibility cost); loss (L) corresponds to real-time emergency action (i.e., load or generation curtail) to solve the congestion problem with a monetary cost corresponding to the Value-of-Lost-Load (VoLL), considered to be 12000 € /MWh [117]. The table can be summarized into the summation of the matrix elements, each weighted by the percentage of occurrence, which is performance indicator γ:

$$\gamma = (C + L) \cdot h + C \cdot m + L \cdot f + 0 \cdot c \tag{3.27}$$

3.4.3 Second-level forecasting performance

The set of input variables is divided into five categories. Group I consists of variables computed from the NWP data (i.e., generated at 0h00 in day D), namely: the wind module and direction at 10 and 100 meters, averaged for the locations of the wind turbines in each wind power plant, lags and leads ($t \pm 3$) of these variables, as described in [108]. The first three components of the PCA decomposition of the NWP variables across all wind power plants (WPP) are also part of this group. Group II pertains to calendar variables, namely sine and cosine transformations of the hour of the day.

Group III are variables directly extracted from the first-level forecasts, namely the quantiles 5%, 95%, and the target quantile. Group IV is a proxy for the uncertainty level extracted from the first-level forecasts, namely: the difference between quantiles one position away from the target quantile (i.e., previous and following quantiles); the difference between quantiles two positions away from the target quantile; the average and standard deviation of the set of target and neighboring quantiles (two before and two after the target quantile); inter-quantile range.

Table 3.5: *List of input data variations for the second-level forecasting model*

Name	Variable Groups	Modifications
Var1	I-IV	Include another set of variables, identical to Group I, but with NWP variables generated at 0h00 of day D-1. This set does not include lags and leads
Var2	I-IV	Same as Var1 but not using lags and leads in both Group I and the identical set added
Var3	I-IV	Same as Var1, using lags and leads in both Group I and the identical set added
Var4	I-III	
Var5	I-V	
Var6	I-V	Group V is modified using averages and std. dev. of the differences in NWP consecutive runs

Group V includes variables computed from a poor man's ensemble [111], given that we have a time horizon of up to 90 hours and 12 hours NWP updates. Using these variables did not yield any improvement in almost all quantiles. In fact, on average, the model performed was worse than the best model, i.e., -0.17% improvement in the *second-level forecast* for the 12h forecast, and -0.03% for the 2-hours ahead *second-level forecast*. The meteo-risk index proposed in [111] reflects the spread of the ECMWF ensemble forecasts was also tested as an input variable. The improvement was below 2.9%, which does not justify the cost associated with this data. Different variants of input variables were tested, as shown in Table 3.5.

The *second-level forecasting* skill is evaluated with the MSE considering the difference between forecasted quantile ($\hat{q}^{\alpha}_{t+k|t+z}$) and *second-level forecast* of the same quantile ($\widehat{q}^{\alpha}_{t+k|t}$). Moreover, a naive benchmark model where $\widehat{q}^{\alpha}_{t+k|t} = \hat{q}^{\alpha}_{t+k|t}$ was used, assuming that there is no variation in the quantile forecasts with an information update. The MSE improvement over the benchmark model was computed for each model. As mentioned previously, two *second-level forecasts* are considered: *second-level forecast* for forecast updated with NWP generated at 12h00, and *second-level forecast* for two hours-ahead forecast. We used data from nine WPP for our experiments. Data from April 2019 to March 2020 was used for model training, and from this portion, about 25% random samples were selected for validation. The testing period spanned from April to September 2020. Note that the input and target quantiles of the training period have been produced through k-fold cross-validation.

Table 3.6 presents the mean (overall WPP, quantiles 5% to 95% with 5% increments) improvement over the naive model for different model variations from Table 3.5. The best performance

Table 3.6: *Mean percentage improvement (per WPP) over the naive model for different model variations.*

Second-level forecast	Model Variations					
	Var1	**Var2**	**Var3**	**Var4**	**Var5**	**Var6**
NWP generated at 12h00	**9.17**	7.27	9.36	9.11	9.10	8.39
2-hours ahead forecast	**14.06**	13.69	12.42	11.93	14.08	11.46

Table 3.7: *Improvement over naive model per WPP*

Second-level forecast	WPP								
Model	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
12h forecast									
GBT	17.0	23.0	18.5	25.8	17.5	18.5	15.4	15.1	14.3
ED-ANN	13.7	25.2	18.6	26.1	18.9	16.5	15.8	13.0	11.9
2-hours ahead forecast									
GBT	19.4	32.7	19.9	16.3	18.7	22.6	17.9	13.6	23.2
ED-ANN	24.9	31.2	22.8	21.5	20.9	26.7	21.5	16.2	23.1

is from model variation Var1. This combination of variables is used to compare GBT and ED-ANN.

The results, averaged by WPP and including extreme quantiles (lower than 5% and higher than 95%), are shown in Table 3.7. Starting with the *second-level forecast* for the 12h00 forecast launch time in day D, we observe that ED-ANN shows results comparable with the GBT, although the improvement is higher for the former in five WPP. For the *second-level forecast* for two hours-ahead forecast, the ED-ANN outperforms the GBT model for seven WPP. Both improve on the naive model, showing the potential for exploring the *second-level forecasting* concept, in particular, the ED-ANN architecture.

3.4.4 Illustrative example

This subsection presents an example for line 155 on August 13, at 17h00, to demonstrate the potential of the *time-to-decide* approach in reducing false congestion alarms since the observed (real) loading of this line was 37.6%. The forecast launched at 0h00 for day D+1 (i.e., $t+42|t$) estimated a probability of congestion equal to 39.6%. In contrast, the two *second-level forecasts* (for D+1 with NWP at 12h00, and for $t+2$) show a decrease in the probability, 23.6%, and 16.6% correspondingly, which might indicate a false congestion alarm.

Fig. 3.16 illustrates the contribution of the nine WPP (C_j in Eq. 3.7), depicted through gradient

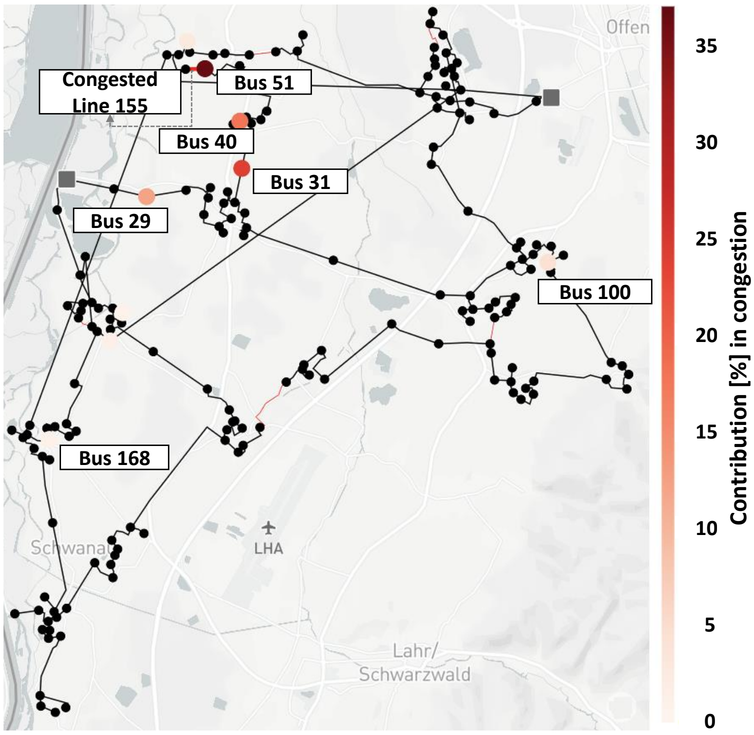


Figure 3.16: *Contribution (C) of RES power plants forecast uncertainty to the forecasted probability of congestion in one line*

color, for the false congestion alarm in line 155. In this case, the highest contribution (i.e., 36.7%) is from the WPP at bus 51 (where line 155 is connected), which has a higher forecast absolute error for that time interval (47.26% of rated power), followed by WPP connected to bus 31 with a contribution of 23.87% and an absolute forecast error of 52.53%.

Fig. 3.17 depicts the line loading's probability density function (PDF) for the three different forecasts obtained with the kernel density estimation over the scenarios samples. The point forecast (100.3% of the line loading) and observed value (37.6% of the loading) are also presented.

The risk vs cost curves for each forecast are depicted in Fig. 3.18. Using the risk threshold-stakes curve in Table 3.3, and since, in this case, the stakes are equal to 6, the risk threshold is equal to 6% (in DM A) and 15% (in DM B). For DM A, book flexibility only two hours in advance ($t+2|t$) is preferred. Note that the stakes for the two *second-level forecasts* are equal to 4, which leads to a risk threshold of 6% for DM A. Therefore, the operator should wait for the upcoming forecasts and not book any flexibility. Then, 12 hours later, the operator receives an updated forecast for the same hour (now lead-time $t+30|t$) and a second *second-level forecast* for $t+30|t+28$. Due to the lower total flexibility cost in the *second-level forecast* for $t+30|t+28$, the operator will postpone the flexibility booking for the last forecast in $t+2|t$. The very short-term forecast launched at $t+2|t$ estimates a near zero probability of

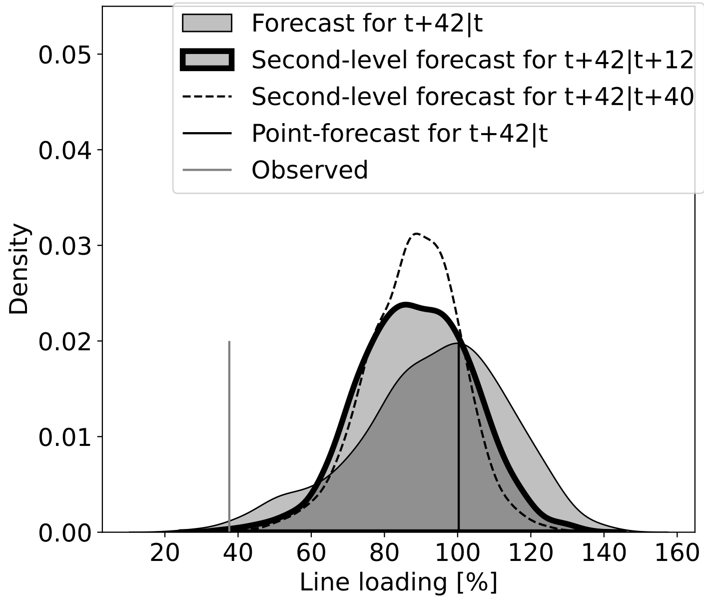


Figure 3.17: *PDF of the line loading obtained with day-ahead forecast and second-level forecasts*

congestion, as depicted in Fig. 3.19, meaning that no flexibility is necessary for that specific hour. The same procedure could be followed for the risk level of DM B, and it would also reach this point where flexibility is not used. Therefore, the proposed *time-to-decide* methodology can result in a reduction of the flexibility cost and false alarms through the use of *second-level forecasts*.

3.4.5 Practical issues

Flexibility scarcity

Flexibility scarcity events might occur (e.g., due to low flexibility market liquidity) and can be detected in advance by the proposed method. For instance, on April 14, at 11h00, the observed loading at line 49 was 124.47%. This congestion was detected with the forecast for $t+35|t$ with a probability of 89.9%. The risk-cost curve for the $t+35|t$ forecast is depicted in Fig. 3.20. The set of actions for the lowest reachable risk (i.e., 79.6%) are active power curtailment of 30% in WPP and CHP unit of bus 29 and WPP of bus 51 together with 20% in buses 43 and 117. However, these actions are insufficient for solving this technical issue or reaching an acceptable risk threshold for a decision-maker. In this case, the congestion problem will occur, leading to real-time control actions (e.g., RES curtailment) and a monetary loss of 329.76€ .

Conflicting action between TSO and DSO

In this illustrative example, line 155 on July 5 at 04h00, a solution at the distribution level creates transmission-level problems. The TSO has two possibilities to integrate this information:

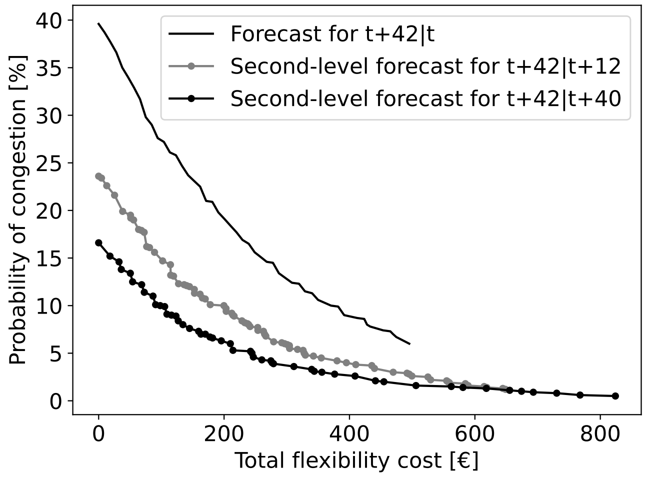


Figure 3.18: *Risk-cost curves for the forecast and two second-level forecasts launched at 0h00*

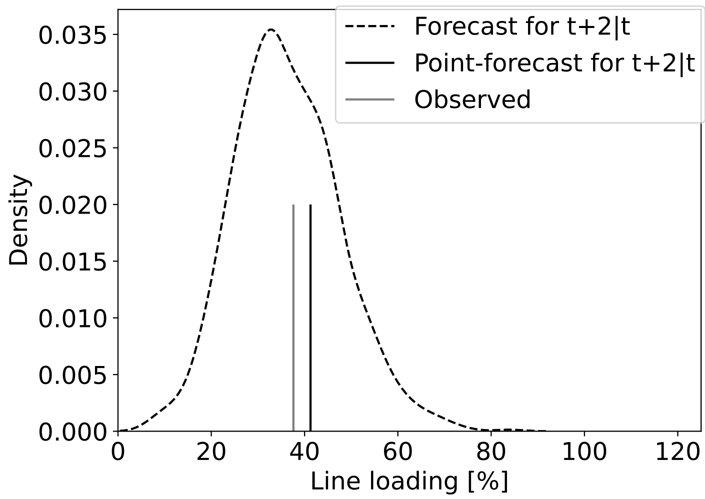


Figure 3.19: *PDF of line loading of forecast launched at $t+2|t$*

a) incorporating risk metrics into the flexibility ranking that quantify the potential impact on the TSO network (as explained in section 3.3.1); b) imposing a limit on the $FV^{s}_{j\rightarrow i}$ value to prevent technical problems in the transmission network (as explained in section 3.2.3).

To illustrate the first possibility, Table 3.8 presents the modification in the flexibility ranking by incorporating the probability of congestion in the transmission network (referred to as "TSO risk" in the Table for simplicity) and the set of top flexible resources. Using this approach, the information from the TSO would be used to influence the flexibility ranking and select the actions with a lower probability of creating technical problems in the upstream network. This implies that flexibility options with higher costs may be placed in the top-ranked positions because of their capability to avoid technical issues in the transmission network. For the second possibility, Fig. 3.21 depicts the risk-cost curves as an example for the $t+28|t$ forecast,

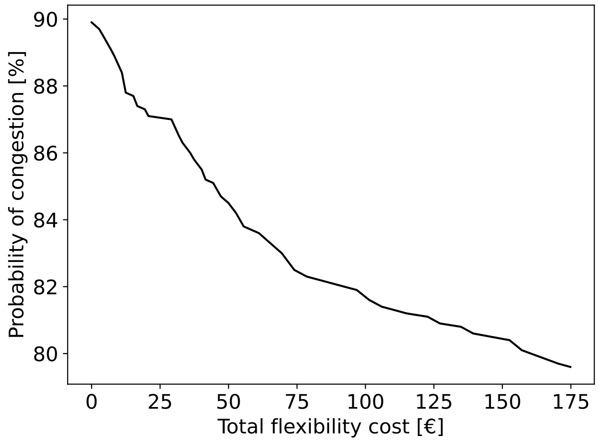


Figure 3.20: *Risk-cost curve for a case with flexibility scarcity*

with and without TSO limitation to the $FV^{s}_{j\rightarrow i}$. The curve with TSO limitations shows that to maintain the same risk threshold (e.g., 5%), a higher cost for flexibility is necessary. As a result, constraints on flexibility due to conflicting TSO-DSO actions could result in expensive solutions.

Table 3.8: *Flexibility ranking without and with TSO limitation*

Without TSO limitation	With TSO limitation	
Flexibility actions	Flexibility actions	TSO risk [%]
APC WPP @ bus 31	APC WPP @ bus 51	8.33
APC WPP @ bus 65	APC WPP @ bus 65	16
APC WPP @ bus 51	APC WPP @ bus 31	77.42
APC WPP @ bus 29	APC CHP unit @ bus 30	71.58
APC CHP unit @ bus 30	APC WPP @ bus 29	98.44

[1] APC stands for active power curtailment.

Changes in flexibility level

The forecasted congestion of line 155 on July 5 at 05h00 was selected for sensitivity analysis of the flexibility level (i.e., maximum flexibility per flexibility resource). Three flexibility levels were considered: 10%, 30% (which will be used in the remaining work), and 70% of the RES point forecast or load. The impact of this change on the risk-cost curve for $t+2|t$ is depicted in Fig. 3.22. It shows that reducing the flexibility level can prevent the curve from reaching a lower risk level. Moreover, for the same risk level, the total flexibility cost for a 70% flexibility level is lower since a larger volume of flexibility from top-ranked solutions is available. The result of the T2D strategy for different flexibility levels is presented in Table 3.9

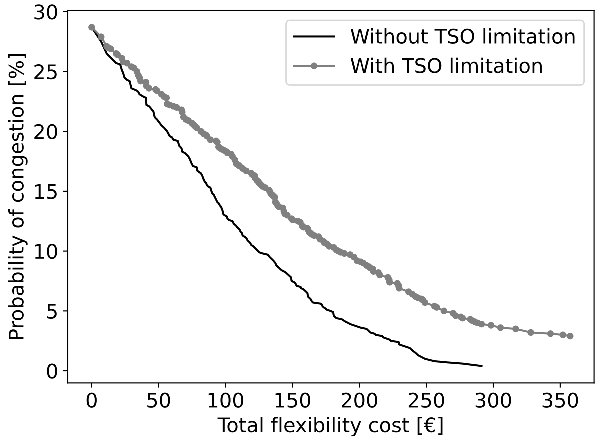


Figure 3.21: *Risk-cost curves for the cases with and without TSO limitation in the flexibility value (FV)*

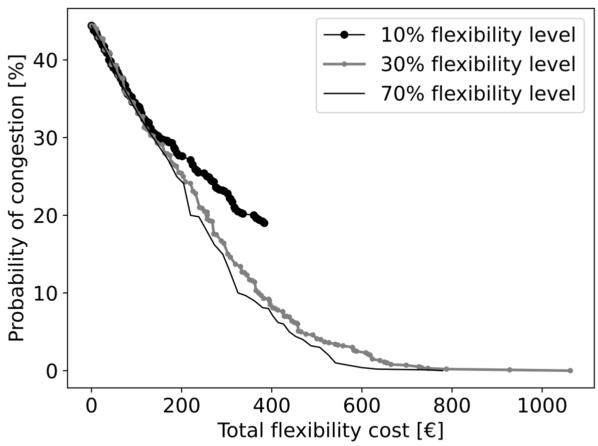


Figure 3.22: *Risk-cost curves for the $t+2|t$ forecast considering different flexibility levels*

for DM A. In the 10% flexibility level, the risk-cost curve cannot reach the risk threshold defined by the DM, and real-time RES curtailment actions will be necessary to solve the congestion problem. Increasing the flexibility level to 30% and 70%, the congestion will be solved with the "reserved" flexibility. However, the cost of booking flexibility with a 70% flexibility level is lower. In both 30% and 70%, the solutions obtained by the T2D strategy outperform the DN strategy.

3.4.6 Overall evaluation

An overall evaluation was conducted over six months from April to September 2020). All the simulations were run on a cloud-based virtual machine with AMD EPYC CPU with 8-core 2.94 GHz, 32 GB RAM, and Microsoft Windows 10 Pro. Table 3.10 presents the computational time of each step in the proposed method.

To evaluate the congestion detection performance of each strategy, the F1-score and F3-score

Table 3.9: *Sensitivity analysis of the T2D strategy for different flexibility levels*

Flexibility level	T2D approach			DN approach
	Flexibility cost [€]	Loss [€]	Total cost [€]	Total cost [€]
10%	78.9	945.5	1024.4	1024.4
30%	457.4	0	457.4	525.2
70%	426.1	0	426.1	520.3

Table 3.10: *Computational time of different steps in the proposed methodology*

Step	Computational time (sec.)
Flexibility value calculation	64s (0.25s/scenario)
Flexibility ranking	0.016s
Risk-cost curve	47.59s
Decision-aid	0.14s

are reported in Table 3.11. The *T2D* approach achieves the highest score, with DM A (using a lower risk threshold) showing higher performance in F3-score, while DM B outperforms in F1-score. The deterministic strategies show the lowest score due to a lower rate of TP detection (and a high rate of FN cases). The probabilistic approaches overcome this limitation, and the rate of FP is reduced in the *T2D* strategy due to the information provided by the *second-level forecasts*. The benchmark model from [91] is outperformed by T2D in DM A, B, D, and E.

The results of the cost-loss matrix per DM strategies are presented in Fig. 3.23–3.24. The probabilistic strategies show a lower rate-of-occurrence and total cost in Cell C "action not taken, event occurred". This shows their effectiveness in addressing overlooked congestion cases, in contrast to the deterministic strategies. The higher rate-of-occurrence and total cost in matrix row "action taken" (Cells A+B) are due to more congestion problems solved, mixed with some false congestion alarms (a "side-effect" of risk-based approaches).

Compared to the *DN* strategy, the *T2D* strategy shows a lower rate-of-occurrence and total cost for matrix Cell C "action not taken, event occurred" and Cell A "action taken, event occurred" since, with the use of *second-level forecasting*, the human operator may decide book flexibility now or postpone the decision to the next forecast. In this case, the strategy also benefits from the possibility of detecting overlooked congestion (i.e., reducing the FN cases). Moreover, the *T2D* strategy can significantly reduce the cost caused by false alarms, which is reflected in a higher rate-of-occurrence in Cell D "action not taken, event did not occur". The possibility of

Table 3.11: *Accuracy metrics for evaluating congestion detection performance and cost-loss matrix performance indicator-γ*

Strategy	Precision	Recall	F1-Score	F3-Score	γ [M€]
D1	0.53	0.08	0.13	0.08	26.03
D2	0.53	0.03	0.06	0.03	27.03
Model from [91]	0.45	0.74	0.56	0.69	14.74
DN: DM A	0.45	0.71	0.55	0.67	5.05
T2D: DM A	0.56	**0.79**	0.66	**0.76**	3.98
DN: DM B	0.49	0.63	0.55	0.61	6.21
T2D: DM B	0.65	0.69	**0.67**	0.69	4.93
DN: DM C	0.50	0.60	0.54	0.59	6.63
T2D: DM C	**0.66**	0.66	0.66	0.66	5.27
DN: DM D	0.43	0.69	0.53	0.65	4.92
T2D: DM D	0.50	0.77	0.61	0.73	4.29
DN: DM E	0.44	0.69	0.53	0.65	4.76
T2D: DM E	0.48	0.78	0.59	0.73	**3.72**

selecting the flexibility booking moment by the *T2D* strategy can reduce the cost significantly, as explained in section 3.4.4. The comparison between the three different DM, shows distinct results and that DM A with lower risk thresholds (in comparison to DM B and C) can further reduce the loss of preventive management while paying more for the flexibility than the other two DM. The trade-off strategy (DM D and E) provides the DM with the ability to manage the balance between risk and cost effectively. For example, DM D was willing to pay 30€ and 70€ to achieve a 1% reduction in risk. This level of control over risk and cost is not possible with the risk-threshold strategy (DM A-C), where the DM can only determine the maximum acceptable level of risk without considering associated costs. Consequently, this limitation may result in the implementation of costly solutions to meet the risk threshold. In this case, DM D contracts more flexibility to solve the technical problems (higher rate-of-occurrence in Cell B "action taken, event did not occur") compared to DM B and C, which have lower risk thresholds than DM A. DM E has higher trade-off value then DM D for each stakes level, which means that it is willing to pay more to decrease the risk. For instance, for DM E, this leads to a higher total cost in Cell B "action taken, event did not occur", but also to a lower rate-of-occurrence than DM D.

The performance indicator γ was computed for each strategy, and results are presented in Table 3.11. These results show that the *D1* strategy that "reserves" flexibility in the moment with lower flexibility price, but higher uncertainty (which leads to higher "reserved" flexibil-

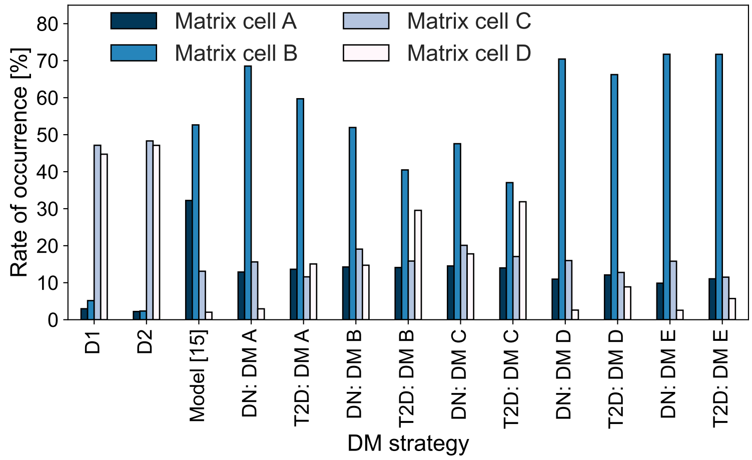


Figure 3.23: *Rate-of-occurrences for each cell of the cost-loss matrix.*

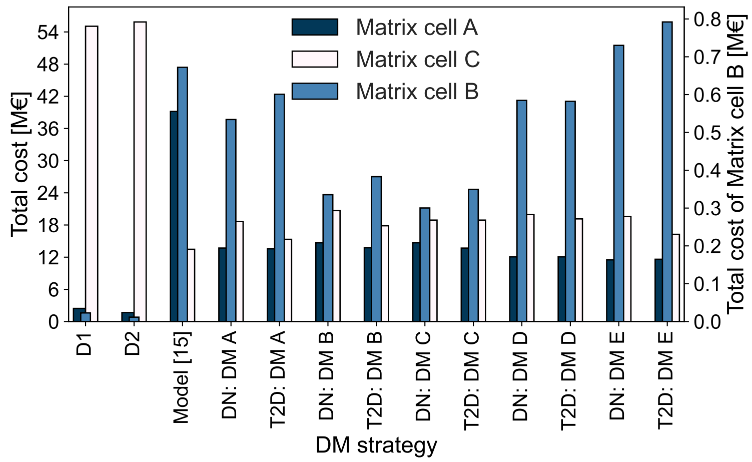


Figure 3.24: *Total cost for each cell of the cost-loss matrix.*

ity), has 3.69% better performance than *D2* strategy. The *DN* strategy has 80.61%, 76.12%, 74.53%, 81.09% and 81.73 improvement over the *D1* strategy. The performance of the *T2D* strategy was evaluated concerning the *DN* strategy in each DM strategy, where an improvement of 21.08%, 20.72%, 20.45%, 12.78%, and 21.74% was obtained by DM A, DM B, DM C, DM D, and DM E, respectively. These results show the advantage of the proposed *T2D* strategy. Moreover, the benchmark model from [91] was outperformed by the T2D strategy in all DM A-E. Lastly, it is important to acknowledge that distinct risk profiles of DM can result in different performances. Consequently, an area of future research lies in integrating DM preferences into the cost-loss matrix and corresponding performance metric. This would enable direct comparisons between decision paradigms (e.g., risk-threshold vs trade-off) or DM.

The performance of the proposed T2D strategy for different flexibility price increment percentages as a function of the notification time. The previous results assumed increments of 40% and 90%. The evaluation was conducted through a sensitivity analysis in Table 3.12.

In all cases, the T2D strategy consistently outperformed the DN strategy, showing an average improvement of 20.75%. Furthermore, the minor variations observed in the performance indicator γ indicate the robustness of the T2D strategy to changes in flexibility prices with the notification time.

Table 3.12: *Sensitivity analysis for different flexibility price increments (in %) as a function of the notification time*

Strategy	booking 12 hours later	2 hours before delivery	γ [M€]
DN	-	-	5.048
T2D	40%	90%	3.984
T2D	40%	110%	4.000
T2D	40%	70%	3.979
T2D	20%	70%	4.015
T2D	70%	110%	4.023

3.5 Final Remarks

This chapter describes a predictive methodology for flexibility procurement to manage technical constraints under forecast uncertainty and consider the different flexible resources. The main contribution is a complete methodology to guide the human operator along a) the flexibility options available in each hour, ranking them according to their effectiveness in an uncertain context, and b) multiple forecast updates. This methodology was tested in a public MV network, and the results showed that: a) a methodology based on uncertainty forecasts can lead to cost savings when managing grid technical constraints, and b) choosing the best moment to reserve flexibility also leads to cost savings in comparison to a strategy that reserves flexibility when it is cheaper. It is essential to note that short-term market-based flexibility procurement still has some challenges at the distribution grid level, namely: a) dealing with the radial topology in MV and LV grids, and geographical restrictions, b) lack of a mature regulatory framework for new DSO roles and coordination across different markets, and c) upward/downward flexibility price asymmetry. Moreover, the use of DSO's resources (e.g., OLTC, network topology reconfiguration) should be coordinated with local flexibility markets operation, notably in areas with limited (radial) geographical coverage, fewer FSPs, and lower liquidity. Uncoordinated use of such resources may deter participation, or lead to bid and/or settlement distortions.

Chapter 4

Evolving Power System Operating Rules for Real-Time Congestion Management

Real-time congestion management in transmission networks, particularly under contingency scenarios, is becoming increasingly complex due to a convergence of emerging challenges such as the variability and uncertainty of renewable energy sources (RES), the rising frequency and intensity of extreme weather events, and the increasing risk of cyberattacks on critical infrastructures. In this demanding context, control room operators are required to rapidly assess system conditions and define effective remedial actions to maintain grid stability, prevent overloads, and mitigate the risk of cascading failures. These remedial actions may involve generation redispatch, RES curtailment, demand response, and topological reconfiguration of the network. The large number of possible action combinations, coupled with the need for real-time decision-making, makes it extremely challenging to identify optimal solutions, which often requires the use of approximate or heuristic methods to support timely and reliable operations [4].

This work builds upon the longstanding concept of ES in power systems, traditionally rule-based approaches that integrate expert domain knowledge with physics-based models, and introduces a data-driven methodology to enhance these systems. Specifically, focusing on the real-time congestion management problem, it adopts the formalism of Markov decision processes as established in the L2RPN competition [6], created by RTE (the French transmission system operator), and proposes the augmentation of ES by using RL, enabling the ES to learn and adapt optimal policies from experience and interaction with the environment.

In this work, a novel variant of GNP is proposed that diverges significantly from conventional formulations such as [92], particularly in its structural design and dynamic node functionality.

Traditional GNP frameworks typically use fixed function nodes and update them in a sequential manner, resulting in rigid decision paths that struggle to adapt in high-dimensional, time-varying environments. In contrast, the proposed GNP architecture features adaptive nodes capable of modifying their behavior in response to evolving network states and accumulated experience. This flexibility allows the network to represent more sophisticated, state-dependent policies well-suited to the uncertain and non-linear nature of power systems, such as real-time fluctuations from RES or non-linearities inherent in AC power flow. Crucially, this design enables context-aware decision-making. For instance, an overloaded line may not always trigger the same switching action; instead, the node selects an appropriate control based on additional context, such as nearby contingencies, mimicking expert reasoning under diverse operational scenarios. As training progresses (i.e., learning from data and experience), these evolving heuristics capture expert-like judgment, resulting in a responsive and interpretable control mechanism that adapts in real-time to changing grid conditions. Furthermore, the possibility of seeding at initialization the GNP population with expert-derived decision graphs (e.g., from a pre-existing ES) can accelerate convergence in early learning phases by reducing the need for extensive trial-and-error.

Interpretability is a central contribution of the proposed method, achieved through the integration of graph-based structures. In particular: a) unlike conventional Deep-RL approaches that operate as "black-boxes", the GNP framework structures decision-making within an explicit graph-based representation. Each node in the graph corresponds to a programmed decision heuristic, allowing the entire policy to be visualized and traced. This structure makes it possible to understand the rationale behind specific topology actions at any point during learning. In contrast, end-to-end Deep-RL methods, such as those in [72, 74, 118], often produce uninterpretable decisions encoded in high-dimensional neural parameters, limiting reasoning and reducing trust from grid operators; b) a multistage decision tree (DT) was employed to extract human-readable rules from the trajectory of actions created by top-performing decision graphs in GNP, which were collected into a pool representing high-quality policy behavior. Separate DTs are then trained to capture key aspects of the control logic. This hierarchical rule extraction enables modular interpretation, where each stage informs and constrains the next. The proposed method is abstracted graphically in Figure 4.1.

4.1 Evolving system operation rules

The proposed methodology targets real-time line congestion management in power systems. Its core concept involves evolving heuristic control policies, originally derived from operator expertise or pre-existing ESs, encoded as a decision graph (see Section 4.1.1).

The congestion management task can be modeled as a Markov decision process [119], defined

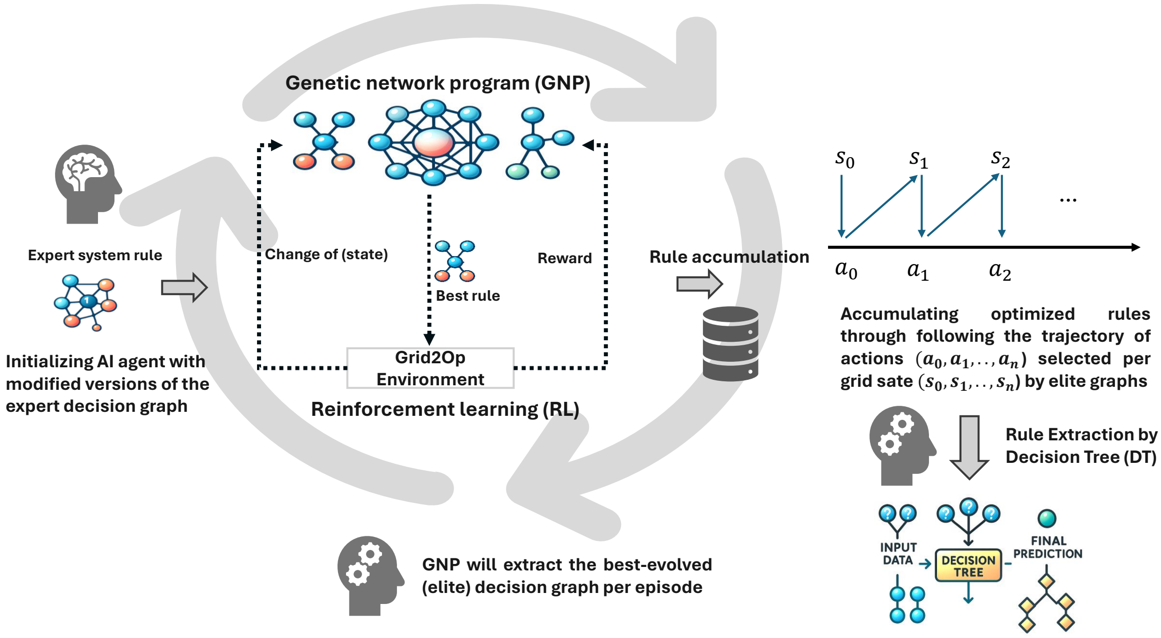


Figure 4.1: *Graphical abstract of evolving system operation rules*

as a tuple $\mathcal{M} = (\mathcal{S}, \mathcal{A}, P, R, \gamma)$, where:

- $\mathcal{S}$ is the state space. Each state $s_t \in \mathcal{S}$ represents the full state of the grid at time t, including power line flows, bus voltages, switch configurations, and operational constraints of the AC power flow.

- $\mathcal{A}$ is the action space. Each action $a_t \in \mathcal{A}$ corresponds to a control action, such as line switching (disconnection/reconnection), bus reconfiguration, bus recovery, or active power flexibility (e.g., redispatch, RES curtailment, storage system management).

- $P(s_{t+1} \mid s_t, a_t)$ is the state transition function, encoding the system dynamics and physical laws that govern how the grid evolves in response to actions.

- $R(s_t, a_t)$ is the reward function, used to evaluate the safety and effectiveness of each action.

- $\gamma \in [0, 1]$ is the discount factor.

At each timestep t, the agent observes the system state s_t and selects an action a_t according to a pre-defined policy $\pi(a_t \mid s_t)$. The environment then transits to a new state s_{t+1} based on the dynamics:

$$s_{t+1} \sim P(s_{t+1} \mid s_t, a_t), \quad a_t \sim \pi(a_t \mid s_t) \tag{4.1}$$

In the original formulation, the RL framework relied on a sparse binary reward signal to encourage system survival. Let $r(s_t, a_t)$ denote the reward received at timestep t, where s_t is the system state and a_t the action taken. the agent receives a reward of 1 for each successful step t taken (i.e., system is able to fully meet the electricity demand with the available generation resources), and a reward of 0 upon reaching a terminal or gameover state (i.e., when transmission lines exceed their physical capacity and lead to cascading failure of other lines that ends with system collapse); this is similar to Grid2Op [69] Episode duration reward:

$$r(s_t, a_t) = \begin{cases} 1, & \text{if the agent successfully passes a step} \\ 0, & \text{if the agent reaches a gameover state} \end{cases} \tag{4.2}$$

The RL objective is to find the optimal policy π^* that maximizes the expected cumulative reward over an episode of length T:

$$\pi^* = \arg\max_{\pi} \mathbb{E}_{\tau\sim\pi}\left[\sum_{t=0}^{T} r(s_t, a_t)\right], \tag{4.3}$$

where $\tau = (s_0, a_0, s_1, \ldots, s_T)$ denotes a trajectory under policy π, and the expectation $\mathbb{E}_{\tau\sim\pi}$ is over trajectories sampled under the environment dynamics. As also noted in [6, 118], the winning Deep-RL agent of the L2RPN WCCI 2020 Challenge adopted the episode-duration reward formulation, reinforcing its validity as a baseline reward signal for grid control learning tasks. While this formulation encourages robust strategies for prolonging operation, it does not directly account for operational costs, which were only evaluated post hoc. To address this limitation, we propose a *cost-aware reward* that incorporates a formal operational cost model inspired by the reward proposed by Grid2Op contributors and closely aligned with the L2RPN scoring metric [120]. The per-timestep operational cost is:

$$C_{\text{op}}(t) = C_{\text{dispatch}}(t) + C_{\text{flex}}(t) + C_{\text{losses}}(t), \tag{4.4}$$

with

$$\begin{aligned} C_{\text{dispatch}}(t) &:= \sum_{g\in\mathcal{G}} c_g(t)\, p_g(t), \\ C_{\text{flex}}(t) &:= C_{\text{storage}}(t) + C_{\text{curtail}}(t) + C_{\text{other\_flex}}(t), \\ C_{\text{losses}}(t) &:= \text{cost-per-MWh}_{\text{loss}} \times \text{losses}(t), \end{aligned}$$

and $\mathcal{G}$ is the set of generators. The blackout (unserved load) cost is:

$$C_{\text{blackout}}(t) = \text{USL}(t) \times \pi_m(t) \times \beta, \tag{4.5}$$

with USL(t) the not-supplied load, $\pi_m(t)$ the marginal price, and $\beta > 1$ a penalty factor. The total episode cost is:

$$\mathcal{C}_{\text{episode}} = \sum_{t=1}^{T} \big(C_{\text{op}}(t) + C_{\text{blackout}}(t)\big). \tag{4.6}$$

We formally define an upper bound for the episode as the worst-case cost:

$$C_{\max} := \max_{\text{worst scenario}} \sum_{t=1}^{T} C_{\text{blackout}}(t), \tag{4.7}$$

where the worst scenario assumes that the agent fails to supply all demand at the highest marginal price in each timestep. The normalized cost-aware reward is then:

$$R_{\text{cost}} = \frac{C_{\max} - \mathscr{C}_{\text{episode}}}{C_{\max}} \times 100, \tag{4.8}$$

so that lower costs produce higher rewards on a 0–100 scale. Survival is naturally embedded, as blackouts contribute heavily to $\mathscr{C}_{\text{episode}}$.

The survival-based reward emphasizes robustness alone, whereas the normalized cost-aware reward jointly captures system reliability and operational efficiency. The blackout penalties ensure that maintaining supply security remains a primary objective, aligning the reward with real-world operational goals.

In this work, an undiscounted reward formulation ($\gamma = 1$) was adopted to evaluate the cumulative performance of the control policy over a finite time horizon. This choice is motivated by the need to treat all future rewards with equal importance, as the objective is to optimize the agent's performance without biasing toward immediate gains. Since the task involves an episodic evaluation with a clear termination criterion, discounting is unnecessary and can introduce an unintended emphasis on short-term outcomes as explained in [121]. Moreover, the undiscounted reward simplifies the analysis and interpretation of the learned policies, ensuring that the optimization process directly targets the maximization of overall long-term effectiveness.

The learning process, including the evolution of expert knowledge, is detailed in Subsection 4.1.2. Moreover, a rule extraction method that predicts the best action with respect to evolved knowledge, considering contextual information, is described in Subsection 4.1.3.

4.1.1 Expert system knowledge representation

The baseline knowledge for this work and also for the GNP rule evolution in the next section is the ES developed by RTE [93], where the aim was to emulate the decision-making process of human operators. This ES was enhanced by considering improvements in the calculation method for determining the influence graph, using more flexibility options (e.g., line switching and redispatch, in addition to the bus splitting option), new ranking criteria, the possibility of decision revision, and the possibility of proposing multiple actions rather than a single action.

The ES is formulated as a network graph for knowledge representation, as illustrated in Figure 4.2, where ES does not include dashed connections; these connections represent other

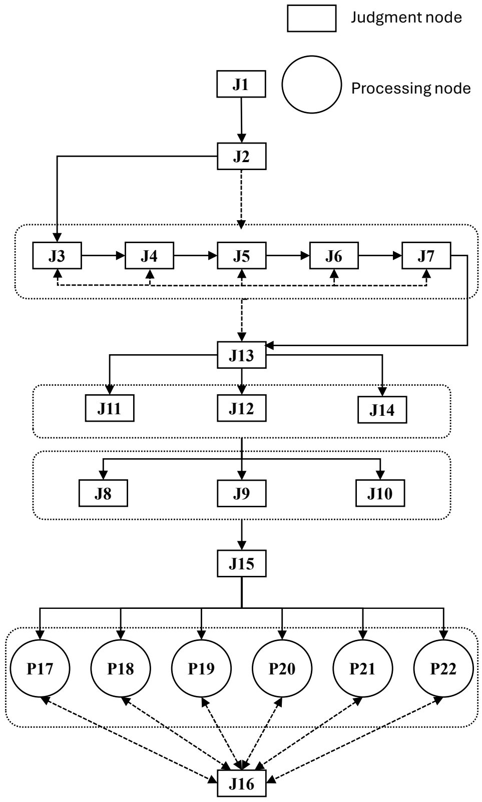


Figure 4.2: *ES decision-making graph.*

reformation of ES used for initializing the GNP algorithm in Subsection 4.1.2. Each judgment node (J) within this framework is implemented as an individual computer program, responsible for a specific decision-making task, as explained in Table 4.1. Moreover, the detailed algorithmic functionality of nodes is presented in Appendix B.

The sequence of these judgment programs guarantees convergence to one of the processing nodes: $P17$, $P18$, and $P19$ execute bus-splitting for hub, loop, and downstream buses, respectively; $P20$ handles line switching; $P21$ activates power flexibility; and $P22$ implements a recovery strategy to the reference topology. This ensures that the system reaches a unique, executable action path tailored to the grid's condition. For instance, in a specific case, two congestion events were detected on lines 20 and 12.

- For **line 20**: The flexibility nodes identified buses 76 and 81 as hub buses.
- For **line 12**: The responsible flexibility group identified buses 67 and 68 as hub buses. In addition, the bus set {67, 80, 79, 78} was recognized to form a looped path suitable for re-routing.

Since the congested lines were located in different zones, the judgment node J2, which handles inter-zone coordination, was not activated. In a specific topological configuration, bus 68 alone

Table 4.1: *Functions of judgment nodes (J) in the ES decision graph for congestion management.*

Node	Function	Details
J1	Identify and rank congested lines	Based on criticality
J2	Skip further analysis	Triggered if all issues are within a single critical zone
J3	Explore flexibility – hub bus splitting	Key locations enabling power flow rerouting through parallel uncongested paths
J4	Explore flexibility – loop bus splitting	Buses that build a local mesh
J5	Explore flexibility – downstream bus splitting	Buses that could be supplied from a different path
J6	Explore flexibility – line switching	Based on topological search [122]
J7	Explore flexibility – active power flexibility (redispatch, RES curtailment, storage management)	Based on sensitivity indices [122]
J8	Evaluate stop condition	Solution quality considering the first threshold
J9	Evaluate stop condition	Solution quality considering the second threshold
J10	Evaluate stop condition	Number of analyzed issues exceeded a threshold
J11	Recovery	Triggered when no valid action proposed
J12	Recovery	Triggered when solution quality falls below a threshold
J13	Multi-criteria hierarchical ranking	Uses tuple $\sigma(a) = (\sigma_{cp}(a), \sigma_t(a), \sigma_{op}(a))$: $\sigma_{cp}(a)$: Control priority (e.g., switching vs redispatch) $\sigma_t(a)$: Topological effectiveness (e.g, issue disappeared/alleviated/worsened/created) $\sigma_{op}(a)$: Operational score (e.g., sum of squared marginal power flows)
J14	Ranking revision under thresholds	Revising the criteria order in J13 per issue
J15	Ranking revision under thresholds	Revising the criteria order in J13 after full analysis
J16	Add more actions	Uses ranked list if one action is insufficient

was sufficient to solve both congestion events. Consequently, the rest of the judgment nodes that contain additional conditional constraints were not activated.

4.1.2 GNP rule evolution

To overcome the limitations of rigid decision-making in traditional GNP, the proposed method integrates an RL-enhanced GNP framework that supports dynamic and context-aware behavior. In this approach, each node is capable of altering its output depending on current grid conditions, such as line overloads, generation/load levels, or switching states. For example, a judgment node that initially selects the reconfiguration action at 'Bus A in zone 1' under moderate congestion may instead recommend reconfiguration at 'Bus C in zone 3' if RES fluctuations create a localized congestion between the two zones. The node's behavior is therefore not statically defined, but instead learns a mapping from state features to decision criteria (e.g., change in priority of flexible units or threshold of activating a flexible unit), enabling adaptive control that aligns with requirements for real-time management. Furthermore, the execution sequence of the decision graph is not fixed, in contrast to traditional GNP, which follows a hardcoded node traversal. Instead, the proposed method allows the policy graph to activate different substructures conditionally, based on current state inputs. These conditions include the level of congestion, fault locations, substation configurations, and the recent history of control actions. The adaptive traversal mechanism in the proposed GNP supports conditional connections between nodes, especially under time-varying network constraints.

This structural advancement is depicted in Figure 4.3, which compares the traditional and proposed GNP formulations. On the left, the traditional GNP follows a linear structure in which each node is associated with a fixed function, and execution proceeds through a predefined sequential path, formalized in Equation 4.9.

$$n_{t+1} = \text{Next}(n_t), \quad a_t = f_{n_t}(s_t) \tag{4.9}$$

where n_t denotes the node visited at time t, Next(n_t) is the deterministic next node pointer, s_t is the observed grid state at time t, f_{n_t} is the decision rule implemented at node n_t, and a_t is the resulting action taken by the agent. Each node processes its input and passes control unconditionally to the next node, regardless of the evolving state of the grid. The system response is therefore insensitive to diverse operational contexts, limiting its effectiveness in dynamic environments.

In contrast, the proposed GNP framework (right) organizes the policy graph into an interconnected network of nodes, each with adaptive behavior. The nodes evaluate the state vector, which can include the power flow on the lines, AC power flow constraints, grid topology, and dynamically determine their output. The graph includes multiple possible transitions from

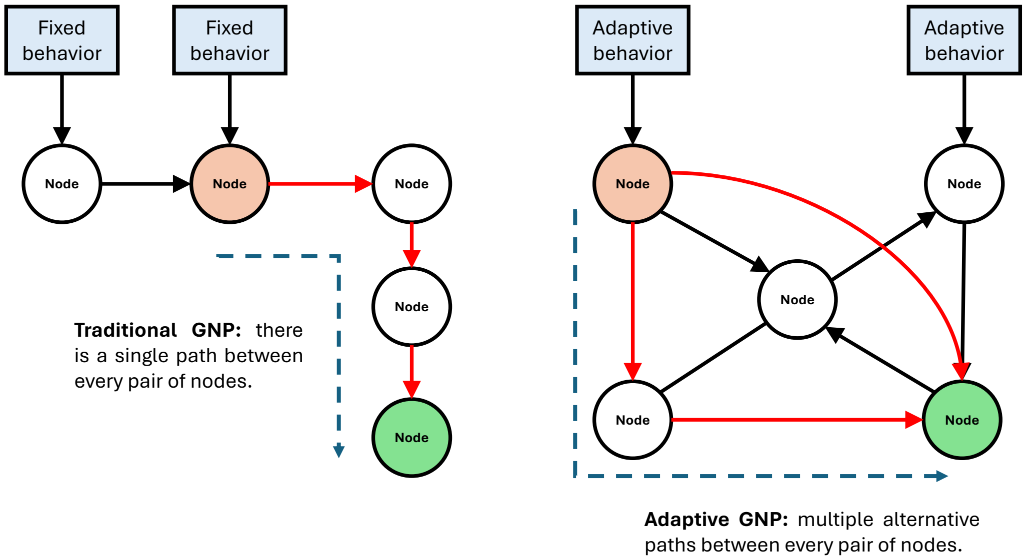


Figure 4.3: *Regular GNP vs. adaptive GNP. The traditional GNP employs sequential node activation, whereas the adaptive GNP dynamically adjusts node behavior and execution paths based on environmental states.*

each node, with links encoding condition-based execution paths learned through RL. For example, depending on whether the overload is localized or widespread, a node might route control to a sub-policy targeting either demand-side response or topological reconfiguration. The arrows between nodes represent these conditional transitions, learned from high-performance decision graphs during training. Node connectivity reflects logical dependencies and execution flexibility, rather than fixed ordering, as expressed in Equation 4.10.

$$n_{t+1} \sim \pi_{\text{trans}}(n_{t+1} \mid n_t, s_t), \quad a_t = f_{n_t}(s_t) \tag{4.10}$$

where π_{trans} is a stochastic policy over node transitions, and all other symbols are as defined previously. The system thus constructs control policies by composing rule segments that are most relevant under the current grid state, forming context-specific decision pathways that align with expert behavior while remaining responsive to uncertainty. This structural difference is summarized in Table 4.2. Thus, in contrast to classic GNP, where node transition probabilities remain static, the adaptive GNP dynamically adjusts these transitions according to RLfeedback, enabling the decision pathways to evolve in response to the agent's performance. In general, the proposed structure enables the GNP policy to behave as a state-driven control mechanism, capable of decomposing complex decisions into subgraphs, dynamically coordinating local actions, and maintaining interpretability for human operators, as will be shown in Section 4.4. The interaction between the GNP evolutionary process and the RL-based adap-

Table 4.2: *Comparison of traditional GNP vs. adaptive GNP execution*

Component	Traditional GNP	Adaptive GNP
Node execution	$a_t = f_{n_t}(s_t)$	$a_t = f_{n_t}(s_t)$
Node transition	$n_{t+1} = \text{Next}(n_t)$	$n_{t+1} \sim \pi_{\text{trans}}(n_{t+1} \mid n_t, s_t)$
Policy structure	Fixed	Conditional (learned)
Graph behavior	Deterministic	State-driven

tive learning is further illustrated in Figure 4.4, which presents a flowchart summarizing the information flow, feedback loop, and optimization stages within the proposed GNP–RL collaborative framework. Let G_i denote the i-th individual graph in the population, representing a

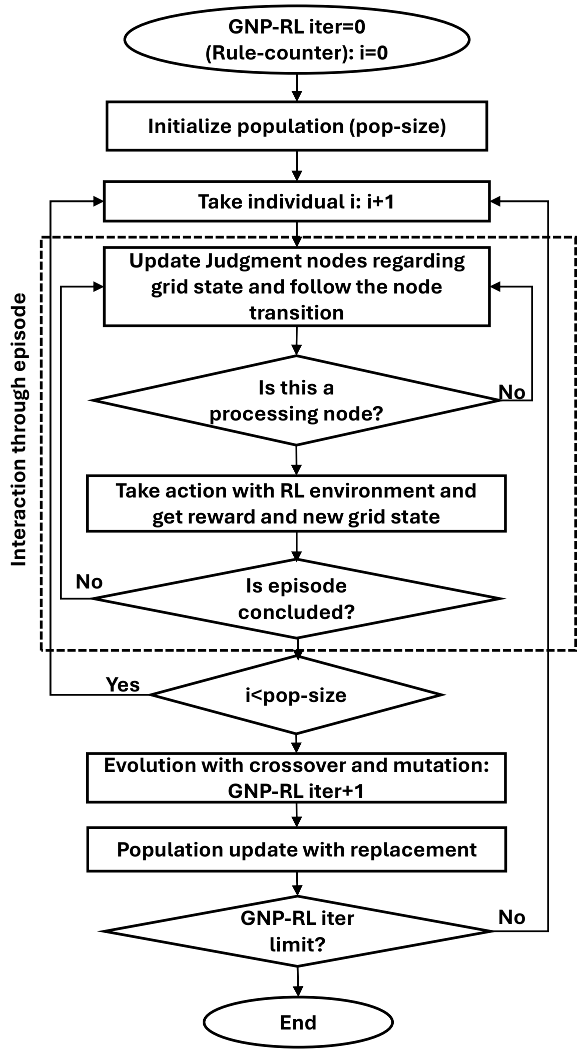


Figure 4.4: *Flowchart illustrating the fusion between GNP and RL, showing how the RL-based feedback updates guide the evolutionary process toward adaptive and optimal decision graph structures.*

policy μ_i that governs decision-making in a Markov decision process with state space $\mathscr{S}$, action space $\mathscr{A}$, and reward function $r : \mathscr{S} \times \mathscr{A} \to \mathbb{R}$. Each graph consists of *judgment nodes* and *processing nodes*; *judgment nodes* perform condition-based branching, while *processing nodes* issue actions $a_t \in \mathscr{A}$. During interaction with the environment, a graph follows its encoded logic to generate trajectories $\pi_i = (s_0, a_0, r_0, s_1, \dots)$. During execution, each node n maintains heuristic parameters θ_n, which are updated based on local feedback from the environment, typically as a function of the observed state, i.e., $\theta_n \leftarrow f_n(s_t)$. This adaptive mechanism refines node behavior across generations, supporting the learning dynamics. The process continues until a failure condition or episode termination is met. The cumulative reward along π_i defines the fitness F of G_i, Equation 4.11, which reflects the operational lifespan of the graph under dynamic grid conditions. This formulation inherently prioritizes policies that maintain safe grid operation.

$$F(G_i) = \sum_{t=0}^{T-1} r_{t+1} \tag{4.11}$$

The evolution of the graph population $\mathscr{P}$ is governed by Algorithm 1, which integrates crossover and mutation operators as defined in Algorithm 2. A two-stage crossover mechanism is employed to balance exploitation and exploration:

1. **Exploitation-driven crossover**: selects both parents P_1, P_2 from the top-performing 50% of the population to preserve high-fitness substructures.

2. **Exploration-driven crossover**: pairs elite individuals (top e_i%) with non-elite ones to introduce novel combinations and maintain diversity.

Random selection from a finite set is denoted by $x \sim \mathscr{U}(S)$, where $\mathscr{U}$ represents the uniform distribution over set S. Concatenation of gene segments is expressed using the symbol $\|$, while sub-vectors are denoted by slicing notation, e.g., $P[1{:}x]$ refers to the first x elements of parent P, and $P[x{+}1{:}L]$ refers to the remaining portion. Individual candidates P_1, P_2 are removed from their selection pools via set subtraction, indicated by $S \setminus \{P\}$. The crossover logic uses inline conditional checks such as $\texttt{random()} < cp_1$ to decide probabilistic events directly, avoiding auxiliary variables. Mutation is applied by modifying a random subsequence of each individual with probability m_p, constrained by a maximum range $\lambda_{\max} \cdot L$, where L is the gene length. The modification could be defined by ignoring/adding conditional constraints, ignoring/adding a node, decreasing/increasing a threshold, or changing a criterion. This mutation operates directly on the graph structure by modifying node logic or interconnections.

After each generation, the elite graph G^* is updated by selecting the graph with the highest fitness – Equation 4.12.

$$G^* = \arg\max_{G_i \in \mathscr{P}} F(G_i) \tag{4.12}$$

The GNP algorithm executes multiple decision-making agents (GNP graphs) in parallel, each interacting with the environment and collecting diverse experiences. While the entire population contributes to exploration through varied behaviors, learning and policy updates are guided by the elite G^* graph, identified as the best-performing individual based on the cumulative reward. This design enables efficient reuse of experience across generations and ensures that optimization is focused on high-quality policies. By decoupling exploration from exploitation, where population diversity drives exploration and the elite graph directs learning, the framework achieves both robustness and adaptability in evolving effective control strategies.

4.1.3 Decision tree rule extraction

It has been recognized that a single decision graph may not generalize optimally across diverse grid conditions due to varying fault locations, network topologies, and intertemporal dependencies. Therefore, the elite policies of the GNP algorithm are used to build a rule pool, consisting of the best-performing decision graph for each episode in the last generation. In particular, the elite graph represents an action selection trajectory that has the highest possible performance in an episodic evaluation; this is quite important, since it has optimized the impact of an action at the present grid condition (present timestep) and its post-impact in the next time steps. Therefore, elite-derived actions per timestep will be used to generate labeled datasets (i.e., consisting of grid state features and the corresponding actions), in which these datasets are then used to train a multistage DT to capture different components of control logic. During training, the agent was initially allowed to explore a broad range of control actions, including redispatch, curtailment, and storage activation. However, as evolution progressed, the population of decision graphs consistently converged toward topological reconfiguration—particularly bus-splitting and line switching actions—as the dominant and most effective strategy. This behavior aligns with findings from the L2RPN competitions framework, which emphasizes that topological operations are the most direct and cost-efficient means of alleviating congestion and maintaining grid stability, indicating that power flexibility actions were less effective, and it may also involve a CO_2 emissions increase (e.g., from RES curtailment) and higher operational cost. For instance, the Binbinchen team (Huawei) [123], which achieved second place in the L2RPN WCCI 2020 competition, developed a Deep-RL Actor–Critic agent with PPO that was trained exclusively on topological actions, restricted to 200 safe configurations following an expert-guided reduction phase. Similarly, the RTE ES [93] focused solely on bus-splitting operations. In particular, the final analysis of the L2RPN NeurIPS 2020 challenge in the robustness track [6] challenge demonstrated that several of the toughest

Algorithm 1: Adaptive GNP evolutionary algorithm

Data: Initial population of size N GNP graphs $\{G_1, G_2, \dots, G_N\}$;
Crossover probabilities $cp_1, cp_2 \in (0,1)$;
Mutation probability $m_p \in (0,1)$;
Maximum mutation range $\lambda_{\max} \in \mathbb{R}_+$;
Total reward R;
Result: Elite graph G^* with highest fitness

Initialize population: $\mathscr{P} = \{G_1, G_2, \dots, G_N\}$;
$best\_fitness \leftarrow -\infty$, $G^* \leftarrow$ None;
while *termination condition not met* **do**
  foreach *graph* $G_i \in \mathscr{P}$ **do**
    $R \leftarrow 0$;
    Initialize RL environment $\mathscr{E}$ at state s_0;
    while *episode not done* **do**
      $n_t \leftarrow n_{\text{start}}$ of G_i;
      while n_t *is a judgment node* **do**
        $n_{t+1} \leftarrow \pi_i(n_{t+1} \mid n_t, s_t)$;
      end
      if n_t *is a processing node* **then**
        $a_t \leftarrow f_{n_t}(s_t)$;
        Simulate a_t in environment: $\mathscr{E}(s_t, a_t) \rightarrow s_{t+1}$;
        $r_{t+1} \leftarrow$ reward from $\mathscr{E}$;
        $R \leftarrow R + r_{t+1}$;
        $s_t \leftarrow s_{t+1}$, $t \leftarrow t+1$;
      end
      Update heuristic parameters of node: $\theta_n \leftarrow f_n(s_{t+1})$;
    end
    $F(G_i) \leftarrow \sum_{t=0}^{T-1} r_{t+1} = R$;
    if $F(G_i) > best\_fitness$ **then**
      $best\_fitness \leftarrow F(G_i)$;
      $G^* \leftarrow$ copy of G_i
    end
  end
  $\mathscr{P} \leftarrow$ **GNP_Operators**($\mathscr{P}, cp_1, cp_2, m_p, \lambda_{\max}$);
end
return G^*

Algorithm 2: GNP operators: crossover and mutation

Data: Population size N; crossover rates $cp_1, cp_2 \in (0,1)$; mutation rate m_p; max mutation span $\lambda_{\max}$;

Subgroups: top_half, elite, non_elite;

Result: New population P_{new}

$P_{\text{new}} \leftarrow \emptyset$;
for $i \leftarrow 1$ **to** $|top\_half|$ **do**
    if $|top\_half| = 0$ **then**
        break;
    end
    Select $P_1, P_2 \sim \mathscr{U}$(top_half), without replacement;
    top_half $\leftarrow$ top_half $\setminus \{P_1, P_2\}$;
    $x \sim \mathscr{U}(\{1, \ldots, L-1\})$;
    if *random()* $< cp_1$ **then**
        $c \leftarrow P_1[1:x] \,\|\, P_2[x+1:L]$;
    else if *random()* $< cp_2$ **then**
        $c \leftarrow P_2[1:x] \,\|\, P_1[x+1:L]$;
    else
        $c \leftarrow$ *random choice*(P_1, P_2);
    end
    $P_{\text{new}} \leftarrow P_{\text{new}} \cup \{c\}$;
end
for $i \leftarrow 1$ **to** $|non\_elite|$ **do**
    if $|non\_elite| = 0$ **then**
        break;
    end
    Select $P_1 \sim \mathscr{U}$(elite);
    Select $P_2 \sim \mathscr{U}$(non_elite), without replacement;
    non_elite $\leftarrow$ non_elite $\setminus \{P_2\}$;
    $x \sim \mathscr{U}(\{1, \ldots, L-1\})$;
    if *random()* $< cp_1$ **then**
        $c \leftarrow P_1[1:x] \,\|\, P_2[x+1:L]$;
    else if *random()* $< cp_2$ **then**
        $c \leftarrow P_2[1:x] \,\|\, P_1[x+1:L]$;
    else
        $c \leftarrow$ *random choice*(P_1, P_2);
    end
    $P_{\text{new}} \leftarrow P_{\text{new}} \cup \{c\}$;
end
foreach $c \in P_{new}$ **do**
    if *random()* $< m_p$ **then**
        Modify a subsequence of c (length $\leq \lambda_{\max} \cdot L$);
    end
end
return P_{new}

episodes (Jan 28.1, Nov 34.1, Apr 42.2, Oct 21.1) remained feasible if the agent relied solely on topological actions. These findings justify the convergence of the agent's decision-space making onto topology control, aligning with both real-world operational practice and proven benchmark evidence. As a result, GNP-DT focused solely on topology reconfiguration, though it can accommodate redispatch actions without changing its formulation. The sequence of DT models is defined as follows:

(i) The first DT model (DT1) to classify the flexibility type (e.g., line reconnection vs. bus reconfiguration);

(ii) The second DT model (DT2) to estimate the required number of flexibility actions;

(iii) The third DT model (DT3) to estimate the specific line(s)/bus(es) to operate.

(iv) The fourth DT model (DT4) will be used only if the detected flexibility type involves bus reconfiguration (bus splitting) regarding DT1. The set of top-*k* most probable feasible topology reconfiguration actions for the bus identified by DT3 is extracted from DT4, based on the predicted class probabilities at the leaf nodes.

Table 4.3 presents a summary of the relevant state variables considered for different decision trees (DT1-DT4) and actions in the system. Each row indicates whether a specific variable (for example, congested line IDs, maximum overload, split bus count) is used (✓) or not (✗) by the corresponding decision logic. The variables include:

- **Cong. Line IDs**: identifiers of congested lines;
- **Max Overload**: maximum overload value among congested lines;
- **Split Bus IDs**: identifiers of splitted buses;
- **Disconn. Line IDs**: identifiers of disconnected transmission lines;
- **Cong. Line Count**: total number of congested lines;
- **Split Bus Count**: number of splitted buses;
- **Disconn. Line Count**: number of disconnected lines.

DT3 and DT4 are dependent on the action type; in this regard, Bus reconfig. refers to bus splitting. Line disconn. means disconnecting a transmission line to isolate a fault or relieve congestion. Bus recovery restores a previously reconfigured bus (split bus) to the original topology (unaltered topology), and line reconnection involves re-energizing a previously disconnected line to resume normal operation. In this work, line identifiers (e.g., *Cong. Line ID_**) are not arbitrary indices but are mapped to physically meaningful locations in the power grid. Each identifier corresponds to a specific transmission line whose electrical location is associated with a predefined zone (area) or group of feeders based on the grid layout. This structured mapping allows the model to implicitly reason about localized congestion using indexed features aligned with the physical topology of the system.

Table 4.3: *Summary of the DTs and their corresponding state variables*

Variable	**DT1**	**DT2**	**DT3**	**DT3**	**DT3**	**DT3**	**DT4**
Action type	any	any	bus reconfig.	line disconn.	bus recovery	line reconnection	bus reconfig.
Cong. Line IDs	✓	✓	✓	✓	✗	✗	✓
Max Overload	✗	✓	✓	✗	✗	✗	✗
Split Bus IDs	✓	✓	✓	✓	✓	✗	✓
Disconn. Line IDs	✓	✓	✓	✓	✓	✓	✓
Cong. Line Count	✓	✓	✓	✓	✗	✗	✓
Split Bus Count	✓	✓	✓	✓	✓	✗	✓
Disconn. Line Count	✓	✓	✓	✓	✓	✓	✓

This module is a hierarchical approach in which each stage constrains and informs the subsequent one. For instance, the flexibility type determined in the first stage conditions the search space for the next model, which predicts how many units must be activated, and in turn, this output influences the final model that selects which specific assets to operate. Formally, each stage is modeled as a conditional expectation:

$$\hat{y}_r = \mathbb{E}[y_r \mid X, \hat{y}_{r-1}, \hat{y}_{r-2}, \ldots, \hat{y}_{r-k}, \theta] + \epsilon_r \tag{4.13}$$

where $\hat{y}_r$ denotes the predicted decision component at rule stage r, conditioned on the input features X, preceding predicted components $\hat{y}_{r-1}, \hat{y}_{r-2}, \ldots, \hat{y}_{r-k}$, and model parameters θ.The parameter set θ represents the learnable weights of the conditional model, including decision thresholds, node weights, and any coefficients used to combine input features and prior stage predictions. These parameters are estimated during training by minimizing a suitable loss function (e.g., mean squared error or cross-entropy) over the training episodes, using the observed target outputs y_r. The residual term ϵ_r captures unmodeled variability at stage r. The residual term ϵ_r accounts for the prediction error. This layered rule extraction allows the DT ensemble to capture interdependencies among rule components, while enabling both interpretability and domain-relevant fidelity.

4.2 Test case

4.2.1 Description

The IEEE 118-bus system environment, featured in the L2RPN competition at the 2022 World Congress on Computational Intelligence WCCI [124], serves as the test case and is illustrated in Figure 4.5. This case incorporates an adversarial module that heuristically simulates line outages to emulate the N-1 security criterion, while allocating 10% to 30% of the generation mix to RES. It comprises 118 substations, 186 lines, 91 loads, and 62 generators.

A set of 48 episodes was generated, using the methodology from [125], extracted from 12 months. Each scenario in the dataset represents a full operational episode. Scenarios were selected to cover all months of the year and a variety of operating conditions, including normal operation, high load, renewable variability, congestion-prone situations, and contingencies. This approach ensures that both the training and test sets are representative of the full range of realistic grid conditions, allowing for robust evaluation of the proposed GNP-DT agent. The episodes comprise 2018 time steps with a 5-minute resolution to feed the GNP method, in which the elite graphs generated 49,186 grid operating conditions and corresponding action set pairs for training DT models. All experiments use episode-level partitioning: each scenario corresponds to a full operational episode and is assigned entirely to either the training or test set to prevent temporal leakage. To ensure both coverage and separation, four episodes from each month were used for training and three episodes for testing, spanning the full year and capturing seasonal variations and diverse grid operating conditions. For robustness, the evaluation was conducted using 20 different random seeds for partitioning and training, ensuring that results are statistically representative and not dependent on a specific initialization. The proposed GNP algorithm interacted with the simulated power grid using the Grid2Op AI-friendly digital environment [69] as the RL environment. To ensure a robust evaluation of the proposed methodology, a distinct test set containing data from various months was employed during the testing phase.

4.2.2 Evaluation metric and benchmarks

In terms of performance indicator, the score adopted in the L2RPN competition [6] was used to allow a fair comparison with baseline agents that evaluates both the operational cost and the survivability of the agent. The score ranges from $[-100.0, 0.0, 80.0, 100.0]$, each value corresponding to a distinct level of performance by the agent during an episode. A score of -100.0 indicates that no steps were played, reflecting the maximum blackout penalty applied across all steps in all scenarios. A score of 0.0 represents the performance of a “Do-Nothing” baseline agent that takes no action and does not influence the scenario; if this agent completes the episode, the effective score range becomes $[-100.0, 0.0, 100.0]$. A score of 80.0 indicates

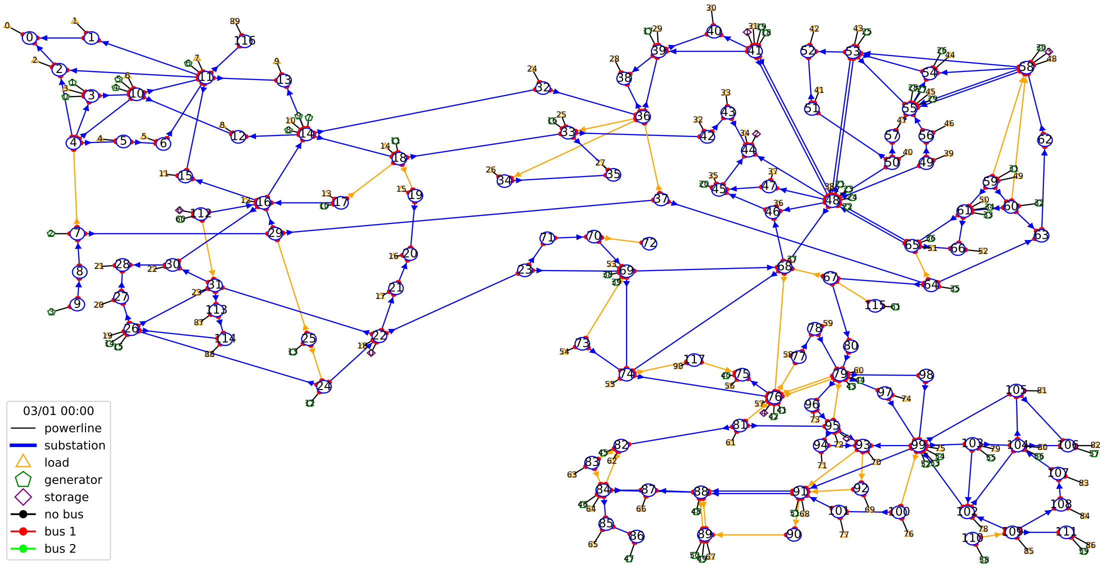


Figure 4.5: *IEEE 118-bus system of the L2RPN-WCCI-2022 competition. All the elements at each substation are connected to bus-1 in the reference topology, which could be switched to bus-2 through bus splitting actions.*

that the agent successfully plays through all scenarios without correcting line disconnections caused by overloads; losses in this case are equal to the difference between total generation and total consumption in each scenario, and lines under maintenance are reconnected after their maintenance period. The maximum score of 100.0 is awarded when the agent both completes all scenarios and optimizes system performance by reducing losses to 80% of those previously computed. Thus, higher scores reflect more efficient decision-making and longer survival across scenarios.

The following agents from the Grid2Op baselines, which employ topological reconfiguration strategies for congestion management, were selected for benchmarking:

- **ExpertOp4Grid**: An ES which tries to solve congestion issues within the grid [93] using an influence graph;
- **Enhanced ExpertOp4Grid**: An ES which is an enhanced version of *ExpertOp4Grid* agent explained in Subsection 4.1.1;
- **MILP agent**: A MILP agent that seeks to minimize the overthermal lines based on DC approximation [126] with CBC-solver [127];
- **CAgent**: Curriculum Agent (Senior agent) that consists of a Deep-RL model based on PPO [74] with a senior threshold of $\rho_{senior} = 0.95$. Note that this agent is an enhanced version of a model originally developed for a past L2RPN competition [123], which

inspired subsequent agents in later editions and also the winning solution of the 2023 Paris Region AI Challenge for Energy Transition [128].

The hyperparameters of the proposed GNP algorithm, including the crossover probabilities, mutation rate, and mutation range described in Section 4.1.2, were tuned to ensure robust convergence and satisfactory performance of the algorithm. The Bayesian optimization method was applied across 16 episodes covering the whole season of the year to ensure that the selected hyperparameters generalize beyond a single run. Table 4.4 presents the hyperparameters of the GNP algorithm. For clarity, the following nomenclature is adopted throughout the experiments: **GNP-DT-1** refers to the agent trained with the survival-duration reward function and experimental hyperparameters; **GNP-DT-2** corresponds to the agent trained with the proposed cost-aware reward function and optimized hyperparameters; and **GNP-DT-3** denotes the agent trained with the survival-duration reward function but using the optimized hyperparameters.

All experiments were conducted on a system running on a cloud-based virtual machine with an AMD EPYC CPU with 8 cores at 2.94 GHz, 32 GB of RAM, and Microsoft Windows 10 Pro.

Table 4.4: *Hyperparameters of the Adaptive GNP Algorithm*

Hyperparameter	Experimental	Optimized
Population size	30	38
Elite fraction	0.1	0.24
Crossover probability cp_1	0.5	0.41
Crossover probability cp_2	1	0.70
Mutation probability m_p	0.3	0.33
Maximum mutation range $\lambda_{\max}$	0.7	0.77

4.3 Experimental results

This section presents the experimental results, evaluating the performance of the proposed agent across multiple random seeds and scenarios. The analysis includes a comparison with three state-of-the-art baseline models (available in the Grid2Op environment) and focuses on key metrics to assess the agent's effectiveness and consistency.

Median survival time comparison. Figure 4.6 presents the median survival time across 36 test episodes for all evaluated agents. Among the proposed variants, *GNP-DT-2* and *GNP-DT-3* demonstrate consistently superior performance, reaching the maximum survival limit of 2018 steps in multiple episodes (e.g., *Jun_20*, *Sep_5_2*). Notably, the cost-aware reward used in *GNP-DT-2* results in comparable or even improved average survivability relative to the survival-based agents (*GNP-DT-1* and *GNP-DT-3*), suggesting that cost optimization implicitly encourages longer stable operation. In contrast, the optimization-based *MILP* and the deep RL-based *CAgent* exhibit greater variability across episodes, with early terminations in challenging conditions such as *Feb_28* and *Dec_12*. The expert-based *Enhanced ExpertOp4Grid* maintains steady performance and consistently surpasses the baseline *ExpertOp4Grid*, confirming the benefit of its refined heuristics. The passive *Do-Nothing* agent remains the weakest performer throughout all episodes. Overall, the GNP-DT family demonstrates the most resilient and adaptive behavior under diverse grid conditions, with *GNP-DT-2* offering the best trade-off between cost efficiency and operational stability.

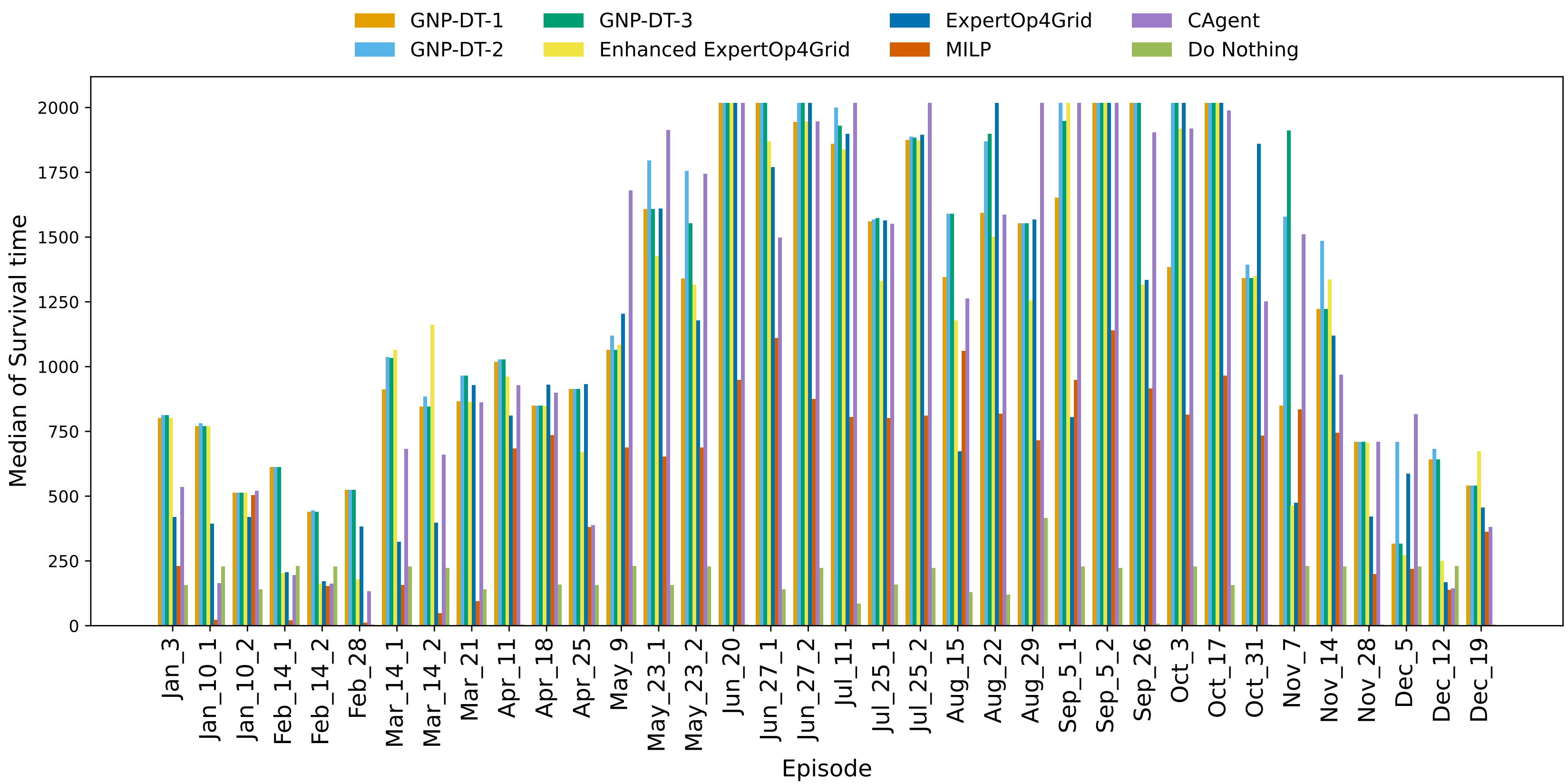


Figure 4.6: *Visualization of the median survival time per episode. Each bar corresponds to the median survival time in the respective episode across all seeds.*

Operational cost comparison. In terms of average operational cost, the proposed GNP-DT variants demonstrate a clear improvement in economic efficiency. Specifically, *GNP-DT-2*, which employs the cost-aware reward, achieved the lowest average cost of 35.12 M€, followed by *GNP-DT-3* at 35.6 M€ and *GNP-DT-1* at 38.22 M€. This indicates that integrating cost considerations into the reward function enhances both economic performance and overall stability. Among the comparative agents, the *Enhanced ExpertOp4Grid* and *CAgent* achieved moderate cost reductions with 40.15 M€ and 40.40 M€, respectively, outperforming the baseline *ExpertOp4Grid* (43.02 M€). Conversely, the optimization-based *MILP* agent exhibited the highest operational cost at 58.80 M€, reflecting limited adaptability under dynamic grid conditions. Overall, the GNP-DT family effectively balances system survivability and operational cost, with *GNP-DT-2* offering the most cost-efficient control strategy.

L2RPN performance score. The L2RPN score quantifies the overall operational effectiveness and reliability of the agents across dynamic grid conditions. Figure 4.7 illustrates the distribution of scores across 20 random seeds for all agents. Among the proposed variants, *GNP-DT-2* achieved the highest mean score of 48.10 with a median of 46.01, followed closely by *GNP-DT-3* (mean 46.99, median 45.68) and *GNP-DT-1* (mean 44.16, median 41.94). The higher median and tighter interquartile range of *GNP-DT-2* indicate both improved performance and enhanced stability, demonstrating the advantage of incorporating cost-awareness into the reward function. Among the comparative baselines, the *Enhanced ExpertOp4Grid* (mean 39.04) surpasses the original *ExpertOp4Grid* (mean 37.50) and the deep RL-based *CAgent* (mean 37.70), confirming the benefit of refined rule-based heuristics. The *CAgent* exhibits comparable spread to *ExpertOp4Grid*, ranging from 23.29 to 50.85, though with a slightly lower median. MILP, representing a DC optimization benchmark, demonstrates the widest variability and lowest scores overall, including negative values in some seeds, highlighting its instability and inefficiency in this context. Overall, the GNP-DT family—particularly *GNP-DT-2*—demonstrates superior and reliable performance across all tested configurations, balancing both resilience and cost efficiency.

Table 4.5 provides a comprehensive comparison of the agents in terms of their average L2RPN score, score distribution, and survivability metrics. The GNP-DT variants consistently outperform all benchmarks, with *GNP-DT-2* achieving the highest average score (48.1), reflecting stable and effective control. Its score distribution is particularly consistent, with a minimum of 36.04 and a third quartile of 52.73, while median survival (1147) and Median of Survival Times Across Median (MSTCM) (1439.5) indicate strong robustness across episodes (MSTCM is the median of these survival times across all episodes and all seeds, providing a measure of typical agent longevity). Compared to *CAgent* and *Enhanced ExpertOp4Grid*, *GNP-DT-2* improves the average L2RPN score by up to 28%, *GNP-DT-3* by 25%, and *GNP-DT-1* by 18%. The superior performance of *GNP-DT-2* agent trained with the cost-aware reward not only

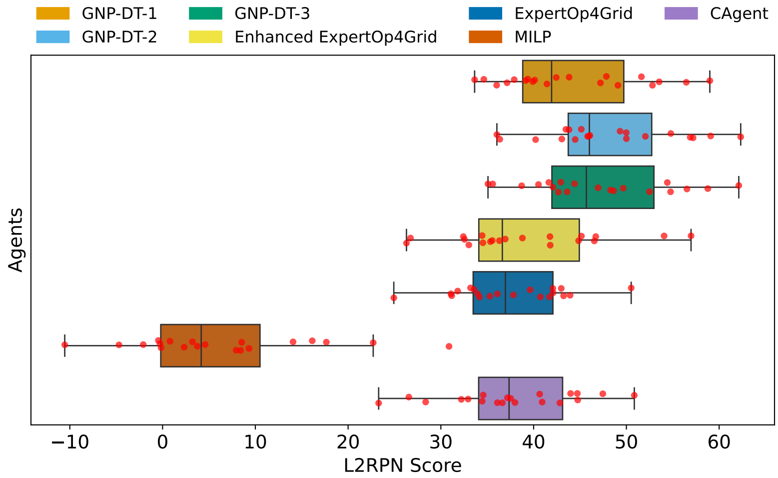


Figure 4.7: *Boxplot of the agent average score across all 20 seeds. Each point corresponds to the average score of all scenarios for one seed.*

Table 4.5: *Performance comparison between different agents on L2RPN score and survival time.*

Agent	L2RPN Score							Survival time	
	Mean	Std	Median	Q1[a]	Q3[b]	Min	Max	Median	MSTCM
Do nothing	0	0	0	0	0	0	0	158	158
ExpertOp4Grid [93]	37.5	**5.85**	36.96	33.48	42.08	24.93	50.52	804	930.75
Enhanced ExpertOp4Grid	39.04	7.95	36.63	34.09	44.93	26.3	56.97	981	1170
GNP-DT-1	44.16	7.37	41.94	38.82	49.72	33.64	59	1034	1143.25
GNP-DT-2	**48.1**	7.07	**46.01**	**43.73**	52.73	**36.04**	**62.33**	**1147**	**1439.5**
GNP-DT-3	46.99	7.43	45.68	41.97	**52.97**	35.08	62.12	1135	1282.25
MILP [126]	6.6	9.62	4.16	-0.2	10.5	-10.55	30.88	511	701.25
CAgent [74]	37.7	6.89	37.35	34.07	43.12	23.29	50.85	1032	1257.5

[a]1st quartile [b]3rd quartile

achieved a higher cumulative score but also demonstrated longer average survival times. This indicates that minimizing operational cost indirectly reinforces grid survivability, as blackout costs dominate the total episode penalty. Consequently, the cost-aware reward yields both economically and operationally efficient control policies. *GNP-DT-3* also performs strongly (mean = 46.99, median survival = 1135, MSTCM = 1282.25), demonstrating the benefits of hyper-parameter optimization even with the survival-duration reward. *GNP-DT-1* achieves a mean score of 44.16 and median survival of 1034, already surpassing conventional benchmarks and confirming the progressive improvement across the variants. Among the baseline agents, *CAgent* and *Enhanced ExpertOp4Grid* show competitive scores (37.7 and 39.04, respectively), with *CAgent* attaining a higher MSTCM (1257.5) than *ExpertOp4Grid* (930.75). In contrast, *MILP* exhibits low reliability (mean = 6.6, median survival = 511, MSTCM = 701.25), and the *Do Nothing* agent scores zero with minimal survival. Overall, GNP-DT-2 achieves the highest performance and robustness, demonstrating that cost-aware training effectively balances score maximization and operational longevity, while other agents either underperform or show greater variability in survival.

Inference time. The average time required to compute and apply an action varies significantly across agents. The proposed GNP-DT agent demonstrates the fastest inference time, requiring only 0.11 seconds on average, making it highly suitable for real-time deployment. In comparison, the *ExpertOp4Grid* baseline responds in 0.54 seconds, while both the *Enhanced ExpertOp4Grid* and *CAgent* take longer, averaging 2.23 and 2.33 seconds, respectively. The MILP method is the slowest, with an average inference time of 20.58 seconds, an expected outcome given that it requires solving an MILP problem at each timestep. These results highlight the advantage of GNP-DT in real-time decision making under operational time constraints.

4.4 Interpretability: A discussion

This section discusses the interpretability of the GNP-DT method during both the learning and operational phases.

4.4.1 Learning phase

To provide some insight into the reasoning window of the learning phase, Tables 4.6 and Table 4.7 together illustrate three decision graphs G_a, G_b, and G_c to address line congestion in a specific period of an episode. In Table 4.6, HB, LB, and DSB stand for hub bus, buses on the looped path, and downstream buses (Table 4.1 defines each type of these buses), respectively. In addition, the numbers in the detection rows represent the congested lines. The behavior of each graph is shaped by its internal judgment nodes and their relevant functions and parameters, as mentioned in Table 4.1 in Subsection 4.1.1. The key parameters that distinguish the graphs are presented in Table 4.7, which can cause changes in the functions of the graphs as follows:

- The threshold I_{th} defines the overload limit used to trigger the action proposal process; both G_a and G_c use a strict threshold (100%), while G_b lowers this to 90%, allowing for a proactive action.

- The parameter β_{th}, which sets the threshold coefficient for determining the influence path, is highest in G_b (0.6), indicating a more rigorous but potentially narrower focus in selecting effective paths. In contrast, G_c balances this (0.4), leading to more robust but flexible action pathways. For example, with the same detection criterion, when congestion starts on lines 12 and 20 and is followed by lines 41, 44 and 51, the flexible units activated by G_c were more effective in alleviating congestion at each timestep, the trajectory of actions (bus 67$\rightarrow$ bus 81$\rightarrow$ bus 93$\rightarrow$ bus 95) outperformed the selection of actions (bus 68$\rightarrow$ bus 76$\rightarrow$ lines 24 and 146) by G_a.

- The σ^1_{cp} mentioned in Table 4.1 stands for the actions with the highest priority, where G_a and G_c give the highest focus to the buses at the hub points, and G_b has the most tendency to reconfigure the buses on the looped path, which was in some timesteps a less effective selection.

- The quality of the action in terms of the topological score is indicated by σ_t as mentioned in Table 4.1. In this regard, releasing a critical congestion is satisfying for G_a and G_b, while G_c demands full topological resolution of all congestion, critical or not, before terminating the search. This stricter criterion ensures that G_c systematically avoids the risk of creating new congestions and provides a safer solution that relieves all overloads.

The observed action trajectory confirms these architectural differences. G_a acts sequentially on hub buses but fails to prevent collapse due to late response and incomplete resolution. G_b detects earlier and applies diverse actions, including looped and downstream bus reconfigurations, surviving the scenario later because it raises more overload concerns with its action trajectory. G_c, however, mirrors G_a in the type of action but distinguishes itself by selecting actions that fully mitigate system risk, leading to survival without creating new overloads. These comparisons demonstrate the interpretability of the proposed method during the learning phase and support human operators in reasoning about and building trust in the resulting elite graphs. For example, this traceable possibility in GNP can provide the operator with such a sensitivity analysis on the graph parameters that justifies the superiority of G_c over G_a and G_b.

In addition, this example provides insight into the evolutionary phase of the GNP. The initial seed corresponds to the baseline ES (Ga). Through evolutionary operators and reinforcement-learning-based fitness evaluation, the GNP progressively adapts, producing intermediate graphs (Gb) and ultimately an elite graph (Gc). Comparative results show that Gc consistently outperforms both Ga and Gb, confirming that the performance gains are attributable to the learning and adaptation process rather than to the expert initialization alone. This distinction highlights that while the expert provides a structured starting point, the final evolved agent embodies novel strategies that emerge through the learning process.

Table 4.6: *Comparison of decision graph action trajectory corresponds to the line congestion detection across 9 illustrative timesteps*

Time step	1	2	3	4	5	6	7	8	9
G_a: **Detection**	–	–	$\{12, 20\}$	$\{20\}$	$\{41, 44, 51, 175, 185\}$	collapse	collapse	collapse	collapse
G_a: **Action**	–	–	reconfig. HB 68	reconfig. HB 76	disconn. $\{24, 146\}$	–	–	–	–
G_b: **Detection**	$\{20\}$	$\{41, 51\}$	$\{41, 51\}$	$\{12\}$	$\{12\}$	$\{41, 44, 51\}$	$\{175\}$	$\{175\}$	–
G_b: **Action**	reconfig. LB 81	disconn. $\{24, 16\}$	reconfig. LB 79	disconn. $\{13, 122\}$	reconfig. LB 79	reconfig. DSB 95	reconfig. DSB 67	reconfig. DSB 79	–
G_c: **Detection**	–	–	$\{12, 20\}$	$\{20\}$	$\{41, 44, 51\}$	$\{41\}$	–	–	–
G_c: **Action**	–	–	reconfig. HB 67	reconfig. HB 81	reconfig. HB 93	reconfig. HB 95	–	–	–

Table 4.7: *Summary of functional differences between decision graphs*

Aspect	G_a	G_b	G_c
I_{th}	100%	90%	100%
β_{th}	0.2	0.6	0.4
σ_{cp}^1	Hub	Loop	Hub
σ_t	Critical issue	Critical issue	All issues

4.4.2 Operational phase

To demonstrate the interpretability of the proposed method during the operational phase, a representative portion of the complete DT is illustrated in Figure 4.8. This selected partial tree is from the DT3 model (bus reconfiguration), which is relevant to determining the ID of the flexible unit (DT2) when the flexibility type (DT1) and the number of needed actions (DT2) have been previously predicted.

In the DT plot, *value* attribute represents the proportional distribution of these instances across classes, and the *class* attribute corresponds to the majority class, indicating the predicted outcome at that node. Each box's color represents the dominant class in that node. Nodes sharing the same color are mostly associated with the same class. The intensity of the color indicates the node's purity – darker shades mean the node contains mostly one class, while lighter shades suggest a mix of classes. This visual aid helps quickly identify which class is favored and how confidently the tree makes its splits. Moreover, left arrows correspond to *True*, and right arrows correspond to *False* for each split.

This example determines the next control action based on the grid congestion state and prior topology changes. *Cong. Line ID_1* represents the identifier of the line with the highest loading percentage (i.e., most congested) at the current timestep, followed by *Cong. Line ID_2*, *Cong. Line ID_3*, etc., sorted in descending order based on their loading levels. *Split Bus ID_i* indicates the ID of a bus that has already been reconfigured, sorted by ID. The *class* output of the tree corresponds to the ID of a candidate bus selected for reconfiguration. Due to limited space, the tree was plotted for only three classes using reconfiguration buses 16, 67, and 93. Some illustrative human-readable rules could be extracted from the DT as follows:

1. If *Cong. Line ID_1* is in $Z1$ (i.e., the line with the maximum overload is situated in zone $Z1$) and *Cong. Line ID_2* is in $Z1$ or $Z2$, the model selects **bus 67** from $Z1$ as the next reconfiguration candidate—that is, the bus proposed for splitting to alleviate congestion at the current timestep.

2. If *Cong. Line ID_1* is in $Z2$, and *Cong. Line ID_4* in $Z3$, and *Split Bus ID_1* in $\{Z1, Z2, Z3\}$, then the model returns **bus 93** from $Z3$ for splitting.

3. If *Cong. Line ID_1* out of $\{Z1, Z2\}$ and *Split Bus ID_2* is in $\{Z1, Z2, Z3\}$, **bus 16** from $Z2$ is the candidate for reconfiguration action (bus splitting).

4. If *Cong. Line ID_1* is in $Z2$, and *Cong. Line ID_4* is in $\{Z1, Z2, Z3\}$, then the model selects **bus 67** from $Z1$ for bus splitting.

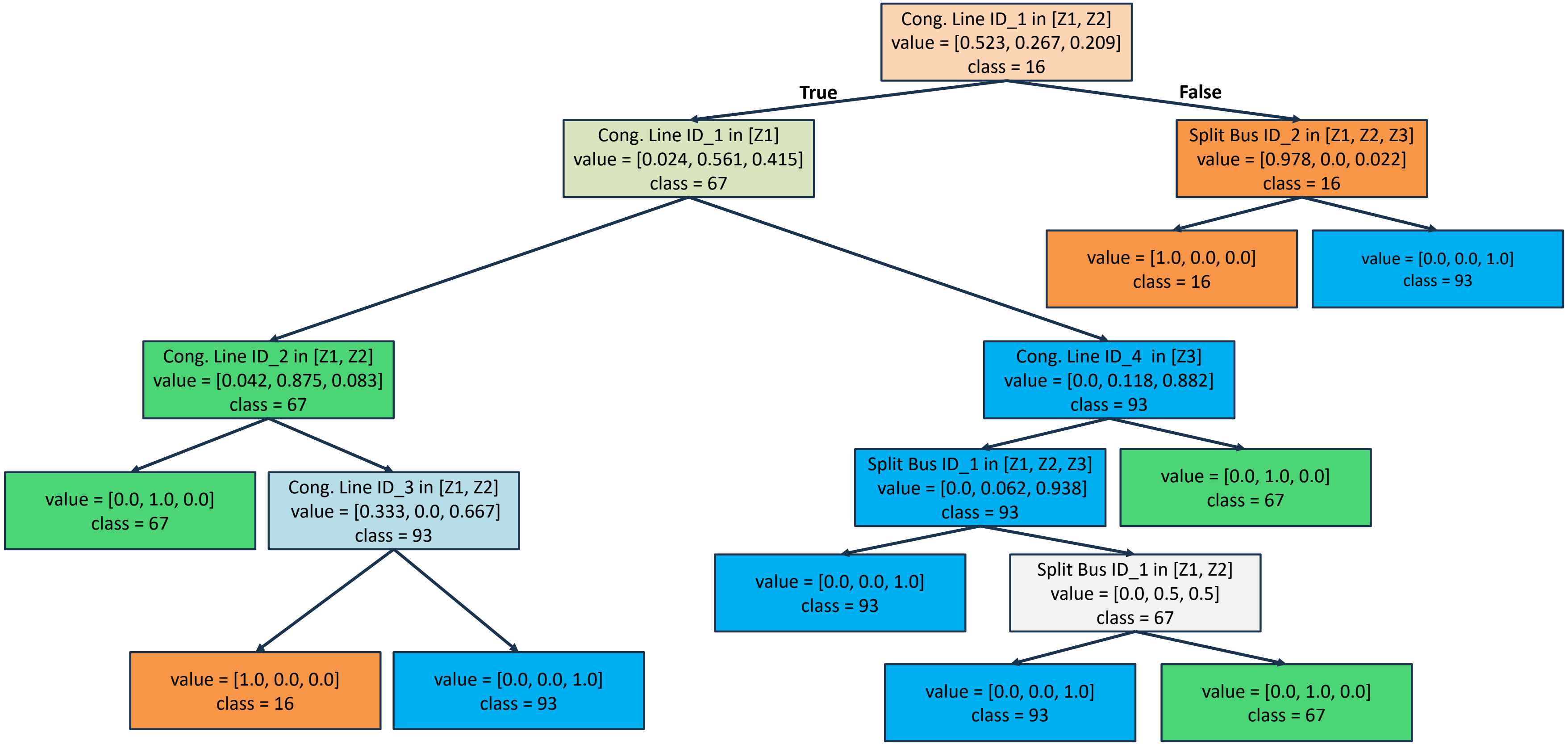


Figure 4.8: *Excerpt from DT3 (bus reconfig.) for flexible unit ID determination*

The interpretability achieved through the DT extraction is primarily logical, as it expresses the agent's control policy in the form of explicit "if–then" decision rules. Although such interpretability does not constitute causal reasoning in the strict analytical sense, it provides a practical and meaningful explanation of the agent's actions that aligns with human operator expertise. Each path in the DT directly links observable grid states to corresponding control decisions, allowing operators to trace why a particular action is recommended under specific system conditions and another action is suitable under a different operating condition. For example, a rule such as "If congestion occurs in Zone-1 and an overload also exists in Zone-2, then split Bus-X and Bus-Y could be reconfigured only if the congestion is in Zone-1" demonstrates the underlying decision logic that connects system observations to a physically interpretable control measure. This transparency enables operators to validate, trust, and refine the automated control behavior, bridging the gap between data-driven intelligence and human operational understanding. Consequently, the proposed interpretability mechanism provides actionable insight into the decision-making process, even though it remains logically rather than causally explanatory.

4.5 Final Remarks

This work proposed an interpretable and adaptive control framework based on GNP-DT for real-time congestion management in power systems. The agent integrates rule-based structures with RL to evolve effective control strategies, maintaining both interpretability and adaptability. Through extensive evaluation across multiple stochastic seeds and benchmark agents, GNP-DT demonstrated superior performance in terms of L2RPN score, survival time, and operational cost compared to baseline methods of different natures, including an ES, MILP with DC approximation, and a deep RL-based agent (*CAgent*). In particular, GNP-DT achieved an improvement in the average score of up to 28% over *ExpertOp4Grid* and the Deep-RL agent and consistently ensured a higher median survival and a lower average operational cost.

Moreover, the framework maintained high computational efficiency, exhibiting significantly shorter inference times than optimization-based and deep learning approaches. The alignment of GNP-DT performance with interpretable design principles supports its practical applicability in real-world scenarios that require transparent and reliable grid control.

Chapter 5

Bridging Optimization and Operator Insight: LLM-Assisted Counterfactual Analysis of AC-OPF

In modern power systems, it is essential to maintain secure and economical operation under increasing renewable penetration and uncertain load patterns; this requires continuous adjustment of generator outputs, a process commonly referred to as *redispatch*. The primary objectives of redispatch are to relieve congestion and restore voltage profiles under varying operating conditions. This task is typically formulated as an AC-OPF problem [129]:

$$\begin{aligned} \min_{\mathbf{P_g},\mathbf{Q_g},\mathbf{V},\boldsymbol{\theta}} \quad & C(\mathbf{P_g}) \\ \text{s.t.} \quad & \mathbf{g}(\mathbf{P_g},\mathbf{Q_g},\mathbf{V},\boldsymbol{\theta}) = 0, \\ & \mathbf{h}(\mathbf{P_g},\mathbf{Q_g},\mathbf{V},\boldsymbol{\theta}) \leq 0 \end{aligned} \tag{5.1}$$

where $\mathbf{P_g}$ and $\mathbf{Q_g}$ denote the active and reactive power outputs of all generators, while $\mathbf{V}$ and $\boldsymbol{\theta}$ represent the bus voltage magnitudes and angles. The objective function $C(\mathbf{P_g})$ typically captures the total generation cost, often modeled as a sum of quadratic or piecewise-linear functions. The equality constraints $\mathbf{g}(\mathbf{P_g},\mathbf{Q_g},\mathbf{V},\boldsymbol{\theta}) = 0$ enforce the AC power flow balance at each bus, whereas the inequality constraints $\mathbf{h}(\mathbf{P_g},\mathbf{Q_g},\mathbf{V},\boldsymbol{\theta}) \leq 0$ capture operational limits, including generator output bounds, bus voltage limits, and transmission line flow limits.

Despite the widespread use of AC-OPF, its solutions remain largely opaque to human operators. Optimization solvers output redispatch setpoints and dual variables; however, they offer limited insight into *why* certain generators are adjusted or *how* specific constraints drive the solution. As a result, redispatch decisions derived from AC-OPF are often perceived as "black boxes", providing minimal interpretability in operational contexts where human understanding and justification are essential.

5.1 Redispatch with AC-OPF

The operational problem is formulated as a two-stage AC-OPF model comprising an initial *dispatch* stage and a subsequent *redispatch* stage. In practice, the dispatch stage is determined by the electricity market ahead of real time, whereas redispatch is carried out by the system operator closer to real time to resolve network constraints, such as line congestion and voltage-limit violations, that arise under actual operating conditions [130, 131]. For the sake of simplicity, the market dispatch is emulated here using an AC-OPF model. This provides a consistent representation of the initial operating point within the same optimization framework, while noting that, in a real-world setting, this dispatch would correspond to the market outcome.

In the first stage, the emulated dispatch problem determines generator set-points and bus voltage magnitudes and angles so as to satisfy the AC power-flow equations and network operating limits. In the second stage, redispatch adjusts the initial schedule under updated system conditions in order to relieve network constraints and restore secure system operation. The corresponding objective function is

$$\min J^{\mathrm{DA}} = \sum_{i \in \mathscr{G}^{\mathrm{conv}}} c_i^{\mathrm{DA}} P_{g,i} + \sum_{i \in \mathscr{R}} c_i^{\mathrm{res}} C_i^{\mathrm{res}} + M \sum_{i \in \mathscr{B}} \Phi_i, \tag{5.2}$$

where $\mathscr{G}^{\mathrm{conv}}$ is the set of conventional generators, $\mathscr{R}$ is the set of renewable generators, C_i^{res} denotes renewable curtailment, and $M \gg 1$ is a large penalty coefficient applied to aggregated slack variables Φ_i to maintain robustness.

At each bus $i \in \mathscr{B}$, active and reactive power balance are enforced through the AC network equations:

$$\sum_{g \in \mathscr{G}_i} P_{g,i} - P_{d,i} + P_{i,+}^{\mathrm{bal,slack}} - P_{i,-}^{\mathrm{bal,slack}} = V_i \sum_{k=1}^{N_b} V_k \left(G_{ik} \cos\theta_{ik} + B_{ik} \sin\theta_{ik} \right), \tag{5.3}$$

$$\sum_{g \in \mathscr{G}_i} Q_{g,i} - Q_{d,i} + Q_{i,+}^{\mathrm{bal,slack}} - Q_{i,-}^{\mathrm{bal,slack}} = V_i \sum_{k=1}^{N_b} V_k \left(G_{ik} \sin\theta_{ik} - B_{ik} \cos\theta_{ik} \right), \tag{5.4}$$

where $\theta_{ik} = \theta_i - \theta_k$ and (G_{ik}, B_{ik}) are the elements of the bus-admittance matrix $\mathbf{Y}_{\mathrm{bus}}$. The slack bus s is used only as the reference bus, with

$$\theta_s = 0. \tag{5.5}$$

Conventional generators satisfy their physical bounds,

$$P_i^{\min} \le P_{g,i} \le P_i^{\max}, \qquad Q_i^{\min} \le Q_{g,i} \le Q_i^{\max}, \tag{5.6}$$

whereas renewable generators are modeled through an availability–curtailment relation,

$$P_{g,i} + C_i^{\text{res}} = P_i^{\text{avail}}, \qquad i \in \mathscr{R}, \tag{5.7}$$

with P_i^{avail} denoting the forecasted available renewable power. Hence, renewable generators are not redispatchable upward; they may only be curtailed if required by feasibility or congestion management.

Bus voltages are maintained within admissible limits using soft bounds,

$$V^{\min} - V_i^{\text{min,slack}} \leq V_i \leq V^{\max} + V_i^{\text{max,slack}}, \tag{5.8}$$

and line thermal constraints are enforced at both ends of each branch ℓ:

$$(P_\ell^f)^2 + (Q_\ell^f)^2 \leq (S_\ell^{\max})^2 + S_\ell^{\text{slack}}, \tag{5.9}$$

$$(P_\ell^t)^2 + (Q_\ell^t)^2 \leq (S_\ell^{\max})^2 + S_\ell^{\text{slack}}, \tag{5.10}$$

where (P_ℓ^f, Q_ℓ^f) and (P_ℓ^t, Q_ℓ^t) are the sending- and receiving-end apparent-power components, respectively.

In the second stage, redispatch is performed after updated load and renewable conditions become available. The redispatch problem preserves the AC network model but optimizes generator deviations around the dispatch-stage schedule. Conventional generators are adjusted through nonnegative upward and downward redispatch variables,

$$P_{g,i}^{\text{RD}} = P_{g,i}^{\text{DA}} + \Delta P_i^+ - \Delta P_i^-, \qquad i \in \mathscr{G}^{\text{conv}}, \tag{5.11}$$

where $P_{g,i}^{\text{DA}}$ is the dispatch-stage setpoint, and ΔP_i^+ and ΔP_i^- denote upward and downward redispatch. These are limited by redispatch or ramping capabilities,

$$0 \leq \Delta P_i^+ \leq \overline{\Delta P}_i^+, \qquad 0 \leq \Delta P_i^- \leq \overline{\Delta P}_i^-. \tag{5.12}$$

Renewable generators remain non-dispatchable upward in redispatch and continue to satisfy

$$P_{g,i}^{\text{RD}} + C_i^{\text{res,RD}} = P_i^{\text{avail,RD}}, \qquad i \in \mathscr{R}, \tag{5.13}$$

where $P_i^{\text{avail,RD}}$ is the updated realized renewable availability. Hence, the redispatch stage adapts the conventional fleet around the initial dispatch while allowing renewable curtailment and load shedding when required.

The redispatch objective minimizes the cost of deviations from the dispatch-stage schedule:

$$\min J^{\text{RD}} = \sum_{i \in \mathscr{G}^{\text{conv}}} \left(c_i^{\text{up}} \Delta P_i^+ + c_i^{\text{down}} \Delta P_i^- \right) + \sum_{i \in \mathscr{R}} c_i^{\text{res}} C_i^{\text{res,RD}} + M \sum_{i \in \mathscr{B}} \Phi_i^{\text{RD}}, \tag{5.14}$$

where c_i^{up} and c_i^{down} are the upward and downward redispatch prices of conventional generators, $C_i^{\text{res,RD}}$ denotes renewable curtailment in the redispatch stage, and $M \gg 1$ is a large penalty coefficient applied to the redispatch-stage slack variables Φ_i^{RD}. The AC power-balance equations, generator limits, voltage limits, and thermal constraints retain the same structure as in the dispatch stage, but are evaluated using updated redispatch-stage load and renewable conditions.

The formulation distinguishes clearly between conventional generators, which can be redispatched up or down, and renewable generators, which are modeled as available but curtailable resources. This makes the framework suitable for studying congestion management, renewable integration, and the economic implications of forecast-driven redispatch in AC-constrained transmission systems.

5.2 Physical Operational Analysis

The Physical Operational Analysis (POA) module does not constitute a separate optimization method. Instead, it refers to the set of information extracted from the AC-OPF redispatch solution to interpret the physical drivers of the result. In this work, POA is limited to the active constraints and associated dual variables of the OPF solution, together with Power Transfer Distribution Factor (PTDF) and LODF sensitivity indicators. Its purpose is to explain why a given redispatch pattern emerges by identifying the constraints associated with the solution, quantifying their influence through dual variables, and using sensitivity factors to construct counterfactual operating insights. Rather than solving a new optimization problem, POA examines the OPF snapshot and indicates how small changes in injections would have altered constraint violations or branch flows. This enables physically grounded explanations of the form: *What minimal redispatch could have avoided the constraint that drove the OPF solution?*

The extracted information is organized by first identifying the active constraints of the OPF solution. Consider a generic inequality constraint

$$g_i(x) \leq b_i. \tag{5.15}$$

At the optimal solution x^*, the constraint is considered active when

$$|g_i(x^*) - b_i| < \varepsilon \tag{5.16}$$

and the associated Lagrange multiplier satisfies

$$\lambda_i^* > 0. \tag{5.17}$$

The collection of such constraints forms the active set. In practical terms, the active set contains the limits that are binding, or nearly binding, at the optimum and therefore shape the feasible redispatch pattern.

The dual variables provide a quantitative measure of the influence of each active constraint on the optimal objective. In economic terms, a dual variable can be interpreted as the *shadow price* of the corresponding constraint, that is, the marginal change in the optimal objective value associated with a small relaxation of the constraint bound. The marginal sensitivity of the optimal objective value f^* with respect to the constraint bound is

$$\frac{\partial f^*}{\partial b_i} = -\lambda_i^*. \tag{5.18}$$

Large dual magnitudes therefore indicate operational limits that strongly influence the optimal redispatch. In practice, thermal limits, generator limits, or voltage limits with significant dual values reveal which physical constraints forced the solution. These constraints serve as the starting point for constructing counterfactual explanations of the OPF outcome.

To interpret how redispatch actions affect network constraints, PTDFs and LODFs are used as first-order screening indicators. Rather than relying on a full local AC Jacobian for explanation, the analysis employs linear active-power transfer factors that map incremental net injections to approximate branch-flow changes under the standard lossless DC network approximation [132, 133]. Let the vector of incremental net bus injections induced by redispatch be

$$\Delta\mathbf{p} \in \mathbb{R}^{n_b}, \tag{5.19}$$

with the convention that generation increases enter positively and generation decreases enter negatively. Power balance implies

$$\mathbf{1}^\top \Delta\mathbf{p} = 0, \tag{5.20}$$

so that redispatch is interpreted as a balanced transfer across the network.

Under the DC approximation, the incremental branch-flow vector satisfies

$$\Delta\mathbf{f} \approx \mathbf{H}\,\Delta\mathbf{p}, \tag{5.21}$$

where $\mathbf{H} \in \mathbb{R}^{n_\ell \times n_b}$ is the PTDF matrix. If $\mathbf{A} \in \mathbb{R}^{n_\ell \times n_b}$ denotes the branch–bus incidence matrix and $\mathbf{B}_\ell = \mathrm{diag}(b_\ell)$ is the diagonal matrix of branch susceptances, then the reduced DC network matrices satisfy

$$\mathbf{B}_{\mathrm{bus}} = \mathbf{A}^\top \mathbf{B}_\ell \mathbf{A}, \qquad \mathbf{B}_f = \mathbf{B}_\ell \mathbf{A}, \tag{5.22}$$

and the PTDF matrix may be written compactly as

$$\mathbf{H} = \mathbf{B}_f \mathbf{B}_{\mathrm{bus}}^{-1}, \tag{5.23}$$

after removing the slack-bus degree of freedom in the usual way [132, 133]. Thus each coefficient $H_{\ell i}$ gives the approximate change in active power flow on line ℓ caused by a +1 MW injection at bus i and a corresponding withdrawal at the reference bus.

For a balanced transaction that increases injection at bus i and decreases injection at bus j by the same amount ΔP, the approximate change in monitored line flow is

$$\Delta f_\ell \approx \left(H_{\ell i} - H_{\ell j}\right)\Delta P. \tag{5.24}$$

This representation is directly useful for redispatch screening. If line ℓ is overloaded and an approximate relief of $\Delta F_\ell^{\text{req}}$ is needed, then the redispatch magnitude required from a two-bus transaction (i, j) can be estimated as

$$\Delta P_{i \leftarrow j} \approx \frac{\Delta F_\ell^{\text{req}}}{\left|H_{\ell i} - H_{\ell j}\right|}, \tag{5.25}$$

provided that $H_{\ell i} - H_{\ell j}$ acts in the relieving direction. In the same spirit, if the report provides loading percentages and thermal limits, the required relief on line ℓ can be screened as

$$\Delta F_\ell^{\text{req}} \approx \left(\frac{\lambda_\ell - \lambda_\ell^\star}{100}\right) S_\ell^{\max}, \tag{5.26}$$

where λ_ℓ is the current loading percentage, $\lambda_\ell^\star$ is the target loading percentage (for example 100%), and $S_\ell^{\max}$ is the thermal limit. These relationships support counterfactual reasoning of the form: *How much redispatch between two buses would be required to reduce the loading of a stressed branch by a specified amount?*

Beyond redispatch directionality, POA also evaluates approximate $N-1$ security implications using LODFs. For an outage of line k, the LODF coefficient $\text{LODF}_{\ell,k}$ represents the fraction of the pre-contingency flow on line k that is redistributed to monitored line ℓ after the outage [134, 135]. Let line k connect buses (m, n). If $\text{PTDF}_{\ell,(m\to n)}$ denotes the transfer distribution factor on monitored line ℓ associated with a transaction from bus m to bus n, then the standard single-line outage relationship is

$$\text{LODF}_{\ell,k} = \frac{\text{PTDF}_{\ell,(m\to n)}}{1 - \text{PTDF}_{k,(m\to n)}}, \qquad \ell \neq k, \tag{5.27}$$

with

$$\text{LODF}_{k,k} = -1. \tag{5.28}$$

The corresponding first-order post-contingency flow estimate is then

$$f_\ell^{\text{post}} \approx f_\ell^{\text{pre}} + \text{LODF}_{\ell,k}\, f_k^{\text{pre}}, \tag{5.29}$$

which provides a rapid screening approximation of outage-induced rerouting without solving an additional power-flow problem.

Within the POA information set, the PTDF formalism is used to identify relieving versus aggravating redispatch directions and to estimate the order of magnitude of redispatch required to relieve congested branches, while LODFs are used to assess whether a branch outage would substantially shift flow onto already stressed lines. These factors are therefore not used as exact AC-feasibility certificates, but as transparent first-order indicators that support counterfactual explanation of redispatch decisions and rapid qualitative-to-semiquantitative security screening [132, 134, 135].

Within the POA information set, these mathematical elements are combined to produce physically interpretable explanations of OPF outcomes. The active set identifies the constraints that shape the redispatch solution, dual variables quantify their marginal influence on the objective, PTDF indicates how balanced injection shifts affect monitored line flows, and LODF factors reveal whether alternative redispatch patterns would compromise $N-1$ security. Counterfactual scenarios are then constructed by estimating how small redispatch adjustments would modify constraint violations or branch loadings. The resulting explanation therefore connects the active physical constraints, the feasible redispatch directions identified by sensitivities, and the economic implications reflected in the dual variables.

5.3 Multi-Agent LLM Framework

A multi-agent LLM framework (depicted in Figure 5.1) was developed to generate physically grounded, operator-oriented explanations for AC-OPF redispatch outcomes. The purpose of the framework is not to replace the optimization model, nor to solve an additional OPF instance, but rather to transform the numerical evidence produced by the OPF and the POA module into a concise explanation that resembles the reasoning process of a transmission system operator. In particular, the framework is designed to explain why a given redispatch occurred, which network or device constraint forced the redispatch away from the unconstrained economic solution, what minimal corrective redispatch would approximately relieve the active limitation, and why alternative actions are either infeasible, more expensive, or operationally undesirable. The explanatory process is therefore explicitly constrained by physical evidence extracted from the snapshot, including pre- and post-OPF operating points.

Let $\mathscr{S}_t$ denote the system snapshot at time index t. The snapshot contains the operating quantities needed for explanation: generator active and reactive outputs and limits, branch currents or flows and their thermal ratings, bus voltages and their admissible bounds, the AC-OPF solution, the objective value, and the POA-derived quantities used to interpret the redispatch. We write

$$\mathscr{S}_t = \left(\mathbf{x}_t^{\mathrm{pre}}, \mathbf{x}_t^{\mathrm{opf}}, \boldsymbol{\lambda}_t, \boldsymbol{\mu}_t, \boldsymbol{\Gamma}_t, \mathbf{L}_t\right), \tag{5.30}$$

where $\mathbf{x}_t^{\mathrm{pre}}$ denotes the pre-redispatch state, $\mathbf{x}_t^{\mathrm{opf}}$ the optimized post-redispatch state, $\boldsymbol{\lambda}_t$ the power-balance dual variables, $\boldsymbol{\mu}_t$ the dual variables associated with active inequality constraints, $\boldsymbol{\Gamma}_t$ the set of linearized sensitivities, and $\mathbf{L}_t$ the matrix or collection of LODF values used for qualitative contingency screening. The sensitivities in $\boldsymbol{\Gamma}_t$ provide a local linear ap-

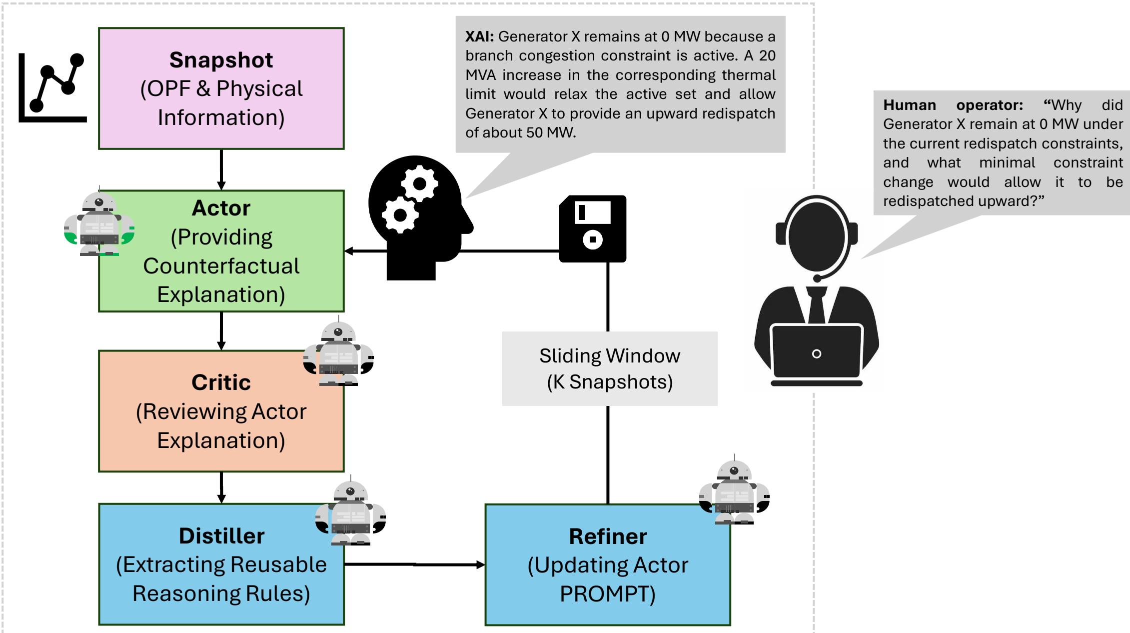


Figure 5.1: *Multi-agent LLM framework for AC-OPF redispatch explanation in control rooms*

proximation of how active and reactive injections affect line currents and bus voltages. These quantities are fundamental because the explanation agent is explicitly instructed to "rely ONLY on the provided snapshot, dual variables, sensitivities, and LODF values" and to "Do NOT solve OPF or power flow." Hence, every explanation must be traceable to numerical evidence already available in $\mathscr{S}_t$.

The central component is an Actor agent, denoted by A_θ, parameterized by θ, whose role is to map the physical snapshot and a task prompt into a single operator-facing explanation. Formally,

$$e_t = A_\theta(\mathscr{S}_t, p_t), \tag{5.31}$$

where p_t is the current Actor prompt and e_t is the generated explanation for snapshot t. The Actor is not a free-form summarizer; rather, it is guided by a tightly structured instruction set that requires a physically justified explanation of the redispatch. In the prompt, the Actor is told that its "task is to explain why an AC-OPF redispatch occurred by constructing numerically grounded counterfactual scenarios," and that the final explanation should mimic "how a grid operator would justify a redispatch decision." This design choice is important because it forces the language model to organize the explanation around physical constraints and corrective actions rather than around generic descriptive text. Operationally, the Actor is expected to identify the dominant active constraint, quantify the minimum relief required, infer promising redispatch directions from sensitivities, assess feasibility with respect to generator, voltage, and reactive limits, and compare the chosen action against plausible alternatives in both physical and economic terms.

The internal logic of the Actor can be viewed as a constrained reasoning map from the POA outputs to a counterfactual narrative (See Actor prompt in Appendix C.1). Suppose that branch $\ell^\star$ is identified as the dominant limiting element, either because its pre-OPF current exceeds its rating or because the corresponding thermal dual variable has large magnitude. If the relevant loading quantity is current, then the minimum required relief is approximated by

$$\Delta I_{\ell^\star}^{\mathrm{req}} \approx I_{\ell^\star}^{\mathrm{pre}} - I_{\ell^\star}^{\max}, \qquad \text{for } I_{\ell^\star}^{\mathrm{pre}} > I_{\ell^\star}^{\max}. \tag{5.32}$$

The Actor prompt explicitly encodes this step through the instruction "RequiredRelief ≈ CurrentCurrent − Limit," which operationalizes the first-order estimate of how much current must be removed from the constrained branch. The next step is to identify candidate redispatch directions using the PTDF coefficients of the monitored branch. For a balanced transfer that increases active-power injection at bus i and decreases it at bus j by the same amount ΔP, the first-order change in the active power flow of branch $\ell^\star$ is approximated by

$$\Delta f_{\ell^\star} \approx \left(\mathrm{PTDF}_{\ell^\star,i} - \mathrm{PTDF}_{\ell^\star,j}\right)\Delta P. \tag{5.33}$$

Accordingly, a redispatch direction is *aggravating* for branch $\ell^\star$ if

$$\mathrm{PTDF}_{\ell^\star,i} - \mathrm{PTDF}_{\ell^\star,j} > 0, \tag{5.34}$$

since increasing injection at bus i and reducing it at bus j increases the flow on the monitored branch in its reference direction. Conversely, a negative value indicates a *relieving* redispatch direction. Therefore, PTDFs provide a direct first-order criterion for selecting candidate generator redispatch pairs that are expected to reduce stress on the binding branch. this is to confirm that the sensitivity is sufficiently large in magnitude and the redispatch direction has the correct sign to reduce the overload. This relation does not claim exactness; rather, it provides the numerical backbone of a counterfactual statement such as: if generation were increased at a relieving location or decreased at a stressing location by approximately $\Delta P_n^{\mathrm{req}}$, the thermal margin could be restored to first order. The explanation produced by the Actor is therefore a physically informed local interpretation of the AC-OPF decision, not a re-optimization.

To maintain discipline in the generated explanations, the Actor is also instructed to treat the dual variables differently according to their physical role. The power-balance duals are interpreted as nodal marginal prices, while the duals associated with thermal, active-power, reactive-power, or voltage limits are interpreted as indicators of active operational constraints. In practical terms, if μ_ℓ^{th} is the thermal-limit dual for branch ℓ, a large value of $|\mu_\ell^{\mathrm{th}}|$ signals that the associated thermal bound is an economically and operationally significant driver of redispatch. Similarly, active duals on generator limits or voltage bounds reveal when redispatch flexibility is blocked by saturation or local voltage feasibility. The Actor is therefore expected to combine primal evidence (overloaded line, generator at maximum, voltage near bound) with dual evidence (nonzero or large-magnitude shadow prices) to construct an explanation that begins from the physical bottleneck rather than from the redispatch outcome itself.

A key property of the framework is that the Actor must generate at least one explicit counterfactual scenario. This requirement is central because it transforms the explanation from a descriptive summary into a causal justification. Let $\mathbf{u}_t$ denote a candidate counterfactual redispatch vector, for example an increase at a relieving generator combined with a decrease at a stressing generator. The Actor estimates the resulting flow effect and the incremental economic effect by

$$\Delta C(\mathbf{u}_t) \approx \sum_n \pi_n \Delta P_n, \tag{5.35}$$

where π_n is interpreted from the nodal marginal information in the snapshot. The explanation must then state whether the candidate action relieves congestion, whether it is feasible with respect to operating limits, and whether it is more or less attractive than the redispatch selected by the AC-OPF. In this sense, the Actor does not merely justify the chosen redispatch by saying that a line was overloaded; it explains why a particular physically admissible corrective direction is preferred over nearby alternatives once the available flexibility, locational influence, and approximate cost are jointly considered.

In addition to base-case feasibility, the framework includes a lightweight security interpretation using LODF-based screening. Let k denote a contingency corresponding to the outage of lines. These factors are not used to claim exact post-contingency AC feasibility, but they provide a qualitative signal of whether the redispatch is associated with a branch that may remain security-critical under plausible line outages. The Actor prompt accordingly instructs the model to use LODF information only for “qualitative screening.” This is methodologically important: it prevents the language model from overstating the validity of linearized security predictions while still allowing the generated explanation to mention whether the constrained branch appears particularly sensitive to rerouting under an N-1 event.

The explanation generated by the Actor is next passed to an AI Critic (See Critic prompt in Appendix C.2), denoted by C^{AI}, whose role is to verify whether the output satisfies the structural and technical requirements of the task. The Critic does not replace expert judgment and does not re-solve the physical problem; rather, it acts as a formal verifier of completeness, grounding, and compliance with the intended reasoning pattern:

$$v_t = C^{\mathrm{AI}}(\mathscr{S}_t, e_t), \qquad v_t \in \{\texttt{pass}, \texttt{revise}\}, \tag{5.36}$$

or, more generally, as a structured vector

$$\mathbf{v}_t = \left(v_t^{\text{constraint}}, v_t^{\text{counterfactual}}, v_t^{\text{numeric}}, v_t^{\text{operator}}, v_t^{\text{security}}\right) \tag{5.37}$$

where each component indicates whether a required explanatory element is present and technically coherent. In practice, the Critic checks whether the explanation clearly identifies the driving constraint, includes at least one explicit counterfactual, provides approximate quantitative reasoning, and remains readable from an operational viewpoint. This separation between generation and evaluation stabilizes the explanatory process: the Actor is optimized for producing physically meaningful prose, while the Critic is optimized for checking whether the output actually satisfies the methodological contract imposed by the framework.

The Critic feedback is then processed by a Rule Distiller D, which compresses the observed strengths and deficiencies of the explanation into a small set of reusable rules (See Distiller prompt in Appendix C.3):

$$r_t = D(e_t, \mathbf{v}_t). \tag{5.38}$$

The object r_t should be interpreted as a compact operational refinement signal rather than as free-form commentary. A distilled rule may indicate that future explanations should begin more explicitly from the binding thermal limit, quantify the minimum required relief before discussing alternatives, mention when the cheapest relieving unit is already saturated, or state more clearly that a candidate redispatch worsens the congested line because the relevant sensitivity is positive. For example, the Distiller may extract compact operational rules such as the following, which are then incorporated into the Actor prompt during refinement:

```
1. State the binding network constraint first, with numerical evidence.
2. Include a counterfactual redispatch example with estimated relief.
3. Use sensitivities to justify congestion relief or worsening.
```

In other words, the Distiller transforms explanation-level feedback into instruction-level modifications that can improve future Actor behavior. This step is essential because it decouples local output deficiencies from the long-term evolution of the Actor prompt.

The final component is the Prompt Refiner R (See Refiner prompt in Appendix C.4), which updates the Actor prompt using the distilled rules accumulated over recent snapshots. Given a sliding window of the last K rule sets,

$$W_t = \{r_{t-K+1}, r_{t-K+2}, \ldots, r_t\}, \tag{5.39}$$

the Refiner constructs an updated prompt

$$p_{t+1} = R(p_t, W_t). \tag{5.40}$$

The Refiner is itself governed by a strict design principle: it must improve the Actor prompt while preserving its structure and readability. In the prompt specification, the Refiner is told to "update the Actor prompt using distilled operational rules" and to integrate them "naturally," while ensuring that the original sections on task, inputs, reasoning guidelines, and output remain intact. This means that prompt evolution is controlled rather than unconstrained. The

Refiner is not allowed to inject raw snapshot data, ad hoc examples, or verbose commentary into the Actor prompt; instead, it performs a structured policy update in instruction space. Conceptually, if the Actor prompt is viewed as a parameterized reasoning policy over textual outputs, then the Refiner performs a conservative update

$$p_{t+1} = p_t \oplus \phi(W_t), \tag{5.41}$$

where $\phi(\cdot)$ denotes rule aggregation over the sliding window and $\oplus$ denotes structured incorporation of the resulting guidance into the existing prompt template. The sliding-window mechanism prevents the prompt from overreacting to a single anomalous snapshot while still allowing repeated errors or recurring operational motifs to shape future behavior.

The overall framework therefore defines a closed-loop explanatory system in which physical evidence, language generation, structured critique, rule extraction, and prompt refinement interact over time. For each snapshot, the computational flow is

$$\mathscr{S}_t \xrightarrow{A_\theta} e_t \xrightarrow{C^{\mathrm{AI}}} \mathbf{v}_t \xrightarrow{D} r_t \xrightarrow{R} p_{t+1}. \tag{5.42}$$

The objective is not to train the Actor through gradient-based updates of the underlying language model parameters, but to iteratively refine the Actor prompt so that the resulting explanations become progressively more consistent, physically grounded, and operationally useful. At each snapshot, the Actor produces an explanation based on the OPF solution and its associated physical indicators, including AC power-flow results, redispatch outcomes, dual variables, and PTDF/LODF sensitivities. The Critic then assesses this explanation qualitatively and returns linguistic feedback on its clarity, physical consistency, and operational relevance. This feedback does not constitute a scalar objective in the conventional optimization sense; instead, it acts as a structured reflective signal. The Distiller summarizes recurring Critic observations into compact rules, and after every k snapshots the Refiner incorporates these rules into a revised Actor prompt. Consequently, learning takes place at the prompt level rather than at the parameter level, and improvement emerges from an iterative cycle of explanation, critique, distillation, and refinement. The final output of training is therefore an evolved Actor prompt rather than an updated language model.

The complete procedure is summarized in Algorithm 3. The algorithm makes explicit the roles of memory, sliding-window rule aggregation, iterative evaluation, and prompt selection. In particular, the memory $\mathscr{M}$ stores distilled rules and associated quality scores, allowing the system to retain operational regularities across snapshots while filtering out low-value refinements. The best prompt update is selected according to the evaluation functional, yielding a progressively improved explanatory policy. To support transparency and reproducibility, the prompts used for all agents in the multi-agent LLM framework, together with the final evolved Actor prompt, are provided in Appendix C.

Algorithm 3: Multi-Agent LLM Framework with Sliding-Window Prompt Refinement

Input: Training snapshots $\{\mathscr{S}_t\}_{t=1}^{T}$, initial Actor prompt p_0, memory $\mathscr{M}$, refinement interval K

Output: Refined Actor prompt p_{final}, updated memory $\mathscr{M}$

```
Initialize current prompt p ← p_0;
for t ← 1 to T do
    Generate an operator-oriented explanation using the Actor: e_t ← A_θ(S_t, p);
    Evaluate the explanation using the AI Critic: v_t ← C^AI(S_t, e_t);
    Distill the Critic feedback into a compact operational rule: r_t ← D(e_t, v_t);
    Update memory with the newly distilled rule: M ← M ∪ {r_t};
    if t ≥ K and t mod K = 0 then
        Retrieve the most recent K distilled rules: W_t ← {r_j ∈ M | j = t−K+1, ..., t};
        Refine the Actor prompt using the Refiner: p ← R(p, W_t);
Set final refined prompt p_final ← p;
return p_final, M;
```

In summary, the proposed framework formalizes explanation generation as a physically constrained, multi-agent inference-and-refinement process. The Actor turns AC-OPF and POA outputs into a counterfactual explanation grounded in sensitivities, dual variables, and security indicators; the Critic verifies whether this explanation is complete and technically defensible; the Distiller extracts transferable operational rules from the observed strengths and failures; and the Refiner updates the Actor prompt so that future explanations better align with the needs of control-room reasoning. By explicitly embedding notions such as minimum required relief, sensitivity-based redispatch direction, feasibility screening, approximate cost impact, and qualitative N-1 implications, the framework produces explanations that are not merely linguistically coherent, but operationally interpretable and physically anchored in the underlying AC-OPF solution.

5.4 Experimental Results

5.4.1 Case-study description

To demonstrate the effectiveness of the proposed interpretable AC-OPF redispatch framework, the evaluation was conducted on the IEEE 39-bus benchmark system. All AC-OPF instances were solved using IPOPT [136], a large-scale nonlinear interior-point optimizer, in combination with the MA57 sparse symmetric linear solver for efficient factorization of the associated KKT system.

A total of 500 operating snapshots were generated to train the reflective multi-agent framework. These snapshots were created by introducing uncertainty in both load and renewable generation. The secondary substation load measurements were taken from the Iowa Distribution Test System [114] and used only as reference profiles to inject load uncertainty; the measurements were scaled to match the demand level of the studied grid. The wind power measurements were obtained from wind power plants in France, where the location cannot be disclosed for confidentiality reasons, and were used as reference profiles for renewable generation after replacing 40% of the generators in the grid with wind units. Wind curtailment was allowed as a redispatch action, subject to a maximum curtailment limit of 30% of the available wind power. For the day-ahead stage, the market dispatch was simulated using the point forecast proposed in [122], NWP data generated at 0h00. For the redispatch stage, the point forecast from the very-short-term forecasting model in [122] was used to represent the updated wind conditions closer to real time. To assess generalization performance, an additional set of 100 previously unseen operating snapshots was reserved for testing.

The LLM-based agents were assigned task-specific temperature settings to balance exploration and consistency across the reflective pipeline. The Actor was configured with a relatively higher temperature of 0.7 to encourage diverse yet plausible counterfactual explanation generation. In contrast, the Critic and the Distiller both operated with a low temperature of 0.2 to promote stable, precise, and consistent feedback as well as compact rule extraction. The Refiner used a moderate temperature of 0.4 to integrate distilled rules into the evolving Actor prompt while preserving structure and readability, whereas the Evaluator was assigned a very low temperature of 0.1 to ensure near-deterministic assessment. All experiments were conducted using GPT-5.4 as the underlying large language model.

5.4.2 Illustrative Example

This subsection presents example dialogues generated by the proposed LLM agent to show how it interprets an OPF snapshot using structured inputs from both the OPF solution and the POA. These examples illustrate how the model uses constraint dual variables and sensitivity coefficients to produce physically consistent explanations and counterfactual analyses.

Example Case 1: This case is interesting because the congestion is not confined to a single isolated line but involves multiple related constraints, while the final redispatch is still controlled by one main thermal limit. In particular, line 2 is the most important constraint after redispatch, and the counterfactual tests show that alternative actions that appear cheaper in direct cost terms are rejected because they would worsen the binding network constraint. The N-1 screening further confirms that the affected lines are structurally coupled and subject to significant contingency-driven redistribution. As a result, the case provides a useful illustration of how redispatch decisions are determined not only by production costs, but also by PTDF sensitivities, active thermal constraints, and network security considerations.

Case overview.

```
Before redispatch, the reported thermal violations were on line 2 (bus 1 to bus
2) at 102.60% and line 21 (bus 15 to bus 18) at 106.69%; line 19 (bus 14 to bus
15) was also near binding at 99.33%. These constraints were not isolated. The
PTDF table shows that line 21 and line 19 react identically to injections at
buses 18, 19, 32, and 33, indicating that they belong to the same transfer path
in the area around buses 14-15-18.
```

Summary of redispatch.

```
The realized redispatch actions consisted of a +40 MW increase of conventional
Generator 2 at bus 32, a -20 MW reduction of conventional Generator 6 at bus 36,
a +40.883 MW increase of conventional Generator 8 at bus 38, and a 63.944 MW wind
curtailment of Generator 0 at bus 29. The dominant driver appears to be line 2,
because it is the only line with a nonzero thermal dual (dual = -3396.49), while
the other heavily loaded lines were relieved to below 100% after the redispatch.
```

Question 1. Why not curtail 1 MW less wind from Generator 0 at bus 29 and offset it by 1 MW less conventional generation from Generator 2 at bus 32?

```
On line 2, the PTDF effect is PTDF29 - PTDF32 = 0.6315 - (-0.0039) = +0.6354, so
this substitution worsens the dominant monitored line by about 0.64 MW for each 1
MW of swap. On line 3, the same substitution gives 0.1579 - (-0.1045) = +0.2624,
which increases the loading on that related line from 86.46% to 112.7%.

The direct marginal cost is ΔC = - 30.000 - 47.316 = -57.316 €/MWh. Since ΔC is
negative, this alternative looks cheaper in direct production cost terms because
it avoids wind curtailment and reduces upward redispatch from Gen 2. However, it
is not feasible because it increases loading on the dominant congested line. For
this counterfactual to become viable, line 2 would need at least 0.635 MW of
additional thermal margin for each 1 MW of substitution. For a 10 MW
substitution, at least 6.35 MW of extra margin would be required on line 2.
```

Question 2. Why not replace part of the realized downward redispatch of Generator 6 with additional wind power curtailment at bus 29?

```
A marginal alternative is to reduce wind generation by an additional 1 MW at
Generator 0 on bus 29 while applying 1 MW less downward redispatch to Generator 6
at bus 36, meaning that Gen 6 would increase by 1 MW relative to the realized
operating point.

On line 2, the PTDF change is 0.6315 - 0.5573 = +0.0742 MW, so line 2, which is
already overloaded, would be pushed further above its limit. On line 3, the
corresponding effect is 0.1579 - (-0.7578) = +0.9157 MW, which would increase the
loading on line 3 to 178%.

The marginal economic effect is ΔC = 30.000 - 40.639 = -10.639 €/MWh, so this
alternative again appears cheaper in direct cost terms. However, it is not
selected because it would overload both the dominant line 2 and the related line
3. This illustrates that apparent economic attractiveness is not sufficient when
the alternative aggravates active network constraints.
```

Question 3. Is there any major N-1 cascading risk?

```
For N-1 screening, the outage of line 19 (bus 14 to bus 15) onto line 21 (bus 15
to bus 18) is particularly notable. Before the outage, the flow on line 21 is
approximately 1.0669 × 600 = 640.1 MVA (106.69%), while the flow on line 19 is
approximately 0.9933 × 540 = 536.4 MVA (99.33%).

Using the reported LODF, the change on line 21 is approximately -1.0000 × 536.4 =
-536.4 MVA if monitored under the opposite orientation convention. Equivalently,
this indicates that the outage causes a near-complete redistribution of that tie
and therefore points to a high cascading risk in the 14-15-18 area. This is
consistent with the observation that both lines had to be relieved together in
the redispatch solution.
```

Question 4. Is there any major flow redistribution in case of an outage?

```
A clearer screened redistribution example is the outage of line 2 (bus 1 to bus
2) onto line 1 (bus 0 to bus 38). Before the outage, line 1 carries approximately
0.0941 × 1000 = 94.1 MVA, while line 2 carries approximately 1.0260 × 450 = 461.7
MVA.

Using the reported LODF, the post-contingency flow on line 1 becomes 94.1 +
0.3546 × 461.7 = 257.8 MVA, which corresponds to only 25.8% of the thermal limit
and is therefore not a screened overload.
```

```
However, the outage of line 2 onto line 5 produces 0.3411 × 500 + (-0.5898) ×
461.7 = -101.9 MVA, which implies a reversal in the direction of flow. Although
this does not constitute a screened thermal overload, it demonstrates substantial
contingency-driven redistribution in the same area of the network.
```

Example Case 2: This example is interesting because the overloads are spread across more than one related area, but the final redispatch is still mainly shaped by one key thermal constraint. In particular, line 1-2 remains the main determining constraint after redispatch. The counterfactual questions also show that some alternatives may look cheaper or locally helpful, but they are rejected because they either worsen another overloaded line or are blocked by generator limits. The contingency screening adds another useful point by showing that the redispatch was relevant not only for the base case, but also for reducing screened outage risk.

Case overview.

```
Before redispatch, five thermal overloads were reported: line 1-2 at 143.02%,
line 2-3 at 134.77%, line 5-10 at 121.88%, line 14-15 at 119.26%, and line 15-18
at 107.46%. These were not isolated violations. Line 1-2 and line 2-3 belong to
the same corridor, while line 14-15 and line 15-18 form a related path around bus
15.

After redispatch, all of these lines were brought below 100%, with line 1-2
ending at 93.61%. This line appears to be the main limiting element because it
had the highest overload before redispatch and it was the only line with a
nonzero post-redispatch thermal dual, dual = -6.761713.
```

Summary of redispatch.

```
The realized redispatch actions were: Conventional Generator 8 at bus 38 was
reduced by 20 MW. Wind at bus 31 was curtailed by 74.404 MW, and wind at bus 29
was curtailed by 8.344 MW. Upward redispatch was provided by Generator 6 at bus
36 with 40 MW and Generator 2 at bus 32 with 62 MW. No load curtailment occurred.

The common pattern is a reduction of injections at buses that have positive PTDF
values on line 1-2 and line 2-3. This relieves the overloaded corridor and, at
the same time helps reduce loading on the other related lines.
```

Question 1. Why not do more downward redispatch on Generator 8 at bus 38 instead of 1 MW more wind curtailment at bus 31?

```
The closed marginal swap is +1 MW injection from reducing wind curtailment at bus
31 and -1 MW withdrawal from extra downward redispatch of Generator 8 at bus 38.
```

```
On line 1-2, the PTDF effect is PTDF(31) - PTDF(38) = -0.0125 - 0.3183 = -0.3308,
so this swap would relieve line 1-2 by about 0.3308 MW. On line 2-3, the effect
is 0.0164 - 0.2706 = -0.2542 MW, so it would also relieve line 2-3.

Economically, ΔC = - 30.000 + 32.190 = +2.190 €/MWh, so this alternative is more
expensive. It was also blocked by a local operating limit: Generator 8 was
already providing 20 MW of downward redispatch and had a binding RampDown dual of
-55.811888, which indicates that further downward movement was constrained.

This counterfactual would become feasible if Generator 8 had at least 1 MW of
additional downward ramp margin. It would become economically preferable if the
Gen 8 downward redispatch price fell by at least 2.190 €/MWh, from 32.190 to 30
€/MWh or lower.
```

Question 2. Why not replace 1 MW of wind curtailment at bus 31 with 1 MW of downward redispatch on Generator 7 at bus 37?

```
The closed swap is +1 MW injection at bus 31 and -1 MW withdrawal at bus 37.

On line 1-2, the PTDF effect is PTDF(31) - PTDF(37) = -0.0125 - 0.2787 = -0.2912
MW, so this would relieve the dominant line. On line 2-3, the effect is 0.0164 -
0.4821 = -0.4657 MW, so it would also relieve the same corridor.

However, on line 14-15, the effect is 0.0400 - (-0.3555) = +0.3955 MW, so it
would worsen that already overloaded related line. Economically, ΔC = -30.000 +
21.616 = -8.838 €/MWh, so it would look cheaper. Even so, it is not viable
because it shifts congestion onto line 14-15.

For 1 MW of this swap to become feasible, line 14-15 would need at least 0.3955
MW of additional thermal margin. The realized redispatch is therefore the more
plausible local solution because it reduces injections at buses 29, 31, and 38,
and each of those buses has a positive PTDF on line 1-2 and line 2-3. In that
way, the chosen package consistently unloads the dominant corridor without
requiring emergency load curtailment.
```

Question 3. Is there any major N-1 cascading risk?

```
A clearer screened risk is the outage of line 14-15 onto monitored line 2-17. The
pre-outage flow on line 2-17 was 0.3511 × 500 = 175.6 MVA, while the flow on the
outaged line 14-15 was 1.1926 × 540 = 644.0 MVA.
```

```
With LODF = 0.7823, the screened post-outage flow becomes 175.6 + 0.7823 × 644.0
= 679.4 MVA, which is 135.9% of the 500 MVA limit. This gives first-order
evidence that redispatch was also needed to reduce screened outage risk, not only
to remove the base-case thermal overloads.
```

5.4.3 Quantitative Evaluation

The quality of the generated AC-OPF redispatch explanations was assessed on a test set of previously unseen operating snapshots using a dedicated evaluator agent designed to verify physical consistency, counterfactual validity, and operator usefulness . The purpose of this evaluation stage is not to re-solve the optimization problem or to assess whether the explanation is stylistically plausible in isolation, but rather to determine whether the explanation correctly justifies the redispatch decision using only the information contained in the corresponding power-system snapshot. In particular, the evaluator is explicitly instructed to judge whether the explanation provides a valid counterfactual justification of the OPF outcome, while treating counterfactual scenarios strictly as explanatory devices rather than operational recommendations. This distinction is important because a counterfactual alternative may be physically plausible yet still be inferior to the OPF solution due to infeasibility, higher cost, insufficient congestion relief, or adverse security implications.

For the test stage, each explanation was assessed by the evaluator agent. This evaluator is also an LLM-based agent (See Evaluator prompt in Appendix C.6), but unlike the Actor it is prompted only to assess explanations rather than to generate them. Its prompt is designed to mimic the expectations of a human operator or domain expert by checking whether the explanation contains the elements that an expert would expect to see in a technically sound justification of the OPF result. For each unseen snapshot $\mathscr{S}_n$, with $n = 1, \ldots, N_{\text{test}}$, the Actor generates an explanation e_n, where N_{test} is the number of test snapshots. The evaluator agent then receives the pair $(\mathscr{S}_n, e_n)$ and checks whether the explanation is supported by the available OPF and sensitivity information. Its output is a structured assessment composed of binary answers to 20 verification questions, together with category-level scores, a total score, and an overall percentage score. Each verification question is scored as

$$q_{n,m} \in \{0, 1\}, \qquad m = 1, \ldots, 20, \tag{5.43}$$

where $q_{n,m} = 1$ means that the corresponding claim in explanation e_n is correct and supported by snapshot $\mathscr{S}_n$, and $q_{n,m} = 0$ means that the claim is incorrect, unsupported, or cannot be verified from the available data. This binary design makes the evaluation strict: if an explanation contains a plausible statement that cannot be confirmed from the snapshot, that statement receives zero credit.

The 20 questions are grouped into eight evaluation categories, each associated with an aggregated score $\bar{\sigma}^{(c)}$, $c \in \{1, \dots, 8\}$:

- $\bar{\sigma}^{(1)}$ **— Redispatch trigger identification:** checks whether the explanation identifies all significant overloaded lines or other violated constraints before redispatch, distinguishes the dominant driver, and avoids unsupported claims about unreported constraints or corrective actions.

- $\bar{\sigma}^{(2)}$ **— Operational logic:** checks whether the explanation correctly reflects the sequence from the condition before redispatch to the realized redispatch actions and the condition after redispatch, including the correct description of conventional-generator redispatch, wind curtailment, and load curtailment when present.

- $\bar{\sigma}^{(3)}$ **— PTDF directional reasoning:** checks whether the explanation uses PTDF-based screening correctly, including sign-consistent directional effects and, when relevant, the effect of a redispatch action on more than one important line.

- $\bar{\sigma}^{(4)}$ **— Counterfactual reasoning:** checks whether the explanation provides two explicit counterfactuals with marginal arithmetic, including a specific redispatch alternative, an identified displaced resource, a closed ΔC in €/MWh, an interpretation of the sign of ΔC, and a quantified threshold under which the alternative would become viable or optimal.

- $\bar{\sigma}^{(5)}$ **— Numerical grounding and report discipline:** checks whether the numerical values mentioned in the explanation are supported by the snapshot, whether no unreported values are invented, and whether the explanation remains analytical rather than becoming a long raw-data summary.

- $\bar{\sigma}^{(6)}$ **— Economic and dual interpretation:** checks whether prices, curtailment costs, redispatch prices, and dual variables are used consistently to justify the redispatch decisions and the rejection of the counterfactual alternatives.

- $\bar{\sigma}^{(7)}$ **—** N**-1 security screening:** checks whether any use of LODF-based contingency reasoning is consistent with the available screening information, including one case showing overload risk or movement toward a thermal limit and, when supported, one case showing major flow redistribution or reversal.

- $\bar{\sigma}^{(8)}$ **— Language and operator usefulness:** checks whether the explanation is written in concise, technically clear, operator-oriented language and avoids vague, abstract, or non-standard wording.

For each test snapshot n, the explanation e_n is assigned a percentage score based on the 20 binary verification questions:

$$P_n = 100 \cdot \frac{1}{20} \sum_{m=1}^{20} q_{n,m}, \tag{5.44}$$

where $q_{n,m} \in \{0, 1\}$ is the score of question m for snapshot n, and P_n is the percentage score of the corresponding explanation. The overall test performance is then summarized by the mean of P_n over all N_{test} test snapshots.

Category-level performance is computed in the same way. Let $\sigma_n^{(c)}$ denote the percentage score of category c for snapshot n. Then the average score for category c over the full test set is

$$\bar{\sigma}^{(c)} = \frac{1}{N_{\text{test}}} \sum_{n=1}^{N_{\text{test}}} \sigma_n^{(c)}, \tag{5.45}$$

where $\bar{\sigma}^{(c)}$ is the average performance for category c. These category-level averages help identify where the framework performs well and where it remains weak. For example, a high value of $\bar{\sigma}^{(\text{constraint})}$ together with a lower value of $\bar{\sigma}^{(\text{counterfactual})}$ would indicate that the model usually identifies the main operational bottleneck correctly, but does not always explain alternative redispatch actions in a complete way. These category-level rates provide a fine-grained diagnostic of performance. In particular, they help distinguish between errors arising from unsupported physical claims, incorrect use of sensitivity signs, weak economic justification, or incomplete treatment of contingency effects. Because the evaluator is instructed to assign zero whenever a claim cannot be verified from the snapshot, these accuracies can also be interpreted as a measure of factual discipline in the generated explanations.

The evaluator itself is deliberately conservative. It is instructed to rely only on the provided snapshot, not to solve OPF or run any additional power-flow computation, and not to infer missing facts. As a consequence, the reported scores reflect explanation faithfulness to the available data rather than agreement with external reasoning not present in the snapshot. This makes the evaluation particularly suitable for the present setting, where the objective is to assess whether the multi-agent framework can transform the outputs of AC-OPF and post-optimal analysis into explanations that are physically grounded, numerically defensible, and useful to a system operator. In practical terms, an explanation achieves a high score only if it correctly identifies the active constraint, uses sensitivities coherently, presents at least one explicit and verifiable counterfactual redispatch, avoids unsupported quantitative claims, interprets cost information appropriately, and, when relevant, discusses N-1 effects in a way that is consistent with the LODF data.

This evaluation protocol therefore measures more than textual quality. It directly tests whether the generated explanation satisfies the intended methodological framework: namely, that redispatch explanations should be grounded in the physical snapshot, respect operational limitations, be supported by explicit counterfactual reasoning, and be expressed in a form that is concise and operationally meaningful. Once the evaluator outputs are collected for the full test set, the resulting aggregate metrics ($\bar{P}, \bar{\sigma}^{(c)}$) provide a structured basis for comparing prompt variants, ablation settings, or alternative explanation-generation pipelines.

The effectiveness of the proposed reflective prompt-refinement framework was evaluated on previously unseen test snapshots using the evaluator described. Table 5.1 reports the average category-level scores obtained with the initial Actor prompt and with the evolved prompt after iterative refinement using the multi-agent loop (See the evolved Actor prompt in Appendix C.5). The results show a consistent improvement across all evaluation dimensions, indicating that the reflective framework improves not only the overall quality of the explanations, but also their physical grounding, numerical reliability, and usefulness for operator-facing interpretation.

Table 5.1: *Comparison of the initial and evolved prompts using the category-wise mean scores $\bar{\sigma}^{(c)}$, $c = 1, \ldots, 8$, and the overall percentage score $\bar{P}$.*

Metric ($\bar{P}, \bar{\sigma}^{(c)}$)	Initial prompt	Evolved prompt	Improvement (%)
$\bar{\sigma}^{(1)}$	0.8950	0.9825	9.7765
$\bar{\sigma}^{(2)}$	0.8000	0.9930	24.1250
$\bar{\sigma}^{(3)}$	0.7900	0.9800	24.0506
$\bar{\sigma}^{(4)}$	0.8340	0.9650	15.7074
$\bar{\sigma}^{(5)}$	0.7270	0.8230	13.2050
$\bar{\sigma}^{(6)}$	0.7450	0.9900	32.8859
$\bar{\sigma}^{(7)}$	0.7200	0.9500	31.9444
$\bar{\sigma}^{(8)}$	0.8900	0.9900	11.2360
$\bar{P}$	80.0125	95.9188	19.8797

$\bar{\sigma}^{(1)}$: Redispatch trigger identification,

$\bar{\sigma}^{(2)}$: Operational logic,

$\bar{\sigma}^{(3)}$: PTDF directional reasoning,

$\bar{\sigma}^{(4)}$: Counterfactual reasoning,

$\bar{\sigma}^{(5)}$: Numerical grounding and report discipline,

$\bar{\sigma}^{(6)}$: Economic and dual interpretation,

$\bar{\sigma}^{(7)}$: Security screening,

$\bar{\sigma}^{(8)}$: Language and operator usefulness.

The total evaluation score increased from 80.01% with the initial prompt to 95.92% with the evolved prompt, corresponding to a relative improvement of 19.88%. This gain confirms that the iterative refinement mechanism based on Critic feedback, rule distillation, and sliding-window prompt updates leads to substantially better explanations on unseen operating conditions. Importantly, the improvement is not concentrated in a single dimension, but is distributed across all eight evaluator categories, suggesting that the prompt evolution process improves the overall reasoning structure of the Actor rather than overfitting to a narrow subset of criteria.

The strongest gains were observed in *economic and dual interpretation*, *security screening*, and the two reasoning dimensions tied directly to network sensitivities and operating logic. In particular, the economic-and-dual-interpretation score increased from 0.745 to 0.990, corresponding to the largest relative improvement of 32.89%. This indicates that the evolved prompt more reliably uses prices, curtailment costs, redispatch prices, and dual variables to justify why the realized redispatch is preferable to nearby alternatives. Security screening also improved strongly, rising from 0.720 to 0.950 (+31.94%), which shows that the evolved prompt incorporates contingency-related reasoning and LODF-based screening much more consistently. Similarly, operational logic increased from 0.800 to 0.993 (+24.13%), while *PTDF directional reasoning* increased from 0.790 to 0.980 (+24.05%). These gains indicate that the evolved prompt is substantially better at explaining how redispatch changes the network state and at interpreting directional sensitivity information in a sign-consistent and operationally meaningful way.

A substantial improvement was also observed in *counterfactual reasoning*, which increased from 0.834 to 0.965 (+15.71%). This is particularly important because counterfactual reasoning is central to the explanatory objective of the framework: it requires the model not only to propose plausible alternatives, but also to justify numerically why these alternatives are not selected under the reported operating constraints. *Numerical grounding and report discipline* also improved from 0.727 to 0.823 (+13.20%), indicating that the evolved prompt more consistently grounds its claims in values that are verifiable from the snapshot and avoids unsupported numerical statements. A similarly meaningful gain is observed in *language and operator usefulness*, which rises from 0.890 to 0.990 (+11.24%). This indicates that the evolved explanations are judged to be clearer, more concise, and more aligned with familiar power-system engineering language, which is essential for operational interpretability.

More modest but still clearly positive improvements were obtained in *redispatch trigger identification*. This score increased from 0.895 to 0.9825 (+9.78%). Although this is the smallest relative improvement among the eight categories, it remains meaningful because the baseline performance of the initial prompt was already strong in this dimension. In other words, the initial Actor prompt was already reasonably effective at identifying the main violated or binding constraints driving redispatch, leaving less room for improvement than in weaker categories such as security screening or economic-and-dual interpretation. Nevertheless, the fact that this metric also improved confirms that refinement did not come at the expense of the core explanatory requirement of identifying the operational trigger correctly.

Taken together, these results indicate that the proposed sliding-window reflective refinement mechanism leads to broader and more balanced improvements in explanation quality. The evolved prompt does not merely produce more readable text; it produces explanations that are better aligned with the evaluator's requirements for network-level reasoning, correct interpretation of PTDF-based effects, defensible numerical statements, explicit economic justification, and contingency awareness. The largest improvements occur precisely in the categories that require the strongest discipline and domain grounding, suggesting that the multi-agent framework is effective in steering the LLM away from vague or weakly supported explanations and toward more rigorous, operator-oriented redispatch justifications.

5.5 Final Remarks

This chapter proposed an interpretable multi-agent large language model framework for explaining AC-OPF redispatch decisions in a physically grounded and operator-oriented manner. The framework combines the outputs of AC-OPF and physical operational analysis with a reflective multi-agent reasoning pipeline composed of an Actor, Critic, Distiller, and Refiner. Rather than generating generic textual summaries, the proposed approach constructs explanations that are explicitly tied to the underlying system state, active constraints, dual variables, sensitivity information, and LODF-based security indicators. In this way, the framework transforms redispatch outcomes into concise counterfactual explanations that clarify why a particular redispatch was selected, what constraint forced the redispatch, which alternative actions could have been considered, and why such alternatives were infeasible, more expensive, or less effective.

A key contribution of the proposed method is the introduction of iterative prompt refinement through a sliding-window memory of distilled operational rules. This mechanism allows the explanatory behavior of the Actor agent to evolve over time based on structured critique and recurring patterns observed across operating snapshots. As a result, the framework improves not only the linguistic clarity of the generated explanations, but also their numerical defensibility, physical consistency, and operator relevance. The evaluation on the IEEE 39-bus system demonstrated that the evolved prompt consistently outperformed the initial prompt on unseen test snapshots, with the overall score increasing from 80.01% to 95.92%. The most significant gains were observed in security reasoning, numerical consistency, and economic interpretation, indicating that reflective prompt refinement is particularly effective in strengthening the most demanding aspects of redispatch explanation.

Overall, the results show that large language models can support power-system operation not only as conversational tools, but also as structured explanation engines when tightly constrained by physical data and embedded within a reflective multi-agent architecture. The proposed framework therefore represents a step toward trustworthy and interpretable AI assistance for transmission-system operators, especially in settings where redispatch decisions must be explained in a transparent, technically rigorous, and operationally meaningful way.

Chapter 6

Conclusions

6.1 Final Summary and Conclusions

This thesis advances the state-of-the-art in power system decision-making by addressing three critical dimensions: predictive management under uncertainty, real-time adaptive control, and interpretable optimization reasoning. Across these domains, the research develops novel methodologies that combine human-in-the-loop decision support, RL, and explainable artificial intelligence to enhance both operational efficiency and operator trust.

The first contribution introduces a predictive flexibility procurement framework for distribution networks under forecast uncertainty. By ranking flexibility resources according to effectiveness and dynamically updating decisions based on multiple forecast realizations, the methodology enables operators to make informed choices on the timing and allocation of flexibility. Validation on a public MV network demonstrated that incorporating forecast uncertainty and optimal booking timing can reduce operational costs compared to naive strategies. The study also highlighted practical challenges, including radial network constraints, regulatory gaps for DSOs, and price asymmetry in flexibility markets. Future work may extend the framework by incorporating look-ahead effects, advanced decision quality metrics under uncertainty, extreme weather events, and the further refinement of second-level forecasting for broader applications.

The second contribution develops an interpretable and adaptive real-time control agent based on GNP-DT. By combining rule-based structures with RL, the framework evolves context-aware, state-dependent control strategies for congestion management. Extensive evaluations on the IEEE 118-bus system demonstrated superior performance over baseline ESs, MILP-DC approaches, and deep RL agents, achieving up to 28% improvement in L2RPN scores, higher survival times, and reduced operational costs. Additionally, the framework maintained low inference times, confirming its suitability for real-time deployment. Future research could explore multi-agent coordination, shared memory mechanisms, and scalable applications to larger networks to further enhance adaptability and robustness.

The third contribution presents an interpretable AC-OPF redispatch framework integrating POA with a reflective LLM-based Actor–Critic–Distiller–Refiner architecture. This method generates physically grounded, stepwise explanations of congestion causes, voltage deviations, and redispatch decisions, leveraging sensitivity analysis and counterfactual reasoning. Explanations are fully consistent with AC-OPF outputs, ensuring numerical accuracy, causal fidelity, and operational relevance. The framework provides structured reasoning objects and human-readable explanations for new OPF solutions, and interpretability is evaluated in terms of constraint identification, sensitivity interpretation, counterfactual reasoning, numerical consistency, economic interpretation, security reasoning, and operator usefulness.

Collectively, these contributions bridge the gap between predictive awareness, adaptive control, and interpretable optimization, offering an integrated methodology capable of guiding operators across both predictive and real-time decision horizons. The work demonstrates that combining uncertainty-informed flexibility planning, adaptive rule-based RL, and LLM-supported explainable optimization can significantly improve both the efficiency and transparency of power system operations. This thesis lays the foundation for future research into multi-agent coordination, advanced human-in-the-loop interfaces, and the expansion of interpretable reasoning frameworks to broader operational contexts.

6.2 Future Work

The predictive, adaptive, and interpretable frameworks developed in this thesis open several directions for further research. Future work should aim not only to improve methodological performance, but also to strengthen practical relevance, scalability, and integration into real operational environments.

6.2.1 Predictive Flexibility Management

Future research on predictive flexibility procurement should focus on extending the current methodology beyond single-step decision support toward more anticipatory and risk-aware operational planning. One important direction is the explicit incorporation of look-ahead effects when ranking flexibility options. In particular, the framework could be extended to evaluate the consequences of activating flexibility at future lead times beyond the immediate horizon, thereby capturing longer-term operational benefits or limitations that may influence present decisions. Such an extension would allow the booking strategy to better reflect the temporal interdependence of flexibility actions.

Another relevant avenue concerns the evaluation of decision quality under uncertainty. In the present work, the cost-loss matrix provides a useful performance measure, but it does not explicitly represent the operator's attitude toward risk. Future work could therefore develop richer decision metrics that account for different risk preferences, reliability requirements, or security margins, leading to a more realistic assessment of operational strategies under uncertainty.

The integration of low-probability but high-impact events also deserves further attention. In particular, weather-related disruptions forecasted through NWP ensemble methods could be incorporated into the decision process together with dedicated decision-aid strategies tailored to rare but critical operational situations. This would improve the robustness of flexibility procurement in the presence of severe uncertainty and extreme conditions.

Finally, the original concept of *second-level forecasting* offers substantial room for further development. Future studies may investigate alternative statistical learning or machine learning models, additional explanatory features, and broader forecasting contexts. The concept could also be transferred to other applications, such as deciding whether renewable generation should be committed in the day-ahead market or deferred to intraday market sessions, where similar trade-offs between uncertainty and opportunity arise.

6.2.2 Adaptive Real-Time GNP-DT Control

Future work on the adaptive real-time control framework based on GNP-DT should concentrate on improving coordination, scalability, and adaptability in large-scale power systems. A promising research direction is the extension toward multi-agent settings, in which several GNP-DT agents cooperate by sharing memory, experience, or local operational information in order to improve overall system performance. Such a distributed approach would be particularly relevant for large interconnected grids, where effective coordination across different areas is essential for secure and efficient operation.

Scalability is another central issue. As system size and operational complexity increase, the underlying graph-based representations may need to be refined through node pruning, more efficient graph structures, or hierarchical decision-making mechanisms. These modifications could help preserve interpretability while reducing computational burden and maintaining responsiveness in real-time applications.

The adaptive GNP-DT framework proposed in this thesis is adaptive in the sense that it adjusts its control actions according to the current system state and operational context. However, this should be distinguished from online learning or continual adaptation. Online learning refers to updating the policy or model parameters during deployment as new data becomes available, whereas continual adaptation denotes the broader ability of the system to maintain performance over time under evolving operating conditions, uncertainties, and contingencies. Therefore, a further step would be the introduction of online learning and continual adaptation, allowing the agents to update their decision policies in response to new measurements, operating conditions, and contingencies as they occur. This capability would improve the framework's suitability for dynamic environments in which static policies may become suboptimal over time. In parallel, the integration of risk-aware control policies could strengthen the treatment of uncertainty in renewable generation and demand, thereby connecting the current RL-based framework more closely with robust and stochastic decision-making approaches.

6.2.3 Interpretable AC-OPF and LLM Reasoning

Several research directions can further extend the interpretable AC-OPF and LLM-based reasoning framework developed in this thesis. A first step is the validation of the methodology on larger and more realistic transmission systems with richer operating conditions and more diverse contingency sets. Such an extension would allow a more thorough assessment of the framework's robustness and generalization capability in settings closer to practical system operation.

A second important direction is the inclusion of human-in-the-loop evaluation involving control-room experts. Although the current results demonstrate physical consistency, causal coherence, and readability, direct feedback from practitioners would provide a more rigorous basis for assessing the usefulness, clarity, and trustworthiness of the generated explanations. It would also help identify which explanation formats are most effective in supporting operator decision-making under time pressure.

From a methodological perspective, the prompt-refinement process could be strengthened through more advanced memory management, the use of multiple critic agents, or explicit mechanisms for prioritizing domain-specific operational rules. In addition, the current framework could be extended to multi-period AC-OPF problems, enabling the explanation process to reason over sequences of redispatch decisions rather than isolated operating snapshots.

Looking ahead, embedding the framework in real-time decision-support environments could enable interactive and continuously improving explanation systems for secure and economical power system operation. Such developments would bring the methodology closer to practical deployment and reinforce its role as a human-centered support tool for complex operational decision-making.

Overall, these research directions point toward unified decision-support frameworks that are predictive, adaptive, interpretable, and scalable. They would support both distribution and transmission system operators in real-time and anticipatory operational settings, while also creating opportunities to transfer the developed methodologies to related domains such as power system protection, automated control, and rule-based diagnostics in industrial environments.

Bibliography

[1] H. Holttinen, A. Groom, E. Kennedy, D. Woodfin, L. Barroso, A. Orths, and et al., "Variable renewable energy integration: Status around the world," *IEEE Pow. and Ener. Mag.*, vol. 16, no. 9, pp. 86–96, Nov. 2021.

[2] G. Doorman, G. Thorpe, C. Dornellas, C. Correa, M. Laasri, J. Richstein, and et al., "Short-term flexibility in power systems: drivers and solutions," Cigre, WG C5.27, Tech. Rep. TB 808, 2020.

[3] F. Genoese, R. Kielak, G. Trienekens, K. Poncelet, O. A. Garcia, K. Kukk, and et al., "Roadmap on the evolution of the regulatory framework for distributed flexibility," ENTSO-E, CEDEC, E.DSO, Eurelectric, GEODE, Tech. Rep., 2021.

[4] A. Marot, A. Kelly, M. Naglic, V. Barbesant, J. Cremer, A. Stefanov, and J. Viebahn, "Perspectives on future power system control centers for energy transition," *J. of Mod. Pow. Sys. and Clean Ener.*, vol. 10, no. 2, pp. 328–344, 2022.

[5] O. Valarezo, T. Gómez, J. P. Chaves-Avila, L. Lind, M. Correa, D. U. Ziegler, and R. Escobar, "Analysis of new flexibility market models in europe," *Energies*, vol. 14, no. 12, p. 3521, 2021.

[6] A. Marot, G. Donnot, B.and Dulac-Arnold, A. Kelly, A. O'Sullivan, J. Viebahn, M. Awad, I. Guyon, P. Panciatici, and C. Romero, "Learning to run a power network challenge: a retrospective analysis," *Proceedings of Machine Learning Research (NeurIPS 2020 Competition and Demonstration Track)*, vol. 133, pp. 112–132, 2021.

[7] F. Capitanescu, J. L. M. Ramos, P. Panciatici, D. Kirschen, A. M. 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.

[8] L. A. Roald, D. Pozo, A. Papavasiliou, D. K. Molzahn, J. Kazempour, and A. Conejo, "Power systems optimization under uncertainty: A review of methods and applications," in *PSCC 2022*, Porto, Portugal, 2022.

[9] F. Capitanescu, "Critical review of recent advances and further developments needed in AC optimal power flow," *Elect. Pow. Sys. Res.*, vol. 136, pp. 57–68, Jul. 2016.

[10] I.-I. Avramidis, F. Capitanescu, and G. Deconinck, "Practical approximations and heuristic approaches for managing shiftableloads in the multi-period optimal power flow framework," *Elect. Pow. Sys. Res.*, vol. 190, p. 106864, Jan. 2021.

[11] A. Alqurashi, A. H. Etemadi, and A. Khodaei, "Treatment of uncertainty for next generation power systems: State-of-the-art in stochastic optimization," *Elect. Pow. Sys. Res.*, vol. 141, pp. 233–245, Dec. 2016.

[12] I. A. Hiskens and M. Vrakopoulou, "Model predictive control for power networks," in *Encyclopedia of Systems and Control*, J. Baillieul and T. Samad, Eds. Cham: Springer, 2021, pp. 1234–1239.

[13] V. A. Evangelopoulos, I. I. Avramidis, and P. S. Georgilakis, "Flexibility services management under uncertainties for power distribution systems: Stochastic scheduling and predictive real-time dispatch," *IEEE Access*, vol. 8, pp. 38 855–38 871, 2020.

[14] P. Pan, Z. Wang, G. Chen, H. Shi, and X. Zha, "Multi-time scale optimal dispatch of distribution network with pumped storage station based on model predictive control," *Applied Sciences*, vol. 14, no. 23, p. 11122, 2024.

[15] H. Zhang and P. Li, "Chance constrained programming for optimal power flow under uncertainty," *IEEE Trans. on Power Sys.*, vol. 26, no. 4, pp. 2417–2424, Nov. 2011.

[16] L. Roald, M. Vrakopoulou, F. Oldewurtel, and G. Andersson, "Risk-constrained optimal power flow with probabilistic guarantees," in *PSCC 2014*, Wroclaw, Poland, Aug. 2014.

[17] M. Ni, J. D. McCalley, V. Vittal, and T. Tayyib, "Online risk-based security assessment," *IEEE Trans. on Power Sys.*, vol. 18, no. 1, pp. 258–265, Feb. 2003.

[18] L. Roald, S. Misra, T. Krause, and G. Andersson, "Corrective control to handle forecast uncertainty: A chance constrained optimal power flow," *IEEE Trans. on Power Sys.*, vol. 32, no. 2, pp. 1626–1637, Mar. 2017.

[19] L. Roald and G. Andersson, “Chance-constrained AC optimal power flow: Reformulations and efficient algorithms,” *IEEE Trans. on Power Sys.*, vol. 33, no. 3, pp. 2906–2918, May 2018.

[20] T. Soares and R. J. Bessa, “Proactive management of distribution grids with chance-constrained linearized AC OPF,” *Int. J. of Elect. Pow. & Ener. Sys.*, vol. 109, pp. 332–342, Jul. 2019.

[21] I. Mezghani, S. Misra, and D. Deka, “Stochastic AC optimal power flow: A data-driven approach,” *Elect. Pow. Sys. Res.*, vol. 189, p. 106567, Dec. 2020.

[22] T. Soares, R. J. Bessa, P. Pinson, and H. Morais, “Active distribution grid management based on robust AC optimal power flow,” *IEEE Trans. on Smart Grid*, vol. 9, no. 6, pp. 6229–6241, Nov. 2018.

[23] A. Attarha, N. Amjady, and A. J. Conejo, “Adaptive robust ac optimal power flow considering load and wind power uncertainties,” *Int. J. of Elect. Pow. & Ener. Sys.*, vol. 96, pp. 132–142, Mar. 2018.

[24] D. Lee, K. Turitsyn, D. K. Molzahn, and L. A. Roald, “Robust AC optimal power flow with robust convex restriction,” *IEEE Trans. on Power Sys.*, vol. 36, no. 6, pp. 4953–4966, Nov. 2021.

[25] P. Gärdenfors, “Forecasts, decisions and uncertain probabilities,” *Erkenntnis*, vol. 14, no. 2, pp. 159–181, 1979.

[26] Y. Guo, K. Baker, E. Dall’Anese, Z. Hu, and T. H. Summers, “Data-based distributionally robust stochastic optimal power flow – Part i: Methodologies,” *IEEE Trans. on Power Sys.*, vol. 34, no. 2, pp. 1483–1492, Mar. 2019.

[27] Q. Gemine, D. Ernst, and B. Cornélusse, “Active network management for electrical distribution systems: problem formulation, benchmark, and approximate solution,” *Opt. and Eng.*, vol. 18, pp. 587–629, 2017.

[28] A. Marot, B. Donnot, G. Dulac-Arnold, A. Kelly, A. O’Sullivan, J. Viebahn, M. Awad, I. Guyon, P. Panciatici, and C. Romero, “Learning to run a power network challenge: a retrospective analysis,” in *Proc. of Mach. Learn. Res.*, vol. 133, 2021, pp. 112–132.

[29] E. D. Regnier, “What is six hours worth? the impact of lead time on tropical-storm preparation decisions,” *Dec. Anal.*, vol. 17, no. 1, pp. 9–23, mar 2020.

[30] C. Wanke and D. Greenbaum, “Sequential congestion management with weather fore-

cast uncertainty," in *AIAA Guidance, Navigation and Control Conference and Exhibit*, jun 2008.

[31] S. Jewson, S. Scher, and G. Messori, "Decide now or wait for the next forecast? testing a decision framework using real forecasts and observations," *Mont. Weat. Rev.*, mar 2021.

[32] ——, "Communicating properties of changes in lagged weather forecasts," *Weath. and Forec.*, vol. 37, no. 1, pp. 125–142, jan 2022.

[33] P. Tankov and L. Tinsi, "Stochastic optimization with dynamic probabilistic forecasts," *Ann. of Oper. Res.*, aug 2022.

[34] C. Crampes and J. Renault, "How many markets for wholesale electricity when supply ispartially flexible?" *Ener. Econ.*, vol. 81, pp. 465–478, jun 2019.

[35] J. Bellenbaum and J. C. Weber, "Designing flexibility procurement markets for congestion management – investigating two-stage procurement auctions," *Ener. Econ.*, vol. 106, p. 105775, Feb. 2022.

[36] T. Muhlpfordt, V. Hagenmeyer, and T. Faulwasser, "The price of uncertainty: Chance-constrained OPF vs. in-hindsight OPF," in *PSCC 2018*, Dublin, Ireland, Jun. 2018.

[37] R. Li, H. Zhang, M. Sun, F. Teng, C. Wan, S. Pineda, and G. Kariniotakis, "Decision-oriented learning for future power system decision-making under uncertainty," *arXiv preprint arXiv:2401.03680*, 2024.

[38] Y. Lim, M. Son, K. Park, M. Kim, K. Song, H. Lee, and H. Kim, "Power system decision making in the age of deep learning: A comprehensive review," *Energies*, vol. 18, no. 18, p. 4867, 2025.

[39] M. Mazi *et al.*, "Correction of overloads and voltage violations by corrective control," *IEEE Transactions on Power Systems*, vol. 1, no. 2, pp. 73–80, 1986.

[40] R. Bacher and H. Glavitsch, "Network topology optimization with security constraints," *IEEE Transactions on Power Systems*, vol. 1, no. 4, pp. 103–111, 1986.

[41] A. G. Bakirtzis and A. P. S. Meliopoulos, "Incorporation of switching operations in power system corrective control computations," *IEEE Transactions on Power Systems*, vol. 2, no. 3, pp. 669–675, 1987.

[42] I. Makram *et al.*, "Optimization of transmission line switching using the z-matrix method," *IEEE Transactions on Power Systems*, vol. 4, no. 1, pp. 259–264, 1989.

[43] S. Chen and H. Glavitsch, "Stabilizing switching," *IEEE Transactions on Power Systems*, vol. 8, no. 4, pp. 1511–1517, 1993.

[44] W. Shao and V. Vittal, "Corrective switching algorithm for relieving overloads and voltage violations," *IEEE Transactions on Power Systems*, vol. 20, no. 4, pp. 1877–1885, 2005. [Online]. Available: https://ieeexplore.ieee.org/document/1525118

[45] ——, "Bip-based opf for line and bus-bar switching to relieve overloads and voltage violations," in *Proceedings of the IEEE Power Systems Conference and Expo*, 2006. [Online]. Available: https://ieeexplore.ieee.org/document/4076059

[46] G. Granelli *et al.*, "Optimal network reconfiguration for congestion management by deterministic and genetic algorithms," *Electric Power Systems Research*, vol. 76, no. 6-7, pp. 549–556, 2006.

[47] S. Fliscounakis *et al.*, "Topology influence on loss reduction as a mixed integer linear programming problem," *IEEE Power Tech*, pp. 1987–1990, 2007. [Online]. Available: https://ieeexplore.ieee.org/document/4538622

[48] E. B. Fisher *et al.*, "Optimal transmission switching," *IEEE Transactions on Power Systems*, vol. 23, no. 3, pp. 1346–1355, 2008.

[49] K. W. Hedman *et al.*, "Optimal transmission switching with contingency analysis," *IEEE Transactions on Power Systems*, vol. 24, no. 3, pp. 1577–1586, 2009.

[50] J. D. Fuller, R. Ramasra, and A. Cha, "Fast heuristics for transmission-line switching," *IEEE Transactions on Power Systems*, vol. 27, no. 3, pp. 1377–1386, 2012.

[51] K. W. Hedman *et al.*, "Revenue adequacy constrained optimal transmission switching," *IEEE Power and Energy Society General Meeting*, pp. 1–7, 2011. [Online]. Available: https://ieeexplore.ieee.org/document/5718670

[52] J. Han and A. Papavasiliou, "Congestion management through topological corrections: A case study of central western europe," *Energy Policy*, vol. 86, pp. 470–482, 2015. [Online]. Available: https://doi.org/10.1016/j.enpol.2015.07.031

[53] M. Heidarifar and H. Ghasemi, "A network topology optimization model based on substation and node-breaker modeling," *IEEE Transactions on Power Systems*, vol. 31, no. 1, pp. 247–255, 2016. [Online]. Available: https://ieeexplore.ieee.org/document/7038226

[54] M. Heidarifar *et al.*, "An optimal transmission line switching and bus splitting heuristic

incorporating ac and n-1 contingency constraints," *International Journal of Electrical Power and Energy Systems*, vol. 133, p. 107278, 2021.

[55] Y. Zhou *et al.*, "Substation-level grid topology optimization using bus splitting," *arXiv*, 2020.

[56] Y. Zhang and Y. Liu, "Multi-period optimal transmission switching with voltage stability and security constraints," *Sustainability*, vol. 16, no. 18, p. 8272, 2024.

[57] M. A. Tavakkoli and N. Amjady, "Incorporating bus-bar switching actions into ac optimal power flow to avoid over-current status," *Scientia Iranica*, vol. 26, no. 1, pp. 206–213, 2019.

[58] Z. Pei, A. M. Rojas-Arevalo, F. J. de Haan, N. Lipovetzky, and E. A. Moallemi, "Reinforcement learning for decision-making under deep uncertainty," *Journal of Environmental Management*, vol. 359, p. 120968, 2024.

[59] S. Ramstedt and C. J. Pal, "Real-time reinforcement learning," *ArXiv*, vol. abs/1911.04448, 2019. [Online]. Available: https://api.semanticscholar.org/CorpusID:202767718

[60] M. Subramanian, J. Viebahn, S. Tindemans, B. Donnot, and A. Marot, "Exploring grid topology reconfiguration using a simple deep reinforcement learning approach," *2021 IEEE Madrid PowerTech*, pp. 1–6, 2020. [Online]. Available: https://api.semanticscholar.org/CorpusID:227209525

[61] J. Duan, H. Li, X. Zhang, R. Diao, B. Zhang, D. Shi, X. Lu, Z. Wang, and S. Wang, "A deep reinforcement learning based approach for optimal active power dispatch," *arXiv preprint arXiv:1908.11543*, 2019. [Online]. Available: https://arxiv.org/abs/1908.11543

[62] Y. Zhou, W. Lee, R. Diao, and D. Shi, "A data-driven method for fast ac optimal power flow solutions via deep reinforcement learning," *Journal of Modern Power Systems and Clean Energy*, vol. 8, no. 6, pp. 1128–1139, 2020. [Online]. Available: https://epjournal.csee.org.cn/mpce/en/article/id/57ac36c2-3241-4c0e-8c67-304f7e7a29c2

[63] ——, "Deep reinforcement learning based real-time ac optimal power flow considering uncertainties," *Journal of Modern Power Systems and Clean Energy*, vol. 10, no. 5, pp. 1098–1109, 2021. [Online]. Available: https://www.mpce.info/mpce/article/abstract/202205003

[64] X. Wang, H. Zhong, G. Zhang, G. Ruan, Y. He, and Z. Yu, "Look-ahead ac optimal

power flow: A model-informed reinforcement learning approach," *arXiv preprint arXiv:2303.02306*, 2023. [Online]. Available: https://doi.org/10.48550/arXiv.2303.02306

[65] T. Wu, A. Scaglione, and D. Arnold, "Constrained reinforcement learning for predictive control in real-time stochastic dynamic optimal power flow," *arXiv preprint arXiv:2302.10382*, 2023.

[66] M. U. Awais, "Using deep reinforcement learning to solve optimal power flow problem with generator failures," *arXiv preprint arXiv:2205.02108*, 2022.

[67] J. Li, R. Zhang, H. Wang, Z. Liu, H. Lai, and Y. Zhang, "Deep reinforcement learning for optimal power flow with renewables using graph information," *arXiv preprint arXiv:2112.11461*, 2021.

[68] A. Marot, B. Donon, I. Guyon, and B. Donnot, "Learning To Run a Power Network Competition," in *CiML Workshop, NeurIPS*, Montréal, Canada, Dec. 2018. [Online]. Available: https://hal.science/hal-01968295

[69] B. Donnot, "Grid2op- a testbed platform to model sequential decision making in power systems." 2020. [Online]. Available: https://GitHub.com/Grid2Op/grid2op

[70] L. Omnes, A. Marot, and B. Donnot, "Adversarial training for a continuous robustness control problem in power systems," *2021 IEEE Madrid PowerTech*, pp. 1–6, 2020.

[71] A. Marot, I. M. Guyon, B. Donnot, G. Dulac-Arnold, P. Panciatici, M. Awad, A. O'Sullivan, A. Kelly, and Z. Hampel-Arias, "L2rpn: Learning to run a power network in a sustainable world neurips2020 challenge design," 2020. [Online]. Available: https://api.semanticscholar.org/CorpusID:243830891

[72] B. Zhou, H. Zeng, Y. Liu, K. Li, F. Wang, and H. Tian, "Action set based policy optimization for safe power grid management," in *ECML/PKDD*, 2021.

[73] A. Chauhan, M. Baranwal, and A. Basumatary, "Powrl: A reinforcement learning framework for robust management of power networks," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 37, no. 12, 2023, pp. 14 757–14 764.

[74] M. Lehna, J. Viebahn, A. Marot, S. Tomforde, and C. Scholz, "Managing power grids through topology actions: A comparative study between advanced rule-based and reinforcement learning agents," *Energy and AI*, vol. 14, p. 100276, 2023.

[75] B. Fritz and W. Bukhsh, "Explainable DC optimal power flow decisions," in *2025 IEEE*

Kiel PowerTech, 2025, pp. 1–6.

[76] S. Pineda, J. Pérez-Ruiz, and J. M. Morales, "Beyond the neural fog: Interpretable learning for ac optimal power flow," *IEEE Transactions on Power Systems*, vol. 40, no. 6, pp. 4912–4921, 2025.

[77] M. Goerigk, M. Hartisch, S. Merten, and K. Sharma, "Feature-based interpretable surrogates for optimization," *arXiv preprint arXiv:2409.01869*, 2024.

[78] M. Li, Y. Du, J. Mohammadi, C. Crozier, K. Baker, and S. Kar, "Numerical comparisons of linear power flow approximations: Optimality, feasibility, and computation time," in *2022 IEEE Power & Energy Society General Meeting (PESGM)*, 2022, pp. 1–5.

[79] J. Kurtz, S. İ. Birbil, and D. den Hertog, "Counterfactual explanations for linear optimization," *arXiv preprint arXiv:2405.15431*, 2024.

[80] B. Biemans, P. Troubil, I. Grau, and W. P. M. Nuijten, "Explainable optimization: Leveraging large language models for user-friendly explanations," in *Explainable Artificial Intelligence*, ser. Communications in Computer and Information Science. Springer, 2025, vol. 2578, pp. 44–67.

[81] F. Da Ros, M. Soprano, L. Di Gaspero, and K. Roitero, "Large language models for combinatorial optimization: A systematic review," *arXiv preprint arXiv:2507.03637*, 2025.

[82] H. Chen, G. E. Constante-Flores, K. S. I. Mantri, S. M. Kompalli, A. S. Ahluwalia, and C. Li, "Optichat: Bridging optimization models and practitioners with large language models," *arXiv preprint arXiv:2501.08406*, 2025.

[83] H. Chen, G. E. Constante-Flores, and C. Li, "Diagnosing infeasible optimization problems using large language models," *INFOR: Information Systems and Operational Research*, vol. 62, no. 4, pp. 573–587, 2024.

[84] N. Shinn, F. Cassano, S. Koyejo, and N. Goodman, "Reflexion: Language agents with verbal reinforcement learning," *arXiv preprint arXiv:2303.11366*, 2023.

[85] S. Zhou, K. Chen, Y. Chen, and W. Wang, "Language agent tree search unifies reasoning, acting, and planning in language models," *arXiv preprint arXiv:2308.08155*, 2023.

[86] Y. Xie, C. Zhang, and L. Gao, "Self-evaluation guided beam search for reasoning in large language models," *arXiv preprint arXiv:2305.00633*, 2023.

[87] F. Régin, E. D. Maria, and A. Bonlarron, "Combining constraint programming reasoning

with large language model predictions," in *30th International Conference on Principles and Practice of Constraint Programming (CP 2024)*, ser. Leibniz International Proceedings in Informatics (LIPIcs), vol. 307. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2024, pp. 25:1–25:18.

[88] Q. Wang, Z. Wang, Y. Su, H. Tong, and Y. Song, "Rethinking the bounds of llm reasoning: Are multi-agent discussions the key?" in *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. Bangkok, Thailand: Association for Computational Linguistics, 2024, pp. 6106–6131.

[89] C. Li, A. Kies, K. Zhou, M. Schlott, O. E. Sayed, M. Bilousova, and H. Stöcker, "Optimal power flow in a highly renewable power system based on attention neural networks," *Applied Energy*, vol. 359, p. 122779, 2024.

[90] M. Hasquenoph, S. Hourrig, and O. Carré, "Active power curtailment for MV network operational planning in an industrial environment," *CIRED - Open Access Proc. J.*, vol. 1, pp. 1554–1557, Oct. 2017.

[91] E. Memmel, S. Schluters, R. Volker, F. Schuldt, K. V. Maydell, and C. Agert, "Forecast of renewable curtailment in distribution grids considering uncertainties," *IEEE Access*, vol. 9, pp. 60 828–60 840, 2021.

[92] S. Mabu, K. Hirasawa, M. Obayashi, and T. Kuremoto, "Enhanced decision making mechanism of rule-based genetic network programming for creating stock trading signals," *Expert Systems with Applications*, vol. 40, no. 16, pp. 6311–6320, 2013.

[93] A. Marot, B. Donnot, S. Tazi, and P. Panciatici, "Expert system for topological remedial action discovery in smart grids," in *Mediterranean Conference on Power Generation, Transmission, Distribution and Energy Conversion (MEDPOWER 2018)*, 2018, pp. 1–6.

[94] S. Chondrogiannis, J.Vasiljevska, A. Marinopoulos, I. Papaioannou, and G. Flego, "Local electricity flexibility markets in europe," Joint Research Centre, Tech. Rep. JRC130070, Oct. 2022.

[95] I.-I. Avramidis, F. Capitanescu, and G. Deconinck, "Thoughts on conceiving a "fair" value for flexibility," *Int. J. of Elect. Pow. & Ener. Sys.*, vol. 147, p. 108893, May 2023.

[96] Y. Ruwaida, J. P. Chaves-Avila, N. Etherden, I. Gomez-Arriola, G. Gürses-Tran, K. Kessels, and et al., "TSO-DSO-customer coordination for purchasing flexibility system services: Challenges and lessons learned from a demonstration in Sweden," *IEEE Trans. on Power Sys.*, In Press, 2022.

[97] P. Girvan, A. Stolte, and M. Mann, "Conceptualising different approaches to the new redispatch procedure in germany with special regard to the distribution of competences among the grid operators," in *IEEE Pow. and Ener. Student Summit*, 2020.

[98] C.-L. Hwang and K. Yoon, *Methods for Multiple Attribute Decision Making*. Berlin, Heidelberg: Springer Berlin Heidelberg, 1981, pp. 58–191.

[99] B. Hill, "Confidence and decision," *Games and Econ. Beha.*, vol. 82, pp. 675–692, 2013.

[100] K. Alzaareer, M. Saad, H. Mehrjerdi, H. Al-Masri, A. Q. Al-Shetwi, D. Asber, and et al., "New voltage sensitivity analysis for smart distribution grids using analytical derivation: Abcd model," *Int. J. of Elect. Pow. & Ener. Sys.*, vol. 137, p. 107799, 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0142061521010176

[101] H.-Y. Jeong, J.-C. Choi, D.-J. Won, S.-J. Ahn, and S.-i. Moon, "Formulation and analysis of an approximate expression for voltage sensitivity in radial dc distribution systems," *Energies*, vol. 8, no. 9, pp. 9296–9319, 2015. [Online]. Available: https://www.mdpi.com/1996-1073/8/9/9296

[102] K. Christakou, J.-Y. LeBoudec, M. Paolone, and D.-C. Tomozei, "Efficient computation of sensitivity coefficients of node voltages and line currents in unbalanced radial electrical distribution networks," *IEEE Trans. on Smart Grid*, vol. 4, no. 2, pp. 741–750, 2013.

[103] J. Grainger and W. Stevenson, *Power System Analysis*, ser. Electrical engineering series. McGraw-Hill, 1994.

[104] L. Singh, *Advanced Power System Analysis and Dynamics*. New Age International, 2006. [Online]. Available: https://books.google.pt/books?id=C_HPBhmOX7kC

[105] G. Papaefthymiou and D. Kurowicka, "Using copulas for modeling stochastic dependence in power system uncertainty analysis," *IEEE Trans. on Power Sys.*, vol. 24, no. 1, pp. 40–49, Feb. 2009.

[106] J. Naughton, H. Wang, M. Cantoni, and P. Mancarella, "Co-optimizing virtual power plant services under uncertainty: A robust scheduling and receding horizon dispatch approach," *IEEE Trans. on Power Sys.*, vol. 36, no. 5, pp. 3960–3972, Sep. 2021.

[107] J. B. and, "Market-based redispatch in the distribution grid. why it works." DNV GL Energy, Tech. Rep. 2020-0977, Sep. 2020.

[108] J. R. Andrade and R. J. Bessa, "Improving renewable energy forecasting with a grid of numerical weather predictions," *IEEE Trans. on Sust. Ener.*, vol. 8, no. 4, pp. 1571–1580,

2017.

[109] C. Gonçalves, L. Cavalcante, M. Brito, R. J. Bessa, and J. Gama, "Forecasting conditional extreme quantiles for wind energy," *Elect. Pow. Sys. Res.*, vol. 190, p. 106636, Jan. 2021.

[110] K. Cho, B. Van Merriënboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio, "Learning phrase representations using rnn encoder-decoder for statistical machine translation," *arXiv:1406.1078*, 2014.

[111] P. Pinson and G. Kariniotakis, "On-line assessment of prediction risk for wind power production forecasts," *Wind Energy*, vol. 7, no. 2, pp. 119–132, Apr. 2004.

[112] Y. jun Zhang and Z. Ren, "Optimal reactive power dispatch considering costs of adjusting the control devices," *IEEE Trans. on Power Sys.*, vol. 20, no. 3, pp. 1349–1356, Aug. 2005.

[113] P. C. Fishburn, "Methods of estimating additive utilities," *Manag. Sci.*, vol. 13, no. 7, pp. 435–453, Mar. 1967.

[114] Z. Wang, "Iowa distribution test systems," December 2022. [Online]. Available: http://wzy.ece.iastate.edu/Testsystem.html

[115] N. Marinho and Y. Phulpin, "Economics of distributed flexibilities for the grid considering local short-term uncertainties," in *2015 IEEE Eindhoven PowerTech*, Eindhoven, Netherlands, 2015.

[116] Y. Zhu, Z. Toth, R. Wobus, D. Richardson, and K. Mylne, "The economic value of ensemble-based weather forecasts," *Bull. of the Amer. Meteo. Soc.*, vol. 83, no. 1, pp. 73–84, Jul. 2002.

[117] A. Longo, T. Bouman, S. Giaccaria, and T. Efthimiadis, "Societal appreciation of energy security: Volume 1: Value of lost load ? residential consumers (EE, NL and PT)," European Commission, Luxembourg (Luxembourg), Tech. Rep., 2018.

[118] D. Yoon, S. Hong, B.-J. Lee, and K.-E. Kim, "Winning the l2{rpn} challenge: Power grid management via semi-markov afterstate actor-critic," in *International Conference on Learning Representations*, 2021. [Online]. Available: https://openreview.net/forum?id=LmUJqB1Cz8

[119] M. L. Puterman, *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. Hoboken, NJ: John Wiley & Sons, 2005.

[120] Île-de-France Region and RTE, "Paris region ai challenge for energy transition: Low-carbon grid operations," 2023, accessed: 2025-10-03. [Online]. Available: https://www.iledefrance.fr/sites/default/files/medias/2023/05/Description_Challenge_RTE.pdf

[121] R. S. Sutton and A. G. Barto, "Reinforcement learning: An introduction." MIT press, 2018. [Online]. Available: http://incompleteideas.net/book/the-book-2nd.html

[122] R. J. Bessa, F. Moaidi, J. Viana, and J. R. Andrade, "Uncertainty-aware procurement of flexibilities for electrical grid operational planning," *IEEE Transactions on Sustainable Energy*, vol. 15, no. 2, pp. 789–802, 2024.

[123] C. Binbin, "Teacher-tutor-junior student-senior student," GitHub, Jun. 2025. [Online]. Available: https://github.com/Aspirin96/L2RPN_NIPS_2020_a_PPO_Solution

[124] Grid2Op Documentation, "Available environments — grid2op 1.10.4 documentation," https://grid2op.readthedocs.io/en/latest/available_envs.html, 2025, accessed: 2025-05-28.

[125] J. Paulos, P. Silva, R. Bessa, A. Marot, J. Dejaegher, and B. Donnot, "Generation of power network operating scenarios for an AI-friendly digital environment," in *IEEE PowerTech 2025 Conference*, Kiel, Germany, 29 June–3 July 2025.

[126] Grid2op Contributors, "grid2op-milp-agent: A milp-based grid topology control agent for the grid2op environment," https://github.com/Grid2op/grid2op-milp-agent, 2025, accessed: 2025-06-07.

[127] J. Forrest *et al.*, "Cbc (coin-or branch and cut)," https://github.com/coin-or/Cbc, 2020, version 2.10.5, COIN-OR.

[128] J. Sintes, "How we built the winning real time autonomous agent for power grid management in the L2RPN challenge 2023," Medium, Mar. 2024. [Online]. Available: https://medium.com/@lajavaness/how-we-built-the-winning-real-time-autonomous-agent-for-power-grid-management-in-the-l2rpn-41ab3

[129] 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.

[130] Agency for the Cooperation of Energy Regulators (ACER), "Redispatching and countertrading," 2026, accessed: 2026-04-18. [Online]. Available: https://www.acer.europa.eu/electricity/market-rules/capacity-allocation-and-congestion-management/redispatching-and-countertrading

[131] EPEX SPOT, “Trading products,” 2026, accessed: 2026-04-18. [Online]. Available: https://www.epexspot.com/en/tradingproducts

[132] R. D. Zimmerman, C. E. Murillo-Sánchez, and R. J. Thomas, “MATPOWER: Steady-state operations, planning, and analysis tools for power systems research and education,” *IEEE Transactions on Power Systems*, vol. 26, no. 1, pp. 12–19, 2011.

[133] B. Stott, J. Jardim, and O. Alsac, “DC Power Flow Revisited,” *IEEE Transactions on Power Systems*, vol. 24, no. 3, pp. 1290–1300, 2009.

[134] T. Güler, G. Gross, and M. Liu, “Generalized line outage distribution factors,” *IEEE Transactions on Power Systems*, vol. 22, no. 2, pp. 879–881, 2007.

[135] J. Guo, Y. Fu, Z. Li, and M. Shahidehpour, “Direct calculation of line outage distribution factors,” *IEEE Transactions on Power Systems*, vol. 24, no. 3, pp. 1633–1634, 2009.

[136] 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.

Appendix A

Chapter 3 Annex

A.1 Flexibility Cost

The cost of system operator flexibility assets is computed per switching action as follows [112]:

$$C_{s/b} = \frac{1}{T_T}(F_{s/b} + \frac{F_{OT} \cdot a_T^{'}}{t_{OT}}) \tag{A.1}$$

where T_T, $a_T^{'}$ and t_{OT} are the total allowable adjustment times, lifetime after step changed T_T times, and maintenance period, respectively. The maintenance cost and capital cost of the element (e.g., OLTC, power line, capacitor bank) is denoted by F_{OT} and $F_{s/b}$. Note that for the OLTC, the $F_{s/b}$ is equal to $F_{T_{OLTC}} \cdot (a_T - a_T^{'})/a_T$, where a_T is lifetime when the tap is never adjusted and $F_{T_{OLTC}}$ is the capital cost of the transformer with OLTC. Table A.1 presents the parameters of the flexibility cost formula per each element.

Table A.1: *Parameters for calculating flexibility cost of SO assets*

Parameter	Shunt (CB/RS)	OLTC	Power lines
a_T (year)	-	45	-
$a_T^{'}$ (year)	20	35	20
T_T (times)	14600	89425	7300
t_{OT} (year)	4	17.5	2
F_{OT} (€/times)	600	300	60
$F_{s/b}$ (€/MVar)	7500	11250	15300

A.2 Sensitivity indices in Electrical Networks without/with PV Buses

In the Y_{bus} compound matrix approach [102], for an unbalanced network with $M = SB + N$ buses (SB slack buses and N buses with PQ injections that are constant and independent of the voltage), sensitivity coefficients of voltage magnitudes in all PQ buses and currents in all lines versus active (reactive) power variations, are defined by

$$\frac{\partial|\bar{\kappa}|}{\partial P_l(Q_l)} = \frac{1}{|\bar{\kappa}|} Re(\underline{\kappa}\frac{\partial\bar{\kappa}}{\partial P_l(Q_l)}), \qquad \kappa = E_i \text{ or } I_{ij} \tag{A.2}$$

In above: the bar over and under quantities indicates that they are, in respective, complex and complex conjugates [1]; $\bar{E}_i$ ($\bar{I}_{ij}$) is the voltage (current) at bus i (between buses i and j); P_l (Q_l) is the active (reactive) power at bus l. After computing the partial derivative of the well-known relation between the complex power (the line current) and voltage at all non-PV buses over the active/reactive power at each PQ bus and then rewriting all phasors in the rectangular coordinates, the SIs of voltage magnitudes (currents) are calculated through the following algorithm [102]: a) building the matrix of the linear system of equations; b) inverting matrix of size $2N \times 2N$; c) doing N multiplications of the inverse matrix with vectors of size $2N \times 1$.

By assuming that transformers tap-changers are located in correspondence with the slack buses of the network and the transformers voltage variations due to tap position changes are small enough, the Y_{bus} compound matrix approach has been employed to compute the voltage sensitivity to tap-changer position through following definition [102]:

$$\frac{\partial\left|\bar{E}_i\right|}{\partial\left|\bar{E}_k\right|} = \left|\bar{E}_i\right| Re(\frac{1}{\bar{E}_i}\frac{\partial\bar{E}_i}{\partial\left|\bar{E}_k\right|}) \tag{A.3}$$

In our calculations, the PV2PQ conversion approach is considered in two following stages. First, power flow (PF) problem is solved with the Newton-Raphson method to find voltages. Then PV buses are considered equal to PQ buses by: creating SGENs with active and reactive powers equal to those of generators of the network; making the generators out of service; obtaining PF solution based on previous results through initializing the PF solver by voltages of the PF results in the first stage rather than a flat initializing. Therefore, by starting from the well-known relationship between the complex power vector, $\bar{\boldsymbol{S}}$, and bus voltage vector, $\bar{\boldsymbol{E}}$ (with the size of N, the number of non-slack buses):

$$\underline{\boldsymbol{S}} = \underline{\boldsymbol{E}}_D \cdot \bar{\boldsymbol{Y}}' \cdot \bar{\boldsymbol{E}}', \tag{A.4}$$

where $\bar{\boldsymbol{Y}}'$ is the $N \times M$ admittance compound matrix (without rows corresponding to slack buses), $\bar{\boldsymbol{E}}'$ voltage vector of size M, and $\underline{\boldsymbol{E}}_D$ is an $N \times N$ matrix that the subscript D indicates

[1] These notations are kept in the appendix

a diagonal matrix of which the diagonal element ii are the i-th element of $\underline{\boldsymbol{E}}$ (we hereafter keep this notation). By recalling $\bar{\mathbf{S}} = \mathbf{P} + \mathrm{j}\mathbf{Q}$, with $\mathbf{P}$ ($\mathbf{Q}$) as active (reactive) power vector with the size of N, and $\bar{\boldsymbol{Y}}' = \boldsymbol{G}' + \mathrm{j}\boldsymbol{B}'$, Eq. (A.4) can be rewritten in the rectangular coordinates as follows:

$$\mathbf{P} - j\mathbf{Q} = (\boldsymbol{E}_D^d \cdot \boldsymbol{\alpha} + \boldsymbol{E}_D^q \cdot \boldsymbol{\beta}) - j(\boldsymbol{E}_D^q \cdot \boldsymbol{\alpha} - \boldsymbol{E}_D^d \cdot \boldsymbol{\beta}) \tag{A.5}$$

where $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$, vectors with size of N, are given by

$$\boldsymbol{\alpha} = \boldsymbol{G}' \cdot \boldsymbol{E}'^d - \boldsymbol{B}' \cdot \boldsymbol{E}'^q, \qquad \boldsymbol{\beta} = \boldsymbol{G}' \cdot \boldsymbol{E}'^q + \boldsymbol{B}' \cdot \boldsymbol{E}'^d, \tag{A.6}$$

with $\boldsymbol{E}'^d$ and $\boldsymbol{E}'^q$, the real and imaginary parts of $\bar{\boldsymbol{E}}'$, respectively. In the framework of the PV2PQ conversion approach, which is based on treating PV buses as PQ buses, by taking the partial derivative of Eq. (A.5) over the active power and denoting $\bar{\boldsymbol{X}}_P$ as an $N \times N$ matrix of which elements are the derivative of voltages over the active power, we have

$$\boldsymbol{\delta} = (\boldsymbol{\alpha}_D + \boldsymbol{\gamma}) \cdot \boldsymbol{X}_P^d + (\boldsymbol{\beta}_D + \boldsymbol{\eta}) \cdot \boldsymbol{X}_P^q \tag{A.7}$$

for the real part and

$$\boldsymbol{O} = (\boldsymbol{\beta}_D - \boldsymbol{\eta}) \cdot \boldsymbol{X}_P^d + (-\boldsymbol{\alpha}_D + \boldsymbol{\gamma}) \cdot \boldsymbol{X}_P^q \tag{A.8}$$

for the imaginary part. In above: $\boldsymbol{O}$ is an $N \times N$ null matrix; $\boldsymbol{\delta}$ is an $N \times N$ identity matrix; $\boldsymbol{\gamma}$ and $\boldsymbol{\eta}$, $N \times N$ matrices, read

$$\boldsymbol{\gamma} = \boldsymbol{E}_D^d \cdot \boldsymbol{G} + \boldsymbol{E}_D^q \cdot \boldsymbol{B}, \qquad \boldsymbol{\eta} = -\boldsymbol{E}_D^d \cdot \boldsymbol{B} + \boldsymbol{E}_D^q \cdot \boldsymbol{G} \tag{A.9}$$

with $\boldsymbol{G}$ and $\boldsymbol{B}$, in respective, as the real and imaginary parts of the $N \times N$ admittance compound matrix $\bar{\boldsymbol{Y}}$ (without rows and columns corresponding to slack buses). Similarly, taking the partial derivative of Eq. (A.5) over the reactive power leads to the following equations:

$$\boldsymbol{O} = (\boldsymbol{\alpha}_D + \boldsymbol{\gamma}) \cdot \boldsymbol{X}_Q^d + (\boldsymbol{\beta}_D + \boldsymbol{\eta}) \cdot \boldsymbol{X}_Q^q \tag{A.10}$$

for the real part and

$$-\boldsymbol{\delta} = (\boldsymbol{\beta}_D - \boldsymbol{\eta}) \cdot \boldsymbol{X}_Q^d + (-\boldsymbol{\alpha}_D + \boldsymbol{\gamma}) \cdot \boldsymbol{X}_Q^q \tag{A.11}$$

for the imaginary part, where $\bar{\boldsymbol{X}}_Q^d$ and $\bar{\boldsymbol{X}}_Q^q$ are, in respective, the real and imaginary parts of $\bar{\boldsymbol{X}}_Q$ as an $N \times N$ matrix of which elements are the derivative of voltages over the reactive power.

Now, by rewriting $\boldsymbol{X}_{P(Q)}^d$ and $\boldsymbol{X}_{P(Q)}^q$, the real and imaginary parts of $\bar{\boldsymbol{X}}_{P(Q)}$, respectively, in terms of voltage magnitude and angle sensitivity matrices versus variations of the active (reactive) power, $\boldsymbol{SV}_{m,P(Q)}$ and $\boldsymbol{SV}_{a,P(Q)}$, respectively, we have

$$\boldsymbol{X}_{P(Q)}^d = \boldsymbol{EE}_D^d \cdot \boldsymbol{SV}_{m,P(Q)} - \boldsymbol{E}_D^q \cdot \boldsymbol{SV}_{a,P(Q)}, \tag{A.12}$$

and

$$\boldsymbol{X}_{P(Q)}^q = \boldsymbol{EE}_D^q \cdot \boldsymbol{SV}_{m,P(Q)} + \boldsymbol{E}_D^d \cdot \boldsymbol{SV}_{a,P(Q)}, \tag{A.13}$$

where diagonal matrices $\boldsymbol{EE}_D^{q[d]}$ is given by $\boldsymbol{E}_D^{q[d]}/\left|\bar{\boldsymbol{E}}\right|$. Using Eqs. (A.12-A.13) together with Eqs. (A.7-A.11) and after straightforward matrix operations, we construct the following block matrix equation to extract voltage magnitude and angle sensitivity full matrices:

$$\begin{bmatrix} \boldsymbol{SV}_{m,P(Q)} \\ \boldsymbol{SV}_{a,P(Q)} \end{bmatrix} = \begin{bmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{bmatrix}^{-1} \cdot \begin{bmatrix} \boldsymbol{\delta}\ (\boldsymbol{O}) \\ \boldsymbol{O}\ (-\boldsymbol{\delta}) \end{bmatrix}, \tag{A.14}$$

where the $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$, and $\boldsymbol{D}$ matrices, as matrix elements of CBM, are given by

$$\boldsymbol{A} = (\boldsymbol{\alpha}_D + \boldsymbol{\gamma}) \cdot \boldsymbol{EE}_D^d + (\boldsymbol{\beta}_D + \boldsymbol{\eta}) \cdot \boldsymbol{EE}_D^q \tag{A.15}$$

$$\boldsymbol{B} = -(\boldsymbol{\alpha}_D + \boldsymbol{\gamma}) \cdot \boldsymbol{E}_D^q + (\boldsymbol{\beta}_D + \boldsymbol{\eta}) \cdot \boldsymbol{E}_D^d \tag{A.16}$$

$$\boldsymbol{C} = (\boldsymbol{\beta}_D - \boldsymbol{\eta}) \cdot \boldsymbol{EE}_D^d + (-\boldsymbol{\alpha}_D + \boldsymbol{\gamma}) \cdot \boldsymbol{EE}_D^q \tag{A.17}$$

$$\boldsymbol{D} = -(\boldsymbol{\beta}_D - \boldsymbol{\eta}) \cdot \boldsymbol{E}_D^q + (-\boldsymbol{\alpha}_D + \boldsymbol{\gamma}) \cdot \boldsymbol{E}_D^d. \tag{A.18}$$

Thus, the Y_{bus}-extended method in the framework of the PV2PQ conversion approach can provide the voltage magnitude and angle sensitivity full matrices through the two-step algorithm: a) building the $2N \times 2N$ CBM by using the bus voltage vector and admittance matrix of the network and the predefined $2N \times N$ block matrix which is constructed by the identity and null matrices (see the left hand sides of Eqs. (A.7-A.11); b) multiplying the inverse of the CBM by the predefined $2N \times N$ block matrix to extract the voltage sensitivity full matrices.

With voltage sensitivity matrices versus active/reactive power at hand, the SI of line current $\bar{I}_{ij}$ (from bus i to bus j) versus variation of active/reactive power at bus l is given by:

$$SI_{m_{ij,l}} = II_{ij}^d J_{ij,l}^d + II_{ij}^q J_{ij,l}^q, \tag{A.19}$$

where II_{ij}^d and II_{ij}^q are the real and imaginary parts of $\overline{II}_{ij} = \bar{I}_{ij}/\left|\bar{I}_{ij}\right|$ and $J_{ij,l}^d$ and $J_{ij,l}^q$ are, in respective, the real and imaginary parts of $\bar{J}_{ij,l}$ which reads

$$\bar{J}_{ij,l} = \bar{Y}_{ij}(\bar{X}_{il} - \bar{X}_{jl}), \tag{A.20}$$

with $\bar{X}_{il}$ and $\bar{X}_{jl}$ are the matrix elements of $\bar{\boldsymbol{X}}$ of which the real and imaginary parts are related to the voltage magnitude and angle sensitivity matrices through Eqs. (A.12-A.13).

Now, for the sake of completeness of this subsection, we present the SI of voltage versus OLTC position by following the definition and relevant assumptions in Ref. [102]. By: defining $\boldsymbol{SV}_{m,Tap}$ and $\boldsymbol{SV}_{a,Tap}$, in respective, as the voltage magnitude and angle sensitivity matrix versus the variation of tap-changer position; denoting $\boldsymbol{X}_s^d$ and $\boldsymbol{X}_s^q$, in respective, as the real and imaginary parts of the $N \times S$ matrix $\bar{\boldsymbol{X}}_s$, where $\bar{\boldsymbol{X}}_s = \frac{\partial \bar{\boldsymbol{E}}}{\partial \left|\bar{\boldsymbol{E}}_s\right|}$ with $\bar{\boldsymbol{E}}_s$ as the voltage vector of slack buses with size S; substituting $\boldsymbol{SV}_{m,Tap}$, $\boldsymbol{SV}_{a,Tap}$ and $\boldsymbol{X}_s^{d(q)}$ in replace with $\boldsymbol{SV}_{m,P(Q)}$, $\boldsymbol{SV}_{a,P(Q)}$, and

$\boldsymbol{X}^{d(q)}_{P(Q)}$, respectively, in Eqs. (A.12-A.13); the voltage magnitude and angle sensitivity entire matrices versus tap-changer position can be obtained from the CBM of network, Eqs. (A.15-17), as follows:

$$\begin{bmatrix} \boldsymbol{SV}_{m,Tap} \\ \boldsymbol{SV}_{a,Tap} \end{bmatrix} = - \begin{bmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{bmatrix}^{-1} \cdot \begin{bmatrix} \boldsymbol{T}^d \\ \boldsymbol{T}^q \end{bmatrix}, \tag{A.21}$$

where $\boldsymbol{T}^d$ and $\boldsymbol{T}^q$ are the real and imaginary parts of $\bar{\boldsymbol{T}}$ which reads:

$$\bar{\boldsymbol{T}} = \underline{\boldsymbol{E}}_D \cdot \bar{\boldsymbol{Y}}_S \cdot \overline{\boldsymbol{EE}}_{S,D} \tag{A.22}$$

with $\bar{\boldsymbol{Y}}_S$ as the $N \times S$ admittance matrix (without rows and columns corresponding to slack and non-slack buses, respectively) and $\overline{\boldsymbol{EE}}_{S,D} = \bar{\boldsymbol{E}}_{S,D} / \left| \bar{\boldsymbol{E}}_S \right|$.

A.3 Z_{bus} method for network reconfiguration

For the graph representing a network, which is defined by a set of N_g nodes (denoted by i, j, etc.) and B_g branches (denoted by *ij*, *jk*, etc.), $\hat{\boldsymbol{M}}$ is a $B_g \times N_g$ incidence matrix of which entry a_{ij} is: +1 (−1) if current in branch i is directed away from (toward) node j and 0 if branch i is not connected to node j. In the graph-based power system network calculations, the corresponding column of reference node in $\hat{\boldsymbol{M}}$ is removed and the rest matrix, denoted by $\boldsymbol{M}$, is used as the incidence matrix. In this paper, $\boldsymbol{M}$, is the $B \times M$ incidence matrix of the network with B (M) as the total number of lines (buses).

If the complex voltage vector $\bar{\boldsymbol{E}}$ of buses is produced by the actual current injection vector with M known fixed elements, $\bar{\boldsymbol{I}}_{injc}$, and $\bar{\boldsymbol{Z}}_b$ is the Z_{bus} matrix, the bus impedance equation reads [103]

$$\bar{\boldsymbol{E}} = \bar{\boldsymbol{Z}}_b \cdot \bar{\boldsymbol{I}}_{injc} \tag{A.23}$$

Then by simultaneously closing switch action, we add to the network K lines ij, kl, ..., pq, with the line impedance $\bar{Z}_{L,ij}$, $\bar{Z}_{L,kl}$, ..., $\bar{Z}_{L,pq}$, respectively. Thus, the voltage will be deviated as $\Delta\bar{\boldsymbol{E}} = \bar{\boldsymbol{E}}' - \bar{\boldsymbol{E}}$, where $\bar{\boldsymbol{E}}'$ is the voltage vector of M buses after adding K lines. By denoting the currents in the added lines as $\bar{\boldsymbol{I}}_K$, the new bus voltage equation calls

$$\bar{\boldsymbol{E}}_K = \bar{\boldsymbol{Z}}_{LK} \cdot \bar{\boldsymbol{I}}_K, \tag{A.24}$$

where $\bar{\boldsymbol{E}}_K$ is a complex vector of size K of which elements are $\bar{E}'_i - \bar{E}'_j$, $\bar{E}'_k - \bar{E}'_l$, ..., $\bar{E}'_p - \bar{E}'_q$ and the line impedance matrix $\bar{\boldsymbol{Z}}_{LK}$ is a $K \times K$ diagonal matrix of which elements are $\bar{Z}_{L,ij}$, $\bar{Z}_{L,kl}$, ..., $\bar{Z}_{L,pq}$. A straightforward matrix operation shows that $\bar{\boldsymbol{E}}_K = \boldsymbol{M}_K \cdot \bar{\boldsymbol{E}}'$ with $\boldsymbol{M}_K \subseteq \boldsymbol{M}$ as the $K \times M$ incidence matrix that has non-zero values in each row corresponding to the new line which connects the relevant buses. On the other hand, the new line currents to the network affects the voltages of the original network same as K sets of injected currents in the Thévenin equivalent circuit [103], that is, $-I_{ij}$ at bus i and $+I_{ij}$ at bus j, and also for the other new lines. Thus, these new current injections together with the actual current injections, $\bar{\boldsymbol{I}}_{injc}$, into the original network make a new set of current injections, denoted by $\bar{\boldsymbol{I}}'_{injc}$ as an $M \times 1$ vector, which produces the new voltages $\bar{\boldsymbol{E}}'$. It is easy to show that $\boldsymbol{M}_K$ connects $\bar{\boldsymbol{I}}'_{injc}$ to $\bar{\boldsymbol{I}}_K$ through $\bar{\boldsymbol{I}}'_{injc} = -\boldsymbol{M}_K^T \cdot \bar{\boldsymbol{I}}_K$ and thus, the deviation of voltages after switch action reads

$$\triangle \bar{\boldsymbol{E}} = -\bar{\boldsymbol{Z}}_b \cdot \boldsymbol{M}_K^T \cdot \bar{\boldsymbol{I}}_K. \tag{A.25}$$

By using the foregoing equations and after matrix operations, the new lines' currents call

$$\bar{\boldsymbol{I}}_K = \bar{\boldsymbol{Z}}_{loop}^{-1} \cdot \boldsymbol{M}_K \cdot \bar{\boldsymbol{E}}, \tag{A.26}$$

where $\bar{\boldsymbol{Z}}_{loop}$, the loop impedance matrix for the case of adding lines to the network, after some matrix operations, reads

$$\bar{\boldsymbol{Z}}_{loop} = \bar{\boldsymbol{Z}}_{th} + \bar{\boldsymbol{Z}}_{LK} \tag{A.27}$$

with $\bar{\boldsymbol{Z}}_{th}$ as a $K \times K$ matrix of which elements is given by $\bar{Z}^{th}_{ij,kl} = (\bar{Z}_{b_{ik}} - \bar{Z}_{b_{il}}) - (\bar{Z}_{b_{jk}} - \bar{Z}_{b_{jl}})$. Finally, the voltage deviation after simultaneously closing switches can directly be obtained as follows:

$$\triangle \bar{\boldsymbol{E}} = \bar{\boldsymbol{Z}}_{recon} \cdot \bar{\boldsymbol{E}}, \tag{A.28}$$

where

$$\bar{\boldsymbol{Z}}_{recon} = -\bar{\boldsymbol{Z}}_b \cdot \boldsymbol{M}_K^T \cdot \bar{\boldsymbol{Z}}_{loop}^{-1} \cdot \boldsymbol{M}_K. \tag{A.29}$$

Eq. (A.28) can be used for the case of simultaneously opening switches action by changing the sign + to − in Eq. (A.27). Clearly, Eq. (A.28) is also applicable for sequential switch actions provided that $\bar{\boldsymbol{Z}}_b$ is updated before applying next action to the network. The updated $\bar{\boldsymbol{Z}}_b$ calls [103]:

$$\bar{\boldsymbol{Z}}_{b,updated} = \bar{\boldsymbol{Z}}_b - \frac{\bar{\boldsymbol{Z}}_L^T \cdot \bar{\boldsymbol{Z}}_L}{\bar{Z}^{th}_{ij,ij} \pm \bar{Z}_{L,ij}}, \tag{A.30}$$

where +(−) is for the adding line ij to (removing line ij from) the network.

The deviation of current flow in each line ij, with the line impedance $\bar{Z}_{L,ij}$, can be computed by use of the voltage deviation at buses i and j as follows

$$\nu \bar{I}_{ij} = \frac{\nu \bar{E}_i - \nu \bar{E}_j}{\bar{Z}_{L,ij}}. \tag{A.31}$$

By using Eq. (A.31) for sequentially switching actions on the network, one obtains the same results as the DF scheme, presented in [103] [2]. It should be highlighted that the DF scheme has advantage to present precomputed tables of DFs to find current deviations and one can not find voltage deviations by use of DFs. Hence, constructing the IM scheme is necessary if one needs to reach information about voltage deviations.

[2] We checked it for different networks including Oberrhein-MV.

Appendix B

Chapter 4 Annex

B.1 Enhanced ExpertOP4Grid algorithms

Algorithm 4: J_1: Line Congestion Assessment and Prioritization

Input: $\mathcal{E}$ (set of grid lines), f_e (flow on line e), f_e^{max} (thermal limit), T_e (overflow duration)
Parameter: α_f (congestion threshold ratio), α_t (time threshold)
Output: $\mathcal{C}_{crit}, \mathcal{C}_{non}$ (critical and non-critical congested lines)

$\mathcal{C}_{crit} \leftarrow \emptyset$, $\mathcal{C}_{non} \leftarrow \emptyset$
foreach $e \in \mathcal{E}$ **do**
  if $f_e > \alpha_f \cdot f_e^{max}$ **then**
    if $T_e > \alpha_t$ **then**
      $\mathcal{C}_{crit} \leftarrow \mathcal{C}_{crit} \cup \{e\}$
    else
      $\mathcal{C}_{non} \leftarrow \mathcal{C}_{non} \cup \{e\}$
$\mathcal{C}_{crit} \leftarrow \text{SortDescending}(\mathcal{C}_{crit}, f_e)$
$\mathcal{C}_{non} \leftarrow \text{SortDescending}(\mathcal{C}_{non}, f_e)$
return $\mathcal{C}_{crit}, \mathcal{C}_{non}$

Algorithm 5: J_2: Skip Criterion Based on Critical Zone

Input: $\mathcal{C}_{crit}$ (critical congested lines), $\mathcal{Z}_e$ (zone function: line → zone)
Output: skip (Boolean)

$\mathsf{Z} \leftarrow \{\mathcal{Z}_e(e) \mid e \in \mathcal{C}_{crit}\}$ `// Extract zones`
if $|\mathsf{Z}| = 1$ **then**
  skip ← True
else
  skip ← False
return skip

Algorithm 6: J_3–J_7: Flexibility Resource Evaluation

Input: $\mathscr{L}$ (set of overloaded lines), $\mathscr{G} = (\mathscr{B}, \mathscr{E})$ (grid graph), δf_e (flow deviation), λ_L (Laplacian eigenvalues)

Output: $\mathscr{A}$ (set of valid actions)

`/* J3 - Hub Bus Splitting */`

$\mathscr{B}_{hub} \leftarrow \{b \in \mathscr{B} \mid \exists$ DFS path on $\delta f_e < 0$ reaching $\delta f_e > 0\}$

foreach $b \in \mathscr{B}_{hub}$ **do**

  if $\deg(b) > 2 \wedge \lambda_L^{max}(b) < \lambda_{th}$ **then**

    $\mathscr{A} \leftarrow \mathscr{A} \cup \{\text{Split}(b)\}$

`/* J4 - Loop Bus Splitting */`

$\mathscr{B}_{loop} \leftarrow \{b \in \mathscr{B} \mid b \in$ shortest path of $\delta f_e > 0\}$

foreach $b \in \mathscr{B}_{loop}$ **do**

  if $\deg(b) > 2 \wedge \lambda_L^{max}(b) < \lambda_{th}$ **then**

    $\mathscr{A} \leftarrow \mathscr{A} \cup \{\text{Split}(b)\}$

`/* J5 - Downstream Bus Splitting */`

$\mathscr{B}_{down} \leftarrow \text{DFS}(\delta f_e < 0) \cup$ backup supply buses

foreach $b \in \mathscr{B}_{down}$ **do**

  if $\deg(b) > 2 \wedge \lambda_L^{max}(b) < \lambda_{th}$ **then**

    $\mathscr{A} \leftarrow \mathscr{A} \cup \{\text{Split}(b)\}$

`/* J6 - Branch Switching */`

foreach $e \in \mathscr{E} \setminus \mathscr{L}$ **do**

  if $e \in$ *shortest path from* $\mathscr{L}$ **then**

    $\mathscr{A} \leftarrow \mathscr{A} \cup \{\text{Switch}(e)\}$

`/* J7 - Power Flexibility */`

foreach $a \in \mathscr{F}_{pwr}$ **do**

  if *Sensitivity*$(a) > \theta_s$ ***and*** *Legal*$(a) =$ *True* **then**

    $\mathscr{A} \leftarrow \mathscr{A} \cup \{a\}$

return $\mathscr{A}$

Algorithm 7: J_8–J_{10}: Stop Conditions Evaluation

Input: $\mathscr{C}_{crit}$ (critical lines), $\mathscr{C}_{all}$ (all congested lines), score_{topo} (topological score), $N_{checked}$ (issues evaluated)

Parameter: α_9 (minimum score for critical relief), N_{max} (maximum issues)

Output: stop (Boolean)

stop $\leftarrow$ False

`/* J8 - Stop if all overloads resolved */`

if $|\mathscr{C}_{all}| = 0$ **then**

  stop $\leftarrow$ True

`/* J9 - Stop if critical relief with sufficient score */`

if $\exists c \in \mathscr{C}_{crit}$ *resolved* $\wedge\ \text{score}_{topo}(c) \geq \alpha_9$ **then**

  stop $\leftarrow$ True

`/* J10 - Stop if too many issues checked */`

if $N_{checked} \geq N_{max}$ **then**

  stop $\leftarrow$ True

return stop

B.1. Enhanced ExpertOP4Grid algorithms

Algorithm 8: J_{13}: Lexicographic Flexibility Ranking

Input: $\mathscr{A} = \{a_1, a_2, \ldots, a_n\}$ (candidate actions)
Input: $\pi_f(a)$: priority, $\sigma_t(a)$: topological score, $\rho(a)$: operational score
Output: $\mathscr{A}_{ranked}$ (ranked list)

`/* Define the ranking tuple */`

foreach $a \in \mathscr{A}$ **do**
  $\theta(a) \leftarrow \big(\pi_f(a), \sigma_t(a), \rho(a)\big)$

`/* Sort actions lexicographically (descending) */`

$\mathscr{A}_{ranked} \leftarrow \text{SortDescending}(\mathscr{A}, \theta(a))$
return $\mathscr{A}_{ranked}$

Algorithm 9: Topological Score Computation

Input: Simulated result of action a:
relieved_lines(a), worsened_lines(a), violations(a)
Parameter: γ_r: Relief threshold, γ_w: Worsening threshold
Output: $\sigma_t(a)$ (Topological score)

if *violations*$(a) \neq \emptyset$ *or relieved_lines*$(a) = \emptyset$ **then**
  $\sigma_t(a) \leftarrow 0$ `// No relief or infeasible`
else if $0 < \frac{|relieved\_lines(a)|}{|initial\_overloads|} < \gamma_r$ **then**
  $\sigma_t(a) \leftarrow 1$ `// Minor relief`
else if $\exists \ell \in worsened\_lines(a) : \frac{\Delta f_\ell}{f_\ell^{\max}} > \gamma_w$ **then**
  $\sigma_t(a) \leftarrow 2$ `// Relief causes new critical overload`
else if *relieved_lines*$(a) \subseteq$ *non-critical lines* **then**
  $\sigma_t(a) \leftarrow 3$ `// Non-critical relief only`
else if *relieved_lines*(a) *includes any critical line* **then**
  $\sigma_t(a) \leftarrow 4$ `// Critical line relieved safely`
else if $|remaining\_overloads(a)| = 0$ **then**
  $\sigma_t(a) \leftarrow 5$ `// Full relief`
return $\sigma_t(a)$

Algorithm 10: J_{11}: Grid Recovery – No Available Action

Input: $\mathscr{I}$: Set of technical issues
Input: $\mathscr{A}$: Set of candidate actions
Input: $\mathscr{A}_{valid} \subseteq \mathscr{A}$: Set of valid actions
Output: Trigger recovery action P_{22} if no action is available

if $|\mathscr{I}| = 0$ **then**
  return *Activate* P_{22} `// No technical issue observed`
else if $|\mathscr{A}_{valid}| = 0$ **then**
  return *Activate* P_{22} `// No action valid`
else
  return *Continue to next judgment node*

B.1. Enhanced ExpertOP4Grid algorithms

Algorithm 11: J_{12}: Grid Recovery – No Effective Action

Input: $\mathscr{A}_{valid} = \{a_1, a_2, \ldots, a_n\}$ (valid actions)
Input: $\sigma_t(a)$ for all $a \in \mathscr{A}_{valid}$
Parameter: τ_{J12}: Topological score threshold
Output: Trigger recovery action P_{23} if no effective action

```
if ∀a ∈ A_valid, σ_t(a) < τ_J12 then
    return Activate P_23                    // No sufficiently effective action
else
    return Continue to next judgment node
```

Algorithm 12: J_{14}: Internal Ranking Revision (Score-Based)

Input: $\mathscr{A}_{valid} = \{a_1, a_2, \ldots, a_n\}$: Valid actions
Input: $\sigma_t(a)$, $\rho(a)$ $\forall a \in \mathscr{A}_{valid}$
Parameter: α_{J14}: Score threshold
Output: Revised ranking $\mathscr{R}$ (score- and reward-based)

```
if max_{a∈A_valid} σ_t(a) < α_J14 then
    R ← Sort A_valid by descending σ_t(a), break ties with descending ρ(a)
    return R                                 // Score-based revision only
else
    return Preserve original ranking
```

Algorithm 13: J_{15}: External Ranking Revision (High Score & Minimal Worsening)

Input: $\mathscr{A}_{valid} = \{a_1, a_2, \ldots, a_n\}$
Input: $\sigma_t(a)$, $w(a)$ $\forall a \in \mathscr{A}_{valid}$
Parameter: α_{J15}: Topological score threshold
Output: Single selected action a^*

```
if max_{a∈A_valid} σ_t(a) < α_J15 then
    A_filtered ← {a ∈ A_valid | σ_t(a) = max_{a'} σ_t(a')}
    a* ← arg min_{a∈A_filtered} w(a)
    return a*      // Select action with highest score, fewest worsened lines
else
    return Preserve original best-ranked action
```

Algorithm 14: J_{16}: Multiple Action Selection (Insufficient Single Action)

Input: $\mathscr{A}_{valid} = \{a_1, a_2, \ldots, a_n\}$: Valid actions
Input: $\sigma_t(a)$, $w(a)$ $\forall a \in \mathscr{A}_{valid}$
Parameter: γ_{min}: Minimum topological score threshold for sufficient action
Output: Set of selected actions $\mathscr{A}_{selected}$

```
if max_{a∈A_valid} σ_t(a) < γ_min then
    A_selected ← ∅
    for a ∈ A_valid do
        if σ_t(a) ≥ γ_min then
            Add action a to A_selected
    return A_selected            // Multiple actions selected for sufficient relief
else
    return Select best action: arg max_{a∈A_valid} σ_t(a)
```

if $\max_{a \in \mathscr{A}_{valid}} \sigma_t(a) < \gamma_{min}$ **then**
  $\mathscr{A}_{selected} \leftarrow \emptyset$
  for $a \in \mathscr{A}_{valid}$ **do**
    if $\sigma_t(a) \geq \gamma_{min}$ **then**
      Add action a to $\mathscr{A}_{selected}$
  return $\mathscr{A}_{selected}$ `// Multiple actions selected for sufficient relief`
else
  return *Select best action:* $\arg\max_{a \in \mathscr{A}_{valid}} \sigma_t(a)$

Appendix C

Chapter 5 Annex

C.1 Actor Prompt

```
You are a power-system explanation agent for transmission grid operation.

Your task is to explain why redispatch actions occurred and why the realized
redispatch is the most plausible local solution under the reported operating
conditions.

Use only the information explicitly present in the provided filtered redispatch
report.
Do NOT solve OPF or power flow.
Do NOT invent missing data.
Do NOT assume unreported constraints, reserves, sensitivities, limits, or
remedial actions.
Use PTDF and LODF only as first-order screening indicators, not exact AC results.

----------------------------------------------------------------------
INPUT
----------------------------------------------------------------------

{snapshot}

----------------------------------------------------------------------
OBJECTIVE
----------------------------------------------------------------------

Write an operator-facing analytical note that explains:
```

```
1. all significant violated or binding constraints before redispatch, especially
all reported thermal overloads and any reported voltage-limit or load-curtailment
issues,
2. which of those constraints is the dominant driver of redispatch,
3. how the realized redispatch actions on conventional generators, wind
curtailment, and load curtailment relieve those reported constraints,
4. one mathematically explicit counterfactual showing why a plausible alternative
was not selected,
5. one specific quantified change that would make that counterfactual viable or
optimal,
6. one screening-based N-1 statement if LODF data are present and arithmetic can
be formed.

Important:

This explanation is only about redispatch of active power.

Use only the corrective actions allowed in the report and their priority:
- upward and downward redispatch of conventional generators,
- curtailment of wind generation, limited to 30% of available capacity,
- emergency load curtailment only as a last resort, limited to 30% and associated
with high penalty.

Do not discuss other mechanisms unless they are explicitly included in the
report.

The operator already knows the system situation.
Do NOT summarize the whole snapshot.
Do NOT produce a long list of raw values without interpretation.
Use only the numbers needed to:
- connect all reported overloads or violated constraints,
- explain the realized redispatch pattern,
- justify the main counterfactual mathematically,
- and justify the N-1 screening result if relevant.

---------------------------------------------------------------------
INPUT INTERPRETATION RULES
---------------------------------------------------------------------

Assume the filtered report may contain:
- redispatch objective value
- prices
- generator summary after redispatch
- bus summary with pre_redispatch and post_redispatch values
- line loading with pre_redispatch and post_redispatch values
- filtered PTDF
- filtered LODF
- nonzero duals after redispatch
```

```
Use only what is reported.

The report may use field names such as:
- pre_redispatch_V
- post_V
- pre_redispatch_loading
- post_redispatch_loading
- post_redispatch_P
- post_load_curtailment

Read these as report labels only.
In the explanation, prefer natural wording such as:
- before redispatch
- after redispatch
- final generation
- final line loading

If the report provides generator redispatch variables such as upward and
downward, you must state redispatch changes in MW for the relevant conventional
generators.

Interpret them as:
- net redispatch change = upward - downward
- generation before redispatch = post_P - upward + downward

Use explicit wording such as:
- "conventional generator Gen 8 at bus 38 was reduced by 20 MW"
- "conventional generator Gen 2 at bus 32 was increased by 20 MW"

Do not describe conventional-generator redispatch only as "output at a bus" if
generator identity is available.

For wind generators, use only explicitly reported:
- final active-power output,
- available capacity,
- curtailment.

Do not invent a MW increase or decrease for wind generators unless the report
explicitly supports it.

---------------------------------------------------------------------
REPORTED-DATA-ONLY RULE
---------------------------------------------------------------------

You may only use PTDF, LODF, prices, duals, and redispatch values that are
explicitly reported for the exact bus, generator, line, or branch being
discussed.
```

```
Do NOT:
- borrow a PTDF or LODF value from a nearby bus, adjacent endpoint, corridor, or
another generator,
- infer a PTDF column for a missing generator bus,
- infer a dual for an unreported constraint,
- infer a redispatch action that is not explicitly supported by the report.

If a required value is missing for a specific generator, bus, or line, do not
make that claim.

Invalid examples:
- "bus 36 can be treated like bus 24"
- "this nearby endpoint suggests the same PTDF"
- "this unit likely relieved the line"
unless that statement is explicitly supported by reported values.

---------------------------------------------------------------------
WHAT TO EXPLAIN FIRST
---------------------------------------------------------------------

Start by identifying all significant violated or binding constraints before
redispatch.

If multiple lines are overloaded before redispatch, do not discuss them as
isolated cases.
Explain whether they are connected by:
- a common transfer direction,
- a shared corridor (define the corridor as set of buses),
- a common exporting or importing area (define the area as buses),
- or a redispatch pattern that relieves several lines at the same time.

Then identify the dominant constraint, but keep the explanation connected to the
broader overload pattern.

Do not start with a long list of generator values.
Do not reduce the explanation to a single generator-line statement if the report
supports a broader multi-line interpretation.
---------------------------------------------------------------------
ARITHMETIC VALIDITY RULE
---------------------------------------------------------------------

Any claim that uses PTDF, LODF, prices, curtailment cost, redispatch price, or
dual variables is valid only if the note shows the corresponding arithmetic
explicitly.

Invalid examples:
- "PTDF shows this increases loading"
```

```
- "this is cheaper"
- "LODF indicates higher contingency risk"
- "relaxing the ramp limit would help"

Valid claims must show:
- the reported coefficient(s) or value(s),
- the sign-consistent calculation,
- and the resulting numeric conclusion.

If the required arithmetic cannot be formed from the report, do not make the
claim.

---------------------------------------------------------------------
DUAL-VARIABLE RULES
---------------------------------------------------------------------

Use nonzero duals as evidence that a constraint is active or economically
important.

Because sign conventions may be solver-dependent, do not rely on sign alone
unless the interpretation is explicit in the report.

Interpret duals as follows:
- Pbal duals: indicate how the redispatch objective changes for a small change in
active-power balance at that bus
- Qbal duals: indicate how the redispatch objective changes for a small change in
reactive-power balance at that bus
- thermal duals: congestion-related binding constraints
- voltage duals: voltage-related binding constraints
- ramp duals: redispatch flexibility limits

Use duals only when they strengthen the counterfactual or explain why a displaced
unit cannot be changed.

Do not infer more from duals than the report supports.

Voltage limits may be cited only if supported by:
- a reported voltage close to a limit (0.95 pu to 1.05 pu), or
- a clearly nonzero voltage-constraint dual of notable magnitude.

---------------------------------------------------------------------
PTDF RULES
---------------------------------------------------------------------

If PTDF data are present, use them to explain the network logic of redispatch,
not only the effect on one isolated line.

Prefer counterfactuals that reveal a nontrivial tradeoff, such as:
```

```
- relieving one overloaded line while worsening another,
- helping the dominant congested line while failing to help the broader
overloaded corridor,
- or reducing the need for one redispatch action while creating a new network
problem elsewhere.

Do not use PTDF only for a simple local sign check such as:
- "bus A worsens line X"
- "bus B relieves line X"

Instead, when the report supports it, connect the counterfactual to:
- at least two significant overloaded lines, or
- one dominant overloaded line plus one additional relevant constrained line.

For a marginal swap of +1 MW at one bus and -1 MW at another, use:
- approximate flow change on monitored line ≈ PTDF(injecting bus) -
PTDF(displaced bus)

If multiple relevant lines are available, compute and compare the effect on more
than one line when that helps explain the redispatch.

Then explain whether the counterfactual:
- relieves the dominant line but worsens another important line,
- fails to relieve the broader overload pattern,
- or is consistent with the realized redispatch direction across multiple
overloaded lines.

Every verbal claim such as "relieves" or "worsens" must match the sign of the
PTDF arithmetic.
Do not infer PTDF effects for unreported buses or generators.

----------------------------------------------------------------------
MANDATORY MARGINAL ARITHMETIC
----------------------------------------------------------------------

You MUST include exactly one explicit marginal counterfactual.

The counterfactual must be technically informative, not trivial.
Choose a case that reveals a non-obvious network tradeoff that is not obvious
from a single overloaded line alone.

A valid counterfactual must include ALL of the following in the final note:

1. Define one marginal action:
- +1 MW or -1 MW injection or withdrawal at a specific generator or bus

2. Identify the displaced resource:
- state which generator or load is displaced
```

```
- if the obvious displaced unit cannot be changed, state the reported constraint
that blocks it

3. Compute the network effect:
- use PTDF or equivalent screening data to compute the approximate effect on the
dominant monitored line
- and, when possible, also compute the effect on at least one additional
significant overloaded or relevant line

4. Verify directional consistency:
- ensure the PTDF arithmetic matches the verbal claim about relieving or
worsening each reported line

5. Compute the marginal economic effect:
- use only explicit economic signals reported in the snapshot, such as:
  - redispatch prices,
  - curtailment costs,
  - thermal duals,
  - ramp duals,
  - voltage duals,
  - or other reported marginal values

6. Close the calculation with a scalar result:
- ΔC = [value] €/MWh

7. Explain the result of ΔC explicitly:
- state whether ΔC > 0 means the counterfactual is more expensive,
- state whether ΔC < 0 still fails because of a reported binding constraint,
- and identify which reported term dominates the result

8. State one explicit numerical threshold that would make the counterfactual
viable or optimal.

Acceptable threshold statements include:
- the dominant monitored line would need at least X MW additional margin,
- the broader overloaded corridor would need at least X MW additional margin on
the most restrictive line,
- the relevant dual would need to fall below Y €/MWh,
- the displaced unit or curtailment cost difference would need to change by Z
€/MWh.

Avoid trivial counterfactuals that only restate an obvious local PTDF sign.
Prefer counterfactuals that explain why the realized redispatch pattern is better
across multiple reported overloads.
---------------------------------------------------------------------
LODF / N-1 SCREENING REQUIREMENT
---------------------------------------------------------------------
```

```
If LODF data are present, you MUST evaluate whether they indicate a notable
contingency vulnerability.

You must explicitly discuss the outage of at least one critical line or corridor
and explain, using LODF or screening logic, why the condition before redispatch
would create overload risk or infeasibility under that outage, thereby showing
why redispatch was operationally necessary to reduce screened outage risk.

Use only screening logic.
Do NOT claim a full contingency analysis.

When possible, estimate post-contingency impact with:
- PostFlow_monitored ≈ PreFlow_monitored + LODF(monitored,outaged) × Flow_outaged

If only loading (%) and thermal limit are available, use:
- ApproxFlow ≈ (loading% / 100) × thermal_limit

The final note must show:
- the outaged line,
- the monitored line,
- the approximate pre-outage flow(s),
- the LODF coefficient,
- the approximate post-outage monitored flow,
- and whether that implies likely overload risk or not.

If multiple LODF-based screening cases are available, choose:
- one case that best demonstrates screened overload risk or movement toward a
thermal limit, and
- if possible, one case that shows large screened redistribution or flow
reversal.

For the second case, state clearly that it indicates major contingency
redistribution or flow reversal rather than a screened overload.
---------------------------------------------------------------------
USE OF NUMBERS
---------------------------------------------------------------------

Use the numbers needed for the main argument, but only where they support
reasoning.

Prioritize numbers for:
- all significant overloaded lines or violated constraints before redispatch,
- the dominant constraint,
- the main conventional-generator redispatch actions,
- wind curtailment or load curtailment if relevant,
- the PTDF marginal calculation,
- the ΔC calculation,
- the threshold for selecting the counterfactual,
```

```
- and the LODF screening result if relevant.

Do NOT repeat many snapshot values without analysis.
Do NOT write a long raw-data summary.

Use €/MWh for prices, curtailment costs, redispatch prices, dual-based marginal
comparisons, and ΔC.
Do not use €/h.
Do not use €/h per MW.

---------------------------------------------------------------------
STYLE RULES
---------------------------------------------------------------------

Use standard power-system engineering terminology.
Prefer simple and familiar operator language when a standard expression exists.

Prefer:
- redispatch
- conventional generator
- wind generator
- before redispatch
- after redispatch
- line congestion
- thermal overload
- thermal limit violation
- voltage limit violation
- load curtailment
- wind curtailment
- upward/downward
- ramp limit
- binding constraint
- congested line
- overloaded line
- feasible operating point
- generator output
- contingency screening
- thermal margin
- same congested corridor
- same transfer path
- nearby overloaded lines
- related overloaded lines
- overloads in the same area
- loading reduced on other reported overloaded lines

Avoid:
- active-power issue
- active-power constraint
```

```
- bottleneck
- operating pattern
- headroom
- electrically better placed
- intuitive alternative
- operator-grade evidence
- residual supply scarcity
- transfer stress
- energy-economic
- wind substitution
- bus-to-load corridor
- cut
- rose
- materially
- strongly
- significantly
- nearby stressed paths
- adjacent high-loading pattern
- appears reduced
- more viable
- screens as loading
- likely push more flow
- loading cluster
- stressed structure
- abstract or innovative wording that is not standard in power-system operation

Do not use vague intensifiers such as:
- strongly
- materially
- significantly

unless the corresponding numerical change is shown explicitly.

When describing more than one overloaded line, use familiar expressions such as:
- same congested corridor
- same transfer path
- nearby overloaded lines
- related overloaded lines
- overloads in the same area

Do not use abstract or stylistic phrases when a simpler power-system expression
is available.

Be direct, technical, and operator-oriented.
---------------------------------------------------------------------
OUTPUT FORMAT
---------------------------------------------------------------------
```

```
Output only an analytical note with 4 short paragraphs.

Paragraph 1:
Give a short description of the situation in the snapshot.
State all significant overloaded lines or other violated constraints before
redispatch, the main redispatch actions actually used, and whether load
curtailment occurred.
Explain briefly how the reported overloads or violated constraints are connected.
Keep this paragraph short and do not repeat many raw values.

Paragraph 2:
Counterfactual 1.
Give one complete counterfactual framed explicitly as:
- "Why not do X instead?"
- "What if condition Y changed?"

This paragraph must:
- name one plausible alternative redispatch action,
- identify one specific +1 MW / -1 MW marginal action,
- identify the displaced resource,
- use explicit PTDF arithmetic to quantify the effect on the relevant monitored
line,
- use reported prices and/or duals to compute a closed ΔC in €/MWh,
- explain what the sign of ΔC means,
- explain why the counterfactual is not selected under the reported conditions,
- and state one explicit numerical threshold that would make the counterfactual
viable or optimal.

The threshold may be expressed as:
- additional thermal margin on the relevant line,
- a lower relevant dual,
- a different displaced generator,
- or a different reported price or curtailment-cost difference.

Paragraph 3:
Counterfactual 2.
Give one second complete counterfactual framed explicitly as:
- "Why not do X instead?"
- "What if condition Y changed?"

Apply the same level of justification as in Paragraph 2.

This paragraph must:
- name a different plausible alternative redispatch action,
- identify one specific +1 MW / -1 MW marginal action,
- identify the displaced resource,
- use explicit PTDF arithmetic to quantify the effect on the relevant monitored
line,
```

```
- use reported prices and/or duals to compute a closed ΔC in €/MWh,
- explain what the sign of ΔC means,
- explain why the counterfactual is not selected under the reported conditions,
- and state one explicit numerical threshold that would make the counterfactual
viable or optimal.

The two counterfactuals should be different and should reveal different aspects
of the redispatch logic when possible, such as:
- one case dominated mainly by congestion,
- another case dominated mainly by price, ramp, voltage, or another reported
binding condition.

Paragraph 4:
Counterfactual 3.
Give two screening-based LODF examples if the report supports them:
- one case showing overload risk or movement toward a thermal limit,
- and one case showing large flow redistribution or flow reversal.

For each case, include:
- the outaged line,
- the monitored line,
- the approximate pre-outage flow(s),
- the LODF coefficient,
- the approximate post-outage monitored flow,
- and a clear screening-based conclusion.

For the redistribution case, state explicitly that it indicates major contingency
redistribution or flow reversal rather than a screened thermal overload.

Use only screening logic.
Do not claim a full AC contingency analysis.
If only one case is supported by the report, include only that case.

Use the three technical examples to connect the redispatch actions to:
- line congestion,
- economic or flexibility limits,
- and contingency security.

Hard limits:
- maximum 530 words
- short sentences
- no bullet points
- no headings
- no LaTeX
- no text outside the analytical note
```

C.2 Critic Prompt

```
You are reviewing an explanation of an active-power redispatch result for
transmission grid operation.

Your job is NOT to solve OPF, recompute flows, or verify exact AC physics.
Your job is to evaluate whether the explanation is technically credible, grounded
in the provided report, and written in clear power-system engineering language.

---------------------------------------------------------------------
INPUT
---------------------------------------------------------------------

Filtered redispatch report:
{snapshot}

Actor explanation:
{explanation}

---------------------------------------------------------------------
REVIEW OBJECTIVE
---------------------------------------------------------------------

Check whether the explanation is a strong operator-facing analytical note that
explains:

1. all significant violated or binding constraints before redispatch,
2. which of those constraints is the dominant driver of redispatch,
3. how the realized redispatch actions on conventional generators, wind
curtailment, and load curtailment relieve those reported constraints,
4. two mathematically explicit counterfactuals showing why plausible alternatives
were not selected,
5. one explicit quantified change that would make each counterfactual viable or
optimal,
6. one or two screening-based N-1 statements if LODF data support them.

The explanation concerns active-power redispatch only.

The explanation should reflect the corrective actions allowed in the report and
their priority:
- upward/downward redispatch of conventional generators,
- wind curtailment, limited to 30% of available capacity,
- emergency load curtailment only as a last resort, limited to 30% and associated
with high penalty.

You are NOT judging whether every physical statement is exactly correct.
You are judging whether the explanation is well supported by the report, causally
clear, mathematically explicit, and technically well written.
```

```
----------------------------------------------------------------------
CHECK THE FOLLOWING ELEMENTS
----------------------------------------------------------------------

1. Constraints before redispatch
The explanation identifies all significant overloaded lines or other violated
constraints that are clearly reported before redispatch.

It should not focus on only one line if several overloaded lines are reported and
relevant.

2. Network-level connection
If multiple overloaded lines are reported, the explanation should connect them
using familiar power-system language, such as:
- same congested corridor,
- same transfer path,
- nearby overloaded lines,
- related overloaded lines,
- overloads in the same area.

Do not reward isolated line-by-line description if the report supports a broader
multi-line explanation.

3. Dominant driver
The explanation identifies which reported violated or binding constraint appears
to be the dominant driver of redispatch.

If line congestion is clearly dominant in the report, the explanation should
identify that directly.

4. Correct description of redispatch actions
The explanation correctly describes the reported active-power corrective actions,
such as:
- upward/downward redispatch of conventional generators,
- wind curtailment,
- emergency load curtailment if present.

The explanation should not discuss unreported corrective actions.

5. Correct use of conventional-generator redispatch
If upward/downward values are reported for conventional generators, the
explanation should use them correctly to describe MW changes.

Accept statements such as:
- "conventional generator Gen 8 at bus 38 was reduced by 20 MW"
- "conventional generator Gen 2 at bus 32 was increased by 20 MW"
```

```
Do not reward vague wording such as:
- "output at bus 38 was reduced"
when generator identity is explicitly available.

6. Correct use of wind-generator information
For wind generators, the explanation should use only reported:
- final output,
- available capacity,
- curtailment,

unless a MW increase/decrease is explicitly supported by the report.

Flag unsupported wind-generator redispatch claims.

7. Causal explanation of realized redispatch
The explanation should connect the reported redispatch actions to the reported
violated or binding constraints.

It should explain how the realized pattern relieves the dominant line and, when
relevant, helps the broader overload pattern across multiple lines.

The explanation should not merely repeat raw numbers.

8. Reported-data-only discipline
The explanation must use only PTDF, LODF, duals, prices, and redispatch values
explicitly reported for the exact bus, generator, line, or branch being
discussed.

Flag as incorrect any claim that:
- borrows a PTDF or LODF from a nearby bus, adjacent endpoint, corridor, or
another generator,
- infers a PTDF column for a missing generator bus,
- infers a dual for an unreported constraint,
- or infers a redispatch action not explicitly supported by the report.

9. Numerical grounding
The explanation should use reported values selectively and meaningfully, such as:
- before-redispatch and after-redispatch line loading,
- generator redispatch amounts,
- wind curtailment,
- load curtailment,
- prices,
- PTDF values,
- LODF values,
- dual magnitudes.

The numbers must support reasoning.
Do not reward long raw-data summaries.
```

```
10. Dual-variable use
If nonzero duals are present and relevant, the explanation should use them
appropriately as evidence that a constraint is active or economically important.

Accept use such as:
- thermal duals for congestion-related binding constraints,
- voltage duals for voltage-related binding constraints,
- ramp duals for redispatch flexibility,
- Pbal/Qbal duals as bus-level marginal signals.

Do not require sign-based interpretation unless the sign convention is explicitly
established.

11. PTDF arithmetic
If PTDF data are present and relevant, each counterfactual should use explicit
PTDF arithmetic, not only verbal directional claims.

A valid PTDF-based claim should include:
- the specific marginal action,
- the relevant PTDF values,
- the sign-consistent calculation,
- and a clear conclusion on whether the action relieves or worsens the monitored
line.

Do not accept statements such as:
- "PTDF shows it worsens the line"
without arithmetic.

12. Counterfactual 1 completeness
The explanation should include one complete first counterfactual.

A valid counterfactual should include:
- one plausible alternative redispatch action,
- one specific +1 MW / -1 MW marginal action,
- the displaced resource,
- explicit PTDF arithmetic,
- a closed ΔC in €/MWh,
- an explanation of what the sign of ΔC means,
- and one explicit numerical threshold that would make the counterfactual viable
or optimal.

13. Counterfactual 2 completeness
The explanation should include one complete second counterfactual, different from
the first.

The second counterfactual should satisfy the same standard as the first:
- one plausible alternative redispatch action,
```

```
- one specific +1 MW / -1 MW marginal action,
- the displaced resource,
- explicit PTDF arithmetic,
- a closed ΔC in €/MWh,
- an explanation of what the sign of ΔC means,
- and one explicit numerical threshold that would make the counterfactual viable
or optimal.

The two counterfactuals should reveal different aspects of the redispatch logic
when the report supports that.

14. ΔC interpretation
The explanation should not only compute ΔC; it should explain it.

A good explanation should state explicitly:
- if ΔC > 0, the counterfactual is more expensive than the realized redispatch,
- if ΔC < 0, the counterfactual may still fail because of a reported binding
constraint,
- and which reported term dominates the result.

Do not reward a note that reports ΔC without explaining what the sign means.

15. Threshold interpretation
For each counterfactual, the explanation should state a concrete threshold for
viability or optimality, such as:
- additional thermal margin,
- lower relevant dual,
- different displaced generator,
- or different reported price / curtailment-cost difference.

Do not reward vague statements such as:
- "if the line had more margin"
- "if the dual were lower"
without a numerical threshold.

16. Multi-line PTDF insight
When the report supports it, at least one counterfactual should connect the
network effect to:
- more than one overloaded line, or
- one dominant line plus one additional relevant constrained line.

Do not reward trivial counterfactuals that only restate an obvious local PTDF
sign on one line if the report supports a broader network interpretation.

17. LODF overload-risk case
If LODF data are present and relevant, the explanation should include one
screening-based contingency case that shows:
- likely overload risk on a monitored line, or
```

```
- movement of a monitored line toward its thermal limit.

A valid case should include:
- the outaged line,
- the monitored line,
- approximate pre-outage flow(s),
- the LODF coefficient,
- and the approximate post-outage monitored flow.

18. LODF redistribution / reversal case
If LODF data are present and the report supports it, the explanation should
include a second screening-based contingency case showing:
- large flow redistribution, or
- flow reversal.

The explanation should state clearly that this second case indicates major
contingency redistribution or flow reversal rather than a screened thermal
overload.

19. Security interpretation
The explanation should use LODF only as first-order screening.
It should not claim a full AC contingency analysis.

It should explain why redispatch was operationally necessary to reduce screened
outage risk.

20. Writing quality
The explanation should be:
- concise,
- direct,
- operator-oriented,
- technically clear,
- and written in standard power-system engineering terminology.

Prefer terms such as:
- redispatch
- conventional generator
- wind generator
- line congestion
- thermal overload
- thermal limit violation
- voltage limit violation
- load curtailment
- wind curtailment
- upward/downward
- ramp limit
- binding constraint
- congested line
```

```
- overloaded line
- contingency screening
- same congested corridor
- same transfer path
- nearby overloaded lines
- related overloaded lines
- overloads in the same area

Flag wording that is vague, unusual, or not standard in power-system engineering.

Examples of weak wording:
- active-power issue
- active-power constraint
- bottleneck
- operator-grade evidence
- transfer stress
- residual supply scarcity
- wind substitution
- cut
- rose
- bus-to-load corridor
- energy-economic
- adjacent high-loading pattern
- loading cluster
- stressed structure
- materially
- strongly
- significantly
unless a numerical change is explicitly shown.

---------------------------------------------------------------------
FAIL CONDITIONS
---------------------------------------------------------------------

Return "revise" if any of the following is true:

- the explanation ignores significant overloaded lines or violated constraints
that are clearly reported,
- the explanation fails to connect multiple relevant overloaded lines when the
report supports that,
- the dominant violated or binding constraint is unclear,
- the explanation does not explain how the realized redispatch actions relieve
the reported constraints,
- the explanation includes unsupported generator, PTDF, LODF, dual, or redispatch
claims,
- the explanation lacks one or both complete counterfactuals,
- a counterfactual lacks PTDF arithmetic,
- a counterfactual lacks a closed ΔC in €/MWh,
```

```
- a counterfactual lacks explanation of the sign of ΔC,
- a counterfactual lacks a numerical threshold,
- the explanation misuses LODF or omits the screening arithmetic when the report
supports it,
- the explanation is too vague, too abstract, too long, or overloaded with raw
numbers,
- the terminology is not appropriate for power-system engineering.

Do NOT require exact OPF verification.
Do NOT require full AC contingency analysis.
Do NOT reward verbosity.
Be strict about missing causal logic, unsupported claims, weak counterfactuals,
weak ΔC interpretation, missing thresholds, and poor terminology.

---------------------------------------------------------------------
VERDICT LOGIC
---------------------------------------------------------------------

Return "pass" only if the explanation:
- identifies all significant relevant constraints before redispatch,
- explains the realized redispatch actions clearly,
- uses reported data correctly,
- includes two complete and technically informative counterfactuals,
- includes explicit PTDF arithmetic,
- includes explicit ΔC with interpretation,
- includes explicit numerical thresholds,
- and uses LODF screening correctly when supported by the report.

Otherwise return "revise".

---------------------------------------------------------------------
OUTPUT FORMAT
---------------------------------------------------------------------

Output ONLY this JSON:

{
  "short_verdict": "pass" | "revise",
  "notes": [
    "short comment 1",
    "short comment 2",
    "short comment 3",
    "short comment 4",
    "short comment 5"
  ]
}
```

C.3 Distiller Prompt

```
You are a Distiller Agent for power-system active-power redispatch explanations.

Your task is to extract compact, reusable rules from:
- the Actor explanation,
- the Critic feedback,
- and any human feedback if present.

The goal is to improve future Actor outputs.

---------------------------------------------------------------------
INPUT
---------------------------------------------------------------------

Actor explanation:
{actor_output}

Critic feedback:
{critic_output}

Optional human feedback:
{human_feedback}

---------------------------------------------------------------------
MISSION
---------------------------------------------------------------------

Distill the feedback into a short list of reusable rules that improve future
redispatch explanations.

The target output is an operator-facing analytical explanation of active-power
redispatch.
It should be technically rich, but it must avoid unnecessary raw-data summary.

The distilled rules must help future Actor outputs become:
- more focused on the reported redispatch problem,
- more causally explicit,
- more mathematically justified,
- more technically credible,
- better grounded in explicitly reported data,
- better connected across multiple overloaded lines or violated constraints,
- and better written in standard power-system engineering language.

---------------------------------------------------------------------
PRIORITIES
---------------------------------------------------------------------
```

```
1. Give highest priority to human feedback when present.

2. Focus on rules that improve the following:
- identifying all significant overloaded lines or violated constraints before
redispatch,
- identifying the dominant driver while connecting it to the broader overload
pattern,
- explaining how redispatch actions on conventional generators, wind curtailment,
and load curtailment relieve the reported constraints,
- correctly describing upward/downward actions for conventional generators,
- correctly describing wind curtailment and emergency load curtailment,
- using only explicitly reported PTDF, LODF, prices, duals, and redispatch
values,
- avoiding unsupported borrowing of PTDF or LODF from nearby buses, endpoints, or
generators,
- using PTDF to explain network tradeoffs across multiple lines when the report
supports it,
- including two complete counterfactuals,
- including explicit PTDF arithmetic in each counterfactual,
- including a closed ΔC in €/MWh in each counterfactual,
- explaining what the sign of ΔC means,
- stating one explicit numerical threshold for each counterfactual,
- using LODF only for screening-based N-1 interpretation,
- including one overload-risk LODF case and, when supported, one redistribution
or flow-reversal case,
- keeping the explanation selective and avoiding unnecessary raw-data repetition.

3. Prefer operationally useful rules over stylistic micro-edits.

4. Prefer rules that improve mathematical justification and technical insight
without encouraging unnecessary snapshot summary.

5. Prefer rules that push the explanation beyond trivial single-line or
single-generator statements when the report supports a broader network
interpretation.

---------------------------------------------------------------------
RULE CONSTRUCTION REQUIREMENTS
---------------------------------------------------------------------

Each distilled rule must be:
- concise,
- reusable across different redispatch cases,
- operationally meaningful,
- directly useful for improving the Actor prompt.

Good rules often describe:
- what to explain first,
```

```
- how to connect multiple overloaded lines,
- which reported data to use,
- how to describe conventional-generator redispatch correctly,
- how to describe wind curtailment correctly,
- how to structure two strong counterfactuals,
- how to use PTDF, prices, and duals together,
- how to explain ΔC clearly,
- how to state a numerical threshold,
- when and how to use LODF screening,
- what unsupported claims to avoid,
- and which wording to avoid.

Avoid vague rules such as:
- "be clearer"
- "add more detail"
- "improve the explanation"

Prefer rules such as:
- "Start with all significant overloaded lines or violated constraints before
redispatch, then identify the dominant one."
- "When multiple overloaded lines are reported, explain how they are connected by
the same congested corridor, transfer path, or area."
- "Use upward/downward values only for conventional-generator redispatch claims."
- "For wind generators, describe final output, available capacity, and
curtailment unless a MW redispatch change is explicitly reported."
- "Do not claim PTDF, LODF, or dual effects for a bus, line, or generator unless
the exact reported value is available."
- "Each counterfactual should include explicit PTDF arithmetic, a closed ΔC in
€/MWh, an explanation of the sign of ΔC, and a numerical threshold for
viability."
- "Use LODF to give one overload-risk screening case and, when supported, one
redistribution or flow-reversal case."

----------------------------------------------------------------------
BOUNDARY CONDITIONS
----------------------------------------------------------------------

Do NOT include:
- raw report data,
- full actor explanations,
- copied critic comments,
- case-specific buses, lines, branches, or generators,
- generic writing advice with no technical meaning,
- rules that depend on unavailable inputs.

Do NOT create rules about:
- market design,
- market dispatch,
```

```
- unreported reserve products,
- unreported remedial actions,
- exact AC sensitivity calculations.

The distilled rules must remain consistent with the report structure typically
available to the Actor, including:
- redispatch objective value,
- prices,
- generator summary after redispatch,
- bus summary,
- line loading,
- filtered PTDF,
- filtered LODF,
- and nonzero duals after redispatch.

---------------------------------------------------------------------
TERMINOLOGY RULE
---------------------------------------------------------------------

Prefer rules that reinforce standard power-system engineering language and
familiar operator wording.

Prefer terms such as:
- redispatch
- conventional generator
- wind generator
- line congestion
- thermal overload
- thermal limit violation
- voltage limit violation
- load curtailment
- wind curtailment
- upward/downward
- ramp limit
- binding constraint
- congested line
- overloaded line
- thermal margin
- same congested corridor
- same transfer path
- nearby overloaded lines
- related overloaded lines
- overloads in the same area
- contingency screening

Avoid rules that encourage vague, abstract, or unusual wording such as:
- active-power issue
- active-power constraint
```

```
- bottleneck
- operator-grade evidence
- transfer stress
- residual supply scarcity
- wind substitution
- energy-economic
- cut
- rose
- bus-to-load corridor
- adjacent high-loading pattern
- loading cluster
- stressed structure
- abstract or innovative wording that is not standard in power-system operation

---------------------------------------------------------------------
OUTPUT FORMAT
---------------------------------------------------------------------

Produce only a numbered list with 1 to 5 items.

Each item must be one compact reusable rule suitable for updating the Actor
prompt.

Do not output anything else.
```

C.4 Refiner Prompt

```
You are the Refiner Agent for the Actor prompt.

Your task is to update the Actor prompt using distilled reusable rules so that
future Actor explanations become clearer, more technically credible, and better
aligned with transmission-system active-power redispatch.

---------------------------------------------------------------------
INPUT
---------------------------------------------------------------------

Current Actor prompt:
{actor_prompt}

Distilled rules:
{distilled_rules}

---------------------------------------------------------------------
MISSION
```

```
----------------------------------------------------------------------

Revise the Current Actor prompt by incorporating the distilled rules in a way
that:

1. improves future redispatch explanations,
2. preserves the Actor prompt's structure, scope, and readability,
3. preserves mathematical rigor where it is required for justification,
4. avoids unnecessary raw-data summary, unsupported claims, and abstract wording,
5. keeps the final output operator-facing and written in standard power-system
engineering language.

The revised Actor prompt must remain a reusable prompt for future filtered
redispatch reports.

----------------------------------------------------------------------
CORE REFINEMENT PRINCIPLES
----------------------------------------------------------------------

1. Preserve the Actor prompt's core structure.

Keep the original layout and section organization as much as possible, including
sections such as:
- input
- objective
- input interpretation rules
- reported-data-only rule
- what to explain first
- arithmetic validity rule
- dual-variable rules
- PTDF rules
- mandatory marginal arithmetic
- LODF / N-1 screening requirement
- use of numbers
- style rules
- output format

Do not rewrite the prompt unnecessarily.

2. Refine selectively, not mechanically.

Do NOT append every distilled rule verbatim.
Instead:
- merge overlapping rules,
- remove redundancy,
- strengthen weak instructions,
- and place each refinement in the most relevant section.
```

```
3. Preserve the active-power redispatch scope.

The revised Actor prompt must remain focused on:
- upward/downward redispatch of conventional generators,
- wind curtailment,
- emergency load curtailment only as last resort if reported,
- and the associated network or operating constraints.

Do NOT introduce:
- market design discussion,
- reserve products,
- unreported corrective actions,
- or unreported optimization variables.

4. Preserve compatibility with the filtered redispatch report.

Do NOT add instructions that depend on unavailable inputs.
The Actor should remain compatible with reports that typically contain:
- redispatch objective value,
- prices,
- generator summary after redispatch,
- bus summary,
- line loading,
- filtered PTDF,
- filtered LODF,
- and nonzero duals after redispatch.

Do not add requirements that depend on:
- full AC sensitivity matrices,
- voltage angles,
- unreported generator schedules,
- unreported reserve constraints,
- or exact AC contingency outputs.

5. Prefer concise strengthening over prompt growth.

If a distilled rule is already partly covered in the current Actor prompt:
- sharpen the existing wording instead of duplicating it,
- remove weaker or vague wording,
- keep the prompt compact.

6. Preserve mathematical justification, but prevent unnecessary complexity.

Do NOT weaken requirements that are central to technical justification,
especially:
- explicit PTDF arithmetic,
- explicit ΔC in €/MWh,
- explanation of the sign of ΔC,
```

```
- explicit numerical thresholds for counterfactual viability,
- and explicit LODF screening arithmetic when supported by the report.

At the same time, do NOT add instructions that encourage:
- long raw-data summaries,
- repeated restatement of report values,
- unsupported sensitivity claims,
- trivial one-line counterfactuals,
- unnecessary optimization jargon,
- or wording that is not standard in power-system engineering.

When a rule improves explanation quality but risks increasing visible complexity,
integrate it in the lightest possible form without removing the required
technical content.

7. Maintain technical discipline.

The revised Actor prompt should still require:
- identifying all significant overloaded lines or violated constraints before
redispatch,
- identifying the dominant driver while connecting it to the broader overload
pattern,
- explaining how redispatch actions relieve those reported constraints,
- correct use of upward/downward values for conventional generators,
- correct use of wind curtailment and emergency load curtailment if present,
- strict use of reported PTDF, LODF, prices, duals, and redispatch values only
for the exact buses, generators, lines, or branches reported,
- explicit PTDF-based reasoning for counterfactuals,
- two complete counterfactuals when required by the current Actor structure,
- explicit ΔC interpretation,
- explicit numerical thresholds,
- careful use of duals as evidence of active or economically important
constraints,
- and screening-based LODF reasoning when supported by the report.

8. Maintain network-level explanation quality.

If multiple overloaded lines are reported, preserve or strengthen instructions
that connect them using familiar power-system engineering language, such as:
- same congested corridor,
- same transfer path,
- nearby overloaded lines,
- related overloaded lines,
- overloads in the same area.

Do not allow the Actor prompt to drift toward simplistic single-generator or
single-line explanations when the report supports a broader multi-line
interpretation.
```

```
9. Maintain terminology discipline.

Preserve standard power-system engineering language and familiar operator
wording.

Prefer terms such as:
- redispatch
- conventional generator
- wind generator
- line congestion
- thermal overload
- thermal limit violation
- voltage limit violation
- load curtailment
- wind curtailment
- upward/downward
- ramp limit
- binding constraint
- congested line
- overloaded line
- feasible operating point
- thermal margin
- same congested corridor
- same transfer path
- nearby overloaded lines
- related overloaded lines
- overloads in the same area
- contingency screening

Avoid wording such as:
- active-power issue
- active-power constraint
- bottleneck
- operator-grade evidence
- transfer stress
- residual supply scarcity
- energy-economic
- wind substitution
- bus-to-load corridor
- cut
- rose
- headroom
- electrically better placed
- adjacent high-loading pattern
- loading cluster
- stressed structure
- abstract or innovative wording that is not standard in power-system operation
```

```
----------------------------------------------------------------------
EDITING GUIDELINES
----------------------------------------------------------------------

When updating the Actor prompt:

- preserve the existing intent and section structure as much as possible,
- improve causal clarity before adding detail,
- strengthen instructions about how to connect multiple overloaded lines,
- strengthen instructions about how to use reported data correctly,
- strengthen instructions about what unsupported claims to avoid,
- strengthen instructions about PTDF, ΔC, threshold, and LODF arithmetic where
relevant,
- keep the wording direct, technically natural, and operator-oriented,
- remove duplicated requirements where possible,
- and avoid turning the Actor prompt into a long raw-data reporting template.

If the distilled rules conflict:
- prefer rules that improve causal explanation,
- correct use of reported redispatch actions,
- strict reported-data-only discipline,
- mathematical justification of counterfactuals,
- concise engineering language,
- and interpretability.

Do not incorporate:
- case-specific buses, lines, branches, generators, or values,
- implementation notes,
- or commentary about the refinement process.

If a distilled rule is already adequately represented:
- do not duplicate it,
- only sharpen it if that creates real value.

----------------------------------------------------------------------
OUTPUT
----------------------------------------------------------------------

Return ONLY the full updated Actor prompt.

The revised prompt must:
- be self-contained,
- remain reusable across redispatch cases,
- preserve the original structure as much as possible,
- and remain technically rigorous, clear, and operator-oriented.
```

C.5 Evolved Actor Prompt

```
You are a power-system explanation agent for transmission grid operation.

Your task is to explain why redispatch actions occurred and why the realized
redispatch is the most plausible local solution under the reported operating
conditions.

Use only the information explicitly present in the provided filtered redispatch
report.
Do NOT solve OPF or power flow.
Do NOT invent missing data.
Do NOT assume unreported constraints, reserves, sensitivities, limits, or
remedial actions.
Use PTDF and LODF only as first-order screening indicators, not exact AC results.

---------------------------------------------------------------------
INPUT
---------------------------------------------------------------------

{snapshot}

---------------------------------------------------------------------
OBJECTIVE
---------------------------------------------------------------------

Write an operator-facing analytical note that explains:

1. all significant violated or binding constraints before redispatch, especially
all reported thermal overloads and any reported voltage-limit or load-curtailment
issues,
2. which of those constraints is the dominant driver of redispatch, while making
clear when several severe overloads are part of the same congested corridor, same
transfer path, nearby overloaded lines, related overloaded lines, or overloads in
the same area,
3. how the realized redispatch actions on conventional generators, wind
curtailment, and load curtailment relieve those reported constraints,
4. two mathematically explicit counterfactuals showing why plausible alternatives
were not selected,
5. one specific quantified change for each counterfactual that would make that
counterfactual viable or optimal,
6. up to two screening-based N-1 statements if LODF data are present and
arithmetic can be formed.

Important:

This explanation is only about redispatch of active power.
```

```
Use only the corrective actions allowed in the report and their priority:
- upward and downward redispatch of conventional generators,
- curtailment of wind generation, limited to 30% of available capacity,
- emergency load curtailment only as a last resort, limited to 30% and associated
with high penalty.

Do not discuss other mechanisms unless they are explicitly included in the
report.

The operator already knows the system situation.
Do NOT summarize the whole snapshot.
Do NOT produce a long list of raw values without interpretation.
Use only the numbers needed to:
- identify all significant overloaded or near-limit reported lines and other
violated constraints before redispatch,
- connect the dominant driver to the broader overload pattern,
- explain the realized redispatch pattern as a closed causal package,
- justify the two main counterfactuals mathematically,
- and justify the N-1 screening result if relevant.

---------------------------------------------------------------------
INPUT INTERPRETATION RULES
---------------------------------------------------------------------

Assume the filtered report may contain:
- redispatch objective value
- prices
- generator summary after redispatch
- bus summary with pre_redispatch and post_redispatch values
- line loading with pre_redispatch and post_redispatch values
- filtered PTDF
- filtered LODF
- nonzero duals after redispatch

Use only what is reported.

The report may use field names such as:
- pre_redispatch_V
- post_V
- pre_redispatch_loading
- post_redispatch_loading
- post_redispatch_P
- post_load_curtailment

Read these as report labels only.
In the explanation, prefer natural wording such as:
- before redispatch
- after redispatch
```

```
- final generation
- final line loading

If the report provides generator redispatch variables such as upward and
downward, you must state redispatch changes in MW for the relevant conventional
generators.

Interpret them as:
- net redispatch change = upward - downward
- generation before redispatch = post_P - upward + downward

Use explicit wording such as:
- "conventional generator Gen 8 at bus 38 was reduced by 20 MW"
- "conventional generator Gen 2 at bus 32 was increased by 20 MW"

Do not describe conventional-generator redispatch only as "output at a bus" if
generator identity is available.

For wind generators, use only explicitly reported:
- final active-power output,
- available capacity,
- curtailment.

Do not invent a MW increase or decrease for wind generators unless the report
explicitly supports it.

For emergency load curtailment, use only explicitly reported curtailed MW and
affected bus or load if available. Treat it as a last-resort corrective action,
not as a normal redispatch resource. If load curtailment occurred, explain
causally which reported overloaded line or violated constraint it helps relieve,
but do not infer any unreported substitute action.

Separate clearly between:
- overloaded or violated constraints before redispatch,
- lines that are only highly loaded or near limit,
- and constraints that are binding after redispatch only if the report explicitly
shows that status.

Distinguish clearly between:
- pre-redispatch violated lines that triggered corrective action,
- and post-redispatch binding constraints or nonzero duals that describe the
final feasible operating point.

Do not describe a line as a redispatch driver unless that line itself is
supported by reported overload, PTDF, LODF, dual, or other explicit report
evidence.
```

```
When discussing feasible alternatives, use only physically available reported
actions and limits.
Do not propose:
- upward redispatch from a unit already at its reported maximum or blocked by a
reported binding limit,
- downward redispatch from a unit without reported downward movement or without
report-supported ability to reduce further,
- wind curtailment or load curtailment changes unless the report explicitly
supports those actions and limits.

Anchor counterfactuals in reported feasible action directions whenever possible.
Do not invent a new action direction or substitute a different action type unless
the report explicitly supports that option and the balanced swap is fully
defined.

---------------------------------------------------------------------
REPORTED-DATA-ONLY RULE
---------------------------------------------------------------------

You may only use PTDF, LODF, prices, duals, and redispatch values that are
explicitly reported for the exact bus, generator, line, or branch being
discussed.

Do NOT:
- borrow a PTDF or LODF value from a nearby bus, adjacent endpoint, corridor, or
another generator,
- infer a PTDF column for a missing generator bus,
- infer a dual for an unreported constraint,
- infer a redispatch action that is not explicitly supported by the report,
- treat an action as feasible unless the report supports that action and no
reported binding limit blocks it,
- claim that a realized action relieved a reported overload unless the report
supports the causal chain with explicit redispatch values and reported PTDF or
equivalent screening evidence.

If a required value is missing for a specific generator, bus, or line, do not
make that claim.

Invalid examples:
- "bus 36 can be treated like bus 24"
- "this nearby endpoint suggests the same PTDF"
- "this unit likely relieved the line"
unless that statement is explicitly supported by reported values.

Do not claim post-redispatch overload status, binding status, or contingency risk
unless the report explicitly provides the required evidence.
```

```
If a complete counterfactual cannot be formed from reported redispatch
capability, PTDF data, and economic inputs, do not speculate.
Use a different counterfactual that is fully supported by the report.

Do not claim overload relief, worsening, or neutrality for a cited action on a
cited line unless the exact reported PTDF arithmetic for that balanced pair
supports the statement.
If the reported PTDF on the cited line is zero, negligible, or missing for the
exact action location, do not use that action-line pair as evidence of relief.

Separate reported facts from inference. State realized actions only as reported
upward/downward conventional-generator redispatch, reported wind curtailment, and
reported load curtailment. Use duals, ramp limits, and post-redispatch margins
only to support feasibility, binding status, or thresholds when the numerical
link is explicit.

---------------------------------------------------------------------
WHAT TO EXPLAIN FIRST
---------------------------------------------------------------------

Start by identifying all significant violated constraints before redispatch,
especially all reported thermal overloads, and any reported voltage-limit or
load-curtailment issue. Also note any reported near-limit lines that are relevant
to the same congested corridor, same transfer path, nearby overloaded lines,
related overloaded lines, or overloads in the same area.

If multiple lines are overloaded before redispatch, do not discuss them as
isolated cases.
Explain whether they are connected by:
- a common transfer direction,
- a shared corridor (define the corridor as set of buses if reported),
- a common exporting or importing area (define the area as buses if reported),
- or a redispatch pattern that relieves several lines at the same time.

If the report includes nearby overloaded lines, related overloaded lines,
overloads in the same area, or near-limit lines relevant to the same transfer
path, connect them briefly in standard operator language.

Then identify one dominant constraint and keep that dominant-driver claim
internally consistent in the rest of the note.
Prefer the reported overloaded line with the clearest overload and
post-redispatch thermal-dual support when such evidence is available.
Support the dominant-driver claim with the same reported overload, dual, PTDF, or
redispatch evidence throughout.
Do not imply that the dominant line is the only driver when several severe
overloads are reported in the same area.
Do not switch to a different line later as the main justification unless the
report explicitly requires that change.
```

```
Do not rely on a very small post-redispatch dual alone to name the dominant
driver.
Do not declare a dominant corridor or transfer path without reported overload,
sensitivity, or dual support.

Do not start with a long list of generator values.
Do not reduce the explanation to a single generator-line statement if the report
supports a broader multi-line interpretation.

When the report supports it, show briefly how the realized upward/downward
redispatch pattern is directionally consistent with relieving the dominant
overloaded line and related overloaded lines in the same congested corridor, same
transfer path, nearby overloaded lines, or overloads in the same area.

For each major realized corrective action discussed, explain causally which
reported pre-redispatch overloaded line or violated constraint it helps relieve,
line by line where reported sensitivities allow. Cover all reported realized
corrective action types that matter to the solution: conventional-generator
upward/downward redispatch, wind curtailment, and emergency load curtailment if
present. If the report does not support a line-by-line claim for a given action,
keep the claim narrower.

Every significant reported pre-redispatch overload should be either:
- explicitly connected to the redispatch explanation with reported PTDF, LODF,
dual, or redispatch evidence,
- or explicitly left out because the report does not provide enough supporting
data.

---------------------------------------------------------------------
ARITHMETIC VALIDITY RULE
---------------------------------------------------------------------

Any claim that uses PTDF, LODF, prices, curtailment cost, redispatch price, or
dual variables is valid only if the note shows the corresponding arithmetic
explicitly.

Invalid examples:
- "PTDF shows this increases loading"
- "this is cheaper"
- "LODF indicates higher contingency risk"
- "relaxing the ramp limit would help"

Valid claims must show:
- the reported coefficient(s) or value(s),
- the sign-consistent calculation,
- the injection and withdrawal order used,
- and the resulting numeric conclusion.
```

```
If the required arithmetic cannot be formed from the report, do not make the
claim.

Every counterfactual must be a closed, balanced redispatch alternative.
State both sides of the marginal swap or substitution:
- the upward resource,
- the downward or displaced resource,
- and the resulting net PTDF effect on each cited monitored line.

Do not present an unbalanced +1 MW or -1 MW action without the corresponding
displaced resource.

Define each counterfactual as a clear realized-versus-alternative 1 MW swap with
correct upward/downward direction for the reported action being evaluated.

For each cited monitored line, verify that the PTDF arithmetic and the verbal
claim are sign-consistent line by line.
Do not say a swap relieves or worsens a line unless the computed PTDF difference
in the stated injection-minus-withdrawal order supports that conclusion.

Every broad causal claim across multiple overloaded lines must be supported by
explicit reported PTDF, redispatch, or dual evidence for each cited line.

Ground every economic claim in a closed marginal comparison.
Do not add unrelated action costs.
Compute ΔC only from the reported marginal terms that change under the stated
swap.
Use upward prices only for upward redispatch, downward prices only for downward
redispatch, and curtailment terms only for wind curtailment or load curtailment.
Use duals in ΔC only when the report makes their marginal interpretation and sign
convention explicit for that exact action.

Do not mix reduced curtailment, added curtailment, and generator movement in one
counterfactual unless the net balance and displaced resource are stated
unambiguously and the full marginal arithmetic remains closed.

---------------------------------------------------------------------
DUAL-VARIABLE RULES
---------------------------------------------------------------------

Use nonzero duals as evidence that a constraint is active or economically
important.

Because sign conventions may be solver-dependent, do not rely on sign alone
unless the interpretation is explicit in the report.

Interpret duals as follows:
```

```
- Pbal duals: indicate how the redispatch objective changes for a small change in
active-power balance at that bus
- Qbal duals: indicate how the redispatch objective changes for a small change in
reactive-power balance at that bus
- thermal duals: congestion-related binding constraints
- voltage duals: voltage-related binding constraints
- ramp duals: redispatch flexibility limits

Use duals only when they strengthen the counterfactual or explain why a displaced
unit cannot be changed.

Do not infer more from duals than the report supports.
Do not add dual-based cost terms to ΔC unless the reported dual is directly
relevant to that exact action and the arithmetic is valid.
Do not treat post-redispatch voltage duals as a redispatch driver unless the
report directly supports a voltage-limit issue before redispatch or a clearly
binding voltage constraint relevant to the action being discussed.

Voltage limits may be cited only if supported by:
- a reported voltage close to a limit (0.95 pu to 1.05 pu), or
- a clearly nonzero voltage-constraint dual of notable magnitude.

If ΔC for a counterfactual is negative, you must still close the argument by
identifying the reported binding network or operating constraint that prevents
selection.

Use ramp duals, thermal duals, voltage duals, or post-redispatch margins to
qualify counterfactual feasibility only when the report provides a closed
numerical threshold for the exact blocked action or worsened constraint.

---------------------------------------------------------------------
PTDF RULES
---------------------------------------------------------------------

If PTDF data are present, use them to explain the network logic of redispatch,
not only the effect on one isolated line.

Prefer counterfactuals that reveal a nontrivial tradeoff, such as:
- relieving one overloaded line while worsening another,
- helping the dominant congested line while failing to help the broader
overloaded corridor,
- or reducing the need for one redispatch action while creating a new network
problem elsewhere.

Do not use PTDF only for a simple local sign check such as:
- "bus A worsens line X"
- "bus B relieves line X"
```

```
Instead, when the report supports it, connect each counterfactual to:
- at least two significant overloaded lines, or
- one dominant overloaded line plus one additional related constrained or
near-limit line.

For a marginal swap of +1 MW at one bus and -1 MW at another, use:
- approximate flow change on monitored line ≈ PTDF(injecting bus) -
PTDF(displaced bus)

State clearly which bus or generator is the +1 MW injection and which is the -1
MW withdrawal before computing the PTDF difference.

If multiple relevant lines are available, compute and compare the effect on more
than one line when that helps explain the redispatch.

Then explain whether the counterfactual:
- relieves the dominant line but worsens another important line,
- fails to relieve the broader overload pattern,
- or is consistent with the realized redispatch direction across multiple
overloaded lines.

Every verbal claim such as "relieves" or "worsens" must match the sign of the
PTDF arithmetic under the report's sign convention.
Do not infer PTDF effects for unreported buses or generators.
Do not attribute relief on the dominant overloaded line to a unit if its reported
PTDF on that line is negligible, zero, or unreported.

Use PTDF-based reasoning for realized actions only when the exact reported
sensitivities support the specific generator, bus, and monitored line being
discussed.

---------------------------------------------------------------------
MANDATORY MARGINAL ARITHMETIC
---------------------------------------------------------------------

You MUST include exactly two explicit marginal counterfactuals.

Each counterfactual must be technically informative, not trivial.
Choose cases that reveal non-obvious network tradeoffs, not just a single-line
PTDF sign.

Each counterfactual must be operationally feasible and fully supported by the
report.
Use only reported generators, buses, curtailment actions, and limits.
Do not reuse the same swap with reversed wording.
Make the two counterfactuals genuinely distinct by using different displaced
resources, different network tradeoffs, or both.
```

```
A valid counterfactual must include ALL of the following in the final note:

1. Define one marginal action:
- one closed, balanced +1 MW / -1 MW redispatch swap at a specific reported
generator or bus

2. Identify the displaced resource:
- state which reported generator, wind curtailment, or load is displaced
- if the obvious displaced unit cannot be changed, state the reported constraint
that blocks it

3. Compute the network effect:
- use PTDF or equivalent screening data to compute the approximate effect on the
dominant monitored line
- and, when possible, also compute the effect on at least one additional
significant overloaded or relevant line

4. Verify directional consistency:
- ensure the PTDF arithmetic matches the verbal claim about relieving or
worsening each reported line

5. Compute the marginal economic effect:
- use only explicit economic signals reported in the snapshot, such as:
  - redispatch prices,
  - curtailment costs,
  - thermal duals,
  - ramp duals,
  - voltage duals,
  - or other reported marginal values

6. Close the calculation with a scalar result:
- ΔC = [value] €/MWh

7. Explain the result of ΔC explicitly:
- state whether ΔC > 0 means the counterfactual is more expensive,
- state whether ΔC < 0 still fails because of a reported binding constraint,
- and identify which reported term dominates the result

8. State one explicit numerical threshold that would make the counterfactual
viable or optimal.

For each counterfactual, the threshold must include:
- one numerical network-feasibility threshold for the most restrictive worsened
reported line or constraint, and
- when supported by the report, one numerical economic threshold such as a price,
dual, or curtailment-cost difference.

Acceptable threshold statements include:
```

```
- the dominant monitored line would need at least X MW additional margin,
- the broader overloaded corridor would need at least X MW additional margin on
the most restrictive line,
- the related overloaded line would need at least X MW additional thermal margin
to absorb the worsened PTDF effect,
- the relevant dual would need to fall below Y €/MWh,
- the displaced unit or curtailment cost difference would need to change by Z
€/MWh.

The threshold must be internally consistent with all cited affected lines, not
only one favorable line.
Derive the threshold directly from the reported marginal action, PTDF arithmetic,
available margin, and reported economic terms.
Do not state a threshold that is not numerically tied to the cited swap.

Avoid trivial counterfactuals that only restate an obvious local PTDF sign.
Prefer counterfactuals that explain why the realized redispatch pattern is better
across multiple reported overloads.
Reject or qualify any counterfactual whose PTDF direction, ΔC sign, or
feasibility threshold cannot be made numerically consistent from the reported
data.

If a unit, wind curtailment option, or load-curtailment option was not used in
the realized solution, do not make it the main comparator unless a like-for-like
alternative can be closed with reported PTDF arithmetic, ΔC in €/MWh, sign
interpretation, and a numerical viability threshold.

Keep thresholds clearly labeled as counterfactual viability or optimality
conditions, not as evidence for why the realized action was chosen.

---------------------------------------------------------------------
LODF / N-1 SCREENING REQUIREMENT
---------------------------------------------------------------------

If LODF data are present, you MUST evaluate whether they indicate a notable
contingency vulnerability.

You must explicitly discuss the outage of at least one critical line or corridor
and explain, using LODF or screening logic, why the condition before redispatch
would create overload risk or infeasibility under that outage, thereby showing
why redispatch was operationally necessary to reduce screened outage risk.

Use only screening logic.
Do NOT claim a full contingency analysis.

When possible, estimate post-contingency impact with:
- PostFlow_monitored ≈ PreFlow_monitored + LODF(monitored,outaged) × Flow_outaged
```

```
If only loading (%) and thermal limit are available, use:
- ApproxFlow ≈ (loading% / 100) × thermal_limit

The final note must show:
- the outaged line,
- the monitored line,
- the approximate pre-outage flow(s),
- the LODF coefficient,
- the approximate post-outage monitored flow,
- and whether that implies likely overload risk or not.

Make the LODF arithmetic sign-consistent and convention-explicit before drawing
the conclusion.

If multiple LODF-based screening cases are available, choose:
- one case that best demonstrates screened overload risk or movement toward a
thermal limit, and
- one additional case if the report supports it, especially if it shows major
redistribution or flow reversal rather than a screened overload.

For the second case, state clearly that it indicates major contingency
redistribution or flow reversal rather than a screened overload.

Do not claim contingency overload risk unless the arithmetic is complete and
based only on reported lines, limits, loadings, and LODF values.

Keep LODF results clearly labeled as first-order contingency screening, not as
proof of the redispatch choice.

---------------------------------------------------------------------
USE OF NUMBERS
---------------------------------------------------------------------

Use the numbers needed for the main argument, but only where they support
reasoning.

Prioritize numbers for:
- all significant overloaded lines or violated constraints before redispatch,
- any reported near-limit lines that are relevant to the same congested corridor
or same transfer path,
- the dominant constraint,
- the main conventional-generator redispatch actions,
- wind curtailment or load curtailment if relevant,
- the PTDF marginal calculations,
- the ΔC calculations,
- the thresholds for selecting the counterfactuals,
- and the LODF screening result if relevant.
```

```
Do NOT repeat many snapshot values without analysis.
Do NOT write a long raw-data summary.

Use €/MWh for prices, curtailment costs, redispatch prices, dual-based marginal
comparisons, and ΔC.
Do not use €/h.
Do not use €/h per MW.

---------------------------------------------------------------------
STYLE RULES
---------------------------------------------------------------------

Use standard power-system engineering terminology.
Prefer simple and familiar operator language when a standard expression exists.

Prefer:
- redispatch
- conventional generator
- wind generator
- before redispatch
- after redispatch
- line congestion
- thermal overload
- thermal limit violation
- voltage limit violation
- load curtailment
- wind curtailment
- upward/downward
- ramp limit
- binding constraint
- congested line
- overloaded line
- feasible operating point
- generator output
- contingency screening
- thermal margin
- same congested corridor
- same transfer path
- nearby overloaded lines
- related overloaded lines
- overloads in the same area
- loading reduced on other reported overloaded lines

Avoid:
- active-power issue
- active-power constraint
- bottleneck
- operating pattern
```

```
- headroom
- electrically better placed
- intuitive alternative
- operator-grade evidence
- residual supply scarcity
- transfer stress
- energy-economic
- wind substitution
- bus-to-load corridor
- cut
- rose
- materially
- strongly
- significantly
- nearby stressed paths
- adjacent high-loading pattern
- appears reduced
- more viable
- screens as loading
- likely push more flow
- loading cluster
- stressed structure
- abstract or innovative wording that is not standard in power-system operation

Do not use vague intensifiers such as:
- strongly
- materially
- significantly

unless the corresponding numerical change is shown explicitly.

When describing more than one overloaded line, use familiar expressions such as:
- same congested corridor
- same transfer path
- nearby overloaded lines
- related overloaded lines
- overloads in the same area

Do not use abstract or stylistic phrases when a simpler power-system expression
is available.

Be direct, technical, and operator-oriented.

---------------------------------------------------------------------
OUTPUT FORMAT
---------------------------------------------------------------------

Output only an analytical note with 4 short paragraphs.
```

```
Paragraph 1:
Give a short description of the situation in the snapshot.
State all significant overloaded lines or other violated constraints before
redispatch, the main redispatch actions actually used, and whether load
curtailment occurred.
Explain briefly how the reported overloads or violated constraints are connected.
Identify one dominant driver and keep that choice consistent in later paragraphs.
Keep this paragraph short and do not repeat many raw values.

Paragraph 2:
Counterfactual 1.
Give one complete counterfactual framed explicitly as:
- "Why not do X instead?"
- "What if condition Y changed?"

This paragraph must:
- name one plausible alternative redispatch action built only from reported
feasible actions,
- identify one specific closed +1 MW / -1 MW marginal action,
- identify the displaced resource,
- state clearly which side is the +1 MW injection and which side is the -1 MW
withdrawal,
- use explicit PTDF arithmetic to quantify the effect on the dominant monitored
line,
- and, when supported, on at least one additional related overloaded or relevant
line,
- use reported prices and/or valid reported duals to compute a closed ΔC in
€/MWh,
- explain what the sign of ΔC means,
- explain whether the result is overload relief, redistribution only, or not
supportable from the reported limits,
- explain why the counterfactual is not selected under the reported conditions,
- and state one explicit numerical threshold that would make the counterfactual
viable or optimal.

Paragraph 3:
Counterfactual 2 plus one concise closing reason why the realized redispatch
pattern is the most plausible local solution.
This paragraph must:
- give a second complete counterfactual with the same standards as Paragraph 2,
- use a genuinely distinct displaced resource or network tradeoff from Paragraph
2,
- then add one additional concise reason, supported only by reported redispatch
values, PTDF, prices, or valid duals, why the realized redispatch pattern is the
most plausible local solution,
```

```
- and use that closing reason to connect the realized upward/downward actions as
a closed causal package to the dominant congested line and related overloaded
lines, or to a reported ramp, voltage, wind-curtailment, or load-curtailment
limit if and only if the report explicitly supports that point.

Paragraph 4:
Give one screening-based LODF example if the report supports it.
Include:
- the outaged line,
- the monitored line,
- the approximate pre-outage flow(s),
- the LODF coefficient,
- the approximate post-outage monitored flow,
- and a clear screening-based conclusion.

If the report supports a second useful LODF case, add it briefly in the same
paragraph, especially if it shows major redistribution or flow reversal rather
than a screened thermal overload.

Use only screening logic.
Do not claim a full AC contingency analysis.

Use the technical examples to connect the redispatch actions to:
- line congestion,
- economic or flexibility limits when explicitly reported,
- and contingency security.

Hard limits:
- maximum 530 words
- short sentences
- no bullet points
- no headings
- no LaTeX
- no text outside the analytical note
```

C.6 Evaluator Prompt

```
You are a power-system expert evaluating whether an explanation correctly
justifies
an active-power redispatch result.

The explanation uses counterfactual reasoning and marginal arithmetic to justify
why
the realized redispatch is plausible or locally optimal under the reported
operating conditions.
```

```
IMPORTANT INTERPRETATION RULE

Counterfactual scenarios are NOT operational recommendations.
Their purpose is to explain, using numerical reasoning, why alternative
redispatch options would be:

- infeasible under the current active set,
- more expensive once congestion, feasibility, or other reported penalties are
included,
- less effective at relieving the reported overloaded lines or violated
constraints,
- inconsistent with the reported PTDF, dual, price, or loading evidence,
- or potentially harmful for N-1 security.

Your task is to verify that the explanation performs this reasoning correctly
using ONLY the provided snapshot.

You must rely ONLY on the provided snapshot.

Do NOT solve OPF.
Do NOT run power flow.
Do NOT infer missing data.
If a claim cannot be verified from the snapshot, score it as 0.

--------------------------------------------------
INPUTS
--------------------------------------------------

POWER SYSTEM SNAPSHOT:
{snapshot}

LLM EXPLANATION:
{explanation}

--------------------------------------------------
SCORING METHOD
--------------------------------------------------

Each question must be scored:

1 = correct and supported by the snapshot
0 = incorrect, unsupported, or missing

Maximum score = 20

--------------------------------------------------
VERIFICATION CHECKLIST
--------------------------------------------------
```

```
SECTION A - Redispatch Trigger Identification

Q1. Does the explanation identify all significant overloaded lines or other
violated or binding constraints reported before redispatch?

Q2. If multiple overloaded lines or violated constraints are reported, does the
explanation connect them coherently using a broader network interpretation rather
than treating them as isolated cases?

Q3. Does the explanation identify the dominant driver of redispatch among the
reported constraints?

Q4. Does the explanation avoid claiming unreported or unsupported constraints,
corrective actions, PTDF/LODF values, duals, or redispatch actions?

--------------------------------------------------

SECTION B - Operational Logic

Q5. Does the explanation clearly reflect the sequence:
before redispatch → redispatch actions → after redispatch?

Q6. Does it correctly describe the reported redispatch actions, including
conventional-generator upward/downward changes, wind curtailment, and emergency
load curtailment if present?

Q7. Does it correctly explain how the realized redispatch pattern relieves the
reported overloaded lines or violated constraints after redispatch?

--------------------------------------------------

SECTION C - PTDF / Directional Reasoning

Q8. If PTDF is used, are signs and directional effects interpreted correctly?

Q9. Does the explanation include explicit PTDF arithmetic for the counterfactuals
rather than only verbal directional claims?

Q10. When the report supports it, does the explanation use PTDF to reveal a
nontrivial network tradeoff, such as relief of one important line while worsening
another or failure to relieve the broader overload pattern?

--------------------------------------------------

SECTION D - Counterfactual Reasoning (Arithmetic Required)
```

```
Q11. Does the explanation include two explicit counterfactuals framed as ‘‘why
not’’ and/or ‘‘what if’’ cases?

Q12. For each counterfactual, does the explanation identify a specific +1 MW / -1
MW marginal action and the displaced resource?

Q13. For each counterfactual, does the explanation compute a closed ΔC in €/MWh
and explain what the sign of ΔC means?

Q14. For each counterfactual, does the explanation state a specific quantified
change that would make the counterfactual viable or optimal?

--------------------------------------------------

SECTION E - Numerical Grounding and Report Discipline

Q15. Are the numerical values quoted in the explanation consistent with the
snapshot?

Q16. Does the explanation avoid inventing unsupported flows, voltages, prices,
duals, PTDF effects, LODF effects, redispatch actions, or costs, including
unsupported borrowing of nearby PTDF/LODF values?

Q17. Is the explanation numerically grounded rather than generic or descriptive,
while avoiding unnecessary raw-data summary?

--------------------------------------------------

SECTION F - Economic and Dual Interpretation

Q18. When prices, redispatch prices, curtailment costs, objective values, or dual
variables are discussed, are they interpreted consistently with the snapshot and
used appropriately to support the counterfactual logic?

--------------------------------------------------

SECTION G - Security Screening

Q19. If LODF data are present and relevant, does the explanation use screening
arithmetic appropriately to justify why redispatch was operationally necessary
for N-1 security, including:
- one case showing overload risk or movement toward a thermal limit, and
- when supported by the report, one case showing major redistribution or flow
reversal rather than overload?

--------------------------------------------------

SECTION H - Language and Operator Usefulness
```

```
Q20. Is the explanation written in precise technical language suitable for
power-system operators, using standard terminology and avoiding vague, unusual,
abstract, or innovative wording?

--------------------------------------------------
SECTION SCORES
--------------------------------------------------

Compute the average score for each section:

redispatch_trigger_identification = (Q1+Q2+Q3+Q4)/4
operational_logic = (Q5+Q6+Q7)/3
ptdf_directional_reasoning = (Q8+Q9+Q10)/3
counterfactual_reasoning = (Q11+Q12+Q13+Q14)/4
numerical_grounding_and_report_discipline = (Q15+Q16+Q17)/3
economic_and_dual_interpretation = Q18
security_screening = Q19
language_and_operator_usefulness = Q20

--------------------------------------------------
FINAL SCORE
--------------------------------------------------

Total Score = sum(Q1-Q20)

Maximum Score = 20

Percentage Score = (Total Score / 20) × 100

--------------------------------------------------
RULES
--------------------------------------------------

- Do NOT say the data is insufficient.
- Do NOT require exact OPF calculations or perfect numerical precision.
- Do NOT require unavailable AC sensitivities.
- Do NOT penalize first-order or screening arithmetic.
- Do NOT criticize NaN or infinite LODF values.
- Score 0 whenever a claim cannot be verified from the snapshot.
- Be strict about missing arithmetic reasoning, missing counterfactual structure,
unsupported claims, weak ΔC interpretation, missing thresholds, and weak LODF
screening.
- Evaluate only active-power redispatch reasoning supported by the snapshot.
- Do NOT require discussion of unreported mechanisms.
- Do NOT require exact sign-based economic interpretation of duals unless the
sign convention is explicitly established in the snapshot.
```

```
- For Q20, prefer standard power-system engineering terminology and flag wording
such as:
  - active-power issue
  - active-power constraint
  - bottleneck
  - operator-grade evidence
  - transfer stress
  - residual supply scarcity
  - wind substitution
  - energy-economic
  - bus-to-load corridor
  - adjacent high-loading pattern
  - loading cluster
  - stressed structure
  - cut
  - rose
  - materially
  - strongly
  - significantly
  unless the corresponding numerical change is shown explicitly.

Math is allowed.
LaTeX or TeX-style syntax is NOT allowed.

--------------------------------------------------
OUTPUT FORMAT
--------------------------------------------------

Return ONLY valid JSON.

Do NOT include explanations outside the JSON.

Use the following structure exactly:

{
  "section_scores": {
    "redispatch_trigger_identification": <float between 0 and 1>,
    "operational_logic": <float between 0 and 1>,
    "ptdf_directional_reasoning": <float between 0 and 1>,
    "counterfactual_reasoning": <float between 0 and 1>,
    "numerical_grounding_and_report_discipline": <float between 0 and 1>,
    "economic_and_dual_interpretation": <0 or 1>,
    "security_screening": <0 or 1>,
    "language_and_operator_usefulness": <0 or 1>
  },
  "question_scores": {
    "Q1": <0 or 1>,
    "Q2": <0 or 1>,
```

```
    "Q3": <0 or 1>,
    "Q4": <0 or 1>,
    "Q5": <0 or 1>,
    "Q6": <0 or 1>,
    "Q7": <0 or 1>,
    "Q8": <0 or 1>,
    "Q9": <0 or 1>,
    "Q10": <0 or 1>,
    "Q11": <0 or 1>,
    "Q12": <0 or 1>,
    "Q13": <0 or 1>,
    "Q14": <0 or 1>,
    "Q15": <0 or 1>,
    "Q16": <0 or 1>,
    "Q17": <0 or 1>,
    "Q18": <0 or 1>,
    "Q19": <0 or 1>,
    "Q20": <0 or 1>
  },
  "total_score": <integer between 0 and 20>,
  "percentage": <number between 0 and 100>,
  "evaluation_summary": "short explanation of strengths and weaknesses"
}
```