



**BIRMINGHAM CITY
University**

EXPLAINABLE
DEEP REINFORCEMENT LEARNING
FOR MICROGRID
RESILIENT ENERGY MANAGEMENT

By

MOHAMMAD HOSSEIN NEJATI AMIRI 

A thesis submitted as part of the requirements for
the degree of Doctor of Philosophy (Ph.D.) in Electrical Engineering
at the Faculty of Computing, Engineering and the Built Environment (CEBE)
Birmingham City University, Birmingham, UK

February 2026

Supervisor(s) Dr. Florimond Guénat , Dr. Mario De Oliveira , and Dr. Fawaz Annaz 

ABSTRACT

Microgrids enable the integration of distributed renewable generation and can support decarbonisation, particularly in rural or weakly connected areas where main grid access is limited or unreliable. This thesis focuses on a standalone rural microgrid in a cyclone-prone coastal region, where weather-driven High-Impact, Low-Probability (HILP) events are the main resilience concern. In this context, HILP events can substantially reduce renewable generation, interrupt normal operation, and increase the risk of unserved priority loads. The studied system relies on photovoltaic generation, wind generation, and a Battery Energy Storage System (BESS), with the battery treated as the only energy storage and fault-ride-through resource during disruptive periods. The central challenge is therefore to keep critical loads energised during weather-related HILP events while limiting battery ageing and providing explanations that operators can trust. This thesis examines whether an interpretable, data-driven controller can match a tuned Model Predictive Control (MPC) benchmark with lower computational cost and improved battery preservation within this defined standalone microgrid context.

A cyclone-prone coastal site in India is adopted as the case study, with asset sizing derived in HOMER Pro from historical demand and weather records that include the Cyclone Laila period to ensure realistic constraints. The benchmark is a multi-objective MPC formulation implemented as a mixed-integer linear programme in Pyomo. It balances resilience, power imbalance, switching effort, and life-cycle degradation through adaptive State Of Charge (SOC) limits and optional grid trading. The proposed alternative is a Deep Rein-

forcement Learning (DRL) policy based on Proximal Policy Optimisation (PPO), trained with a four-stage curriculum that progressively increases renewable perturbations and scenario difficulty, and interpreted post hoc with SHapley Additive exPlanations (SHAP) and Local Interpretable Model-Agnostic Explanations (LIME). Robustness is assessed through Monte Carlo experiments under smoothly correlated Perlin noise uncertainty.

Under deterministic conditions, the learned policy achieves a Resilience Index (RI) of 0.996 versus 0.999 for MPC and extends battery life Expected Year (EY) from 14.44 to 15.90 years by smoothing charge-discharge behaviour. The rate of change and the acceleration of SOC fall by 12–19% and 26–38%, respectively. Under uncertainty, resilience remains within ~ 0.3 percentage of MPC (RI 0.996 versus 0.999) and EY improves by $\sim 5\%$ (15.88 versus 15.06 years). Across 4,000 runs, performance concentrates tightly with mean RI 0.992 and EY 15.95 years. Inference is $5,000\times$ lighter per step than MPC, and lifetime compute is almost $3\times$ lower, indicating suitability for edge deployment. Explanations agree with power system intuition: charge under surplus, discharge under deficit, and use medium-priority loads as the primary source of flexibility.

Keywords: microgrid, resilience, battery degradation, model predictive control, deep reinforcement learning, explainability (SHAP, LIME).

DEDICATION

First and foremost, I want to thank my parents (Mohammad Kazem and Fariba). Their love and support have meant everything to me. From the very beginning, they've always believed in me, offering both emotional and financial support that has made this journey possible. Without their encouragement and sacrifices, none of this would have been possible. I am profoundly grateful for their unwavering support, which made this journey possible.

To my wife, Fatemeh, thank you for being by my side through these three years, despite the distance between us. Your patience, understanding, and constant support have kept me going. Every day, your belief in me pushed me to work harder and finish sooner. I feel incredibly lucky to have you as my partner, and I couldn't have done this without you.

Finally, I want to thank my Director of Studies, Florimond Guéniat. Your guidance and support throughout my Ph.D. journey have been invaluable. You've been a constant source of encouragement, helping me stay on track and guiding me when I needed it most. I extend my sincere gratitude for your continuous guidance and invaluable advice throughout this PhD.

ACKNOWLEDGMENTS

I would like to express my heartfelt gratitude to Dr. Sandeep Dhundhara for the fruitful collaboration throughout my PhD. His expertise, support, and insightful discussions have been incredibly valuable in shaping my research.

I am also deeply grateful to Birmingham City University for providing me with the opportunity to teach and interact with students. This experience has not only enriched my academic journey but also allowed me to grow as an educator.

A special thanks to the SmartSIP+ project team for allowing me to be a part of their impressive initiative. Although my time with them was short, being involved in such a large and impactful project was an invaluable experience, and I truly appreciate the opportunity to contribute.

I would like to acknowledge my PhD colleagues, with whom I've had the privilege of learning and growing. Their friendship, support, and intellectual exchanges have been integral to my research journey.

I also want to extend my thanks to the prestigious journals, from IEEE to Elsevier, for giving me the opportunity to review their papers. This experience has been invaluable in helping me learn the details of high-quality journal submissions and has significantly contributed to my understanding of academic publishing.

PUBLICATIONS RELATED TO THIS THESIS

The publications listed below are based on research conducted during the course of this PhD programme. They support the conceptual foundations, methodological development, and evaluation presented in this thesis. In particular, the review of resilience metrics informs Chapter 2; the MPC-based studies contribute to the modelling and optimisation framework developed in Chapters 3 and 4; and the DRL and explainability-focused works underpin the learning-based framework and evaluation presented in Chapters 3 and 5.

Published and Accepted Works

1. **Mohammad Hossein Nejati Amiri**, Fawaz Annaz, Mario De Oliveira, and Florimond Guéniat. *Strategies for Resilience and Battery Life Extension in the Face of Communication Losses for Isolated Microgrids*. In *Proceedings of the IEEE International Conference on Omni-layer Intelligent Systems (COINS)*, London, UK, 2024. (Contributes to Chapters 3 and 4)
2. **Mohammad Hossein Nejati Amiri** and Florimond Guéniat. *Towards a framework for measurements of power systems resiliency: Comprehensive review and development of graph and vector-based resilience metrics*. *Sustainable Cities and Society*, vol. 109, p. 105517, 2024. (Contributes to Chapter 2)
3. **Mohammad Hossein Nejati Amiri**, Sandeep Dhundhara, Fawaz Annaz, Mario De Oliveira, and Florimond Guéniat. *Two-stage microgrid resilience and battery life-aware planning and operation for cyclone-prone areas in India*. *Sustainable Cities and*

Society, vol. 124, p. 106290, 2025. (Contributes to Chapters 3 and 4)

4. **Mohammad Hossein Nejati Amiri**, Fawaz Annaz, Mario De Oliveira, and Florimond Guéniat. *Deep Reinforcement Learning with Local Interpretability for Transparent Microgrid Resilience Energy Management*. In *Proceedings of the V International Conference on Electrical, Computer and Energy Technologies (ICECET)*, Paris, France, 2025. (Contributes to Chapters 3 and 5)
5. **Mohammad Hossein Nejati Amiri** and Florimond Guéniat. *Quicker to Train and Cheaper to Run: Edge-Ready DRL for Microgrid Resilient Energy Management*. In *Proceedings of the 9th International Conference on Environment Friendly Energies and Applications (EFEA)*, Newcastle upon Tyne, UK, 2025. (Contributes to Chapters 3 and 5)
6. **Mohammad Hossein Nejati Amiri** and Florimond Guéniat. *Explainable Deep Reinforcement Learning for Resilient and Battery-Aware Microgrid Control*. *Energy Conversion and Management*, vol. 353, p. 121215, 2026. (Contributes to Chapters 3 and 5)

Manuscripts Under Review or In Preparation

1. **Mohammad Hossein Nejati Amiri**, Fawaz Annaz, Mario De Oliveira, and Florimond Guéniat. *Ethics and Explainability in AI: A Focus on Explainable Deep Reinforcement Learning*. Book chapter, Springer, in preparation, 2025. (Conceptually informs Chapters 2 and 5)

Contents

	Page
1 Introduction	1
1.1 Background and Motivation	3
1.2 Challenges: Uncertainty, Battery Ageing, and HILP Events	5
1.3 Problem Statements	7
1.4 Research Questions	8
1.5 Research Aim and Objectives	9
1.6 Scope and Boundaries	10
1.7 Thesis Structure	11
2 Literature Review	14
2.1 Thematic Structure of the Review	14
2.2 Concept of resilience in power systems	15
2.2.1 Motivation and definitions	15
2.2.2 Brief note on literature selection and context	21
2.2.3 Framework, applications, and metric taxonomy	23
2.3 Review of control techniques in microgrids	31
2.3.1 Model Predictive Control	32
2.3.2 Deep Reinforcement Learning	36
2.3.2.1 DRL in Microgrid Resilience	43
2.4 Explainable AI with a focus on XDRL	51

2.5	Research Gaps and Thesis Contributions	56
3	Research Methodology and Case Study	60
3.1	Research Philosophy and Approach	60
3.2	Methodological Approach	61
3.3	Simulation-Based Modelling Strategy	64
3.3.1	System Overview and Microgrid Architecture	65
3.3.2	Case Study Site: Kothapatnam, India	68
3.4	Software, Tools, and Modelling	73
3.4.1	HOMER Pro Modelling: Formulations, Assumptions, and Planning Results	74
3.4.1.1	Solar Photovoltaic Generation	74
3.4.1.2	Wind Turbine Generator	76
3.4.1.3	Bidirectional Converter	77
3.4.1.4	Battery Energy Storage	77
3.4.1.5	Grid Interaction	78
3.4.1.6	Economic Metrics	79
3.4.1.7	Planning Assumptions	79
3.4.1.8	Planning Results and Discussion	81
3.4.1.9	Static comparison with a diesel-supported ride-through option.	85
3.4.2	Model Predictive Control Model (PYOMO Implementation)	90
3.4.2.1	Principles of MPC	90
3.4.2.2	From MPC to MILP at Each Time Step	91
3.4.2.3	Smart Resilient Control Flowchart	108
3.4.3	Deep Reinforcement Learning Modelling (PPO SB3 and Gymnasium Implementation)	110
3.4.3.1	PPO Foundation	110

3.4.3.2	Policy Gradient Methods	110
3.4.3.3	Trust Region Policy Optimization	112
3.4.3.4	Proximal Policy Optimization	113
3.4.3.5	PPO Algorithm and Implementation Framework	115
3.4.3.6	Gymnasium Environment and Decision Model	117
3.4.3.7	Microgrid Gymnasium Environment Model	118
3.4.4	Shap and Lime Python Libraries for Explainability	123
3.4.4.1	Local Interpretable Model-agnostic Explanations (LIME) . .	125
3.4.4.2	SHapley Additive exPlanations (SHAP)	128
3.5	Data Preparation and Uncertainty Modelling	130
3.5.1	Training Dataset Construction via Controlled Perturbations	131
3.5.1.1	Dataset perturbation and assembly	131
3.5.2	Post-Training Uncertainty Injection with Perlin Noise	138
3.6	Computational Cost Evaluation Protocol	141
3.7	Summary of Methodology and Transition to Results	142
4	Model Predictive Control Framework	144
4.1	Introduction	144
4.2	MPC Operational Results and Discussion	145
4.2.1	Model Validation Against HOMER Pro	145
4.2.2	Scenario Design and Classification	147
4.2.3	Sensitivity to Objective Weighting	147
4.2.4	Overall Optimisation Performance	150
4.2.5	Sensitivity of MPC Results to BLC Linearisation Accuracy	153
4.2.6	Impact of Demand Response	158
4.2.7	Smart Resilient Control under Reduced Battery Capacity	160
4.2.8	Grid-Connected Scenarios and Profit Analysis	162

4.3	Discussion and Conclusion	166
5	Deep Reinforcement Learning and Explainability Framework	168
5.1	Introduction	168
5.2	Performance Evaluation and Comparative Results	169
5.2.1	Deterministic Evaluation	169
5.2.2	MPC Baseline under Imperfect Forecasts	172
5.2.3	Performance under Renewable Uncertainty	174
5.2.4	Stage-wise SOC Analysis and Derivative Comparison	176
5.2.5	Computational Cost Benchmarking	177
5.3	Monte Carlo-Based Uncertainty Quantification and Robustness Analysis . .	180
5.4	Explainability Assessment	188
5.4.1	Global Feature Importance using SHAP	189
5.4.2	Local Decision Interpretation using LIME	194
5.4.3	Cross-Validation between SHAP and LIME Insights	199
5.4.4	Implications for Operator Trust and Responsible Deployment	202
5.4.4.1	Safeguards considered in this thesis	202
5.4.4.2	Practical guidelines and future extensions	202
5.5	Chapter Summary	205
6	Conclusion and Future Work	206
6.1	Conclusion	206
6.2	Limitations and Future Directions	209
6.2.1	Limitations	209
6.2.2	Future Directions	211
	References	214

List of Figures

1.1	Total disaster costs by decade with polynomial fit (1980s–2010s). The 2020–2024 period is omitted because it does not cover a full decade [1].	2
1.2	Share of total cost by disaster type (1980–2024), excluding Freeze for legibility. Values are shown in USD billions [1].	2
1.3	Challenges for microgrids and why they matter.	6
1.4	Logical mapping from Research Questions (RQ) to Objectives (Obj) and Chapters (Ch).	13
2.1	The underlying origins and types of HILP events [2].	17
2.2	Context diagram linking common HILP event types with affected microgrid components, stakeholders, and resilience concerns.	18
2.3	Methodology for selecting and including resilient metric articles.	21
2.4	a): Number of articles obtained via bibliometric analysis, published by countries. b): Global hazard total economic loss risk distributions (adapted from [27], [28]). Colour encodes the magnitude of losses.	22
2.5	Resilience Framework illustrating the key stages of Absorb, Survive, Recover and Adapt, with a specific emphasis on Resilience Metrics [2].	24
2.6	Applications of resilience metrics [2].	25
2.7	Categories of resilience metrics and associated percentage of reviewed works [2].	27
2.8	Evolution of resilience curves.	27

2.9	Comparison of the proposed resilience curve with real HILP event (Hurricane Harvey).	28
2.10	A comparison of different classical control methods.	35
2.11	Agent and Environment Interactions.	36
2.12	Popular Techniques to Solve MDPs.	39
2.13	Popular Deep Reinforcement Learning Techniques.	41
2.14	Applications of Deep Reinforcement Learning.	42
2.15	Distribution of algorithms used in DRL-based microgrid resilience studies. Percentages reflect algorithm mentions across the reviewed literature (Table 2.6). (Unit: individual algorithm mentions; if a paper used multiple algorithms, each was counted once).	51
2.16	XDRL agents speak: IA on simplicity, IE on design integration, PHE on insightful analysis.	53
2.17	Accuracy versus explainability across XDRL families. PPO as high-accuracy control is applied and post-hoc explanations is adopted to meet transparency needs without degrading performance.	56
3.1	Methodology structured by research questions (RQ1–RQ5). Each stream follows a waterfall process of literature review, data preparation, model development, evaluation, and documentation, contributing to the overall thesis aim.	61
3.2	Microgrid planning stage model in HOMER Pro.	66
3.3	Microgrid operational stage models: (a) MPC design and (b) XDRL design.	67
3.4	Annual renewable resource inputs (NASA POWER) for the selected site in Andhra Pradesh, used in the planning stage.	70
3.5	Average daily load profile used as a planning input.	70
3.6	Tropical Cyclone Laila soaking southeastern India, observed on 19 May 2010, showing rainfall distribution and storm track [175].	71

3.7	Load demand types during Cyclone Laila (19–20 May 2010) for the selected site.	71
3.8	Raw weather data and renewable generation profiles during Cyclone Laila (16–25 May 2010) obtained from PVGIS and processed using <code>pvlib</code> and <code>windpowerlib</code>	72
3.9	Monthly renewable energy production (PV and WT). The profile exhibits strong complementarity: higher WT generation during early monsoon and higher PV generation in shoulder and dry months.	82
3.10	Battery utilisation over one year (SOC heat map).	83
3.11	TNPC, initial cost, and COE sensitivity to generator capital costs.	84
3.12	TNPC, initial cost, and COE sensitivity to battery cost and sell-back price.	85
3.13	Sliding-window MPC with a 24-hour prediction horizon.	92
3.14	Battery life versus depth of discharge: (a) non-linear curve, (b) Piecewise-linear BLC–DOD approximation.	98
3.15	Smart resilient control flowchart for MPC-based microgrid operation.	109
3.16	Procedure for creating and using a custom Gymnasium environment.	119
3.17	Curriculum-based PPO training sequence under progressively perturbed scenarios.	124

3.18	EX scenario (Cyclone Laila) — augmented datasets used in training: (a) loads and total demand, (b) renewable generation, and (c) total load versus total renewables. Ribbons show the 10–90% range; solid lines indicate the mean values.	134
3.19	NO_V1 scenario (normal–HOMER Pro) — augmented datasets used in training: (a) loads and total demand, (b) renewable generation, and (c) total load versus total renewables. Ribbons show the 10–90% range; solid lines indicate the mean values.	135
3.20	NO_V2 scenario (normal–historical) — augmented datasets used in training: (a) loads and total demand, (b) renewable generation, and (c) total load versus total renewables. Ribbons show the 10–90% range; solid lines indicate the mean values.	136
3.21	Pearson correlation between perturbed and original time series across perturbation factors for (a) EX, (b) NO_V1, and (c) NO_V2 scenarios. Shaded bands denote one standard deviation across 10 perturbed variants, indicating how similarity to the original data decreases as perturbation strength increases.	137
3.22	Illustration of white noise (i.i.d.) versus Perlin noise (correlated). Perlin maintains smooth, temporally coherent variations.	139
3.23	Methodological flow: HOMER-based planning, MILP-formulated MPC baseline, DRL training and evaluation under data augmentation, post-training explainability and Monte-Carlo robustness analyses, and final results and discussion.	143
4.1	Validation of the proposed resilient operational method through comparison with HOMER PRO over a 10-day horizon (1–10 May): (a) total generation and load, (b) battery SOC profiles.	146
4.2	Expected Resiliency Index under different objective-weighting configurations.	151

4.3	Battery SOC under islanded operation: (a) resiliency and total imbalance only (Scenario 9), (b) all objectives equally weighted (Scenario 1).	152
4.4	Battery SOC under grid-connected operation: (a) without profit consideration (Scenario 25), (b) with profit optimisation (Scenario 26).	153
4.5	Normalised values of the four objective-function terms in Scenario 1, demonstrating proper scaling and balanced optimisation.	154
4.6	Load imbalance for Scenario 1: (a) total system imbalance, (b) essential loads.	155
4.7	Load imbalance for Scenario 1: (a) business loads, (b) deferred (agricultural) loads.	156
4.8	Comparison of the nonlinear BLC–DOD curve with the coarse, reference, and refined PWL approximations used in the MPC sensitivity analysis.	157
4.9	Relative approximation error of the coarse, reference, and refined BLC–DOD PWL models. The reported area error is the mean relative area error over the full DOD range.	158
4.10	Adjusted essential and business load profiles after applying a 10% demand-response shift (Scenario 16).	159
4.11	Effect of smart resilient control on the ERI with 25% of the designed battery capacity (Scenarios 22–24).	161
4.12	Electricity exchange with the grid: (a) without profit optimisation (Scenario 25), (b) with profit optimisation (Scenario 26).	164
4.13	Electricity exchange with the grid for 25% battery capacity: (a) without profit optimisation (Scenario 27), (b) with profit optimisation (Scenario 28).	165
5.1	SOC for Different MPC Weighting.	170
5.2	SOC MPC VS DRL.	171
5.3	Battery SOC comparison between perfect and imperfect forecast MPC operation during the HILP test period.	173

5.4	Perlin noise effects on a) renewable generation and b) battery SOC.	175
5.5	SOC comparison: Stage 3 prioritises resilience, whereas Stage 4 balances resilience and battery life.	177
5.6	Comparison of a) SOC, b) first derivative, and c) second derivative between DRL and MPC.	178
5.7	Comparative spider diagram of DRL and MPC under deterministic conditions across resilience, battery life, curtailment by load tiers, inference cost per episode, and total computation cost.	180
5.8	a) SOC and b) renewable generation under uncertainty across 4000 Monte Carlo scenarios. The blue shaded region denotes the 5th–95th percentile range, and the black line represents the deterministic baseline.	182
5.9	Statistical distributions of (a) Resilience Index and (b) Expected Years under 4000 Monte Carlo simulations.	183
5.10	Relationship between Resilience Index and Expected Years across 4000 Monte Carlo runs.	184
5.11	Correlation of Resilience Index with (a) average SOC, (b) SOC standard deviation, and (c) average renewable generation.	186
5.12	Correlation of Expected Years with (a) average renewable generation and (b) SOC standard deviation.	187
5.13	Global SHAP summary for critic network showing the top ten most influential features. Each dot represents one observation, with colour denoting the feature value and horizontal position showing its impact on the model output.	191

5.14	Global SHAP explanations for (a) charging and (b) discharging actions in the actor network.	192
5.15	Global SHAP explanations for the weighting of supply decisions across (a) essential, (b) business, and (c) agricultural loads. Only the top ten most influential features are shown.	194
5.16	SOC trajectory showing selected instances for charging (hour 13), discharging (hour 94), and idle (hour 162) actions.	195
5.17	LIME explanations for charging instance (hour 13): (a) charging and (b) discharging actions.	196
5.18	LIME explanations for charging instance (hour 13): supply weights for (a) essential, (b) business, and (c) agricultural loads.	197
5.19	LIME explanations for discharging instance (hour 94): (a) charging and (b) discharging actions.	198
5.20	LIME explanations for idle instance (hour 162): (a) charging and (b) discharging actions.	198
5.21	Qualitative alignment between SHAP and LIME.	200

List of Tables

2.1	Distribution of disaster impacts by type, 1980–2024 (costs in USD billion) [1].	16
2.2	Trends in disaster frequency, impact, and cost by period, 1980–2024 (costs in USD billion) [1].	16
2.3	Reliability versus resilience in power systems.	20
2.4	Comparison of resilience curves against a real HILP event (Hurricane Harvey) using the cumulative impact $\mathcal{I}_{RC}$ (outage-hours) from Equation (2.1). Errors are relative to the real-outage baseline.	28
2.5	Comparative Analysis of Microgrid Design and Operational Literature [49]. .	32
2.6	Comparative Review of DRL-based Microgrid Resiliency Studies.	44
2.6	<i>Continue</i>	45
2.6	<i>Continue</i>	46
2.6	<i>Continue</i>	47
3.1	Techno-economic performance parameters of the microgrid system.	81
3.2	List of parameters and their definition.	93
3.3	Variables and their definitions	94
3.4	Core libraries and pinned versions used in the DRL and explainability part. .	117
3.5	Variables and Parameters Used in the Environment Model.	118
4.1	Simulation Scenarios and Results.	148
4.2	Sensitivity of MPC performance to BLC–DOD PWL approximation fidelity.	156
5.1	Comparison of DRL and MPC results under certain and uncertain scenarios.	172

5.2	Computation-time comparison of MPC and DRL controllers [FLOPs].	179
5.3	Statistical summary of 4000 Monte Carlo runs for DRL policy performance. .	183
5.4	Performance of the DRL policy in selected extreme scenarios.	184

Acronyms

A2C	Advantage Actor-Critic. 41, 44
A3C	Asynchronous Advantage Actor-Critic. 41, 45, 46
AHP	Analytic Hierarchy Process. 26
AMI	Advanced Metering Infrastructure. 1, 16
API	Application Programming Interface. 116–118
AUC	Area Under The Curve. 97, 99, 102
BESS	Battery Energy Storage System. 4, 5, 9, 10, 12, 13, 17, 18, 20, 23, 32, 33, 44–47, 62, 65, 80–83, 85, 91, 143
BLC	Battery Life Cycle. xiii, xv, xviii, 9, 12, 13, 34, 57, 60, 63, 65, 83, 91–93, 97, 98, 101, 102, 106, 108, 150, 152, 154, 156–158, 166, 172, 209, 210
CAIDI	Customer Average Interruption Duration Index. 20, 44
CEENS	Conditional Expectation Of Energy Not Supplied. 19, 20, 23, 25
CHP	Combined Heat and Power. 32, 33, 44
CIA	Confidentiality Integrity Availability. 17
CNN	Convolutional Neural Network. 44, 47, 55
COE	Cost Of Energy. xiii, 69, 79, 81–85, 88, 89
CPARM	Cyber-Physical Resilience Metric. 26
CRF	Capital Recovery Factor. 79
CVaR	Conditional Value At Risk. 19, 20, 25
DARPA	Defense Advanced Research Projects Agency. 51, 55

DDPG	Deep Deterministic Policy Gradient. 41, 43–48, 50
DDQL	Double Deep Q-Learning. 44, 46, 50
DDQN	Double Deep Q-Network. 43, 45, 47, 50
DE	Differential Evolution. 32
DER	Distributed Energy Resource. 1, 20, 23, 26, 45–48
DG	Distributed Generation. 32, 44–47
DNN	Deep Neural Network. 4, 39, 46
DOD	Depth Of Discharge. xiii, xv, xviii, 4, 5, 7, 9, 11–13, 34, 49, 50, 57, 60, 63, 65, 80, 83, 88, 91–93, 97–99, 101, 102, 107, 108, 122, 123, 154–158, 161, 166, 172, 176, 209, 210
DP	Dynamic Programming. 37, 38, 47
DQN	Deep Q-Network. 40, 44–47, 50
DR	Demand Response. 90, 91, 102, 145, 147, 148, 158, 159, 166, 207, 208
DRL	Deep Reinforcement Learning. vii, xii, xiv–xvi, xviii, xix, 3, 4, 6, 8–14, 29, 31, 32, 35, 38–46, 48–52, 55–60, 63–66, 81, 83, 89, 102, 105, 108, 110, 116, 117, 123, 133, 141–143, 166–169, 171–174, 176–181, 183–185, 188, 189, 191, 193, 194, 197, 204, 205, 207, 209, 210, 212
EAGLE-I	Environment for Analysis of Geo-Located Energy Information. 26, 30
EENS	Expected Energy Not Supplied. 15, 20, 23, 25
EIU	Energy Index Of Unreliability. 15, 20, 25
EMS	Energy Management System. 33, 34, 43, 44, 48–51, 55, 56
ENS	Energy Not Supplied. 20, 29
ERC	Expected Resilience Curve. 30

ERI	Expected Resilience Index. xv, 30, 150, 160, 161, 166, 207
ESPC	Effective Shortest Paths Between Critical Nodes. 26
EY	Expected Years. 108, 118, 148, 156, 157, 170–172, 181, 183–185, 206–208
FDI	False Data Injection. 45, 48
FLOPs	Floating-Point Operations. xix, 142, 177, 179
GA	Genetic Algorithm. 46, 49
GAE	Generalized Advantage Estimation. 111, 116
GAMS	General Algebraic Modeling System. 92
GDPR	General Data Protection Regulation. 52, 55
GHI	Global Horizontal Irradiance. 69, 70
Grad-CAM	Gradient-Weighted Class Activation Mapping. 55
HIL	Hardware In the Loop. 48
HILP	High-Impact Low-Probability. vii, xi, xii, xviii, 1, 3, 5–10, 12–20, 23, 28–35, 43–51, 55, 60, 62, 63, 81, 85
IA	Interpretable Agents. xii, 52, 53, 56
IBMDP	Iterative Bounding Markov Decision Process. 53
IE	Intrinsic Explainability. xii, 52, 53, 56
IFRL	Interpretable Fuzzy Reinforcement Learning. 53
IWOA	Improved Whale Optimization Algorithm. 32
KL	Kullback–Leibler. 112–114
LF	Load Following. 79

LIME	Local Interpretable Model-Agnostic Explanations. ix, x, xvii, 3, 4, 9, 12, 13, 15, 44, 50, 51, 55, 56, 58, 60, 74, 125–127, 129, 169, 189, 194–205, 208, 210
LOLE	Loss Of Load Expectation. 15, 20, 25
LOLF	Loss Of Load Frequency. 15, 20, 25
LOLP	Loss Of Load Probability. 25, 44
LSI	Line Strength Index. 26
LSTM	Long Short-Term Memory. 44, 47
MA	Multi Agent. 45–48
MAB	Multi-Armed Bandit. 37
MAPPO	Multi Agent Proximal Policy Optimisation. 46
MC	Monte Carlo. 37, 38, 44–46, 58
MDP	Markov Decision Process. xii, 37, 39, 53
MER	Mobile Energy Resource. 20, 23
MG	Micro Grid. 23, 26, 29, 44
MILP	Mixed Integer Linear Programming. viii, xiv, 9, 10, 12–14, 32, 34, 44–47, 49, 57, 60, 63, 65, 73, 91–93, 95, 99, 102, 104, 105, 109, 142–145, 172, 206, 208
MIMO	Multi Input Multi Output. 90
MINLP	Mixed Integer Non Linear Programming. 92
MIQP	Mixed Integer Quadratic Programming. 103, 105
ML	Machine Learning. 6
MMG	Multi Micro Grid. 46, 47
MPC	Model Predictive Control. viii, ix, xii–xvi, xviii, xix, 3–5, 8–14, 29–31, 33–35, 44, 46, 49–51, 56–60, 62–67, 73, 81, 83, 89–93, 102, 103, 105, 107–109,

	141–147, 154, 156–158, 166–174, 176–180, 188, 202, 203, 205–210, 212
MTTR	Mean Time To Repair. 26
NOCT	Nominal Operating Cell Temperature. 75
OOD	Out-of-Distribution. 203
OPF	Optimal Power Flow. 44, 46, 48
PHE	Post-Hoc Explainability. xii, 53
PMU	Phasor Measurement Unit. 16
POWER	Prediction of Worldwide Energy Resources. xii, 62, 69, 70, 73, 79
PPO	Proximal Policy Optimisation. viii, ix, xii, xiii, 3, 39, 41, 43–48, 50, 55, 56, 60, 63–66, 73, 110, 112–116, 123, 124, 126, 168, 188, 205, 207
PSI	Pole Strength Index. 26
PV	Photovoltaic. xiii, 4, 9–13, 18, 32, 33, 44–47, 62, 65, 69, 73–75, 79, 81–84, 88, 118, 133, 138, 140, 143
PYOMO	Python Optimization Modeling Objects. viii, 63, 64, 73, 90, 92, 144, 206
RES	Renewable Energy Source. 4, 33, 45, 46, 78, 83
RI	Resilience Index. 106, 118, 143, 148–150, 156, 157, 170–174, 181, 183–185, 206–208
RICD	Resilience Index Considering the Duration of Extreme Events. 25
RL	Reinforcement Learning. 29, 36, 37, 40, 45–47, 55
RMS	Root Mean Square. 207
SAC	Soft Actor-Critic. 41, 44–48, 50
SAIDI	System Average Interruption Duration Index. 15, 20, 25

SAIFI	System Average Interruption Frequency Index. 15, 20, 25, 44
SARSA	State-Action-Reward-State-Action. 38
SB3	Stable Baselines3. viii, 63, 64, 73, 110, 116, 118
SCADA	Supervisory Control And Data Acquisition. 16
SHAP	SHapley Additive exPlanations. ix, x, xvi, xvii, 3, 4, 9, 12–14, 44, 50, 51, 54–56, 58, 60, 74, 125, 128–130, 169, 189–194, 199–205, 208, 210
SOC	State Of Charge. x, xiii–xvii, 4, 5, 7, 10, 32–34, 45, 46, 49, 50, 57, 65, 78, 80–83, 89–91, 93–95, 108, 109, 117, 118, 121, 123, 140, 143–146, 148, 150–154, 160, 161, 166, 169–178, 181–183, 185–187, 190, 195, 202, 203, 205, 207, 208, 210, 212
SRI	Severity Risk Index. 20, 23
SSA	Salp Swarm Algorithm. 32
STC	Standard Test Conditions. 74
TD	Temporal-Difference. 37, 38, 40, 111, 112
TD3	Twin Delayed Deep Deterministic Policy Gradient. 41, 44–46, 48, 50
TNPC	Total Net Present Cost. xiii, 69, 79, 81–85, 88
TRPO	Trust Region Policy Optimisation. 50, 110, 112, 113
TS	Total Switches. 108
VaR	Value At Risk. 19, 20, 25
VoLL	Value Of Lost Load. 23
WT	Wind Turbine. xiii, 4, 9, 10, 12, 13, 18, 32, 45–47, 62, 65, 69, 73, 74, 77, 79, 81–84, 118, 143
XAI	eXplainable Artificial Intelligence. 51

XDRL eXplainable Deep Reinforcement Learning. vii, xii, 3, 5, 11, 12, 51, 53, 56, 64, 66, 67, 141, 169, 189, 205

Chapter 1

Introduction

The rapid electrification of multiple sectors, the transition towards net-zero emissions, and the growing frequency of High-Impact Low-Probability (HILP) events are reshaping modern power systems. In parallel, digitalisation through Advanced Metering Infrastructures (AMIs) enables two-way interaction with distribution and transmission system operators. This transformation reshapes conventional grids into smart grids, comparable to the restructuring that followed liberalisation in the 1990s. Within this shift, microgrids are local clusters of loads and Distributed Energy Resources (DERs) that can operate grid-connected or islanded. They provide a practical means to integrate renewables while enhancing local resilience during disruptions. However, their daily operation is challenging due to uncertain generation, fluctuating demand, cyber-physical vulnerabilities, and battery ageing, all of which directly influence cost and sustainability. These factors motivate control approaches that are high-performing and explainable to operators and stakeholders.

To illustrate the importance of resilience studies, Figure 1.1 shows the rise of “billion-dollar” disasters in the United States since 1980. Events per year have increased from 3.3 in the 1980s to 23 in the last five years, and the annual cost rose from \$22 B to \$149.3 B, with \$182.7 B in 2024 alone. The detailed values are provided in Table 2.2. The distribution of losses across hazards in Figure 1.2 demonstrates that “Tropical Cyclones” dominate total cost over 1980–2024 [1]. The underlying values are summarised in Table 2.1. While HILP events may include natural hazards, cyberattacks, communication failures, and large grid outages,

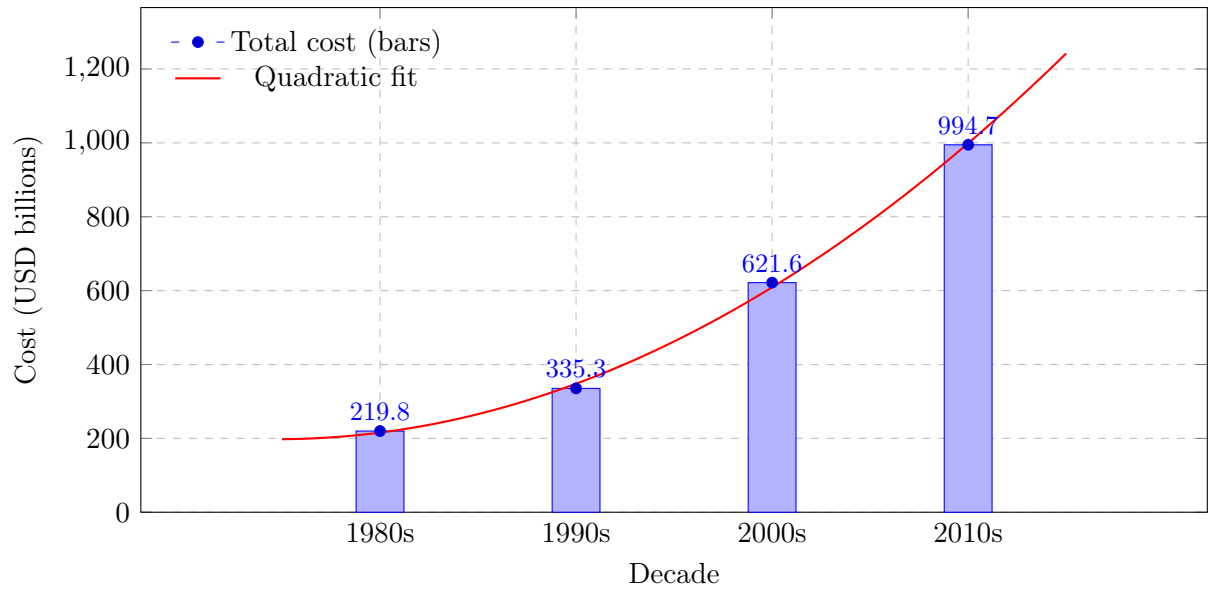


Figure 1.1: Total disaster costs by decade with polynomial fit (1980s–2010s). The 2020–2024 period is omitted because it does not cover a full decade [1].

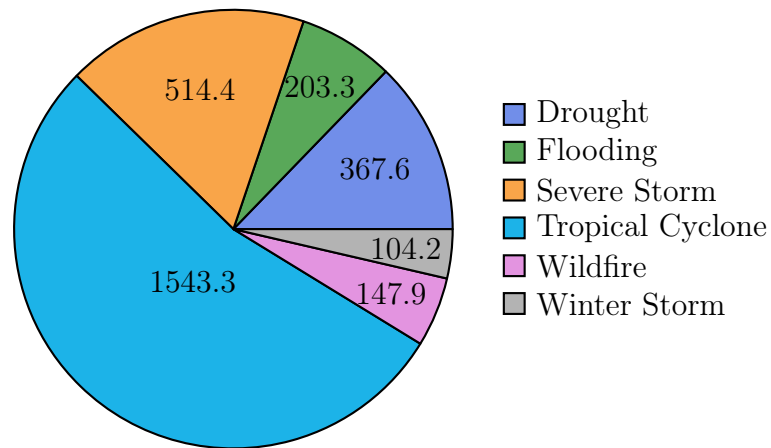


Figure 1.2: Share of total cost by disaster type (1980–2024), excluding Freeze for legibility. Values are shown in USD billions [1].

the main focus of this work is weather-driven disruption, particularly cyclone-related stress affecting a rural standalone microgrid. In such events, renewable availability may be reduced, storage dependence may increase, and priority loads may be exposed to curtailment risk. Taken together, these trends justify a resilience-centred approach and motivate the use of real cyclone data to evaluate how the proposed controllers perform under extreme events.

In response, this thesis advances an eXplainable Deep Reinforcement Learning (XDRL) framework for resilient microgrid energy management, with Model Predictive Control (MPC) as the benchmark. Two complementary directions are pursued. First, microgrid design and MPC-based operation are developed using real cyclone data from India to anchor sizing and operation in observed extremes. Second, a Deep Reinforcement Learning (DRL) controller based on Proximal Policy Optimisation (PPO) is equipped with post hoc explanations, SHapley Additive exPlanations (SHAP) and Local Interpretable Model-Agnostic Explanations (LIME), to support transparency, validation, and operator trust.

1.1 Background and Motivation

Microgrids are valuable in remote and rural areas with weak or unreliable grid access and in developed regions seeking resilience and decarbonisation. They can operate in islanded mode to maintain essential services during disturbances. Yet planning and operation must account for HILP events that disrupt generation, storage, and local networks.

In power systems, reliability has been extensively studied for decades and is now considered a mature field, with established methods, standards and metrics. It studies the ability to supply demand under single or limited faults, often using $N-1$ or $N-2$ contingency analysis. However, the more pressing challenge today lies in resilience, which addresses rare and severe events beyond the scope of traditional reliability studies. It consists of preparation, operation during the event, and restoration after the event. Despite its growing importance, resilience remains less formalised than reliability, with fewer standardised metrics and limited validation in real-world events. For microgrids, resilience is crucial because local shocks can rapidly disrupt essential services such as health care, water supply, and communications, unlike large interconnected grids where redundancy can limit the impact of disturbances.

Renewable Energy Source (RES) are inherently intermittent and variable, which in small systems can lead to energy imbalances as well as voltage and frequency stress. To maintain stability, backup resources are therefore essential. Diesel generators can provide effective backup, but they introduce noise and pollution that conflict with sustainability goals. Battery Energy Storage System (BESS), by contrast, offer a cleaner alternative and enable fast response during disturbances. Hybrid systems that combine Photovoltaic (PV) and Wind Turbine (WT) are also beneficial, as their generation profiles can complement one another depending on site conditions. When integrated with BESS, such hybrid RES configurations reduce curtailment and enhance the ability to serve priority loads.

While resource portfolios improve supply stability, effective control remains the next critical challenge. Conventional rule-based or deterministic scheduling methods often fail to cope with uncertainty. MPC can address constraints and anticipate future states, but it relies on accurate models and forecasts. Moreover, its computational and communication demands make deployment difficult at the edge, a limitation that is common in remote sites. In addition, many MPC formulations consider battery life only indirectly through simple State Of Charge (SOC) limits. This means that, in practice, battery degradation is often underestimated, limiting the long-term sustainability of existing control strategies. Since battery ageing is highly sensitive to Depth Of Discharge (DOD), explicit battery-aware control is required.

DRL, by contrast, does not require a detailed model and can handle uncertainty more flexibly. Although it generally yields near-optimal rather than strictly optimal solutions, once trained, it can act quickly and deliver robust performance in practice. The drawback is that DRL relies on Deep Neural Networks (DNNs), which are often treated as black boxes. This hinders adoption in critical infrastructure where transparency is essential. Post hoc explainability methods, such as SHAP and LIME, can reveal why a trained agent selects particular actions and thereby support trust, governance, and validation.

Accordingly, this thesis establishes an MPC baseline with explicit battery life and resilience objectives and develops an XDRL controller that targets the same goals with fast inference and operator-oriented explanations. This combination of performance under uncertainty, battery-aware operation, and transparency frames the contributions of the work.

1.2 Challenges: Uncertainty, Battery Ageing, and HILP Events

Operational decisions in isolated or weakly connected microgrids face persistent challenges:

1. **Uncertainty:** Renewable generation and demand are volatile, and forecasts remain imperfect. As the planning horizon or time resolution increases, scenario counts grow rapidly and solution times expand. This expansion raises the risk that re-optimised schedules become outdated, particularly where computing capacity and communications are limited at the edge.
2. **Battery ageing:** The lifetime of BESS depends on depth of discharge, cycling frequency, C-rate, temperature, and calendar fade. Approximations based on wide SOC bands or simple penalties mask real life-cycle costs and can shift expenditure to earlier replacements. Explicit, DOD-aware modelling is therefore necessary to capture the operational impact of control policies.
3. **HILP conditions:** Severe, rare events (e.g., cyclones) create multi-contingency stress and partial data loss. Local records are limited, yet a credible resilience assessment requires validation against real disturbance data. Using location-specific cyclone periods provides a realistic basis for validation.
4. **Metric ambiguity:** Resilience is measured using diverse indices and normalisations.

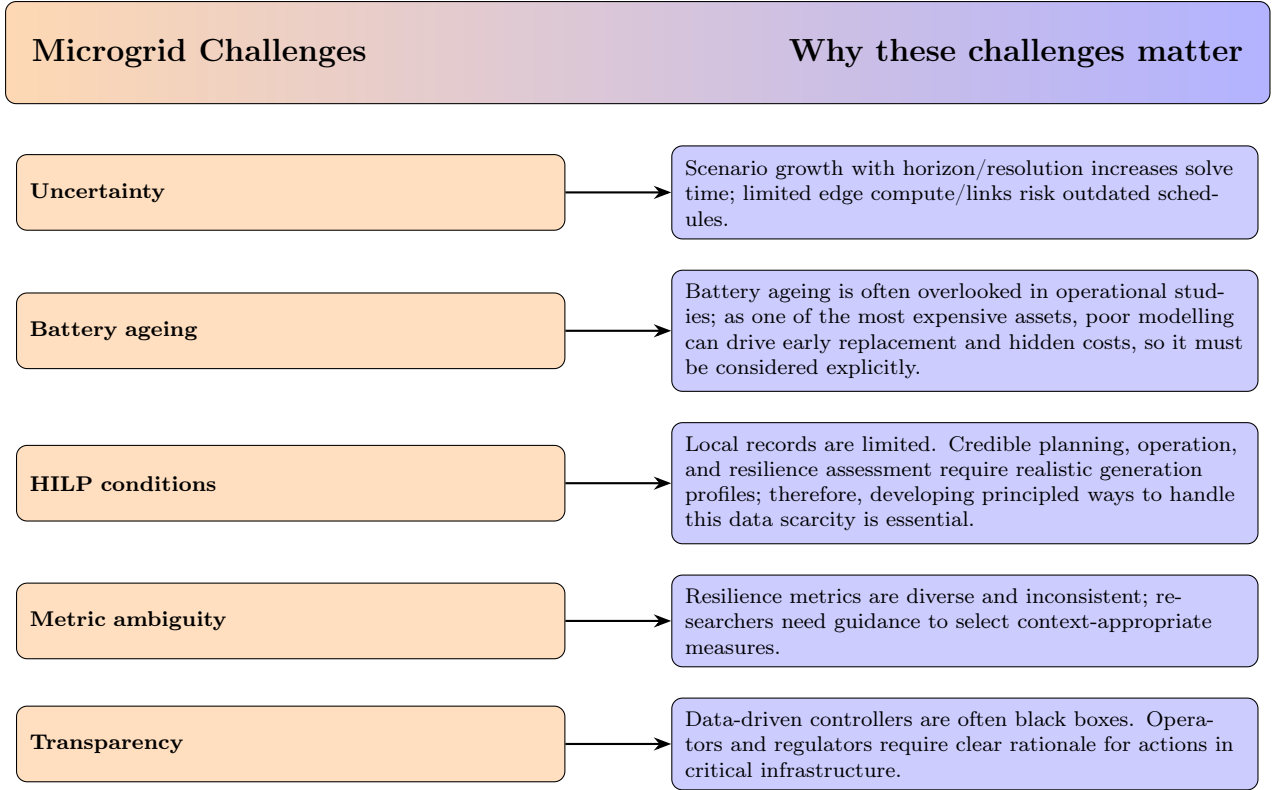


Figure 1.3: Challenges for microgrids and why they matter.

This diversity complicates cross-study comparison and makes metric selection difficult for new researchers. Recent reviews offer structure, but ambiguity persists across systems and contexts [2].

5. **Transparency:** Machine Learning (ML)-based controllers, including DRL, are often treated as black boxes. In safety-critical infrastructure, operators and regulators require clear rationales for actions to support accountability and governance. Post hoc explainability can help bridge this gap by revealing the drivers behind selected actions.

Figure 1.3 summarises these challenges and their implications for practice. These challenges highlight the limitations of current microgrid control approaches and frame the central problem addressed in this thesis.

1.3 Problem Statements

Problem Statements

Isolated or weakly connected microgrids require an energy management approach that:

1. Clarifies metric choice: select resilience metrics appropriate to the context so results are interpretable and comparable for researchers (including new researchers).
2. Addresses HILP conditions: remain effective under rare, severe events (e.g., cyclones) using location-specific, real disturbance data when available.
3. Handles uncertainty and variability: produce timely decisions when forecasts are imperfect and conditions change rapidly.
4. Protects battery life: incorporate DOD-aware ageing rather than wide SOC bands or simple penalties.
5. Ensures transparency: provides justifiable decisions for stakeholders, supporting governance and trust in critical infrastructure.

The proposed design framework is guided by several essential requirements that respond directly to the operational challenges of microgrids. First, resilience must be reported in a consistent and interpretable way. This calls for the adoption of a clear, priority-weighted resilience index, ensuring that results remain comparable across different studies and case applications. Second, evaluation should not rely solely on generic or synthetic scenarios. Instead, it needs to be grounded in real, location-specific disturbance data, which can then be complemented with principled synthetic perturbations to explore a wider range of conditions. Third, battery ageing should not be treated with oversimplified assumptions. Instead, explicit modelling of DOD dependent effects is needed to capture the true costs and replacement dynamics of storage assets. Fourth, the framework should achieve computational

efficiency by delivering fast inference suitable for edge deployment once offline training has been completed. This feature is particularly important in applications where access to high-performance computing is limited. Finally, the framework must go beyond performance metrics alone by ensuring transparency. Post hoc explainability methods are therefore required to reveal the drivers of decisions, enabling operators, regulators, and stakeholders to validate actions and place trust in the system’s governance.

1.4 Research Questions

This thesis pursues the following questions:

Research Questions

- RQ1.** How can resilience be measured for microgrids under HILP conditions, and which metrics are most suitable for operations with priority loads while remaining comparable across studies?
- RQ2.** How can realistic, location-specific generation and demand profiles be developed for cyclone-prone microgrid case studies, ensuring credible representation of extreme events despite limited metered data?
- RQ3.** How should an MPC baseline be structured to operate while accounting for battery life, load priorities, and an expected-resilience perspective?
- RQ4.** To what extent can a PPO-based DRL controller enhance resilience and responsiveness compared with MPC under uncertain forecasts and HILP stress tests? What trade-offs arise for expected battery life, and how do their computational characteristics differ in terms of solve, training, and inference time?
- RQ5.** How interpretable and practically useful are DRL-based energy management decisions in isolated microgrids, and to what extent can explainability techniques support operator trust and governance?

1.5 Research Aim and Objectives

Aim

To develop and evaluate an explainable, optimal or near-optimal framework for micro-grid energy management in isolated settings. The framework should enhance resilience to HILP events, protect BESS lifetime, and remain transparent to stakeholders and operators.

Objectives

1. **Clarify resilience measurement.** Review and organise quantitative resilience metrics, then adopt a priority-weighted served energy index and resilience curves suitable for operations and HILP analysis. Provide selection guidance to reduce metric ambiguity.
2. **Build a location-specific case.** Plan and size a PV–WT–BESS microgrid using real load data and historical weather from a cyclone-prone region in India (HOMER Pro). Derive realistic PV/WT generation from irradiance and wind-speed records to anchor the case study.
3. **Formulate an MPC baseline.** Implement an MPC/Mixed Integer Linear Programming (MILP) operation model with explicit battery life modelling via a piece-wise linear Battery Life Cycle (BLC)–DOD relation. Include resilience terms linked to power imbalance and resilience curves; model load priorities.
4. **Design the DRL controller.** Create a Gym-style environment and reward that combine resilience with battery life terms. Train a PPO agent, using data augmentation to expose the policy to uncertainty and HILP variability.
5. **Provide transparency.** Use post hoc explanations with LIME (local) and SHAP (global) to reveal the drivers of actions and support operator trust and validation.

... Objectives

6. **Benchmark and stress-test.** Compare MPC and DRL under deterministic and stochastic tests (including a real cyclone profile), reporting resilience indices, expected battery life, and edge-relevant computational characteristics (solve, training, inference time). Assess robustness via Monte Carlo analysis.
7. **Extract operational insight.** Use explainability and sensitivity analyses to identify patterns that guide reserve setting and prioritisation (e.g., links between average SOC, net energy balance, and resilience).

1.6 Scope and Boundaries

The study is structured in two stages. Stage 1 establishes a common foundation for both MPC and DRL by planning and sizing the hybrid PV–WT–BESS microgrid using HOMER Pro, together with local load estimates and historical weather records. This stage also identifies the disturbance profile, including the Cyclone Laila period, to anchor the case study. Stage 2 evaluates operation on this basis using two alternative controllers. First, an MPC–MILP model with explicit resilience and battery life terms is implemented. Second, a DRL controller is trained with uncertainty-augmented scenarios. Their performance is then compared on resilience, battery life, and computational efficiency, with explainability methods applied to the DRL policies.

The analysis is bounded by practical choices consistent with the research aim. The case centres on one rural standalone microgrid at a cyclone-prone coastal site in India, where weather-driven HILP events, particularly Cyclone Laila, provide the main disturbance context. The studied system consists of PV generation, wind generation, prioritised loads, a bidirectional converter, and a BESS; within the operational analysis, the battery is treated

as the only energy storage and fault-ride-through resource during disruptive periods. Diesel generation, fuel cells, hydrogen storage, and other backup technologies are therefore outside the main dynamic control analysis. PV and wind outputs are reconstructed from irradiance and wind records; battery ageing in the MPC model is represented by a piecewise DOD relation; and network-level phenomena, including voltages, reactive power, line limits, and protection, are abstracted, with the focus placed on energy balancing and priority-aware scheduling. DRL training is offline, stress tests focus on Cyclone Laila and structured perturbations, explanations are post hoc rather than causal, and computation is benchmarked on general-purpose hardware. These boundaries are mitigated by using location-specific disturbance records, curriculum and perturbation-based training, and Monte Carlo robustness analysis.

Findings at a glance

At a high level, the proposed XDRL controller matches the MPC baseline on resilience during cyclone stress tests while improving expected battery life and keeping inference lightweight for edge-class hardware. The learned behaviours are interpretable: charge during surplus, discharge to secure essential loads during deficit, and modulate medium-priority demand as the primary flexibility lever. Across uncertainty augmentations and Monte Carlo runs, results remain stable, indicating robust policy behaviour and operator-aligned explanations.

1.7 Thesis Structure

The thesis is organised as follows. **Chapter 2** reviews resilience in power systems, resilience metrics for priority loads, control approaches (MPC and DRL), battery ageing models, and explainability methods. It highlights gaps in meaningful resilience metrics, explicit battery-

aware control under HILP conditions, benchmarking of DRL against MPC, and transparency in learning-based controllers. **Chapter 3** sets out the modelling and evaluation framework: data sources and preprocessing (including reconstruction of PV/WT generation from irradiance and wind records), scenario design with cyclone profiles, resilience indices and battery life approximations, uncertainty augmentations, and an edge-aware benchmarking protocol for computation (solve, training, inference). This chapter also presents the case-study microgrid (PV–WT–BESS) for a coastal village near Ongole/Kothapatnam, India, and reports sizing results using HOMER Pro with location-specific resource and load data. **Chapter 4** formulates the MPC baseline as a MILP with priority-weighted loads, an expected-resilience construct, demand-response elements, and an explicit piecewise-linear BLC–DOD relation to moderate ageing. **Chapter 5** develops the XDRL controller: a PPO-based DRL agent with a composite reward (resilience and battery terms), curriculum to balance objectives, and post hoc explanations using LIME (local) and SHAP (local/global). This chapter also carries out the comparative study with the MPC baseline under deterministic and stochastic tests (including cyclone conditions), reporting resilience indices, expected battery life, and computational characteristics for edge-feasible inference. **Chapter 6** summarises findings, notes limitations, and outlines future directions.

To provide an overview of the research structure, Figure 1.4 maps the research questions to the thesis objectives and the chapters where each objective is delivered.

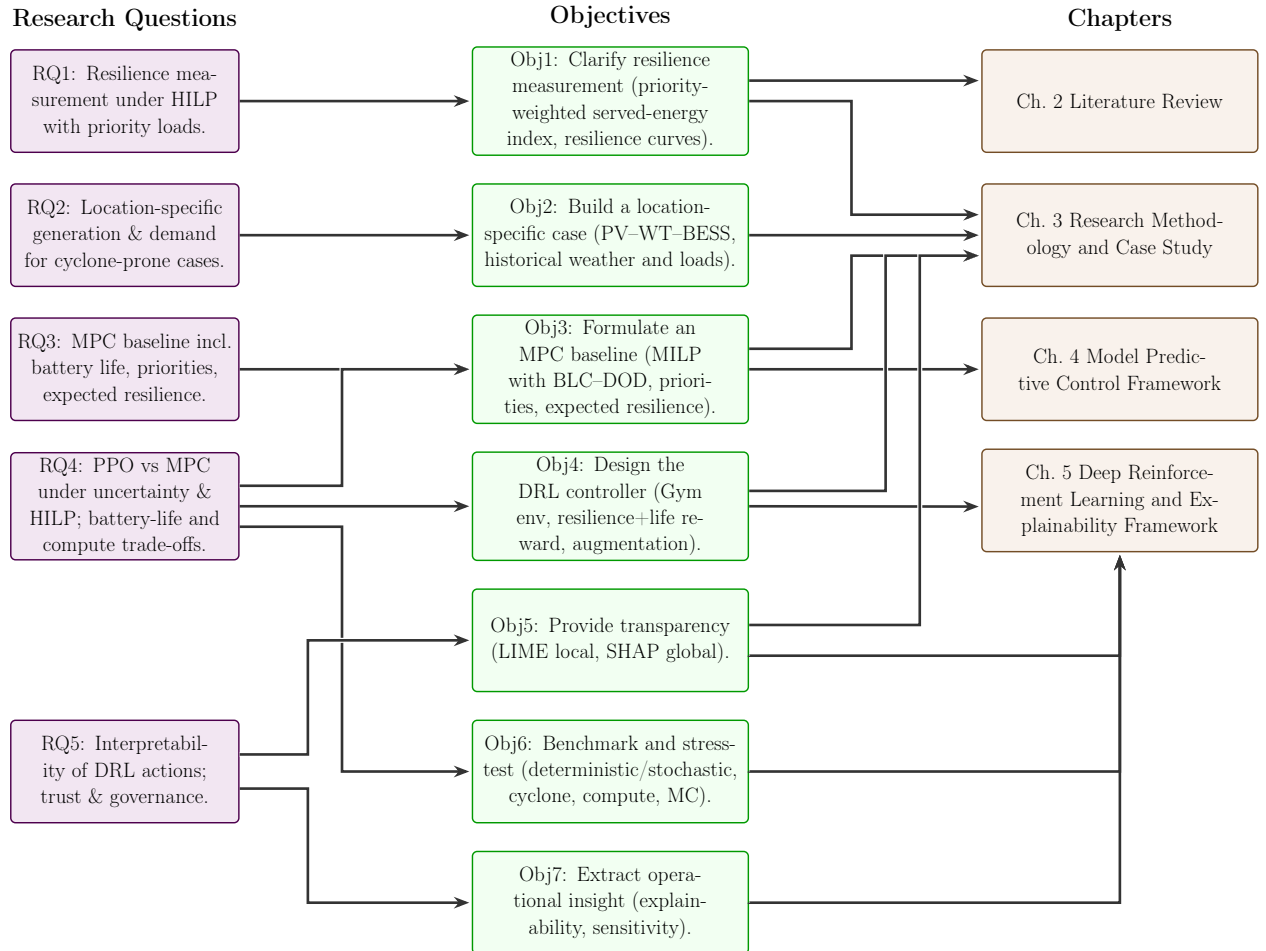


Figure 1.4: Logical mapping from Research Questions (RQ) to Objectives (Obj) and Chapters (Ch).

Chapter 2

Literature Review

2.1 Thematic Structure of the Review

This thesis integrates three strands: **Explainability**, **Deep Reinforcement Learning (DRL)**, and **Resilient Energy Management**.¹

Although these themes appear sequentially in the title, the literature review intentionally proceeds “bottom-to-top,” beginning with fundamentals and advancing toward implementation. It first introduces the concept of **resilience**, explaining how it differs from reliability, and then outlines the three stages of absorptive, adaptive, and restorative capacity. This part concludes with a review of resilience metrics and curves that capture performance under disruption and recovery.

The second part examines **control approaches**, comparing MPC, often formulated as MILP, with reinforcement learning methods. The analysis focuses on model formulation, objective design for energy balance and battery life, and comparative performance under uncertainty and HILP events.

The final part explores **transparency and governance** in learning-based controllers, highlighting the ethical and regulatory motivations for explainability. It contrasts interpretable, intrinsically explainable, and post-hoc attribution techniques such as SHAP and

¹To support navigation throughout the literature review, colour is used as a consistent visual marker for the three thematic strands: **resilience and its metrics**, **control approaches**, and **explainability**.

LIME, emphasising their role in supporting trust and accountability in decision-making.

The following section defines resilience and its operational metrics before examining control approaches and explainability frameworks.

2.2 Concept of resilience in power systems

Section Summary

Resilient energy management begins with a clear understanding of what resilience means in power systems and how it differs from reliability. This section defines the concept, outlines its key attributes, and reviews the indices used to quantify it, establishing the foundation for subsequent optimisation and control design.

2.2.1 Motivation and definitions

In power system engineering, *reliability* refers to the sustained ability to supply electricity within quality standards and with minimal interruptions, typically decomposed into *adequacy* (sufficient resources to meet demand) and *security* (ability to withstand $N - 1$ or $N - 2$ contingencies) [3], [4], [5].

Reliability is assessed with metrics like System Average Interruption Duration Index (SAIDI)/System Average Interruption Frequency Index (SAIFI), Loss Of Load Expectation (LOLE)/Loss Of Load Frequency (LOLF), and Expected Energy Not Supplied (EENS)/Energy Index Of Unreliability (EIU) under normal or moderately disturbed operating conditions [2].

HILP events, by contrast, are infrequent and system-scale disturbances can trigger

Table 2.1: Distribution of disaster impacts by type, 1980–2024 (costs in USD billion) [1].

Disaster Type	Events	Events/Year	Freq. (%)	Total Cost	Cost Share (%)	Cost/Event	Cost/Year	Deaths	Deaths/Year
Drought	32.0	0.7	7.9	367.6	12.6	11.5	8.2	4,658.0	104.0
Flooding	45.0	1.0	11.2	203.3	7.0	4.5	4.5	765.0	17.0
Freeze	9.0	0.2	2.2	37.4	1.3	4.2	0.8	162.0	4.0
Severe Storm	203.0	4.5	50.4	514.4	17.6	2.5	11.4	2,145.0	48.0
Tropical Cyclone	67.0	1.5	16.6	1,543.3	52.9	23.0	34.3	7,211.0	160.0
Wildfire	23.0	0.5	5.7	147.9	5.1	6.4	3.3	537.0	12.0
Winter Storm	24.0	0.5	6.0	104.2	3.6	4.3	2.3	1,463.0	33.0
All Disasters	403	9.0	100.0	2918.1	100.0	7.2	64.8	16941	376

Table 2.2: Trends in disaster frequency, impact, and cost by period, 1980–2024 (costs in USD billion) [1].

Time Period	Events	Events/Year	Cost	Cost Share (%)	Cost/Year	Deaths	Deaths/Year
1980s (1980–1989)	33.0	3.3	219.8	7.5	22.0	3,017.0	302.0
1990s (1990–1999)	57.0	5.7	335.3	11.5	33.5	3,075.0	308.0
2000s (2000–2009)	67.0	6.7	621.6	21.3	62.2	3,102.0	310.0
2010s (2010–2019)	131.0	13.1	994.7	34.1	99.5	5,227.0	523.0
Last 5 Years (2020–2024)	115.0	23.0	746.7	25.6	149.3	2,520.0	504.0
Last 3 Years (2022–2024)	73.0	24.3	461.6	15.8	153.9	1,534.0	511.0
Last Year (2024)	27.0	27.0	182.7	6.3	182.7	568.0	568.0
All Years (1980–2024)	403	9.0	2918.1	100.0	64.8	16941	376

multi-element failures and prolonged restoration times. Empirical data show both the frequency and cost of such events have increased [1]. Table 2.1 shows the cost share is dominated by tropical cyclones (over half of total losses), while severe storms account for the largest share of events (more than 50%). Table 2.2 further highlights this escalation: the annual number of events increased from 3.3 in the 1980s to 23.0 in the five years ending 2024, with annual fatalities rising from 302 to 504. These trends motivate going beyond steady-state reliability to explicitly address HILP.

Figure 2.1 classifies the underlying causes of extreme events in power systems into three main categories [2]. First, natural disasters (e.g., floods, hurricanes, wildfires, droughts, winter storms, earthquakes) have become increasingly frequent and costly, as shown by the trends in Tables 2.1 and 2.2. Second, cyber attacks, once uncommon in grid operations, have expanded sharply with increasing digitalisation. Technologies such as AMI, Phasor Measurement Units (PMUs), and networked Supervisory Control And Data Acquisition (SCADA)

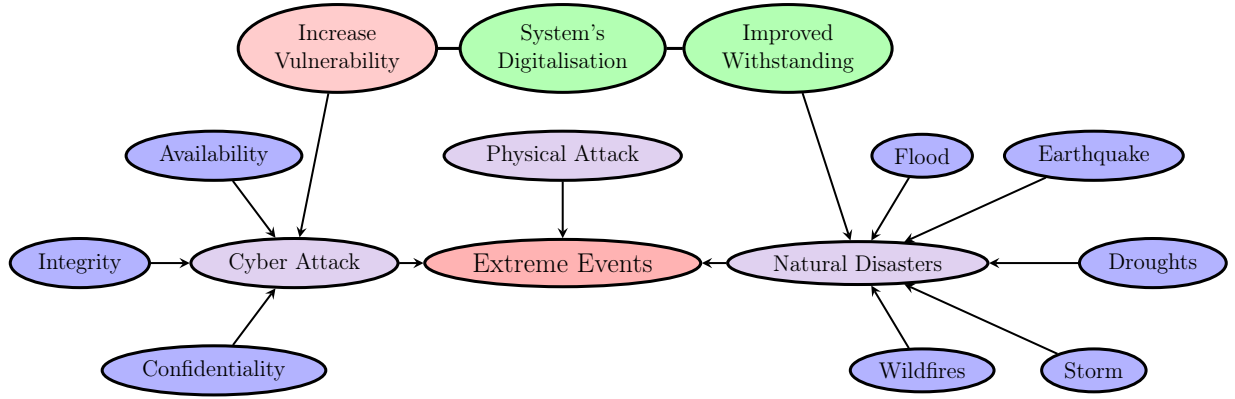


Figure 2.1: The underlying origins and types of HILP events [2].

systems enhance observability and control but simultaneously enlarge the attack surface across the Confidentiality Integrity Availability (CIA) triad [6], [7], [8]. Third, physical attacks (e.g., sabotage, vandalism, or conflict-driven damage) can disrupt substations, lines, or communication channels, leading to widespread contingencies. At the same time, digitalisation strengthens the system's ability to withstand and adapt to natural hazards by enabling faster situational awareness, network reconfiguration, controlled islanding, and coordinated microgrid or BESS dispatch [9], [10]. This duality motivates the need for a resilience-oriented approach that anticipates and mitigates HILP events across natural, cyber, and physical domains while digital tools are leveraged to prioritise critical loads and accelerate recovery.

A context diagram is useful because HILP events can affect a microgrid through different physical, cyber, operational, and stakeholder pathways. Figure 2.2 links common HILP event types with the microgrid components they may affect, the stakeholders concerned, and the resulting resilience issues. This provides a system-level view before the following sections discuss resilience metrics and control approaches in more detail.

Figure 2.2 shows that resilience assessment depends on both the origin of the disruption and the part of the microgrid affected. Weather-driven events mainly stress renewable generation, storage reserve, and priority-load supply. Grid-level events affect the grid interface and require islanded balancing. Cyber and communication events reduce observability

HILP event type	Microgrid components affected	Stakeholders	Main resilience concern
Weather-driven events Cyclones, storms, floods, heatwaves	PV, WT, lines, BESS, priority loads	Consumers, critical-service users, microgrid operator	Reduced renewable availability, load curtailment, reserve depletion
Grid-level events Outage, forced islanding, load shedding	Grid interface, converter, BESS, controller/EMS	Microgrid operator, grid operator, consumers	Loss of import/export support, islanded operation, higher local balancing need
Cyber/communication events Data loss, false data, controller disruption	Communication layer, sensors, controller/EMS, smart loads	Operator, regulator, consumers	Reduced observability, poor dispatch, trust and accountability concerns
Physical/component events Line damage, converter fault, generation outage	PV/WT units, converter, feeder, storage system	Operator, planner/investor, consumers	Loss of capacity, repair cost, redundancy and investment trade-offs

Figure 2.2: Context diagram linking common HILP event types with affected microgrid components, stakeholders, and resilience concerns.

and may compromise the controller or data exchange, while physical/component failures directly reduce available capacity. Therefore, resilience metrics and control strategies are not universally interchangeable; they should be selected with awareness of the event type, affected components, time stage of the event, and stakeholder priorities.

The geographical context of a microgrid also shapes how resilience should be framed. For example, recent work on South African microgrids emphasises load-shedding as a frequent and often scheduled grid disconnection problem, where resilience depends on predictive preparation, such as pre-charging storage, pre-synchronising resources, and coordinating hierarchical control before the outage occurs [11]. This differs from the cyclone-prone coastal context considered in this thesis, where the main stressor is a weather-driven HILP event that can reduce renewable availability, increase dependence on BESS reserve, and expose priority loads to curtailment. It also differs from regions such as the north of the United Kingdom, where high wind availability may be a recurring operating condition rather than a low-probability event, although extreme storms may still create resilience concerns through

network damage or forced curtailment. Therefore, resilience studies should specify the geographical and operational context, because the relevant event type, affected components, stakeholder concerns, and suitable control strategy may change from one region to another.

The term *resilience* is used inconsistently across the literature and institutions. For clarity, this thesis follows the IEEE PES definition:

Resilience Definition

“The ability to withstand and reduce the magnitude and/or duration of disruptive events, which includes the capability to anticipate, absorb, adapt to, and/or rapidly recover from such an event” [12].

In practical terms, resilience encompasses three sequential stages pre-event, mitigation and preparedness, during-event absorption and response (performance degradation and containment), and post-event recovery and adaptation [13]. Within this framework, microgrids play a pivotal role: they can prioritise essential loads, operate in islanded mode, enable black-start capability, and accelerate local restoration following HILP events [14].

Methodologically, reliability and resilience address different questions. Reliability evaluates long-term average performance under defined contingencies, whereas resilience captures the dynamic trajectory of system response (degradation depth, duration, and recovery rate) during and after disruptive events. It also incorporates critical load priorities within the optimisation objective (e.g., priority-weighted served energy) and quantifies tail risks through resilience curves and risk-aware measures such as Value At Risk (VaR), Conditional Value At Risk (CVaR), or Conditional Expectation Of Energy Not Supplied (CEENS) [13], [15], [16], [17]. A side-by-side summary of their respective scopes, objectives, metrics, modelling perspectives, and microgrid roles is provided in Table 2.3. This distinction forms the basis of the proposed framework, in which the controllers optimise resilience-oriented operational indices and deliver interpretable decisions.

Table 2.3: Reliability versus resilience in power systems.

Attribute	Reliability	Resilience
Definition	Long-run ability to supply electricity within standards and with few interruptions; includes adequacy and security considerations [3], [4], [5].	Ability to anticipate, withstand, adapt to, and rapidly recover from high-impact disruptions (natural or cyber) [12], [13].
Event scope	Frequent, low-impact, credible contingencies (e.g., $N-1/N-2$), routine faults [3], [5].	HILP hazards; multi-element failures; cyber-physical attacks; cascading effects [9], [18], [19].
Time frame / phases	Assessed over long-run averages; steady-state adequacy/security during normal operation [3].	Covers pre-event (mitigation/preparedness), during-event (absorption/response), and post-event (recovery/adaptation) with explicit trajectories [13], [15].
Primary objectives	Minimise expected outage frequency/duration; maintain power quality; satisfy adequacy margins [3], [4].	Limit impact and duration of extreme events; prioritise critical services; accelerate restoration; learn/adapt post-event [12], [13].
Representative metrics	SAIDI, SAIFI, CAIDI (distribution); LOLE, LOLF, ENS/EENS, EIU (adequacy) [2].	Resilience curves (triangle/trapezoid/holistic) and areas; recovery time/rate; priority-weighted served-energy indices; SRI; risk-based VaR/CVaR of EENS; CEENS [9], [15], [16], [17].
Uncertainty treatment	Typically deterministic or average-probabilistic, centred on credible contingencies [3], [4].	Scenario-based and tail-risk focused (e.g., CVaR of EENS, CEENS); resilience metrics sometimes treated as distributions [2], [16], [17], [20].
Modelling lenses	Power-flow adequacy/security analysis; probabilistic resource adequacy and contingency analysis [3], [5].	Stress-testing with graph- and flow-based models; curve-based performance trajectories; stochastic/risk optimisation under HILP [15], [20], [21].
Planning actions	Capacity expansion; maintenance; protection coordination; redundancy for likely failures [3], [4].	Hardening (e.g., selective undergrounding), vegetation management, spares logistics, sectionalisation, <i>microgrid formation</i> and DER siting [22], [23].
Operational actions (during event)	N-1 security actions, remedial re-dispatch, voltage/reactive support; typically within intact topology [2].	Controlled islanding; priority load shedding; DER/BESS dispatch; MER; black-start and staged restoration [9], [10], [14], [24].
Role of micro-grids	Contribute to day-to-day reliability (local support, peak shaving, power quality).	Enable islanded operation for critical loads; contain disruptions; support black-start and faster local recovery; improve served-energy to priorities [14], [25], [26].

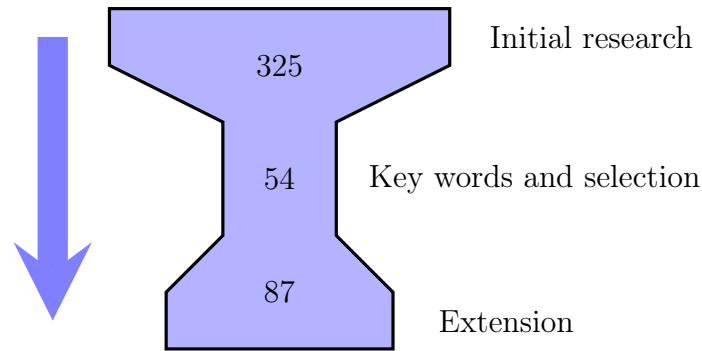


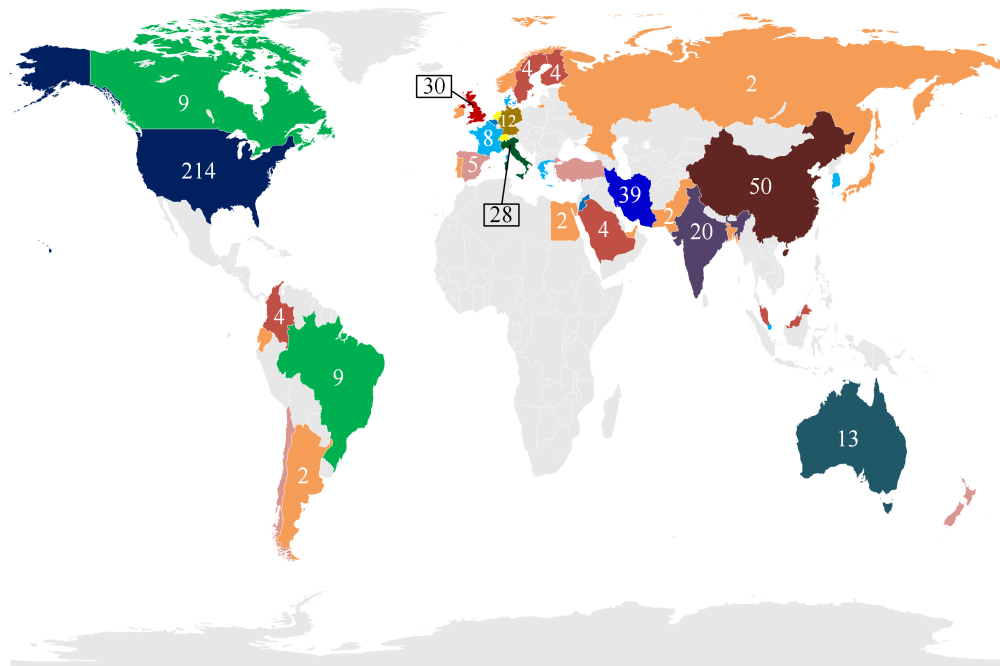
Figure 2.3: Methodology for selecting and including resilient metric articles.

2.2.2 Brief note on literature selection and context

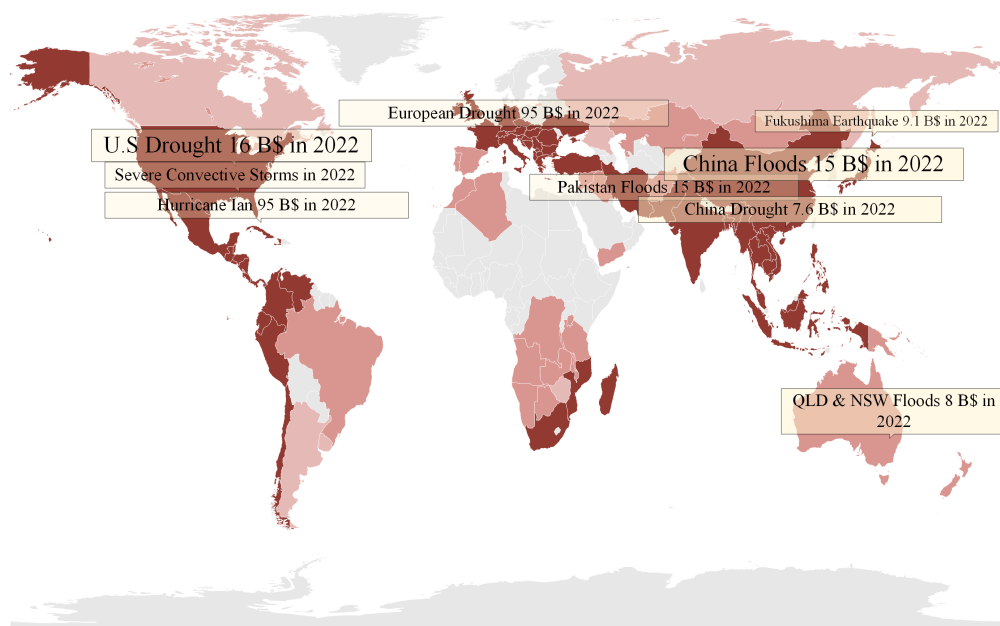
This section is based on our previously published review paper [2]. The literature base established in that study is adopted here as the foundation for discussing resilience metrics.

The earlier review [2] queried the Scopus database using *TITLE-ABS-KEY* (*("power system" OR "electrical grid" OR "microgrid") AND ("resilience" OR "resilient" OR "resiliency") AND ("quantitative" OR "quantification" OR "metric" OR "quantity")*) with filters (2019–2023, English). From an initial set of 325 records, 54 primary quantitative studies were included after full-text review, and the pool was expanded to 87 through forward and backward snowballing (Figure 2.3). Consistent with that review, the emphasis was placed on studies that define or apply quantitative resilience metrics, while purely qualitative checklists and narrative works were excluded.

Figure 2.4a shows that the United States, China, Iran, the United Kingdom, Italy, and India account for most recent publications on resilience metrics. This pattern aligns with global exposure: as shown by the loss-risk map in Figure 2.4b, these countries also face high economic risk from natural hazards. The overlap between research activity and vulnerability highlights the growing relevance of resilience-oriented studies.



(a)



(b)

Figure 2.4: a): Number of articles obtained via bibliometric analysis, published by countries. b): Global hazard total economic loss risk distributions (adapted from [27], [28]). Colour encodes the magnitude of losses.

2.2.3 Framework, applications, and metric taxonomy

A staged resilience framework (Figure 2.5) provides a reference for analysing power-system resilience across the pre-event, during-event, and post-event phases [13]. It operates across multiple decision layers such as planning, operations, protection, and logistics. Typical pre-event measures include asset hardening, vegetation management, spares preparation, and microgrid design. During the event, corrective actions such as controlled islanding, priority-based shedding, and DER/BESS/MER dispatch maintain continuity of supply and remedial reconfiguration. Post-event activities cover black-start sequencing, safe re-synchronisation, and restoration logistics. Figure 2.6 links these phases to practical applications of resilience metrics.

Studies on power-system resilience cluster around several themes. Most focus on operational assessments of HILP events, often adapting indices derived from reliability analysis. An increasing number incorporate risk awareness, uncertainty, or curve-based formulations. Graph-theoretic approaches dominate islanding and topology studies, while explicitly cyber-physical metrics remain in early development. To structure this diversity, the main approaches identified in the literature have been grouped into the following categories.

1. **Operational and planning metrics.** Operational metrics quantify service continuity during events, including priority-weighted served energy or curtailment, restoration time or rate, and triggers for corrective actions such as SRI-guided defensive islanding [9], [14], [22], [24], [25], [29]. Planning studies embed resilience in design loops using CEENS/EENS/Value Of Lost Load (VoLL) and cost-based indices (e.g., multi-energy Micro Grid (MG)s, BESS placement under seismic risk) [4], [16], [30], [31]. FLEP/trapezoid links both by tying degradation depth, duration, and recovery speed to objectives [15].

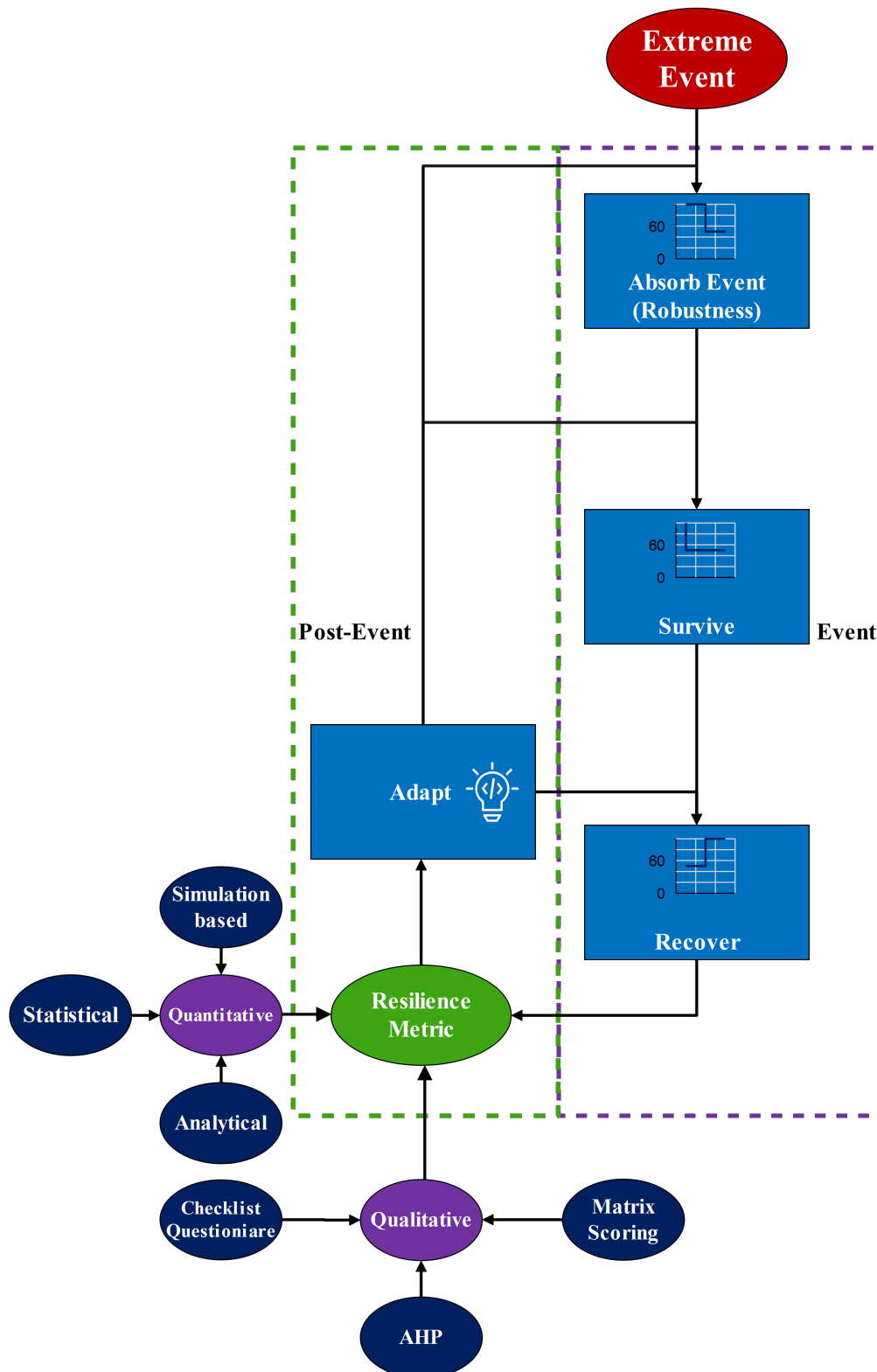


Figure 2.5: Resilience Framework illustrating the key stages of Absorb, Survive, Recover and Adapt, with a specific emphasis on Resilience Metrics [2].

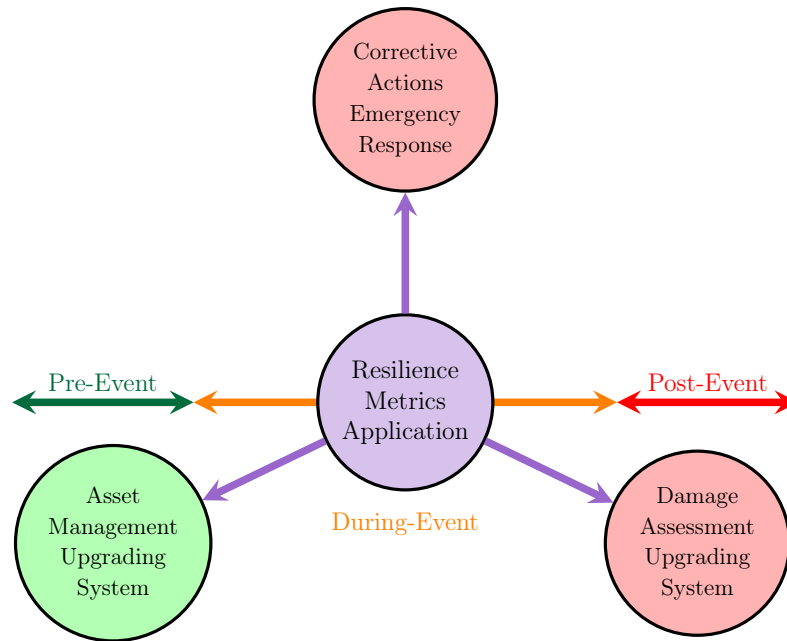


Figure 2.6: Applications of resilience metrics [2].

2. **Reliability-based metrics.** Classical adequacy and security measures (LOLE, LOLF /Loss Of Load Probability (LOLP), EENS, EIU, SAIFI, and SAIDI) summarise outage frequency, duration, and unsupplied energy [3], [4], [5], [21]. These metrics remain interpretable and comparable but underrepresent cascading or tail risk unless extended with risk terms.
3. **Uncertainty- and risk-aware metrics.** Risk-shifted adequacy (CEENS), tail measures (VaR/CVaR of EENS), or distribution-based formulations over scenarios/stages emphasise extremes outcomes rather than averages and support tail-robust optimisation [16], [17], [20], [32].
4. **Temporal staging.** Metrics align with pre-/during-/post-event phases: pre-event robustness/vulnerability (failure/repair models, islanding possibility), during-event absorption and degraded operation, and post-event restoration speed/efficiency/economy [22]. Extensions include Resilience Index Considering the Duration of Extreme Events (RICD) (initial resistance and effective hazard duration) and code-based encodings of

operational timescales capture temporal features [33], [34].

5. **System performance composites.** Composite vectors capture outage counts (K), recovery difficulty (G combining Mean Time To Repair (MTTR) and logistics), weakness margins, and economics (outage penalties, repair costs) useful for cross-scenario ranking and restoration planning [10], [21], [35], [36]. Cyber-physical composites (e.g., Cyber-Physical Resilience Metric (CPARM); attack success ratios plus physical penalty terms) reflect information and communication dependencies [37], [38].
6. **Graph-theoretic/topological metrics.** Algebraic connectivity, redundancy, Effective Shortest Paths Between Critical Nodes (ESPC), centrality/criticality, and switching feasibility screen vulnerable structures, inform islanding partitions, and rank hardening options; multi-criteria schemes (Analytic Hierarchy Process (AHP))/Choquet combine these with feasibility checks [18], [39], [40], [41], [42]. Hybrid schemes combine these with physical indicators (Line Strength Index (LSI)/Pole Strength Index (PSI) for lines/poles) enrich purely topological views; MG-specific vulnerability/criticality ranks appear in operations [43], [44].
7. **Curve-based trajectories.** Triangular and trapezoidal curves summarise degradation depth, duration, and recovery speed, while smoothness/area-based measures capture total losses [15], [45], [46]. Our previous work [2] synthesised these into a holistic resilience curve (Figure 2.8) incorporating early DER support, multiple post-shock degradations, and distinct operational vs. structural recovery. Applied to Environment for Analysis of Geo-Located Energy Information (EAGLE-I) outage data from Hurricane Harvey (Jefferson County, TX), the holistic curve (Figure 2.9) achieved the smallest deviation from real outage data, outperforming triangular and trapezoidal forms (Table 2.4) [2], [15], [19], [46].

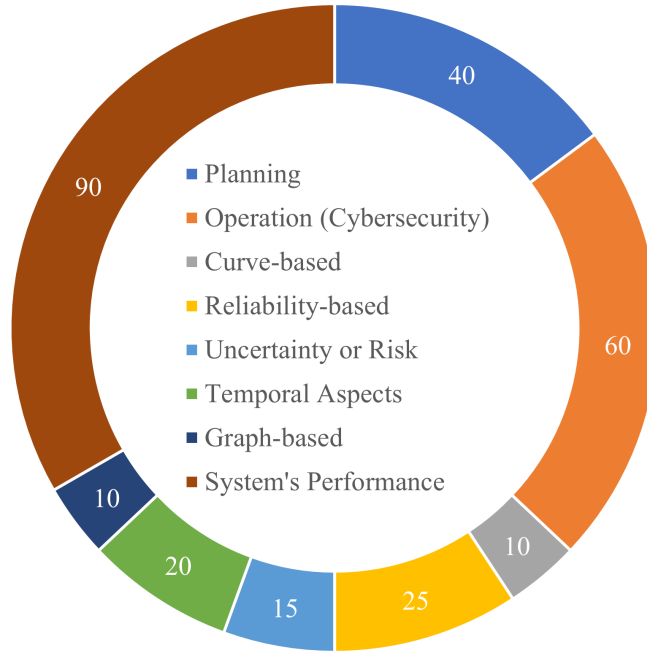


Figure 2.7: Categories of resilience metrics and associated percentage of reviewed works [2].

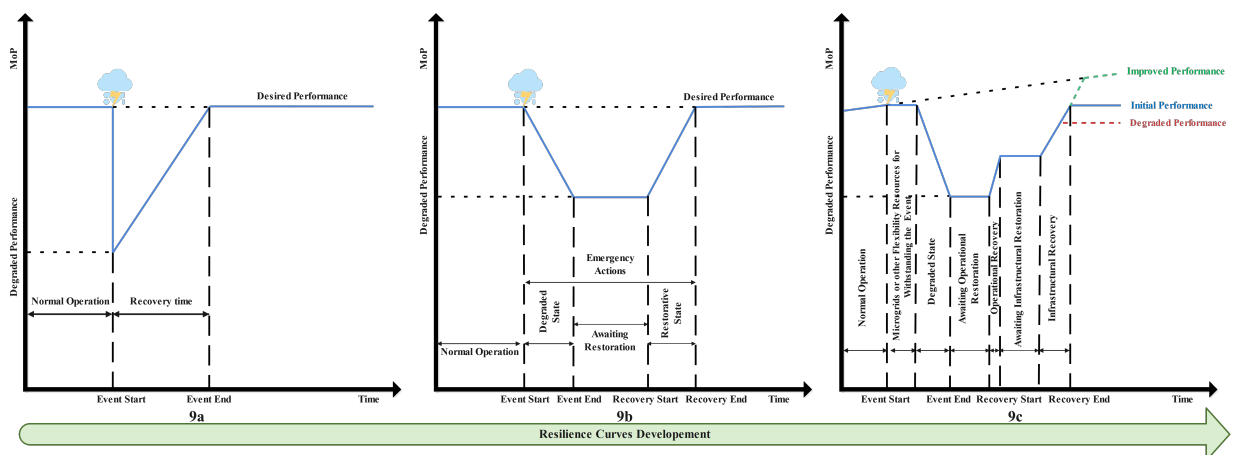


Figure 2.8: Evolution of resilience curves.

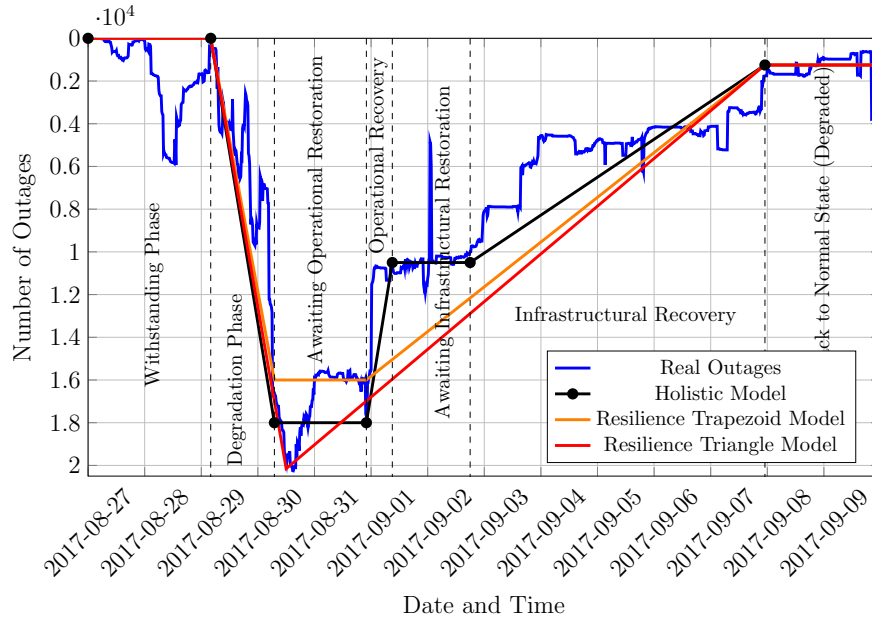


Figure 2.9: Comparison of the proposed resilience curve with real HILP event (Hurricane Harvey).

Table 2.4: Comparison of resilience curves against a real HILP event (Hurricane Harvey) using the cumulative impact $\mathcal{I}_{RC}$ (outage-hours) from Equation (2.1). Errors are relative to the real-outage baseline.

Resilience curve	$\mathcal{I}_{RC}$ (outage-hours)	Error (%)
<i>Real outage (baseline)</i>	2,090,939.0	—
Resilience triangle	2,556,000.0	22.2%
Resilience trapezoid	2,358,976.0	12.8%
Holistic resilience	2,244,064.0	7.3%

For a resilience curve (RC), the cumulative impact $\mathcal{I}_{RC}$ is given by

$$\mathcal{I}_{RC} = \int_{t_d}^{t_e} RC(t)dt, \quad (2.1)$$

representing the total outage impact (outage-hours) between the start of degradation t_d and the end of the event t_e . In practice, it is often difficult to identify the event (t_d) before cascading failures begin. The deviation between the fitted curve and real outage trajectory quantifies how accurately the curve captures the event's full impact.

Microgrid-centric, operational lens. Given the thesis focus on resilient microgrid energy management, evaluation adopts operational metrics that explicitly value continuity of critical services. The priority-weighted served-energy index is therefore defined as

$$\text{RI} = \frac{\sum_{t=t_1}^{t_2} [\omega_{CL}(P_t^{\text{Demand,CL}} - P_t^{\text{ENS,CL}}) + \omega_{RL}(P_t^{\text{Demand,RL}} - P_t^{\text{ENS,RL}})]}{\sum_{t=t_1}^{t_2} P_t^{\text{Demand}}}, \quad (2.2)$$

where $P_t^{\text{ENS},\cdot}$ is ENS (curtailment) and $\omega_{CL} > \omega_{RL} \geq 0$ encode explicit priority for critical vs. regular loads. The numerator aggregates served energy (demand minus ENS) with priority weights; the denominator is the total demand over the horizon $[t_1, t_2]$. $\text{RI} \in [0, 1]$ equals one only when all loads are met and is more sensitive to shortfalls at critical loads when $\omega_{CL} \gg \omega_{RL}$ [47], [48], [49]. This class of metrics is widely adopted for microgrid operation because it is priority- and horizon-aware, and compatible with MPC or DRL formulations for islanding, dispatch, and shedding decisions.

Why this deep dive, and what it changes. This survey was essential for identifying appropriate resilience metrics for the research objectives. Transitioning from reliability to risk- and curve-based perspectives established a vocabulary of resilience tailored to microgrid operations. Our earlier review [2] highlighted ambiguity and urged use of standard, quantitative metrics; still sound advice, but one lesson stands out: *context matters*. The “right” metric depends on the controller (rule-based, MPC, Reinforcement Learning (RL)), the targeted phase (pre-event planning vs. during-event operation), the study scope (single vs. networked MGs), and the HILP origin (natural, cyber, physical).

Problem Statements Linked to Research Gaps

- ✓ **Gap 1:** Clarifies metric choice: select resilience metrics appropriate to the context so results are interpretable and comparable for researchers (including new researchers).

Accordingly, the thesis standardises on a simple operational core, the priority-weighted served-energy index (RI, Equation (2.2)), because it directly values continuity of critical services. With MPC in the loop, each update yields an H -hour look-ahead of demand and curtailment; forecast counterparts, the Expected Resilience Index (ERI) and Expected Resilience Curve (ERC), summarise anticipated performance over that horizon (formal definitions in Sections 3.4.2 and 4.2). For trajectory comparison, the holistic resilience curve, validated against EAGLE-I outage data, replaces the traditional triangle/trapezoid forms. The achievement of this first objective is summarised below.

Objective Achieved

✓ **Objective 1 – Clarify resilience measurement.** This objective asked: *How can resilience be measured for microgrids under HILP conditions, and which metrics are most suitable for operations with priority loads while remaining comparable across studies?*

The objective has been achieved through two complementary outcomes. First, the published review [2] classified resilience metrics, reduced ambiguity, and recommended Equation 2.2 as a practical measure for operational assessment. Second, the thesis extends this foundation by introducing expected resilience metrics (ERI, ERC) within the MPC framework to evaluate forecast-based resilience over short-term operating horizons.

Together, these form a consistent basis for quantifying resilience in both assessment and control contexts.

The following sections focus on designing and analysing microgrid control policies that protect critical loads during HILP events.

2.3 Review of control techniques in microgrids

Section Summary

Section 2.2 established an operational view of resilience in microgrids, clarifying performance objectives under HILP events and formalising priority-weighted service continuity. Building on that foundation, this section reviews control approaches that can realise those objectives in practice. The focus is on two complementary families: *MPC*, representing model-based, constraint-aware optimisation over a receding horizon; and *DRL*, representing data-driven, model-free policies with near-instantaneous inference.

The review highlights three core aspects:

1. the progression from conventional and heuristic methods to MPC, with emphasis on how the proposed MPC framework uniquely combines demand response, resilience, and battery-life modelling,
2. the strengths and limitations of MPC, setting up DRL as a response to its main shortcomings; and
3. the emerging role of DRL in resilience-oriented microgrid control, particularly its adaptability and computational efficiency but lack of interpretability, which motivates the final part of this chapter.

The discussion proceeds with MPC in Section 2.3.1, followed by DRL in Section 2.3.2, and concludes with a synthesis that motivates the methodological choices adopted in this thesis.

Table 2.5: Comparative Analysis of Microgrid Design and Operational Literature [49].

Ref	Technology	Mode	DR	Resiliency	Battery Life	Software	Real Case	Planning (P) Operation (O)
[50]	WT-PV-MT	IS/GC	-	-	-	CPLEX(MILP)	-	P
[66]	WT-PV-BESS	IS	-	-	✓	MATLAB(YALMIP)	-	O
[62]	WT-PV-BESS-MT-DiG-FC	IS/GC	-	-	-	MATLAB(SSA)	-	O
[56]	ESS	IS	-	✓	-	IPOPT(AGA)	-	P
[55]	WT-PV-ESS	GC	-	✓	-	MATLAB-CPLEX	-	P
[54]	WT-PV-BESS	IS	-	-	-	MATLAB	-	P
[64]	WT-PV-GT-FC-CG	IS/GC	✓	-	-	MATLAB(CSAJAYA)	-	O
[63]	WT-PV-BESS-DiG-FC	GC	-	-	-	-(IWOA)	-	O
[57]	Distributed Generation (DG)	IS	✓	✓	-	GAMS(CPLEX)	-	P
[67]	WT-PV-DG	GC	-	-	✓	CPLEX(MILP)	-	P
[68]	WT-PV-BESS	IS	✓	-	✓	MATLAB-GAMS	-	O
[69]	WT-PV-BESS	IS	-	✓	✓	Pyomo (MILP)	-	O
[58]	PV-BESS	GC	-	-	-	GAMS(NLP/MILP)	✓	P
[52]	WT-PV-BESS-BM	IS/GC	-	-	-	HOMER Pro	✓	P
[51]	WT-PV-BESS-EV	GC	-	-	-	-(MINLP)	✓	P
[59]	WT-BESS-DiG	GC	-	-	-	MATLAB(Heuristic)	✓	O
[53]	WT-PV-BESS-DiG	IS	-	-	-	MATLAB(DE)	✓	P
[60]	WT-PV-BM-EV-HP	GC	✓	-	-	-(MILP)	✓	O
[70]	WT-PV-BESS-DiG	IS	-	-	-	HOMER Pro DiGSILENT	✓	P/O
[71]	WT-PV-BSS-BM	IS	-	-	-	MATLAB-MILP	✓	P/O
[72]	WT-PV-ESS-DiG-HP-Combined Heat and Power (CHP)	IS	-	✓	-	-(DRL)	✓	P
Part of Thesis	WT-PV-BESS	IS/GC	✓	✓	✓	HOMER Pro Pyomo (MILP)	✓	P/O

2.3.1 Model Predictive Control

Conventional microgrid control and scheduling are typically formulated as deterministic or heuristic optimisation problems over fixed horizons (day-ahead or hour-ahead) [49]. Mixed-integer formulations capture unit commitment, charging and discharging logic, and network constraints and are widely applied in both planning and operational contexts [50], [51], [52], [53], [54], [55]. Multi level or stage designs separate long-term investment from short-term dispatch, often adding robustness against outages or cyber threats [56], [57], [58], [59], [60], [61]. Meta-heuristics (e.g., Salp Swarm Algorithm (SSA), Improved Whale Optimization Algorithm (IWOA), Differential Evolution (DE)) are also employed for cost and emission minimisation or siting decisions [62], [63], [64]. These approaches are attractive for their modelling flexibility and their ability to handle discrete decisions; however, resilience is often treated implicitly (through generic reliability indices or ad-hoc shedding penalties), and battery longevity is commonly approximated by SOC bounds or switching costs. As summarised in Table 2.5, location-specific stress testing under HILP conditions is rare, and simplified battery life variants dominate [65], [66].

MPC Explanation

MPC, on the other hand, formulates operation as a receding-horizon optimisation: at each step an optimisation problem is solved to obtain a sequence of future control actions; only the first action is applied, and the problem is resolved at the next step with updated states and forecasts [73].

In microgrid applications, MPC has been deployed in centralised, decentralised, and distributed forms to coordinate distributed renewable energy sources, batteries, and controllable loads. Consensus-based schemes let local controllers for RES and BESS solve coupled MPC problems [74]. Household-scale Energy Management System (EMS) designs use MPC to jointly optimise battery SOC, a micro-CHP unit, and an auxiliary gas boiler [75]. Decentralised MPC with day-ahead PV/load forecasts schedules thermal generation in rural communities [76]. Mixed-integer MPC for smart-grid scheduling maximises self-consumption under time-varying tariffs and co-ordinates shared resources and grid exchanges [77]. Two-stage MPC variants also appear, decoupling unit commitment from economic dispatch to reduce online complexity while preserving near-optimal operation [78].

Across these studies, MPC objectives usually combine operating cost, curtailment, and power-imbalance penalties, sometimes with soft priorities for load classes. However, resilience is frequently *indirect*, emerging as a by-product of economic scheduling or ad-hoc shedding penalties rather than as an explicit, priority-weighted objective tuned to HILP conditions [49]. As indicated in Table 2.5, operation-stage analyses with location-specific, real-case disturbance profiles are limited, whereas real data are more common in the planning stage.

Battery degradation modelling also varies in complexity [49]. Common methods constrain the SOC within fixed bands or penalise switching, which only indirectly relates to cycle ageing [66]. Discharged energy-based linear variants are also employed in distributed

MPC and scheduling settings to approximate degradation at low computational cost [68]. More explicit formulations embed DOD dependence: a prevalent route discretises DOD into bands and assigns a constant cycle-life cost per band (a *stepwise* linear model) to retain mixed-integer tractability [67], [79]. Building on this line of work, this thesis introduces a *piecewise-linear* BLC–DOD formulation ([49]) that preserves MILP/MPC compatibility. The full model is presented in the later modelling Section 3.4.2.2. Overall, detailed ageing costs remain the exception rather than the norm in microgrid energy management studies. As reflected in Table 2.5, location-specific operation under HILP conditions is seldom combined with battery ageing models.

Problem Statements Linked to Research Gaps

- ✓ **Gaps 2 and 6:** Handles uncertainty and variability: produce timely decisions when forecasts are imperfect and conditions change rapidly.
- ✓ **Gap 3:** Protects battery life: incorporate DOD–aware ageing rather than wide SOC bands or simple penalties.
- ✓ **Gap 4:** Addresses HILP conditions: remain effective under rare, severe events (e.g., cyclones) using location-specific, real disturbance data when available.

MPC’s principal strength is dynamic optimisation with feedback: it anticipates future conditions over a prediction horizon and corrects actions as new information arrives, which enhances robustness to forecast errors and operating variability [80]. Centralised MPC offers globally optimal schedules, whereas distributed variants scale better for larger networks. However, MPC performance depends on model accuracy and forecasts; online solution time grows rapidly with horizon length, decision dimensionality, and discrete variables, which can challenge large-scale or fast-sampling deployments [81], [82]. Although specific reports show sub-second horizons or year-long simulations executed efficiently for certain problem sizes [77], [78], many studies do not report EMS computation times, and assume ideal commu-

nication infrastructures. Hence, despite its strengths, MPC can become computationally burdensome for real-time microgrid operation, especially under HILP events.

In summary, classical optimisation (including MPC) offers foresight and principled constraint handling but remains sensitive to model and forecast errors, and its online solve time scales poorly with horizon length, discreteness, and network size. Robust and stochastic versions alleviate some issues at the cost of conservatism or heavy distributional assumptions, while meta-heuristics trade optimality for flexibility. Figure 2.10 compares these trade-offs [83]. These limitations, especially under HILP conditions, motivate shifting the computational burden offline and using DRL for fast (inference-time) decisions, then benchmarking and constraining it against MPC.

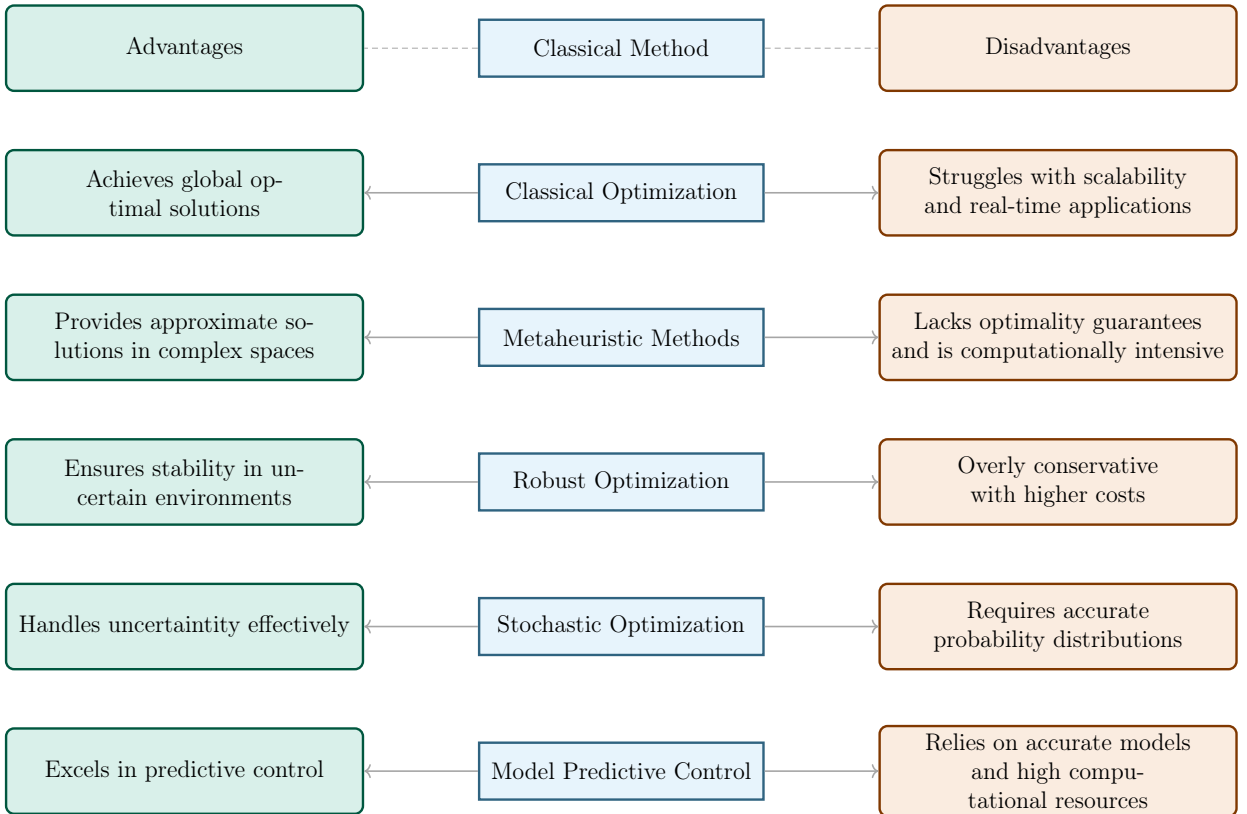


Figure 2.10: A comparison of different classical control methods.

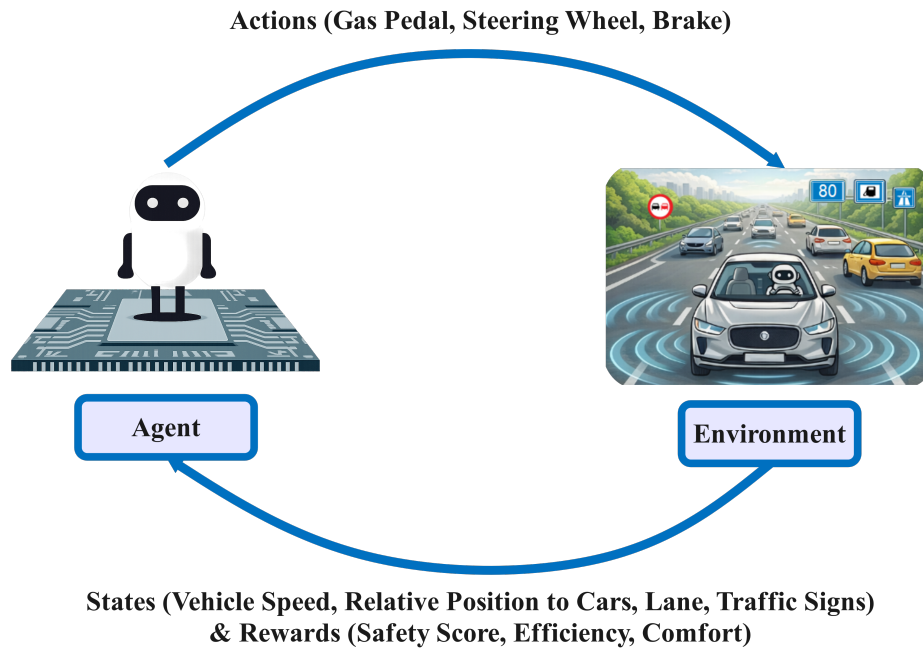


Figure 2.11: Agent and Environment Interactions.

2.3.2 Deep Reinforcement Learning

RL Explanation

RL formalises sequential decision-making as an interaction between an *agent* and an *environment*: at each time step, the agent observes a state, takes an action, and receives a scalar reward that guides its long-term behaviour (Figure 2.11).

Unlike supervised learning (which relies on labelled data) or unsupervised learning (which uncovers patterns), RL *learns by interacting*, optimising behaviour directly from experience over time [84], [85], [86]. Consider autonomous driving. The *state* can include speed, lane position, distances to other vehicles, traffic lights/signs, and road friction. The *actions* are continuous control signals for gas and brake pedal, and the steering wheel angle. The *reward* balances safety (no collisions, lane keeping), comfort (smooth control), and efficiency (progress, energy use). By interacting with the road dynamics (the environment), the agent

discovers a policy that gets passengers to their destination safely and efficiently, without being given labeled “optimal” actions.

RL problems are typically modelled as *Markov Decision Process (MDP)* [85], defined by the tuple $(\mathcal{S}, \mathcal{A}, p, r, \gamma)$, where $\mathcal{S}$ denotes the state space, $\mathcal{A}$ the action space, $p(s', r \mid s, a)$ the transition probability distribution giving the likelihood of moving to the next state s' and receiving reward r after taking action a in state s , $r : \mathcal{S} \times \mathcal{A} \rightarrow \mathbb{R}$ the reward function, and $\gamma \in [0, 1]$ the discount factor weighting future rewards. The objective is to find a policy $\pi(a \mid s)$, a mapping from states to a probability distribution over actions, that maximises the expected discounted return:

$$G_t = \sum_{k=0}^{\infty} \gamma^k R_{t+k+1}, \quad (2.3)$$

where R_{t+k+1} is the reward received $k+1$ steps after time t .

Classical solution methods fall into several categories. Monte Carlo (MC) methods estimate value functions from complete episodes by averaging returns, relying only on sampled experience and are thus *model-free*. In contrast, Dynamic Programming (DP) methods such as policy iteration and value iteration assume full knowledge of the environment’s transition dynamics and rewards, making them *model-based*. The Multi-Armed Bandit (MAB) problem represents a single-state decision setting in which an agent repeatedly chooses among actions with unknown reward distributions, highlighting the fundamental exploration–exploitation trade-off [85].

Temporal-Difference (TD) learning combines principles from MC and DP by updating value estimates from partial experience while *bootstrapping*, that is, by using existing value predictions to update current estimates rather than waiting for complete returns. The

simplest case, TD(0), updates the value function as

$$V(S_t) \leftarrow V(S_t) + \alpha [R_{t+1} + \gamma V(S_{t+1}) - V(S_t)], \quad (2.4)$$

where α is the learning rate and the term $\gamma V(S_{t+1})$ represents the bootstrapped estimate of future rewards. This enables incremental, online prediction of long-term returns [85].

For control, Q-learning [85] is an *off-policy* method: it learns the value of the optimal greedy policy even while following a different (typically exploratory) behaviour policy. Its update is

$$Q(S_t, A_t) \leftarrow Q(S_t, A_t) + \alpha [R_{t+1} + \gamma \max_a Q(S_{t+1}, a) - Q(S_t, A_t)]. \quad (2.5)$$

State-Action-Reward-State-Action (SARSA), by contrast, is an *on-policy* method: it evaluates and improves the same policy that is generating behaviour, updating with the actual next action A_{t+1} chosen under the behaviour policy:

$$Q(S_t, A_t) \leftarrow Q(S_t, A_t) + \alpha [R_{t+1} + \gamma Q(S_{t+1}, A_{t+1}) - Q(S_t, A_t)]. \quad (2.6)$$

In summary, MC, TD, Q-learning, and SARSA are *model-free*, as they require only sampled experience $(S_t, A_t, R_{t+1}, S_{t+1})$ and not prior knowledge of transition probabilities or rewards. DP is *model-based*, since it presumes full access to the environment's dynamics. Within model-free control, the distinction between *on-policy* and *off-policy* is central: on-policy methods (e.g., SARSA) learn the value of the policy actually executed, including exploratory actions, whereas off-policy methods (e.g., Q-learning) learn the value of an optimal target policy independently of the behaviour used to explore. Together, model-based and model-free approaches backing modern (DRL) [86], as illustrated in Figure 2.12.

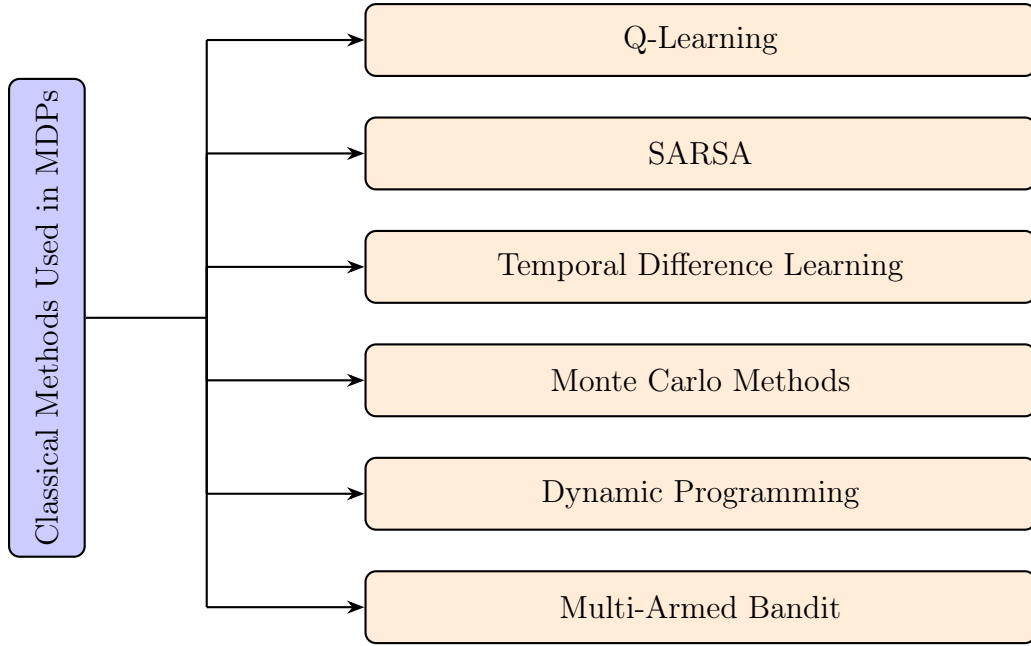


Figure 2.12: Popular Techniques to Solve MDPs.

DRL Explanation

DRL extends classical reinforcement learning by using DNNs as function approximators for value functions and policies in high-dimensional or partially observed environments [86]. In this setting, a policy is represented as a parameterised function,

$$\pi_{\theta}(a \mid s) = \mathbb{P}(A_t = a \mid S_t = s; \theta), \quad (2.7)$$

where a DNN with parameters θ maps observations (or states) to a probability distribution over actions. This parameterisation enables the agent to learn complex, non-linear relationships between observations and actions that are infeasible to capture with tabular or linear methods.

The power of DRL is well illustrated by game benchmarks. For example, OpenAI Five employed PPO with deep neural policies to achieve expert-level performance in the real-time strategy game *Dota 2*, a multi-agent environment characterised by delayed rewards, partial

observability, and long-horizon decision-making [87]. In this context, the agent learned to trade short-term losses for long-term advantage, such as sacrificing immediate resources or accepting early setbacks (like early “buyback” from death) to improve overall team success. These behaviours can be attributed to the ability of deep networks to encode rich policies, meaning that they can learn intricate correlations and interactions between high-dimensional observations and action choices across extended time horizons.

The same representational flexibility, however, leads to opacity: the internal network representations and the rationale behind specific actions are not directly interpretable. This “black-box” nature of DRL motivates the need for explainability methods, which are discussed later in this chapter.

To illustrate the difference between classical RL and DRL, consider Q-learning (Equation 2.5), which updates action values in a tabular form. In DRL, this table is replaced by a parameterised function approximator, typically a neural network, which leads to the formulation of the *Deep Q-Network (DQN)* [86], [88].

A DQN comprises three core components. First, a neural network $Q(S_t, A_t; \theta)$ with parameters θ approximates the action-value function; for instance, in Atari benchmarks, raw pixel observations are mapped to Q-values for each discrete action. Second, an *experience replay* buffer stores transitions $(S_t, A_t, R_{t+1}, S_{t+1})$, from which random mini-batches are sampled to break temporal correlations and improve generalisation. Third, a separate *target network* with parameters θ^- computes stable targets, with weights periodically synchronised from the main network. Training minimises the mean-squared TD loss:

$$L(\theta) = \mathbb{E} \left[\left(R_{t+1} + \gamma \max_a Q(S_{t+1}, a; \theta^-) - Q(S_t, A_t; \theta) \right)^2 \right], \quad (2.8)$$

using stochastic gradient descent. An ϵ -greedy policy selects a random action with probability ϵ and otherwise chooses the greedy action $\arg \max_a Q(S_t, a; \theta)$; the target network is

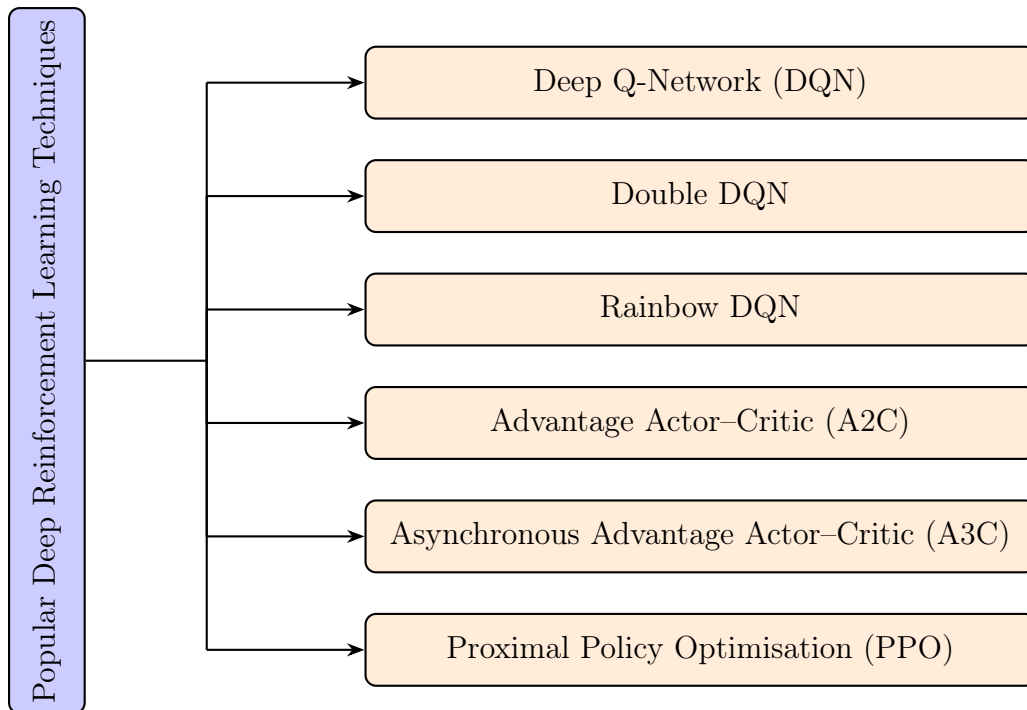


Figure 2.13: Popular Deep Reinforcement Learning Techniques.

periodically synchronised; and ϵ is gradually decayed to shift from exploration to exploitation. This recipe enabled human-level Atari performance and has since been adapted to domains such as robotics, autonomous driving, finance, and healthcare [86].

For continuous control and set-points, *actor-critic* methods dominate. On-policy such as Advantage Actor-Critic (A2C), Asynchronous Advantage Actor-Critic (A3C), and PPO, and off-policy such as Deep Deterministic Policy Gradient (DDPG), Twin Delayed Deep Deterministic Policy Gradient (TD3), and Soft Actor-Critic (SAC) balance sample efficiency, stability, and exploration [86], [89], [90]. Among these, PPO is discussed in detail in Chapter 3, as it forms a key component of this thesis framework. Figures 2.13 and 2.14 provide a compact overview of major DRL families and typical application areas.

With the foundations in place, it is natural to turn from general purpose DRL to its stage in microgrid operation and resilience. To illustrate this perspective, consider a microgrid as a living system: weather conditions change, loads fluctuate, components fail

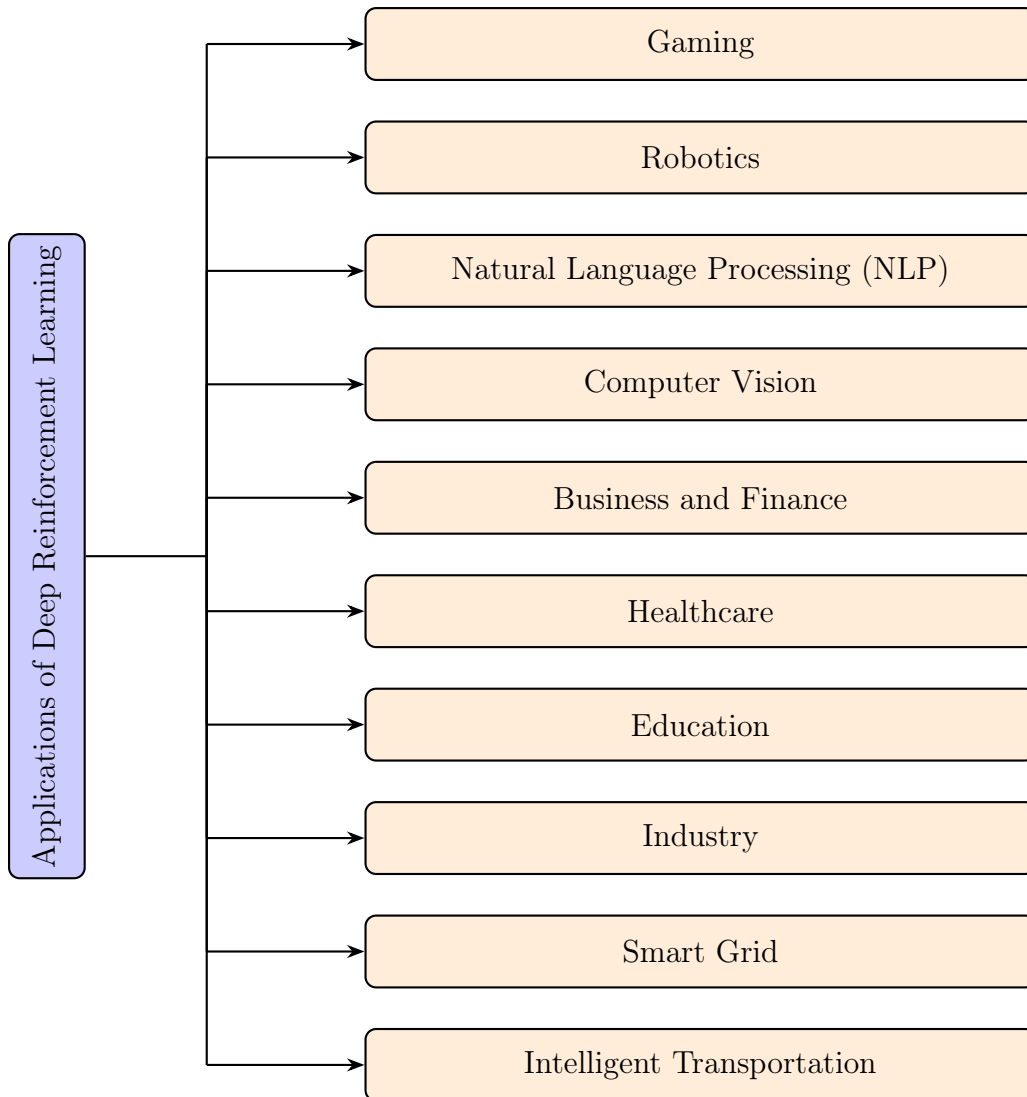


Figure 2.14: Applications of Deep Reinforcement Learning.

and recover, and topology is restructured through switching. DRL fits this rhythm. The *state* gathers forecasts, sensors, and network status; the *actions* range from reconfiguration and dispatch to storage scheduling and market bids; and the *reward* weaves together reliability, resilience, cost, and emissions. In other words, the same ingredients that made DRL compelling for strategic games and robotics reappear, only now the field is critical infrastructure. The next subsection reviews DRL applications for microgrid resilience, linking algorithm families to specific operational tasks and identifying the research gaps addressed in this thesis.

2.3.2.1 DRL in Microgrid Resilience

DRL has become a central approach for enhancing microgrid resilience because it can learn policies that react quickly to uncertainty and disturbances without relying on fully specified system models. In practice, applications fall into two categories aligned with resilience goals: *operational resilience* (service restoration, topology reconfiguration, cyberattack response, robust EMS under volatility) and *planning for resilience* (capacity and technology choices that reduce risk and cost over the long term). A Scopus search (TITLE-ABS-KEY: (*reinforcement AND learning OR deep AND reinforcement AND learning OR DRL OR machine AND learning*) AND (*microgrid OR micro AND grid*) AND (*resilience OR resilient OR resiliency*)) returned 61 items in total: 36 journal articles, 19 conference papers, 5 review articles, and 2 conference reviews, of which 58 were accessible. Together with Table 2.6, this provides a current overview of the field, covering algorithmic approaches, power generation types, objectives and rewards, resiliency domains, operational versus planning focus, benchmarking against classical methods, explainability, real-world case studies, and uncertainty quantification.

On the energy management side, the focus here, DRL co-optimises storage usage, load balancing, and imports/exports under volatile generation, demand, and prices [97], [102], [104], [109], [115]. Many studies prioritise fast recovery and continuity of supply during HILP events often by shaping rewards around load restoration and operating costs [91], [94], [119]. Robustness to connectivity changes is examined in [96] across grid-connected, intermittent, and islanded modes. Uncertainty is injected during training via domain randomisation and auxiliary exploration signals [93], or coupled with forecasting modules in distributed PPO dispatch [104]. Bayesian variants expose epistemic/aleatoric uncertainty, e.g., Bayesian DDPG and DDQN for coordination under partial observability and communication failures [101], [110].

Table 2.6: Comparative Review of DRL-based Microgrid Resiliency Studies.

Ref.	Algorithm	Power Generation	Objective(s) Reward(s)	Resiliency Domain	Operation (O) Planning (P)	Comparison with Classical Methods	Explainability	Real-world Case Study	Uncertainty Quantification
Part of Thesis	PPO	Solar PV, Wind, BESS	Resiliency, Battery Life	Energy Management	O	Yes (MPC)	Yes (SHAP, LIME)	Yes (India, Cyclone Layla)	Yes (Perlin Noise, Scenario Augmentation)
[91]	Improved TD3	PV, Wind, Diesel DGs, Building AC systems	Load restoration maximization, DG cost minimization, Occupants' comfort maintenance	Energy Management, Load Restoration	O	No (Comparing Different DRL Algorithms)	No	No (IEEE 33 and 123-node test systems)	Yes (Endogenous from load flexibility, exogenous from renewable, load, ambient temperature)
[92]	Long Short-Term Memory (LSTM)-SAC	PV, Wind, CHP, Diesel DGs, Storage, EVs	Profit/Loss, Surplus/Deficit power, Fault Duration	Energy Management, Energy Trading	O	No (DC-Optimal Power Flow (OPF) with PSO)	No	No (IEEE 14-bus system)	Yes (stochastic load, generation, faults)
[93]	PPO + DDPG + Domain Randomization + Random Network Distillation	PV, Diesel Genset, Macrogrid, Storage (Battery)	Cost minimization (operation cost, loss load penalty, overgeneration penalty)	Energy Management	O	Yes (Compared with MPC and RBC)	No	No (simulated microgrid, no specific real HILP event, only assumed uncertainty)	Yes (Battery model uncertainty, Forecast uncertainty via noise injection)
[94]	DQN	PV, FDGs, MEGs	Restored load, Post-restoration reliability of critical loads	MG Formation and Emergency Restoration	O	YES (MILP, deterministic & stochastic optimization for comparing RI)	No	No (IEEE 13-, 37-, and 123-bus systems; HILP source assumed but not specified)	Yes (Load and PV uncertainty modeled via MCS and forecast errors)
[95]	Deep Q-Learning	DGs	Maximizing defender's utility, minimizing system instability against rootkit attacks	Cybersecurity (rootkit detection and mitigation)	O	NO (Compared with static ANN predictor and classical Q-learning)	No	No (simulation on 4-DG and 10-DG microgrid models)	Yes (attack randomness, stochastic attack modeling)
[96]	PPO, A2C	PV, ESS	Cost and Demand Coverage under Connectivity Scenarios	Energy Management	O	No (Greedy Heuristic and Random EMS)	No	No (Simulated MG with real PV/load data; no specific HILP like weather)	Yes (Stochastic PV and Load using NREL data)
[97]	Q-learning	PV, Wind, Micro-turbine, BESS	Cost minimization, fault-tolerant operation	Energy Management	O	No	No	No (Simulation without physical or event-driven real case; HILP failure assumed abstractly)	No (Real-time policy negates prediction/uncertainty modeling)
[98]	HARIMA-DRL	Not explicitly stated (Synthetic test system)	Resilience improvement, RMSE minimization	Energy Management, Post-Disaster Restoration	O	NO (Compared with Convolutional Neural Network (CNN)-LSTM and GRM, not classical optimization)	No	No (Simulated disasters: floods, cyclones, storms on ACTIVSg2000)	Yes (Data-driven RMSE)
[72]	Double Deep Q-Learning (DDQL)	Solar, Wind, Diesel, Hydro, CHP, Li-ion, Lead-acid, Vanadium redox, Flywheel, PSH	Minimize total cost, Maximize resilience (via LOLP, CCP), Minimize GHG (CO ₂ , SO ₂ , NO _x)	Long-term Expansion Planning (Resilience, Environmental Goals)	P	Yes (Greedy Algorithm)	No	Yes (Des Moines, USA; HILP from SAIFI/CAIDI stats, but impact simulation not explored)	Yes (Price and Load Scenarios, Degradation, Weather data, GHG penalties)
[99]	DDQL	PV, Onshore Offshore Wind, CHP, Diesel, BESS (Li-ion, VRB, Lead-acid, Flywheel)	Minimize total cost, Enhance resilience, Account for degradation and outages	Planning	P	No	No	Yes (Synthetic microgrid, Atlantic City, based on historical SAIFI and CAIDI, 2013–2020)	Yes (Battery degradation modeling and renewable failure probability with uniform distribution, the root of HILP is not mentioned)

Table 2.6: *Continue . . .*

Ref.	Algorithm	Power Generation	Objective(s) Reward(s)	Resiliency Domain	Operation (O) Planning (P)	Comparison with Classical Methods	Explainability	Real-world Case Study	Uncertainty Quantification
[100]	A3C	IESO Data, IEEE 9-bus, 14-bus	DR profit maximization, Cyberattack detection (False Data Injection (FDI)), Demand-supply pricing balance	Cyber Resilience, Energy Management	O	No	No	No (Synthetic simulations on IEEE bus systems, IESO dataset)	Yes (Cyberattack simulation: FDI, stochastic pricing and demand patterns)
[101]	BDDQN	RES (Solar), ESS, Main Grid	Profit maximization (RES, ESS), Cost minimization (DSL), Energy trading balance	Energy Management with Communication Failure	O	No (Compared with uncoordinated DQN)	No	No (Simulated scenario with communication failure, no real-world HILP event)	Yes (Bayesian belief updating under uncertain agent observability)
[102]	PPO, SAC	Wind, BSS, Heaters, Main Grid	Cost minimization, Load shifting, Efficient battery use	Energy Management	O	No (Baseline: no battery-load scheduling)	No	Yes (Finland, real wind/price/temp data from Fingrid Oyj; no specific HILP)	Yes (Real stochastic input from actual datasets, multiple what-if scenarios)
[103]	SAC	ESS (DC source + inverter), Cordova microgrid	Minimize frequency deviation and ROCOF during disturbances	Frequency Control (Fast Frequency Response)	O	NO (vs. C-code training RL setup)	No	No (Simulated Cordova benchmark, synthetic load step events)	Yes (Kalman Filter-based estimation, step load disturbances, co-simulation uncertainty)
[104]	PPO (Distributed)	PV, Wind, Hydrogen (via Methanol FC), NG, BESS Solar, Wind, Li-ion, Flywheel, Load types (Residential, Commercial, Industrial)	Minimize cost, Maximize RE utilization, Stability	Energy Management	O	Yes (MILP, PSO, DQN, DDPG)	No	Yes (Island in South China, Renewable intermittency, No HILP)	Yes (MC sampling, LightTS prediction model)
[105]	Multi Agent (MA)RL	Load types (Residential, Commercial, Industrial)	Improve reliability, cost savings, renewable utilization, scalability	Energy Management	O	Yes (vs. Rule-Based Control)	No	No (Simulation-based; realistic config but no specific HILP event)	Yes (Demand/load variation, agent coordination, scalability testing)
[106]	PPO	Solar, Wind, EVs (V2G)	Increase RES utilization, Improve grid frequency and voltage stability	Energy Management (V2G Scheduling)	O	Yes (vs. deterministic and rule-based scheduling)	No	No (Simulation-based; no real HILP event source or case)	No (Sensitivity to learning rate, discount factor; grid/load variation modeled)
[107]	Distributed DRL	Traditional DGs, Wind Farms	Voltage stability, economic cost, and system recovery under FDI attacks	Cybersecurity (FDI attack defense)	O	No (Compared with other DRL-based approaches, not classical methods)	No	No (Hypothetical FDI scenarios on 6-microgrid test system)	Yes (FDI modeled probabilistically with RL-based online estimation)
[108]	TD3 + SVM	Diesel Generator, EVs	Frequency regulation under FDI, maintaining stability	Cybersecurity (FDI detection and mitigation)	O	Yes (vs PID and optimal controller)	No	No (simulated system only; FDI inspired by Ukraine attack, no impact modeled)	Yes (FDI attack strength, wind power uncertainty from Sweden)
[109]	RS, TL, DDPG	PV, WT, DG, BESS	Operational cost minimization, load shedding avoidance, SOC/power balance constraint satisfaction	Energy Management	O	Yes (Compared with DQN, Double Deep Q-Network (DDQN), standard DDPG, and MIP benchmark)	No	Yes (Gasa Island, Korea – general renewable/load uncertainty, no specific HILP source)	Yes (RES and load stochasticity from real 2020 data, Ninja app)
[110]	BDDPG (Bayesian DDPG)	PV, Wind, DGs, ESS, Main Grid	Minimize cost, avoid load loss, account for degradation under HILP	Energy Management (Resilience under HILP)	O	Yes (Compared with SP and naïve DDPG)	No	No (Simulated HILP: DER loss, grid outage, but no real-world case)	Yes (Bayesian modeling of epistemic and aleatoric uncertainties)

Table 2.6: *Continue . . .*

Ref.	Algorithm	Power Generation	Objective(s) Reward(s)	Resiliency Domain	Operation (O) Planning (P)	Comparison with Classical Methods	Explainability	Real-world Case Study	Uncertainty Quantification
[111]	DRLN	PV, Wind, DG, BESS	Minimize voltage deviation, optimize power flow, reduce DRES overload	Energy Management + Fault Resilience	O	Yes (Compared with fuzzy droop controller)	No	Yes (HIL) implementation; HILP: communication link failure, more reliability)	Yes (Simulation under source outage, bus failure, and real-time fluctuations)
[112]	Multi Agent Proximal Policy Optimisation (MAPPO)	PV, Diesel, Gas-fired Generators	Load restoration, Minimizing travel and shedding costs	RC Dispatch Coordination (Power-Gas-Transport)	O	Yes (Compared with MARL baselines like MADDPG, MASAC, and GA and MPC) Yes (Compared with rigid design baselines (Genetic Algorithm (GA)))	No	No (Synthetic testbeds: coupled 6-bus and 33-bus networks)	Yes (Traffic volume, PV generation, loads modeled as stochastic processes)
[113]	ACER	PV, Wind, BESS, Electric Heating	Expected Net Present Cost (ENPC), CO ₂ emissions, Total Unmet Load	Planning under Climate-Induced Mobility	P	Yes (Compared with rigid design baselines (Genetic Algorithm (GA)))	No	Yes (Mongolia, Nomadic Migration; no specific HILP impact analysis)	Yes (Geometric Brownian Motion for RES, load, grid, and spatial dispersion)
[114]	A3C++ PPO	Wind, Grid, ESS	Revenue maximization, Cost reduction, Noise-resilient forecasting	Energy Management + Forecasting	O	No (Compared A3C++ with PPO)	No	No (Simulated residential microgrid; no real HILP context)	Yes (Wind data noise, uncertainty in load and generation modeled)
[115]	DQL, DDQL	PV, Wind, BESS, Gas Generator, Grid	Energy cost minimization, Production throughput, Battery health	Energy Management in Manufacturing Systems	O	Yes (DQL vs. DDQL)	No	No (simulated case study; normal operations without HILP event)	Yes (Stochastic solar/wind and machine failure modeled in transitions)
[116]	FedSAC	GFM Inverters (GridLAB-D Simulation)	Voltage recovery, Attack mitigation	Cyber Resilience	O	No (Compared with Decentralized SAC)	No	No (IEEE 123-bus test system; cyber attack simulated but not linked to real HILP source)	Yes (Adversarial signal pool, privacy-preserving federated data handling)
[117]	MC RL + DNN	Grid, DGs, ESS (generic)	Cost minimization, PAR reduction, DSO profit	Energy Management (Multi Micro Grid (MMG) Coordination)	O	Yes (Compared with DQN, DDQP, and model-based solver)	No	No (Simulated MMG; no explicit HILP or real event)	Yes (Load/generation profile disturbance, model-free RL)
[118]	SOC-Droop + Fuzzy Logic (NO DRL)	PV, BESS	SOC balancing, Transient voltage response improvement	Voltage and Power Sharing Control	O	Yes (PI controller)	No	No (Hardware-in-the-loop on testbed, no real-world HILP source)	No (only deterministic performance analysis)
[119]	DQN	DGs (Voltage Source Modeled)	Maximize cumulative restored critical load, penalize violations	Microgrid Formation for Load Restoration	O	Yes (Compared with MILP from [2])	No	No (IEEE 13, 37, and 123-node feeders; source of HILP is abstract main grid failure, unspecified real event)	Yes (Random load fluctuation, topology violation considered in training)
[120]	SOSA (NO DRL)	PV, WT, ESSs	Minimize load shedding, investment/operation cost, Maximize DER profit	Infrastructure planning	P	No	No	No (Standard 33-, 54-, 119-, 136-node test networks), HILP source: earthquake simulated via fragility curves	Yes (Attenuation functions, fragility curves, SSGM-based scenarios)
[121]	PPO (MARL)	PV, Diesel, BESS	Load balancing, Profile-following, Storage compliance, Reliability and Resilience Score	Energy Management	O	Yes (LP benchmark, but not classical comparison)	No	Yes (Austria PV and industrial load, no real-world HILP event)	Yes (Synthetic profiles, stressor-based analysis)
[122]	TD3	EV-integrated DERs	Voltage stabilization, reactive power control	No (OPF)	O	No (Compared with DDQP and DQN)	No	No (Modified IEEE 30-bus system, synthetic loads)	No

Table 2.6: *Continue . . .*

Ref.	Algorithm	Power Generation	Objective(s) Reward(s)	Resiliency Domain	Operation (O) Planning (P)	Comparison with Classical Methods	Explainability	Real-world Case Study	Uncertainty Quantification
[123]	CM-DDQN (CNN-based DQN)	Multiple DERs (unspecified types)	Feasible topology, load support, voltage/branch constraints, switching cost	MMG Formation	O	Yes (compared to DDQN and conventional MMG Formation [4,10])	No	No (IEEE 7-bus, IEEE 123-bus; Fault assumed, source not specified)	Yes (DG uncertainty modeled by Gaussian distribution with forecast error)
[124]	DDPGfD	PV, DGs, ES	Minimize load curtailment, DG cost, and battery degradation	Service Restoration	O	Yes (Compared with standard DDPG and model-based MILP)	No	No (IEEE 123-node test system, assumed outage)	Yes (Multiple outage scenarios, PV/load uncertainty)
[125]	Deep LSTM-based RL	Solar PV, Battery Storage	Minimize energy purchase and optimize battery usage	No (Energy Management)	O	No (ILP and Rule-Based methods used as environment)	No	Yes (IIoT microgrid on real buildings)	Yes (Time series prediction uncertainties from real data)
[126]	MA-SAC	Hybrid AC/DC	Maximize restored load	Network Reconfiguration	O	Yes (Genetic Algorithm, MILP)	No	No (Modified IEEE 37 and 123-bus systems, assumed extreme outage events)	Yes (Load and contingency scenarios)
[127]	PPO	DER, switches in microgrid	Minimize network loss, maximize load restoration	Network Reconfiguration	O	No (BPSO)	No	No (Modified IEEE 34-bus microgrid, assumed line faults)	Yes (Random load conditions, line outages)
[128]	DDQN	PV, Diesel Genset, Batteries	Minimize operational cost under rare events	Energy Management	O	Yes (Rule-based, DP, RL-DT)	No	Yes (Ecole Polytechnique, France)	Yes (Significant Rare Events Handling with MemC and CER)
[129]	Value Iteration, Q-Learning	DG, PV, ESS, Controllable Loads	Maximize load serviceability, minimize operational cost	Energy Management	O	No	No	No (Benchmark MG3)	Yes (Extreme event probability incorporated)
[130]	Curriculum Learning PPO	Micro-turbine, Battery, Wind, PV	Maximize critical load restoration, minimize voltage violation	Energy Management	O	No	No	No (IEEE 13-bus, Renewable Uncertainty)	Yes (Renewable generation uncertainty)
[131]	Learning Automata (Multi-agent)	PV, Battery Storage	Maximize communication quality and energy availability	Energy Management	O	No	No	No	Yes (Stochastic game)
[132]	Q-learning	Micro-turbines, Fuel Cells, PV, Batteries	Voltage angle and magnitude	Voltage stability	O	No	No	No	No (Black Start Scenario)
[133]	DDPG, SAC	PV, WT, DG, BESS	Energy cost minimization, Reliability maximization in SIPS	Energy Management, System Integrity Protection	O	No	No	Yes (Gasa Island Microgrid, SIPS Alarms)	Yes (Stochastic load, renewables, market price)
[134]	Evolution Strategy RL (ES-RL)	PV, Wind, Micro-turbine, Battery	Maximize prioritized load restoration, penalize load shedding	Energy Management	O	Yes (Model Predictive Control)	No	Yes (Single bus system, Real PV and Wind data)	Yes (Imperfect renewable forecasts)
[135]	MLP-based RL	Diesel Generators, PV	Primary frequency response improvement	Frequency Stability	O	Yes (PI Controller)	No	No	No
[136]	Q-learning (DMQ)	N/A (Communication network resource allocation)	Latency minimization, Fairness	Communication Network (Latency)	O	Yes (PF and DIRA optimization-based)	No	No (Simulated LTE small cell network, no specific HILP event source)	Yes (Dynamic network state)
[137]	DQN	PV, Gas Turbine, BESS	Maximizing weighted load restoration benefit	Load Restoration	O	Yes (Compared with PSO, MILP, enumeration)	No	No (Modified IEEE 13-bus system), Assumed extreme hot weather	Yes (DRO with Wasserstein metric for PV/load uncertainty)

A complementary design path is *architecture*: multi-agent and hierarchical approaches distribute decision making and can scale in complex networks. Examples include hierarchical MAPPO for repair-crew dispatch across coupled energy networks [112], federated reinforcement learning to improve cyber resilience while preserving data privacy [116], and EMS based on market mechanisms or contracts, enabling decentralised negotiation among DERs [97]. Hardware In the Loop (HIL) and real-time simulation platforms are also appearing, e.g., a real-time digital simulator to accelerate training for frequency dynamics [103] and HIL implementations for voltage control [111].

DRL has also been adopted for long-term planning, where the objective is to select capacities and technologies that reduce expected cost and outage risk. [72], [99] use deep Q-learning variants to co-optimize cost, environmental metrics, and resilience (via loss of load and critical-customer metrics), while [113] addresses mobile microgrid design under climate-induced mobility. Although these works make resilience an explicit criterion, they mostly rely on synthetic or historical reliability statistics, rather than replaying measured or historical HILP event trajectories.

Beyond EMS, DRL supports network reconfiguration and OPF for rapid recovery and stability, including MA-SAC for AC/DC hybrid reconfiguration [126], PPO for fault-aware reconfiguration [127], and TD3/DDPG for continuous voltage and reactive power control [111], [122]. A number of studies benchmark against mathematical programming or metaheuristics, which helps ground DRL performance in recognised baselines [119], [126], [137].

Cyber-physical threats are modelled with DRL-based detection and mitigation policies, including rootkit and False Data Injection (FDI) scenarios [95], [100], [107], [108]. These studies show that reward design and partial observability handling are key to stabilising responses under adversarial signals.

When forecasting errors dominate, DRL is paired with signal models and learned predictors [104], [114]; digital twin settings fuse live measurements and simulators to host learning and policy in the loop evaluation, as in [102]. These trends point towards deployable frameworks that are closer to operations.

What is missing and how this thesis responds. Three persisting gaps motivate the design choices and directly justify the methodological design in this thesis (see Table 2.6).

1. **Battery life is often simplified.** Most EMS papers enforce basic SOC and power limits but do not model ageing explicitly, which blurs trade-offs between short-term resilience (e.g., deep cycling during outages) and long-term asset health. Exceptions exist: some include simple degradation substitutions or penalties [99], [115], [124], yet detailed DOD effects is rarely embedded in the learning loop. *This thesis* integrates a battery-life objective alongside resilience so the agent internalises the cost of aggressive cycling during HILP.
2. **Real HILP data and case studies are scarce.** Many studies rely on standard test feeders, assumed faults, or noise-based scenario generation [93], [110], [119], [137], with notable but limited real-world cases (e.g., a campus microgrid [128], Gasa Island operations [109], or national data feeds [102]). *This thesis* utilises real geographical and load data from Ongole, India, and simulates the impact of Cyclone Laila on supply and demand [49]. Renewable profiles are further augmented with Perlin-noise perturbations to expand the training and testing distributions while preserving temporal structure.
3. **Baselines and explainability are limited.** While some works compare against optimisation baselines such as MPC, MILP, Genetic Algorithm (GA), or rule-based control [93], [112], [119], [126], [134], [137], many benchmark only against other DRL variants [99], [100], [101], [107], [114], [122], [129], [130], [131], [132], [133]. Explainabil-

ity is also rarely reported in EMS settings, despite its importance for operator trust [138]. *This thesis* therefore introduces a directly comparable MPC baseline built under the same constraints and forecasts to enable fair per-horizon evaluation of battery life, load served, and computational effort, and supplements DRL results with post-hoc explanations (LIME and SHAP) to expose decision drivers transparently to operators.

Among available actor–critic algorithms, PPO [139] is adopted due to its proven training stability, robustness to hyperparameter tuning, and suitability for continuous action spaces. As an improvement over its predecessor Trust Region Policy Optimisation (TRPO), PPO strikes a balance between policy stability and computational efficiency, making it one of the most widely validated algorithms for complex control tasks. By analysing the reviewed literature (Table 2.6), a total of 50 algorithm mentions were recorded, including 14 using the Q-learning family (Q-learning, DQN, DDQN, DDQL), 11 employing PPO, 8 based on DDPG or TD3, 6 using SAC, and the remainder adopting other variants or hybrids, as shown in Figure 2.15. While the Q-learning family dominates due to its simplicity and effectiveness in discrete action spaces, PPO is more suitable for continuous control problems such as power dispatch and battery management in microgrids, justifying its selection in this thesis.

Problem Statements Linked to Research Gaps

- ✓ **Gap 3:** Protects battery life: incorporate DOD-aware ageing rather than wide SOC bands or simple penalties.
- ✓ **Gap 4:** Addresses HILP conditions: remain effective under rare, severe events (e.g., cyclones) using location-specific, real disturbance data when available.
- ✓ **Gap 5:** Ensures transparency: provides justifiable decisions for stakeholders, supporting governance and trust in critical infrastructure.

Accordingly, this thesis employs PPO for EMS centred resilience, incorporating: ex-

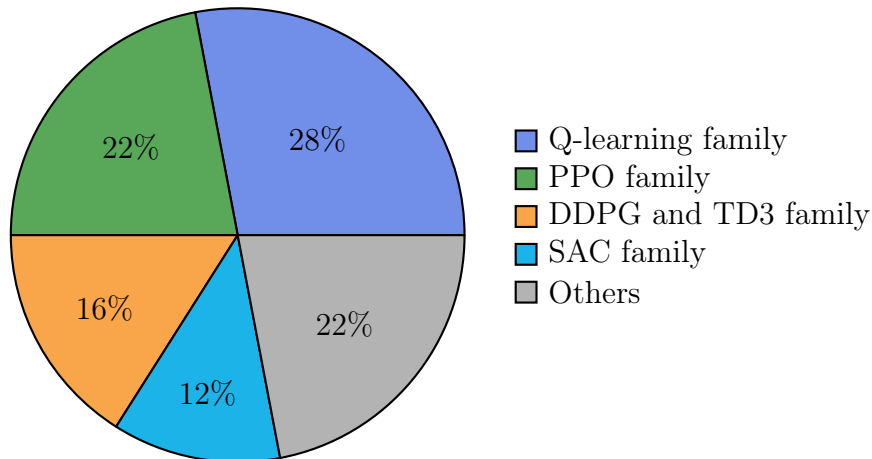


Figure 2.15: Distribution of algorithms used in DRL-based microgrid resilience studies. Percentages reflect algorithm mentions across the reviewed literature (Table 2.6). (Unit: individual algorithm mentions; if a paper used multiple algorithms, each was counted once).

plicit battery-life terms in the reward; location-specific modelling with a real HILP event (Cyclone Laila); structured uncertainty via scenario augmentation; a comparable MPC baseline; and post-hoc explanations using LIME and SHAP [140], [141].

2.4 Explainable AI with a focus on XDRL

Section Summary

Building on the identified gap in transparent DRL decisions for microgrid EMS (Section 2.3.2), this section defines explainability in DRL, distinguishes it from interpretability, reviews the principal interpretable, intrinsic, and post-hoc techniques, and identifies post-hoc attribution as the most appropriate choice for this thesis.

Artificial intelligence is now embedded in daily life and critical infrastructure. As learning systems have grown in scale and autonomy, transparency becomes a practical necessity. Two milestones frame this shift. First, Defense Advanced Research Projects Agency (DARPA)’s eXplainable Artificial Intelligence (XAI) programme promoted models whose

reasoning humans can interpret and trust, catalysing a wave of research in methods that “open up” black boxes [142]. Second, regulation has advanced rapidly: the EU’s General Data Protection Regulation (GDPR) established strong data rights and a “right to explanation” in automated decision settings, while the EU AI Act targets transparency, auditability, and human oversight in high-risk systems [143], [144]. Together with broader ethical principles of fairness, accountability, and safety, make explainability more than an academic nicety for systems like microgrids; it is a practical requirement [145], [146].

It is useful to distinguish *interpretability* from *explainability*. Interpretability refers to models that are inherently understandable by design (e.g., decision trees or rule-based controllers). Explainability, in contrast, uses tools to clarify a complex model’s reasoning after the fact (e.g., explaining a deep neural policy via feature attributions) [147]. In practice, both are useful. Interpretability supports direct reasoning, while explainability allows high-performing yet opaque models to remain usable in sensitive applications.

Explainability has already become a fundamental component of deployment. In healthcare it supports clinical validation and error analysis for diagnostic support tools [148]. In robotics it improves human–robot teaming by clarifying intent and failure modes [149]. In natural language processing and computer vision it helps unpack long-horizon strategies and attention in agents [150], [151]. In finance and business it aids compliance and stakeholder trust [152]. In transport it addresses passenger trust and post-incident analysis for autonomous vehicles [153]. For energy and smart grids, explainability is directly tied to operator confidence and regulation, especially for emergency control and demand-side management [87], [154], [155], [156].

Explainable DRL techniques are commonly grouped into three categories [147]. The first comprises *Interpretable Agents (IA)*, which are transparent by construction and expose their decision logic directly. The second category covers *Intrinsic Explainability (IE)*, where

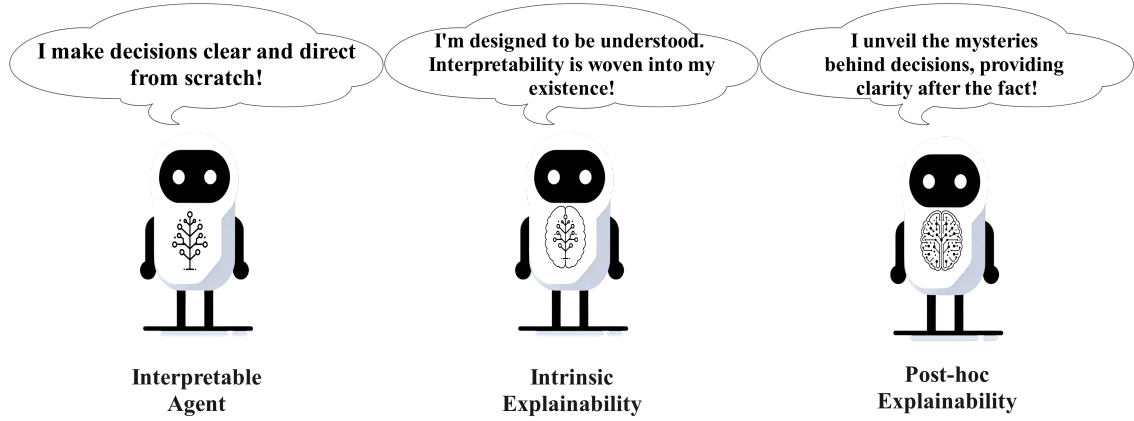


Figure 2.16: XDDL agents speak: IA on simplicity, IE on design integration, PHE on insightful analysis.

explanation mechanisms are embedded within the model architecture prior to training. The third group consists of *Post-Hoc Explainability (PHE)* methods, which aim to explain the behaviour of a trained black-box model without modifying its internal structure. Figure 2.16 (“XDDL agents speak”) visualises these three perspectives.

Interpretable Agents. Here the policy is itself a human-readable object, which eases audit and formal checks, though usually in lower-dimensional settings. Examples include:

- **Interpretable Fuzzy Reinforcement Learning (IFRL)** uses fuzzy rule bases within an actor-critic framework. The IF-THEN rules make policy logic legible (e.g., “if valley and low velocity then full throttle”), while still handling continuous actions [157].
- **Deep Symbolic Policy (DSP)** searches the space of mathematical expressions for control policies, yielding concise symbolic forms rather than neural nets. It also distils NN policies into symbolic surrogates for multi-dimensional actions [158].
- **Iterative Bounding Markov Decision Process (IBMDP)** wrap the MDP to learn neural policies that can be *extracted* as decision trees, preserving verifiability and

feature relevance while training with modern function approximators [159].

Intrinsic Explainability. The agent or its world model is modified so explanations emerge during the forward pass or from compact state abstractions.

- **Saliency/attention** for sequential vision: the DRIVE model fuses bottom-up and top-down attention, learning to anticipate hazards in driving while producing attention maps that show “where the agent looked” when predicting risk [160].
- **Self-Explaining Neural Networks (C-SENN).** By combining supervised contrastive learning with Barlow Twins regularisation, C-SENN learns distinct, object-grounded *concepts* that are tied to decisions and thus more intelligible in complex scenes [161].
- **Generative explanations.** Textual rationales can be produced from the model’s learned concepts to describe the “why” behind an action (e.g., autonomous driving narratives) [162].
- **Representative explanations.** Using the agent’s own structured representation (e.g., decision-tree policy or simplified state space) as the explanation itself improves human alignment in collaborative tasks.

Post-hoc Explainability. These methods sit outside the trained model and explain decisions locally or globally without altering the policy. This is crucial when performance must remain intact.

- **SHAP** unifies feature-attribution via Shapley values, offering theoretically grounded, additive explanations at the instance level and summaries across populations [163].

- **LIME** fits an interpretable surrogate (e.g., sparse linear model) in the neighbourhood of a prediction to reveal which inputs most affected the outcome [164].
- **Gradient-Weighted Class Activation Mapping (Grad-CAM)** highlights the image regions that drive a CNN’s score, useful when visual observations feed an RL policy [165].
- **Anchors** provide high-precision if-then rules that “lock” a prediction when certain conditions are met, turning complex behaviour into compact, testable statements [166].

Why post-hoc methods are adopted here. Critical infrastructures cannot sacrifice control performance simply to gain transparency. As shown in Figure 2.17, interpretable and intrinsic approaches often trade accuracy for transparency, whereas post-hoc methods maintain model fidelity while offering clear diagnostic insight. In microgrid EMS and resilience control, where safety margins, restoration speed, and cost are critical, a high-performance DRL policy (PPO) is desirable, with post-hoc explanations (SHAP and LIME) applied to reveal the drivers of actions without compromising performance [138], [147], [163], [164]. This preserves accuracy and fast inference while meeting operator and regulatory needs for transparent reasoning and post-event analysis. This also aligns naturally with real-world operational workflows: explanations can be computed offline over historical episodes, embedded into dashboards for situational awareness, or attached to alerts during HILP scenarios.

Problem Statements Linked to Research Gaps

- ✓ **Gap 5:** Ensures transparency: provides justifiable decisions for stakeholders, supporting governance and trust in critical infrastructure.

Summary for microgrid context. Anchored in the DARPA, GDPR, and EU AI Act initiatives for trustworthy AI [142], [143], [144], the thesis integrates first, a high-performing

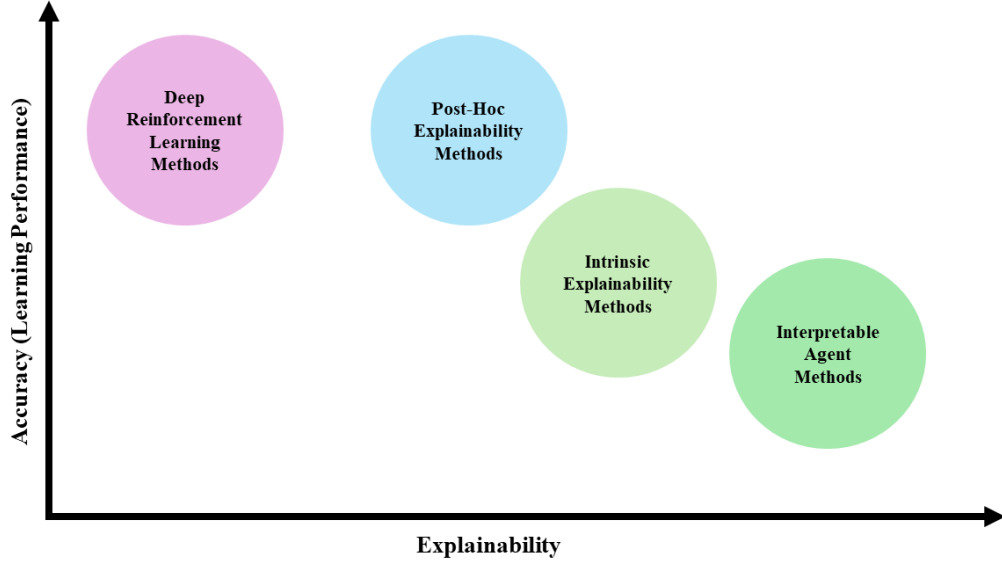


Figure 2.17: Accuracy versus explainability across XDRL families. PPO as high-accuracy control is applied and post-hoc explanations is adopted to meet transparency needs without degrading performance.

PPO-based controller for resilient EMS; second, an MPC benchmark for interpretability and validation; and third, post-hoc XDRL using SHAP and LIME for operator-level transparency. While IA and IE methods remain valuable for compact or attention-driven applications, post-hoc approaches offer the most pragmatic pathway to trustworthy and high-performing microgrid control under uncertainty [87], [146], [156].

Having reviewed resilience metrics, control strategies, and explainability frameworks, the next step is to summarise the key research gaps and presents the contributions of this thesis.

2.5 Research Gaps and Thesis Contributions

The review across resilience metrics, MPC, DRL, and explainability reveals recurring limitations that directly motivate the thesis objectives. These limitations are synthesised below as

six interrelated research gaps, each paired with the corresponding contribution of this work.

Gap 1: Lack of clarity in resilience concepts and metrics. For new researchers, resilience is often difficult to distinguish from reliability, and the literature presents a wide range of metrics with no clear guidance on adoption.

Contribution: A comprehensive review paper [2] was published to synthesise resilience metrics, supported by a holistic resilience curve that captures the full sequence of withstanding, absorption, and recovery phases. This curve was validated against Hurricane Harvey outage data, demonstrating improved accuracy compared with triangular and trapezoidal representations. Furthermore, an Expected Resilience Curve and Index were proposed, enabled by the MPC-based framework.

Gap 2: Resilience treated implicitly rather than explicitly in operation. In many operational studies, resilience is often considered a by-product of the proposed frameworks, rather than an explicit control objective.

Contribution: To maintain the MILP property of MPC, a novel linearisation was introduced to embed resilience directly in the optimisation objective. For DRL, a priority-weighted Resilience Index was designed as part of the reward function, making it an explicit driver of policy learning.

Gap 3: Limited consideration of battery degradation in operation. Battery longevity is strongly influenced by depth of discharge, yet most works ignore this or adopt oversimplified SOC bounds.

Contribution: In the MPC framework, a linearised BLC–DOD formulation was developed that preserves MILP structure while reflecting degradation more realistically. In the DRL framework, the exact nonlinear BLC–DOD curve was incorporated directly, allowing the agent to optimise expected battery years. A curriculum learning strategy was also introduced to resolve the conflict between resilience and degradation objectives, ensuring that

resilience is satisfied first before improving battery lifetime.

Gap 4: Reliance on generic or synthetic case studies. Most operational studies are based on IEEE test feeders or synthetic data, which limits their relevance for real deployments.

Contribution: This thesis focused on a cyclone-prone region in Andhra Pradesh, India. Historical cyclone data were used to reproduce renewable generation patterns during the event, while HOMER Pro was employed to design an economically sized microgrid based on real load profiles and location. This provided a solid case study for validating the proposed controllers under realistic extreme conditions.

Gap 5: Absence of explainability in DRL for microgrids. Although review papers identify explainability as a critical barrier for regulatory acceptance and trust, DRL applications to microgrid resilience have largely neglected it.

Contribution: This work pioneers the application of post-hoc explainability methods (SHAP and LIME) to DRL controllers in this domain, clarifying the controller’s rationale behind charging, discharging, idle decisions, and power allocation among priority loads. This contribution advances transparency and stakeholder confidence in AI-driven control of critical infrastructure.

Gap 6: Limited benchmarking and robustness evaluation. Existing works often compare DRL algorithms only with each other, neglecting rigorous benchmarking against classical methods or robustness testing under uncertainty.

Contribution: MPC was used as a baseline to validate the DRL framework, with a systematic comparison of solution quality, resilience, and inference time. Robustness was further assessed by introducing Perlin-noise perturbations and MC evaluation, demonstrating that the proposed DRL approach generalises across uncertain operating conditions.

Overall, this research addresses six critical gaps. **Gaps 1–3** focus on conceptual clar-

ity and methodological formulation, covering resilience metrics, explicit resilience modelling, and battery degradation awareness in MPC and DRL frameworks. **Gaps 4–6** address validation and practical deployment, including location-specific case studies, benchmarking against classical methods, robustness under uncertainty, and transparency through explainable DRL.

Chapter 3

Research Methodology and Case Study

3.1 Research Philosophy and Approach

This thesis adopts a **quantitative, simulation-based** philosophy. Microgrid behaviour is represented through mathematical models and evaluated via replicable computational experiments. The design is **comparative** and **question-led**. An advanced learning-based controller is benchmarked against an optimisation baseline, with each research question driving a focused methodological stream.

MPC, implemented as a MILP, provides an interpretable baseline whose structure explicitly encodes priorities, resilience terms, and a linearised BLC–DOD relation. DRL, particularly the PPO algorithm, serves as a model-free counterpart designed to operate under uncertainty in renewables, demand, and HILP events. Post hoc explanation tools (e.g., SHAP, LIME) are used to analyse learned policies.

Overall, the work is a **simulation-based evaluation** that combines optimisation, learning, and explainability under controlled test scenarios including real extreme-weather data (Cyclone Laila) and principled perturbations. This framing underpins the methodological streams that follow and aligns directly with RQ1–RQ5 (Section 1.4).

3.2 Methodological Approach

The methodology is organised into five streams, each beginning with a guiding question, followed by a sequence of activities: literature review, data collection and preprocessing, model development, evaluation, and documentation (Figure 3.1). Together, these streams converge to deliver an explainable, optimal or near-optimal framework for microgrid energy management.

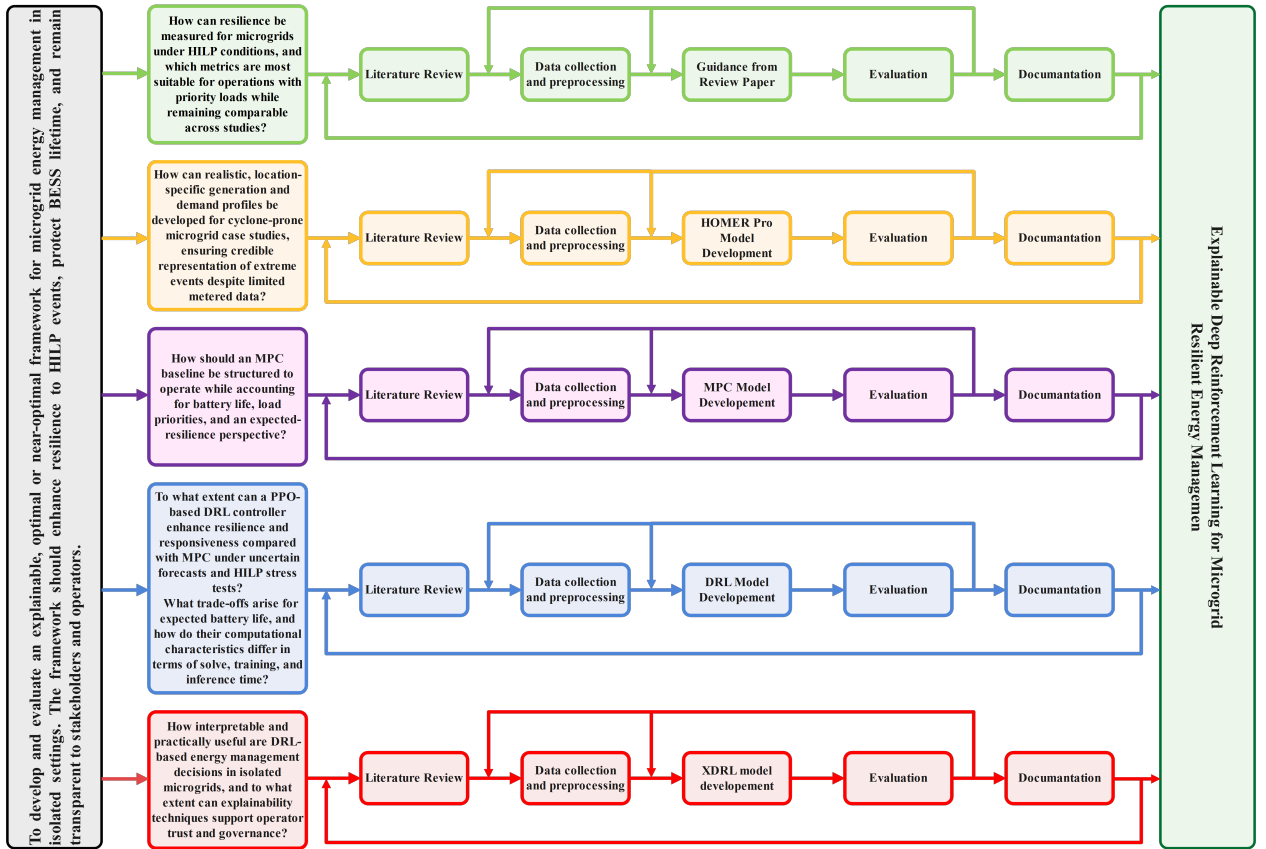


Figure 3.1: Methodology structured by research questions (RQ1–RQ5). Each stream follows a waterfall process of literature review, data preparation, model development, evaluation, and documentation, contributing to the overall thesis aim.

The five streams correspond to the thesis research questions:

1. Resilience Metrics (RQ1):

Question: How can resilience be measured for microgrids under HILP conditions, and which metrics are most suitable for operations with priority loads while remaining comparable across studies?

This stream begins with a review of resilience concepts, highlighting the lack of clarity and consistency in existing definitions. Data analysis and synthesis lead to the construction of a holistic resilience curve and associated indices, validated against real outage data from Hurricane Harvey. The resulting guidance provides a foundation for embedding resilience explicitly into subsequent operational models. Documentation of this work concluded in the published review article of the thesis [2].

2. Location-Specific Case Development (RQ2):

Question: How can realistic, location-specific generation and demand profiles be developed for cyclone-prone microgrid case studies, ensuring credible representation of extreme events despite limited metered data?

This stream develops the case study site in Andhra Pradesh, India. The literature review strengthens the identified gap that in most operational and resilience-oriented studies, the sizing and location of the microgrid as well as the source of the HILP event are often neglected. Data are obtained from NASA/Prediction of Worldwide Energy Resources (POWER) and PVGIS for solar and wind resources, and HOMER Pro is used for techno-economic sizing of the PV–WT–BESS configuration. Cyclone Laila (2010) is employed as a validation scenario, ensuring that operational models are tested under real extreme-weather conditions. Documentation of this work, particularly the planning and sizing aspects, formed part of the first journal paper associated with this thesis [49].

3. MPC Baseline (RQ3):

Question: How should an MPC baseline be structured to operate while accounting for

battery life, load priorities, and an expected-resilience perspective?

This stream formulates a Python Optimization Modeling Objects (PYOMO)-based MPC model, building on insights from the literature. Data preprocessing aligns HOMER-derived sizes and load priorities with operational constraints. Model development embeds a piecewise linearised BLC–DOD relation and a resilience index into the objective, while preserving MILP structure. Evaluation examines performance under deterministic and cyclone scenarios, providing a transparent, interpretable baseline. Documentation of this work, including the MPC design and linearisation strategy, was reported in the first journal and conference paper linked to this thesis [49], [167].

4. **DRL Controller (RQ4):**

Question: To what extent can a PPO-based DRL controller enhance resilience and responsiveness compared with MPC under uncertain forecasts and HILP stress tests? What trade-offs arise for expected battery life, and how do computational characteristics differ in terms of solve, training, and inference time?

The DRL stream leverages Gymnasium and Stable Baselines3 (SB3) to develop a PPO agent. Data preprocessing adapts renewable and load profiles, including synthetic perturbations, for training and testing. The agent’s reward function balances resilience and battery life, while curriculum learning ensures resilience objectives are prioritised before battery longevity. Evaluation compares performance against MPC across multiple dimensions: resilience indices, expected battery years, and computational efficiency. Documentation of this work is presented in the second conference paper [168] and further extended in a second journal paper, currently under review.

5. **Explainability (RQ5):**

Question: How interpretable and practically useful are DRL-based energy management decisions in isolated microgrids, and to what extent can explainability techniques

support operator trust and governance?

This stream focuses on transparency and stakeholder confidence. Literature review identifies explainability gaps in power system AI applications. Data and state representations are adapted for explanation algorithms. The XDRL model integrates post hoc methods with PPO policies. These explanations are used to evaluate why specific decisions were made, supporting operator trust and governance in critical infrastructure contexts. Documentation of this stream was first developed as part of a book chapter drafted during the PhD, then consolidated in the second conference paper [168], and is being further extended into a journal article currently under review.

In summary, the methodological framework is explicitly tied to the research questions, ensuring that each identified gap is addressed systematically. The streams progress from clarifying resilience concepts, to developing a realistic case, to constructing MPC and DRL models, and finally to ensuring explainability. The following section translates this framework into a simulation-based modelling strategy, detailing the planning, operational, and evaluation steps through which the proposed methods are implemented and validated.

3.3 Simulation-Based Modelling Strategy

The methodology adopts a simulation-based modelling strategy that integrates microgrid planning, operational optimisation, and controller benchmarking under realistic disturbance scenarios. The framework is structured in two stages: (i) planning and sizing using HOMER Pro, and (ii) resilient operation and benchmarking using PYOMO-based MPC and SB3-based PPO DRL. This section describes the two-stage framework and case study site.

3.3.1 System Overview and Microgrid Architecture

The proposed microgrid consists of PV modules, WT, a bidirectional converter, and a lithium-ion BESS. The modelling and evaluation framework follows a two-stage workflow that links long-term system design with short-term operational control.

In the *planning stage*, the microgrid is sized using HOMER Pro through techno-economic optimisation. Historical irradiance and wind-speed data, together with representative daily and seasonal load profiles, are used to determine the optimal capacities of PV, WT, BESS, and the power converter under a load-following dispatch strategy. This stage establishes the nominal ratings and configuration of all components and provides a realistic basis for subsequent operational analysis.

In the *operation stage*, the planned system is re-modelled for control and decision-making under uncertainty. First, an MPC formulation based on MILP is implemented within a sliding-window framework. Battery ageing is explicitly represented through a linearised BLC–DOD relationship embedded in the objective function, while resilience is modelled by penalising priority-weighted unserved load. In parallel, a custom DRL environment in Gymnasium is developed where the same system states (SOC, net energy balance, available renewable generation, and loads demand) and actions (charge, discharge, and loads' weights) are used for PPO training. Both MPC and DRL controllers are evaluated under cyclone conditions and additional scenarios.

This two-stage framework ensures that system design decisions are consistent with operational constraints and resilience objectives, addressing gaps in the operational research literature where the sizing of the microgrid and the use of real case studies are often neglected.

The comparative design contrasts two complementary streamline:

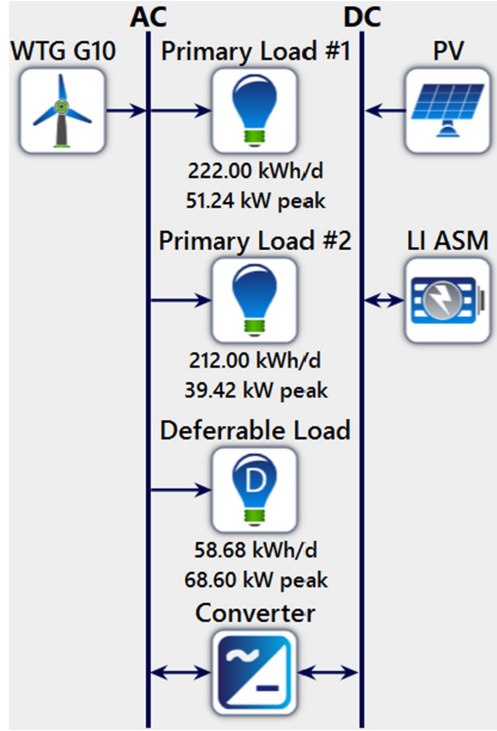
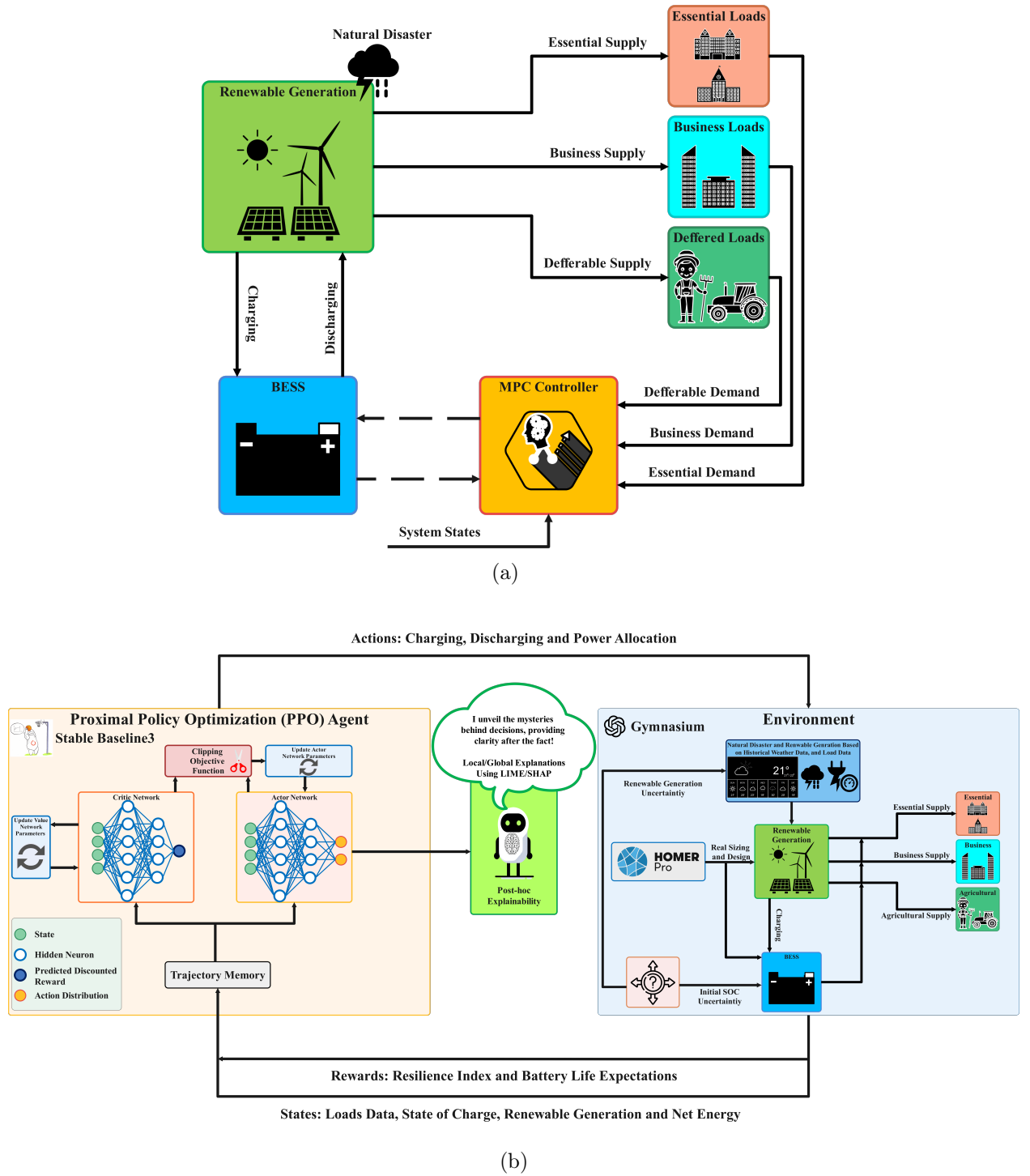


Figure 3.2: Microgrid planning stage model in HOMER Pro.

- **MPC:** A model-based approach relying on mathematical formulation of system constraints, accurate forecasts, and linearised degradation models. MPC yields near-optimal solutions but is computationally intensive and forecast-dependent.
- **DRL:** A model-free approach (PPO) that learns policies directly through interaction with the environment. DRL shifts computation to training and enables rapid inference at run time. However, policy transparency is reduced and strict optimality is not guaranteed. To overcome DRL opacity, post hoc explainability tools are applied to interpret policy decisions without compromising performance.

Figure 3.2 shows the planning stage of the microgrid architecture as obtained in HOMER Pro. The operational stage is presented in Figure 3.3, where Figure 3.3a depicts the MPC-based construction and Figure 3.3b illustrates the XDRL framework.



3.3.2 Case Study Site: Kothapatnam, India

The case study focuses on a rural coastal village in Andhra Pradesh, India (near Ongole/Kothapatnam), which has experienced repeated cyclone events over the past decades. The geographical coordinates of this location are $15^{\circ}28'36.3''$ *N* latitude and $80^{\circ}11'54.0''$ *E* longitude. The location is characterised by intermittent renewable resources, moderate load demands from domestic, agricultural, community, and small-scale industrial activities, and limited access to a reliable grid.

The load composition and magnitude are designed to be representative of a typical rural settlement in India and are consistent with profiles commonly adopted in microgrid planning studies for similar contexts. The categorisation and scaling of loads are informed by rural electrification literature, reported demand patterns for domestic, commercial, and agricultural users, and standard community-level assumptions used in HOMER Pro-based studies [169], [170]. Three categories of demand are considered, reflecting priority levels for resilient operation:

1. **Essential loads:** domestic appliances (lamps, fans, radios), schools, and health centres, with an average daily demand of 222 kWh with a peak of 51.2 kW, and a load factor of 0.181.
2. **Business loads:** community centres, local shops, and street lighting, with an average daily demand of 212 kWh with a peak of 39.4 kW, and a load factor of 0.224.
3. **Deferred loads:** mainly agricultural irrigation pumping, with an average daily demand of 58.6 kWh and a peak of 68.6 kW, having a storage capacity of 30 kW.

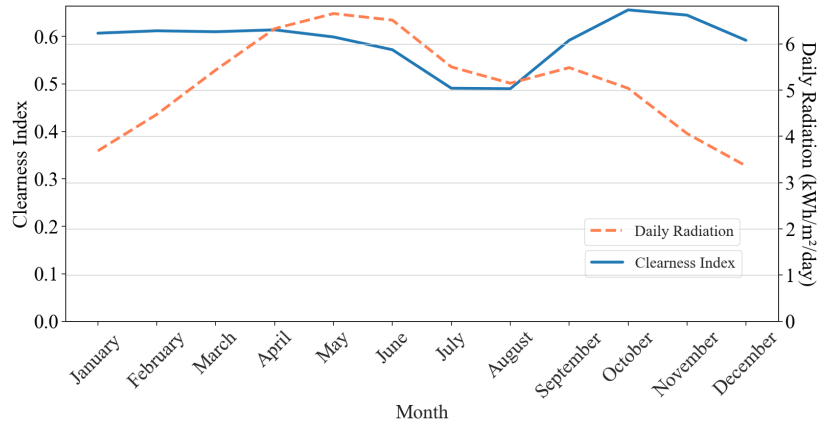
In the remainder of this thesis, **essential loads** are sometimes denoted as Load_1, **business loads** as Load_2, and **deferred (agricultural) loads** as Load_3. This notation highlights

their respective priority levels, with Load_1 holding the highest criticality and Load_3 the lowest.

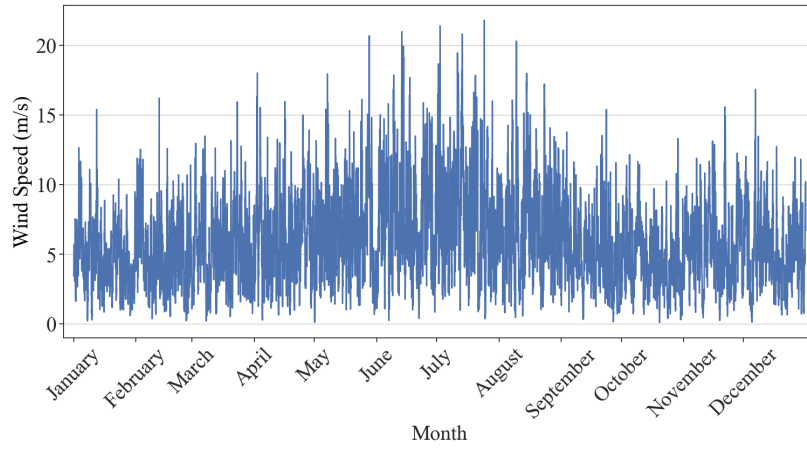
Using HOMER Pro with a load-following dispatch strategy, the adopted configuration comprises 140kW PV, 80kW wind, a 52kW bidirectional converter, and 780kWh of Li-ion storage (1kWh units). Detailed techno-economic outputs (e.g., Total Net Present Cost (TNPC), Cost Of Energy (COE), operating costs, renewable fraction, excess energy) and annual generation plots are reported later in the HOMER Pro software subsection 3.4.1.8.

HOMER Pro is employed to perform techno-economic sizing using long-term weather data and the measured/surveyed load profiles. Solar Global Horizontal Irradiance (GHI) and wind speed are obtained from the NASA POWER database. Figure 3.4 shows the annual GHI and wind speed used as planning inputs, and Figure 3.5 shows the average daily load profile.

For the operational evaluation, real cyclone data (Cyclone Laila, May 2010) was used to validate system performance. Weather data for this period was obtained from the PVGIS database for the same coordinates. Figure 3.6 illustrates the cyclone's track and associated rainfall as observed by NASA's Earth Observatory during landfall in Andhra Pradesh. Figure 3.7 shows different load profiles during the cyclone. From the raw weather data obtained from the PVGIS database [171], [172], PV generation was derived using the `pvlb` library [173], and WT generation was computed using the `windpowerlib` package [174], based on the optimal component sizes obtained from HOMER Pro (Table 3.1). The resulting datasets are illustrated in Figure 3.8: Figure 3.8a shows wind speed alongside the corresponding WT generation, Figure 3.8b presents solar irradiation with derived PV output, and Figure 3.8c compares the aggregated renewable generation against the total load demand during the cyclone period.



(a) Annual solar GHI profile



(b) Annual wind speed profile

Figure 3.4: Annual renewable resource inputs (NASA POWER) for the selected site in Andhra Pradesh, used in the planning stage.

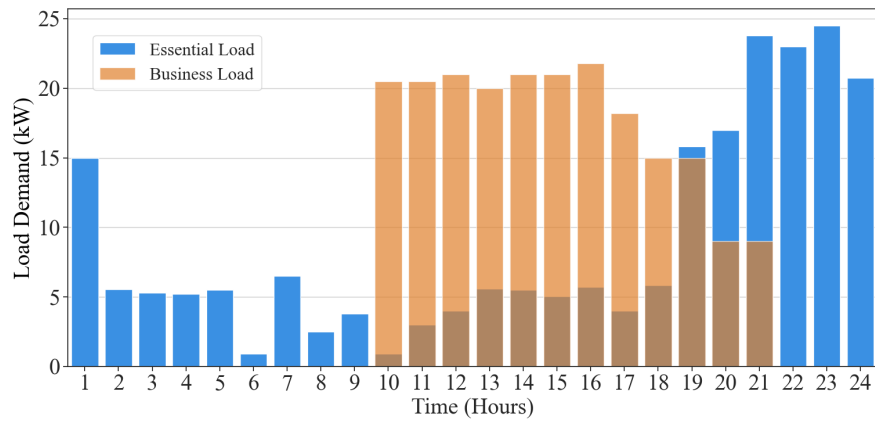


Figure 3.5: Average daily load profile used as a planning input.

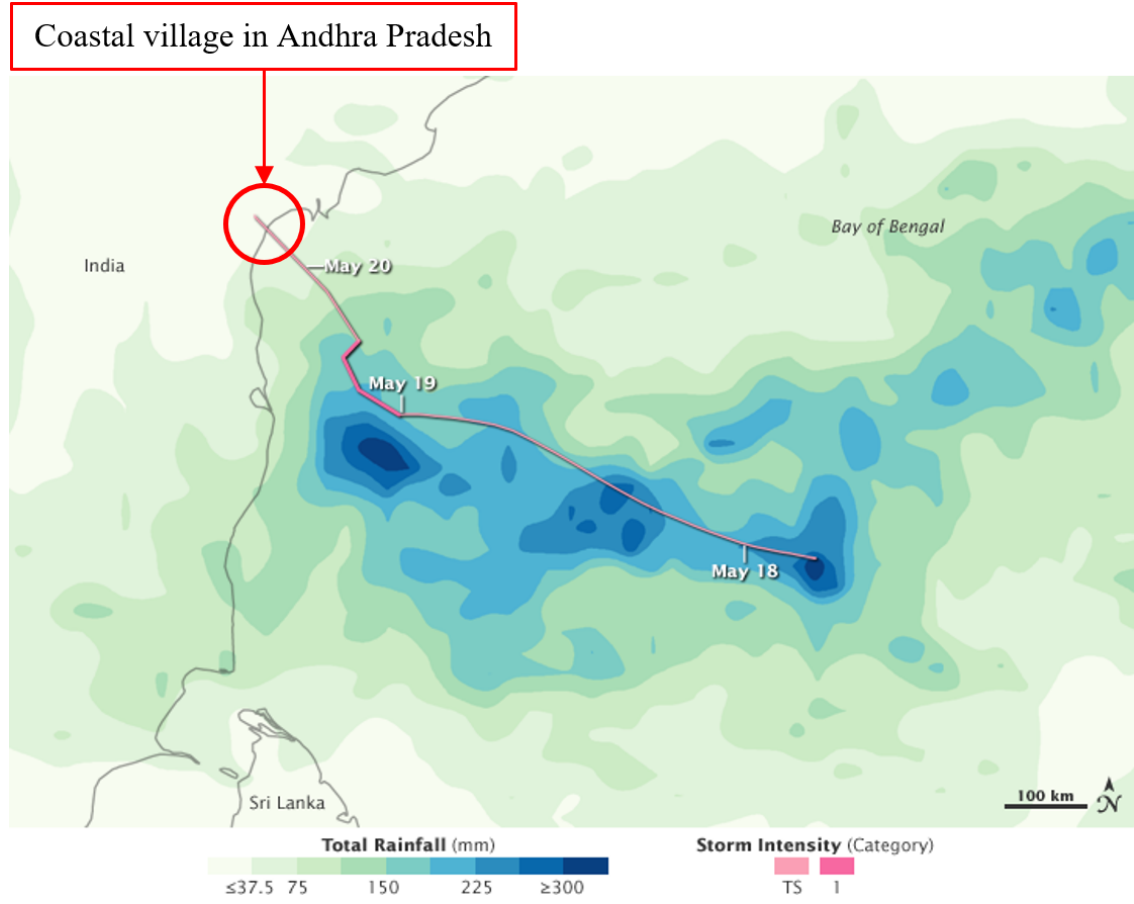


Figure 3.6: Tropical Cyclone Laila soaking southeastern India, observed on 19 May 2010, showing rainfall distribution and storm track [175].

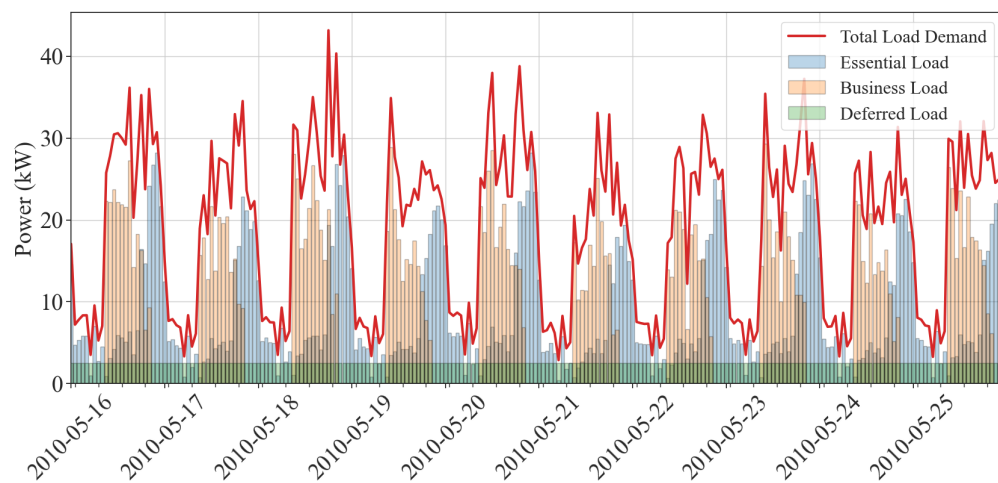


Figure 3.7: Load demand types during Cyclone Laila (19–20 May 2010) for the selected site.

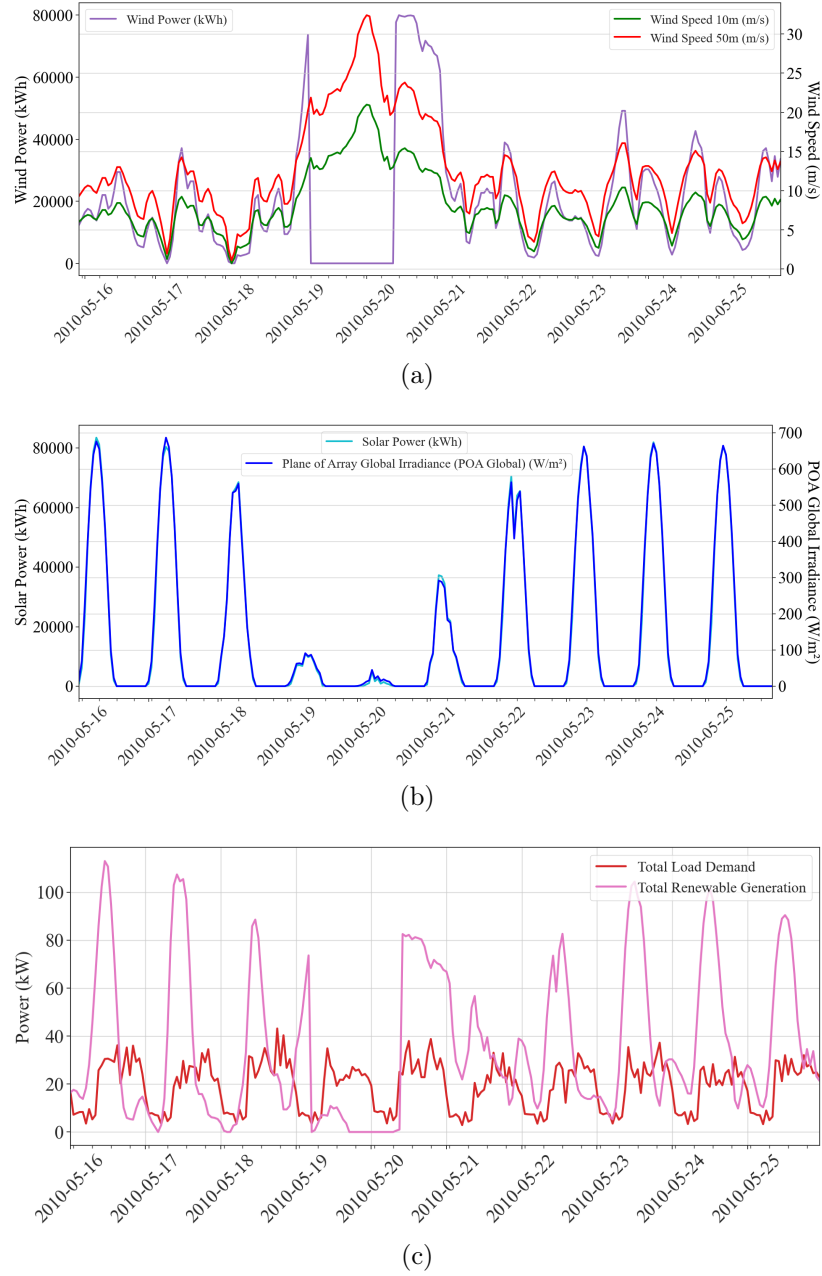


Figure 3.8: Raw weather data and renewable generation profiles during Cyclone Laila (16–25 May 2010) obtained from PVGIS and processed using pvlib and windpowerlib.

Objective Achieved

✓ **Objective 2 – Build a location-specific case.** This objective asked: *How can realistic, location-specific generation and demand profiles be developed for cyclone-prone microgrid case studies, ensuring credible representation of extreme events despite limited metered data?*

It has been achieved by constructing a case study for Kothapatnam, India, using real load surveys and NASA/POWER weather data for planning (HOMER Pro) and PVGIS cyclone records for operation. From this raw data, PV and WT generation were derived using `pvlib` and `windpowerlib`, anchored to the optimal component sizes from HOMER Pro (Table 3.1). As shown in Figure 3.8, renewable generation drops sharply during 19–20 May when Cyclone Laila reached peak intensity, leading to supply shortfalls. Nevertheless, the hybrid configuration demonstrates complementary behaviour: wind generation surges before and after the cyclone until turbine cut-off, while solar PV provides modest output even under cloudy conditions. This complementarity underlines the robustness of the hybrid design and validates the case study as a realistic operational benchmark.

The methodological construction and cyclone validation of this case study have been further detailed in our published article [49].

3.4 Software, Tools, and Modelling

This thesis employs a reproducible, Python-centred toolchain that integrates commercial planning software with open-source optimisation, learning, and explainability libraries. Planning and techno-economic sizing are conducted in HOMER Pro; operational optimisation is formulated in PYOMO and solved as a MILP within an MPC scheme; the learning controller is implemented with SB3 (PPO) in Gymnasium; and post hoc interpretability is provided via

SHAP and LIME. The following subsections detail each component, the interfaces between them, and key configuration choices adopted for reproducibility.

3.4.1 HOMER Pro Modelling: Formulations, Assumptions, and Planning Results

HOMER Pro is employed to obtain techno-economic sizes for the PV, WT, converter, and battery subsystems, and to evaluate levelised cost metrics and annual energy balances under a load-following dispatch strategy. This subsection documents the analytical relations used by HOMER and the parameter choices adopted for the case study.

3.4.1.1 Solar Photovoltaic Generation

The instantaneous PV power output is modelled as

$$P_{PV} = C_{PV} D_{PV} \left[\frac{I_T}{I_T^{\text{STC}}} \right] [1 + \alpha_P (T_C - T_C^{\text{STC}})], \quad (3.1)$$

where C_{PV} (kW) is the rated capacity and D_{PV} is a derating factor that aggregates non-idealities such as soiling, wiring, and ageing. The term I_T/I_T^{STC} scales production by the ratio of in-plane irradiance to its standard-test value, while the bracketed temperature term captures the linear sensitivity of PV power to the cell temperature T_C via the manufacturer's power coefficient α_P (%/°C). Thus, increases in irradiance raise power proportionally, but elevated cell temperature reduces it at a rate given by α_P relative to the Standard Test Conditions (STC) cell temperature T_C^{STC} . STC refer to the laboratory reference settings for photovoltaic evaluation, defined by an irradiance of 1000 W/m², a cell temperature of 25°C, and an air mass of 1.5, which enable consistent measurement and comparison of module performance [176].

The cell temperature is obtained from an energy balance across the PV module cover:

$$\tau_\alpha = \frac{I_T \eta_{PV} + U_L (T_C - T_a)}{I_T}, \quad (3.2)$$

where τ_α is the effective transmittance-absorptance product, η_{PV} is the array efficiency, U_L is the heat transfer coefficient, and T_a is ambient temperature [177]. Physically, a fraction τ_α of incident irradiance I_T is absorbed as useful electrical output $I_T \eta_{PV}$, with the remainder rejected as sensible heat $U_L(T_C - T_a)$.

Manufacturers typically report parameters at Nominal Operating Cell Temperature (NOCT), corresponding to $T_a^{\text{NOCT}} = 20^\circ\text{C}$, $I_T^{\text{NOCT}} = 0.8 \text{ kW/m}^2$, and no electrical load ($\eta_{PV}=0$) [178], [179]. Under these conditions, Equation (3.2) reduces to

$$\tau_\alpha = U_L \frac{T_C^{\text{NOCT}} - T_a^{\text{NOCT}}}{I_T^{\text{NOCT}}}, \quad (3.3)$$

providing a consistent way to estimate τ_α or U_L from NOCT data. HOMER adopts $\tau_\alpha \approx 0.9$ as a default. Substituting this into the rearranged balance yields the operating cell temperature:

$$T_C = T_a + I_T \left(\frac{T_C^{\text{NOCT}} - T_a^{\text{NOCT}}}{I_T^{\text{NOCT}}} \right) \left(1 - \frac{\eta_{PV}}{\tau_\alpha} \right), \quad (3.4)$$

which increases with irradiance I_T and with poorer heat rejection (larger $\Delta T_{\text{NOCT}} = T_C^{\text{NOCT}} - T_a^{\text{NOCT}}$), and decreases with higher conversion efficiency η_{PV} ; for fixed η_{PV} , a larger absorptance-transmittance τ_α slightly increases T_C [177].

Parameterisation. Generic flat-plate multi-crystalline modules with $\eta_{PV} = 14.5\%$ are used [180]. The derating factor is set to $D_{PV} = 0.8$ to capture soiling, wiring, and temperature effects over life [181]. Economic inputs follow [180], [182]: capital \$741/kW, replacement \$741/kW, and operation and maintenance (O&M) \$25/kW/year, with a 20-year lifetime.

3.4.1.2 Wind Turbine Generator

The wind power curve is approximated piecewise as [177]

$$P_{WT}(v) = \begin{cases} \frac{P_{WT}^{\text{rat}}(v - v^{\text{cin}})}{v^{\text{rat}} - v^{\text{cin}}}, & v^{\text{cin}} \leq v \leq v^{\text{rat}}, \\ P_{WT}^{\text{rat}}, & v^{\text{rat}} \leq v \leq v^{\text{co}}, \\ 0, & v \leq v^{\text{cin}} \text{ or } v \geq v^{\text{co}}, \end{cases} \quad (3.5)$$

where v is hub-height wind speed; v^{cin} , v^{rat} , and v^{co} are cut-in, rated, and cut-out speeds; and P_{WT}^{rat} is the rated power.

The wind speed at hub height depends primarily on geographic exposure and the turbine's elevation, and can be estimated by the following relation:

$$v = v^{\text{ref}} \left(\frac{H}{H^{\text{ref}}} \right)^{\alpha}, \quad (3.6)$$

where v is the wind speed at height H , and v^{ref} is the speed measured at the reference height H^{ref} . The exponent α is influenced by factors such as surface roughness, wind speed, time of day, temperature, and season. Typical values range from about 0.05 for smooth, tropical exposures to about 0.4 for very rough terrain in temperate regions [177]. For the planning stage, $\alpha = 0.142$ is adopted to represent the consistent coastal winds at the case study.

To correct the power curve for air-density deviations, HOMER scales output by the density ratio:

$$P_{WT} = \left(\frac{\rho}{\rho_0} \right) P_{WT}^{\text{STP}}, \quad (3.7)$$

where $\rho_0 = 1.225 \text{ kg/m}^3$ denotes standard air density, and P_{WT}^{STP} is the manufacturer curve at Standard Temperature and Pressure (STP) [177], [179].

Parameterisation. A generic 10 kW WT is used as the module, with $v^{\text{cin}} = 3$ m/s, $v^{\text{rat}} = 12$ m/s, $v^{\text{co}} = 25$ m/s, hub height 50 m, and a 20-year life [169]. Economics follow [169], [179]: capital \$11,539 per 10 kW module, replacement \$11,539, and O&M \$230/year.

3.4.1.3 Bidirectional Converter

The power rating of a bidirectional converter interfacing the AC and DC buses can be calculated as [181]:

$$P_{\text{con}} = \frac{P_L}{\eta_{\text{con}}}, \quad (3.8)$$

where P_L and η_{con} represent the peak load demand and the converter efficiency, respectively. In this study, a round-trip efficiency of 90% and a lifetime of 20 years are assumed. The converter capital and replacement costs are \$116/kW each, and the Operation and Maintenance (O&M) cost is \$3/kW/year [180].

3.4.1.4 Battery Energy Storage

The battery is an essential component of a microgrid system as it provides temporal balancing of renewable variability and enables the system to operate independently of the utility grid. Its delivered energy over an interval is

$$E_{\text{bat}} = E_{\text{bat}}^0 + \int_0^t V_{\text{bat}} I_{\text{bat}} dt, \quad (3.9)$$

where E_{bat}^0 is initial stored energy, and $V_{\text{bat}} I_{\text{bat}}$ is instantaneous power [177]. The state of charge evolves with net charge/discharge power $P_{\text{bat}}(t)$:

$$\frac{\text{SOC}(t)}{\text{SOC}(t-1)} = \int_{T-1}^T \frac{P_{\text{bat}}(t) \eta_{\text{bat}}}{V_{\text{bus}}} dt, \quad (3.10)$$

with η_{bat} the (charge or discharge) efficiency and V_{bus} the DC bus voltage. The maximum power capability is

$$P_{\text{bat}}^{\text{max}} = \frac{N_{\text{bat}} V_{\text{bat}} I_{\text{bat}}^{\text{max}}}{1000}, \quad (3.11)$$

with N_{bat} the number of series/parallel modules, V_{bat} the nominal module voltage, and $I_{\text{bat}}^{\text{max}}$ the maximum charging current rating in Amperes [181], [183].

Parameterisation. A generic Li-ion bank is used with SOC bounds $\text{SOC}_{\text{min}} = 10\%$, $\text{SOC}_{\text{max}} = 100\%$, and efficiency $\approx 90\%$ [181]. The \$243/kWh is taken as the capital cost and \$200/kWh as the replacement cost of the battery, while the O&M cost in the present study is taken as \$0.50/kW-year [184].

3.4.1.5 Grid Interaction

Interconnection with the utility grid enables greater utilisation of renewable energy sources, supports emissions reduction, and can stimulate local economic activity. When on-site generation exceeds instantaneous demand, the surplus is exported to the grid and remunerated at the applicable sell-back tariff. This export provides a revenue stream for the microgrid and can enhance system reliability by contributing additional capacity during periods of high demand, thereby helping to stabilise the wider network. In one of the planning scenarios, the microgrid is connected to the local distribution grid for the purpose of selling excess electricity generated by RES; energy imports are not considered. A sell-back price of \$0.080/kWh is assumed for the analyses.

3.4.1.6 Economic Metrics

The TNPC is a measure of the present value of all costs associated with the system over its lifetime, minus any revenues that are generated by the system. HOMER calculates the total cost of the system, including capital costs, operations and maintenance costs, and fuel costs. HOMER reports the TNPC as

$$TNPC = \frac{C_A}{CRF}, \quad (3.12)$$

where C_A is the total annualised cost (\$/year) and Capital Recovery Factor (CRF) [181]. The CRF for nominal interest rate i and project life n years is

$$CRF = \frac{i(1+i)^n}{(1+i)^n - 1}. \quad (3.13)$$

The levelised COE, which represents the cost of producing and delivering one unit of energy, such as 1 kWh of electricity, over the lifetime of the system, is calculated by the following equation:

$$COE = \frac{C_A}{E_T}, \quad (3.14)$$

where E_T is the annual electrical demand served (kWh/year) [181]. These metrics consistently compare isolated and grid-connected modes by accounting for both costs and revenue streams.

3.4.1.7 Planning Assumptions

Analyses are conducted in HOMER Pro v3.14 using an Load Following (LF) dispatch strategy over a 20-year project life with a 2.0% inflation assumption. Weather inputs use NASA/POWER (planning); demand is segmented into priority classes for later operational studies. Technology lifetimes are set to 20 years for PV, WT, and converter; battery replace-

ments follow the economic inputs above.

In the present HOMER Pro analysis, battery ageing is represented at an aggregate planning level through the assumed battery lifetime, replacement schedule, and replacement cost over the project horizon. This aggregate representation captures the financial consequence of battery replacement, but it does not explicitly link degradation to the operating profile, such as SOC, DOD, C-rate, temperature, or cumulative energy throughput.

The work of Gailani and others provides a useful basis for extending this treatment because it shows how calendar- and cycle-induced capacity loss can be translated into a monetary degradation cost for techno-economic assessment [185]. In a planning model, the full physics-based degradation model would not necessarily need to be embedded directly into HOMER Pro. Instead, reduced empirical or semi-empirical degradation factors could be derived from expected annual throughput, equivalent full cycles, DOD, average SOC, C-rate, and site temperature. For each candidate BESS size, an analogous planning-level degradation cost could then be estimated by monetising the predicted capacity loss as

$$C_{\text{deg,plan}} = C_{\text{rep}} \frac{\Delta E_{\text{loss}}}{E_{\text{nom}}}, \quad (3.15)$$

where C_{rep} is the full battery replacement cost, E_{nom} is the nominal battery capacity, and ΔE_{loss} is the estimated capacity loss over the assessment period. This term could be added to the annualised system cost or used to adjust the expected replacement schedule. Such a reduced-order approach is consistent with the energy-throughput and cycle-counting degradation methods reviewed by Vykhotsev et al. [186], while remaining more tractable than embedding a full electrochemical model in the planning optimisation. The DOD-dependent battery-life formulation developed later in this thesis therefore provides a practical basis for a future degradation-aware planning extension, in which BESS sizing, lifecycle cost, resilience, and battery lifetime are co-optimised.

3.4.1.8 Planning Results and Discussion

This section reports detailed results from the planning analysis conducted in HOMER Pro for the proposed microgrid. The focus is on techno-economic sizing, annual generation behaviour, battery utilisation, and sensitivity to key economic parameters. Operational performance under HILP conditions is evaluated later in the MPC/DRL chapters. Two operating modes are considered: (i) isolated (Case 1) and (ii) grid-connected with energy export only (Case 2).

Table 3.1 summarises the optimised configuration and cost metrics. In isolated operation, the system achieves $TNPC = \$494,093$, $COE = \$0.237/\text{kWh}$, and annual operating cost \$8,849. Enabling grid export (with no imports) reduces these to $TNPC = \$442,701$, $COE = \$0.163/\text{kWh}$, and \$4,411, corresponding to decreases of 10.4% in TNPC, 31.2% in COE, and 50.1% in operating cost due to monetising surplus production.

Table 3.1: Techno-economic performance parameters of the microgrid system.

Components/ Case	PV/ (kW)	WT/ (kW)	Batteries/ (Units)	Converter/ (kW)	Total/ NPC (\$)	Operation Cost (\$/year)	COE/ (\$/kWh)
	140	80	780	52	\$494,093	\$8,849	\$0.237
	140	80	780	52	\$442,701	\$4,411	\$0.163

Components/ Case	Initial Cost (\$)	Renewable Fraction (%)	Grid Connection	WT Generation (kWh/year)	PV Generation (kWh/year)	Excess Electricity
	\$391,624	100	No	129,825	204,600	43.5%
	\$391,624	100	Yes	129,825	204,600	25%

Figure 3.9 displays monthly energy production by source. Total annual production is 334,425 kWh, of which PV contributes 204,600 kWh (61.2%) and WT 129,825 kWh (38.8%). A clear seasonal complementarity is observed: wind output peaks in early monsoon (June–July), while PV dominates outside this window. This complementary profile mitigates seasonal energy risk and reduces the depth and duration of storage discharge events.

Figure 3.10 shows the BESS SOC heat map across the year. The bank (780 kWh

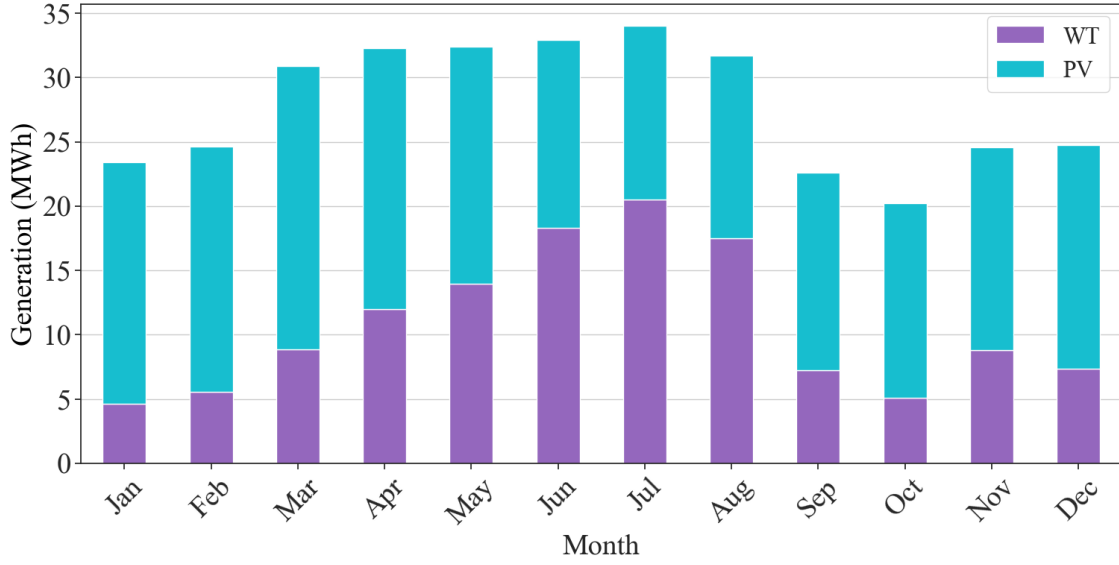


Figure 3.9: Monthly renewable energy production (PV and WT). The profile exhibits strong complementarity: higher WT generation during early monsoon and higher PV generation in shoulder and dry months.

nominal) cycles predominantly within the upper SOC band, with frequent midday charging episodes driven by PV and evening/night discharges meeting demand. The battery’s annual discharged energy estimated by HOMER Pro is 64,949 kWh/yr, corresponding to approximately 178 kWh per day. For a 780 kWh BESS, this is equivalent to roughly 83 full discharge cycles per year. The SOC heat map confirms this pattern, with midday peaks from PV-driven charging and evening/night bands showing routine discharge; low-SOC periods are rare and coincide with simultaneous PV–wind lows. Prolonged high-SOC plateaus indicate persistent surplus and likely curtailment in the absence of flexible demand or export.

HOMER sensitivity studies are conducted for PV, WT, and battery capital costs (multipliers 0.5–2.0), and for the grid sell-back tariff (0.04–0.08 \$/kWh). Figures 3.11a and 3.12b present the TNPC, initial cost, and COE responses.

PV cost sensitivity. For the 140 kW PV subsystem (base capital \$741/kW), halving/doubling capital cost shifts the initial cost between \$339,754 and \$495,364, with TNPC spanning \$390,830.50–\$546,440.50. COE varies from \$0.1435 to \$0.2006/kWh, a range of \$0.057/kWh.

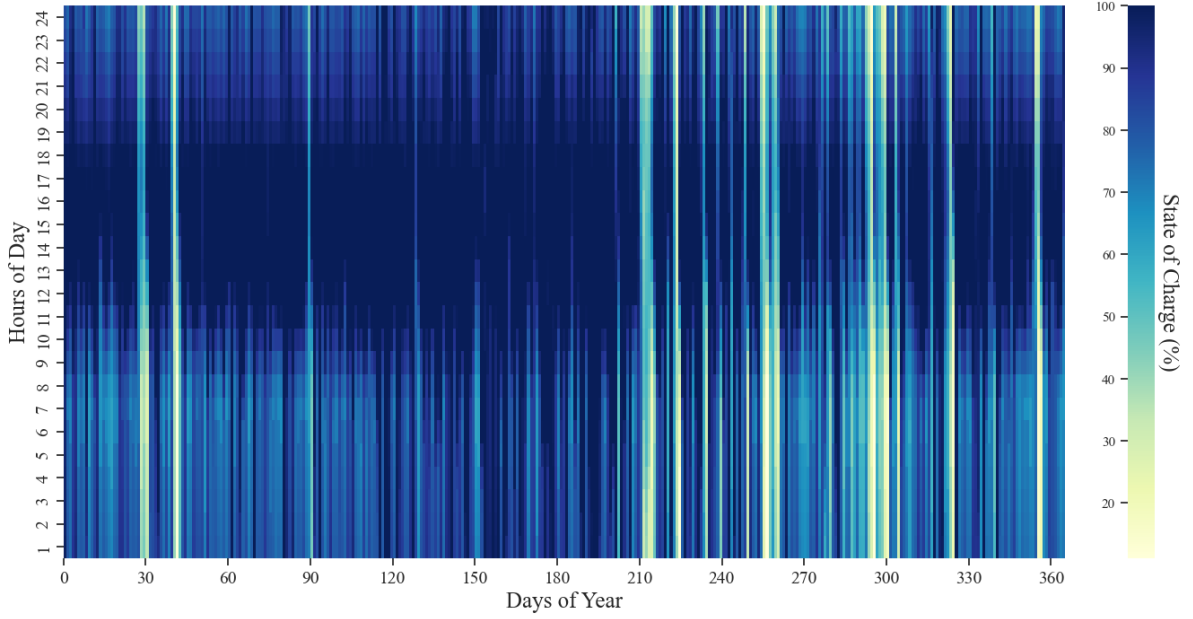
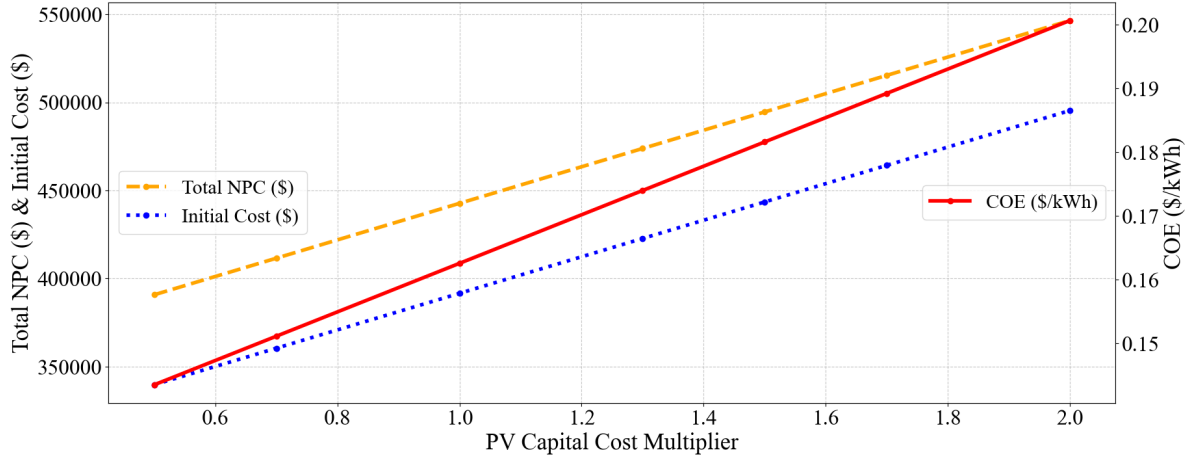


Figure 3.10: Battery utilisation over one year (SOC heat map).

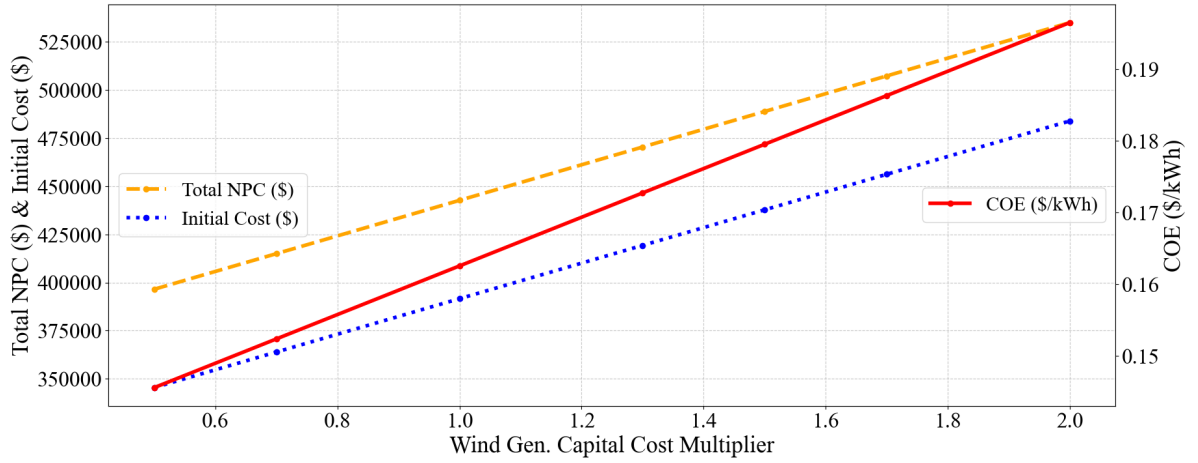
The near-linear trend reflects the capex-dominant nature of a fuel-free, 100% RES design.

WT cost sensitivity. For the 80 kW WT subsystem, capital cost swings yield initial costs of \$345,468–\$483,936 and TNPC of \$396,544.50–\$535,012.50. COE varies from \$0.1456 to \$0.1964/kWh (\$0.051 range). The slope is comparable to PV but marginally smaller, consistent with WT’s lower annual share (38.8%) and the specific cost structure adopted.

Battery cost sensitivity (most influential). For the 780 kWh BESS (base \$243/kWh capex), initial cost ranges from \$296,854 to \$581,164 and TNPC from \$347,930.50 to \$632,240.50. COE spans \$0.1277–\$0.2321/kWh, a \$0.104/kWh swing, the largest among the subsystems. This dominance is expected in a fuel-free configuration where storage capacity directly governs curtailment, excess energy, and the ability to displace backup sources; capital-cost volatility in batteries therefore has the greatest influence on lifecycle costs. *Given this dominant economic leverage, battery longevity becomes an operational priority.* Accordingly, the following chapters model degradation explicitly, via piecewise-linear BLC–DOD terms in the MPC objective, and embed battery expected years in the DRL reward to pursue resilience



(a) PV capital cost multiplier



(b) WT capital cost multiplier

Figure 3.11: TNPC, initial cost, and COE sensitivity to generator capital costs.

without sacrificing battery longevity.

Sell-back tariff sensitivity. Varying the export price between \$0.04 and \$0.08/kWh leaves initial cost unchanged (capex-driven) but monotonically reduces TNPC and COE as the tariff increases. Within this range, COE decreases from \$0.1720 to \$0.1626/kWh (about 5.8% improvement), with a modest TNPC reduction (e.g., to \$442,700.50 at \$0.08/kWh). The approximately linear trend indicates that, for the present sizing, revenue from exports is significant but does not dominate total lifecycle costs; nevertheless, tariff policy remains a material driver of affordability in grid-connected deployments.

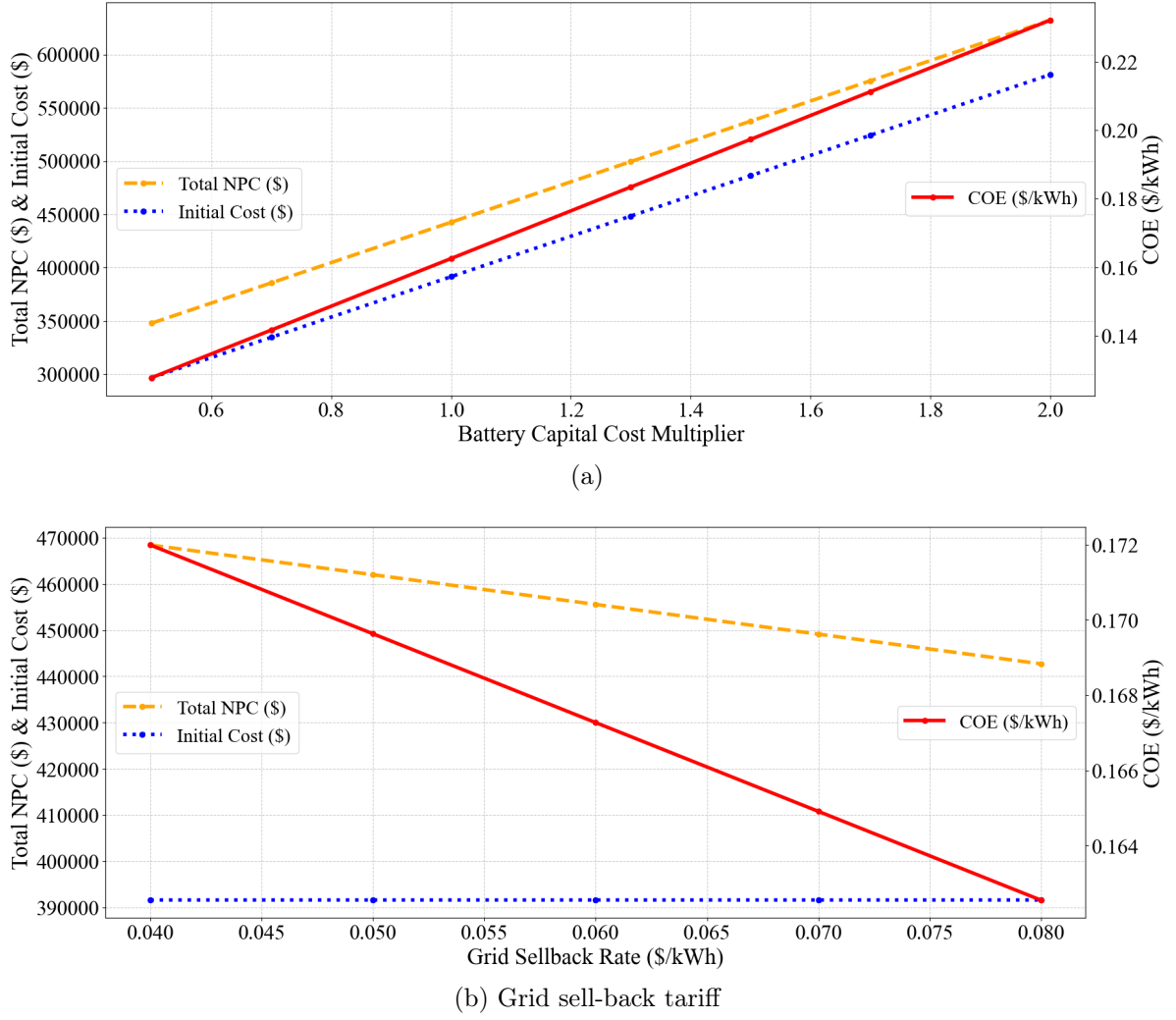


Figure 3.12: TNPC, initial cost, and COE sensitivity to battery cost and sell-back price.

3.4.1.9 Static comparison with a diesel-supported ride-through option.

The selected planning result adopts a renewable-only configuration with a 780 kWh BESS and a 52 kW converter. A practical alternative would be to retain a smaller battery for regular non-HILP balancing and add a diesel generator for extended fault ride-through. A full HOMER re-optimisation with diesel is outside the scope of this thesis; however, a static estimate is useful for quantifying the implication of the battery-only assumption. For illustration, the non-HILP battery is assumed to be 25% of the selected storage capacity, i.e.,

$$E_{\text{BESS, bal}} = 0.25 \times 780 = 195 \text{ kWh.} \quad (3.16)$$

This leaves

$$\Delta E_{\text{BESS}} = 780 - 195 = 585 \text{ kWh} \quad (3.17)$$

as the additional storage capacity associated with the battery-only resilience assumption. Using the NREL 2025 battery storage cost projections for 4-hour lithium-ion systems, the 2024 reference cost is 334 USD/kWh, while the 2035 low–high range is 147–339 USD/kWh [187]. Therefore, the indicative capital value of this additional 585 kWh is approximately

$$585 \text{ kWh} \times 334 \text{ USD/kWh} \approx 195,390 \text{ USD} \quad (3.18)$$

using the 2024 reference value, or

$$585 \times (147\text{--}339) \approx 85,995\text{--}198,315 \text{ USD} \quad (3.19)$$

using the 2035 projected range.

For the diesel-supported alternative, a generator rating close to the converter rating and essential-load peak is required. The essential-load peak is 51.2 kW and the converter rating is 52 kW; therefore, a standard 60 kW diesel generator is used as an indicative comparison. Using the NREL assumption of 800 USD/kW for diesel generator capital cost, the static generator capital cost is

$$60 \times 800 = 48,000 \text{ USD,} \quad (3.20)$$

excluding fuel, maintenance, installation variation, fuel storage, and logistics [188]. This comparison indicates that a diesel-supported design may reduce upfront storage capital cost

relative to a battery-only ride-through design. However, this apparent saving is not equivalent to a full lifecycle advantage, because it introduces fuel dependence and direct emissions during the disruptive event.

The fuel and emissions implications can be estimated using the case-study load data. Supplying only the essential load for a ten-day disruptive period requires

$$E_{\text{ess}} = 222 \times 10 = 2220 \text{ kWh.} \quad (3.21)$$

Supplying all three load categories over the same period requires

$$E_{\text{all}} = (222 + 212 + 58.6) \times 10 = 4926 \text{ kWh.} \quad (3.22)$$

A 60 kW diesel generator data sheet reports fuel consumption of 19.5 L/h at full load, 16.9 L/h at 75% load, and 13.1 L/h at 50% load. Dividing each hourly fuel value by the corresponding electrical output gives approximately 0.325, 0.376, and 0.437 L/kWh, respectively [189]. Using a mid indicative value of 0.38 L/kWh gives

$$F_{\text{ess}} = 2220 \times 0.38 \approx 844 \text{ L,} \quad (3.23)$$

and

$$F_{\text{all}} = 4926 \times 0.38 \approx 1872 \text{ L.} \quad (3.24)$$

Using a diesel emission factor of 2.661 kgCO₂e/L from the UK Government GHG conversion factors gives approximate emissions of

$$844 \times 2.661 \approx 2.25 \text{ tCO}_2\text{e} \quad (3.25)$$

for essential-load support and

$$1872 \times 2.661 \approx 4.98 \text{ tCO}_2\text{e} \quad (3.26)$$

for all-load support over the ten-day event [190]. These estimates are static and indicative; they do not replace a full techno-economic optimisation. They show, however, that diesel support would exchange battery capital cost for fuel-delivery risk, direct emissions, maintenance requirements, and dependence on fuel availability during cyclone conditions.

Accordingly, the battery-only ride-through assumption is retained in this thesis to keep the system boundary consistent with a renewable-only and low-emission microgrid. This does not imply that diesel plus battery systems are technically infeasible or economically unattractive in all contexts. Rather, diesel-supported ride-through is treated as a practical alternative planning option and a relevant direction for future work, especially where lower upfront capital cost or longer ride-through duration is prioritised over emissions, fuel security, and renewable-only operation.

Interpretation of the Planning Results.

- *Resource complementarity* supports supply adequacy (Figure 3.9): wind-intensive monsoon periods offset PV troughs, while PV dominates in shoulder and dry months; thereby lowering the incidence of deep battery cycles.
- *Storage right-sizing and operational care* are pivotal: the strong TNPC/COE elasticity to battery capital cost highlights the importance of procurement timing, technology selection, and demand-side management; equally, it necessitates lifetime-aware operation (limiting unnecessary cycling and DOD).
- *Diesel-supported ride-through* could reduce the amount of battery capacity required for

rare-event support, but it shifts part of the resilience requirement from stored electrical energy to fuel availability, maintenance, and emissions. This is important in cyclone-prone contexts because the same event that disrupts renewable generation and local loads may also disrupt fuel delivery and generator maintenance access.

- *Export pathways and grid access* reduce lifecycle costs without altering the physical configuration; however, this benefit is realisable only when a suitable grid point is nearby or the cost of grid extension is acceptable. Where access is limited or extension is costly, the planning analysis should explicitly include grid-expansion costs.
- *Curtailment and high-SOC plateaus* evident in Figure 3.10 indicate scope for flexible demand (e.g., irrigation pumping, cold-storage pre-cooling) or modest export upgrades to absorb surplus, reduce curtailment, and lower COE.

Remark. The planning results above establish a cost-plausible design baseline and fixed component ratings. Subsequent chapters focus on operational control (MPC and DRL) under deterministic and cyclone-period disturbances, using these ratings and the same demand segmentation to ensure a consistent comparison framework.

Novelty Highlight — Addressing Gap 4

✓ **Contribution: Real, cyclone-grounded case study prior to operation.** Most operational studies depend on IEEE feeders or synthetic data. **Here, a location-specific case in cyclone-prone Andhra Pradesh (near Ongole)** is built using historical cyclone weather and **HOMER Pro** sizing from measured load profiles. The three priority demand classes are preserved from planning to operation, and **fixed capacities** from planning are stress-tested under the actual cyclone window. **This directly addresses Gap 4** by anchoring controller validation in a documented, reproducible real case.

3.4.2 Model Predictive Control Model (PYOMO Implementation)

MPC Multi Input Multi Output (MIMO) formulation captures the interactions among battery operation, class-wise demand, and renewable variability, while its constraint framework enforces converter limits, SOC bounds, Demand Response (DR) limits, and resilience requirements. Combined with its ability to anticipate future conditions through 24-hour load and generation forecasts, and its inherent receding-horizon feedback, MPC is particularly well suited for resilient microgrid operation.

3.4.2.1 Principles of MPC

MPC is an optimisation-based feedback controller that, at each sampling instant, predicts system behaviour over a finite *prediction horizon* using a model, determines a sequence of future control inputs that minimises a specified cost under given constraints, and applies only the first input before re-solving the problem at the next step. This receding-horizon mechanism inherently embeds feedback through state or output re-measurement and effectively manages multivariable interactions, operational constraints, and forecast information. In the state-space formulation, future trajectories are obtained by propagating the model over the prediction horizon and expressing them in compact vector and matrix form, enabling the entire control task to be written as a single optimisation problem at each step.

Key design elements. Prediction horizon (p) should be long enough to capture the dominant dynamics and relevant preview (e.g., demand or generation forecasts). Control horizon ($m \leq p$) limits how many future moves are free variables (often $m \approx 0.1\text{--}0.2p$ or just a few steps), reducing decision dimension without sacrificing look-ahead. The objective typically trades set-point/trajectory tracking against control effort and move-smoothness through weights. Constraints encode hard physical limits (actuators, states) and soft limits

via slack variables when strict feasibility is not guaranteed; excessive hard constraints can lead to infeasible solution. Finally, computational burden grows with model order, number of constraints, and horizon lengths; practical remedies include model-order reduction, shorter p or m , simplifying constraint sets, explicit or parametric MPC, and accepting sub-optimal early termination when appropriate.

How MPC operates in the present microgrid study. Figure 3.13 illustrates the receding-horizon control process with a 24-hour prediction window. At each hourly step, the optimiser determines the control actions for the upcoming 24 hours, including the BESS operating mode (charge, discharge, or idle), the corresponding SOC trajectory, and DR adjustments across the load classes. Only the first-hour action is applied before the window advances by one hour (green box). The measurements and forecasts are then updated, and a new MILP problem is solved to generate the next control input. This sliding-window procedure continues until the end of Day 9; since a 24-hour look-ahead is still required for the final step, forecasts for Day 10 are included to complete the last optimisation. Through this iterative process, feedback is naturally embedded while forecasts are continuously leveraged and operational constraints maintained.

3.4.2.2 From MPC to MILP at Each Time Step

At every hour, the controller solves a mixed-integer linear program on a 24-hour prediction horizon. The MILP selects the battery operating mode through binary variables and sets the corresponding charge or discharge power under converter ratings and SOC dynamics. Demand response is scheduled while keeping the daily energy requirement unchanged. Imbalance and resilience variables quantify unserved energy by load class. The objective combines switching cost, a battery-life term derived from the BLC–DOD relation, a quadratic mismatch penalty, and a normalised resilience score. Each term is scaled to a similar range to

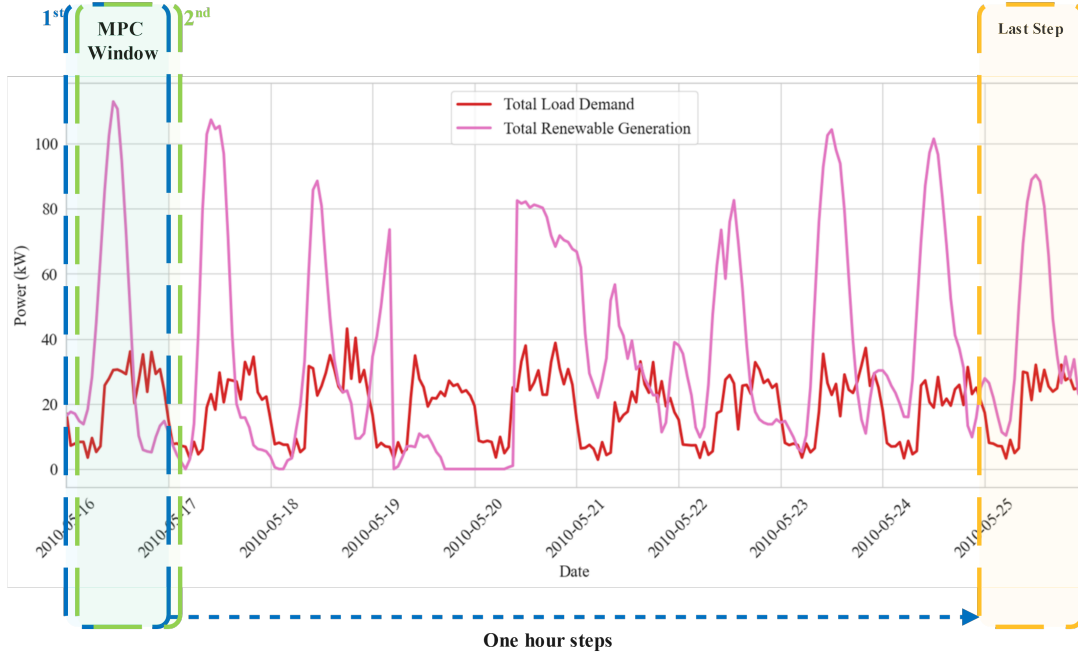


Figure 3.13: Sliding-window MPC with a 24-hour prediction horizon.

keep the optimisation well balanced. The BLC–DOD relation is represented with a piecewise-linear model; segment activation is handled with binaries and Big- M bounds.

The optimisation problem must be written in an algebraic modelling framework that can declare sets, parameters, variables, constraints, and the objective, and then pass the instance to a solver. General packages such as General Algebraic Modeling System (GAMS) are widely used, but they are commercial and need license. This work adopts the open-source PYOMO framework (Version 6.6.2) to encode the MILP model. PYOMO provides a clear, solver-agnostic syntax and seamless interaction with Python for data loading, preprocessing, and post-processing.

For the solver backend, Gurobi (Version 10.0.3) is employed under an academic licence. Gurobi offers robust and reliable performance for large MILP problems. At the beginning of this thesis, general Mixed Integer Non Linear Programming (MINLP) capability in Gurobi was not supported; mature functionality arrived with Version 12 in early 2025. Accordingly,

Table 3.2: List of parameters and their definition.

Parameters	Definition	Parameters	Definition
$Load_1$	Essential Loads (kW)	RE_{Gen}	Total Renewable Generation (kW)
$Load_2$	Business Loads (kW)	η_{ch}	Charging Efficiency
$Load_3$	Agricultural Loads (Deferred Loads) (kW)	η_{dis}	Discharging Efficiency
ω_{bat}	Weighting Factor for Minimizing Total Switching	ω_t	Weighting Factor for Minimizing Total Imbalance
ω_R	Weighting Factor for Minimizing Total Unserved Loads	ω_{BLC}	Weighting Factor for Maximizing Battery Life Cycle
$w_{Essential}$	Prioritizing Factor of Essential Loads	$w_{Business}$	Prioritizing Factor of Business Loads
$w_{Deferred}$	Prioritizing Factor of Deferred Loads	BigM	Big M Factor for Linearization
SOC_{init}	Initial State of Charge	C_{bat}	Cost of Battery (\$/kWh)
$Cycles_{bat}$	Number of Battery Cycles with 100% DOD	C_{CH-DIS}	Cost of each Charging and Discharging (\$)
C_{idle}	Cost of IDLE Mode (\$)	P_{bat}^{min}	Minimum Capacity of Battery Converter (kW)
P_{bat}^{max}	Maximum Capacity of Battery Converter (kW)	P_{bat}^{cap}	Maximum Capacity of Battery (kWh)
DR_1	Allowed Shifting Essential Load (kW)	DR_2	Allowed Shifting Business Load (kW)
DR_3	Allowed Shifting Deferred Load (kW)	$P_{DR-Diff_1}$	Demand Response Differences from Each Iteration for Essential Load (kW)
$P_{DR-Diff_2}$	Demand Response Differences from Each Iteration for Business Load (kW)	$P_{DR-Diff_3}$	Demand Response Differences from Each Iteration for Deferred Load (kW)
Coefficients	Linearized BLC Curve Coefficients	ranges	Different Ranges of Linearized BLC Curve

by using linearisation techniques, the overall problem remains MILP. Beyond tooling constraints, the MILP choice provides certified optimality (or tight optimality gaps), serves as a robust baseline, and imposes a lower computational burden than non-linear formulations.

The index sets structure this optimisation. Let A , K , and B denote the sets in Equation (3.27). Set A contains the battery operating modes (charging, discharging, and idle mode). Set K indexes the 24-hour prediction horizon used by the MPC. Set B indicates the piecewise segments used to linearise the BLC-DOD relation into eight different parts.

$$\text{Set } A : \{'CH', 'DIS', 'IDLE'\} \quad (3.27a)$$

$$\text{Set } K : \{1, 2, 3, \dots, 24\} \quad (3.27b)$$

$$\text{Set } B : \{1, 2, 3, \dots, 8\}. \quad (3.27c)$$

The parameters and variables employed by the MILP are summarised in Tables 3.2 and 3.3. A bar “ $-$ ” symbol denotes auxiliary variables introduced for linearisation.

Table 3.3: Variables and their definitions

Variables	Definition	Variables	Definition
$P_s(k)$	Total Power Supply at Hour k (kW)	$SOC(k)$	State of Charge (0.2, 0.9)
$P_{s_1}(k)$	Power Supply for Essential Loads at Hour k (kW)	$SOC(1)$	SOC_{init}
$P_{s_2}(k)$	Power Supply for Business Loads at Hour k (kW)	$\delta(A, k)$	Showing different Battery Operation Modes at Hour k
$P_{s_3}(k)$	Power Supply for Deferred Loads at Hour k (kW)	$P_{bat}(A, k)$	Battery Power ($P_{bat_{min}}, P_{bat_{max}}$) at Hour k (kW)
$\bar{P}_{bat}(A, k)$	Auxiliary Variable for Linearizing $P_{bat}(A, k)$ (kW)	$BLC(B, k)$	BLC at Hour k
$\sigma(A, A, k)$	Transition of Battery Modes at Hour k	$\delta_{BLC}(B, k)$	Different Linear Equation Lines of BLC Curve at Hour k
$P_{Imb_1}(k)$	Power Imbalance for Essential Loads at Hour k (kW)	$\delta_{Imb_1}(k)$	Binary for Negative Imbalances in Essential Loads at Hour k
$P_{Imb_2}(k)$	Power Imbalance for Business Loads at Hour k (kW)	$\delta_{Imb_2}(k)$	Binary for Negative Imbalances in Business Loads at Hour k
$P_{Imb_3}(k)$	Power Imbalance for Deferred Loads at Hour k (kW)	$\delta_{Imb_3}(k)$	Binary for Negative Imbalances in Deferred Loads at Hour k
$\bar{P}_{Imb_1}(k)$	Negative Imbalance for Essential Loads at Hour k (kW)	$DR_1(k)$	Demand Response Ratio for Essential Loads at Hour k
$\bar{P}_{Imb_2}(k)$	Negative Imbalance for Business Loads at Hour k (kW)	$DR_2(k)$	Demand Response Ratio for Business Loads at Hour k
$\bar{P}_{Imb_3}(k)$	Negative Imbalance for Deferred Loads at Hour k (kW)	$DR_3(k)$	Demand Response Ratio for Deferred Loads at Hour k
$P_{DR_1}(k)$	Essential Loads after Demand Response at Hour k (kW)	$DR_{Change_1}(k)$	Changes in Essential Loads after Demand Response at Hour k (kW)
$P_{DR_2}(k)$	Business Loads after Demand Response at Hour k (kW)	$DR_{Change_2}(k)$	Changes in Business Loads after Demand Response at Hour k (kW)
$P_{DR_3}(k)$	Deferred Loads after Demand Response at Hour k (kW)	$DR_{Change_3}(k)$	Changes in Deferred Loads after Demand Response at Hour k (kW)
$J_{bat}(k)$	Cost of Battery Switching at Hour k	$J_t(k)$	Cost of Power Imbalances at Hour k
$J_R(k)$	Cost of Resiliency Index at Hour k	$J_{BLC}(k)$	Cost of BLC at Hour k

Equation 3.28 shows the battery energy balance update in each hour. The left-hand side, $SOC(k+1) P_{bat}^{cap}$, represents the stored energy at the next hour, while the right-hand side expresses the energy balance as the sum of the current storage level, $SOC(k) P_{bat}^{cap}$, the energy added through charging, $\eta_{ch} \bar{P}_{bat}('CH', k)$, and the energy withdrawn through discharging, $\bar{P}_{bat}('DIS', k)/\eta_{dis}$. Here, $SOC(\cdot) \in [0, 1]$ denotes the per-unit state of charge, and P_{bat}^{cap} is the battery energy capacity in kilowatt-hours. The charge and discharge efficiencies, η_{ch} and η_{dis} , lie within $(0, 1]$. The terms $\bar{P}_{bat}('CH', k)$ and $\bar{P}_{bat}('DIS', k)$ represent *auxiliary* mode-specific powers introduced to maintain linearity. These variables are linked to the actual power flows and binary mode indicators through the big- M relations in (3.29c) and the exclusivity constraint in (3.29a). Operational limits are enforced alongside this balance: the

SOC is bounded each hour, and the charging/discharging powers are capped by the converter rating.

$$SOC(k+1) P_{bat}^{cap} = SOC(k) P_{bat}^{cap} + \eta_{ch} \bar{P}_{bat}('CH', k) - \frac{\bar{P}_{bat}('DIS', k)}{\eta_{dis}}, \quad (3.28a)$$

$$SOC_{min} \leq SOC(k+1) \leq SOC_{max}. \quad (3.28b)$$

The constraint in (3.29a) presents the binary variables $\delta(a, k) \in \{0, 1\}$ and enforces the summation of all modes to be equal to 1. Hence, at each step exactly one operating mode is active and the other two are inactive, any concurrent charging, discharging, or idling is prevented. To ensure that power flows are counted only in the active mode, an auxiliary power $\bar{P}_{bat}(a, k)$ is introduced in Equation(3.29b). This product is non-linear; therefore, it is replaced by the Big- M linearisation in (3.29c). Here, the Big- M value is set to correspond to the maximum physical limit of the converter (52 kW), and wherever Big- M values are used throughout the article, they are selected to be sufficiently large to reflect the maximum physical constraints of the relevant components [191].

The resulting constraints behave as intended. When $\delta('CH', 1) = 1$, the inequalities in (3.29c) simplify to $|\bar{P}_{bat}('CH', 1)| \leq M$, $\bar{P}_{bat}('CH', 1) - P_{bat}('CH', 1) \leq 0$, and $\bar{P}_{bat}('CH', 1) - P_{bat}('CH', 1) \geq 0$. These relations imply that $\bar{P}_{bat}('CH', 1) = P_{bat}('CH', 1)$. Conversely, when $\delta('CH', 1) = 0$, the inequalities reduce to $\bar{P}_{bat}('CH', 1) \leq 0$ and $\bar{P}_{bat}('CH', 1) \geq 0$, which enforce $\bar{P}_{bat}('CH', 1) = 0$; that is, no charging power is recorded at that step. In parallel, the physical mode powers remain bounded each hour by the converter ratings, $P_{bat}^{min} \leq P_{bat}(a, k) \leq P_{bat}^{max}$ for $a \in \{'CH', 'DIS'\}$, and only one mode can be active due to (3.29a). Finally, using $\bar{P}_{bat}(\cdot, k)$ (rather than δP_{bat}) in the SOC update keeps the dynamics linear and compatible with the MILP structure.

$$\delta('CH', k) + \delta('DIS', k) + \delta('IDLE', k) = 1 \quad (3.29a)$$

$$\bar{P}_{bat}(a, k) = \delta(a, k) \times P_{bat}(a, k) \quad (3.29b)$$

$$\left\{ \begin{array}{l} \delta('CH', k) + \delta('DIS', k) + \delta('IDLE', k) = 1 \\ \bar{P}_{bat}(a, k) \leq BigM \times \delta(a, k) \\ \bar{P}_{bat}(a, k) \geq -BigM \times \delta(a, k) \\ \bar{P}_{bat}(a, k) - P_{bat}(a, k) \leq BigM \times (1 - \delta(a, k)) \\ \bar{P}_{bat}(a, k) - P_{bat}(a, k) \geq -BigM \times (1 - \delta(a, k)). \end{array} \right. \quad (3.29c)$$

Mode changes are encoded by the binary variables $\sigma(\alpha, \beta, k) \in \{0, 1\}$ in Equation (3.30).

$$\sigma(\alpha, \beta, k) \geq \delta(\alpha, k-1) + \delta(\beta, k) - 1, \quad \alpha \neq \beta, \quad k > 1, \quad (3.30a)$$

$$\sigma(\alpha, \beta, k) \leq \delta(\alpha, k-1), \quad (3.30b)$$

$$\sigma(\alpha, \beta, k) \leq \delta(\beta, k), \quad (3.30c)$$

$$\sigma(\alpha, \alpha, k) = 0, \quad \forall k, \quad \sigma(\alpha, \beta, 1) = 0, \quad \forall \alpha \neq \beta. \quad (3.30d)$$

Example (idle $\rightarrow$ discharging). If at time $k-1$ the battery is idle and at time k it is discharging, then $\delta('IDLE', k-1) = 1$ and $\delta('DIS', k) = 1$. Substituting into (3.30a)–(3.30c) gives $\sigma('IDLE', 'DIS', k) \geq 1$ and $\sigma('IDLE', 'DIS', k) \leq 1$, hence $\sigma('IDLE', 'DIS', k) = 1$, i.e., a switch is recorded. If either mode is zero (no idle at $k-1$ or no discharge at k), the bounds collapse to $\sigma('IDLE', 'DIS', k) = 0$.

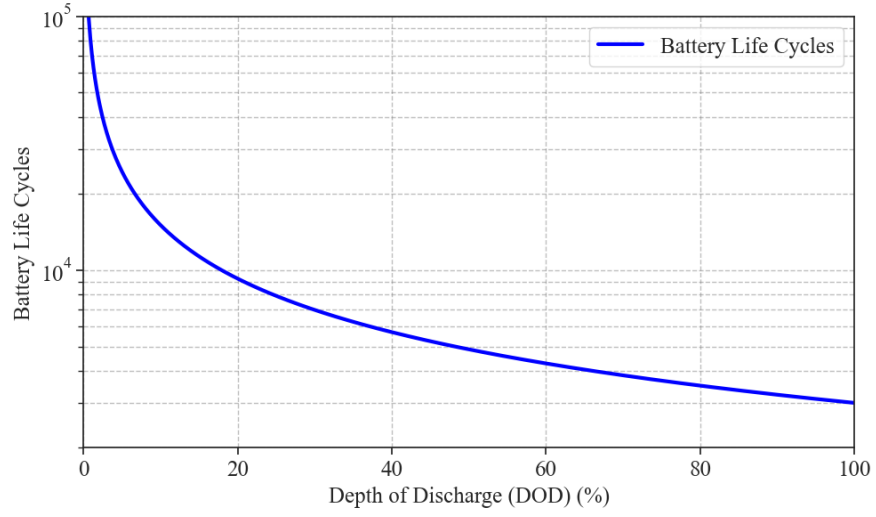
The corresponding battery switching cost term is then

$$\begin{aligned}
 J_{\text{bat}}(k) = & \sum_{k \in K} \frac{\mathcal{C}_{\text{bat}}}{\text{Cycles}_{\text{bat}}} (\delta('CH', k) + \delta('DIS', k)) \\
 & + \mathcal{C}_{\text{CH-DIS}} (\sigma('CH', 'DIS', k) + \sigma('DIS', 'CH', k)) \\
 & + \mathcal{C}_{\text{idle}} (\sigma('IDLE', 'DIS', k) + \sigma('DIS', 'IDLE', k) \\
 & + \sigma('IDLE', 'CH', k) + \sigma('CH', 'IDLE', k)).
 \end{aligned} \tag{3.31}$$

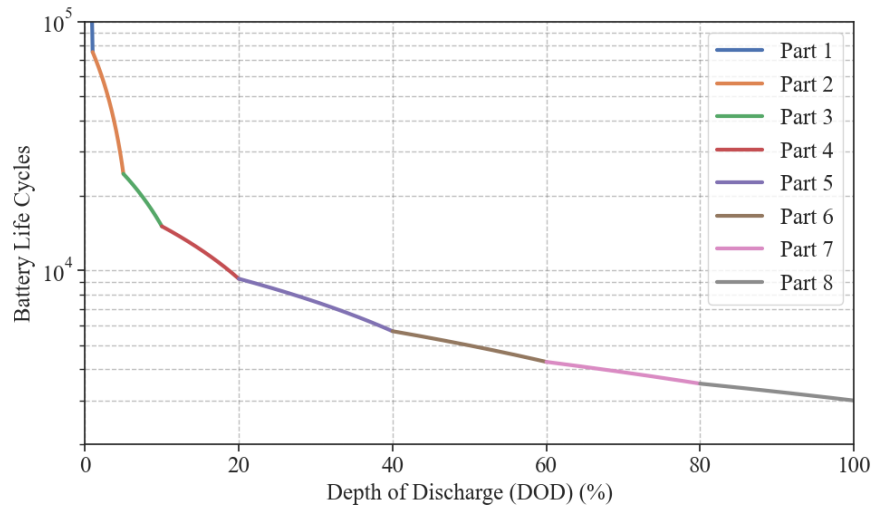
The BLC–DOD characteristic is inherently non-linear: lower DOD yields longer life. To maintain linear-programming compatibility, a piecewise linear approximation with eight segments is adopted as in (3.32). Figure 3.14a displays the non-linear curve, while Figure 3.14b shows its linearised counterpart.

$$y = \begin{cases} -25.35x + 5535, & 80 \leq x \leq 100 \\ -39.1x + 6635, & 60 \leq x < 80 \\ -70.4x + 8513, & 40 \leq x < 60 \\ -177.9x + 12813, & 20 \leq x < 40 \\ -578x + 20815, & 10 \leq x < 20 \\ -1878x + 33815, & 5 \leq x < 10 \\ -12732.75x + 88088.75, & 1 \leq x < 5 \\ -324644x + 400000, & 0 \leq x < 1. \end{cases} \tag{3.32}$$

Eight segments are selected because the aggregate approximation error, measured by an Area Under The Curve (AUC) metric over the DOD domain, remains below 5%. Let $x \in [0, 100]$ denote DOD (%) and $y = f(x)$ the non-linear BLC model used for validation.



(a)



(b)

Figure 3.14: Battery life versus depth of discharge: (a) non-linear curve, (b) Piecewise-linear BLC–DOD approximation.

The reference AUC is

$$A_{\text{nonlinear}} = \int_0^{100} f(x) dx = 1,000,000. \quad (3.33)$$

The linearisation $y \approx m_i x + c_i$ is the piecewise-linear law specified in (3.32). Let the eight segments be denoted by intervals $[a_i, b_i]$ with coefficients (m_i, c_i) for $i = 1, \dots, 8$. For a generic segment $[a_i, b_i]$, the closed-form area is

$$A([a_i, b_i]; m_i, c_i) = \int_{a_i}^{b_i} (m_i x + c_i) dx = \frac{m_i}{2} (b_i^2 - a_i^2) + c_i (b_i - a_i). \quad (3.34)$$

Summing (3.34) over the eight segments of (3.32) yields

$$A_{\text{piecewise}} = \sum_{i=1}^8 A([a_i, b_i]; m_i, c_i) \approx 1,049,750,$$

which is used in the AUC error calculation reported below.

$$\varepsilon_{\text{area}}(B) := \frac{|A_{\text{piecewise}}(B) - A_{\text{nonlinear}}|}{A_{\text{nonlinear}}}, \quad B = \text{number of segments}. \quad (3.35)$$

With $B = 8$,

$$\varepsilon_{\text{area}}(8) = \frac{1,049,750 - 1,000,000}{1,000,000} = 0.04975 (\approx 5\%), \quad (3.36)$$

which justifies the eight-segment partition as a balanced choice between fidelity and computational size. Higher precision can be obtained by introducing additional breakpoints, which reduces $\varepsilon_{\text{area}}(B)$ at the cost of a larger MILP.

Selection of the active line and enforcement of its valid DOD range are given by (3.37) to (3.40).

$$\begin{aligned}
 BLC(b, k) &= \frac{\bar{P}_{\text{bat}}('DIS', k)}{P_{\text{bat}}^{\text{cap}}} \times \text{coefficients}(b, 1) \\
 &+ \delta_{\text{BLC}}(b, k) \times \text{coefficients}(b, 2) \\
 &+ (1 - \delta_{\text{BLC}}(b, k)) \times \text{coefficients}(8, 2).
 \end{aligned} \tag{3.37}$$

$$\sum_{b \in B} \delta_{BLC}(b, k) = \delta('DIS', k). \tag{3.38}$$

$$\begin{aligned}
 \frac{\bar{P}_{\text{bat}}('DIS', k)}{P_{\text{bat}}^{\text{cap}}} \times 100 &\geq \text{ranges}(b, 1) \times \delta_{\text{BLC}}(b, k) \\
 &- \text{BigM} \times (1 - \delta_{\text{BLC}}(b, k)).
 \end{aligned} \tag{3.39}$$

$$\begin{aligned}
 \frac{\bar{P}_{\text{bat}}('DIS', k)}{P_{\text{bat}}^{\text{cap}}} \times 100 &\leq \text{ranges}(b, 2) \times \delta_{\text{BLC}}(b, k) \\
 &+ \text{BigM} \times (1 - \delta_{\text{BLC}}(b, k)).
 \end{aligned} \tag{3.40}$$

$$\left\{ \begin{aligned}
 &-\delta_{BLC}(b, k) \times \text{BigM} \leq B\bar{L}C(b, k) \\
 &B\bar{L}C(b, k) \leq \delta_{BLC}(b, k) \times \text{BigM} \\
 &B\bar{L}C(b, k) - BLC(b, k) \leq (1 - \delta_{BLC}(b, k)) \times \text{BigM} \\
 &-(1 - \delta_{BLC}(b, k)) \times \text{BigM} \leq B\bar{L}C(b, k) - BLC(b, k).
 \end{aligned} \right. \tag{3.41}$$

$$J_{BLC}(k) = \sum_{b \in B} B\bar{L}C(b, k). \tag{3.42}$$

Piecewise BLC selection: step-by-step explanation. Equation (3.37) computes the (per-segment) battery-life contribution at time k if *discharge* occurs. First, the auxiliary discharge power $\bar{P}_{\text{bat}}(\text{'DIS'}, k)$ is nonzero if and only if $\delta(\text{'DIS'}, k) = 1$ (by the Big- M gating), so the term $\bar{P}_{\text{bat}}(\text{'DIS'}, k)/P_{\text{bat}}^{\text{cap}}$ yields the DOD (in per-unit) for that hour. The coefficients satisfy $\text{coefficients}(b, 1) = m_i$ and $\text{coefficients}(b, 2) = c_i$ in the notation of (3.34), with $b \in B$ indexing the piecewise segments. The final term in (3.37) uses $\text{coefficients}(8, 2)$ (i.e., the intercept of the 0–1% segment, 400,000) so that inactive segments remain numerically well-defined; these inactive values are later removed by the gating constraints (3.41).

Detection of active segment. The conditions in Equation (3.38) enforces that no BLC segment is active when there is no discharge, and exactly one segment becomes active during discharge. The range constraints in Equations (3.39) and (3.40) then bind the DOD (%) to the valid interval of the *active* segment: if $\delta_{\text{BLC}}(b, k) = 1$, they reduce to $\text{ranges}(b, 1) \leq 100 \bar{P}_{\text{bat}}(\text{'DIS'}, k)/P_{\text{bat}}^{\text{cap}} \leq \text{ranges}(b, 2)$; if $\delta_{\text{BLC}}(b, k) = 0$, they relax to $-M \leq (\cdot) \leq M$ and impose no restriction.

Introduction of $B\bar{L}C(b, k)$. The Equation (3.41) is the Big- M linearisation of the product $\delta_{\text{BLC}}(b, k) \cdot BLC(b, k)$. When $\delta_{\text{BLC}}(b, k) = 1$, these inequalities force $B\bar{L}C(b, k) = BLC(b, k)$; when $\delta_{\text{BLC}}(b, k) = 0$, they force $B\bar{L}C(b, k) = 0$. Thus only the active segment contributes in the objective via Equation (3.42) (In practice, $B\bar{L}C(b, k)$ is normalised by the upper intercept of the corresponding segment to harmonise magnitudes across lines.)

Illustration. Assume $\delta(\text{'DIS'}, 10) = 1$ and the scheduled discharge implies a DOD of 50% at $k = 10$. Then (3.38) yields activation, and (3.39)–(3.40) select the segment whose range contains 50% (segment $b=3$ with $[40, 60]$). Equation (3.37) evaluates $BLC(3, 10)$ using (m_3, c_3) from (3.32), while (3.41) sets $B\bar{L}C(3, 10) = BLC(3, 10)$ and $B\bar{L}C(b, 10) = 0$ for $b \neq 3$. Consequently, $J_{BLC}(10) = B\bar{L}C(3, 10)$ (up to the stated normalisation), and no other segment contributes.

Novelty Highlight — Addressing Gap 3

✓ **Contribution: Battery longevity embedded in the MPC objective.** Recent studies often *discretise DOD into fixed bands* and assign a constant cycle-life per band (stepwise linearisation), which introduces discontinuities and may distort marginal ageing costs (e.g., [66], [67], [79]). **This work adopts a piecewise-linear, segmented BLC–DOD formulation** that preserves MILP structure and yields smoother behaviour. The eight-segment construction (see (3.32)) is justified by an AUC deviation below 5% (Equation (3.36)), offering a balanced trade-off between precision and model size. This directly addresses *Gap 3* (limited consideration of battery degradation) by reflecting DOD-dependent ageing in operation without sacrificing tractability. For completeness, the DRL framework employs the *exact* non-linear BLC–DOD curve and a curriculum strategy so that resilience objectives are satisfied before optimising battery life.

Before formulating the supply–demand balance, DR is introduced via Equation (3.43). For each 24-hour MPC window, the shiftable portion of demand is temporally reallocated by applying a bounded proportional adjustment to the baseline load. This preserves the intended daily energy requirement within each optimisation window while allowing limited intra-day reshaping. Because the controller is implemented in a receding-horizon manner (one-hour shift per step), a non-negative correction term, $P_{DR-Diff}$, is included to anchor the DR-adjusted profile to the original load profile and prevent cumulative drift across consecutive windows. In the implementation, $P_{DR-Diff}$ is updated as the absolute difference between the total DR-adjusted energy and the baseline energy over the solved window, ensuring that the DR schedule remains strictly referenced to the underlying demand rather than allowing sign-changing residuals to propagate. As the rolling horizon progresses, this anchoring term remains small and does not materially affect the long-horizon energy balance. In the rolling implementation, $P_{DR-Diff}$ is computed as $P_{DR-Diff} = \left| \sum_{k \in K} P_{DR}(k) - \sum_{k \in K} \text{Load}(k) \right|$ and

passed to the next MPC step.

$$\begin{cases} P_{DR}(k) = \text{Load}(k) + \text{DR}_{\text{Change}}(k), \\ \text{DR}_{\text{Change}}(k) = \text{DR}(k) \text{Load}(k), \\ \sum_{k \in K} \text{DR}_{\text{Change}}(k) + P_{DR\text{-}Diff} = 0, \end{cases} \quad (3.43)$$

The supply balance and class allocation are given by (3.44). Supply consists of renewable generation and net battery output (discharge minus charge). For each load class, imbalances may be positive or negative; positive values correspond to surplus energy that cannot be absorbed due to limits or optimisation choices.

$$\begin{cases} P_s(k) = RE_{\text{Gen}}(k) + \bar{P}_{bat}(\text{'DIS'}, k) - \bar{P}_{bat}(\text{'CH'}, k) \\ P_s(k) = P_{s_1}(k) + P_{s_2}(k) + P_{s_3}(k) \\ P_{s_{\text{Load}}}(k) = P_{DR_{\text{Load}}}(k) + P_{\text{Imb}_{\text{Load}}}(k) \\ P_{\text{Imb}}^{\text{Total}}(k) = \sum_{\text{Load}=\{1,2,3\}} \bar{P}_{\text{Imb}_{\text{Load}}}(k). \end{cases} \quad (3.44)$$

The third component of the objective penalises power imbalance via the quadratic expression in (3.45). Although quadratic, this preserves tractability under a Mixed Integer Quadratic Programming (MIQP) framework supported by Gurobi.

$$J_t(k) = \sum_{k \in K} (P_s(k) - P_{DR_{\text{Load}}}(k))^2. \quad (3.45)$$

To identify negative imbalances explicitly, (3.46) is employed.

$$\left\{ \begin{array}{l} -P_{Imb_{Load}}(k) \leq \bar{P}_{Imb_{Load}}(k) \\ -\delta_{Imb_{Load}}(k) \times \text{BigM} \leq \bar{P}_{Imb_{Load}}(k) + P_{Imb_{Load}}(k) \\ \bar{P}_{Imb_{Load}}(k) + P_{Imb_{Load}} \leq \delta_{Imb_{Load}}(k) \times \text{BigM} \\ -(1 - \delta_{Imb_{Load}}(k)) \times \text{BigM} \leq \bar{P}_{Imb_{Load}}(k) \\ \bar{P}_{Imb_{Load}}(k) \leq (1 - \delta_{Imb_{Load}}(k)) \times \text{BigM} \\ P_{Imb_{Load}}(k) \leq \text{BigM} \times \delta_{Imb_{Load}}(k) \\ -P_{Imb_{Load}}(k) \leq \text{BigM} \times (1 - \delta_{Imb_{Load}}(k)). \end{array} \right. \quad (3.46)$$

Using priority weights, the per-step resilience term is expressed in (3.47).

$$J_R(k) = \sum_{k \in K} 1 - \frac{\sum_{Load=1,2,3} (w_{Load} \times \bar{P}_{Imb_{Load}}(k))}{\sum_{Load=1,2,3} (w_{Load} \times P_{Load})}. \quad (3.47)$$

Identification of negative imbalances and resilience metric. The system imbalance for each class is $P_{Imb_{Load}}(k)$, where positive values denote surplus and negative values denote unserved demand. The binary selector $\delta_{Imb_{Load}}(k) \in \{0, 1\}$ and the auxiliary variable $\bar{P}_{Imb_{Load}}(k) \geq 0$ are used in (3.46) to extract the *negative part* of the imbalance in a MILP formulation, i.e., $\bar{P}_{Imb}(k) = \max\{0, -P_{Imb}(k)\}$. Constraints (i)–(ii)–(iii) enforce $\bar{P}_{Imb}(k) = -P_{Imb}(k)$ when $\delta_{Imb}(k) = 0$ (the “deficit” case) and relax otherwise; (iv)–(v) gate $\bar{P}_{Imb}(k)$ to zero when $\delta_{Imb}(k) = 1$ (the “surplus” case); (vi)–(vii) tie the *sign* of $P_{Imb}(k)$ to $\delta_{Imb}(k)$, yielding $P_{Imb}(k) \geq 0$ if $\delta_{Imb}(k) = 1$ and $P_{Imb}(k) \leq 0$ if $\delta_{Imb}(k) = 0$. With a Big- M constant chosen no smaller than the maximum allowable imbalance magnitude, the construction produces:

$$\delta_{Imb}(k) = 0 \Rightarrow \bar{P}_{Imb}(k) = -P_{Imb}(k) (\geq 0), \quad \delta_{Imb}(k) = 1 \Rightarrow \bar{P}_{Imb}(k) = 0.$$

Illustration. If $P_{\text{Imb}}(k) = -3 \text{ kW}$ (deficit), feasibility forces $\delta_{\text{Imb}}(k) = 0$ and the pair of inequalities $\bar{P}_{\text{Imb}}(k) + P_{\text{Imb}}(k) \leq 0$ and $\bar{P}_{\text{Imb}}(k) + P_{\text{Imb}}(k) \geq 0$ yields $\bar{P}_{\text{Imb}}(k) = 3 \text{ kW}$. If $P_{\text{Imb}}(k) = +2 \text{ kW}$ (surplus), feasibility forces $\delta_{\text{Imb}}(k) = 1$ and $(iv)-(v)$ give $\bar{P}_{\text{Imb}}(k) = 0$.

The per-step resilience term (3.47) penalises only *unserved* energy, weighted by class priorities w_{Load} . Consistent with the implementation, the numerator uses the class-weighted deficits at hour k , while the denominator uses the 24-hour sum of class-weighted demands, providing a normalised resilience score $J_R(k) \in [0, 1]$. This construction preserves MILP structure while making resilience an explicit, tunable objective in the optimisation.

Novelty Highlight — Addressing Gap 2

✓ **Contribution: Explicit resilience embedded in the MPC objective.** Operational resilience is modelled *directly* rather than implicitly by: (i) isolating class-wise **negative** imbalances with a Big- M linearisation (see (3.46)), and (ii) optimising a **priority-weighted, normalised** resilience index $J_R(k)$ (see (3.47)) that penalises only unserved energy according to criticality. This design preserves MILP/MIQP tractability while making resilience an explicit, tunable objective in operation. For completeness, the DRL framework employs the exact non-linear resilience term within the reward and uses a curriculum strategy so that resilience targets are satisfied before improving secondary objectives (e.g., battery life). Thus, the approach addresses *Gap 2* by elevating resilience from a by-product to a primary optimisation driver.

Combining all components yields the objective function (3.48); coefficients normalise the terms to comparable scales.

$$OF = \sum_{k \in K} \frac{\omega_{bat} \times J_{bat}(k) \times 18.18}{24} - \frac{\omega_{BLC} \times J_{BLC}(k)}{24} + \frac{\omega_t \times J_t(k)}{24 \times 1600} - \frac{\omega_R \times J_R(k)}{24}. \quad (3.48)$$

The BLC term is segment-normalised and divided by 24 to obtain a daily average in $[0, 1]$, and the resilience index is intrinsically bounded in $[0, 1]$ and likewise averaged over the day. The quadratic imbalance cost is scaled by $1600 = 40^2$ to reflect a conservative 40 kW bound on hourly net mismatch, then averaged by 24. For the battery term, the factor $18.18 = 1/0.055$ maps a charge–discharge switching event (penalty 0.055) to approximately one “unit”, an IDLE transition (0.0275) to ≈ 0.50 , and an active charging/discharging hour ($C_{bat}/Cycles_{bat} = 125/3000 \approx 0.04167$) to ≈ 0.7575 ; the subsequent division by 24 produces a per-day average on the same numerical order as the other normalised components. This calibration equalises magnitudes across criteria and prevents any single term, particularly the switching penalties, from dominating the optimisation.

For performance assessment after optimisation, additional metrics are computed. The overall Resilience Index (RI) over the full study horizon is analogous to (3.47) but aggregated across all hours:

$$RI = \sum_{TimeFrame} 1 - \frac{\sum_{Load=1,2,3} (w_{Load} \times \bar{P}_{Imb_{Load}})}{\sum_{Load=1,2,3} (w_{Load} \times P_{Load})}. \quad (3.49)$$

Although the implemented case study uses three priority load classes, the formulation is not restricted to three classes. Let $\mathcal{L} = \{1, \dots, N_L\}$ denote a general set of load-priority classes, where N_L is the number of classes and ω_i is the priority weight assigned to class i . The priority-weighted resilience index can then be written as

$$\text{RI}_{\mathcal{L}} = 1 - \frac{\sum_{k \in K} \sum_{i \in \mathcal{L}} \omega_i P_{i,k}^{\text{ENS}}}{\sum_{k \in K} \sum_{i \in \mathcal{L}} \omega_i P_{i,k}^{\text{Demand}}} = \frac{\sum_{k \in K} \sum_{i \in \mathcal{L}} \omega_i (P_{i,k}^{\text{Demand}} - P_{i,k}^{\text{ENS}})}{\sum_{k \in K} \sum_{i \in \mathcal{L}} \omega_i P_{i,k}^{\text{Demand}}}. \quad (3.50)$$

The three-class formulation used in this thesis is obtained by setting $N_L = 3$, with the load classes corresponding to essential, business, and agricultural/deferred demand. Additional priority levels could therefore be incorporated by adding new load-indexed demand, curtailed-load, and served-load variables, together with their corresponding weights ω_i , without changing the structure of the optimisation problem.

The same formulation can also be interpreted economically. The dimensionless load-priority weights used in this thesis act as operational proxies for differentiated curtailment penalties. In an economic MPC formulation, each ω_i could be replaced or calibrated using a monetary penalty c_i^{ENS} , expressed for example in \$/kWh or £/kWh of unserved energy, with higher values assigned to more critical loads. Similarly, the battery-life component could be translated into a degradation penalty by assigning a monetary cost to battery throughput or to operation in high-degradation regions, such as deep DOD, high C-rate, or high-temperature cycling. A conceptual economic objective could therefore include terms of the form

$$J_{\text{eco}} = \sum_{k \in K} \sum_{i \in \mathcal{L}} c_i^{\text{ENS}} P_{i,k}^{\text{ENS}} \Delta t + \sum_{k \in K} c_k^{\text{deg}} E_k^{\text{cyc}}, \quad (3.51)$$

where c_i^{ENS} is the curtailment penalty for load class i , E_k^{cyc} is the battery energy cycled at time step k , and c_k^{deg} is a degradation-cost coefficient that may increase under operating patterns associated with faster ageing. This economic interpretation shows how the operational objective could be converted into a cost-based formulation. In this thesis,

the normalised priority-weighted formulation is retained to keep the comparison between MPC and DRL consistent and independent of uncertain monetary penalty assumptions.

Expected Years (EY) are obtained from the BLC–DOD relation by aggregating individual discharge events across the study period:

$$EY = \left(\frac{1}{\sum_{i=1}^{n_{\text{dis}}} \frac{1}{\text{BLC}(\text{DOD}\%_i)}} \right) \times \left(\frac{n_{\text{Days}}}{365} \right). \quad (3.52)$$

Here, $\text{BLC}(\text{DOD}\%)$ is evaluated post-optimisation using the original non-linear BLC–DOD curve (rather than the piecewise-linear surrogate used during MPC), i.e., the realised discharge depths from the optimal schedule are mapped through the non-linear function to obtain the cycle life at each event. The term $\left(\frac{1}{\text{BLC}(\text{DOD}\%_i)} \right)$ represents the fraction of life consumed by that event; the denominator aggregates these fractions over the study period. The overall expression thus returns the number of equivalent full cycles associated with the observed mix of shallow and deep discharges, and the factor $\left(\frac{n_{\text{Days}}}{365} \right)$ converts this quantity to years.

Finally, the Total Switches (TS) is determined by summing all mode transitions:

$$TS = \sum_{i \in \{\text{CH}, \text{DIS}, \text{IDLE}\}} \sum_{\substack{j \in \{\text{CH}, \text{DIS}, \text{IDLE}\} \\ i \neq j}} \sigma(i, j). \quad (3.53)$$

3.4.2.3 Smart Resilient Control Flowchart

Figure 3.15 summarises the supervisory logic that wraps the MPC optimiser presented above. At each hour the controller receives updated measurements and 24-hour forecasts for renewable generation and loads, then evaluates a resilience indicator over the prediction window. If the indicator remains above a preset threshold (e.g., 95%), the *normal* constraint set is used: SOC bounds, battery-life penalties, demand-response budgets, and converter ratings

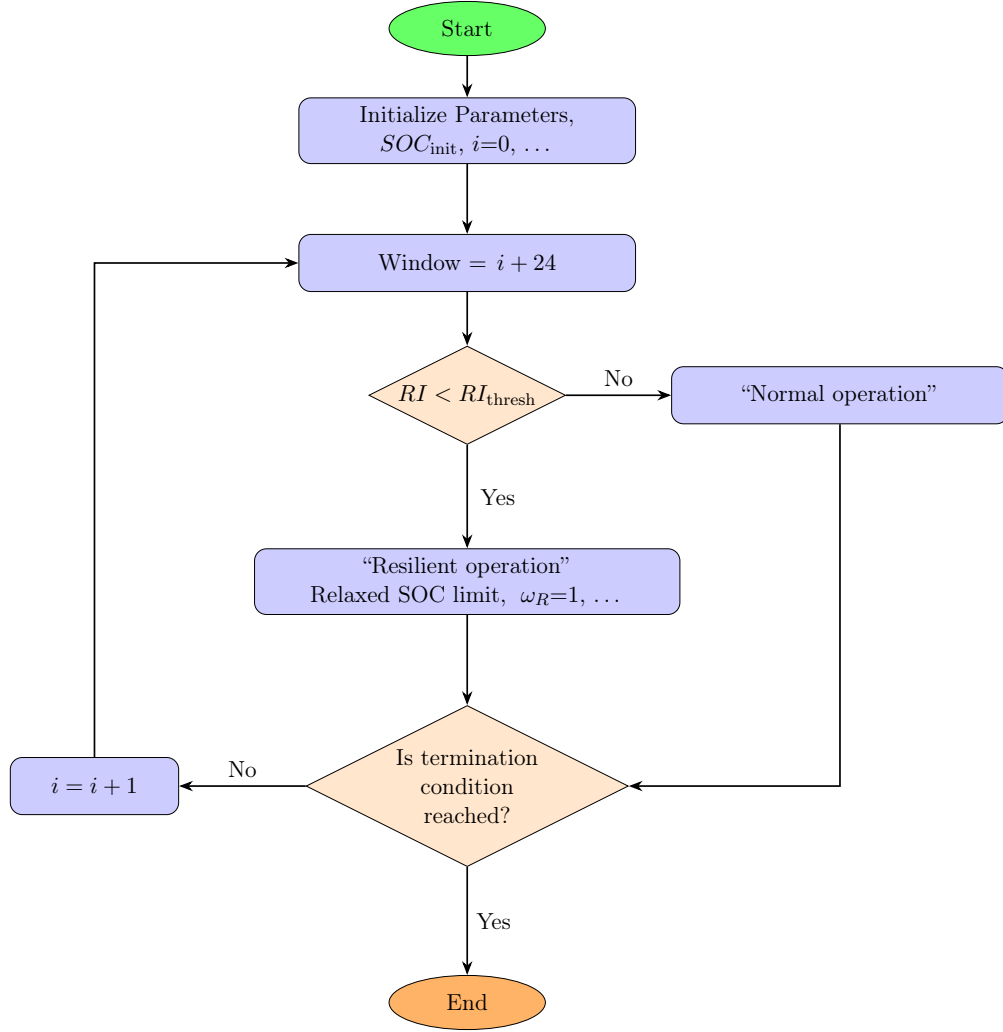


Figure 3.15: Smart resilient control flowchart for MPC-based microgrid operation.

are enforced as specified. If the indicator falls below the threshold, an *emergency* policy is activated that temporarily relaxes SOC margins and battery-life penalties to prioritise essential demand; the class weights in the resilience term ensure that critical loads are served first. The MILP is then solved for the 24-hour horizon, only the first-hour decisions (battery mode and power, demand-response shifts) are implemented, metrics are logged, and the horizon is shifted forward by one hour with refreshed data. This receding-horizon loop continues until the end of the study period.

Despite its strengths, MPC depends on an accurate predictive model and reliable

forecasts. Model mismatches or poorly tuned weights can degrade performance or even cause infeasibility, and effective deployment typically requires expert configuration. Moreover, solving a mixed-integer program at every receding step is computationally demanding, which complicates real-time deployment at the edge and makes large-scale uncertainty studies burdensome. These limitations motivate a complementary, model-free approach that maintains resilience performance while shifting computation off-line. In this thesis, a DRL controller is developed whose inference, once trained, is near-instantaneous and therefore well suited for embedded or remote settings with limited computational resources. It also remains applicable to extensive perturbation analyses under renewable uncertainty.

3.4.3 Deep Reinforcement Learning Modelling (PPO SB3 and Gymnasium Implementation)

3.4.3.1 PPO Foundation

This section is developed based on the foundational article on the PPO algorithm [139]. PPO is a reinforcement learning algorithm that improves policy-gradient methods by ensuring stable and efficient training. PPO strikes a balance between simplicity and performance by using a clipped surrogate objective, which prevents overly large policy updates and improves upon its predecessor, TRPO.

3.4.3.2 Policy Gradient Methods

Policy gradient methods directly optimise the policy $\pi_{\theta}(a|s)$ by maximising the expected return rather than deriving it from value functions. The gradient of the policy parameters

is estimated as

$$\hat{g} = \hat{\mathbb{E}}_t \left[\nabla_{\theta} \log \pi_{\theta}(a_t | s_t) \hat{A}_t \right], \quad (3.54)$$

where $\hat{\mathbb{E}}_t$ denotes the empirical average over a finite batch of samples, ∇_{θ} is the gradient operator with respect to the policy parameters θ , $\pi_{\theta}(a_t | s_t)$ is the probability of selecting action a_t in state s_t under policy π_{θ} , and $\hat{A}_t$ is the advantage function indicating how much better or worse an action performs relative to the expected baseline. When $\hat{A}_t > 0$, the action a_t yields higher than expected performance, leading to an increase in its selection probability for similar states. Conversely, when $\hat{A}_t < 0$, the action performs worse than expected, and its probability of selection is reduced accordingly. The advantage function is defined using Generalized Advantage Estimation (GAE) as:

$$\hat{A}_t = \delta_t + (\gamma\lambda)\delta_{t+1} + (\gamma\lambda)^2\delta_{t+2} + \dots + (\gamma\lambda)^{T-t-1}\delta_{T-1}, \quad (3.55)$$

where δ_t is the TD error at timestep t , given by:

$$\delta_t = r_t + \gamma V(s_{t+1}) - V(s_t). \quad (3.56)$$

The advantage function incorporates several key elements. The term r_t denotes the immediate reward obtained after taking action a_t in state s_t . The discount factor $\gamma \in [0, 1]$ that reduces the weight of future rewards. $V(s_t)$ and $V(s_{t+1})$ represent the estimated values of the current and next states, respectively. Finally, the parameter λ in GAE governs the bias-variance trade-off: when $\lambda = 1$, the full reward trajectory is used, reducing bias at the cost of higher variance; when $\lambda = 0$, only the one-step TD error is considered, lowering variance but increasing bias.

Special Cases of the Advantage Function. Two limiting cases of the GAE can be

identified. When $\lambda = 1$, the advantage uses the full reward trajectory:

$$\hat{A}_t = \sum_{k=0}^{T-t-1} \gamma^k \delta_{t+k}, \quad (3.57)$$

which corresponds to the Monte Carlo estimation approach. Conversely, when $\lambda = 0$, the advantage reduces to the one-step TD estimate:

$$\hat{A}_t = \delta_t. \quad (3.58)$$

In summary, ∇_{θ} determines how to adjust the policy parameters θ to improve the likelihood of better actions based on the advantage $\hat{A}_t$. However, standard policy gradient methods suffer from high variance and are not robust to large policy updates. These limitations motivated the development of constrained and clipped policy optimisation methods, such as TRPO and PPO [139].

3.4.3.3 Trust Region Policy Optimization

TRPO improves policy gradient methods by constraining the update step size to ensure stability. Its objective is formulated as:

$$\max_{\theta} \hat{\mathbb{E}}_t \left[\frac{\pi_{\theta}(a_t|s_t)}{\pi_{\theta_{\text{old}}}(a_t|s_t)} \hat{A}_t \right], \quad (3.59)$$

subject to the constraint:

$$\hat{\mathbb{E}}_t [\text{KL}(\pi_{\theta_{\text{old}}}(\cdot|s_t), \pi_{\theta}(\cdot|s_t))] \leq \delta, \quad (3.60)$$

where $\text{KL}(\cdot)$ denotes the Kullback–Leibler (KL) divergence, a measure of dissimilarity between two probability distributions. The parameter δ defines the maximum allowable divergence, thus controlling the size of the trust region. θ_{old} represents the policy parameters

before the update, and the constraint ensures that the new policy π_θ does not deviate excessively from the previous policy $\pi_{\theta_{\text{old}}}$. This formulation helps maintain monotonic policy improvement while preventing destabilising updates.

The theory justifying TRPO suggests using a penalty on the KL divergence instead of a hard constraint. This leads to the following unconstrained optimization problem:

$$\max_{\theta} \hat{\mathbb{E}}_t \left[\frac{\pi_\theta(a_t|s_t)}{\pi_{\theta_{\text{old}}}(a_t|s_t)} \hat{A}_t - \beta \text{KL}(\pi_{\theta_{\text{old}}}(\cdot|s_t), \pi_\theta(\cdot|s_t)) \right], \quad (3.61)$$

where β is a penalty coefficient that determines how strongly deviations from the old policy are penalized.

Advantages. TRPO ensures stable and monotonic policy improvement by preventing excessively large parameter updates that could destabilise learning. It achieves this by keeping the new policy close to the previous one through the use of KL-based constraints or penalties.

Disadvantages. Despite its stability benefits, TRPO requires solving a constrained optimisation problem that involves second-order derivatives, which significantly increases computational cost. Moreover, selecting an appropriate KL divergence threshold or penalty coefficient can be challenging and task-dependent. As a result, TRPO is often complex to implement and computationally demanding for large-scale applications.

3.4.3.4 Proximal Policy Optimization

PPO simplifies TRPO while maintaining its stability. PPO introduces a clipped surrogate objective that avoids large policy updates without the need for solving constrained optimization problems. PPO uses the probability ratio:

$$r_t(\theta) = \frac{\pi_\theta(a_t|s_t)}{\pi_{\theta_{\text{old}}}(a_t|s_t)}, \quad (3.62)$$

which compares the new policy π_θ to the old policy $\pi_{\theta_{\text{old}}}$. The surrogate objective in PPO is inspired by Conservative Policy Iteration (CPI) Objective:

$$L^{\text{CPI}}(\theta) = \mathbb{E}_t \left[r_t(\theta) \hat{A}_t \right]. \quad (3.63)$$

To ensure stable updates, PPO modifies L^{CPI} by clipping $r_t(\theta)$ to the range $[1 - \epsilon, 1 + \epsilon]$, where ϵ is a small constant (e.g., 0.2). The clipped objective is:

$$L^{\text{CLIP}}(\theta) = \mathbb{E}_t \left[\min \left(r_t(\theta) \hat{A}_t, \text{clip}(r_t(\theta), 1 - \epsilon, 1 + \epsilon) \hat{A}_t \right) \right]. \quad (3.64)$$

An alternative formulation penalizes changes in the policy using the KL divergence:

$$L^{\text{KL PEN}}(\theta) = \mathbb{E}_t \left[r_t(\theta) \hat{A}_t - \beta \text{KL}(\pi_{\theta_{\text{old}}}(\cdot | s_t), \pi_\theta(\cdot | s_t)) \right], \quad (3.65)$$

where β is a penalty coefficient.

The value function loss L^{VF} quantifies the error between the predicted state value $V_\phi(s_t)$ and its target value V_t^{target} . It is expressed as:

$$L^{\text{VF}}(\phi) = \mathbb{E}_t \left[(V_\phi(s_t) - V_t^{\text{target}})^2 \right], \quad (3.66)$$

where $V_\phi(s_t)$ denotes the state value predicted by the critic network parameterised by ϕ . The target value V_t^{target} can be computed either from the full discounted reward trajectory,

$$V_t^{\text{target}} = r_t + \gamma r_{t+1} + \gamma^2 r_{t+2} + \dots + \gamma^{T-t} r_T, \quad (3.67)$$

or using a bootstrapped approach based on the next state's estimated value,

$$V_t^{\text{target}} = r_t + \gamma V_\phi(s_{t+1}), \quad (3.68)$$

where γ is the discount factor that determines how future rewards are weighted.

The entropy loss L^{entropy} promotes exploration by discouraging premature convergence to deterministic policies. It is defined as:

$$L^{\text{entropy}}(\theta) = \hat{\mathbb{E}}_t \left[- \sum_{a \in \mathcal{A}} \pi_\theta(a|s_t) \log \pi_\theta(a|s_t) \right], \quad (3.69)$$

where $\pi_\theta(a|s_t)$ represents the probability of selecting action a in state s_t , and $\mathcal{A}$ is the set of all possible actions. A higher entropy value indicates a more stochastic policy, thereby encouraging diverse exploration; conversely, lower entropy corresponds to more deterministic behaviour.

Overall PPO Objective. The final PPO loss combines the policy surrogate objective, the value function loss, and the entropy term as follows:

$$L_t^{\text{PPO}}(\theta, \phi) = \hat{\mathbb{E}}_t \left[L_t^{\text{CLIP}}(\theta) - c_1 L_t^{\text{VF}}(\phi) + c_2 L_t^{\text{entropy}}(\theta) \right], \quad (3.70)$$

where c_1 and c_2 are weighting coefficients that balance value accuracy and exploration.

3.4.3.5 PPO Algorithm and Implementation Framework

The PPO algorithm alternates between collecting samples and optimising the policy. The pseudocode is:

Algorithm 1 PPO (Actor–Critic Style)

-
- 1: Initialise policy parameters θ (Actor), value function parameters ϕ (Critic), and set $\theta_{\text{old}} \leftarrow \theta$
 - 2: **for** each iteration **do**
 - 3: Run the current policy $\pi_{\theta_{\text{old}}}$ in the environment for T timesteps
 - 4: Collect trajectories (s_t, a_t, r_t, s_{t+1}) and compute advantage estimates $\hat{A}_t$ using the Critic (e.g., via GAE)
 - 5: **for** each epoch **do**
 - 6: Optimise the combined loss

$$L_t^{\text{PPO}}(\theta, \phi) = \hat{\mathbb{E}}_t [L_t^{\text{CLIP}}(\theta) - c_1 L_t^{\text{VF}}(\phi) + c_2 L_t^{\text{entropy}}(\theta)]$$

with respect to the policy parameters θ using stochastic gradient ascent

- 7: Update the value network parameters ϕ by minimising L^{VF}
 - 8: **end for**
 - 9: Update old policy parameters: $\theta_{\text{old}} \leftarrow \theta$
 - 10: **end for**
-

To operationalise PPO in this study, SB3 is adopted as a mature open-source implementation. SB3 provides a robust PyTorch-based realisation of PPO with well-tested default settings and reproducible training utilities. To ensure broad compatibility with widely used reinforcement learning toolchains (e.g., SB3 and CleanRL) and to decouple the control design from any specific library, the environment is implemented using the *Gymnasium* Application Programming Interface (API). This standardised interface for observation, action, and reset/step structure simplifies benchmarking and supports future extensions.

For reproducibility and consistent dependency management, all core library versions are explicitly pinned. This practice ensures seamless integration across the DRL and explainability experiments, preventing API drift and environment conflicts. The libraries and versions used throughout this study are summarised in Table 3.4. These dependencies were installed in the `explainable_ai` Conda environment and used consistently for all PPO training, evaluation, and post-hoc explainability analyses presented in this thesis.

Table 3.4: Core libraries and pinned versions used in the DRL and explainability part.

Library	Version	Library	Version
stable_baselines3	2.3.2	lime	0.2.0.1
gymnasium	0.29.1	torch	2.5.0
numpy	1.26.4	pandas	2.2.3
shap	0.46.0	matplotlib	3.9.2

3.4.3.6 Gymnasium Environment and Decision Model

Gymnasium is a maintained fork of OpenAI Gym that provides a standard, lightweight API for single-agent reinforcement learning environments, together with a registry and a collection of reference tasks. Its interface centres on two calls, `reset()` and `step()`, and standardises the return signature of `step` as (obs, reward, terminated, truncated, info); when an episode ends by either termination or truncation, the agent must call `reset` before continuing. This stable API underpins interoperability across training libraries (e.g., Stable-Baselines3) and simplifies reproducible benchmarking. Figure 3.16 summarises the workflow of custom Gymnasium environment from subclassing to agent training.

Applying the procedure to the microgrid case: Following the above, the microgrid controller is exposed as a Gymnasium *Env*:

1. *Spaces*: the observation is a stacked history of normalised signals (SOC, three different load categories, renewable power, net balance), declared as `Box(6H)`; the action is a `Box(5)` comprising charge/discharge commands and three soft allocation logits.
2. *reset()*: initialises SOC, loads, and renewable profiles for the Cyclone Laila window; returns the first stacked observation and metadata (scenario ID, seeds).
3. *step()*: maps the action to physical power, applies feasibility projection, updates battery SOC, computes delivered power and shortages, and returns the next observation, a shaped reward (resilience + longevity), termination flags, and an `info` dictionary

that aggregates episode metrics (e.g., RI, EY, tier-wise imbalance).

4. *Termination*: episodes end after $T = 216$ steps (truncation) or on safety violations (termination).
5. *Registration*: the class is optionally registered (e.g., "MicrogridXDRL-v0") for use with `gym.make`, enabling drop-in training with SB3.

3.4.3.7 Microgrid Gymnasium Environment Model

The microgrid controller is implemented as a custom `gymnasium.Env` with the standard `reset()` / `step()` API. Episodes span $T = 216$ hourly steps (9 days) drawn from scenario blocks built around the Cyclone Laila window. Symbols and constants follow Table 3.5.

Table 3.5: Variables and Parameters Used in the Environment Model.

Symbol	Description	Value (Unit)
$\omega_t^{(k)}$	Raw allocation preference for load k	–
$L_{k,t} \geq 0$	Demand of load tiers $k \in \{1, 2, 3\}$ (essential, business, agricultural)	– (kW)
$R_t \geq 0$	Renewable generation (PV+WT)	– (kW)
P_t^{net}	Pre-battery net balance $R_t - \sum_{k=1}^3 L_{k,t}$	– (kW)
$S_{k,t}$	Shortage of load k	– (kW)
SOC_t	State of charge in $[\text{SOC}_{\min}, \text{SOC}_{\max}]$	$[0.20, 0.90]$ (p.u.)
$E_{\max}$	Battery energy capacity	780 (kWh)
$P_{b,\max}$	Battery power limit (charge or discharge)	52 (kW)
$\eta_{\text{ch}}, \eta_{\text{dis}}$	Charge and discharge efficiencies	0.90, 0.95
H	History length	5 (steps)
T	Episode length	216 (steps; 9 days)
$f_t^{(k)}$	Softmax allocation fraction for load k	$\in (0, 1)$, $\sum_k f_t^{(k)} = 1$
q_k	Priority weight	$q_1=7, q_2=2, q_3=1$

Observation space: Each step includes a compact summary of states. The six-dimensional observation collects SOC, tiered loads, total renewables, and the pre-battery balance, each normalised for cross-scenario consistency:

$$s_t = [\text{SOC}_t, \hat{L}_{1,t}, \hat{L}_{2,t}, \hat{L}_{3,t}, \hat{R}_t, \hat{P}_t^{\text{net}}], \quad (3.71)$$

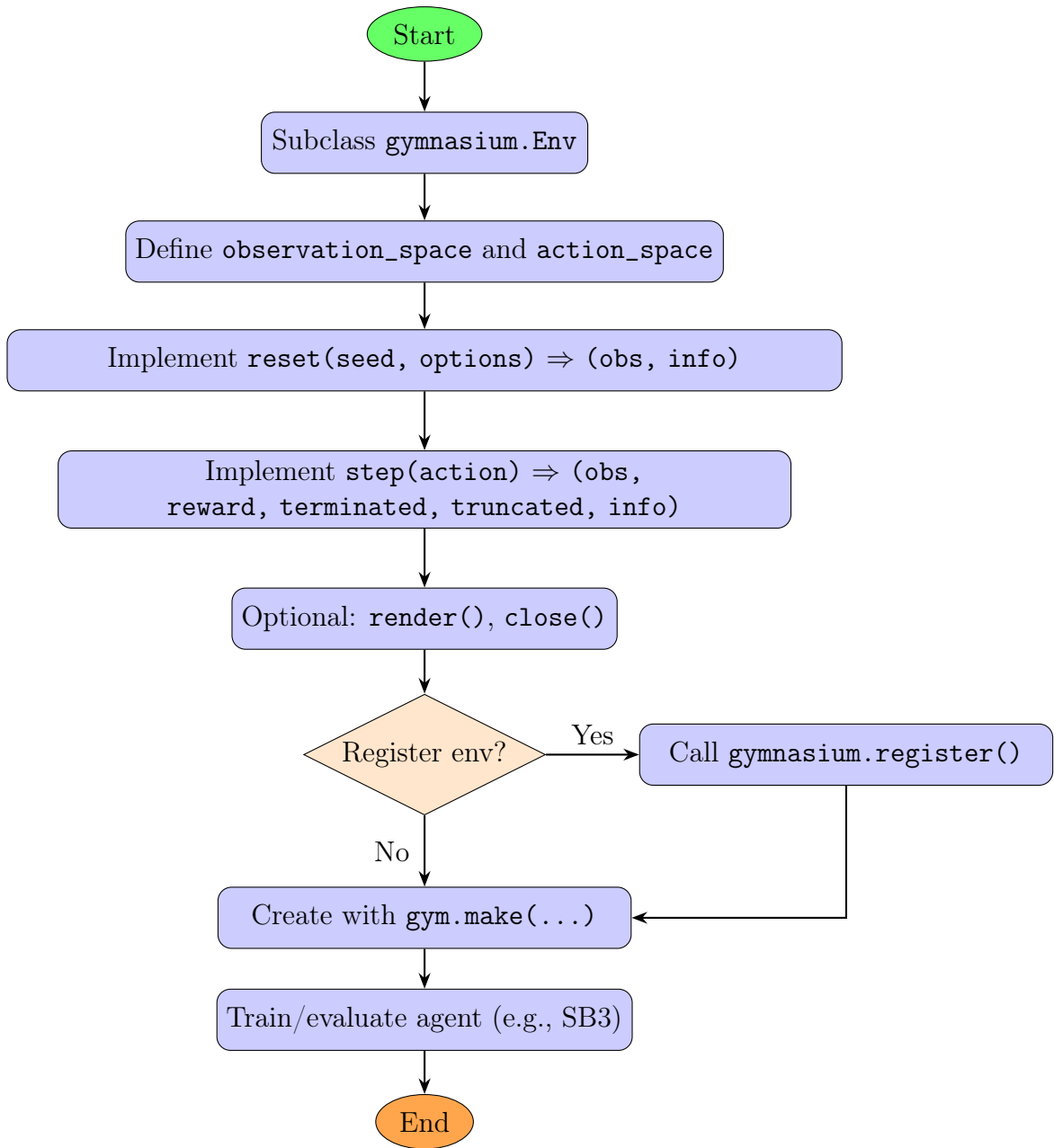


Figure 3.16: Procedure for creating and using a custom Gymnasium environment.

with min-max scaling over the dataset

$$\hat{X}_t = \frac{X_t - X^{\min}}{X^{\max} - X^{\min}}, \quad (3.72)$$

and a stacked history to capture short-term dynamics:

$$\mathbf{o}_t = [s_t, s_{t-1}, \dots, s_{t-H+1}] \in \mathbb{R}^{6H}. \quad (3.73)$$

This representation allows the policy to react to recent trends (evening peaks, windy nights).

Action space: Control consists of a charge command, a discharge command, and three soft allocation logits for the load tiers:

$$\mathbf{a}_t = [u_t^{\text{ch}}, u_t^{\text{dis}}, \omega_t^{(1)}, \omega_t^{(2)}, \omega_t^{(3)}] \in [-1, 1]^5, \quad (3.74)$$

mapped to physical power as

$$P_t^{\text{ch}} = \frac{u_t^{\text{ch}} + 1}{2} P_{b,\text{max}}, \quad P_t^{\text{dis}} = \frac{u_t^{\text{dis}} + 1}{2} P_{b,\text{max}}. \quad (3.75)$$

Mutual exclusivity is imposed by the pre-battery balance, aligning sign with operating mode:

$$P_t^{\text{net}} \geq 0 \Rightarrow P_t^{\text{dis}}=0, \quad P_t^{\text{net}} < 0 \Rightarrow P_t^{\text{ch}}=0. \quad (3.76)$$

Battery power is then clipped by converter limits and remaining capacity to ensure physical feasibility:

$$0 \leq P_t^{\text{ch}} \leq \min \left\{ P_{b,\text{max}}, (\text{SOC}_{\text{max}} - \text{SOC}_t) E_{\text{max}} \right\}, \quad (3.77)$$

$$0 \leq P_t^{\text{dis}} \leq \min \left\{ P_{b,\text{max}}, (\text{SOC}_t - \text{SOC}_{\text{min}}) E_{\text{max}} \right\}. \quad (3.78)$$

Battery state update: SOC evolves under asymmetric efficiencies for charging and discharging:

$$\text{SOC}_{t+1} = \text{SOC}_t + \frac{\eta_{\text{ch}} P_t^{\text{ch}} - \frac{P_t^{\text{dis}}}{\eta_{\text{dis}}}}{E_{\text{max}}}. \quad (3.79)$$

This equation accounts for energy losses during charging and discharging in the system dynamics.

Power allocation and unmet demand: Post-battery power is first formed, then split across loads through a softmax over the logits $\omega_t^{(k)}$:

$$P_t^{\text{s}} = R_t + P_t^{\text{dis}} - P_t^{\text{ch}}, \quad f_t^{(k)} = \frac{e^{\omega_t^{(k)}}}{\sum_{j=1}^3 e^{\omega_t^{(j)}}}. \quad (3.80)$$

Delivered power and tier shortages are

$$P_{k,t}^{\text{s}} = \begin{cases} f_t^{(k)} P_t^{\text{s}}, & P_t^{\text{s}} \geq 0, \\ 0, & P_t^{\text{s}} < 0, \end{cases} \quad S_{k,t} = \max\{0, L_{k,t} - P_{k,t}^{\text{s}}\}. \quad (3.81)$$

The allocation mechanism thus trades among priority tiers while respecting instantaneous power availability.

Per-step reward: Service quality is measured via a priority-weighted coverage at each step. With weights q_k ,

$$D_t = \sum_{k=1}^3 q_k L_{k,t}, \quad N_t = \sum_{k=1}^3 q_k S_{k,t}. \quad (3.82)$$

A curriculum-dependent penalty $\lambda \in \{1, 2, 4\}$ (stages 1–3) amplifies shortfalls under harder scenarios, yielding

$$R_t^{\text{step}} = \begin{cases} 1, & D_t = 0, \\ \max\{0, 1 - \lambda \frac{N_t}{D_t}\}, & D_t > 0, \end{cases} \quad (3.83)$$

and the emitted signal keeps scales comparable across horizons:

$$r_t = \frac{w_{\text{RI}} R_t^{\text{step}}}{T}. \quad (3.84)$$

This shaping concentrates learning on maintaining supply to high-priority tiers as uncertainty increases.

Terminal evaluation: End-of-episode metrics synthesise resilience and battery life. The episode-level Resilience Index is

$$\text{RI}_{\text{ep}} = \begin{cases} 1, & \sum_t D_t = 0, \\ 1 - \frac{\sum_t N_t}{\sum_t D_t}, & \sum_t D_t > 0, \end{cases} \quad (3.85)$$

and expected battery life is calculated from variable-depth cycling. With instantaneous DOD in percent,

$$\text{DOD}_t = \frac{P_t^{\text{dis}}}{E_{\text{max}}} \times 100, \quad (3.86)$$

a harmonic aggregator produces equivalent full cycles and an annualised life:

$$C_{\text{harm}} = \left(\sum_{t: \text{DOD}_t > 0} \frac{1}{\text{BLC}(\text{DOD}_t)} \right)^{-1}, \quad \text{EY} = \frac{C_{\text{harm}} \times n_{\text{days}}}{365}. \quad (3.87)$$

This formula uses the nonlinear curve in Figure 3.14a. In stage 4, a terminal bonus combines RI_{ep} and a normalised lifetime score $\text{EY}/15$ with weights $(w_{\text{RI}}, w_{\text{bat}}) = (0.5, 0.5)$, aligning step-wise behaviour with long-run objectives.

Curriculum and scenario sampling: Although, in principle, one could expose the agent to all perturbation levels and optimise both resilience and battery life simultaneously, in practice this leads to unstable learning because the two objectives pull in opposite directions. Re-

silience rewards supplying all priority loads (encouraging frequent/large discharges), whereas battery life rewards reduced cycling and shallower DOD. Joint training therefore presents the agent with a non-stationary and internally contradictory signal, often yielding policies that are neither reliably resilient nor battery-aware.

To address this, `reset()` implements a four-stage curriculum (Figure 3.17) that first secures robustness in resilience and only then introduces battery preservation. Stages 1–3 sample 216-hour windows with increasing perturbation magnitudes (5–10%, 20–30%, 40–50%), while using reward weights $(w_{\text{RI}}, w_{\text{bat}}) = (1, 0)$ to focus the agent on maintaining served energy under stress. Stage 4 evaluates across the full perturbation range and adds the longevity objective with $(w_{\text{RI}}, w_{\text{bat}}) = (0.5, 0.5)$, allowing the policy to trade small margins of surplus resilience for materially improved battery life. Initial SOC is randomised during training to enhance generalisation and fixed during evaluation for fair, repeatable comparisons.

Termination and logging: Episodes truncate at T steps. The `info` dictionary returns RI_{ep} , EY , tier-wise imbalances, and an action log, enabling consistent evaluation and post hoc analysis.

3.4.4 Shap and Lime Python Libraries for Explainability

With the environment and PPO agent in place, the next task is to clarify *why* the policy selects particular battery and load-allocation actions given a stacked observation. Actor–critic DRL policies based on neural networks are opaque; post hoc model-agnostic tools are therefore used to diagnose decisions without altering the trained controller. Post hoc analysis works well here for two simple reasons. First, it leaves the trained policy exactly as it is, so there is no retraining and no risk of hurting performance. Second, it can be run on any

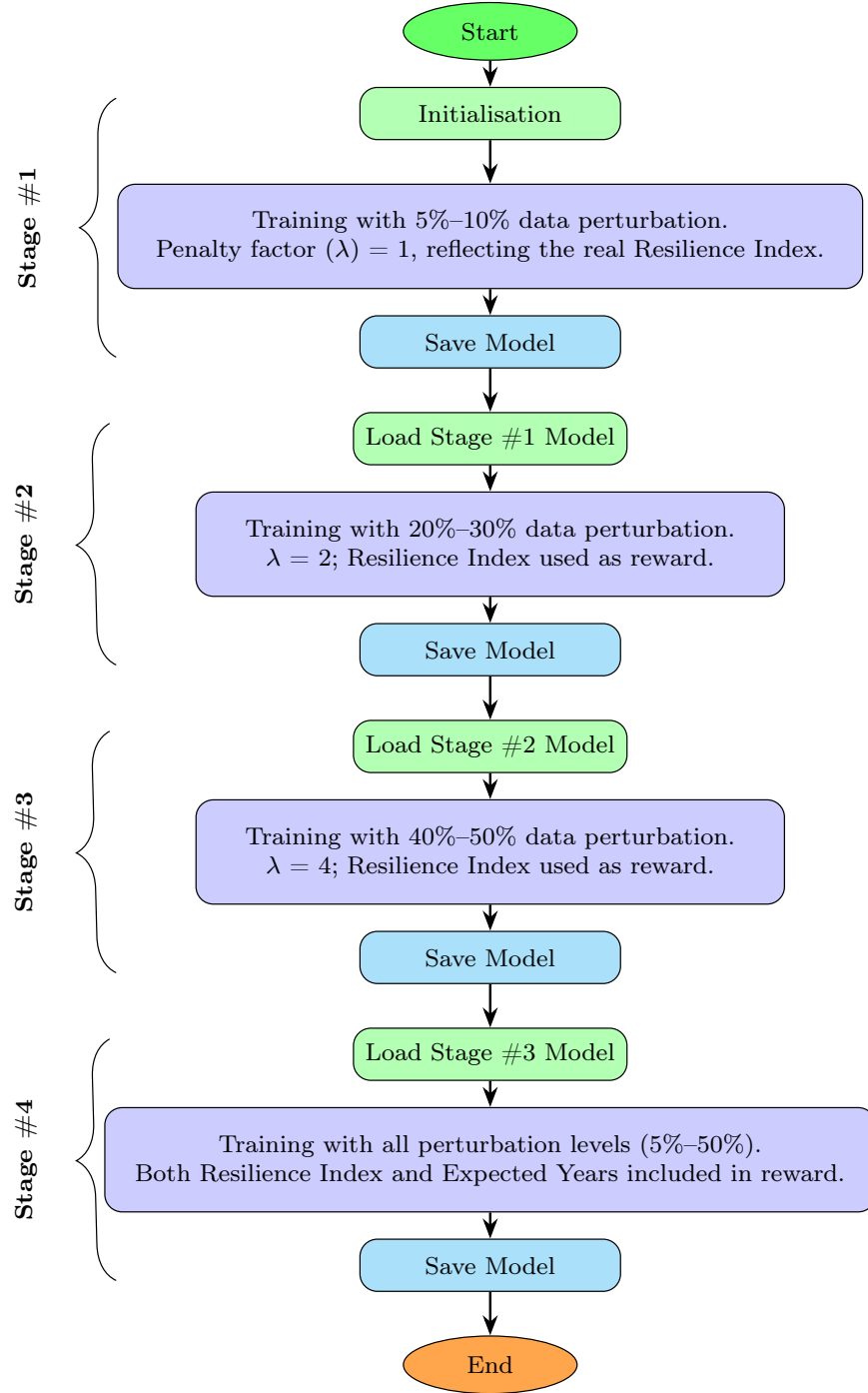


Figure 3.17: Curriculum-based PPO training sequence under progressively perturbed scenarios.

model that is already trained, which makes it easy to audit past checkpoints and explain behaviour across different scenarios. Two complementary libraries are adopted: LIME for local (per-decision) attributions and SHAP for global (dataset-level) attributions.

3.4.4.1 Local Interpretable Model-agnostic Explanations (LIME)

LIME explains a single decision by fitting a simple surrogate model around a target state while keeping the original policy unchanged [140]. For a microgrid decision at time t^* , the target to explain is either one actor output (e.g., charge power) or the critic value. The surrogate is trained only on synthetic perturbations near the target state, so its coefficients can be read as local feature attributions.

Let $\mathbf{o}_t \in \mathbb{R}^{6H}$ denote the stacked observation defined in Section 3.4.3.7. The actor network, parameterised by θ , transforms the observation through a multilayer perceptron $h_\theta(\cdot)$ and applies a hyperbolic tangent activation to bound each action within $[-1, 1]$:

$$\mathbf{a}(\mathbf{o}_t) = \tanh(h_\theta(\mathbf{o}_t)) \in [-1, 1]^5. \quad (3.88)$$

Here, $h_\theta(\cdot)$ represents the neural mapping that combines the hidden layers and the final linear layer of the policy network before activation. The resulting five-dimensional action vector corresponds to the following control components:

$$P^{\text{ch}}(\mathbf{o}_t) = \frac{a_1(\mathbf{o}_t) + 1}{2} P_{b,\text{max}}, \quad P^{\text{dis}}(\mathbf{o}_t) = \frac{a_2(\mathbf{o}_t) + 1}{2} P_{b,\text{max}}, \quad (3.89)$$

$$\omega^{(k)}(\mathbf{o}_t) = a_{k+2}(\mathbf{o}_t), \quad f^{(k)}(\mathbf{o}_t) = \frac{e^{\omega^{(k)}(\mathbf{o}_t)}}{\sum_{j=1}^3 e^{\omega^{(j)}(\mathbf{o}_t)}}, \quad k \in \{1, 2, 3\}. \quad (3.90)$$

The first two outputs correspond to normalised charge and discharge actions, later scaled to kW by the environment. The remaining three logits $\omega^{(k)}$ define relative load preferences that

are converted into allocation fractions $f^{(k)}$ via a softmax, ensuring non-negative proportions that sum to unity. The critic returns $V_\theta(\mathbf{o}_t) \in \mathbb{R}$. LIME is applied in regression mode to each target $\mathcal{F}_\theta^{(i)} \in \{P^{\text{ch}}, P^{\text{dis}}, \omega^{(1)}, \omega^{(2)}, \omega^{(3)}, V_\theta\}$, one at a time.

The interpretable feature vector corresponds to the same tabular input used by PPO: the $6H$ normalised signals,

$$\mathbf{o}_t = [\text{SOC}_t, \hat{L}_{1,t}, \hat{L}_{2,t}, \hat{L}_{3,t}, \hat{R}_t, \hat{P}_t^{\text{net}}, \dots, \text{SOC}_{t-H+1}, \dots]. \quad (3.91)$$

Because the features have different physical units and scales, distances between samples would otherwise be dominated by high-magnitude variables (e.g., loads in kW). To ensure all features contribute comparably to distance calculations, a standard score transformation is applied across the 216-step trajectory:

$$\tilde{o}_i = \frac{o_i - \mu_i}{\sigma_i}, \quad \tilde{\mathbf{o}}^\star = \text{StdScale}(\mathbf{o}_{t^\star}), \quad (3.92)$$

where μ_i and σ_i denote the mean and standard deviation of feature i . This standardisation ensures that the Euclidean distance in the scaled space reflects relative differences between features.

Locality is then defined using an exponential weighting kernel that assigns higher importance to samples near the target state $\tilde{\mathbf{o}}^\star$:

$$W(\tilde{\mathbf{z}}; \tilde{\mathbf{o}}^\star) = \exp\left(-\frac{D(\tilde{\mathbf{z}}, \tilde{\mathbf{o}}^\star)^2}{\sigma^2}\right), \quad (3.93)$$

where $D(\cdot, \cdot)$ is the Euclidean distance and σ controls the spread of the neighbourhood. For example, if $\sigma = 3.0$ and the scaled distance $D = 1$, the corresponding weight is $w = e^{-1/9} \approx 0.89$; if $D = 3$, the weight drops to $w = e^{-1} \approx 0.37$. Thus, nearby samples have stronger influence in the local surrogate fitting, while distant samples contribute less, ensuring that

the explanation remains faithful to the model’s behaviour around the chosen instance.

Given perturbed samples $\{\tilde{\mathbf{z}}_m\}_{m=1}^M$ around $\tilde{\mathbf{o}}^*$, with labels $\mathcal{F}_\theta^{(i)}(\mathbf{z}_m)$ from the trained policy or critic, LIME fits an interpretable model g by:

$$\xi(\mathbf{o}_{t^*}) = \arg \min_{g \in G} \sum_{m=1}^M W(\tilde{\mathbf{z}}_m; \tilde{\mathbf{o}}^*) \left(\mathcal{F}_\theta^{(i)}(\mathbf{z}_m) - g(\tilde{\mathbf{z}}'_m) \right)^2 + \Omega(g), \quad (3.94)$$

where G is a family of simple models and $\Omega(g)$ penalises complexity. The default surrogate in the implementation is a sparse linear model:

$$g(\tilde{\mathbf{z}}') = \mathbf{w}^\top \tilde{\mathbf{z}}' + b, \quad (3.95)$$

with a constraint that limits the number of nonzero coefficients in $\mathbf{w}$, ensuring that the explanation highlights only a few influential features.

Features are indexed to preserve temporal structure; for $k = 0, 1, \dots, H-1$,

$$\{\text{SOC}_{(t-k)}, L_{1,(t-k)}, L_{2,(t-k)}, L_{3,(t-k)}, R_{(t-k)}, P_{(t-k)}^{\text{net}}\}. \quad (3.96)$$

The explainer perturbs the target state $\tilde{\mathbf{o}}^*$ to generate $M=10,000$ synthetic neighbours, evaluates the frozen actor or critic network $\mathcal{F}_\theta^{(i)}(\cdot)$ on these samples, applies locality weights $W(\tilde{\mathbf{z}}_m; \tilde{\mathbf{o}}^*)$ with kernel width $\sigma=3.0$, and fits the surrogate model g . A single 216-step evaluation episode is rolled out to collect states, and a trajectory-specific **StandardScaler** is fitted so that all features are standardised before perturbation. LIME operates in regression mode using these scaled features, where the **StandardScaler** is fitted over the same 216-step episode to ensure consistent local scaling. The resulting coefficients $\mathbf{w}$ quantify the local influence of each feature across recent time steps on the selected target output.

3.4.4.2 SHapley Additive exPlanations (SHAP)

SHAP expresses each model prediction as a sum of feature attributions and a baseline term, using Shapley values from cooperative game theory. This additive form ensures that the individual feature contributions collectively reconstruct the model output, while maintaining local accuracy and global consistency [141].

For a model output at a specific instance $\mathbf{o}_{t^*}$, the Shapley value ϕ_i quantifies the contribution of feature i to that prediction. It is obtained from two components: the marginal contribution of feature i when added to a subset of features, and a weighting factor that ensures fair averaging across all possible orderings. For any subset $S \subseteq \{1, \dots, n\} \setminus \{i\}$ that does not include feature i , the marginal effect is

$$\Delta_i(S) = \mathbb{E}[\mathcal{F}_\theta(\mathbf{o}) \mid \mathbf{o}_{S \cup \{i\}}] - \mathbb{E}[\mathcal{F}_\theta(\mathbf{o}) \mid \mathbf{o}_S], \quad (3.97)$$

measuring how the expected model output changes when feature i is introduced given that only features in S are known. The weighting factor

$$w(S) = \frac{|S|! (n - |S| - 1)!}{n!} \quad (3.98)$$

ensures that contributions are averaged fairly over all possible orderings of features, following the fairness principle of Shapley values. Combining these terms, the final attribution for feature i is expressed as

$$\phi_i = \sum_{S \subseteq \{1, \dots, n\} \setminus \{i\}} w(S) \Delta_i(S), \quad (3.99)$$

where $n=6H$ is the number of input features corresponding to the six signals over the history length H .

Locally, SHAP represents the prediction as an additive decomposition

$$\mathcal{F}_\theta(\mathbf{o}_{t^*}) = \phi_0 + \sum_{i=1}^n \phi_i, \quad \phi_0 = \mathbb{E}[\mathcal{F}_\theta(\mathbf{o})], \quad (3.100)$$

so that the sum of feature attributions equals the deviation from the baseline ϕ_0 .

Two types of outputs are analysed for explainability. The first target is the critic value function, represented as

$$V_\theta(\mathbf{o}_t) \in \mathbb{R}, \quad (3.101)$$

which estimates the expected return from state $\mathbf{o}_t$. The second target is the actor output vector,

$$\mathbf{a}_\theta(\mathbf{o}_t) = [\mathcal{F}_\theta^{(1)}(\mathbf{o}_t), \mathcal{F}_\theta^{(2)}(\mathbf{o}_t), \mathcal{F}_\theta^{(3)}(\mathbf{o}_t), \mathcal{F}_\theta^{(4)}(\mathbf{o}_t), \mathcal{F}_\theta^{(5)}(\mathbf{o}_t)] \in \mathbb{R}^5, \quad (3.102)$$

where $\mathcal{F}_\theta^{(1)}$ and $\mathcal{F}_\theta^{(2)}$ correspond to the normalised charge and discharge control signals, while $\mathcal{F}_\theta^{(3)}$, $\mathcal{F}_\theta^{(4)}$, and $\mathcal{F}_\theta^{(5)}$ represent the three allocation logits, consistent with the formulation used in the LIME analysis. SHAP explanations are generated either for the scalar critic output $V_\theta(\mathbf{o}_t)$ or for each scalar actor component $\mathcal{F}_\theta^{(j)}(\mathbf{o}_t)$ with $j \in \{1, \dots, 5\}$. All actor explanations are reported on their normalised outputs.

A background dataset of 5000 observations ($\mathcal{B}$) is generated by repeatedly resetting the environment (`env.reset()`) to capture a wide range of operating conditions and scenario variations. Using this dataset as the reference distribution, the `shap.Explainer` is applied to both the critic and actor models in a model-agnostic (KernelSHAP-style) configuration. SHAP values are then computed for the same 5000 states, producing feature attributions for the critic output with a shape of $(M, 6H)$ and for the actor outputs with a shape of $(M, 6H, 5)$. Feature names follow the observation structure of the microgrid environment, for example $\{\text{SOC}_{(t)}, \text{Load}_1^{(t)}, \dots, \text{NetEn}^{(t-4)}\}$ when $H = 5$.

For each target, a SHAP summary plot ranks features by their mean absolute attribution $\mathbb{E}[|\phi_i|]$, providing a global importance order; points show local contributions (colour encodes feature value). Dependence plots display (feature value, ϕ_i) pairs to reveal monotonic or interaction effects; heatmaps compactly visualise ϕ_i across many samples. Positive ϕ_i indicates a feature pushes the model output above the baseline, negative ϕ_i the reverse. Because the actor is explained on its normalised outputs, attributions describe how inputs shift the raw action components.

Having specified the planning toolkit, operational controllers, and explainability methods, the study now defines the disturbance scenarios and uncertainty drivers through which these models are trained, stress-tested, and compared.

3.5 Data Preparation and Uncertainty Modelling

This section explains how data variability is incorporated during both training and post-training evaluation. Two complementary approaches are used. First, controlled data augmentation creates multiple training variants by perturbing the baseline time series while enforcing instantaneous and daily limits for each load class. Renewable sums are recalculated, and correlations with the original data are recorded for comparison. Second, structured post-training uncertainty is introduced using one-dimensional Perlin noise to represent smooth and realistic forecast errors during independent evaluation. These methods are applied to three source datasets: a cyclone-grounded window (Cyclone Laila), a normal historical dataset, and a HOMER Pro normal dataset.

3.5.1 Training Dataset Construction via Controlled Perturbations

To enlarge the training set while preserving physical realism, each base dataset is replicated n_{variants} times under multiplicative perturbations. For an original series X_t , a variant applies $X'_t = X_t \times \xi_t$ with $\xi_t \sim \mathcal{U}[1-f, 1+f]$, where $f \in \{0.05, 0.10, 0.20, 0.30, 0.40, 0.50\}$ controls perturbation amplitude. Values below zero are left unchanged to avoid false noise on zeros. Each load profile is first limited by its instantaneous capacity to prevent unrealistic peaks and then adjusted within each 24-hour period to meet its corresponding daily energy constraint. After perturbation, the total renewable output is recalculated as the sum of solar and wind generation to maintain consistency in the energy balance. Finally, each data variant is labelled with its scenario type and perturbation level, along with the correlation values that indicate how closely it resembles the original dataset.

3.5.1.1 Dataset perturbation and assembly

Step-by-step explanation on Algorithm 2.

1. **Random perturbation.** For each dataset, factor level, and variant, every nonzero hourly sample of solar, wind, and each load is multiplied by an independent uniform draw in $[1-f, 1+f]$; values at or below the zero threshold are left unchanged.
2. **Instantaneous limits.** Perturbed load samples are clipped by their instantaneous caps $L_k^{\max}$ to respect per-hour feasible demand.
3. **Renewable recomposition.** The aggregate renewable column is recomputed as the sum of solar and wind after perturbation to maintain additivity.
4. **Daily energy limits.** Within each 24-hour block and for each load class, if the total energy exceeds its cap $E_k^{\max}$, a uniform scaling factor is applied to that day's samples

Algorithm 2 PERTURB_AND_AUGMENT

Require: Datasets $\mathcal{D} = \{\text{EX}, \text{NO_V1}, \text{NO_V2}\}$, factors $F = \{0.05, 0.10, \dots, 0.50\}$, number of variants n , threshold ε , instantaneous limits $L_k^{\max}$, and daily limits $E_k^{\max}$ for $k \in \{1, 2, 3\}$

Ensure: Augmented dataset with all variants, metadata, and correlations

```

1: for all  $D \in \mathcal{D}$  do                                     ▷ iterate over base datasets
2:   for all  $f \in F$  do                                       ▷ iterate over perturbation factors
3:     for  $v \leftarrow 1$  to  $n$  do                                   ▷ generate  $n$  variants
4:        $P \leftarrow \text{copy}(D)$ 
5:       for all row  $i$  in  $P$  do
6:         for all  $c \in \{\text{Solar}, \text{Wind}, \text{Load}_1, \text{Load}_2, \text{Load}_3\}$  do
7:           if  $P_i[c] > \varepsilon$  then
8:             draw  $r \sim \mathcal{U}(1 - f, 1 + f)$ 
9:              $P_i[c] \leftarrow P_i[c] \cdot r$ 
10:          if  $c$  is a load then
11:             $P_i[c] \leftarrow \min(P_i[c], L_c^{\max})$ 
12:          end if
13:        end if
14:      end for
15:    end for
16:    recompute  $P[\text{Sum\_Renewable}] \leftarrow P[\text{Solar}] + P[\text{Wind}]$ 
17:    for all day  $d$  in  $P$  do                                     ▷ 24-hour blocks
18:      for all  $k \in \{1, 2, 3\}$  do
19:         $S \leftarrow \sum_{t \in d} P_t[\text{Load}_j]$ 
20:        if  $S > E_k^{\max}$  then
21:           $\alpha \leftarrow E_k^{\max} / S$ 
22:           $P_t[\text{Load}_k] \leftarrow \alpha P_t[\text{Load}_k] \quad \forall t \in d$ 
23:        end if
24:      end for
25:    end for
26:    compute correlations  $\rho \leftarrow \text{corr}(P, D)$ 
27:    append  $P$  with labels (Scenario,  $f$ ,  $\rho$ ) to master list
28:  end for
29: end for
30: end for
31: concatenate all variants, sort by scenario order ( $\text{EX} \prec \text{NO\_V1} \prec \text{NO\_V2}$ )
32: save result as Augmented_data_all_factors.csv

```

of the corresponding load so that the daily cap is met while preserving intra-day shape.

5. **Correlation logging.** Pearson correlations are computed between the perturbed and original series for the five key signals; the results are stored as metadata to track similarity to the source.
6. **Assembly and export.** All variants are labelled with scenario identifiers and perturbation factors, concatenated, stably ordered by scenario (EX, NO_V1, NO_V2), and written to a single CSV file for training and analysis.

Scenario visualisation and training coverage. Figures 3.18—3.20 summarise the augmented time series used during training to improve DRL generalisation. Within each scenario (EX, NO_V1, NO_V2), panel (a) shows the three load classes (essential, business, agricultural) and their total, panel (b) shows solar, wind, and their total, and panel (c) overlays total load and total renewables. Shaded ribbons depict the 10–90% range across variants, while solid lines show the per-time stamp mean. All variants from each scenario are included during training so the policy is exposed to both typical and stressed operating conditions.

Figure 3.21 illustrates how the Pearson correlation between the perturbed and original datasets declines as the perturbation factor increases from 5% to 50%. Each scenario (EX, NO_V1, and NO_V2) exhibits a consistent monotonic reduction in correlation, confirming that the augmented datasets introduce controlled variability while maintaining structural realism. Among all signals, the agricultural load shows the steepest decline, indicating higher sensitivity to perturbations, whereas the essential and business loads remain strongly correlated even at higher levels. The renewable components (PV and wind) retain relatively stable correlations, reflecting smoother and less abrupt perturbation effects. This systematic variation ensures that the DRL agent is trained over datasets that progressively differ from the baseline, enhancing robustness and generalisation to unseen operating conditions.

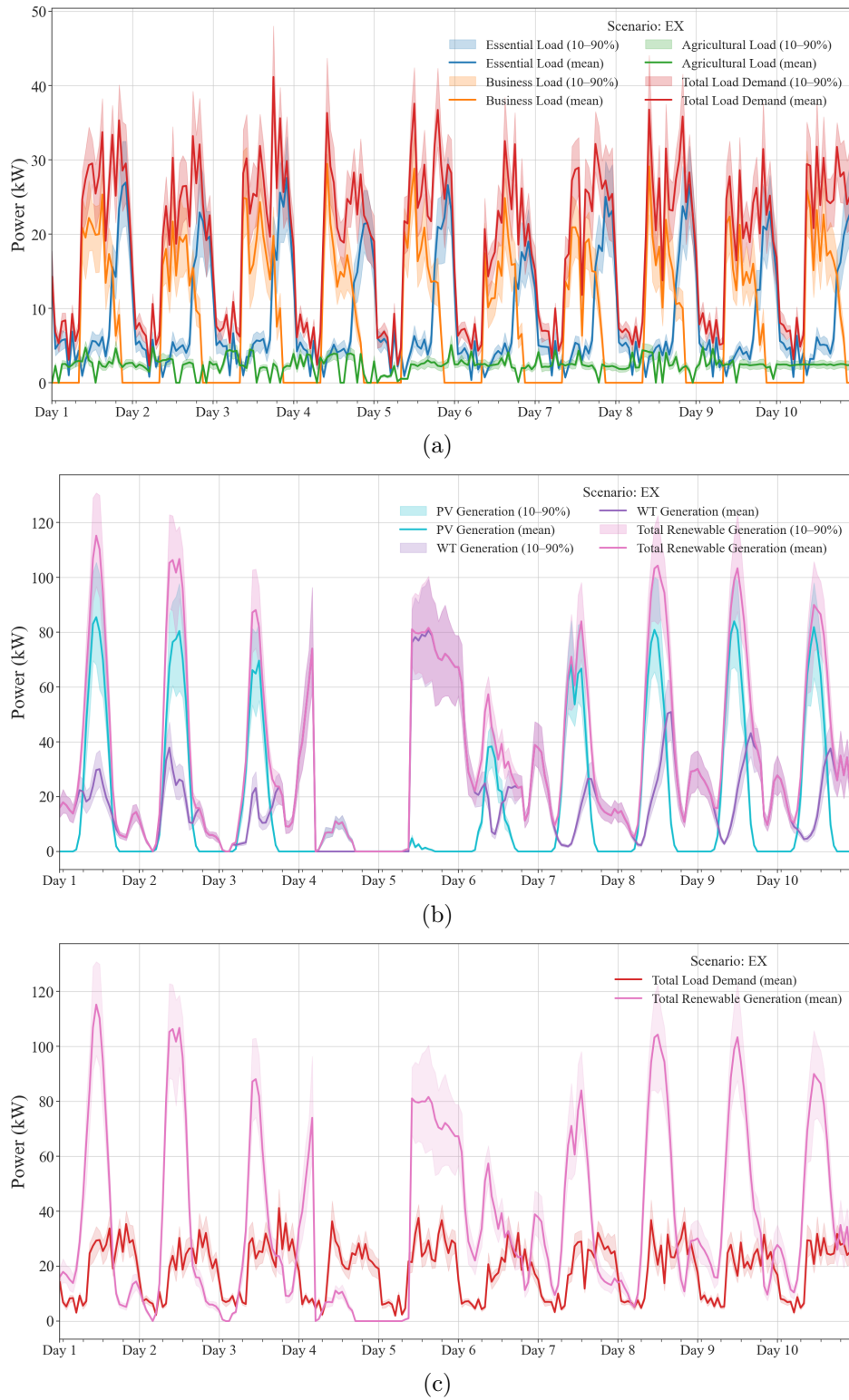


Figure 3.18: EX scenario (Cyclone Laila) — augmented datasets used in training: (a) loads and total demand, (b) renewable generation, and (c) total load versus total renewables. Ribbons show the 10–90% range; solid lines indicate the mean values.

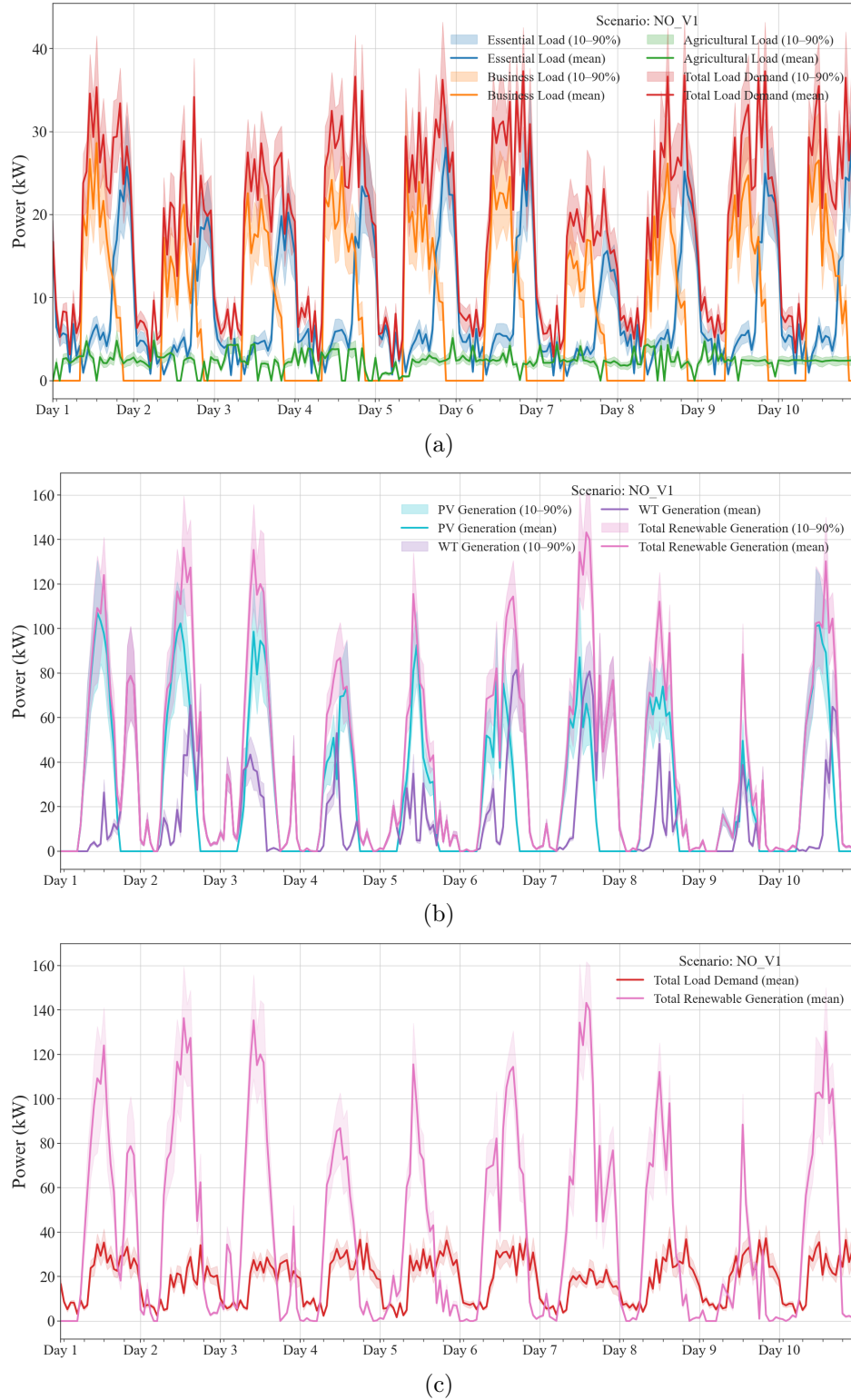


Figure 3.19: NO_V1 scenario (normal-HOMER Pro) — augmented datasets used in training: (a) loads and total demand, (b) renewable generation, and (c) total load versus total renewables. Ribbons show the 10–90% range; solid lines indicate the mean values.

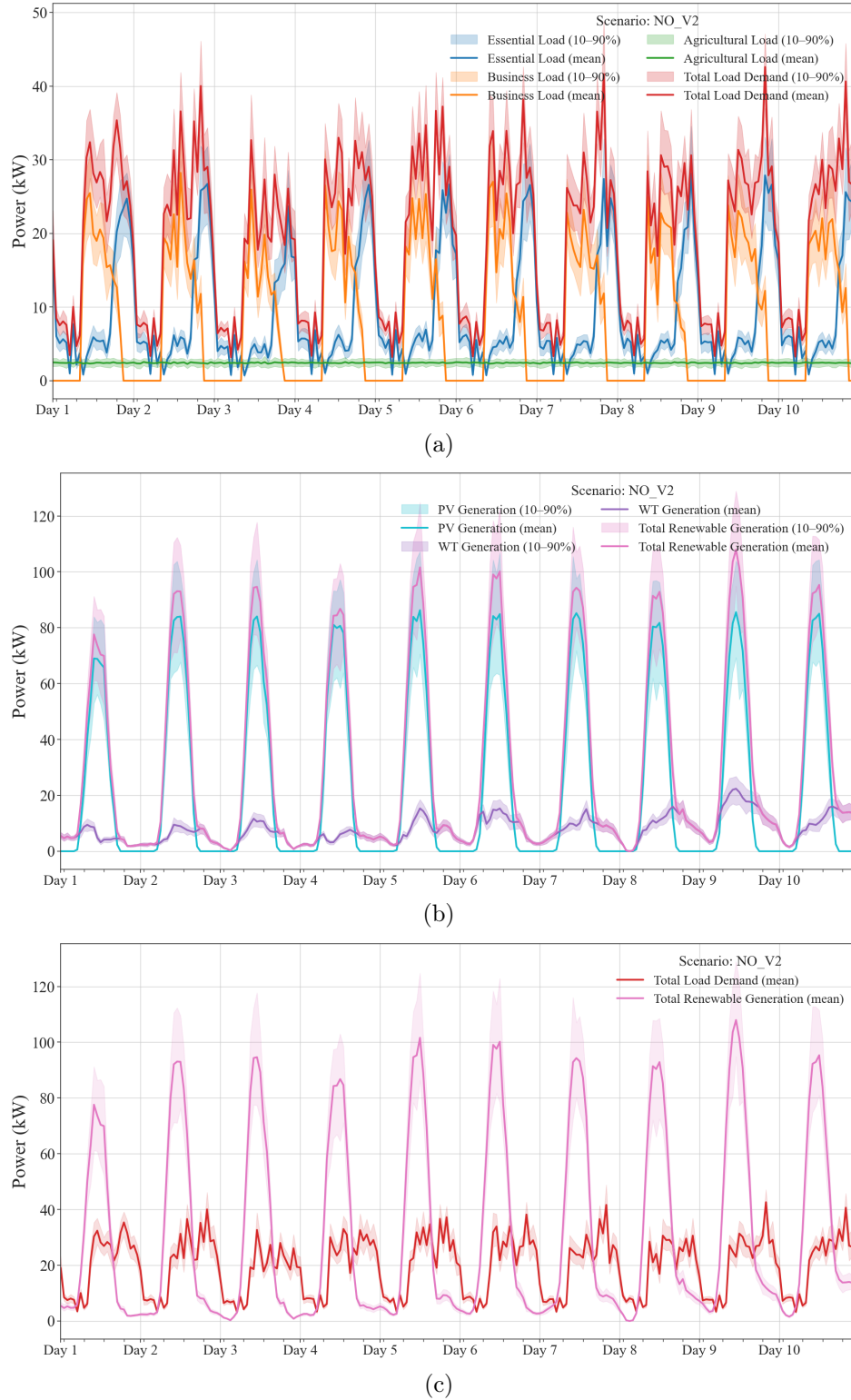


Figure 3.20: NO_V2 scenario (normal-historical) — augmented datasets used in training: (a) loads and total demand, (b) renewable generation, and (c) total load versus total renewables. Ribbons show the 10–90% range; solid lines indicate the mean values.

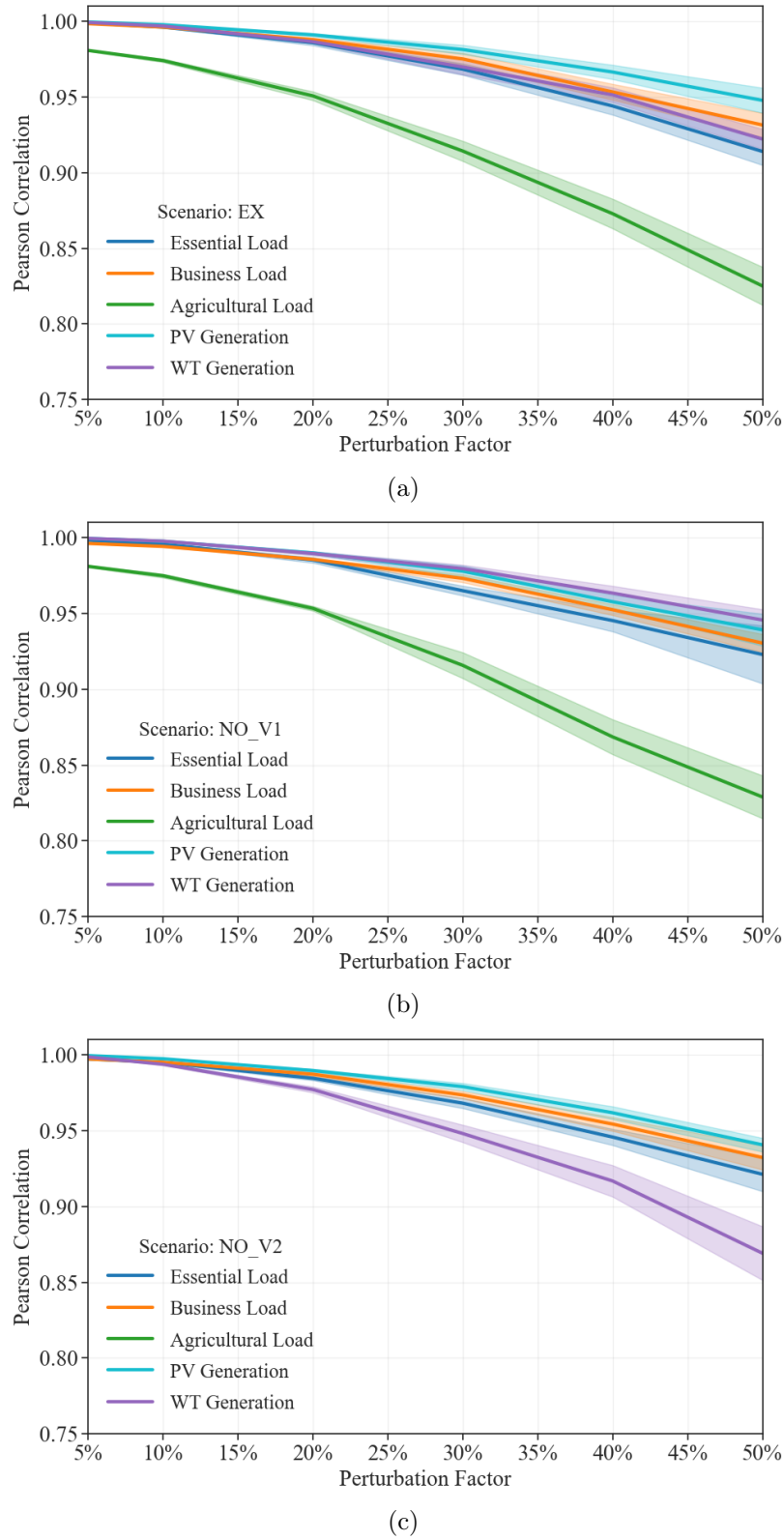


Figure 3.21: Pearson correlation between perturbed and original time series across perturbation factors for (a) EX, (b) NO_V1, and (c) NO_V2 scenarios. Shaded bands denote one standard deviation across 10 perturbed variants, indicating how similarity to the original data decreases as perturbation strength increases.

3.5.2 Post-Training Uncertainty Injection with Perlin Noise

To evaluate robustness under realistic forecast errors, post-training input uncertainty is injected using Perlin noise. Perlin noise is a coherent, smoothly varying random field; unlike white noise, which is independent and identically distributed (i.i.d.) and produces abrupt point-to-point jumps, Perlin noise varies gradually and exhibits short-range correlation, making it suitable for temporal perturbations of renewable profiles where neighbouring hours are physically related (cloud motion, wind gust build-up/decay).

Only renewable inputs (solar and wind) are perturbed during evaluation, while loads remain unchanged. Perturbations are multiplicative, so zeros remain zeros (e.g., night-time PV stays zero), diurnal shape is preserved up to a smooth gain, and negative anomalies are avoided via a non-negativity clamp. Because Perlin noise is low-pass relative to white noise, slow trends such as night-time wind plateaus and multi-day depressions during cyclone periods are preserved and only their amplitudes are modulated.

One-dimensional value-noise Perlin is employed with the standard smoothstep polynomial

$$\text{fade}(t) = 6t^5 - 15t^4 + 10t^3, \quad (3.103)$$

which produces smooth transitions between random lattice values, maintaining continuous gradients throughout. Let $P_s(t)$ and $P_w(t) \in [-1, 1]$ denote the normalised Perlin signals for solar and wind at time index t . For an original series $x(t)$ (solar or wind), the perturbed series $\tilde{x}(t)$ is

$$\tilde{x}(t) = x(t) \max\left(0, 1 + \alpha P(t)\right), \quad (3.104)$$

where $\alpha \in [0, 1)$ is the noise strength. The non-negativity ensures physical realism in low-resource conditions. After perturbing both sources, the aggregate renewable column is re-

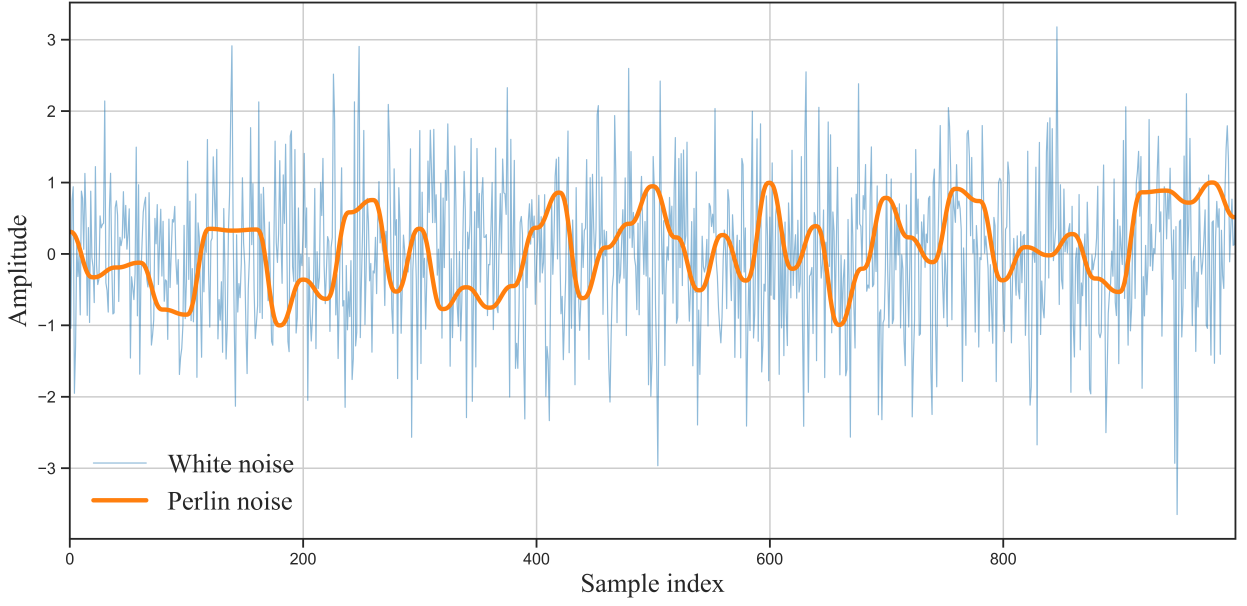


Figure 3.22: Illustration of white noise (i.i.d.) versus Perlin noise (correlated). Perlin maintains smooth, temporally coherent variations.

computed as

$$Sum_Extreme(t) \leftarrow \tilde{Solar}(t) + \tilde{Wind}(t).$$

Figure 3.22 contrasts a white-noise signal with a Perlin-noise signal. The latter exhibits smooth, wave-like fluctuations that better reflect realistic weather variability.

Algorithm 3 summarises the evaluation time, noise injection. Independent seeds are used for solar and wind; a “scale” parameter controls the lattice density (correlation length): larger scales yield slower fluctuations across hours/days, smaller scales yield quicker variations.

Algorithm 3 Perlin-Noise-Based Uncertainty Injection for Post-Training Evaluation

Require: Time series $\mathbf{x}_s(t)$ (solar), $\mathbf{x}_w(t)$ (wind); noise scales s_s, s_w ; scaling factors α_s, α_w ; random seeds (z_s, z_w)

Ensure: Perturbed dataset $\tilde{\mathbf{x}}_s(t), \tilde{\mathbf{x}}_w(t)$ and recomputed $Sum_Extreme(t)$

1: Generate one-dimensional Perlin noise vectors:

$$P_s(t) = \text{PERLIN1D}(t; s_s, z_s), \quad P_w(t) = \text{PERLIN1D}(t; s_w, z_w)$$

where each uses a random gradient lattice and the smooth fade function from Equation 3.103.

2: **for** each time step t in dataset **do**

3: Compute gain factors using the perturbation law (Equation 3.104):

$$g_s(t) = \max(0, 1 + \alpha_s P_s(t)), \quad g_w(t) = \max(0, 1 + \alpha_w P_w(t))$$

4: Apply multiplicative perturbation:

$$\tilde{x}_s(t) = x_s(t) g_s(t), \quad \tilde{x}_w(t) = x_w(t) g_w(t)$$

5: **end for**

6: Recompute the aggregate renewable profile:

$$Sum_Extreme(t) = \tilde{x}_s(t) + \tilde{x}_w(t)$$

7: **return** perturbed dataset with solar, wind, and updated total generation

The multiplicative form in Equation 3.104, combined with the smooth correlation of Perlin noise, ensures that night-time PV remains at or near zero, typical diurnal patterns are preserved, and cyclone-period wind regimes such as sustained ramps are maintained but only modulated in amplitude, avoiding unphysical spikes.

A Monte-Carlo ensemble is formed by drawing independent seed pairs (z_s, z_w) and re-evaluating the trained policy on a fixed 10-day window (e.g., the Laila cyclone period). For each run, resilience and ageing indicators are recorded, along with trajectory-level statistics (SOC mean/std, renewable mean/std). Distributions, regression trends, and ribbon plots (e.g., 5–95% bands) are then used to quantify robustness and sensitivity under uncertainty. The detailed results and analysis of this Monte Carlo evaluation are presented later in the

results section for the XDRL framework.

3.6 Computational Cost Evaluation Protocol

All computational timing experiments were conducted offline. The mixed-integer controller was executed on an Intel Core i7-6700HQ laptop (four physical cores, 3.5 GHz maximum turbo frequency).¹ The deep-learning controller was executed on an Intel Core i7-14700 desktop (sixteen physical cores, 5.3 GHz maximum turbo frequency).

Only the core decision routines were benchmarked: the Gurobi `solve()` call for the MPC formulation and the neural-network forward pass `model.predict()` for the DRL controller. Each evaluation episode contained $N_s = 216$ hourly steps, representing nine days of rolling operation.

Let $t_{r,i}$ denote the wall-clock time of step i in run r , and let R represent the total number of runs. The mean and sample standard deviation of the step-level computation time are given by:

$$t_1 = \frac{1}{RN_s} \sum_{r=1}^R \sum_{i=1}^{N_s} t_{r,i}, \quad \sigma_1 = \sqrt{\frac{1}{RN_s - 1} \sum_{r=1}^R \sum_{i=1}^{N_s} (t_{r,i} - t_1)^2}. \quad (3.105)$$

The total episode time is obtained by summing the step times across all steps and runs:

$$T_1 = \frac{1}{R} \sum_{r=1}^R \sum_{i=1}^{N_s} t_{r,i}, \quad \Sigma_1 = \sqrt{\frac{1}{R - 1} \sum_{r=1}^R \left(\sum_{i=1}^{N_s} t_{r,i} - T_1 \right)^2}. \quad (3.106)$$

To estimate the cumulative computational effort over a fifteen-year operational hori-

¹The academic licence of Gurobi used in this study is node-locked and may be installed on only one machine; therefore, a dedicated laptop was selected as the solver host.

zon, the total Floating-Point Operations (FLOPs) are calculated as:

$$C = T_1 \left(\frac{24 \times 365}{N_s} \right) \times 15 \times F_{\text{peak}}, \quad (3.107)$$

where F_{peak} is the theoretical FP32 peak performance of the host CPU, estimated at run time as $F_{\text{peak}} = f_{\text{max}} n_{\text{cores}} \times 16 \text{ FLOPs}^{-1}$, with f_{max} denoting the maximum reported frequency. All timing tests were repeated $R = 5$ times.

3.7 Summary of Methodology and Transition to Results

This chapter set out the methodological backbone of the study. First, HOMER Pro provides the techno-economic component sizing of the microgrid. Second, an MPC baseline is constructed by reformulating system components as a MILP and evaluated on cyclone and normal periods; results are compared against the HOMER planning outputs. Third, a Gymnasium compatible environment is built for DRL; the agent is trained on the augmented scenarios (EX, NO_V1, NO_V2) and evaluated against the MPC baseline. Finally, explainability methods are used to analyse DRL decisions, and a Monte Carlo study with Perlin-noise uncertainty quantifies robustness of key metrics. The overall workflow is summarised in Figure 3.23. The next two chapters apply this framework to the case study, report quantitative results under cyclone and normal conditions, and discuss the implications of the DRL decisions and robustness findings in detail.

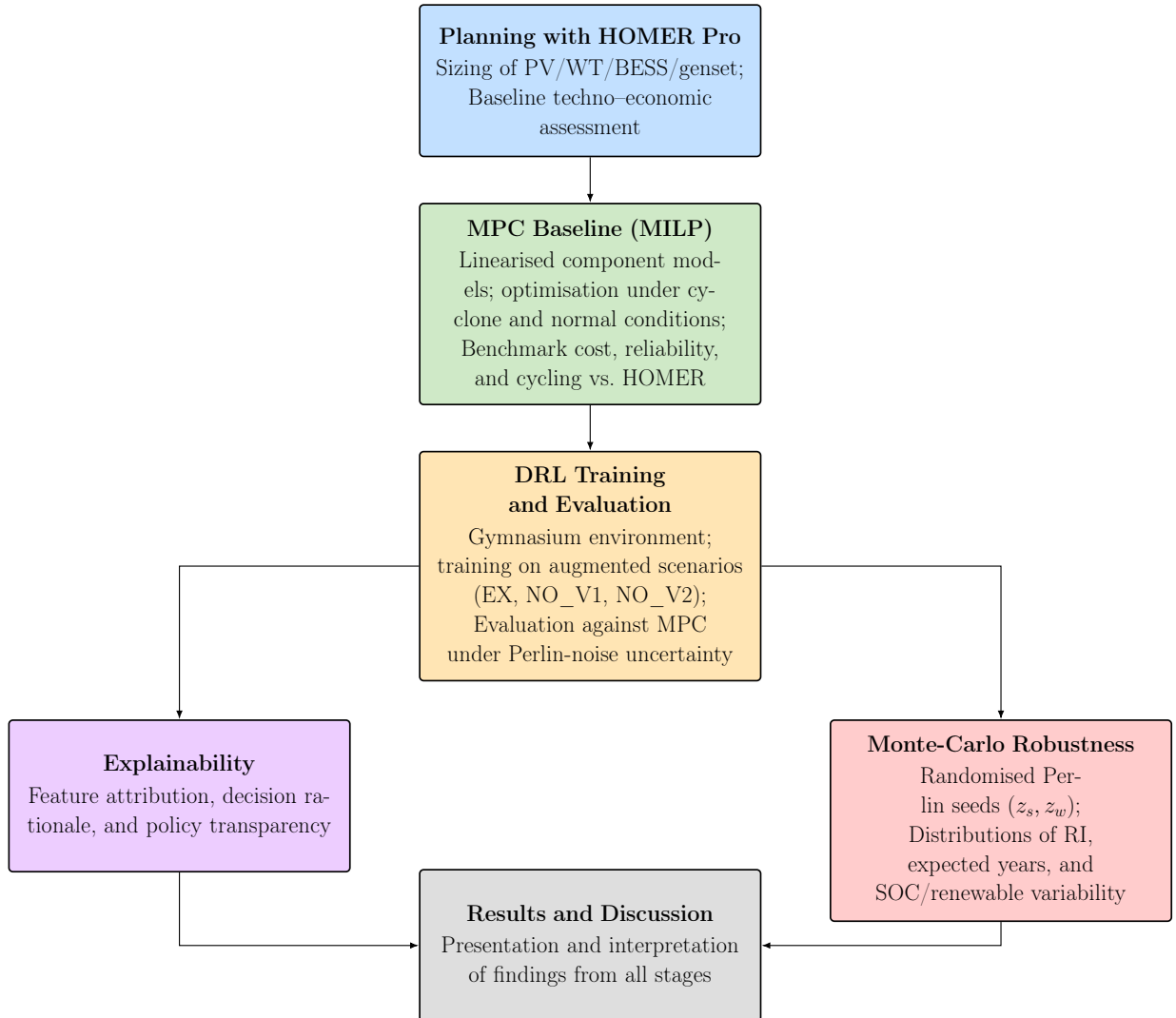


Figure 3.23: Methodological flow: HOMER-based planning, MILP-formulated MPC baseline, DRL training and evaluation under data augmentation, post-training explainability and Monte-Carlo robustness analyses, and final results and discussion.

Chapter 4

Model Predictive Control Framework

4.1 Introduction

Chapter 3 established the overall modelling framework (Section 3.3), including the system architecture (Section 3.3.1), case study site in India, and the Cyclone Laila dataset (Section 3.3.2). Historical weather and demand data were processed to generate consistent inputs for both planning and operation, ensuring realistic and stress-tested simulations.

In the planning stage, system components were sized using HOMER Pro (Section 3.4.1). The selected configuration (Table 3.1), annual SOC trajectory (Figure 3.10), and renewable generation profiles (Figure 3.9) confirmed technical feasibility. A sensitivity analysis on key economic parameters (Figures 3.11a–3.12b) evaluated the robustness of cost performance.

Building on these results, the operational strategy was formulated as an MILP problem in PYOMO (Section 3.4.2). The formulation enforced power balance, converter limits, and battery dynamics while jointly optimising resilience and battery lifetime (Equations 3.27–3.53). The smart resilient control flow in Figure 3.15 integrates data updates, resilience assessment, policy selection, and horizon shifting within the MPC loop.

This chapter now presents and discusses the operational results obtained from that modelling framework under a range of scenarios.

4.2 MPC Operational Results and Discussion

The performance of the proposed MILP-MPC controller is evaluated in four steps. First, the model is validated against HOMER Pro. Second, a structured set of scenarios is used to study sensitivity to objective-function weights. Third, the impacts of DR and reduced battery capacity under the smart resilient controller are examined. Finally, grid-connected cases are analysed to assess the benefits of controlled power exchange and profit-seeking operation.

4.2.1 Model Validation Against HOMER Pro

The MPC formulation was first validated against baseline results from HOMER Pro over a 10-day period (1-10 May), using identical demand and resource conditions. Figures 4.1a and 4.1b show total generation, load, and SOC trajectories.

Both methods successfully meet the total demand, but their operational behaviours differ considerably. The HOMER Pro schedule, which performs daily optimisation without foresight, charges the battery to full capacity whenever surplus renewable generation becomes available. This reactive behaviour often leads to frequent charging and discharging cycles that accelerate battery wear. By contrast, the MPC model employs a rolling prediction horizon that anticipates upcoming variations in demand and renewable supply. As shown in Figure 4.1b, this foresight allows the controller to moderate battery charging during periods of surplus generation, reserving capacity for later demand peaks. For instance, around 3–5 May, the MPC strategy deliberately maintains the SOC at an intermediate level rather than charging to 100%, preventing redundant cycling and reducing stress on the battery.

Two MPC variants are shown: MPC_SOC_Limited (SOC constrained to 20-90%)

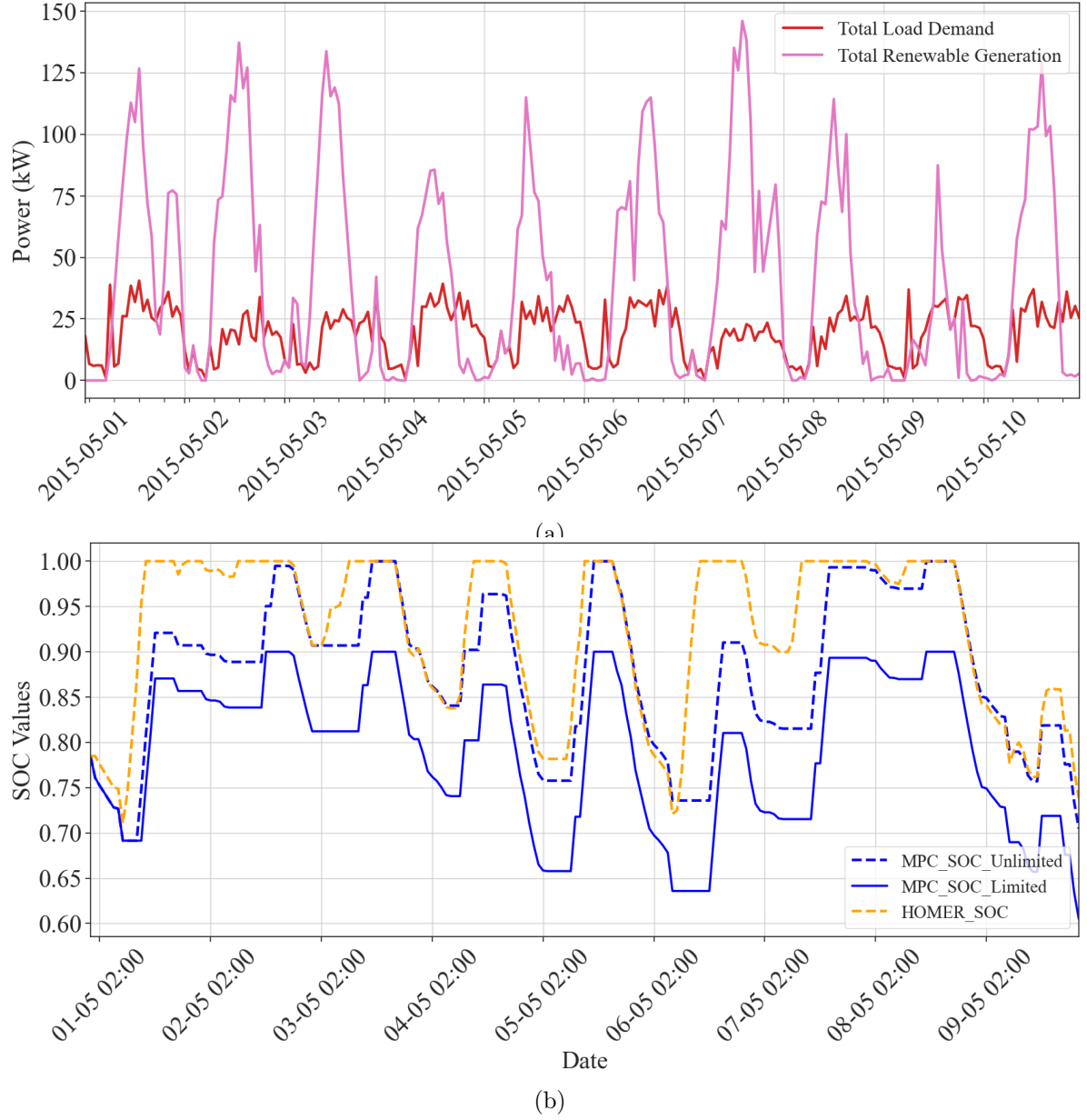


Figure 4.1: Validation of the proposed resilient operational method through comparison with HOMER PRO over a 10-day horizon (1–10 May): (a) total generation and load, (b) battery SOC profiles.

and MPC_SOC_Unlimited (full capacity accessible). Both maintain full load supply, indicating that predictive control sustains reliability even under conservative SOC limits. The limited case further protects against overcharge and deep discharge.

4.2.2 Scenario Design and Classification

To systematically assess the MPC framework, twenty-eight scenarios were designed (Table 4.1). The first fifteen scenarios focus on identifying the optimal weighting configuration of the four main objectives: resilience (ω_R), total power imbalance (ω_t), battery switching cost (ω_{bat}), and battery life cycle (ω_{BLC}). Scenarios 2–5 assess the effect of activating only one objective at a time, while Scenarios 6–11 combine two objectives to study their mutual interactions. Scenarios 12–15 examine combinations of three objectives, and Scenario 1 considers all four simultaneously with equal weighting, serving as the reference configuration for subsequent comparisons.

Scenario 16 introduces limited DR participation, assuming that essential and business loads can shift up to 10% of their demand. This represents a practical case in which critical loads participate only minimally in DR programmes, while deferred loads remain fully flexible. Scenario 17 employs the smart resilient controller proposed earlier (Figure 3.15) to evaluate its performance under identical conditions.

Scenarios 18–24 apply the smart controller under reduced battery capacities (75%, 50%, and 25% of the design value), testing robustness under storage limits. Scenarios 25–28 analyse grid-connected operation with and without an explicit profit term, both for full and 25% battery capacity, to quantify how grid support and energy trading influence resilience and battery ageing.

4.2.3 Sensitivity to Objective Weighting

Scenarios 1–15 clarify how the four objectives shape operation. By adjusting the relative importance of these weights, the controller’s ability to balance resilience and battery longevity can be systematically assessed.

Table 4.1: Simulation Scenarios and Results.

Scenario	Weighting Factors	TS	Essential Loss	Business Loss	Deferred Loss	Total Loss	RI	EY	DIS number
1	$\omega_R = 0.25, \omega_t = 0.25, \omega_{\text{bat}} = 0.25, \omega_{\text{BLC}} = 0.25$	7	$\simeq 0$	0.3978	$\simeq 0$	0.3979	0.9999	14.425	97
2	$\omega_R = 1.0, \omega_t = 0.0, \omega_{\text{bat}} = 0.0, \omega_{\text{BLC}} = 0.0$	45	1.7218	30.566	13.2381	45.526	0.9948	10.810	102
3	$\omega_R = 0.0, \omega_t = 0.0, \omega_{\text{bat}} = 1.0, \omega_{\text{BLC}} = 0.0$	0	1323.0564	973.1779	369.1734	2567.4605	0.36759	137.63	15
4	$\omega_R = 0.0, \omega_t = 0.0, \omega_{\text{bat}} = 0.0, \omega_{\text{BLC}} = 1.0$	0	1177.5799	987.3639	402.5166	2665.4079	0.3106	∞	0
5	$\omega_R = 0.0, \omega_t = 1.0, \omega_{\text{bat}} = 0.0, \omega_{\text{BLC}} = 0.0$	21	344.4674	642.2060	167.8735	1154.5471	0.7699	9.6637	120
6	$\omega_R = 0.0, \omega_t = 0.5, \omega_{\text{bat}} = 0.5, \omega_{\text{BLC}} = 0.0$	10	1045.7357	690.1659	90.2695	1826.1712	0.4765	32.0781	20
7	$\omega_R = 0.5, \omega_t = 0.0, \omega_{\text{bat}} = 0.5, \omega_{\text{BLC}} = 0.0$	0	487.5285	367.4426	145.1345	1000.1058	0.7443	∞	0
8	$\omega_R = 0.0, \omega_t = 0.0, \omega_{\text{bat}} = 0.5, \omega_{\text{BLC}} = 0.5$	0	1311.9293	774.7107	405.9348	2492.5748	0.33671	82.5825	25
9	$\omega_R = 0.5, \omega_t = 0.5, \omega_{\text{bat}} = 0.0, \omega_{\text{BLC}} = 0.0$	24	$\simeq 0$	$\simeq 0$	0.0003	0.000415	0.9999	9.6665	120
10	$\omega_R = 0.0, \omega_t = 0.5, \omega_{\text{bat}} = 0.0, \omega_{\text{BLC}} = 0.5$	22	721.116	910.9397	128.3236	1760.3795	0.5832	13.368	105
11	$\omega_R = 0.5, \omega_t = 0.0, \omega_{\text{bat}} = 0.0, \omega_{\text{BLC}} = 0.5$	0	353.0940	320.5975	183.061	856.7527	0.8037	89.3644	17
12	$\omega_R = 0.33, \omega_t = 0.0, \omega_{\text{bat}} = 0.33, \omega_{\text{BLC}} = 0.33$	0	318.0040	321.1475	204.5931	843.7447	0.81701	85.5490	20
13	$\omega_R = 0.33, \omega_t = 0.33, \omega_{\text{bat}} = 0.0, \omega_{\text{BLC}} = 0.33$	22	$\simeq 0$	0.1010	0.5966	0.6977	0.9999	13.3286	106
14	$\omega_R = 0.33, \omega_t = 0.33, \omega_{\text{bat}} = 0.33, \omega_{\text{BLC}} = 0.0$	11	295.2484	195.3237	50.9615	541.5337	0.85063	30.7044	21
15	$\omega_R = 0.0, \omega_t = 0.33, \omega_{\text{bat}} = 0.33, \omega_{\text{BLC}} = 0.33$	7	626.5864	685.5575	109.3016	1421.4456	0.6506	14.5358	98
16 (DR = 10%)	$\omega_R = 25, \omega_t = 0.25, \omega_{\text{bat}} = 0.25, \omega_{\text{BLC}} = 0.25$	8	0.0001	$\simeq 0$	0.3305	0.3308	0.9999	15.5806	90
17 (Smart SOC)	$\omega_R = 25, \omega_t = 0.25, \omega_{\text{bat}} = 0.25, \omega_{\text{BLC}} = 0.25$	7	$\simeq 0$	0.3978	0.0004	0.3983	0.9999	14.3965	98
18 (75%)	$\omega_R = 25, \omega_t = 0.25, \omega_{\text{bat}} = 0.25, \omega_{\text{BLC}} = 0.25$	6	0.0001	1.4404	7.3999	8.8405	0.9993	11.4466	102
19 (50%)	$\omega_R = 25, \omega_t = 0.25, \omega_{\text{bat}} = 0.25, \omega_{\text{BLC}} = 0.25$	6	24.4823	67.7473	12.8009	105.0305	0.9809	9.3103	96
20 (Soc-50%)	$\omega_R = 25, \omega_t = 0.25, \omega_{\text{bat}} = 0.25, \omega_{\text{BLC}} = 0.25$	6	21.6323	40.7023	115.3986	77.7333	0.9852	8.9895	98
21 (Soc-w-50%)	$\omega_R = 25, \omega_t = 0.25, \omega_{\text{bat}} = 0.25, \omega_{\text{BLC}} = 0.25$	8	17.7714	68.8945	18.0915	104.7576	0.9833	9.0702	96
22 (25%)	$\omega_R = 25, \omega_t = 0.25, \omega_{\text{bat}} = 0.25, \omega_{\text{BLC}} = 0.25$	6	87.8225	140.9595	36.2263	265.0084	0.9444	8.5126	71
23 (SOC-25%)	$\omega_R = 25, \omega_t = 0.25, \omega_{\text{bat}} = 0.25, \omega_{\text{BLC}} = 0.25$	6	60.9311	116.2264	28.4003	205.5578	0.9590	7.3121	82
24 (SOC-w-25%)	$\omega_R = 25, \omega_t = 0.25, \omega_{\text{bat}} = 0.25, \omega_{\text{BLC}} = 0.25$	9	35.5838	182.2255	38.5804	256.3898	0.9611	7.4870	81
25 (Grid)	$\omega_R = 0.25, \omega_t = 0.25, \omega_{\text{bat}} = 0.25, \omega_{\text{BLC}} = 0.25$	0	$\simeq 0$	$\simeq 0$	$\simeq 0$	$\simeq 0$	$\simeq 1$	$\simeq \infty$	7
26 (Grid-Profit)	$\omega_R = 0.25, \omega_t = 0.25, \omega_{\text{bat}} = 0.25, \omega_{\text{BLC}} = 0.25$	10	$\simeq 0$	$\simeq 0$	$\simeq 0$	$\simeq 0$	$\simeq 1$	27.07	58
27 (Grid-25%)	$\omega_R = 0.25, \omega_t = 0.25, \omega_{\text{bat}} = 0.25, \omega_{\text{BLC}} = 0.25$	0	$\simeq 0$	$\simeq 0$	$\simeq 0$	$\simeq 0$	$\simeq 1$	$\simeq \infty$	7
28 (Grid-Profit-25%)	$\omega_R = 0.25, \omega_t = 0.25, \omega_{\text{bat}} = 0.25, \omega_{\text{BLC}} = 0.25$	11	$\simeq 0$	$\simeq 0$	$\simeq 0$	$\simeq 0$	$\simeq 1$	11.18	59

Note: ω_R : Resiliency Weighting Factor in Objective Function, ω_t : Total Imbalance Weighting Factor in Objective Function, ω_{bat} : Battery Switching Cost Weighting Factor in Objective Function, ω_{BLC} : Battery Life Cycle Weighting Factor in Objective Function, DR: Demand Response, TS: Total Switches, RI: Resilience Index, EY: Expected Years.

In Scenarios 2–5, each objective was activated individually to isolate its contribution to system performance. When only the resiliency objective was considered (Scenario 2), the microgrid achieved a high Resilience Index ($RI \simeq 0.995$) with minimal power loss; however, the frequent switching between charging and discharging increased the total number of transitions. In contrast, focusing solely on the switching cost (Scenario 3) produced almost no switching events but resulted in severe load losses and an unrealistic expected battery lifetime exceeding 130 years, as the battery remained largely idle. When the battery life-cycle objective was prioritised (Scenario 4), the system again avoided discharging to preserve the battery, which caused the worst resiliency performance among all single-objective cases. Considering only the total imbalance term (Scenario 5) improved RI relative to Scenarios 3 and 4, but excessive charge–discharge cycles led to premature degradation of the battery.

Scenarios 6–11 combined two objectives to capture possible interactions. The configuration combining resiliency and total imbalance (Scenario 9) produced the best load-supply performance among dual-objective cases ($RI \simeq 1$), although it also exhibited the highest number of battery discharges and switches due to the absence of explicit degradation control. By contrast, the combination of resiliency and battery life-cycle objectives (Scenario 11) achieved the longest expected battery life (89 years) while still maintaining a strong resiliency level ($RI \simeq 0.80$), indicating a more balanced outcome between supply continuity and asset preservation.

Scenarios 12–15 explored combinations of three objectives. Scenario 13, which included resiliency, total imbalance, and battery life-cycle terms, achieved a good compromise, maintaining full supply for most loads while extending battery life to 13 years. Scenario 14, where the battery switching cost replaced the life-cycle term, yielded slightly lower resiliency but significantly reduced switching frequency. Overall, the inclusion of either battery-related term, switching cost or life-cycle degradation, proved essential to control excessive cycling and extend the expected lifespan.

Among all configurations, Scenario 1, which assigned equal weights to all four objectives, delivered the most balanced performance. It achieved an almost perfect Resilience Index (0.9999) while maintaining an expected battery life of 14.4 years, close to the planning-stage value of 15 years. The number of discharges and switching events was also moderate, reflecting an optimal trade-off between resilience and battery preservation.

Figure 4.2 illustrates the ERI for Scenarios 1–5, corresponding to different combinations of objective weights. The curve labelled `RI_values_all` represents Scenario 1, where all objectives are included with equal weighting. The remaining curves isolate individual objectives: `RI_values_R` corresponds to resiliency-only optimisation (Scenario 2), `RI_values_bat` to switching cost only (Scenario 3), `RI_values_BLC` to battery life-cycle only (Scenario 4), and `RI_values_t` to total imbalance only (Scenario 5). A higher and more stable ERI curve indicates stronger and more consistent resilience performance over time. The analysis confirms that equal or near-equal weighting across objectives leads to the most effective operational strategy. Excessive emphasis on a single term, whether resiliency, switching, or degradation, causes either overuse or underuse of the battery, compromising one of the two key goals: resilience or longevity. Hence, Scenario 1 is adopted as the baseline configuration for the subsequent analyses.

4.2.4 Overall Optimisation Performance

Figure 4.3 compares the battery SOC trajectories for islanded operation under two configurations: Scenario 9 and Scenario 1. When only resiliency and total imbalance are considered (Scenario 9), the controller reacts aggressively to short-term fluctuations, charging and discharging almost every hour with minimal idle periods. This behaviour increases cycling frequency and accelerates battery wear, leading to a shorter expected lifetime. In contrast, the SOC profile is smoother, with intentional idle periods and moderated swings, reflecting

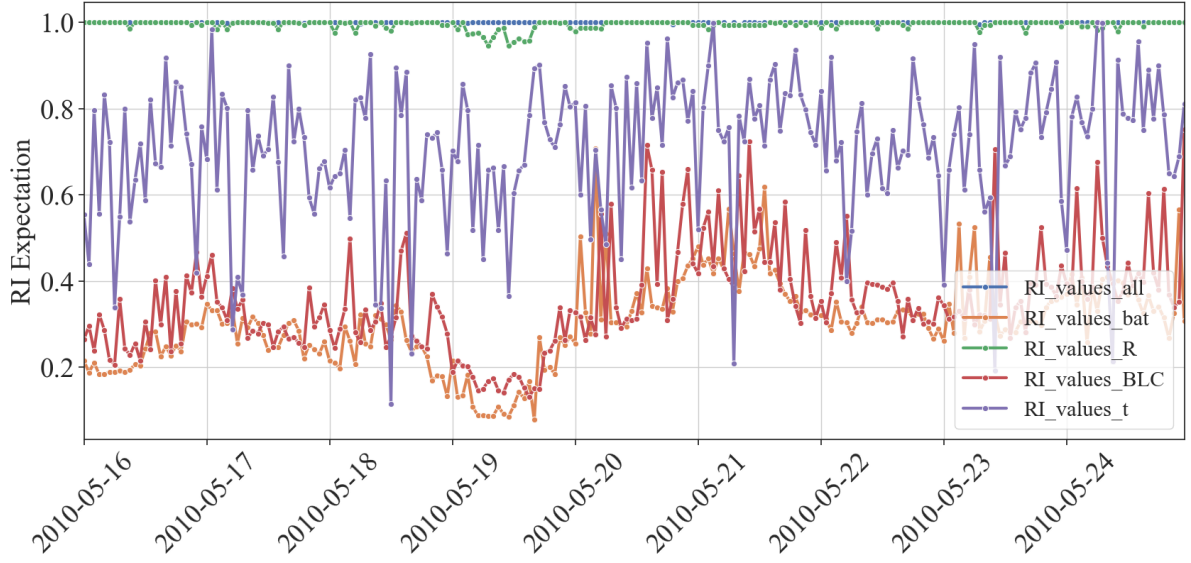


Figure 4.2: Expected Resiliency Index under different objective-weighting configurations.

the influence of degradation-aware terms in Scenario 1.

Figure 4.4 illustrates the SOC trajectories for the grid-connected scenarios. In Scenario 25, where grid interaction is allowed but profit is not included in the objective function, the controller primarily focuses on utilising the available storage capacity. As a result, the battery is discharged steadily to support local demand without strategic consideration of energy pricing or trade-offs, reflecting a behaviour similar to an energy-balancing mode. When profit maximisation is introduced (Scenario 26), the controller shifts towards a more selective strategy: charging during periods of low grid prices and discharging when selling is more beneficial. This reduces deep discharges, increases idle periods, and results in an extended battery lifespan of approximately 27 years. The difference between these two profiles demonstrates that the inclusion of an economic objective leads to more sustainable and cost-effective operation.

To verify the normalisation of the multi-objective formulation, the values of each objective term in Scenario 1 are shown in Figure 4.5. The figure confirms that all four objectives are appropriately scaled within the range of 0 to 1. This ensures that no single

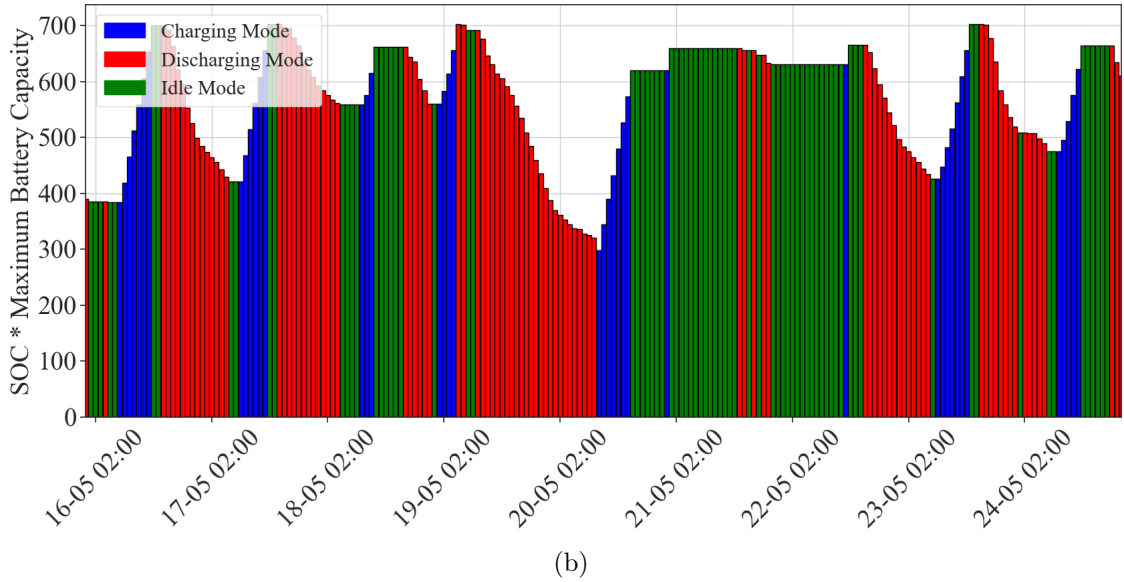
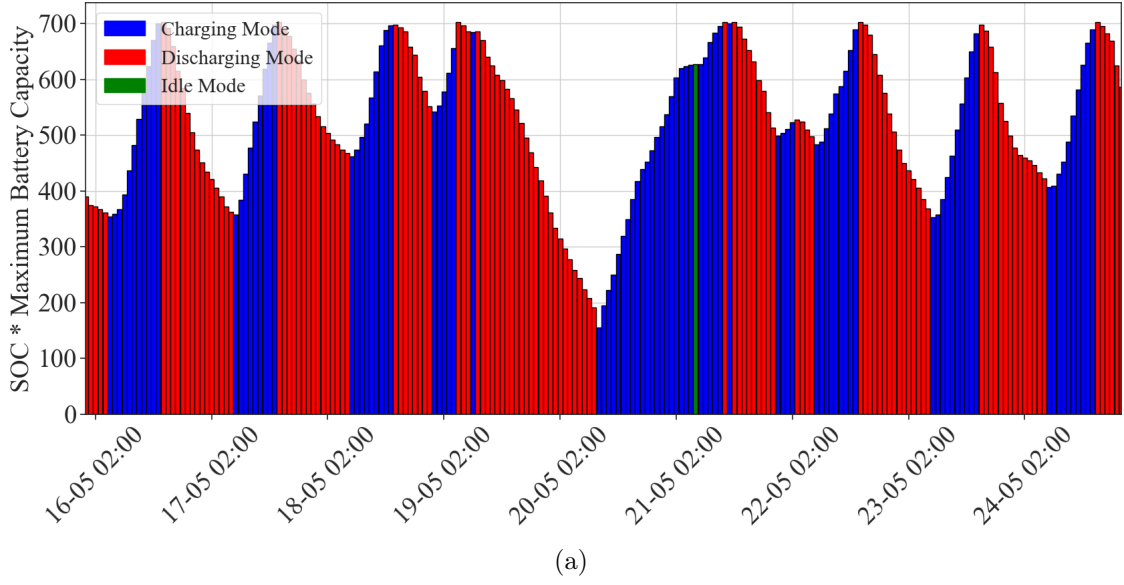


Figure 4.3: Battery SOC under islanded operation: (a) resiliency and total imbalance only (Scenario 9), (b) all objectives equally weighted (Scenario 1).

component dominates the optimisation process. When comparing the SOC behaviour in Figure 4.3b with the normalised objective values, it can be observed that periods of battery idling (for instance, between Days 21 and 22) coincide with increased values of the BLC term. This correlation indicates that the optimisation successfully balances resilience and longevity by reducing stress on the battery whenever possible.

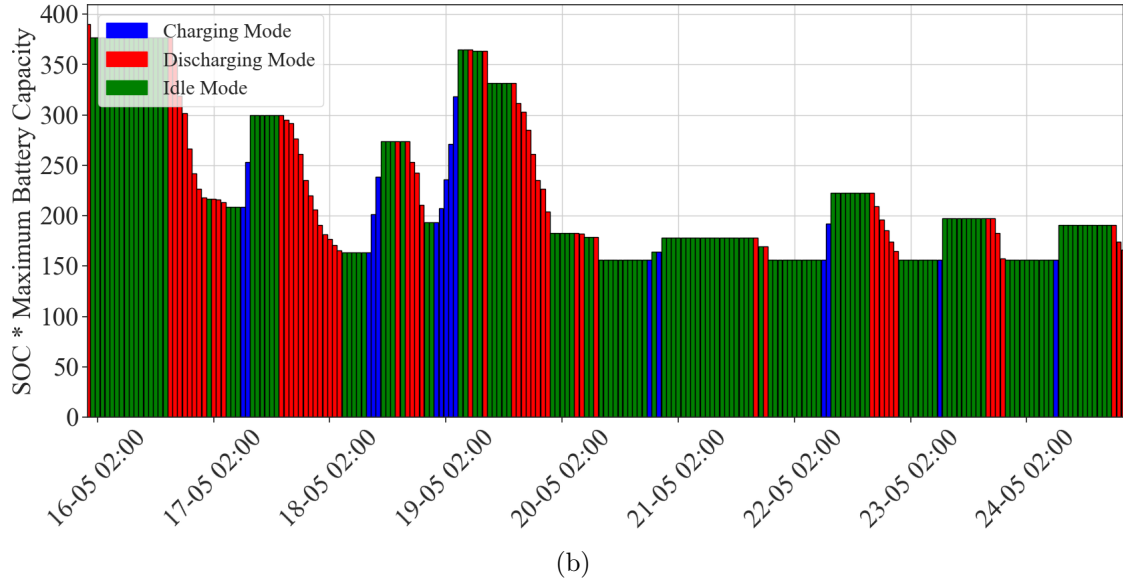
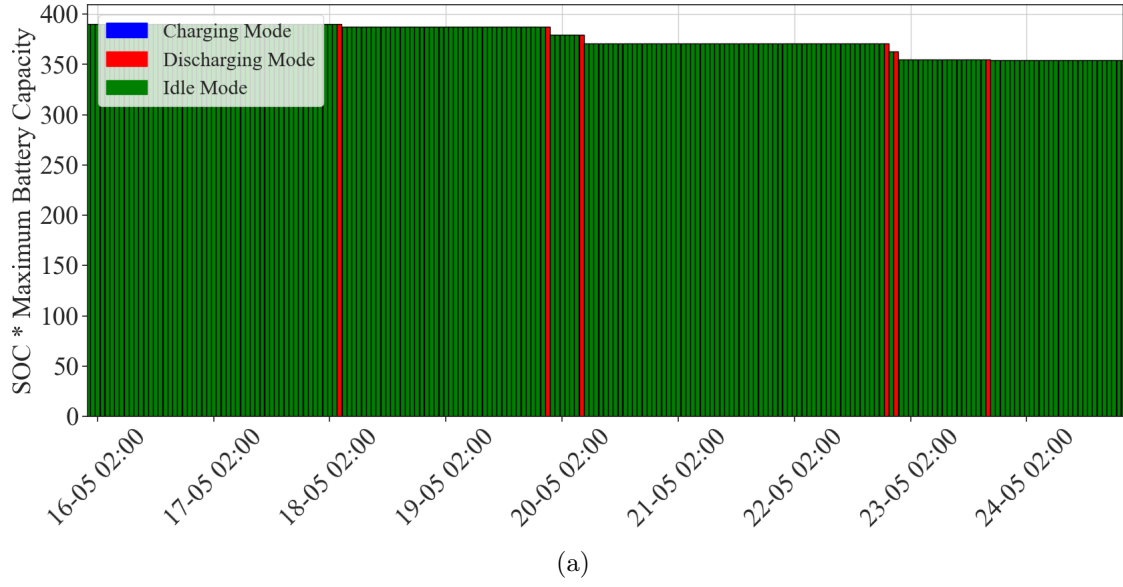


Figure 4.4: Battery SOC under grid-connected operation: (a) without profit consideration (Scenario 25), (b) with profit optimisation (Scenario 26).

Figures 4.6 and 4.7 show that the imbalance across all load categories is almost negligible. Essential and deferred (agricultural) loads are fully supplied throughout the operation, and the business load exhibits only minimal fluctuations that are practically insignificant.

4.2.5 Sensitivity of MPC Results to BLC Linearisation Accuracy

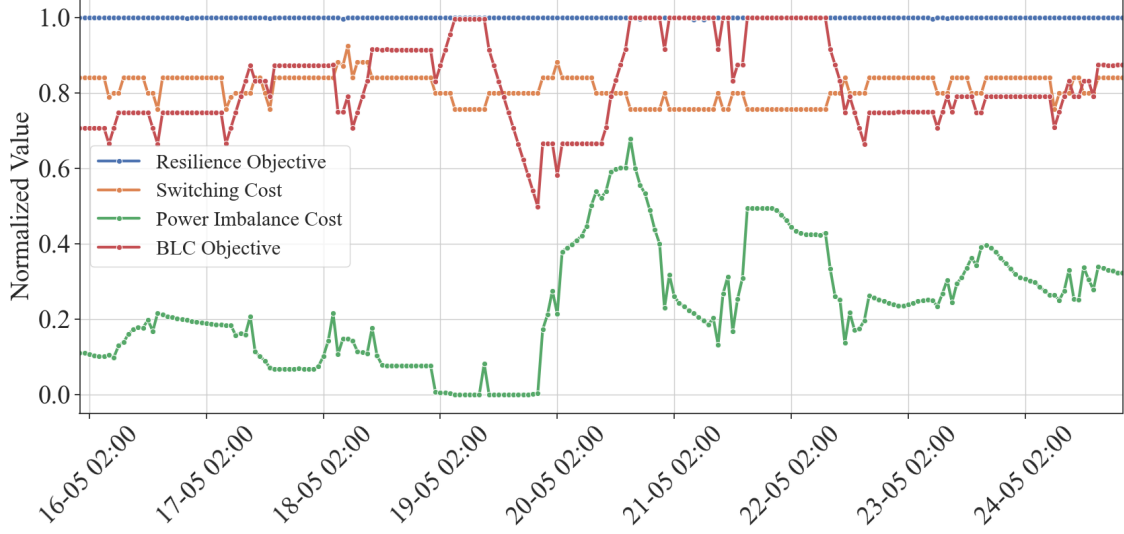


Figure 4.5: Normalised values of the four objective-function terms in Scenario 1, demonstrating proper scaling and balanced optimisation.

An additional sensitivity test was carried out to examine how the BLC–DOD piecewise-linear (PWL) approximation affects the MPC baseline. The same MPC formulation, prediction horizon, operating constraints, initial SOC, and objective-weighting configuration were used in all cases. Therefore, the only change between the tests was the PWL representation of the BLC–DOD curve.

The objective-weighting configuration was selected based on the weighting sensitivity analysis discussed in Section 4.2.3. This allows the PWL sensitivity test to focus specifically on how degradation-model fidelity affects the resulting resilience–battery-life trade-off, rather than changing the optimisation priorities.

The approximation error was measured using the mean relative area error over the full DOD range:

$$E_{\text{area}} = \frac{1}{D_{\text{max}} - D_{\text{min}}} \int_{D_{\text{min}}}^{D_{\text{max}}} \left| \frac{BLC_{\text{PWL}}(D) - BLC_{\text{exact}}(D)}{BLC_{\text{exact}}(D)} \right| \times 100 dD. \quad (4.1)$$

This metric was used because the BLC–DOD curve is highly nonlinear at low DOD, and a

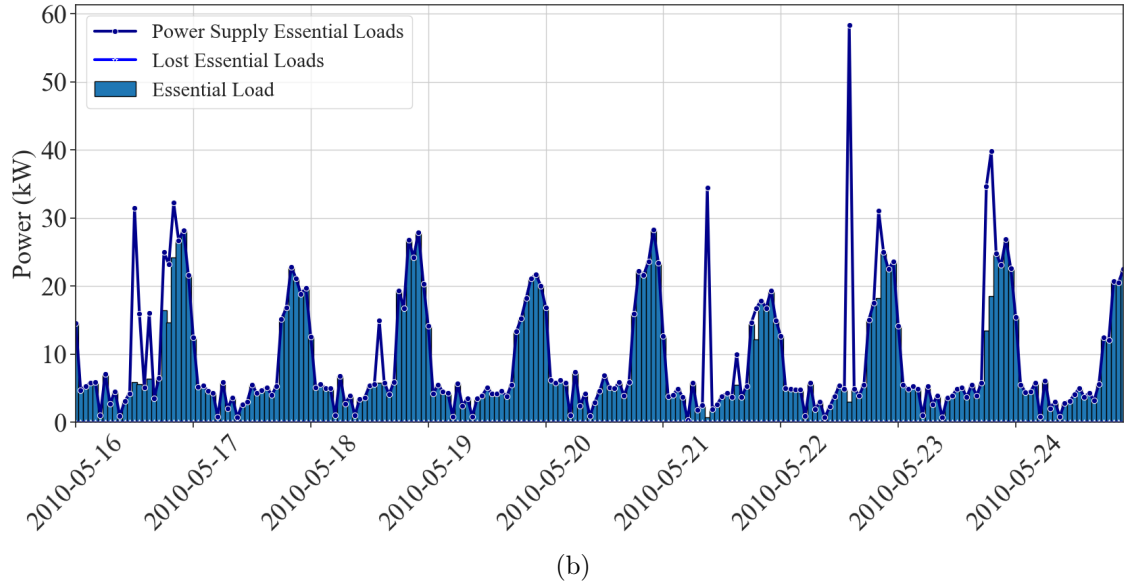
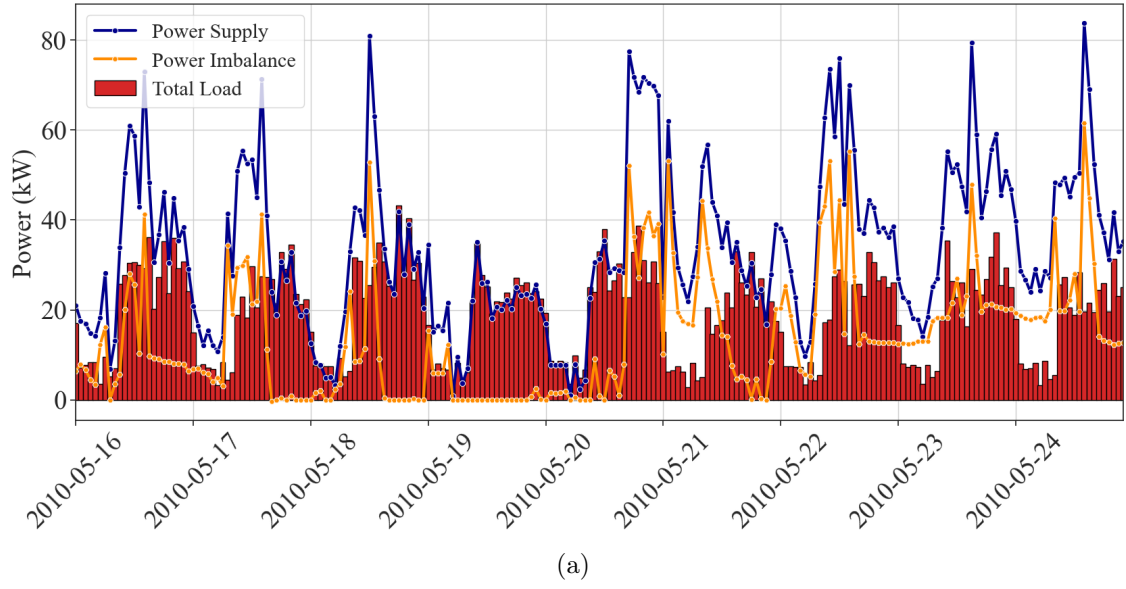


Figure 4.6: Load imbalance for Scenario 1: (a) total system imbalance, (b) essential loads.

local maximum error alone may not represent the overall approximation quality.

Three non-uniform PWL approximations were tested: a coarse six-segment model, the reference eight-segment model used in the thesis baseline, and a refined 17-segment model. Non-uniform segment lengths were used to increase the resolution in the low-DOD region, where the degradation curve changes most rapidly. Figures 4.8 and 4.9 compare the resulting approximations and their relative errors.

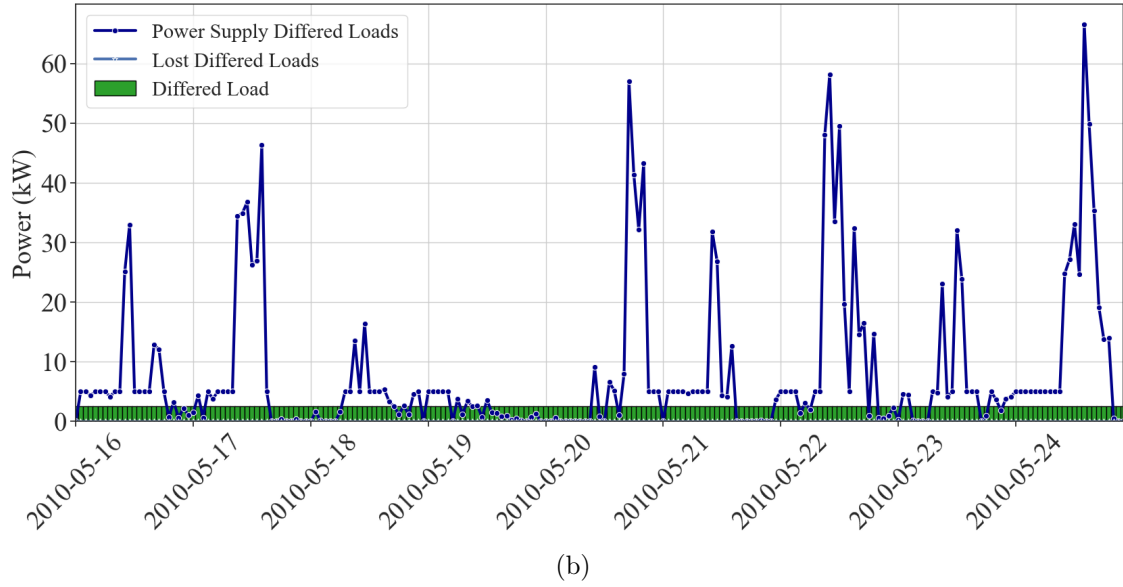
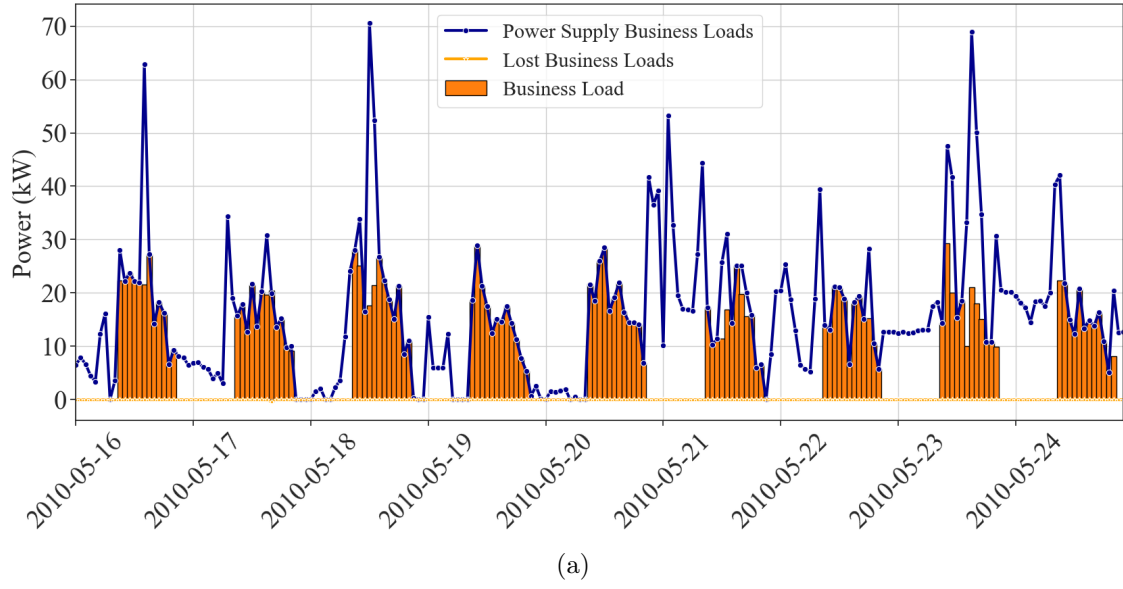


Figure 4.7: Load imbalance for Scenario 1: (a) business loads, (b) deferred (agricultural) loads.

Table 4.2: Sensitivity of MPC performance to BLC–DOD PWL approximation fidelity.

PWL case	Segments	Area error (%)	BLC binaries	RI	Total loss	EY
Reference PWL	8	4.02	1728	0.9978	8.61	16.90
Coarse PWL	6	9.70	1296	0.9959	20.26	16.87
Refined PWL	17	2.31	3672	0.9578	217.18	18.66

Table 4.2 shows that the coarse and reference PWL models produce similar operating behaviour. Reducing the number of segments from eight to six decreases the number of BLC-

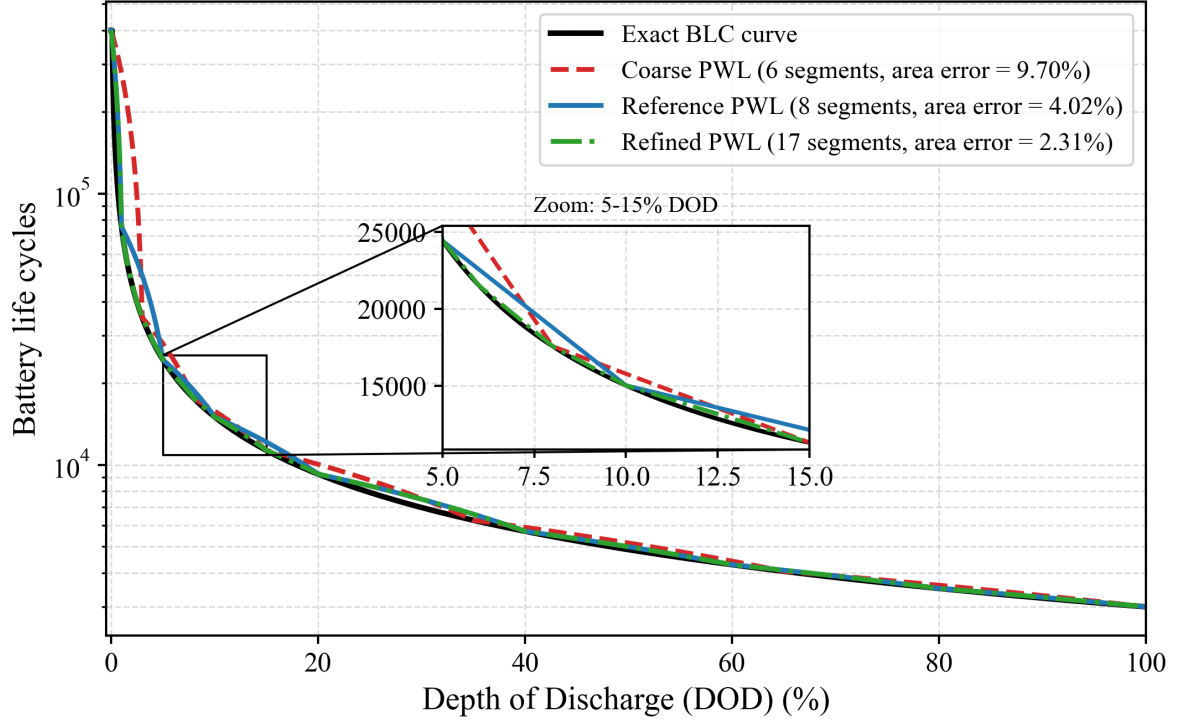


Figure 4.8: Comparison of the nonlinear BLC–DOD curve with the coarse, reference, and refined PWL approximations used in the MPC sensitivity analysis.

related binary variables from 1728 to 1296, while RI decreases only slightly from 0.9978 to 0.9959. The expected battery lifetime also remains almost unchanged, moving from 16.90 to 16.87 years.

The refined PWL model reduces the approximation error to 2.31%, but it also increases the number of BLC-related binary variables to 3672. In this case, the optimiser selects a more battery-preserving dispatch, increasing EY to 18.66 years. However, this improvement in battery lifetime is accompanied by a lower RI of 0.9578 and a higher total loss. This indicates that a more detailed degradation representation can shift the optimiser towards battery preservation when the BLC term is included in the objective function.

Overall, the results show that the MPC baseline is sensitive to the BLC–DOD linearisation. The reference eight-segment PWL model provides a practical compromise between approximation accuracy, model complexity, resilience, and expected battery lifetime. There-

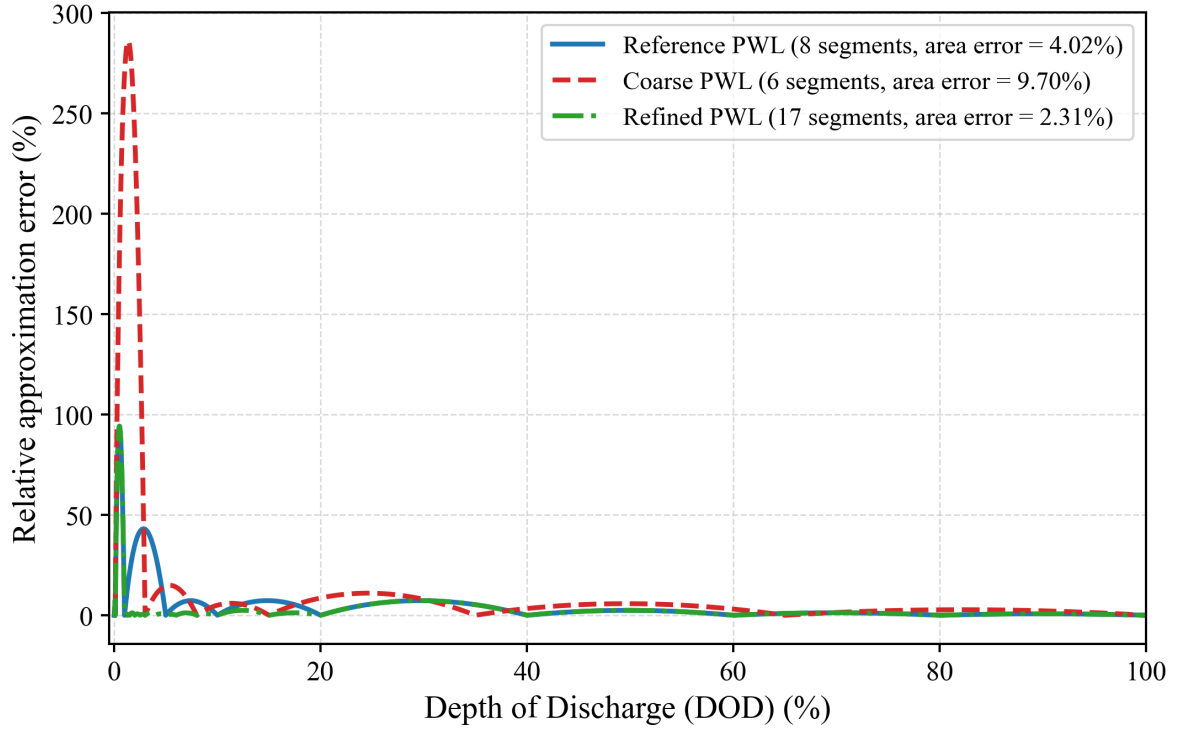


Figure 4.9: Relative approximation error of the coarse, reference, and refined BLC–DOD PWL models. The reported area error is the mean relative area error over the full DOD range.

fore, it is retained as the baseline degradation approximation for the main MPC results.

4.2.6 Impact of Demand Response

Scenario 16 examines the effect of introducing limited demand response participation into the MPC framework. In this case, essential and business loads can shift up to 10% of their demand, representing a realistic flexibility range for high-priority consumers. Deferred (agricultural) loads remain fully flexible, maintaining complete participation as in earlier cases. This setup was tested under identical conditions to the balanced configuration (Scenario 1) to isolate the specific influence of DR.

The inclusion of partial DR results in slightly flatter demand profiles and a measurable improvement in operational stability. Total load loss remains negligible, while the expected battery lifetime extends from 14.4 to about 15.6 years. Figure 4.10 shows the adjusted load patterns after applying the 10% shift to essential and business loads. The small temporal shift alleviates short-term power peaks, improving energy balance and reducing the number of high-rate charge–discharge events.

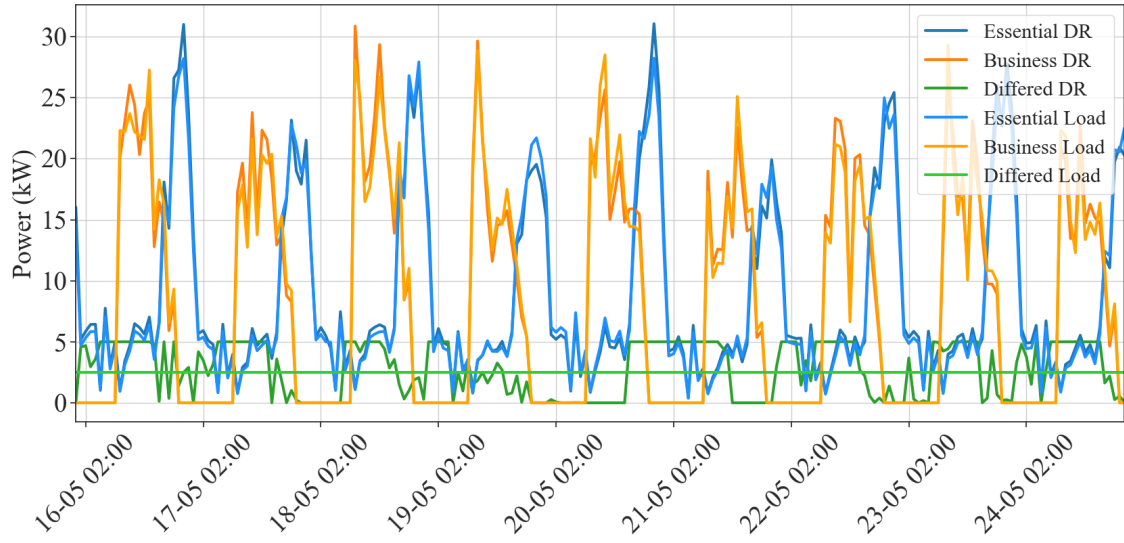


Figure 4.10: Adjusted essential and business load profiles after applying a 10% demand-response shift (Scenario 16).

Overall, the results demonstrate that even modest DR participation can enhance system resilience and extend battery life. Such limited flexibility is achievable in practice, requiring only small scheduling adjustments from consumers, and thus offers a practical means of improving microgrid performance without compromising the continuity of critical loads.

4.2.7 Smart Resilient Control under Reduced Battery Capacity

This subsection evaluates the effectiveness of the smart resilient control strategy when the available battery capacity is reduced. The controller introduced in Figure 3.15 dynamically adjusts the SOC limits and weighting priorities in response to the system's ERI. When resilience begins to deteriorate, the control logic relaxes SOC bounds or reassigns objective weights to prioritise load supply. This feature enables the microgrid to sustain operation under constrained energy storage conditions.

Scenario 17 serves as the reference case for this analysis, implementing the smart control mechanism with the full designed battery capacity. Its results are comparable to Scenario 1, achieving near-perfect resilience and an expected battery life of about 14.4 years, confirming that the adaptive controller performs equivalently to the baseline under nominal conditions.

Scenarios 18–24 then explore progressively reduced battery capacities corresponding to 75%, 50%, and 25% of the design size. At 75% capacity (Scenario 18), the ERI remains above 0.95, with only minor losses in business loads ($\simeq 1.4$ kWh) and a moderate decline in expected battery life to 11.4 years. This suggests that moderate capacity reduction still provides robust performance during HILP events. When capacity is halved (Scenarios 19–21), the ERI occasionally drops below 0.95, prompting the smart controller to adapt its parameters. Scenario 20 allows temporary relaxation of SOC bounds, while Scenario 21 further prioritises resiliency over all other objectives. As a result, Scenario 21 achieves lower essential load loss compared with Scenarios 19 and 20, albeit with a reduced battery lifetime of approximately 9 years. Across these cases, total load loss remains minimal (less than 3 % of demand), and the controller successfully maintains operation even during periods of very low generation.

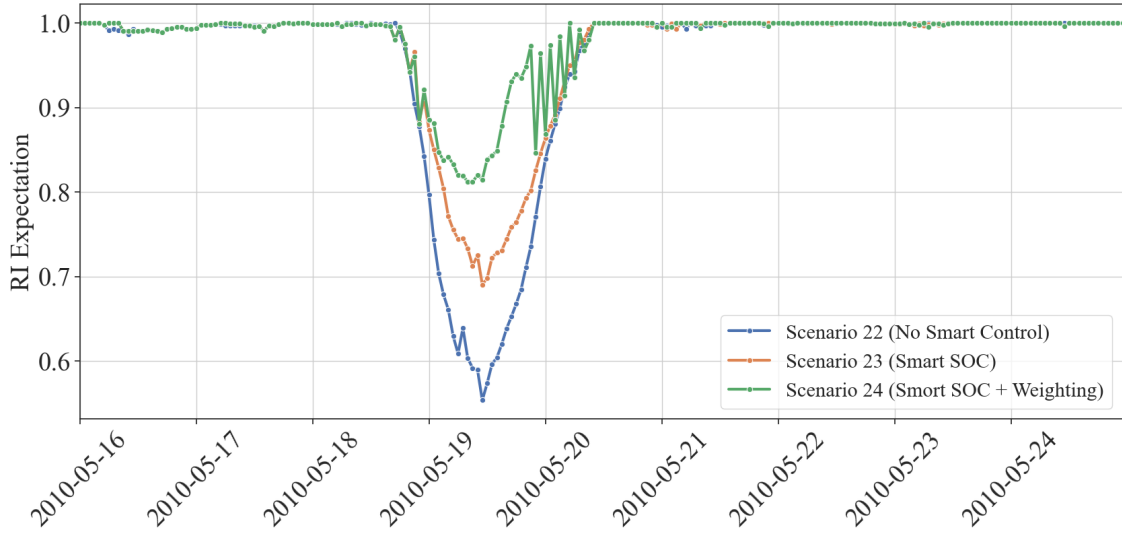


Figure 4.11: Effect of smart resilient control on the ERI with 25% of the designed battery capacity (Scenarios 22–24).

The impact of further reduction to 25 % of the design capacity (Scenarios 22–24) is shown in Figure 4.11. Although total losses increase slightly, the smart controller continues to stabilise system operation by dynamically relaxing SOC limits and reallocating energy among load categories. Without the controller (Scenario 22), the essential load loss reaches 87.8 kWh. When the smart control is fully applied (Scenario 24), this loss decreases to 35.6 kWh, approximately one-third of the uncontrolled case, while the ERI remains at 0.96. Battery life expectancy, however, falls to 7.5 years due to the deeper cycling associated with reduced storage. This reduction in expected life is attributed not to the depth of discharge required for a single event but to the higher relative DOD imposed by limited capacity.

Overall, the results demonstrate that the proposed smart resilient controller effectively preserves load supply and operational stability under capacity constraints. Even with only 25 % of the design storage, the system maintains an ERI above 0.95, preventing large-scale outages. Although the higher cycling depth shortens expected battery life, this trade-off is acceptable given the enhanced continuity of supply during extreme events. These findings confirm that adaptive SOC management and dynamic weighting are key to sustaining

resilience when storage resources are limited.

4.2.8 Grid-Connected Scenarios and Profit Analysis

This subsection examines the effect of connecting the microgrid to the main grid under both technical and economic perspectives. Scenarios 25–28 evaluate how grid interaction influences resilience, battery utilisation, and economic performance. In these cases, the controller is allowed to import or export power depending on system conditions, while maintaining the same multi-objective structure as the islanded configuration. Electricity purchase and sale prices are set to 25 pence kWh^{-1} and 15 pence kWh^{-1} respectively, representing a typical differential between grid import and export tariffs.

Scenarios 25 and 26 employ the full battery capacity designed during the planning stage. Scenario 25 allows unrestricted exchange of power with the grid but does not include profit maximisation in the objective function. As a result, the controller discharges the battery steadily to meet local demand, with little consideration of economic benefit. Although this strategy maintains full load supply, it uses available storage capacity more frequently, yielding an extended but not optimised battery lifetime.

Scenario 26 introduces profit as an explicit objective, encouraging the controller to make more efficient use of energy trading opportunities with the grid. Although the buy and sell prices remain constant throughout the simulation, the inclusion of the profit term leads to a more selective operation pattern. The controller prioritises meeting local demand using renewable generation and exports only the excess energy, while avoiding unnecessary imports. This adjustment reduces deep discharges, lowers cycling frequency, and extends the expected battery life to approximately 27 years. Over the nine-day simulation, total energy purchases decrease from 900.82 kWh to 240.97 kWh, while exports fall slightly from 2,208.46

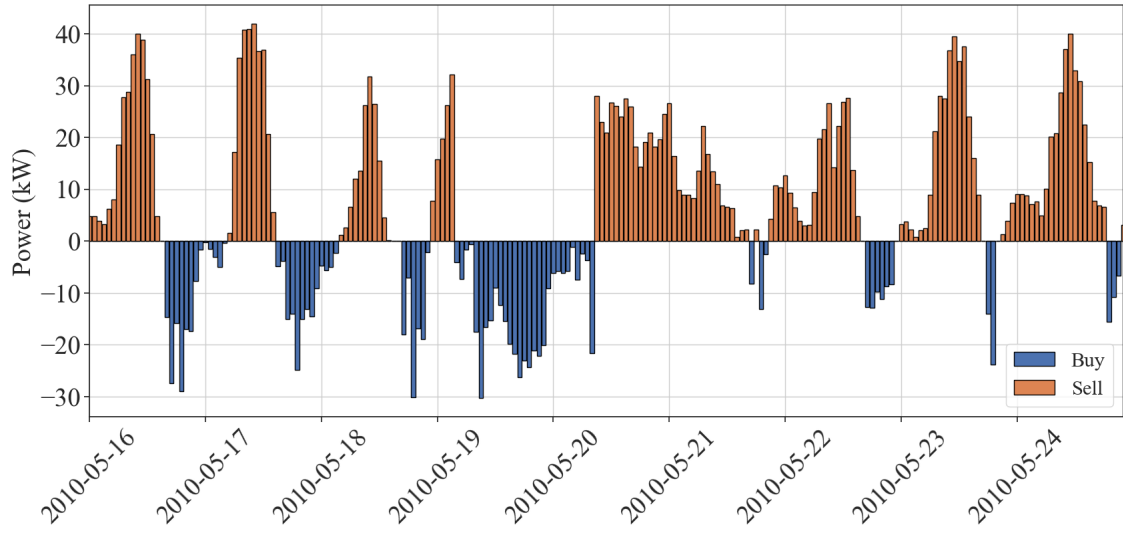
kWh to 2,010.27 kWh. The resulting profit increases from \$106.06 in Scenario 25 to \$250.29 in Scenario 26, representing a 136 % improvement.

Figures 4.12a and 4.12b illustrate the power exchange with the grid for Scenarios 25 and 26 respectively. In Scenario 25, the pattern reflects a nearly continuous discharge behaviour, whereas Scenario 26 demonstrates periodic import–export cycles aligned with the economic optimisation objective. The introduction of the profit term therefore transforms the control strategy from purely technical balancing to cost-aware scheduling, yielding simultaneous gains in both economic and operational performance.

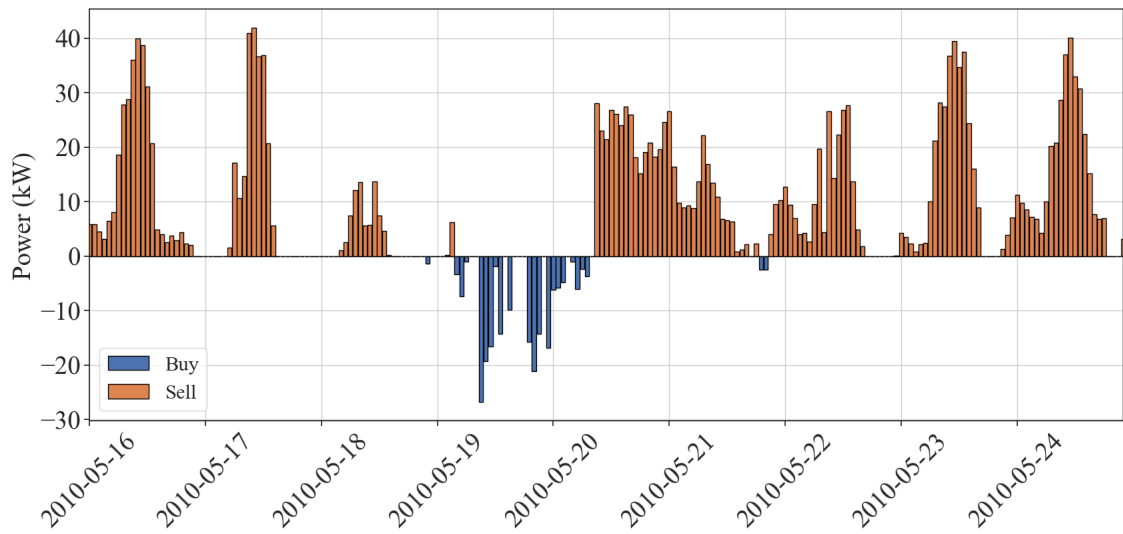
Scenarios 27 and 28 repeat the same configurations but with only 25 % of the original battery capacity to investigate how grid support compensates for limited storage. Despite the reduction, all loads are fully supplied due to the grid’s backup capability. Battery lifetime increases by nearly three years compared with the equivalent islanded scenarios (22–24), since grid imports prevent deep discharges. In Scenario 27, where profit is not considered, 891.03 kWh is purchased and 2,210.85 kWh is exported, leading to a net benefit of \$108.87. Scenario 28, which incorporates profit optimisation, further refines this balance: only 211.31 kWh is purchased and 1,950.16 kWh is exported, producing a total profit of \$239.69 (approximately double that of Scenario 27). These results confirm that even with reduced storage, intelligent grid interaction preserves full resiliency and can generate economic value.

Figure 4.13a and 4.13b show the corresponding import and export patterns. Scenario 27 displays steady discharge behaviour similar to Scenario 25, whereas Scenario 28 achieves a more profitable and storage-friendly pattern through selective energy trading. The ability to sell surplus renewable generation while purchasing energy strategically reduces battery stress and extends its lifespan.

Overall, the grid-connected analyses highlight the dual benefit of external energy support and market participation. Grid access reduces reliance on local storage, thereby



(a)



(b)

Figure 4.12: Electricity exchange with the grid: (a) without profit optimisation (Scenario 25), (b) with profit optimisation (Scenario 26).

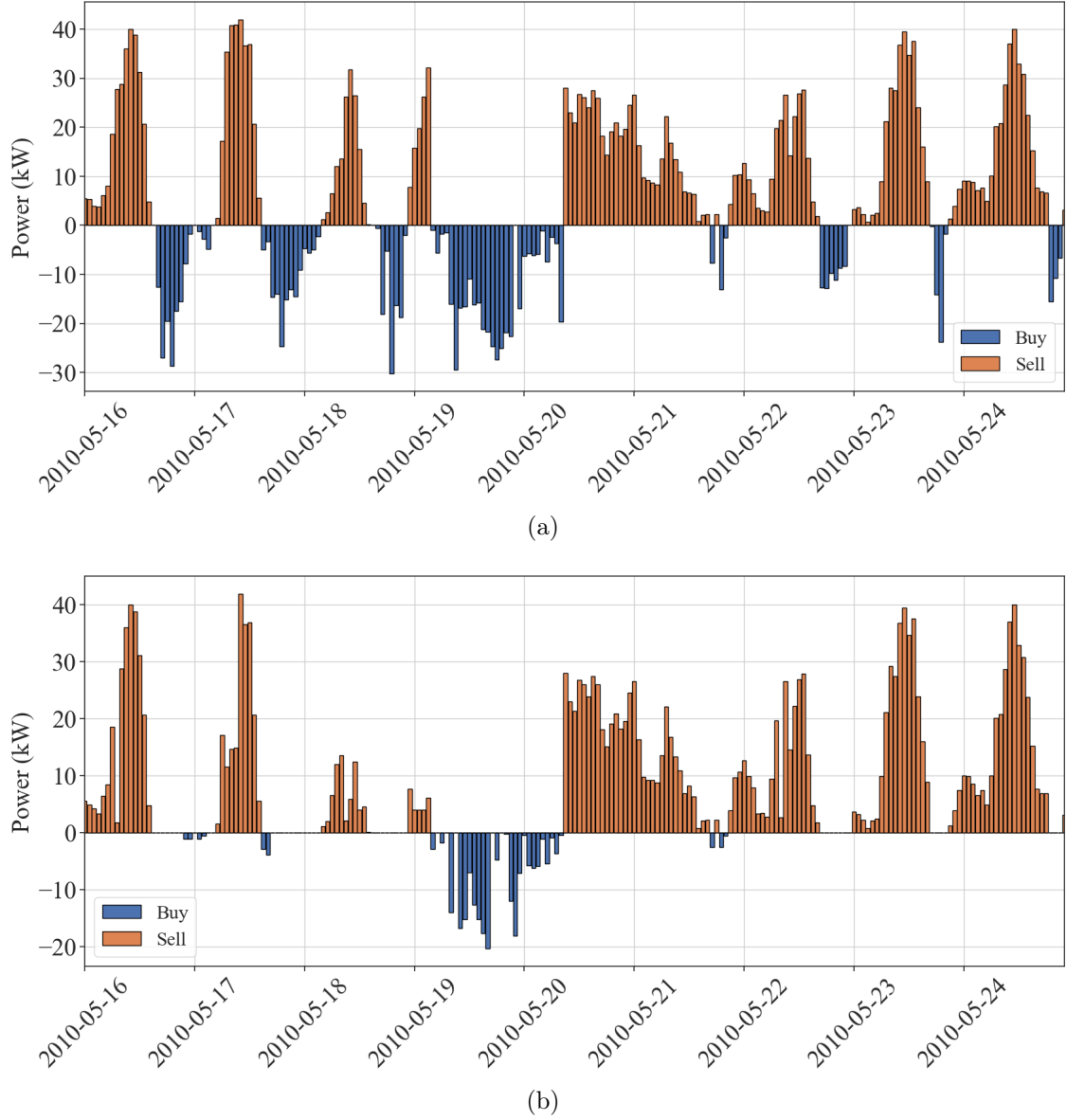


Figure 4.13: Electricity exchange with the grid for 25% battery capacity: (a) without profit optimisation (Scenario 27), (b) with profit optimisation (Scenario 28).

mitigating degradation and improving economic efficiency. While isolated operation guarantees autonomy, strategic grid coupling offers a practical trade-off between independence, resilience, and long-term sustainability. These findings underscore the importance of integrating market-based objectives into future microgrid controllers to enhance both operational and financial resilience.

Objective Achieved

✓ **Objective 3 – Formulate an MPC baseline.** This objective asked: *How should an MPC baseline be structured to operate while accounting for battery life, load priorities, and an expected-resilience perspective?*

It has been accomplished through the development of a mixed-integer linear MPC formulation that integrates a piecewise BLC-DOD relation, priority-weighted loads, and an expected-resilience term derived from resilience curves. Scenario-based simulations confirmed that the controller balances resilience and battery longevity, forming a robust benchmark for the DRL comparison in Chapter 5.

4.3 Discussion and Conclusion

The results presented in this chapter confirm that the proposed MPC framework offers an effective and interpretable solution for resilient and battery life-aware microgrid operation. Across all scenarios, the controller maintained nearly full load supply during both normal and HILP events while safeguarding battery health through foresight-based optimisation. The integration of multiple objectives was essential for balancing short-term operational resilience with long-term storage preservation, with equal weighting among these terms delivering the most stable and sustainable performance.

Incorporating a limited DR programme further improved operational flexibility, extending the expected battery lifetime by approximately 8.4% without compromising the supply of critical loads. Similarly, the smart resilient controller sustained high performance even under reduced storage conditions, maintaining an ERI above 0.95 even with only 25% of the designed battery capacity. These results confirm the adaptive strength of the control structure, which dynamically adjusts SOC bounds and objective priorities to preserve

resilience during resource-constrained conditions.

Grid-connected operation introduced an additional dimension of economic optimisation. By enabling controlled power exchange with the main grid, the system reduced deep discharge cycles and alleviated stress on the battery. Profit-oriented operation increased the total economic return by 136% compared with the non-profit case, while maintaining uninterrupted load supply. These findings demonstrate that coupling the microgrid with the grid can enhance both resilience and cost efficiency without compromising autonomy during critical events.

Overall, the MPC-based framework establishes a robust foundation for resilient microgrid operation by integrating predictive optimisation, adaptive weighting, and flexible coordination across technical and economic layers. However, its dependence on accurate forecasting and computationally intensive mixed-integer programming restricts its real-time scalability. These limitations motivate the development of data-driven approaches, particularly deep reinforcement learning, to achieve comparable or superior performance with significantly reduced computation time. The following chapter builds on this foundation by introducing a DRL-based controller designed to replicate the decision quality of MPC while improving speed, adaptability, and deployability for real-world microgrid applications.

Chapter 5

Deep Reinforcement Learning and Explainability Framework

5.1 Introduction

This chapter presents the DRL framework developed for the resilient and battery-aware operation of the microgrid system introduced in Chapter 4. While Chapter 4 focused on the MPC benchmark, the current chapter transitions to a model-free control strategy based on the PPO algorithm. The proposed framework integrates a curriculum-based training protocol, uncertainty modelling, and post-hoc explainability to ensure both robust performance and transparent decision-making.

The theoretical foundation of the PPO algorithm is introduced in Section 3.4.3.4, supported by its step-wise implementation in Algorithm 1. The Gymnasium environment, which provides a standard interface for DRL studies, is described in Section 3.4.3.6. The complete modelling of the microgrid environment, including state, action, and reward formulation, is detailed in Section 3.4.3.7 through Equations 3.71 to 3.87. These formulations define how the agent perceives the system, interacts with it, and learns to balance resilience with battery longevity.

Training follows a curriculum strategy illustrated in Figure 3.17, where task complex-

ity is increased gradually to improve convergence and adaptability. The data preparation and augmentation procedures used to construct the training datasets are explained in Section 3.5.1. Three distinct datasets are employed: the extreme-weather dataset based on Cyclone Laila, the HOMER Pro-derived dataset representing nominal operating conditions, and a historical dataset of normal operation. Their variations and perturbation effects are shown in Figures 3.18 to 3.20.

To assess post-training robustness, Perlin noise is imposed on the renewable generation profiles as described in Section 3.5.2. This uncertainty model is later employed for the Monte Carlo analysis presented in this chapter to assess the generalisation capability of the trained policy. Finally, the chapter introduces post-hoc explainability techniques, LIME and SHAP, in Section 3.4.4. Together, these components establish the XDRL framework that is later evaluated and benchmarked against the MPC baseline.

5.2 Performance Evaluation and Comparative Results

This section evaluates the trained DRL controller and compares it with the MPC benchmark. The analysis covers resilience, expected battery life, load imbalance, and SOC behaviour under both deterministic and uncertain conditions. A final comparison quantifies computational efficiency to assess the practicality of real-time or edge-level deployment.

5.2.1 Deterministic Evaluation

This part establishes a baseline under ideal conditions by evaluating the trained DRL controller against the MPC benchmark with perfectly known inputs over the 9-day Cyclone Laila window. All runs use the same asset sizing and constraints as in Table 3.1, and the

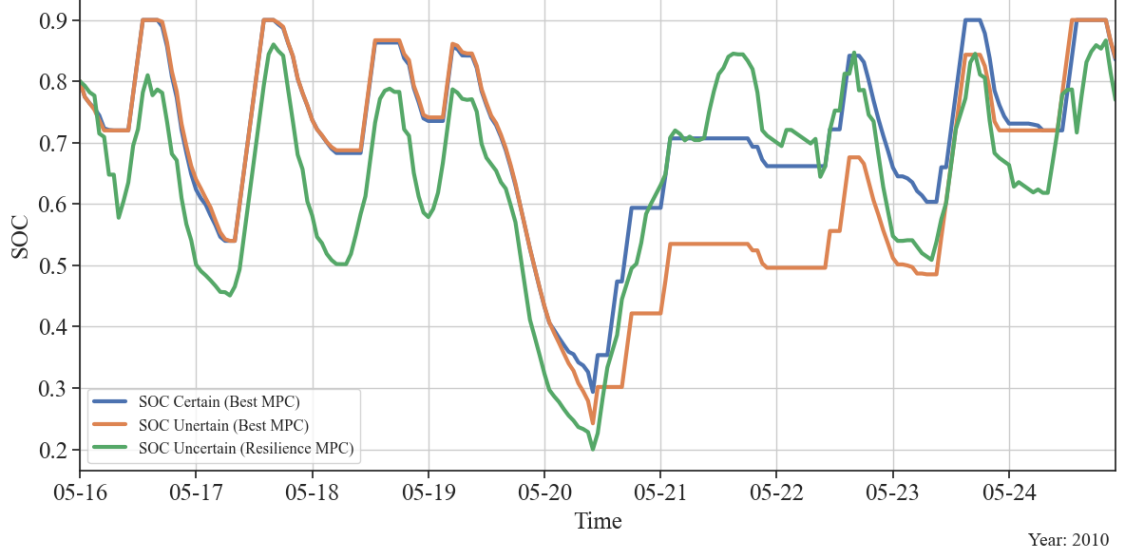


Figure 5.1: SOC for Different MPC Weighting.

initial SOC is fixed at 0.8 to ensure a fair comparison.

Figure 5.1 shows SOC trajectories for MPC under resilience-only and optimally weighted configurations as the baseline for comparison. The resiliency-only tuning (green) cycles almost continuously: rapid climbs to the upper bound after renewable peaks and deep draw-downs (down to about 0.2 around 20 May), with few idle plateaus. By contrast, the best weighting runs maintain longer plateaus in the 0.7–0.8 band and smaller daily amplitude, evidencing more conservative reserve management. The MPC rows of Table 5.1 are consistent with these trajectories. The deterministic best case achieves the highest RI (0.9991) with low total imbalance (14.00 kWh), fully confined to L_3 (zero for L_1 and L_2), and a moderate EY of 14.44 years. Under uncertainty, the *best* case retains high RI (0.9989) while slightly increasing EY to 15.06 years and keeping imbalances low (16.18 kWh, mostly L_3 with just 0.81 kWh on L_2). The resiliency-only setting, despite its objective, depletes the battery more aggressively and later experiences reduction across all classes, yielding the lowest EY (10.78 years) and the largest total imbalance (55.46 kWh; $L_1=1.44$, $L_2=23.36$, $L_3=30.66$) with RI 0.9947.

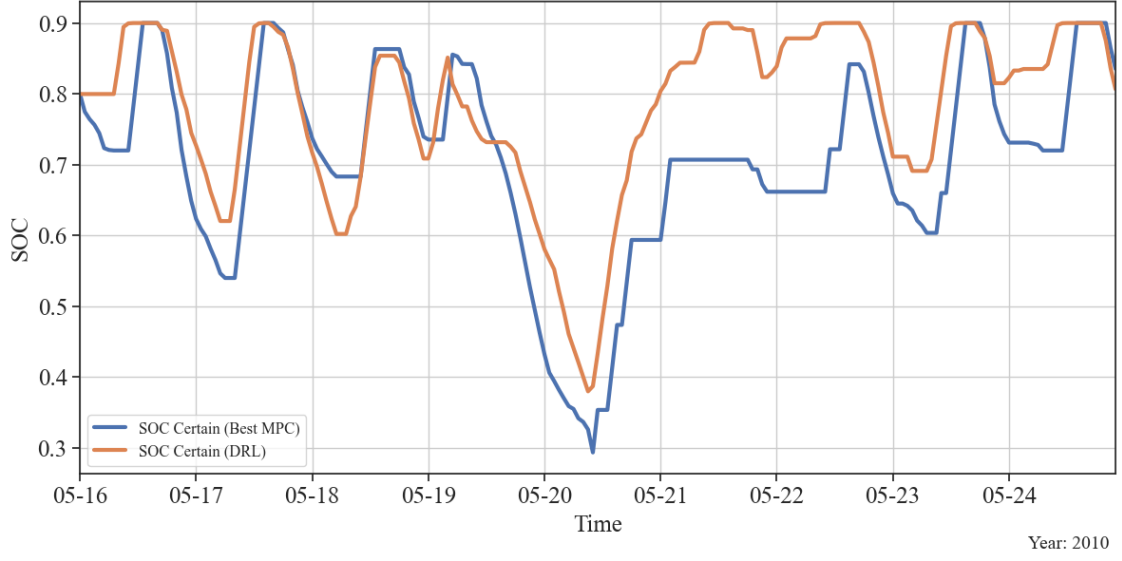


Figure 5.2: SOC MPC VS DRL.

The direct comparison between controllers is shown in Figure 5.2. The trajectories reveal that DRL exhibits smoother charging and discharging transitions over time, whereas MPC performs more aggressive and binary shifts between charging and discharging, which is particularly evident during solar peaks and load dips. This behaviour indicates that DRL learns to moderate energy exchange and avoid unnecessary cycling stress, contributing to improved battery longevity compared with MPC.

Table 5.1 summarises the quantitative outcomes. DRL in Stage 4 (certain) achieves only 0.3% lower RI compared with MPC, but delivers more than 10% improvement in EY, extending from 14.44 to 15.90 years. The load imbalance pattern reflects the learned trade-off within the DRL policy. DRL has learned to prioritise essential demand, maintaining zero imbalance for essential loads (L_1), while flexibly adjusting business (L_2) and agricultural (L_3) loads to balance both objectives. MPC, in contrast, allocates all imbalance to L_3 , achieving perfect short-term supply at the cost of higher cycling. Thus, DRL attains near-optimal resilience with improved longevity through smoother dispatch. A derivative-based quantification of smoothness is provided later in Section 5.2.4.

Table 5.1: Comparison of DRL and MPC results under certain and uncertain scenarios.

Method	Scenario	RI	EY	L_1 Imb.	L_2 Imb.	L_3 Imb.	Tot Imb.
DRL	Stage 3	0.9986	13.49	2.84	0.00	2.20	5.04
DRL	Stage 4 (Certain)	0.9961	15.90	0.00	31.58	1.20	32.78
DRL	Stage 4 (Uncertain)	0.9956	15.88	0.00	35.96	1.59	37.55
MPC	Certain (Best)	0.9991	14.44	0.00	0.00	14.00	14.00
MPC	Uncertain (Best)	0.9989	15.06	0.00	0.81	15.37	16.18
MPC	Resiliency Only	0.9947	10.78	1.44	23.36	30.66	55.46

For consistency, the reported EY values in Table 5.1 are calculated using the exact nonlinear BLC–DOD curve as a common post-evaluation basis for both MPC and DRL. The PWL degradation model is used within the MPC formulation to preserve the MILP structure, whereas the DRL environment can evaluate the nonlinear degradation curve directly. To check the effect of this modelling difference, the saved Stage 4 DRL trajectory was also re-evaluated using the same PWL approximations used for the MPC sensitivity test. The operational results were unchanged, however, the post-calculated EY varied from (15.89) years using the exact nonlinear curve to (28.05), (21.81), and (16.30) years using the coarse, reference, and refined PWL approximations, respectively. This confirms that the PWL approximation can affect the battery-life estimate, particularly for low-DOD trajectories, and therefore supports using the exact nonlinear curve for the final MPC–DRL battery-life comparison.

5.2.2 MPC Baseline under Imperfect Forecasts

To further strengthen the comparison framework, the MPC baseline was also tested under imperfect renewable generation forecasts. The same MPC formulation, objective weights, initial SOC, battery parameters, and operating constraints were retained in both cases. Therefore, the only change between the two MPC cases was the renewable-generation forecast used within the prediction horizon.

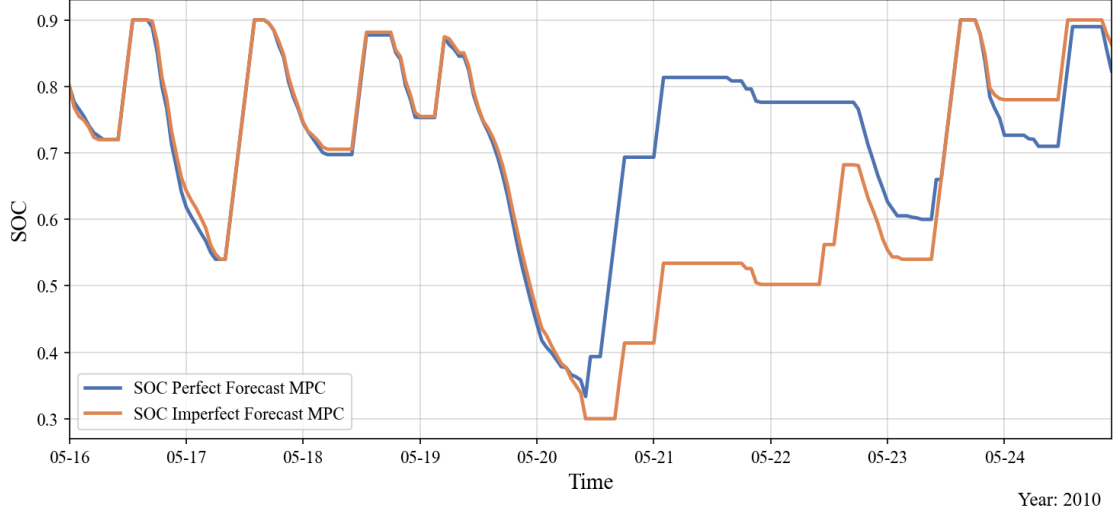


Figure 5.3: Battery SOC comparison between perfect and imperfect forecast MPC operation during the HILP test period.

In the perfect-forecast case, the measured extreme renewable-generation profile was used across the full prediction horizon. In the imperfect-forecast case, the current time step was evaluated using the measured extreme profile, while the future prediction horizon used a Perlin-noise based forecast error model, consistent with the uncertainty representation used in the DRL uncertainty analysis. This provides a direct forecast sensitivity check for the MPC baseline.

Figure 5.3 shows that the imperfect forecast case follows a different battery trajectory during and after the main disturbance period. This supplementary test indicates that forecast imperfection changes the MPC dispatch pattern, particularly by encouraging less aggressive battery use over part of the horizon. Compared with the perfect-forecast run, the imperfect forecast run gives an almost unchanged RI, with only a marginal reduction, while total load loss increases from 7.72 kWh to 10.63 kWh. At the same time, the expected battery lifetime increases from 15.06 to 16.10 years, and the number of discharges decreases from 96 to 89. This confirms that the main effect of forecast imperfection in this additional MPC test is a small shift in the resilience–battery life trade-off, rather than a major loss of

resilience.

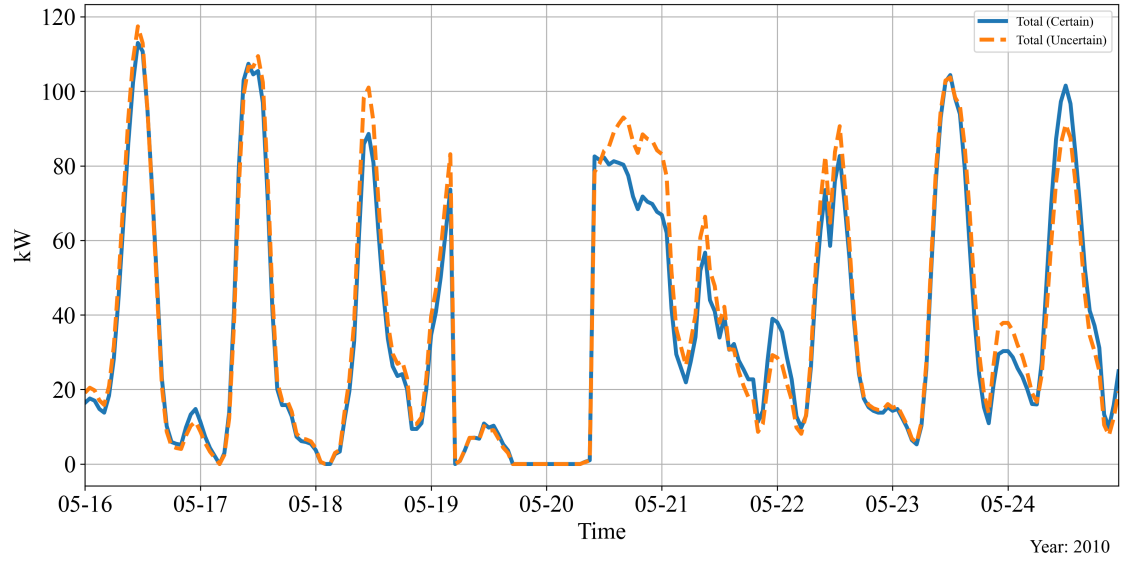
This indicates that forecast imperfection does not materially undermine the resilience of the MPC baseline in the tested HILP case, since both cases retain an RI very close to unity. Its main effect is to shift the resilience–battery life trade-off: the imperfect forecast case produces slightly higher operational loss, but also less aggressive battery use and a longer expected battery lifetime.

5.2.3 Performance under Renewable Uncertainty

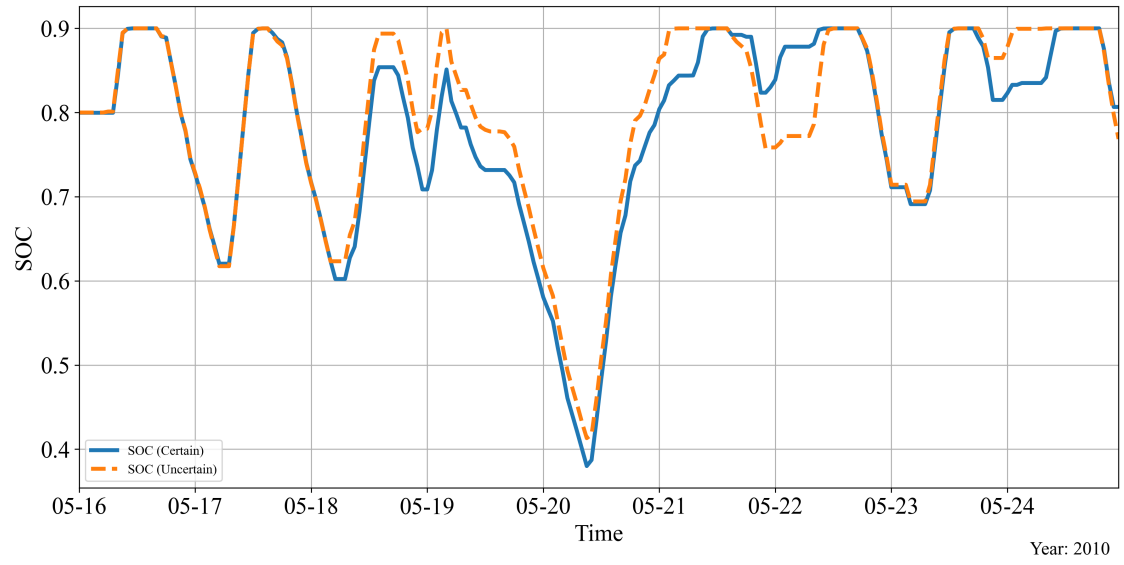
The preceding subsection isolates the effect of forecast imperfection on the MPC baseline only. The following subsection returns to the controller comparison results in Table 5.1, where MPC and DRL are compared under the common uncertainty evaluation setting.

To evaluate robustness under realistic forecasting errors, both controllers were tested using perturbed renewable generation profiles produced by the Perlin noise model described in Section 3.5.2. This approach introduces smooth, temporally correlated variations in solar and wind signals, emulating the uncertainty typically observed in day-ahead forecasts without causing abrupt or unphysical fluctuations. Figure 3.22 illustrates the comparison between white noise and Perlin noise perturbations, highlighting the latter’s gradual and coherent variability.

The effect of the perturbations is shown in Figure 5.4. In Figure 5.4a, the “uncertain” renewable trace (dashed) stays close to the deterministic profile but nudges midday peaks and shoulder periods up or down by a few kilowatts, while preserving the daily shape. In Figure 5.4b, the resulting SOC trajectories almost overlap: both remain mostly in the 0.7–0.9 band, with the uncertain case tracking the same rise/fall pattern and slightly easing the deepest dip around 20 May.



(a)



(b)

Figure 5.4: Perlin noise effects on a) renewable generation and b) battery SOC.

Under uncertainty (Table 5.1), DRL attains $RI = 0.9956$ versus 0.9989 for MPC, a gap of about 0.33%, while keeping a higher expected battery life ($EY = 15.88$ years vs. 15.06 years for MPC, approximately 5% advantage). The curtailment pattern differs: DRL holds the essential load at zero imbalance and uses the business and agricultural loads as flexibility (0.00, 35.96, 1.59 kWh for L_1 , L_2 , and L_3 ; 37.55 kWh total). MPC also protects essentials but concentrates almost all curtailment on the lowest-priority agricultural load (0.00, 0.81, 15.37 kWh; 16.18 kWh total). Notably, DRL’s EY is essentially unchanged from the deterministic case ($15.90 \rightarrow 15.88$ years) with only a minor RI reduction ($0.9961 \rightarrow 0.9956$), indicating that its battery-aware behaviour transfers robustly to perturbed renewables while preserving essential demand.

5.2.4 Stage-wise SOC Analysis and Derivative Comparison

To further illustrate the behavioural differences between the curriculum stages and benchmark control, this section analyses the SOC trajectories and their temporal derivatives. Figure 5.5 compares Stage 3, which focuses solely on resilience, with Stage 4, which balances resilience and battery lifetime through the inclusion of the degradation-aware reward term. The Stage 3 profile exhibits aggressive charge–discharge cycling with frequent transitions between upper and lower bounds, reflecting a purely resilience-driven objective. By contrast, Stage 4 maintains smoother and more moderated SOC evolution, characterised by longer periods of partial charging and reduced DOD. This behavioural shift confirms that the agent learns to trade slight reductions in instantaneous resilience for significantly improved long-term battery preservation.

A finer comparison between MPC and DRL is presented in Figure 5.6, showing the SOC first and second temporal derivatives. As illustrated in Figure 5.2, both controllers start from the same initial SOC of 0.8.

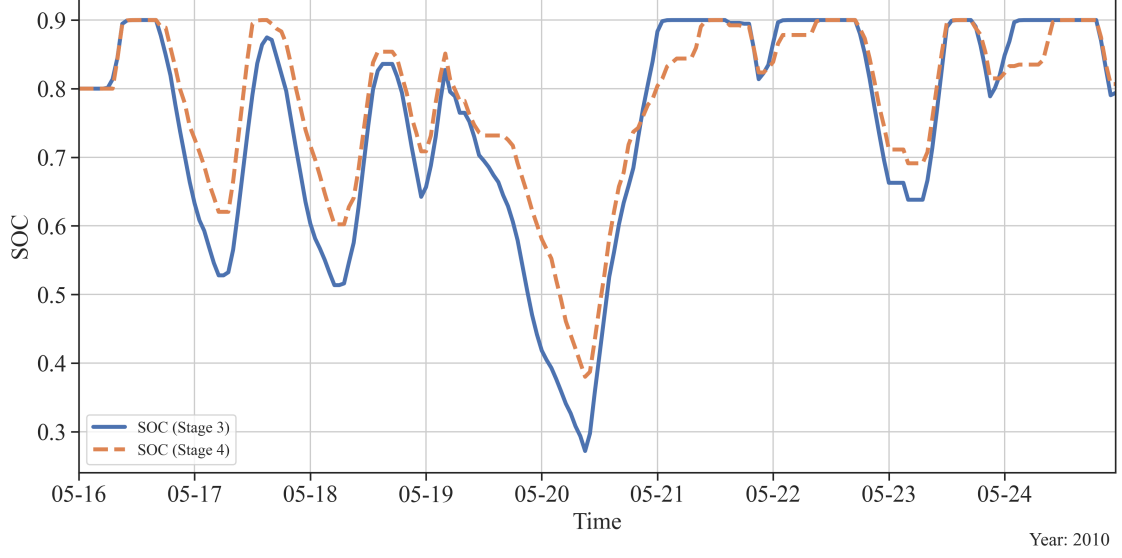
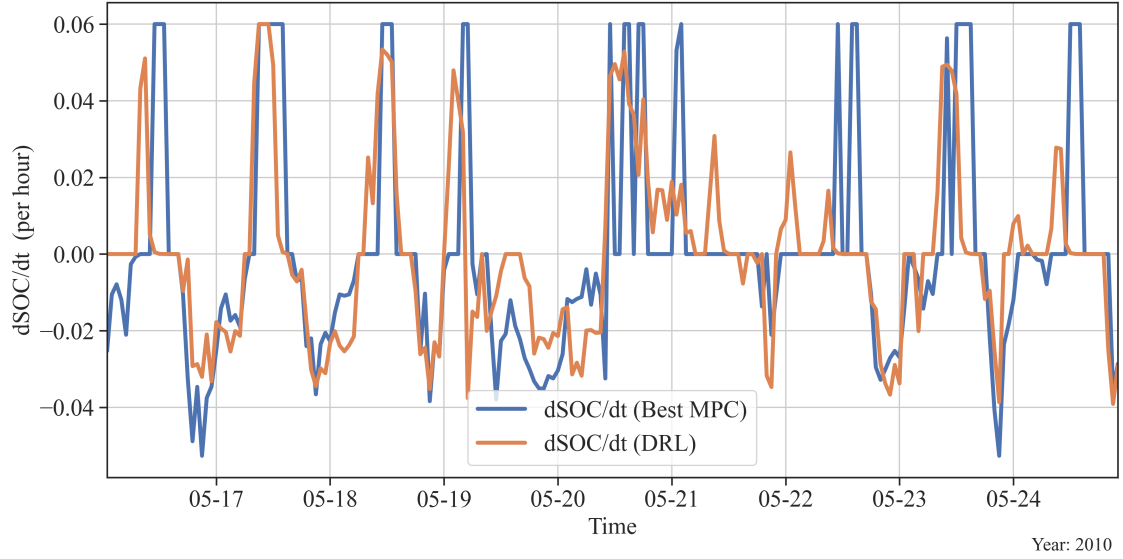


Figure 5.5: SOC comparison: Stage 3 prioritises resilience, whereas Stage 4 balances resilience and battery life.

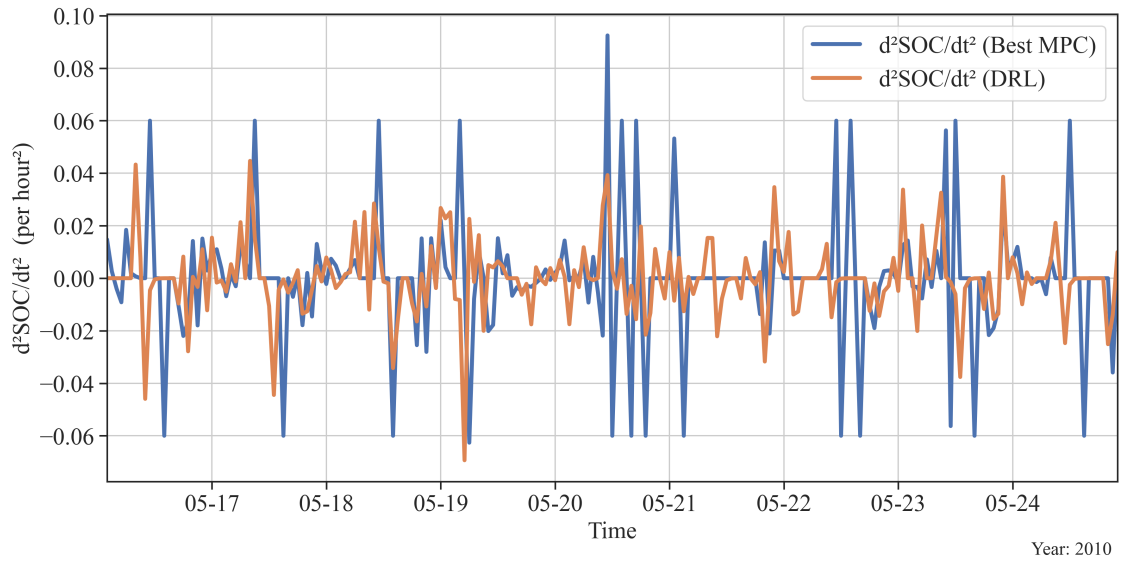
The first derivative of SOC, shown in Figure 5.6a, represents the rate of charge and discharge. Quantitatively, DRL achieves an average $|dSOC/dt|$ of $0.0155h^{-1}$ (RMS = $0.0225h^{-1}$), compared with $0.0177h^{-1}$ (RMS = $0.0277h^{-1}$) for MPC, corresponding to reductions of about 12% in mean and 19% in RMS rate. The second derivative in Figure 5.6b quantifies the smoothness of transitions between charging and discharging modes. DRL maintains lower fluctuation levels with an average $|d^2SOC/dt^2|$ of $0.0087h^{-2}$ (RMS = $0.0141h^{-2}$), whereas MPC shows $0.0118h^{-2}$ (RMS = $0.0229h^{-2}$), indicating reductions of roughly 26% and 38%, respectively. These metrics confirm that DRL moderates high-frequency control oscillations, resulting in smoother energy exchange and lower degradation stress.

5.2.5 Computational Cost Benchmarking

Computational benchmarking was carried out to quantify the processing efficiency of each controller. The total FLOPs were estimated for both training and inference phases to pro-



(a)



(b)

Figure 5.6: Comparison of a) SOC, b) first derivative, and c) second derivative between DRL and MPC.

Table 5.2: Computation-time comparison of MPC and DRL controllers [FLOPs].

Method	Training Time	Inference per Step	Inference per Episode	Lifetime Cost
MPC	–	$(7.594 \pm 21.840) \times 10^{11}$	$(1.650 \pm 0.048) \times 10^{14}$	$(1.004 \pm 0.030) \times 10^{17}$
DRL	$(3.489 \pm 0.059) \times 10^{16}$	$(1.318 \pm 0.047) \times 10^8$	$(2.846 \pm 0.102) \times 10^{10}$	$(3.491 \pm 0.059) \times 10^{16}$

vide a hardware-independent comparison. As summarised in Table 5.2, the MPC controller exhibits substantially higher computational requirements at every stage. MPC’s per-step inference cost is over three orders of magnitude greater than that of DRL, making the latter more than five thousand times lighter for real-time execution. Across an entire operating episode, DRL maintains this efficiency advantage, requiring approximately 2.8×10^{10} FLOPs compared with 1.7×10^{14} FLOPs for MPC, corresponding to a reduction of nearly four orders of magnitude in computational effort. When the fifteen-year operational horizon is considered, DRL achieves a total lifetime cost roughly three times lower than MPC, even after including its one-off training phase. This reduction in computational demand translates directly into faster inference, reduced energy consumption, and lower hardware dependency, factors that are critical for real-time control and practical deployment in edge or embedded microgrid systems.

To conclude the deterministic evaluation presented in this section, Figure 5.7 provides an integrated, normalised comparison between DRL and MPC across key performance indicators. Each axis is scaled such that higher values indicate better performance, with curtailments and computational measures inverted accordingly. The DRL controller achieves a broader coverage overall, highlighting its superior battery longevity, substantially lower inference and total computational cost, and balanced curtailment distribution across the different load tiers. In contrast, MPC exhibits slightly higher resilience but at the expense of higher computational demand and more concentrated curtailment in the lowest-priority load. These results emphasise the efficiency and adaptability of DRL, which maintains near-optimal resilience while delivering significant gains in battery life and execution practicality.

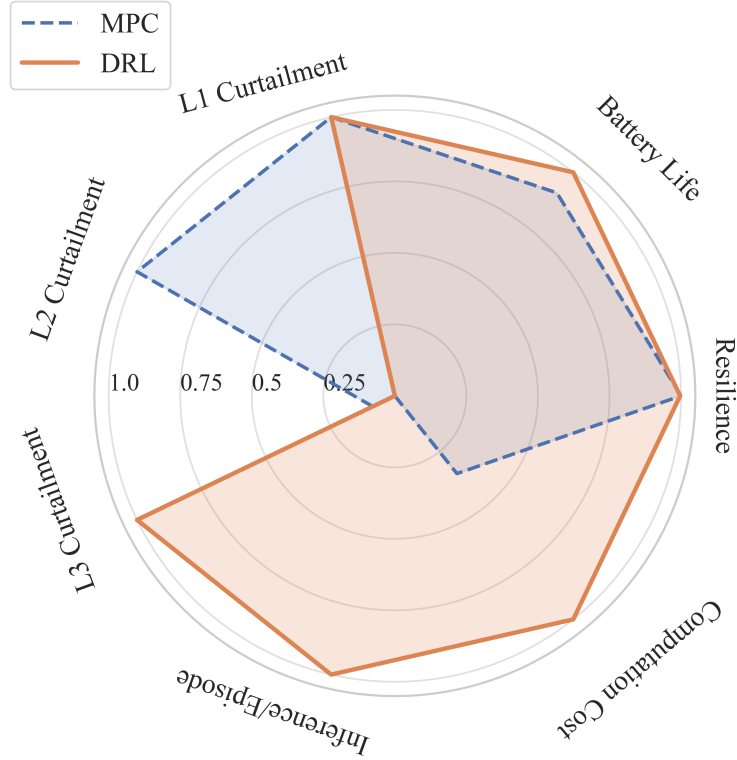


Figure 5.7: Comparative spider diagram of DRL and MPC under deterministic conditions across resilience, battery life, curtailment by load tiers, inference cost per episode, and total computation cost.

5.3 Monte Carlo-Based Uncertainty Quantification and Robustness Analysis

As discussed in Section 5.2.5, the lightweight inference structure of the DRL controller enables large-scale robustness testing that would otherwise be computationally prohibitive for traditional optimisation-based controllers. Once training is complete, each control decision reduces to a single forward pass through the neural network. As a result, the entire Monte Carlo analysis was completed in approximately twenty minutes on a standard laptop. By contrast, the MPC formulation must solve a mixed-integer optimisation at every step, where even a single scenario can take several minutes or more to converge under complex forecast or weighting conditions. This computational advantage facilitates extensive robustness

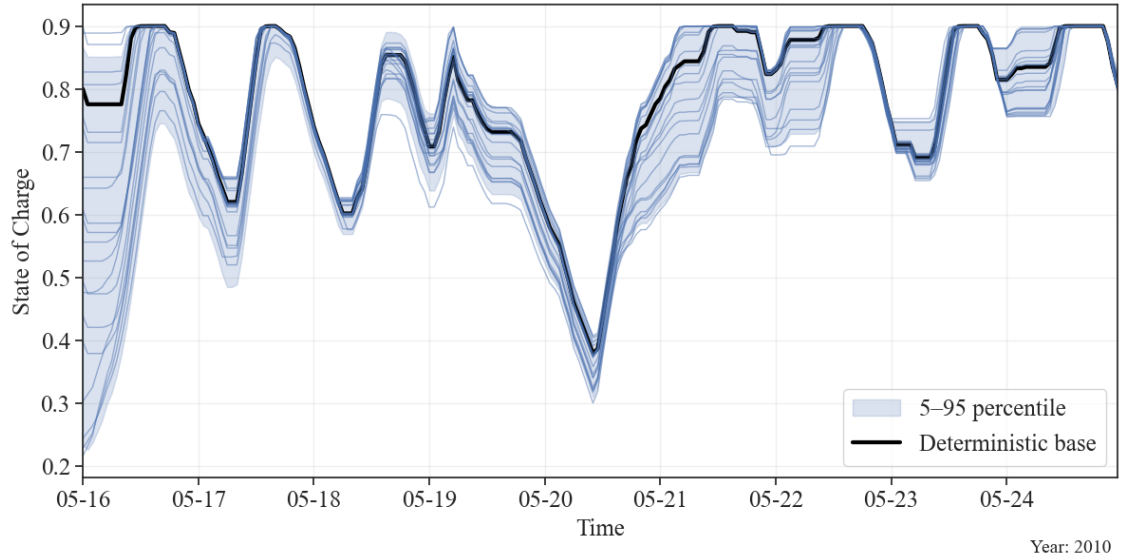
evaluation and scalable uncertainty assessment in practical microgrid settings.

To evaluate the robustness and reliability of the proposed DRL controller under uncertain renewable generation, a Monte Carlo-based Uncertainty Quantification (UQ) analysis was conducted. A total of 4000 stochastic scenarios were generated by perturbing the solar and wind generation profiles using the Perlin noise model. Each scenario was evaluated using the trained DRL policy, and the corresponding performance metrics were recorded.

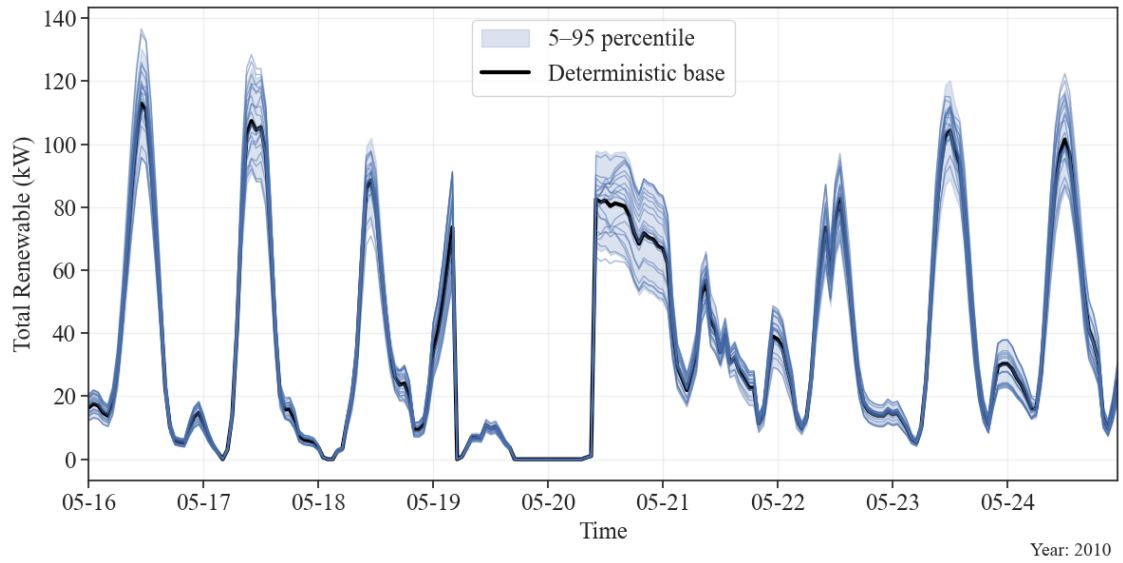
Figure 5.8 illustrates the dynamic response of the system across the 4000 Monte Carlo simulations. Figure 5.8a presents the battery SOC trajectories, while Figure 5.8b shows the total renewable generation profiles. The SOC traces in Figure 5.8a exhibit consistent charging–discharging behaviour aligned with renewable availability, with most trajectories confined within a stable operating band, indicating robust control adaptation. In Figure 5.8b, renewable generation displays notable variability due to the applied Perlin noise, especially during high-irradiance periods. Despite this, the DRL controller maintains stable SOC behaviour and effective energy balance, demonstrating its capacity to manage uncertainty without performance degradation.

Figure 5.9 shows the statistical distributions of RI and EY across the 4000 Monte Carlo runs. The histograms and kernel density estimates indicate that RI values are slightly skewed towards the upper bound, with most samples concentrated around 0.99, confirming the consistently high resilience of the DRL policy. In contrast, EY exhibits a near-symmetric, bell-shaped distribution centred around 16 years, with limited spread, indicating stable long-term battery life expectations despite renewable uncertainty.

The aggregated statistical results are summarised in Tables 5.3 and 5.4. The DRL controller sustains high resilience and consistent battery performance across all scenarios. The mean RI remains above 0.99, with a narrow standard deviation of 0.0047, confirming reliable supply continuity under stochastic fluctuations. The expected battery lifetime also



(a)



(b)

Figure 5.8: a) SOC and b) renewable generation under uncertainty across 4000 Monte Carlo scenarios. The blue shaded region denotes the 5th–95th percentile range, and the black line represents the deterministic baseline.

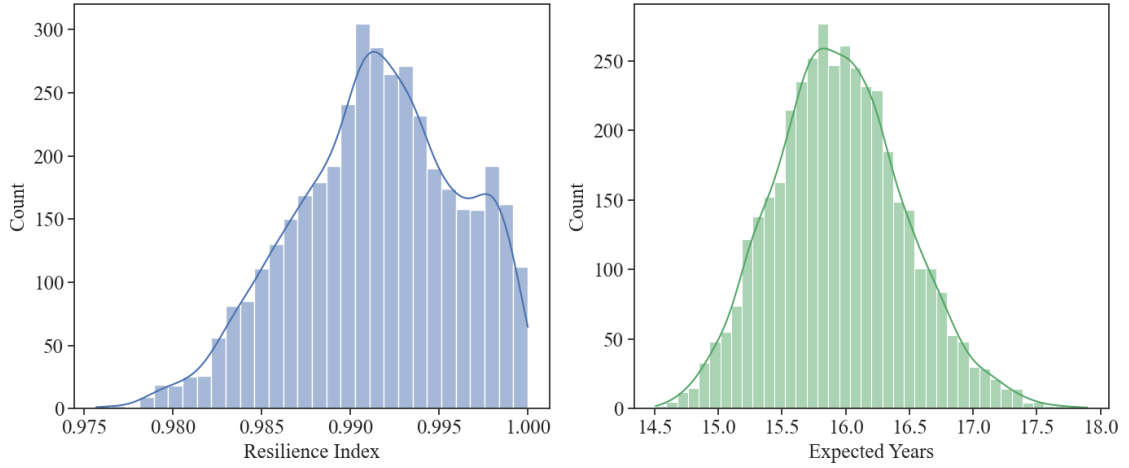


Figure 5.9: Statistical distributions of (a) Resilience Index and (b) Expected Years under 4000 Monte Carlo simulations.

Table 5.3: Statistical summary of 4000 Monte Carlo runs for DRL policy performance.

Metric	RI	EY	Average SOC	Average Renewable Generation
Mean	0.9915	15.95	0.7512	35.67
Std. Dev.	0.0047	0.5165	0.0233	1.32
Min	0.9729	14.37	0.6593	31.25
5% Percentile	0.9836	15.14	0.7090	33.44
25% Percentile	0.9884	15.59	0.7359	34.75
Median (50%)	0.9916	15.93	0.7540	35.66
75% Percentile	0.9950	16.28	0.7688	36.60
95% Percentile	0.9988	16.83	0.7849	37.81
Max	1.0000	17.81	0.8028	39.84

remains stable, averaging around 15.95 years with a standard deviation below 0.52 years, indicating that the learned control strategy generalises effectively across diverse conditions. SOC and renewable generation statistics further confirm smooth operational dynamics and preserved system stability.

Even in the least favourable scenario, the controller maintains a RI of 0.973 and an EY of 14.37 years, demonstrating robust continuity of supply and acceptable asset performance under severe renewable fluctuations. Conversely, the best cases reveal the controller’s adaptability: the highest-resilience case achieves perfect reliability ($RI = 1$) with a battery life exceeding 15 years; the best-longevity case reaches 17.81 years with 99.4% resilience; and

Table 5.4: Performance of the DRL policy in selected extreme scenarios.

Case	RI	EY
Sum-Best (RI + EY)	0.9995	17.40
Best Expected Years	0.9938	17.81
Best Resilience Index	1.0000	15.41
Worst Expected Years	0.9862	14.37
Worst Resilience Index	0.9729	15.43

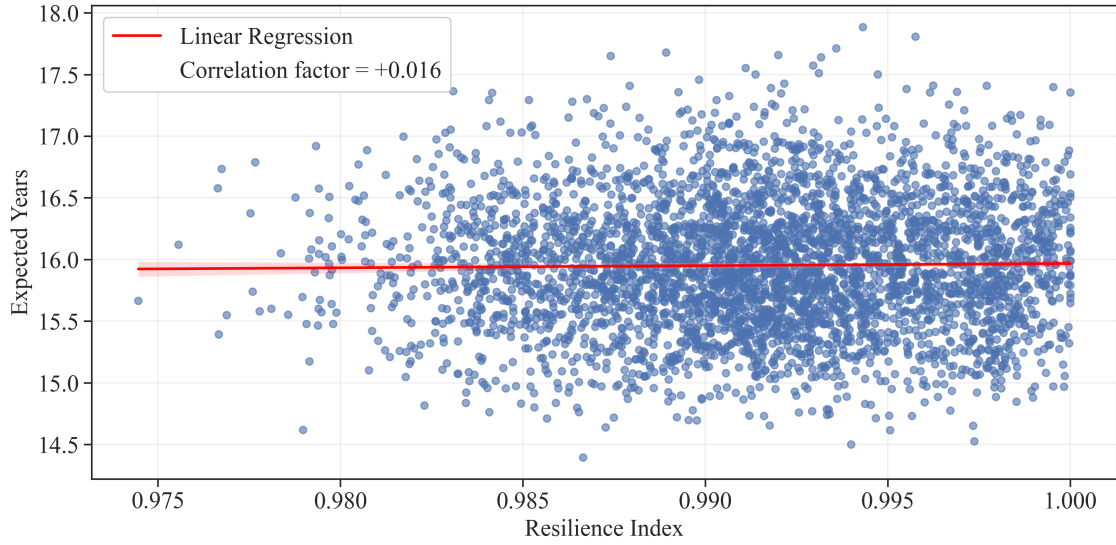


Figure 5.10: Relationship between Resilience Index and Expected Years across 4000 Monte Carlo runs.

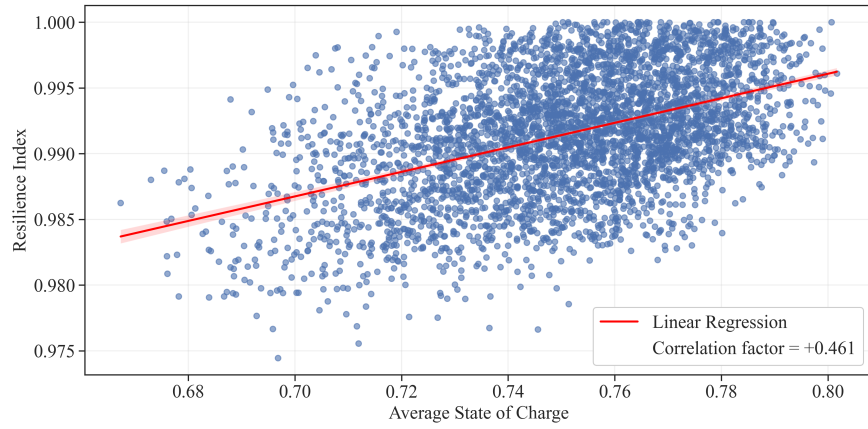
the sum-best scenario achieves an almost ideal balance with 99.95% resilience and 17.4 years of lifetime. Overall, these findings confirm that the DRL controller maintains stability, high resilience, and strong longevity across a wide uncertainty spectrum, validating its suitability for real-world microgrid operation.

Correlation Analysis. Figure 5.10 depicts the relationship between the RI and EY. Although a slight positive trend is visually observable, the overall scatter indicates a weak linear correlation, suggesting that improvements in one metric do not necessarily lead to proportional improvements in the other. This suggests that resilience and battery longevity are somewhat decoupled in performance, motivating a deeper exploration into the underlying influencing factors.

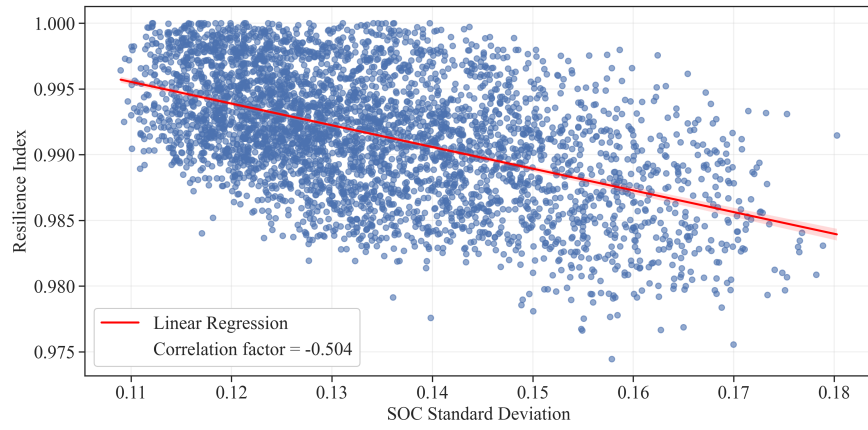
Figure 5.11 presents correlations between RI and internal variables. As shown in Figure 5.11a, higher average SOC correlates positively with higher RI, indicating that systems maintaining a consistently high charge state are more resilient. Figure 5.11b reveals a negative trend between SOC standard deviation and RI, implying that systems with more stable charge profiles tend to achieve greater resilience. In Figure 5.11c, average renewable generation shows a weak but positive association with RI, suggesting that higher renewable availability marginally improves resilience. These relationships highlight the importance of maintaining stable SOC levels and consistent renewable inputs in sustaining operational reliability.

Complementarily, Figure 5.12 examines the dependency of EY on renewable availability and SOC dynamics. Figure 5.12a demonstrates a moderate positive trend between EY and average renewable generation, confirming that greater energy availability supports extended battery life. Interestingly, Figure 5.12b shows a subtle positive tendency between EY and SOC variability, suggesting that limited cycling activity may contribute to more uniform utilisation and reduce long-term degradation. However, this effect remains weak and would require further analysis for full interpretation.

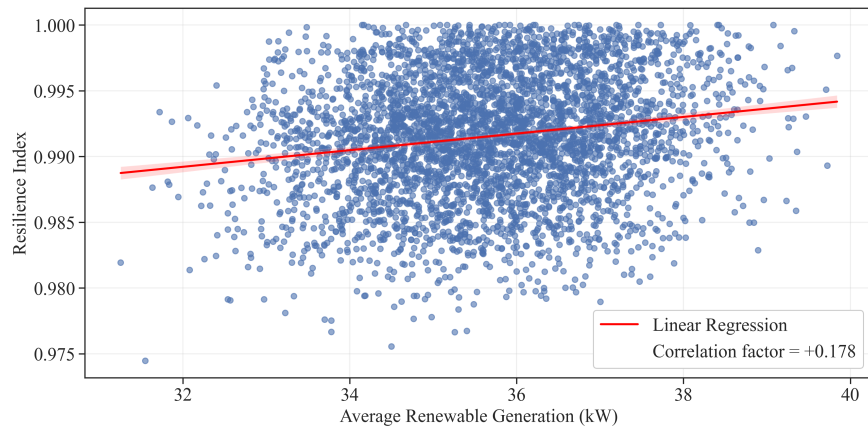
Taken together, these correlation analyses confirm that while resilience and longevity are largely independent in their quantitative behaviour, both benefit from stable SOC management and consistent renewable input. The DRL controller demonstrates the ability to balance these objectives effectively under uncertain conditions, ensuring reliable operation without explicit multi-objective retraining.



(a)

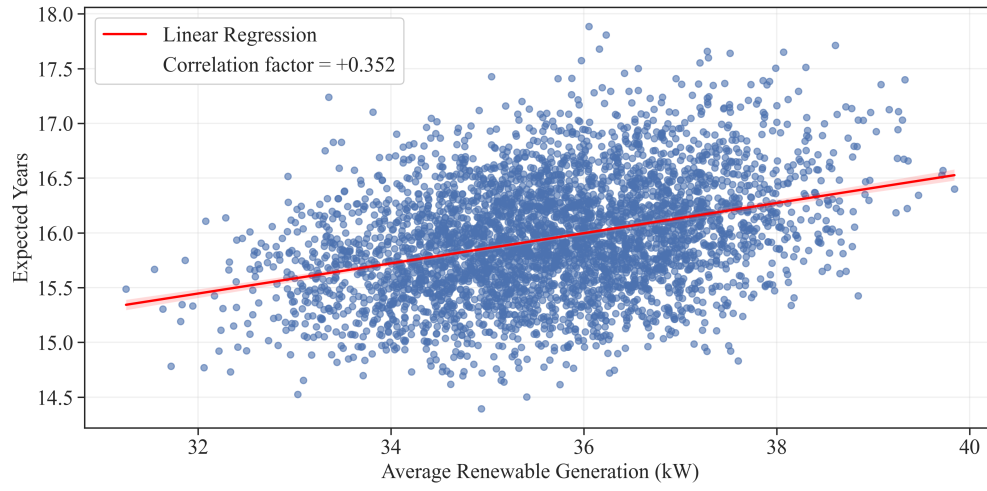


(b)

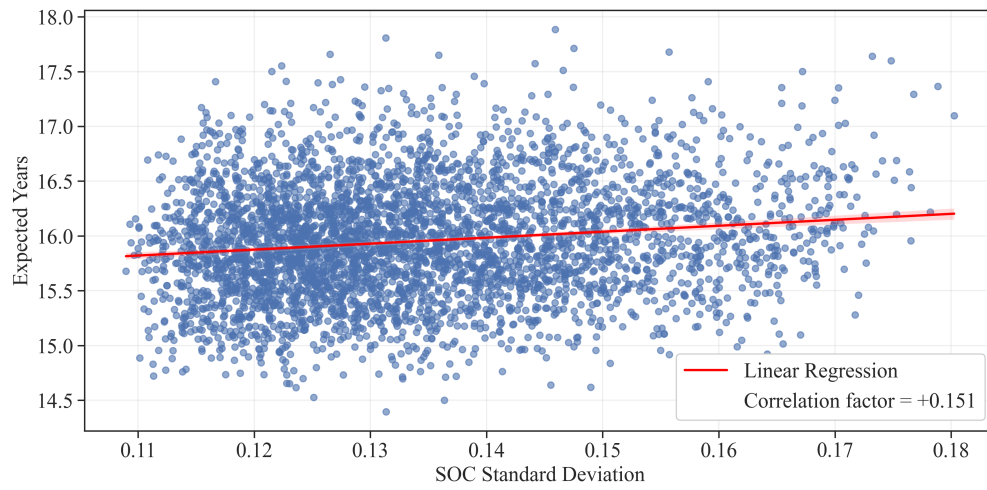


(c)

Figure 5.11: Correlation of Resilience Index with (a) average SOC, (b) SOC standard deviation, and (c) average renewable generation.



(a)



(b)

Figure 5.12: Correlation of Expected Years with (a) average renewable generation and (b) SOC standard deviation.

Objective Achieved**✓ Objectives 4 & 6 – Design, benchmark, and stress-test the DRL controller.**

These objectives asked: *To what extent can a PPO-based DRL controller enhance resilience and responsiveness compared with MPC under uncertain forecasts and HILP stress tests, and what trade-offs arise for expected battery life and computational performance?*

They have been achieved through the full design, training, and evaluation of the DRL framework. A custom **Gymnasium-based environment** was created to emulate the microgrid dynamics, and a composite **reward function** was formulated to balance resilience and battery-life terms. A **curriculum-based PPO training** strategy with data perturbations and Monte Carlo uncertainty exposure enabled robust policy learning under variable renewable and demand conditions. The trained controller was then **benchmarked against the MPC baseline** across deterministic and stochastic scenarios, including real cyclone-derived stress tests. Results showed that the DRL agent maintained near-MPC resilience, extended expected battery life, and achieved orders-of-magnitude faster inference times, confirming its suitability for **edge-level deployment**. Together, these contributions deliver a resilient, efficient, and generalisable control policy that serves as the foundation for the explainability analyses presented next.

5.4 Explainability Assessment

The preceding sections demonstrated that the trained DRL controller achieves high resilience, strong robustness under uncertainty, and performance comparable to the near-optimal MPC benchmark, while offering significantly lower computational cost. Despite these advantages, a major challenge remains: the decision-making process of deep neural networks is inherently opaque, often perceived as a “black box.” For critical applications such as microgrid

operation, it is essential not only to deliver high performance but also to explain and justify the underlying control decisions. This transparency is necessary to enhance operator trust, facilitate validation, and ensure safe deployment in real-world environments.

To address this need, this section employs XDRL, a post-hoc interpretability framework that illuminates how learned policies make decisions. The analysis comprises two complementary layers of explainability: *global interpretability*, using the SHAP method to quantify the overall contribution of each feature to control outcomes, and *local interpretability*, using the LIME method to analyse individual decision instances. The global SHAP analysis reveals feature importance patterns across all operating conditions, while the LIME analysis focuses on three representative decision modes, charging, discharging, and idle, to provide human-understandable explanations at the scenario level. Finally, insights from both methods are cross-examined to identify consistent trends and complementarity, forming a basis for assessing operator trust and safe controller deployment.

5.4.1 Global Feature Importance using SHAP

To interpret the internal reasoning of the trained DRL controller, the SHAP method was applied to both the critic and actor networks. Only the ten most influential features are shown in each plot out of the thirty available inputs, which include the current state (t) and the preceding four time steps ($t-1$ to $t-4$). SHAP quantifies each feature’s marginal contribution to the model output, providing a global understanding of how the agent weighs system variables in its decision-making process.

Critic Network Analysis: Figure 5.13 presents the SHAP summary for the critic network, which approximates the state-value function $V_\phi(s_t)$ defined in Section 3.4.3.4 (Equations 3.67–3.68). This function estimates the expected discounted return under the designed re-

ward structure. As shown, the current and recent values of SOC, Load 2, and NetEn (net energy balance) have the greatest influence on the agent’s estimation of long-term value. Higher SOC values at t , $t-2$, and $t-3$ contribute positively, indicating that maintaining a well-charged battery increases the critic’s predicted return. Because the reward function rewards resilience and battery longevity while penalising curtailment or deep cycling, higher predicted values correspond to more resilient and energy-balanced operation. Accordingly, a positive SHAP value for a feature implies that increasing that feature improves the expected return and, by extension, supports greater supply continuity and operational flexibility.

Conversely, sustained high SOC at $t-4$ contributes negatively, suggesting that prolonged overcharging without discharge may reduce efficiency or limit flexibility. Medium-priority demand (Load 2) also ranks among the key contributors, reflecting its pivotal role in balancing allocation between essential and non-essential loads. Recent renewable indicators (NetEn) show positive contributions, suggesting that the critic network associates short-term renewable surpluses with higher expected system value.

Actor Network Analysis: SHAP was further applied to the actor network to interpret operational decisions for the battery’s charging and discharging actions (Figure 5.14). For charging (Figure 5.14a), the most influential features are the current NetEn and SumRen, both showing strong positive contributions. This indicates that the agent aligns its charging behaviour with the availability of surplus renewable energy. Load 1 and Load 2 also appear among the top features, revealing that the controller accounts for immediate demand alongside renewable surplus, rather than charging indiscriminately under excess generation.

For discharging (Figure 5.14b), medium-priority demand acts as the main brake on discharge. High Load₂ values across the look-back window $(t, \dots, t-4)$ yield predominantly negative contributions, indicating that the policy intentionally limits discharge to preserve capacity for higher-priority needs. In contrast, essential demand serves as a trigger: high

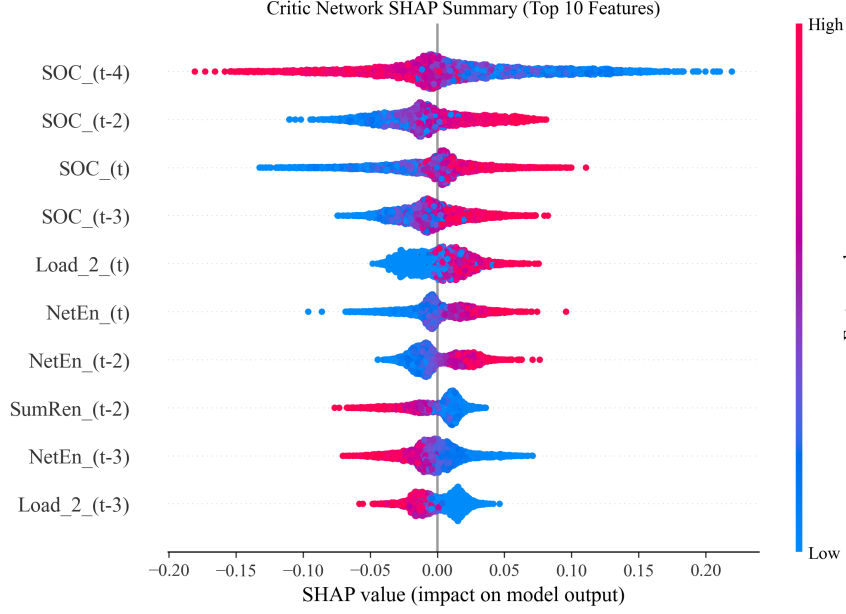
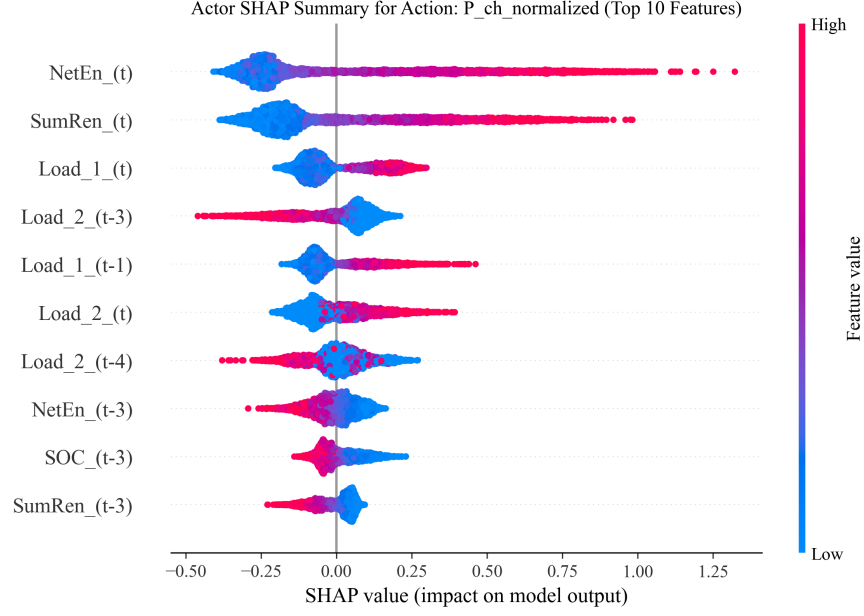


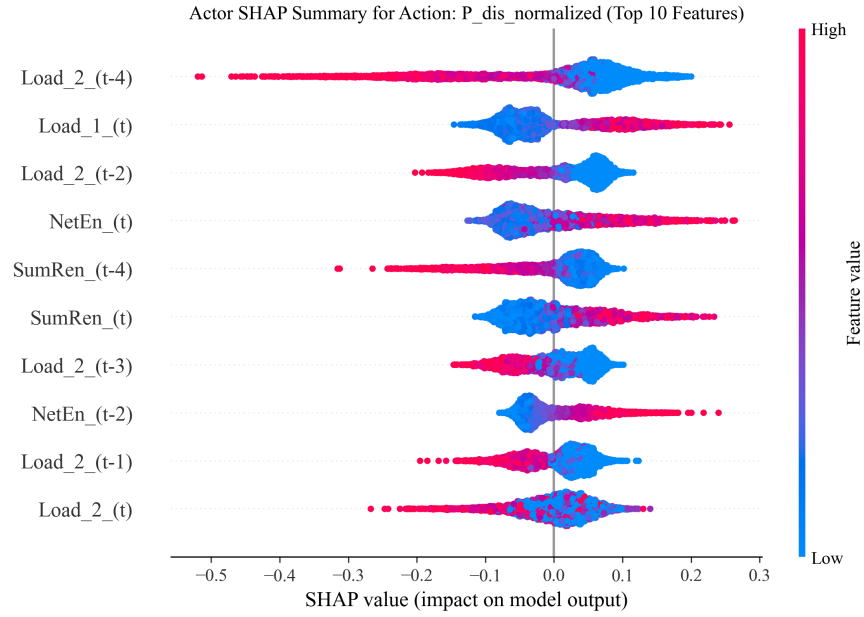
Figure 5.13: Global SHAP summary for critic network showing the top ten most influential features. Each dot represents one observation, with colour denoting the feature value and horizontal position showing its impact on the model output.

$\text{Load}_1(t)$ values produce mostly positive SHAP contributions, prompting discharge to sustain critical supply. This is consistent with Table 5.1, where the DRL agent consistently sacrifices medium-priority loads while maintaining uninterrupted essential supply. Regarding energy-balance variables, $\text{NetEn} = R_t - \sum_k L_{k,t}$ and SumRen , the SHAP results show that recent renewable increases (t and $t-2$) slightly encourage discharge, whereas older surpluses ($t-4$) discourage it. This implies that the controller primarily responds to current supply conditions, releasing stored energy when short-term surpluses occur to maintain operational flexibility. However, when compared with charging SHAP values in Figure 5.14a, the influence of renewable-related features on discharge remains weaker (amplitudes ≈ 0.2 – 0.3 vs. 1.0 – 1.25), confirming that renewable surpluses predominantly drive charging, while discharge is used selectively for essential-load support and capacity maintenance.

Supply–Allocation Brains: The SHAP analysis for supply allocation weights (Figure 5.15) shows that Load 2 consistently ranks among the top features across all load tiers,



(a) Charging action



(b) Discharging action

Figure 5.14: Global SHAP explanations for (a) charging and (b) discharging actions in the actor network.

confirming its function as the main balancing layer mediating energy allocation between essential and lower-priority demands.

For essential loads (Ps_1), high $Load_1$ at the current and recent time steps yields strong positive SHAP contributions, indicating reliable prioritisation of critical demand. High $Load_2(t)$, however, has a negative influence on Ps_1 , suggesting that the policy restricts allocation to essentials when medium-priority demand is simultaneously high, whereas older $Load_2(t-4)$ shows a mild positive effect consistent with delayed balancing.

For business loads (Ps_2), recent values of $Load_2$ at $t, t-1, t-2$ show strong positive impacts, meaning that the controller supplies business demand when its requirement persists over time. In contrast, high $Load_1(t)$ contributes negatively, indicating that essential demand takes precedence over business supply. Positive contributions from $NetEn(t)$ and $SumRen(t)$ confirm that business loads are served when short-term renewable availability is high.

For agricultural loads (Ps_3), SHAP magnitudes are much smaller, and $Load_3$ itself is absent among the top features, indicating minimal direct influence. Instead, Ps_3 depends primarily on the states of $Load_1$, $Load_2$, and renewable indicators ($NetEn$, $SumRen$), showing that agricultural supply is largely conditional on surplus conditions once higher-priority demands are met. These SHAP patterns collectively confirm that the DRL policy reproduces the intended priority hierarchy: maintaining essential reliability, serving business loads under balanced conditions, and allocating residual energy to agricultural loads during renewable surpluses.

In summary, the SHAP-based global interpretation confirms that the DRL agent’s control logic is both physically coherent and aligned with the intended operational priorities: maintaining essential supply, preserving battery health, and utilising renewable surpluses efficiently. However, SHAP primarily provides aggregate insights over the entire dataset and does not reveal how individual decisions evolve under specific operating conditions.

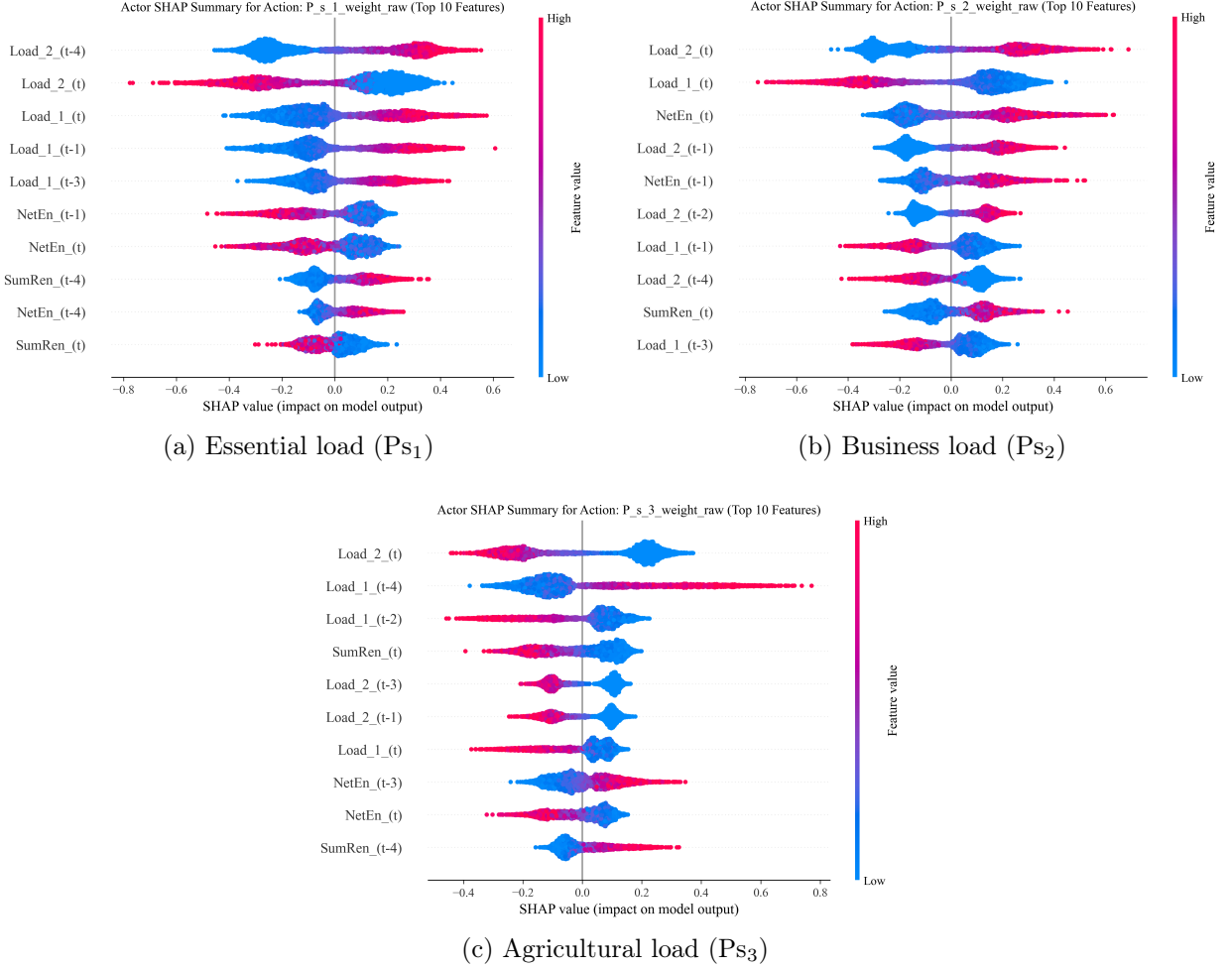


Figure 5.15: Global SHAP explanations for the weighting of supply decisions across (a) essential, (b) business, and (c) agricultural loads. Only the top ten most influential features are shown.

To address this limitation, the following subsection applies LIME to examine local decision pathways for representative cases of charging, discharging, and idle operation, offering complementary insight into the controller's behaviour in real time.

5.4.2 Local Decision Interpretation using LIME

To complement the global interpretation provided by SHAP, the LIME algorithm was applied to explain individual decisions of the trained DRL controller. Three representative time steps

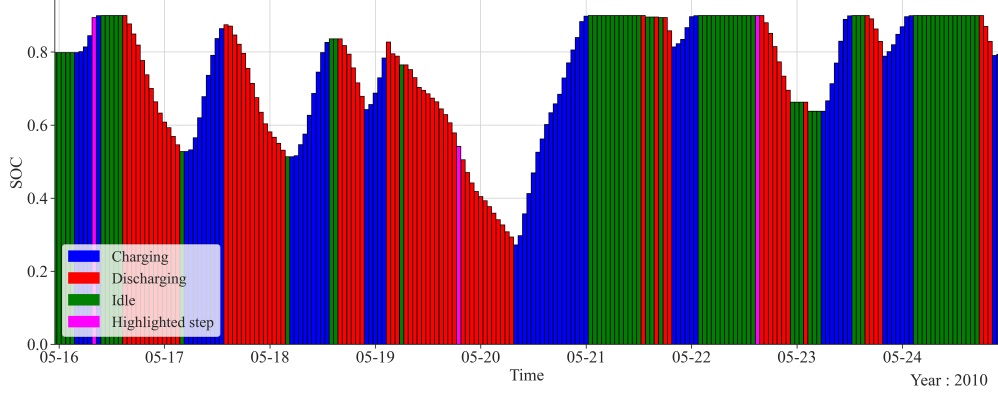


Figure 5.16: SOC trajectory showing selected instances for charging (hour 13), discharging (hour 94), and idle (hour 162) actions.

were selected to capture the main operational modes of the battery: charging, discharging, and idle. In the control framework, the agent manages five continuous actions, two for charging and discharging power, and three scalar weights that determine the allocation of available supply to essential, business, and agricultural loads. By examining these specific decision instances, LIME reveals how local input conditions influence the agent’s real-time choices.

Representative Instances: Figure 5.16 illustrates the battery SOC trajectory, where operating modes are distinguished by colour: blue for charging, red for discharging, and green for idle. The three highlighted points in pink correspond to the representative decisions analysed using LIME, charging (hour 13), discharging (hour 94), and idle (hour 162). These instances provide a clear view of how the controller responds to varying energy availability and demand profiles.

Charging Decision (Hour 13): Figure 5.17 shows the local explanations for the charging instance. The charging action (Figure 5.17a) is driven by high current and recent values of NetEn and SumRen ($t, t-2$), confirming that the agent initiates charging when renewable surplus is available. Conversely, past high loads, particularly Load 2, discourage charging, reflecting anticipation of future demand. The discharging action (Figure 5.17b) exhibits

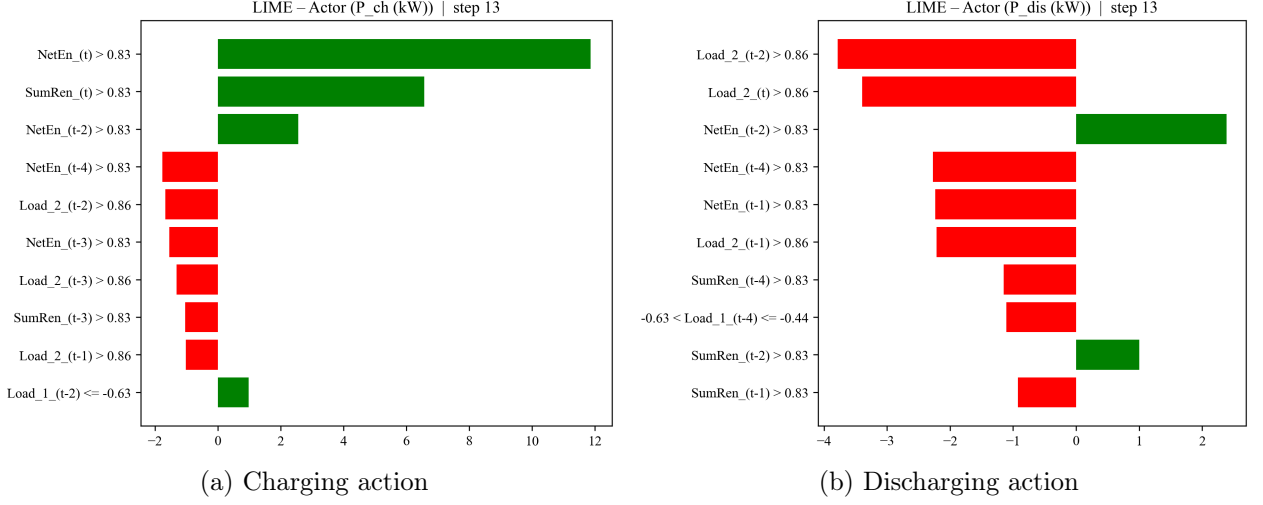


Figure 5.17: LIME explanations for charging instance (hour 13): (a) charging and (b) discharging actions.

mainly negative contributions, confirming that the agent finds no incentive to release energy during surplus conditions.

The corresponding supply weights are shown in Figure 5.18. Among the three load tiers, the business load (Ps_2) receives the highest allocation weight (0.682), followed by essential (0.209) and agricultural (0.109) loads. This pattern aligns with the current system state, where essential demand is low and business demand is moderate, allowing the controller to exploit available renewable energy while ensuring balanced supply distribution.

Discharging Decision (Hour 94): At hour 94, the agent switches to discharging mode (Figure 5.19). The discharging action (Figure 5.19b) is strongly encouraged by high essential demand (Load 1) and low NetEn, indicating a shortage of renewable supply. In this case, the agent responds by releasing stored energy to maintain resilience. The corresponding charging action (Figure 5.19a) is dominated by negative influences from both NetEn and SumRen, together with past high load levels, all of which discourage charging under supply-deficit conditions.

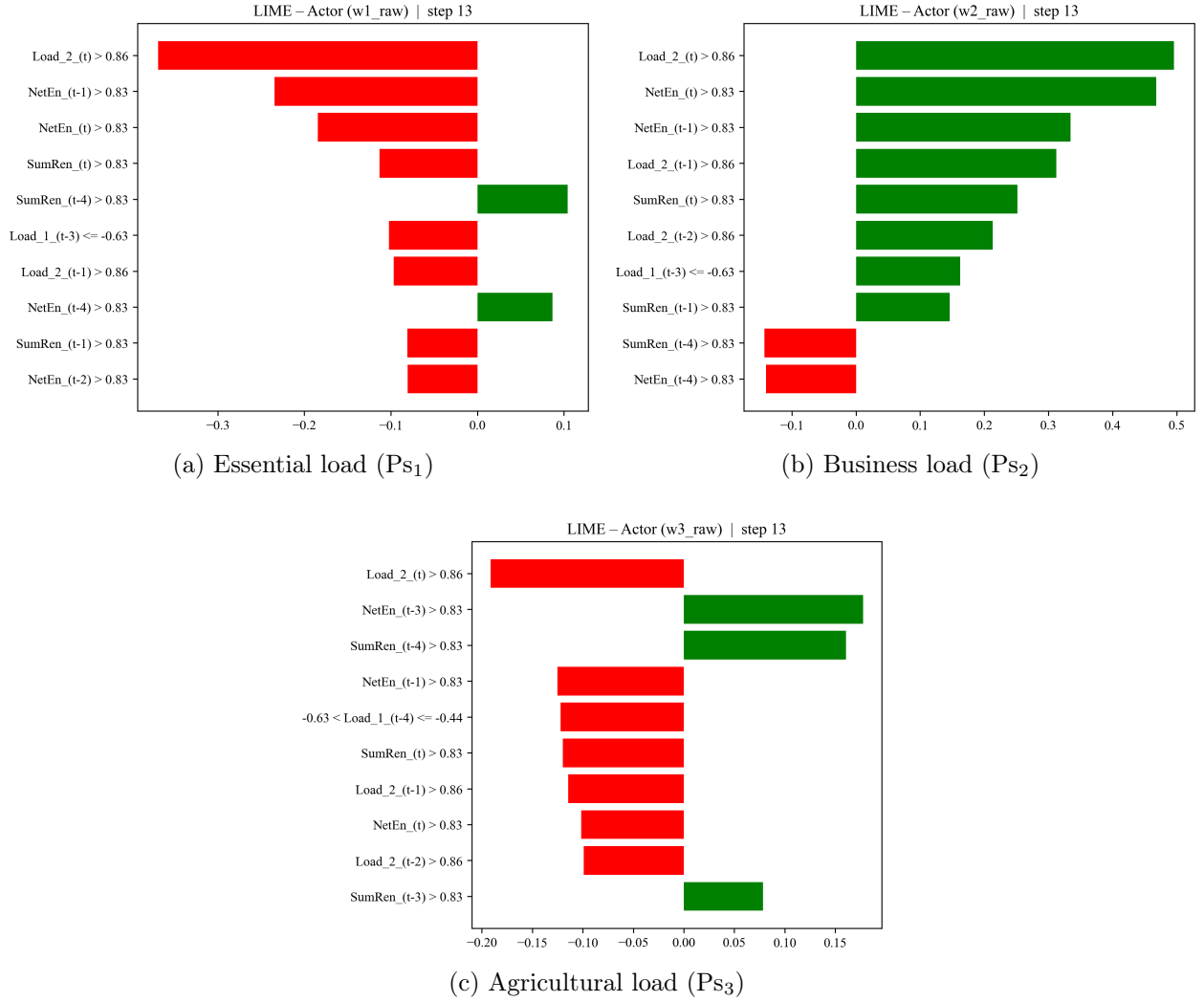


Figure 5.18: LIME explanations for charging instance (hour 13): supply weights for (a) essential, (b) business, and (c) agricultural loads.

Idle Decision (Hour 162): In the final instance (Figure 5.20), the agent remains idle, avoiding both charging and discharging operations. Figure 5.20a shows that charging is discouraged by low NetEn and limited renewable generation, while Figure 5.20b indicates that discharging is simultaneously disincentivised by recent high loads and the absence of surplus energy. This balanced discouragement leads to an idle state, suggesting that the agent opts to preserve battery capacity until more favourable conditions arise.

Discussion: The LIME results confirm that the DRL agent exhibits context-aware and

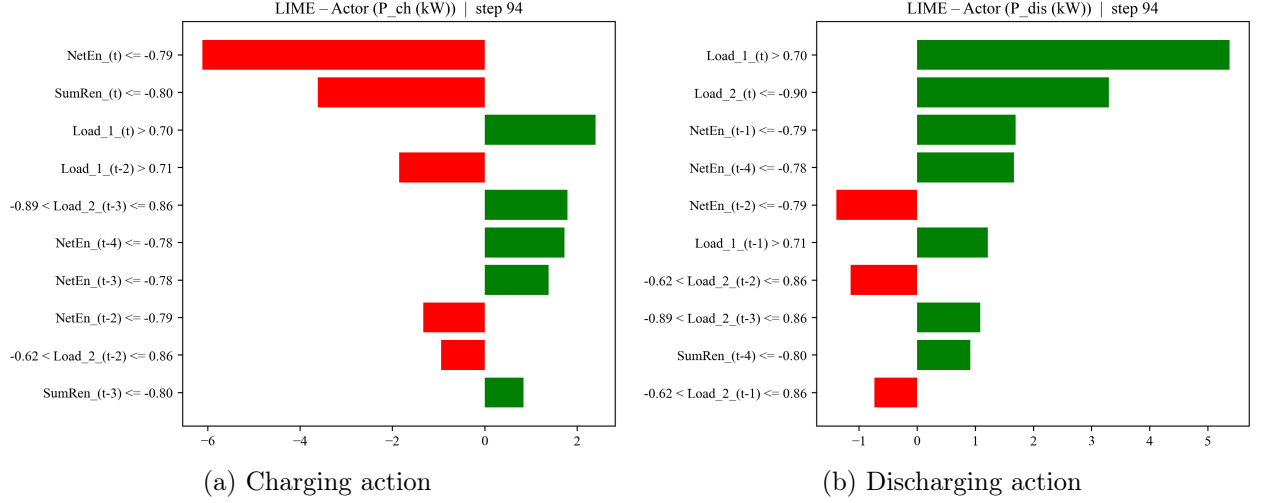


Figure 5.19: LIME explanations for discharging instance (hour 94): (a) charging and (b) discharging actions.

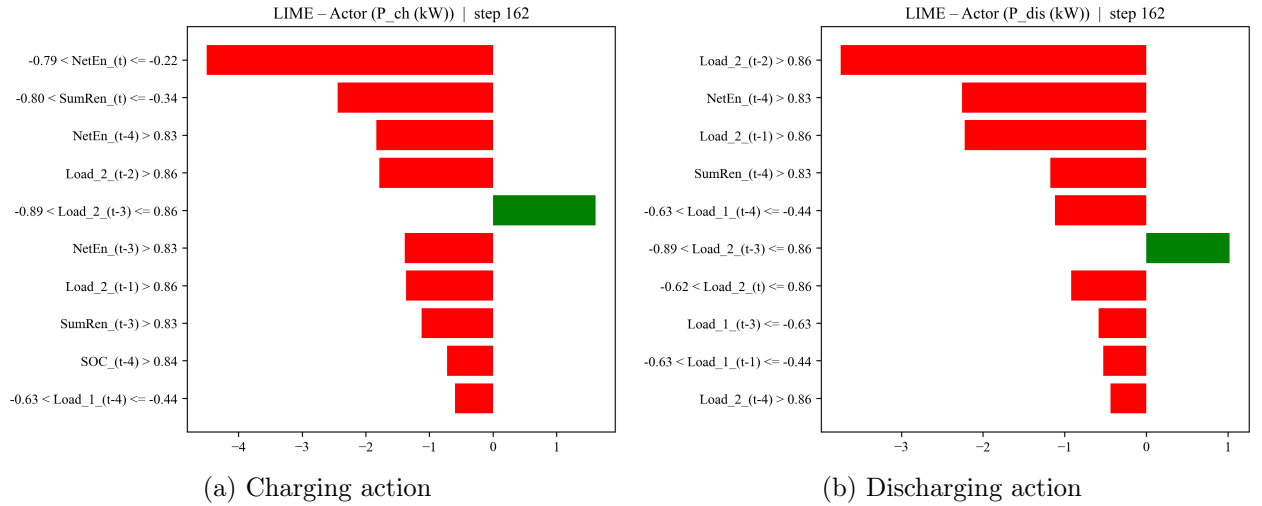


Figure 5.20: LIME explanations for idle instance (hour 162): (a) charging and (b) discharging actions.

physically consistent decision-making across all operational modes. Charging occurs primarily under renewable surplus, discharging under supply deficit driven by essential demand, and idle operation under neutral conditions where both charging and discharging are unfavourable. These local interpretations complement the SHAP findings by revealing how the agent translates global learning into specific, situation-dependent actions. This enhances trust in its reliability for real-world microgrid operation.

5.4.3 Cross-Validation between SHAP and LIME Insights

To verify the reliability of the post-hoc explanations, this subsection cross-checks the global attributions from SHAP with the local explanations from LIME. Agreement between the two methods provides stronger evidence that the agent’s policy follows consistent internal logic rather than coincidental correlations.

Integrated comparison. Figure 5.21 summarises the alignment between the two explainability methods. Cell colours indicate the direction of influence: green for encouragement, red for discouragement, purple for mixed effects, and grey for neutrality. Darker shades represent higher importance. The comparison focuses on four main feature groups identified earlier: short-term energy balance (NetEn and SumRen) and tiered demand levels (Load₁, Load₂) across recent time lags. For reference, the original SHAP and LIME attribution plots are shown in Figures 5.14a–5.15c and Figures 5.17–5.20, with the value-function (critic) importance in Figure 5.13. The figure should therefore be read as a compact synthesis of the preceding SHAP and LIME results rather than as a separate standalone analysis.

Interpretation Based on Figure 5.21

Figure 5.21 allows a direct comparison between the global and local interpretation results.

Case	NetEn	SumRen	Load ₁	Load ₂
SHAP (global)				
Charging	$\uparrow (t)$	$\uparrow (t)$	$\downarrow (t \text{ low})$	$\uparrow (t)$ $\downarrow (t-3)$
Discharging	$\downarrow (t \text{ low})$	$\downarrow (t \text{ low})$	$\uparrow (t)$	$\uparrow (t \text{ low})$ $\uparrow (t-4 \text{ low})$
Idle	—	—	—	—
P _{s1} (essential)	$\downarrow (t)$	$\downarrow (t)$ $\uparrow (t-4)$	$\downarrow (t \text{ low})$	$\downarrow (t)$
P _{s2} (business)	$\uparrow (t)$	$\uparrow (t)$	$\uparrow (t \text{ low})$	$\uparrow (t)$
P _{s3} (agric.)	$\uparrow (t-3)$	$\downarrow (t)$ $\uparrow (t-4)$	$\downarrow (t)$	$\downarrow (t)$
LIME (local)				
Charging (h13)	$\uparrow (t)$	$\uparrow (t)$	$\uparrow (t \text{ low})$	$\downarrow (\text{lags})$
Discharging (h94)	$\uparrow (\text{lags low})$	$\uparrow (t-4 \text{ low})$	$\uparrow (t)$	$\uparrow (t \text{ low})$
Idle (h162)	—	—	—	—
P _{s1} (essential)	$\downarrow (t-1)$ $\downarrow (t)$	$\downarrow (t)$ $\uparrow (t-4)$	$\downarrow (t-3 \text{ low})$	$\downarrow (t)$
P _{s2} (business)	$\uparrow (t)$	$\uparrow (t)$	$\uparrow (t-3 \text{ low})$	$\uparrow (t)$
P _{s3} (agric.)	$\uparrow (t-3)$ $\downarrow (t)$	$\downarrow (t)$ $\uparrow (t-4)$	$\downarrow (t-4)$	$\downarrow (t)$

Figure 5.21: Qualitative alignment between SHAP and LIME.

Charging case. Both methods show that NetEn and SumRen are the most influential features, with high positive values clearly promoting charging. Load₂ lags have moderate negative effects, indicating that the agent avoids charging when medium-priority demand is likely to increase. A small difference appears for Load₁: SHAP shows a medium negative impact, while LIME shows a small positive one. This is because SHAP represents the overall trend where essential demand, even if small, still discourages charging to maintain readiness for critical supply. LIME, however, considers the specific moment when essential demand is very low and renewable surplus is high, allowing charging to proceed. Overall, both analyses confirm that renewable balance mainly drives the charging decision, while demand-related

features fine-tune it to maintain system stability.

Discharging case. In both methods, discharging is mainly influenced by high Load_1 and low Load_2 values, which reflects the priority of essential demand over business loads. The energy-balance features (NetEn and SumRen) show some differences: SHAP links their low current values (t) with reduced renewable availability, while LIME highlights their earlier lags, showing that previous low values lead the agent to discharge. Despite this time shift, both methods agree that low renewable conditions and high essential demand together trigger discharging. SHAP also indicates that low NetEn reduces the likelihood of charging more strongly than it increases discharging, confirming that the agent reacts to low energy availability by releasing stored power.

Idle case. The idle state appears when the agent decides neither to charge nor to discharge. This behaviour is not an explicit action in the model but an outcome of the decision process. As SHAP is based on fixed actions, it cannot display this behaviour directly, whereas LIME identifies it at the time step where both charging and discharging are discouraged. The grey cells in Figure 5.21 therefore represent a neutral state in which neither operation is suitable.

Supply-allocation cases. The supply-allocation results (Ps_1 – Ps_3) show consistent trends between SHAP and LIME. For Ps_1 , high Load_1 values increase the allocation to essential loads, while high Load_2 values reduce it, confirming that medium-priority demand competes for available power. In Ps_2 , high Load_2 values are the main positive driver, indicating that strong business demand raises its share of supply. For Ps_3 , Load_3 has little influence, and its allocation depends mostly on Load_1 , Load_2 , and renewable conditions. This confirms that agricultural loads are supplied only when there is surplus energy after meeting higher-priority needs.

5.4.4 Implications for Operator Trust and Responsible Deployment

The agreement between SHAP and LIME indicates that the trained controller relies on physically meaningful signals, including SOC, short-term energy balance, renewable availability, and tiered demand. These findings support interpretability and provide useful diagnostic evidence for future operator-facing implementation. However, the explainability analysis remains post-hoc and simulation-based; therefore, it should be interpreted as a step towards operational transparency rather than as a complete validation of operator trust or field readiness.

5.4.4.1 Safeguards considered in this thesis

Several safeguards are represented within the controller design and simulation-based evaluation. First, the environment enforces SOC bounds, converter limits, and load-priority rules, reducing the risk of infeasible actions within the modelled operating conditions. Second, battery-aware behaviour is encouraged through the reward formulation and staged training process, leading to smoother SOC trajectories and reduced cycling compared with the MPC benchmark. Third, post-hoc explainability using SHAP and LIME provides consistent evidence that the learned policy follows physically interpretable patterns. Finally, robustness is empirically tested using Monte Carlo simulations with Perlin-noise perturbations, which provides evidence of stable performance under the uncertainty scenarios considered in this thesis.

5.4.4.2 Practical guidelines and future extensions

These items are not implemented in this thesis but follow directly from the findings and may guide future deployments.

- **Human-in-the-loop visibility.** Provide concise, real-time summaries of dominant decision drivers (e.g., “charging due to high $\text{NetEn}(t)$ and $\text{SumRen}(t)$; discharge discouraged by medium-priority demand”) to enhance operator situational awareness.
- **Drift and Out-of-Distribution (OOD) detection.** Monitor key inputs (SOC, NetEn , SumRen , load tiers) for persistent deviations from the training range; upon detection, slow updates or revert to a validated backup policy (e.g., MPC baseline).
- **Explanation-based logic checks.** Verify that explanations remain physics-consistent: Ps_1 rises with Load_1 ; renewable surplus does not trigger discharge; medium-priority demand does not override essentials. Require operator confirmation if violated.
- **Interpretability health metrics.** Track LIME fidelity and SHAP ranking stability to detect conceptual drift; use trends to trigger inspection or retraining.
- **Lifecycle governance.** Maintain action/state logs and top explanatory factors for audit; schedule periodic Monte Carlo re-tests and SHAP-trend reviews to ensure continued alignment with operational priorities.

Operator-facing acceptance criteria. For prospective deployments, three acceptance criteria naturally follow from the analysis. Essential-load curtailment should be absent under normal operation, with any deviation fully explained by physical constraints. During renewable surplus, explanations should primarily justify charging actions, whereas under deficit conditions, discharging should be driven by essential demand. Finally, medium-priority loads should act as the principal balancing mechanism, with their modulation justified by short-term supply–demand conditions.

Limitations and responsible use. Post-hoc explanations are diagnostic rather than causal. SHAP assumes additive effects, and LIME relies on local surrogates whose fidelity

varies by instance. These explanations should therefore complement, not replace, hard operational constraints, out-of-distribution monitoring, fallback control, and operator review. The guidelines proposed here should be understood as future deployment requirements rather than fully implemented field safeguards.

Summary. Within the scope of this thesis, explainability helps show that the learned policy is aligned with power-system intuition and the designed load-priority structure. The proposed governance measures provide a practical route for future validation, but real deployment would require hardware-in-the-loop testing, operator assessment, and site-specific safety verification.

Objective Achieved

✓ Objectives 5 & 7 – Provide transparency and extract operational insight.

These objectives asked: *How interpretable and practically useful are DRL-based energy-management decisions in isolated microgrids, and to what extent can explainability techniques support operator trust and governance?*

They have been achieved through the development of a comprehensive **explainability framework** combining SHAP (global) and LIME (local) analyses for the trained DRL controller. Post-hoc explanations were generated for all key actions, charging/discharging, load allocation, and idle states, and cross-validated to ensure consistency between global trends and instance-level reasoning. The resulting attributions revealed **physically meaningful control logic**: renewable surplus promotes charging, essential demand triggers discharging, and medium-priority demand moderates both to protect storage capacity. From these patterns, **operator-facing rules and acceptance criteria** were derived, defining expected decision behaviour under surplus and deficit conditions.

... Objective Achieved

To translate interpretability into operational reliability, **safety and governance layers** were formalised, including SOC and inverter limits, explanation-based logic checks, out-of-distribution monitoring, and an MPC fallback policy. Finally, **lifecycle procedures** (i.e. record-keeping, explanation-fidelity tracking, and periodic Monte Carlo retesting) were proposed to sustain transparency and trust through deployment.

Overall, the explainability analysis provides diagnostic evidence that the trained DRL controller follows physically meaningful decision patterns. This supports future development of operator-facing monitoring and governance tools, but further validation is required before deployment in real microgrid operation.

5.5 Chapter Summary

This chapter presented the XDRL framework for resilient and battery life-aware microgrid operation. A PPO-based controller was trained using curriculum learning and benchmarked against the MPC reference controller, achieving comparable resilience, smoother control actions, and improved expected battery lifetime under the tested scenarios. Monte Carlo analysis provided empirical evidence of robustness under Perlin-noise renewable perturbations, while SHAP-LIME explainability identified physically meaningful decision drivers. Together, these results show that the proposed XDRL framework combines high technical performance with interpretable decision behaviour, while further validation is required before field deployment.

Chapter 6

Conclusion and Future Work

6.1 Conclusion

This study set out to answer a simple but high-stakes question: *Can a rural microgrid exposed to high-impact, low-probability events be operated in a way that preserves essential supply, protects battery life, and remains computationally practical for real deployments?* Across planning, optimisation, learning, uncertainty quantification, and explainability, the evidence says yes, and shows precisely how.

Beginning with planning, HOMER Pro produced a feasible configuration whose technical behaviour was reproduced by an MILP-MPC benchmark in PYOMO. That benchmark mattered: it gave a transparent, constraint-aware reference against which any “smart” controller would have to compete. Validated over a demanding 10-day window, the MPC matched demand fulfilment and revealed a crucial insight: objective weighting governs whether the battery is over-cycled for short-term resilience or preserved at the cost of service. The balanced case, equal weights on resiliency, total imbalance, switching cost, and life-cycle degradation, achieved an RI near unity while keeping EY close to 14–15 years and avoiding unnecessary cycling; overly single-minded tunings either sacrificed resilience (when saving the battery at all costs) or wore the battery prematurely (when chasing imbalance or resiliency alone). From there the work stepped closer to deployment realities.

Modest DR (10% on essential/business tiers) flattened peaks and nudged EY upward without compromising resilience; a smart resilient controller that adapts SOC bounds and objective priorities based on a live ERI signal kept the lights on even when storage was cut to 25% of its design size ($\text{ERI} \gtrsim 0.95$), a practical lever for communities with budget or supply-chain constraints. Coupling to the grid and explicitly valuing profit transformed “energy balancing” into cost-aware scheduling: profit rose by about 136% and EY extended (e.g., to ~ 27 years in the full-capacity case) while maintaining uninterrupted supply, showing that markets, when embedded into the control objective, can strengthen both resilience and asset health.

Against that benchmark, the PPO controller, trained with curriculum learning and stress-tested with Perlin-noise uncertainty, delivered near-optimal resilience and superior longevity with dramatically lower compute. Under deterministic Cyclone Laila conditions the DRL policy’s RI (0.9961) trailed the best MPC (0.9991) by only $\sim 0.3\%$, yet EY improved from 14.44 to 15.90 years; under uncertainty the pattern held (RI 0.9956 vs. 0.9989; EY 15.88 vs. 15.06 years). Critically, DRL *learned* the intended priority structure: essential load was always protected; business and agricultural tiers acted as the flexibility margin. Derivative-based SOC metrics quantified why EY improved: relative to MPC, DRL reduced mean and Root Mean Square (RMS) of $|d\text{SOC}/dt|$ by $\sim 12\%$ and $\sim 19\%$, and mean and RMS of $|d^2\text{SOC}/dt^2|$ by $\sim 26\%$ and $\sim 38\%$. These reductions result in smoother dispatch, lower degradation stress, and a longer expected battery lifetime.

That battery-aware behaviour came with another, often decisive, advantage: computational efficiency. DRL inference required $>10^3$ times fewer FLOPs per step than MPC (more than a $5,000\times$ reduction), making real-time control and large-scale robustness testing feasible on commodity hardware; even accounting for one-off training, lifetime compute cost was about three times lower than MPC. This efficiency enabled a 4,000-scenario Monte Carlo analysis: RI clustered tightly around 0.99 ($\text{SD} < 0.005$) and EY around 16 years ($\text{SD} \sim 0.5$),

with worst cases still acceptable ($RI \sim 0.973$; $EY \sim 14.4$) and best cases achieving perfect resilience or $EY \sim 17.8$ years.

Because a controller that cannot be explained cannot be trusted in critical infrastructure, the work also treated explainability as a first-class requirement. SHAP (global) and LIME (local) converged on physically meaningful drivers: charging is encouraged by short-term renewable surplus (NetEn, SumRen), discharging is triggered by essential demand under deficits, and business-tier demand regulates both to protect storage capacity; supply allocation respects the designed priority order. Coupled with hard safety envelopes (SOC bounds, inverter limits, essential-load reserves), drift/OOD checks, and an MPC fallback, this explainability layer offers a pathway toward operator trust.

These findings matter beyond the case study. They show that a microgrid need not choose between resilience, battery longevity, and practical deployability: with the right combination of predictive optimisation for validation and explainable learning for execution, it can have all three. They also suggest a policy and design implication: embedding market signals and modest DR into the control objective can deliver simultaneous gains in reliability, cost, and asset life, an especially relevant lever for rural and resource-constrained settings where every cycle and every watt matters.

At the same time, the study’s scope and limitations are clear. The MILP relied on linearised battery terms and recomputed longevity with a nonlinear model, errors were small ($< 5\%$) but merit validation against richer calendar plus cycle ageing models. Perlin noise captured smooth, correlated forecast errors but not all meteorology; broader stochastic weather generators and price uncertainty would further harden conclusions. Tariffs were simplified; dynamic or uncertain markets could change behaviour. Results were proven on a single site; multi-site datasets, human in the loop, and field trials are the natural proving ground.

Taken together, this thesis shows that explainable DRL can provide a promising alternative to optimisation-based control for resilient, battery-aware microgrid operation. The framework achieves resilience close to the MPC benchmark while improving expected battery lifetime and reducing online computational burden in the tested case study. However, the contribution should be understood as a simulation-based framework rather than a field-ready controller. The results should therefore be interpreted as promising but bounded by the scope of this thesis, since only a single case-study microgrid and one recorded HILP event, Cyclone Laila, have been assessed using simulation-based tools. Further validation using multi-site data, richer battery ageing models, hardware-in-the-loop testing, and operator studies is required before practical deployment.

6.2 Limitations and Future Directions

6.2.1 Limitations

Despite its comprehensive scope, this study is not without limitations. The findings were derived from a single cyclone-prone coastal site in India, where the specific mix of renewable resources and demand characteristics shaped both the system design and controller performance. While the DRL environment normalises data to enhance transferability, any application to other regions would require careful re-sizing and tuning to reflect local resource profiles and demand dynamics.

The renewable generation data used in both planning and operation were reconstructed from irradiance and wind records rather than direct power measurements, introducing minor estimation uncertainties. In the MPC formulation, battery ageing was represented through a piecewise BLC-DOD relationship; temperature effects, calendar degradation, and

rate-dependent phenomena were simplified to maintain linearity. These aspects may lead to small deviations from full electrochemical ageing models and could be refined in future work.

The operational layer primarily addressed supply–demand balance, load prioritisation, and storage management. Network-level characteristics such as voltage regulation, reactive power control, line constraints, and protection coordination were abstracted. Similarly, the DRL agent was trained entirely offline using historical and synthesised scenarios; safe online adaptation, transfer learning, or continual learning during deployment remain outside the present scope.

A related limitation concerns controller adaptability when modelling assumptions change. In the MPC baseline, changes to the battery model, degradation curve, SOC limits, converter rating, or load-priority structure can generally be incorporated by revising the corresponding equations, constraints, or objective-function terms and then re-solving the optimisation problem. This makes MPC relatively transparent and structurally adaptable when model assumptions change, although online optimisation may remain computationally demanding. By contrast, the trained DRL policy provides fast inference after training, but its behaviour is tied to the environment, reward function, and degradation model used during training. Therefore, substantial changes in the battery technology, BLC–DOD relationship, component sizing, or operating context would require further validation and may require policy fine-tuning or retraining before the controller could be safely reused.

Uncertainty modelling was limited to structured perturbations and correlated forecast errors generated by Perlin noise around the Cyclone Laila dataset. Although this approach provided a realistic stress test, it does not capture the full complexity of meteorological or market variability. Tariffs were treated as deterministic, and dynamic or uncertain market signals were not considered. Finally, explainability was assessed using post-hoc SHAP and LIME analyses conducted offline. While these methods clarified decision logic, they offer

attribution rather than causality and were not validated through real operator studies.

The computational analysis was performed on general-purpose desktop hardware under ideal communication conditions. Further work is required to evaluate performance on embedded or industrial edge controllers, where limited processing power, latency, and data loss may influence responsiveness.

Overall, these limitations outline clear directions for future research. Building on the findings of this thesis, several promising extensions are identified below.

6.2.2 Future Directions

Future research should advance the framework along three complementary directions, **physical realism**, **learning adaptability**, and **human interpretability**, to bridge the gap between simulation and field deployment.

1. **Hardware-in-the-loop and field validation:**

Experimental implementation on a real microgrid or a HIL setup would provide the most rigorous validation of the proposed framework. Integrating programmable renewable emulators, battery systems, and communication interfaces would enable assessment of latency, sampling delays, and controller robustness under real-time constraints.

2. **Network-aware and multi-microgrid operation:**

Future studies could couple the controller with detailed power-flow models to capture voltage regulation, reactive power dynamics, and line limits. Extending the framework to multi-microgrid or peer-to-peer configurations would allow cooperative resilience and resource sharing, supporting community-level energy autonomy.

3. **Advanced battery ageing models:**

Incorporating temperature-dependent, calendar, and rate-capacity degradation terms into both MPC and DRL formulations would improve life-cycle estimation accuracy. Coupling these models with real-time battery health diagnostics could enable predictive maintenance and adaptive ageing-aware control.

4. Online learning and transferability:

Developing safe online or lifelong learning mechanisms would allow the DRL agent to adapt to changing renewable patterns, demand behaviour, tariff conditions, and battery ageing characteristics without full retraining from scratch. However, in power-system applications, online learning should not be implemented as unrestricted trial-and-error control. Future work should therefore investigate constrained policy updates, out-of-distribution detection, digital-twin validation, and fallback control using the previous validated policy, rule-based logic, or MPC whenever safety limits are approached.

Transferable DRL is also needed to extend the proposed framework beyond the single case-study microgrid. A policy trained on one site could be adapted to another using transfer learning, domain adaptation, domain randomisation, curriculum learning, or fine-tuning with local data. For example, the policy may learn general relationships between renewable surplus, SOC, battery dispatch, and load priority, while site-specific layers or reward parameters could be updated for a new battery model, component rating, weather regime, or load structure. Such transfer would reduce training effort, but it would still require site-specific validation before deployment because changes in battery degradation behaviour or HILP event characteristics may alter the resilience–battery-life trade-off.

5. Uncertainty and market integration:

Incorporating stochastic weather generators and dynamic pricing would strengthen economic and operational realism. This extension would enable the exploration of resilience–profitability trade-offs and demonstrate how market signals can support sus-

tainable operation under uncertainty.

6. Human-centred explainability:

Future research should extend explainability beyond algorithmic analysis toward practical, operator-oriented tools. Interactive visual dashboards or simplified explanation reports could help operators understand the controller's reasoning during critical events. Evaluating these tools through user feedback or expert assessment would ensure that the explanations genuinely support trust and safe adoption in real-world operation.

These directions would extend the current simulation-based framework towards a more physically detailed, adaptive, and operator-informed control system for resilient microgrids operating under real-world constraints.

References

- [1] A. B. Smith, “Us billion-dollar weather and climate disasters, 1980-present,” 2024.
- [2] M. H. N. Amiri and F. Guéniat, “Towards a framework for measurements of power systems resiliency: Comprehensive review and development of graph and vector-based resilience metrics,” *Sustainable Cities and Society*, vol. 109, p. 105 517, 2024.
- [3] M. Panteli and P. Mancarella, “Modeling and evaluating the resilience of critical electrical power infrastructure to extreme weather events,” *IEEE Systems Journal*, vol. 11, no. 3, pp. 1733–1742, 2015.
- [4] S. Espinoza, M. Panteli, P. Mancarella, and H. Rudnick, “Multi-phase assessment and adaptation of power systems resilience to natural hazards,” *Electric Power Systems Research*, vol. 136, pp. 352–361, 2016.
- [5] G. Fu et al., “Integrated approach to assess the resilience of future electricity infrastructure networks to climate hazards,” *IEEE Systems Journal*, vol. 12, no. 4, pp. 3169–3180, 2017.
- [6] M. H. Nejati Amiri, M. Mehdinejad, A. Mohammadpour Shotorbani, and H. Shayanfar, “Heuristic retailer’s day-ahead pricing based on online-learning of prosumer’s optimal energy management model,” *Energies*, vol. 16, no. 3, p. 1182, 2023.
- [7] O. S. Omogoye, K. A. Folly, and K. O. Awodele, “Review of sequential steps to realize power system resilience,” in *2020 International SAUPEC/RobMech/PRASA Conference*, IEEE, 2020, pp. 1–6.
- [8] J. Jasiūnas, P. D. Lund, and J. Mikkola, “Energy system resilience—a review,” *Renewable and Sustainable Energy Reviews*, vol. 150, p. 111 476, 2021.

-
- [9] M. Panteli, D. N. Trakas, P. Mancarella, and N. D. Hatziargyriou, “Power systems resilience assessment: Hardening and smart operational enhancement strategies,” *Proceedings of the IEEE*, vol. 105, no. 7, pp. 1202–1213, 2017.
 - [10] M. M. Hosseini, L. Rodriguez-Garcia, and M. Parvania, “Hierarchical combination of deep reinforcement learning and quadratic programming for distribution system restoration,” *IEEE Transactions on Sustainable Energy*, vol. 14, no. 2, pp. 1088–1098, 2023.
 - [11] R. Wattegama, M. Short, G. Aggarwal, M. Al-Greer, and R. Naidoo, “A systematic review of hierarchical control frameworks in resilient microgrids: South africa focus,” *Energies*, vol. 19, no. 3, p. 644, 2026.
 - [12] A. Stankovic, “The definition and quantification of resilience,” *IEEE PES Industry Technical Support Task Force: Piscataway, NJ, USA*, pp. 1–4, 2018.
 - [13] A. Younesi, H. Shayeghi, Z. Wang, P. Siano, A. Mehrizi-Sani, and A. Safari, “Trends in modern power systems resilience: State-of-the-art review,” *Renewable and Sustainable Energy Reviews*, vol. 162, p. 112 397, 2022.
 - [14] H. Gao, Y. Chen, Y. Xu, and C.-C. Liu, “Resilience-oriented critical load restoration using microgrids in distribution systems,” *IEEE Transactions on Smart Grid*, vol. 7, no. 6, pp. 2837–2848, 2016.
 - [15] M. Panteli, P. Mancarella, D. N. Trakas, E. Kyriakides, and N. D. Hatziargyriou, “Metrics and quantification of operational and infrastructure resilience in power systems,” *IEEE Transactions on Power Systems*, vol. 32, no. 6, pp. 4732–4742, 2017.
 - [16] T. Lagos et al., “Identifying optimal portfolios of resilient network investments against natural hazards, with applications to earthquakes,” *IEEE Transactions on Power Systems*, vol. 35, no. 2, pp. 1411–1421, 2019.

-
- [17] W. Li et al., “Assessing power system resilience to floods: A geo-referenced statistical model for substation inundation failures,” in *2022 IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT-Europe)*, IEEE, 2022, pp. 1–5.
 - [18] M. Panteli, D. N. Trakas, P. Mancarella, and N. D. Hatziargyriou, “Boosting the power grid resilience to extreme weather events using defensive islanding,” *IEEE Transactions on Smart Grid*, vol. 7, no. 6, pp. 2913–2922, 2016.
 - [19] M. Abdelmalak, J. Cox, S. Ericson, E. Hotchkiss, and M. Benidris, “Quantitative resilience-based assessment framework using eagle-i power outage data,” *IEEE Access*, vol. 11, pp. 7682–7697, 2023.
 - [20] E. L. Barrett et al., “A risk-based framework for power system modeling to improve resilience to extreme events,” *IEEE Open Access Journal of Power and Energy*, vol. 10, pp. 25–35, 2022.
 - [21] X. Liu, M. Shahidehpour, Z. Li, X. Liu, Y. Cao, and Z. Bie, “Microgrids for enhancing the power grid resilience in extreme conditions,” *IEEE Transactions on Smart Grid*, vol. 8, no. 2, pp. 589–597, 2016.
 - [22] H. Zhang, H. Yuan, G. Li, and Y. Lin, “Quantitative resilience assessment under a tri-stage framework for power systems,” *Energies*, vol. 11, no. 6, p. 1427, 2018.
 - [23] F. Dehghani, M. Mohammadi, and M. Karimi, “Age-dependent resilience assessment and quantification of distribution systems under extreme weather events,” *International Journal of Electrical Power & Energy Systems*, vol. 150, p. 109 089, 2023.
 - [24] M. Lotfi, M. Panteli, B. V. Venkatasubramanian, M. S. Javadi, L. M. Carvalho, and C. S. Gouveia, “Quantifying the difference between resilience and reliability in the operation planning of mobile resources for power distribution grids,” *Findings*, 2022.

-
- [25] A. Hussain, A. O. Rousis, I. Konstantelos, G. Strbac, J. Jeon, and H.-M. Kim, "Impact of uncertainties on resilient operation of microgrids: A data-driven approach," *IEEE Access*, vol. 7, pp. 14 924–14 937, 2019.
- [26] N. T. Bazargani and S. Bathaee, "A general framework for resiliency evaluation of radial distribution system against extreme events," in *Electrical Engineering (ICEE), Iranian Conference on*, IEEE, 2018, pp. 1179–1184.
- [27] M. Dilley, *Natural disaster hotspots: a global risk analysis*. World Bank Publications, 2005, vol. 5.
- [28] AON. "Weather, climate and catastrophe insight | aon." Accessed Sep. 13, 2023. [Online]. Available: <https://www.aon.com/weather-climate-catastrophe/index.aspx>.
- [29] J. Gui, H. Lei, T. R. McJunkin, B. Chen, and B. K. Johnson, "Operational resilience metrics for power systems with penetration of renewable resources," *IET Generation, Transmission & Distribution*, vol. 17, no. 10, pp. 2344–2355, 2023.
- [30] S. D. Manshadi and M. E. Khodayar, "Resilient operation of multiple energy carrier microgrids," *IEEE Transactions on Smart Grid*, vol. 6, no. 5, pp. 2283–2292, 2015.
- [31] M. Nazemi, M. Moeini-Aghaie, M. Fotuhi-Firuzabad, and P. Dehghanian, "Energy storage planning for enhanced resilience of power distribution networks against earthquakes," *IEEE Transactions on Sustainable Energy*, vol. 11, no. 2, pp. 795–806, 2019.
- [32] A. Poudyal, S. Poudel, and A. Dubey, "Risk-based active distribution system planning for resilience against extreme weather events," *IEEE Transactions on Sustainable Energy*, vol. 14, no. 2, pp. 1178–1192, 2022.
- [33] Y. Yang, W. Tang, Y. Liu, Y. Xin, and Q. Wu, "Quantitative resilience assessment for power transmission systems under typhoon weather," *IEEE Access*, vol. 6, pp. 40 747–40 756, 2018.

-
- [34] S. Chanda, A. K. Srivastava, M. U. Mohanpurkar, and R. Hovsapien, “Quantifying power distribution system resiliency using code-based metric,” *IEEE Transactions on Industry Applications*, vol. 54, no. 4, pp. 3676–3686, 2018.
 - [35] A. Younesi, H. Shayeghi, A. Safari, and P. Siano, “A quantitative resilience measure framework for power systems against wide-area extreme events,” *IEEE Systems Journal*, vol. 15, no. 1, pp. 915–922, 2020.
 - [36] T. Xu, C. Shao, M. Shahidehpour, and X. Wang, “Coordinated planning strategies of power systems and energy transportation networks for resilience enhancement,” *IEEE Transactions on Sustainable Energy*, vol. 14, no. 2, pp. 1217–1229, 2022.
 - [37] A. Rahiminejad et al., “A resilience-based recovery scheme for smart grid restoration following cyberattacks to substations,” *International Journal of Electrical Power & Energy Systems*, vol. 145, p. 108 610, 2023.
 - [38] M. Mccarty et al., “Cybersecurity resilience demonstration for wind energy sites in co-simulation environment,” *IEEE Access*, vol. 11, pp. 15 297–15 313, 2023.
 - [39] P. Dehghanian, S. Aslan, and P. Dehghanian, “Maintaining electric system safety through an enhanced network resilience,” *IEEE Transactions on Industry Applications*, vol. 54, no. 5, pp. 4927–4937, 2018.
 - [40] S. Chanda and A. K. Srivastava, “Defining and enabling resiliency of electric distribution systems with multiple microgrids,” *IEEE Transactions on Smart Grid*, vol. 7, no. 6, pp. 2859–2868, 2016.
 - [41] C. Qin, L. Jia, S. Bajagain, S. Pannala, A. K. Srivastava, and A. Dubey, “An integrated situational awareness tool for resilience-driven restoration with sustainable energy resources,” *IEEE Transactions on Sustainable Energy*, vol. 14, no. 2, pp. 1099–1111, 2023.

-
- [42] P. Bajpai, S. Chanda, and A. K. Srivastava, "A novel metric to quantify and enable resilient distribution system using graph theory and choquet integral," *IEEE Transactions on Smart Grid*, vol. 9, no. 4, pp. 2918–2929, 2016.
- [43] B. R. Naidu, K. K. Saini, P. Bajpai, and C. Chakraborty, "A novel framework for resilient overhead power distribution networks," *International Journal of Electrical Power & Energy Systems*, vol. 147, p. 108 839, 2023.
- [44] A. D. Bintoudi and C. Demoulias, "Optimal isolated microgrid topology design for resilient applications," *Applied Energy*, vol. 338, p. 120 909, 2023.
- [45] M. Mahzarnia, M. P. Moghaddam, P. Siano, and M.-R. Haghifam, "A comprehensive assessment of power system resilience to a hurricane using a two-stage analytical approach incorporating risk-based index," *Sustainable Energy Technologies and Assessments*, vol. 42, p. 100 831, 2020.
- [46] G. Li, G. Huang, Z. Bie, Y. Lin, and Y. Huang, "Component importance assessment of power systems for improving resilience under wind storms," *Journal of Modern Power Systems and Clean Energy*, vol. 7, no. 4, pp. 676–687, 2019.
- [47] S. A. Mansouri, E. Nematbakhsh, A. Ahmarinejad, A. R. Jordehi, M. S. Javadi, and M. Marzband, "A hierarchical scheduling framework for resilience enhancement of decentralized renewable-based microgrids considering proactive actions and mobile units," *Renewable and Sustainable Energy Reviews*, vol. 168, p. 112 854, 2022.
- [48] S. A. Mansouri, E. Nematbakhsh, M. S. Javadi, A. R. Jordehi, M. Shafie-khah, and J. P. Catalão, "Resilience enhancement via automatic switching considering direct load control program and energy storage systems," in *2021 IEEE International Conference on Environment and Electrical Engineering and 2021 IEEE Industrial and Commercial Power Systems Europe (EEEIC/I&CPS Europe)*, IEEE, 2021, pp. 1–6.

-
- [49] M. H. N. Amiri, S. Dhundhara, F. Annaz, M. De Oliveira, and F. Gueniat, “Two-stage microgrid resilience and battery life-aware planning and operation for cyclone prone areas in india,” *Sustainable Cities and Society*, vol. 124, p. 106 290, 2025.
- [50] M. Yin, K. Li, and J. Yu, “A data-driven approach for microgrid distributed generation planning under uncertainties,” *Applied Energy*, vol. 309, p. 118 429, 2022.
- [51] M. S. Abid, H. J. Apon, S. Hossain, A. Ahmed, R. Ahshan, and M. H. Lipu, “A novel multi-objective optimization based multi-agent deep reinforcement learning approach for microgrid resources planning,” *Applied Energy*, vol. 353, p. 122 029, 2024.
- [52] P. Kurukuri, M. R. Mohamed, P. H. Raavi, and Y. Arya, “Optimal planning and designing of microgrid systems with hybrid renewable energy technologies for sustainable environment in cities,” *Environmental Science and Pollution Research*, pp. 1–18, 2024.
- [53] M. M. Kamal, I. Ashraf, and E. Fernandez, “Planning and optimization of microgrid for rural electrification with integration of renewable energy resources,” *Journal of Energy Storage*, vol. 52, p. 104 782, 2022.
- [54] N. N. Ibrahim, J. J. Jamian, and M. M. Rasid, “Optimal multi-objective sizing of renewable energy sources and battery energy storage systems for formation of a multi-microgrid system considering diverse load patterns,” *Energy*, vol. 304, p. 131 921, 2024.
- [55] M. Borghei and M. Ghassemi, “Optimal planning of microgrids for resilient distribution networks,” *International Journal of Electrical Power & Energy Systems*, vol. 128, p. 106 682, 2021.
- [56] Y. Wang, A. O. Rousis, and G. Strbac, “A three-level planning model for optimal sizing of networked microgrids considering a trade-off between resilience and cost,” *IEEE Transactions on Power Systems*, vol. 36, no. 6, pp. 5657–5669, 2021.

-
- [57] M. Ghasemi, A. Kazemi, A. Mazza, and E. Bompard, “A three-stage stochastic planning model for enhancing the resilience of distribution systems with microgrid formation strategy,” *IET Generation, Transmission & Distribution*, vol. 15, no. 13, pp. 1908–1921, 2021.
- [58] E. G. Vera, C. A. Cañizares, M. Pirnia, T. P. Guedes, and J. D. M. Trujillo, “Two-stage stochastic optimization model for multi-microgrid planning,” *IEEE Transactions on Smart Grid*, vol. 14, no. 3, pp. 1723–1735, 2022.
- [59] S. Phommixay, M. L. Dombia, and Q. Cui, “A two-stage two-layer optimization approach for economic operation of a microgrid under a planned outage,” *Sustainable Cities and Society*, vol. 66, p. 102675, 2021.
- [60] S. Yang, J. Fang, Z. Zhang, S. Lv, H. Lin, and L. Ju, “Two-stage coordinated optimal dispatching model and benefit allocation strategy for rural new energy microgrid,” *Energy*, vol. 292, p. 130274, 2024.
- [61] F. Amini, S. Ghassemzadeh, N. Rostami, and V. S. Tabar, “A stochastic two-stage microgrid formation strategy for enhancing distribution network resilience against earthquake event incorporating distributed energy resources, parking lots and responsive loads,” *Sustainable Cities and Society*, vol. 101, p. 105191, 2024.
- [62] S. Mansouri, F. Zishan, O. D. Montoya, M. Azimzadeh, and D. A. Giral-Ramírez, “Using an intelligent method for microgrid generation and operation planning while considering load uncertainty,” *Results in Engineering*, vol. 17, p. 100978, 2023.
- [63] Y. Liu, S. Yang, D. Li, and S. Zhang, “Improved whale optimization algorithm for solving microgrid operations planning problems,” *Symmetry*, vol. 15, no. 1, p. 36, 2022.

-
- [64] B. Dey, S. Misra, and F. P. G. Marquez, “Microgrid system energy management with demand response program for clean and economical operation,” *Applied Energy*, vol. 334, p. 120717, 2023.
- [65] H. Khaloie, A. Abdollahi, M. Rashidinejad, and P. Siano, “Risk-based probabilistic-possibilistic self-scheduling considering high-impact low-probability events uncertainty,” *International Journal of Electrical Power & Energy Systems*, vol. 110, pp. 598–612, 2019.
- [66] M. Shehzad and F. Gueniat, “Modeling and optimal control of energy storage strategy for battery life extension via model predictive control,” in *2021 European Control Conference (ECC)*, IEEE, 2021, pp. 1981–1986.
- [67] M. Amini, M. B. Sanjareh, M. H. Nazari, G. Gharehpetian, and S. H. Hosseinian, “A novel model for battery optimal sizing in microgrid planning considering battery capacity degradation process and thermal impact,” *IEEE Transactions on Sustainable Energy*, 2024.
- [68] A. Nawaz, J. Wu, J. Ye, Y. Dong, and C. Long, “Distributed mpc-based energy scheduling for islanded multi-microgrid considering battery degradation and cyclic life deterioration,” *Applied energy*, vol. 329, p. 120168, 2023.
- [69] M. H. N. Amiri, F. Annaz, M. De Oliveira, and F. Gueniat, “Strategies for resilience and battery life extension in the face of communication losses for isolated microgrids,” in *2024 IEEE International Conference on Omni-layer Intelligent Systems (COINS)*, 2024, pp. 1–5. DOI: [10.1109/COINS61597.2024.10622118](https://doi.org/10.1109/COINS61597.2024.10622118).
- [70] S. A. Shezan, K. N. Hasan, A. Rahman, M. Datta, and U. Datta, “Selection of appropriate dispatch strategies for effective planning and operation of a microgrid,” *Energies*, vol. 14, no. 21, p. 7217, 2021.

-
- [71] H. U. R. Habib et al., “Optimal planning and ems design of pv based standalone rural microgrids,” *IEEE Access*, vol. 9, pp. 32 908–32 930, 2021.
- [72] K. Pang, J. Zhou, S. Tsianikas, D. W. Coit, and Y. Ma, “Long-term microgrid expansion planning with resilience and environmental benefits using deep reinforcement learning,” *Renewable and Sustainable Energy Reviews*, vol. 191, p. 114 068, 2024.
- [73] M. Pereira and R. E. Araújo, “Model-free finite-set predictive current control with optimal cycle time for a switched reluctance motor,” *IEEE Transactions on Industrial Electronics*, vol. 70, no. 8, pp. 8355–8364, 2022.
- [74] S. U. Ali, A. Waqar, M. Aamir, S. M. Qaisar, and J. Iqbal, “Model predictive control of consensus-based energy management system for dc microgrid,” *PloS one*, vol. 18, no. 1, e0278110, 2023.
- [75] D. I. H. Rodríguez and J. M. Myrzik, “Economic model predictive control for optimal operation of home microgrid with photovoltaic-combined heat and power storage systems,” *IFAC-PapersOnLine*, vol. 50, no. 1, pp. 10 027–10 032, 2017.
- [76] J. Sachs and O. Sawodny, “A two-stage model predictive control strategy for economic diesel-pv-battery island microgrid operation in rural areas,” *IEEE Transactions on Sustainable Energy*, vol. 7, no. 3, pp. 903–913, 2016.
- [77] R. Carli, M. Dotoli, J. Jantzen, M. Kristensen, and S. B. Othman, “Energy scheduling of a smart microgrid with shared photovoltaic panels and storage: The case of the ballen marina in samsø,” *Energy*, vol. 198, p. 117 188, 2020.
- [78] I. Novickij and G. Joós, “Model predictive control based approach for microgrid energy management,” in *2019 IEEE Canadian conference of electrical and computer engineering (CCECE)*, IEEE, 2019, pp. 1–4.

-
- [79] M. Amini, M. H. Nazari, and S. H. Hosseinian, “Optimal energy management of battery with high wind energy penetration: A comprehensive linear battery degradation cost model,” *Sustainable Cities and Society*, vol. 93, p. 104 492, 2023.
- [80] A. Parisio, E. Rikos, and L. Glielmo, “A model predictive control approach to microgrid operation optimization,” *IEEE Transactions on Control Systems Technology*, vol. 22, no. 5, pp. 1813–1827, 2014.
- [81] A. Vasilakis, I. Zafeiratou, D. T. Lagos, and N. D. Hatziargyriou, “The evolution of research in microgrids control,” *IEEE Open Access Journal of Power and Energy*, vol. 7, pp. 331–343, 2020.
- [82] H. Zhang, S. Seal, D. Wu, F. Bouffard, and B. Boulet, “Building energy management with reinforcement learning and model predictive control: A survey,” *IEEE Access*, vol. 10, pp. 27 853–27 862, 2022.
- [83] P. Mohammadi, R. Darshi, S. Shamaghdari, and P. Siano, “Comparative analysis of control strategies for microgrid energy management with a focus on reinforcement learning,” *IEEE Access*, pp. 1–1, 2024. DOI: [10.1109/ACCESS.2024.3495032](https://doi.org/10.1109/ACCESS.2024.3495032).
- [84] I. H. Sarker, “Machine learning: Algorithms, real-world applications and research directions,” *SN computer science*, vol. 2, no. 3, p. 160, 2021.
- [85] R. S. Sutton and A. G. Barto, *Reinforcement learning: An introduction*. MIT press, 2018.
- [86] L. Graesser and W. L. Keng, *Foundations of deep reinforcement learning: theory and practice in Python*. Addison-Wesley Professional, 2019.
- [87] V. Hassija et al., “Interpreting black-box models: A review on explainable artificial intelligence,” *Cognitive Computation*, vol. 16, no. 1, pp. 45–74, 2024.
- [88] M. Hessel et al., “Rainbow: Combining improvements in deep reinforcement learning,” in *Proceedings of the AAAI conference on artificial intelligence*, vol. 32, 2018.

-
- [89] V. Mnih et al., “Asynchronous methods for deep reinforcement learning,” in *International conference on machine learning*, PMLR, 2016, pp. 1928–1937.
- [90] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, “Proximal policy optimization algorithms,” *arXiv preprint arXiv:1707.06347*, 2017.
- [91] C. Ma, S. Lei, D. Chen, C. Wang, N. D. Hatziargyriou, and Z. Song, “Sequential service restoration with grid-interactive flexibility from building ac systems for resilient microgrids under endogenous and exogenous uncertainties,” *Applied Energy*, vol. 377, p. 124351, 2025.
- [92] D. D. Sharma and R. C. Bansal, “Lstm-sac reinforcement learning based resilient energy trading for networked microgrid system,” *AIMS Electronics and Electrical Engineering*, vol. 9, no. 2, pp. 165–191, 2025.
- [93] A. Halev, Y. Liu, and X. Liu, “Microgrid control under uncertainty,” *Engineering Applications of Artificial Intelligence*, vol. 138, p. 109360, 2024.
- [94] H. Momen and S. Jadid, “Resilience enhancement of power distribution system using fixed and mobile emergency generators based on deep reinforcement learning,” *Engineering Applications of Artificial Intelligence*, vol. 137, p. 109118, 2024.
- [95] S. Rath, T. Das, and S. Sengupta, “Improvise, adapt, overcome: Dynamic resiliency against unknown attack vectors in microgrid cybersecurity games,” *IEEE Transactions on Smart Grid*, 2024.
- [96] G. Jones, X. Li, and Y. Sun, “Robust energy management policies for solar microgrids via reinforcement learning,” *Energies*, vol. 17, no. 12, p. 2821, 2024.
- [97] M. S. Jonban, L. Romeral, M. Marzband, and A. Abusorrah, “A reinforcement learning approach using markov decision processes for battery energy storage control within a smart contract framework,” *Journal of Energy Storage*, vol. 86, p. 111342, 2024.

-
- [98] D. Kumar and A. Kumar, “A reliable hybrid autoregressive integrated moving average and deep reinforcement machine learning strategy for resiliency enhancement in microgrid,” *Sustainable Energy, Grids and Networks*, vol. 39, p. 101 424, 2024.
- [99] K. Pang, J. Zhou, S. Tsianikas, and Y. Ma, “Deep reinforcement learning for resilient microgrid expansion planning with multiple energy resource,” *Quality and Reliability Engineering International*, vol. 40, no. 1, pp. 34–56, 2024.
- [100] A. Sinha, R. Vyas, F. Alasali, W. Holderbaum, and O. Vyas, “A deep reinforcement learning-based approach for cyber resilient demand response optimization,” *Frontiers in Energy Research*, vol. 12, p. 1 494 164, 2025.
- [101] Q. Hu, Y. Ye, Y. Wu, and X. Xie, “Microgrid energy management considering communication: A bayesian deep q-learning learning approach,” in *2024 3rd International Conference on Power Systems and Electrical Technology (PSET)*, IEEE, 2024, pp. 626–631.
- [102] E. Özkan, İ. Kök, Ā. Suat, et al., “Deeptwin: A deep reinforcement learning supported digital twin model for micro-grids,” *IEEE Access*, 2024.
- [103] P. Aslami et al., “A real-time digital simulator accelerated reinforcement learning training environment for power system frequency dynamics,” in *2024 56th North American Power Symposium (NAPS)*, IEEE, 2024, pp. 1–6.
- [104] S. Guoping, Q. Yece, W. Longsheng, L. Yujuan, and Z. Yufeng, “Dispatching isolated off-grid microgrids considering uncertainty: A prediction-decision integrated approach,” *Journal of Control and Decision*, pp. 1–17, 2024.
- [105] R. S. Nuvvula et al., “Distributed multi-agent reinforcement learning for autonomus management of renewable energy microgrids,” in *2024 12th International Conference on Smart Grid (icSmartGrid)*, IEEE, 2024, pp. 298–304.

-
- [106] P. P. Kumar et al., “Energy-aware vehicle-to-grid (v2g) scheduling with reinforcement learning for renewable energy integration,” in *2024 12th International Conference on Smart Grid (icSmartGrid)*, IEEE, 2024, pp. 345–349.
- [107] H. Zhang, D. Yue, C. Dou, and G. P. Hancke, “Resilient optimal defensive strategy of micro-grids system via distributed deep reinforcement learning approach against fdi attack,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 35, no. 1, pp. 598–608, 2022.
- [108] T. Tabassum, S. Lim, and M. R. Khalghani, “Artificial intelligence-based detection and mitigation of cyber disruptions in microgrid control,” *Electric Power Systems Research*, vol. 226, p. 109 925, 2024.
- [109] L. Tightiz, L. M. Dang, and J. Yoo, “Novel deep deterministic policy gradient technique for automated micro-grid energy management in rural and islanded areas,” *Alexandria Engineering Journal*, vol. 82, pp. 145–153, 2023.
- [110] T. Zhang, M. Sun, D. Qiu, X. Zhang, G. Strbac, and C. Kang, “A bayesian deep reinforcement learning-based resilient control for multi-energy micro-gird,” *IEEE Transactions on Power Systems*, vol. 38, no. 6, pp. 5057–5072, 2023.
- [111] P. R. Jeyaraj, S. P. Asokan, A. C. Kathiresan, and E. R. S. Nadar, “Deep reinforcement learning-based network for optimized power flow in islanded dc microgrid,” *Electrical Engineering*, vol. 105, no. 5, pp. 2805–2816, 2023.
- [112] D. Qiu, Y. Wang, T. Zhang, M. Sun, and G. Strbac, “Hierarchical multi-agent reinforcement learning for repair crews dispatch control towards multi-energy microgrid resilience,” *Applied Energy*, vol. 336, p. 120 826, 2023.
- [113] C. Caputo, M.-A. Cardin, P. Ge, F. Teng, A. Korre, and E. A. del Rio Chanona, “Design and planning of flexible mobile micro-grids using deep reinforcement learning,” *Applied energy*, vol. 335, p. 120 707, 2023.

-
- [114] A. K. Goswami, S. Deb, N. Jha, N. Das, and B. K. Das, “Machine learning based wind energy forecasting for energy mangement in microgrid system,” in *2023 IEEE 3rd International Conference on Smart Technologies for Power, Energy and Control (STPEC)*, IEEE, 2023, pp. 1–6.
- [115] M. Ballouch, O. Souissi, and M. R. El-Fenni, “Enhancing control in manufacturing and microgrid systems: Deep reinforcement learning with double q-learning,” in *2023 14th International Conference on Intelligent Systems: Theories and Applications (SITA)*, IEEE, 2023, pp. 1–7.
- [116] S. Mukherjee et al., “Enhancing cyber resilience of networked microgrids using vertical federated reinforcement learning,” in *2023 IEEE Power & Energy Society General Meeting (PESGM)*, IEEE, 2023, pp. 1–5.
- [117] S. S. Mavuri, J. Nakka, and A. Kotla, “Deep neural network based intelligent multi-microgrid energy management,” in *2023 IEEE 3rd International Conference on Sustainable Energy and Future Electric Transportation (SEFET)*, IEEE, 2023, pp. 1–6.
- [118] M. Panda, O. Al Zaabi, K. Al Jaafari, P. Kumar, K. Al Hosani, and U. R. Muduli, “State-of-charge droop control for efficient power sharing and transient response in microgrids,” in *2023 IEEE IAS Global Conference on Emerging Technologies (Glob-ConET)*, IEEE, 2023, pp. 1–5.
- [119] Y. Huang, G. Li, C. Chen, Y. Bian, T. Qian, and Z. Bie, “Resilient distribution networks by microgrid formation using deep reinforcement learning,” *IEEE Transactions on Smart Grid*, vol. 13, no. 6, pp. 4918–4930, 2022.
- [120] R. Artis, M. Assili, and M. Shivaie, “A seismic-resilient multi-level framework for distribution network reinforcement planning considering renewable-based multi-microgrids,” *Applied Energy*, vol. 325, p. 119 824, 2022.

-
- [121] K. Deshpande, P. Möhl, A. Hämmerle, G. Weichhart, H. Zörrer, and A. Pichler, “Energy management simulation with multi-agent reinforcement learning: An approach to achieve reliability and resilience,” *Energies*, vol. 15, no. 19, p. 7381, 2022.
 - [122] J. Nie, Y. Liu, L. Zhou, X. Jiang, and M. Preindl, “Deep reinforcement learning based approach for optimal power flow of microgrid with grid services implementation,” in *2022 IEEE Transportation Electrification Conference & Expo (ITEC)*, IEEE, 2022, pp. 1148–1153.
 - [123] J. Zhao, F. Li, S. Mukherjee, and C. Sticht, “Deep reinforcement learning-based model-free on-line dynamic multi-microgrid formation to enhance resilience,” *IEEE Transactions on Smart Grid*, vol. 13, no. 4, pp. 2557–2567, 2022.
 - [124] Y. Du and D. Wu, “Deep reinforcement learning from demonstrations to assist service restoration in islanded microgrids,” *IEEE Transactions on Sustainable Energy*, vol. 13, no. 2, pp. 1062–1072, 2022.
 - [125] A. Dridi, H. Affi, H. Moun gla, and J. Badosa, “A novel deep reinforcement approach for iiot microgrid energy management systems,” *IEEE Transactions on Green Communications and Networking*, vol. 6, no. 1, pp. 148–159, 2021.
 - [126] T. Wu, J. Wang, X. Lu, and Y. Du, “Ac/dc hybrid distribution network reconfiguration with microgrid formation using multi-agent soft actor-critic,” *Applied Energy*, vol. 307, p. 118 189, 2022.
 - [127] J. Rahman, R. A. Jacob, S. Paul, S. Chowdhury, and J. Zhang, “Reinforcement learning enabled microgrid network reconfiguration under disruptive events,” in *2022 IEEE Kansas Power and Energy Conference (KPEC)*, IEEE, 2022, pp. 1–6.
 - [128] T. Levent, P. Preux, G. Henri, R. Alami, P. Cordier, and Y. Bonnassieux, “The challenge of controlling microgrids in the presence of rare events with deep reinforcement learning,” *IET Smart Grid*, vol. 4, no. 1, pp. 15–28, 2021.

-
- [129] F. N. Shimim, M. Alali, H. Nehrir, J. Sheppard, M. Bahramipناه, and Z. Shahooei, “Resiliency-aware power management of microgrids using agent-based dynamic programming and q-learning,” in *2021 IEEE PES Innovative Smart Grid Technologies-Asia (ISGT Asia)*, IEEE, 2021, pp. 1–5.
- [130] X. Zhang, A. T. Eseye, M. Reynolds, B. Knueven, and W. Jones, “Restoring critical loads in resilient distribution systems using a curriculum learned controller,” in *2021 IEEE Power & Energy Society General Meeting (PESGM)*, IEEE, 2021, pp. 1–5.
- [131] R. Hu and A. Kwasinski, “Energy management for microgrids using a reinforcement learning algorithm,” in *2021 IEEE green energy and smart systems conference (IGESSC)*, IEEE, 2021, pp. 1–6.
- [132] Y. Li, Z. Xu, K. B. Bowes, and L. Ren, “Reinforcement learning-enabled seamless microgrids interconnection,” in *2021 IEEE Power & Energy Society General Meeting (PESGM)*, IEEE, 2021, pp. 1–5.
- [133] L. Tightiz and H. Yang, “Resilience microgrid as power system integrity protection scheme element with reinforcement learning based management,” *IEEE Access*, vol. 9, pp. 83 963–83 975, 2021.
- [134] X. Zhang, A. T. Eseye, B. Knueven, and W. Jones, “Restoring distribution system under renewable uncertainty using reinforcement learning,” in *2020 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm)*, IEEE, 2020, pp. 1–6.
- [135] N. Radhakrishnan et al., “Improving primary frequency response in networked microgrid operations using multilayer perceptron-driven reinforcement learning,” *IET Smart Grid*, vol. 3, no. 4, pp. 500–507, 2020.

-
- [136] M. Elsayed, M. Erol-Kantarci, B. Kantarci, L. Wu, and J. Li, “Low-latency communications for community resilience microgrids: A reinforcement learning approach,” *IEEE Transactions on Smart Grid*, vol. 11, no. 2, pp. 1091–1099, 2019.
- [137] R. Fan, R. Sun, Y. Liu, and R. u. Hassan, “An online decision-making method based on multi-agent interaction for coordinated load restoration,” *Frontiers in Energy Research*, vol. 10, p. 992966, 2022.
- [138] K. Zhang, J. Zhang, P.-D. Xu, T. Gao, and D. W. Gao, “Explainable ai in deep reinforcement learning models for power system emergency control,” *IEEE Transactions on Computational Social Systems*, vol. 9, no. 2, pp. 419–427, 2021.
- [139] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, “Proximal policy optimization algorithms,” *arXiv preprint arXiv:1707.06347*, 2017.
- [140] M. T. Ribeiro, S. Singh, and C. Guestrin, “" why should i trust you?" explaining the predictions of any classifier,” in *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*, 2016, pp. 1135–1144.
- [141] S. Lundberg, “A unified approach to interpreting model predictions,” *arXiv preprint arXiv:1705.07874*, 2017.
- [142] D. Gunning and D. Aha, “Darpa’s explainable artificial intelligence (xai) program,” *AI magazine*, vol. 40, no. 2, pp. 44–58, 2019.
- [143] Council of the European Union, *The eu general data protection regulation (gdpr)*, Accessed: 2025-08-24, 2023. [Online]. Available: <https://www.consilium.europa.eu/en/policies/data-protection-regulation/>.
- [144] European Parliament, *Artificial intelligence act: Meps adopt landmark law*, Accessed: 2025-08-24, 2024. [Online]. Available: <https://www.europarl.europa.eu/news/en/press-room/20240308IPR19015/artificial-intelligence-act-meps-adopt-landmark-law>.

-
- [145] M. R. Carrillo, “Artificial intelligence: From ethics to law,” *Telecommunications policy*, vol. 44, no. 6, p. 101 937, 2020.
- [146] V. Hassija et al., *Interpreting black-box models: A review on explainable artificial intelligence. cognitive computation*, 2023.
- [147] Y. Bekkemoen, “Explainable reinforcement learning (xrl): A systematic literature review and taxonomy,” *Machine Learning*, vol. 113, no. 1, pp. 355–441, 2024.
- [148] D. Jin, E. Sergeeva, W.-H. Weng, G. Chauhan, and P. Szolovits, “Explainable deep learning in healthcare: A methodological survey from an attribution view,” *WIREs Mechanisms of Disease*, vol. 14, no. 3, e1548, 2022.
- [149] F. Cruz, R. Dazeley, P. Vamplew, and I. Moreira, “Explainable robotic systems: Understanding goal-driven actions in a reinforcement learning scenario,” *Neural Computing and Applications*, vol. 35, no. 25, pp. 18 113–18 130, 2023.
- [150] X. Peng, M. Riedl, and P. Ammanabrolu, “Inherently explainable reinforcement learning in natural language,” *Advances in Neural Information Processing Systems*, vol. 35, pp. 16 178–16 190, 2022.
- [151] N. Le, V. S. Rathour, K. Yamazaki, K. Luu, and M. Savvides, “Deep reinforcement learning in computer vision: A comprehensive survey,” *Artificial Intelligence Review*, pp. 1–87, 2022.
- [152] Y.-J. Hu and S.-J. Lin, “Deep reinforcement learning for optimizing finance portfolio management,” in *2019 Amity International Conference on Artificial Intelligence (AICAI)*, 2019, pp. 14–20. DOI: [10.1109/AICAI.2019.8701368](https://doi.org/10.1109/AICAI.2019.8701368).
- [153] G. Rjoub, J. Bentahar, and O. A. Wahab, “Explainable ai-based federated deep reinforcement learning for trusted autonomous driving,” in *2022 International Wireless Communications and Mobile Computing (IWCMC)*, IEEE, 2022, pp. 318–323.

-
- [154] H. Kumar, P. Mammen, and K. Ramamritham, *Explainable ai reinforcement learning agents for residential cost savings*, 2019.
 - [155] K. Zhang, J. Zhang, P.-D. Xu, T. Gao, and D. W. Gao, “Explainable ai in deep reinforcement learning models for power system emergency control,” *IEEE Transactions on Computational Social Systems*, vol. 9, no. 2, pp. 419–427, 2021.
 - [156] Y. Li, “Deep reinforcement learning: An overview,” *arXiv preprint arXiv:1701.07274*, 2017.
 - [157] J. Huang, P. P. Angelov, and C. Yin, “Interpretable policies for reinforcement learning by empirical fuzzy sets,” *Engineering Applications of Artificial Intelligence*, vol. 91, p. 103 559, 2020.
 - [158] M. Landajuela et al., “Discovering symbolic policies with deep reinforcement learning,” in *International Conference on Machine Learning*, PMLR, 2021, pp. 5979–5989.
 - [159] N. Topin, S. Milani, F. Fang, and M. Veloso, “Iterative bounding mdps: Learning interpretable policies via non-interpretable methods,” in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 35, 2021, pp. 9923–9931.
 - [160] W. Bao, Q. Yu, and Y. Kong, “Drive: Deep reinforced accident anticipation with visual explanation,” in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2021, pp. 7619–7628.
 - [161] Y. Sawada and K. Nakamura, “C-senn: Contrastive self-explaining neural network,” *arXiv preprint arXiv:2206.09575*, 2022.
 - [162] H. Ben-Younes, É. Zablocki, P. Pérez, and M. Cord, “Driving behavior explanation with multi-level fusion,” *Pattern Recognition*, vol. 123, p. 108 421, 2022.
 - [163] S. M. Lundberg and S.-I. Lee, “A unified approach to interpreting model predictions,” *Advances in neural information processing systems*, vol. 30, 2017.

-
- [164] M. T. Ribeiro, S. Singh, and C. Guestrin, "“why should i trust you?” explaining the predictions of any classifier,” in *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*, 2016, pp. 1135–1144.
- [165] R. R. Selvaraju, A. Das, R. Vedantam, M. Cogswell, D. Parikh, and D. Batra, “Grad-cam: Why did you say that?” *arXiv preprint arXiv:1611.07450*, 2016.
- [166] G. Lopardo, F. Precioso, and D. Garreau, “Understanding post-hoc explainers: The case of anchors,” *arXiv preprint arXiv:2303.08806*, 2023.
- [167] M. H. N. Amiri, F. Annaz, M. De Oliveira, and F. Gueniat, “Strategies for resilience and battery life extension in the face of communication losses for isolated microgrids,” in *2024 IEEE International Conference on Omni-layer Intelligent Systems (COINS)*, IEEE, 2024, pp. 1–5.
- [168] M. H. N. Amiri, F. Annaz, M. D. Oliveira, and F. Gueniat, *Deep reinforcement learning with local interpretability for transparent microgrid resilience energy management*, 2025. arXiv: [2508.08132](https://arxiv.org/abs/2508.08132) [eess.SY]. [Online]. Available: <https://arxiv.org/abs/2508.08132>.
- [169] C. Phurailatpam, B. S. Rajpurohit, and L. Wang, “Planning and optimization of autonomous dc microgrids for rural and urban applications in india,” *Renewable and Sustainable Energy Reviews*, vol. 82, pp. 194–204, 2018.
- [170] R. Sen and S. C. Bhattacharyya, “Off-grid electricity generation with renewable energy technologies in india: An application of homer,” *Renewable energy*, vol. 62, pp. 388–398, 2014.
- [171] European Commission, Joint Research Centre, *Photovoltaic geographical information system (pvgis)*, Accessed: 2024-09-03, 2024. [Online]. Available: https://re.jrc.ec.europa.eu/pvg_tools/en/.

-
- [172] T. Huld, R. Müller, and A. Gambardella, “A new solar radiation database for estimating pv performance in europe and africa,” *Solar energy*, vol. 86, no. 6, pp. 1803–1815, 2012.
- [173] K. S. Anderson, C. W. Hansen, W. F. Holmgren, A. R. Jensen, M. A. Mikofski, and A. Driesse, “Pvlib python: 2023 project update,” *Journal of Open Source Software*, vol. 8, no. 92, p. 5994, 2023. DOI: [10.21105/joss.05994](https://doi.org/10.21105/joss.05994). [Online]. Available: <https://doi.org/10.21105/joss.05994>.
- [174] WindPowerLib_Developers, *Windpowerlib: A python library for wind power simulation and analysis*, Accessed: 2024-09-03, 2024. [Online]. Available: <https://windpowerlib.readthedocs.io/en/latest/>.
- [175] N. E. Observatory, *Tropical cyclone laila soaks india*, <https://earthobservatory.nasa.gov/images/44045/tropical-cyclone-laila-soaks-india>, Accessed: 29 August 2025, May 2010.
- [176] N. F. Q. Amran, W. N. A. A. W. Zulkifli, N. Jaalam, F. Arith, and A. S. M. Shah, “Investigation of standard test condition requirement in establishing alternative measurement platform for photovoltaic cell,” in *Journal of Physics: Conference Series*, IOP Publishing, vol. 2928, 2024, p. 012 003.
- [177] B. K. Das, N. Hoque, S. Mandal, T. K. Pal, and M. A. Raihan, “A techno-economic feasibility of a stand-alone hybrid power generation for remote area application in bangladesh,” *Energy*, vol. 134, pp. 775–788, 2017.
- [178] S. Singh and S. C. Kaushik, “Optimal sizing of grid integrated hybrid pv-biomass energy system using artificial bee colony algorithm,” *IET Renewable Power Generation*, vol. 10, no. 5, pp. 642–650, 2016.

-
- [179] S. Singh, M. Singh, and S. C. Kaushik, "Feasibility study of an islanded microgrid in rural area consisting of pv, wind, biomass and battery energy storage system," *Energy Conversion and Management*, vol. 128, pp. 178–190, 2016.
- [180] P. Malik, M. Awasthi, and S. Sinha, "A techno-economic investigation of grid integrated hybrid renewable energy systems," *Sustainable Energy Technologies and Assessments*, vol. 51, p. 101976, 2022.
- [181] S. Dhundhara, Y. P. Verma, and A. Williams, "Techno-economic analysis of the lithium-ion and lead-acid battery in microgrid systems," *Energy conversion and management*, vol. 177, pp. 122–142, 2018.
- [182] Waaree. "1kw on-grid solar system price in india." [Online]. Available: <https://waaree.com/blog/1kw-on-grid-solar-system-price-in-india-2023/>.
- [183] S. C. Bhattacharyya, "Mini-grid based electrification in bangladesh: Technical configuration and business analysis," *Renewable Energy*, vol. 75, pp. 745–761, 2015.
- [184] P. Patel. "Lithium-ion battery prices in india from 1kwh to 10 kwh." [Online]. Available: <https://e-vehicleinfo.com/lithium-ion-battery-prices-in-india/>.
- [185] A. Gailani, "Modelling and degradation cost assessment of batteries in capacity markets," Ph.D. dissertation, Teesside University, 2022.
- [186] A. V. Vykhodtsev, D. Jang, Q. Wang, W. Rosehart, and H. Zareipour, "A review of modelling approaches to characterize lithium-ion battery energy storage systems in techno-economic analyses of power systems," *Renewable and Sustainable Energy Reviews*, vol. 166, p. 112584, 2022.
- [187] W. Cole, V. Ramasamy, and M. Turan, "Cost projections for utility-scale battery storage: 2025 update," National Renewable Energy Laboratory, Tech. Rep. NREL/TP-6A40-93281, Jun. 2025. [Online]. Available: <https://www.osti.gov/servlets/purl/2583471>.

- [188] S. J. Ericson and D. R. Olis, “A comparison of fuel choice for backup generators,” National Renewable Energy Laboratory (NREL), Golden, CO (United States), Tech. Rep., 2019.
- [189] Caterpillar, *Cat xq60 generator set technical data*, Technical data sheet for 60 kW / 75 kVA standby diesel generator fuel consumption, 2024. [Online]. Available: https://www.ers-cat.com/application/files/7116/5548/2907/ERSSpecSheets_60kw_SpecSheet.pdf.
- [190] Department for Energy Security and Net Zero, *Greenhouse gas reporting: Conversion factors 2025*, UK Government conversion factors for greenhouse gas reporting, Jun. 2025. [Online]. Available: <https://www.gov.uk/government/publications/greenhouse-gas-reporting-conversion-factors-2025>.
- [191] H. Qiu et al., “Resilience-directional robust power dispatching of microgrids under meteorological disasters,” *IET Renewable Power Generation*, vol. 13, no. 12, pp. 2084–2093, 2019.