

Master Thesis

Forecasting and Optimization for Gas Turbine Participation in the aFRR Market

Huining Liu

Chalmers University of Technology

Supervisor: Ann-Brith Strömberg
Industrial Supervisor: Awais Salmon

October 2, 2026

Contents

1	Introduction	3
1.1	Background	3
1.2	Problem Statement	4
1.3	Research Objectives	4
1.4	Research Questions	5
1.5	Scope and Limitations	5
1.6	Thesis Structure	6
2	Background and Literature Review	7
2.1	Forecasting in Electricity and Reserve Markets	7
2.2	Optimization under Uncertainty	8
2.3	Research Gap	9
3	Data and Problem Setup	11
3.1	Data Sources	11
3.2	Data Preparation	12
3.3	Definition of Targets and Optimization Inputs	13
3.4	Gas Turbine Parameterization	14
4	Methodology	16
4.1	Overall Framework	16
4.2	Forecasting and Scenario Generation	16
4.3	Deterministic Optimization Model	18
4.3.1	Sets, Parameters, and Variables	18
4.3.2	Objective Function	19
4.3.3	Constraints	20
4.4	Stochastic Optimization Model	20
4.4.1	Sets, Parameters, and Variables	21
4.4.2	Objective Function	21
4.4.3	Constraints	21
4.4.4	Interpretation of the Activation-Ratio Formulation	22
4.5	Comparison of the Deterministic and Stochastic Formulations	22
4.6	Model Assumptions	23
5	Implementation	25
5.1	Model Implementation	25
5.2	Benchmark Strategies	26
5.3	Evaluation Design	27
5.4	Computational Setup and Instance Size	29

6	Results	31
6.1	Multi-Day Backtesting Results	31
6.2	Deterministic vs Stochastic Comparison	35
6.3	Benchmark Strategy Comparison	39
6.4	Sensitivity Analysis	41
6.5	Case Studies	44
7	Discussion	48
7.1	Interpretation of Main Findings	48
7.2	Strengths of the Current Model	49
7.3	Limitations	50
7.4	Implications for Practice	51
7.5	Future Work	52
8	Conclusion	54
8.1	Summary of the Thesis	54
	Bibliography	56

Chapter 1

Introduction

1.1 Background

The ongoing energy transition is fundamentally changing the operation of modern power systems. The increasing penetration of variable renewable energy sources, such as wind and solar power, introduces larger forecast errors and stronger short-term variability in net load. This increases the need for operational flexibility and balancing services [12, 2]. Unlike conventional thermal generation, renewable generation is weather-dependent and only partially controllable, which makes it more difficult for system operators to maintain the real-time balance between electricity supply and demand.

One important way of providing this flexibility is through ancillary services and balancing reserves. In the European electricity-balancing framework, automatic Frequency Restoration Reserve (aFRR) is used to restore the system balance after frequency deviations and is increasingly coordinated through common European balancing platforms such as PICASSO [7, 8]. These markets create opportunities for flexible assets, but they also expose participants to uncertain reserve prices and uncertain activation.

Gas turbines remain highly relevant in this context because they combine relatively fast response capability with operational flexibility. Compared with many other conventional units, gas turbines can change output quickly, provide reserve products, and adapt to fluctuating market signals. This makes them attractive participants in balancing markets, especially in a system with increasing renewable penetration. At the same time, their participation in reserve markets is not trivial: providing reserve requires careful coordination between technical feasibility, operating cost, and market profitability.

From an economic perspective, participation in the aFRR market has become increasingly important because it creates an additional value stream beyond energy-market revenues. A gas turbine operator may earn revenue not only from selling electricity in the day-ahead (DA) market, but also from reserving upward and downward balancing capacity. However, reserve participation also introduces trade-offs. Capacity committed to reserve may reduce the flexibility to optimize energy-market output, and activation uncertainty may affect actual dispatch and operating cost. Therefore, determining an economically attractive and technically feasible strategy for a gas turbine in the aFRR market is a relevant and timely optimization problem.

1.2 Problem Statement

This thesis addresses the problem of how a gas turbine should participate in the aFRR market under uncertainty. In practice, the operator must decide the unit's base generation schedule and reserve offers before the actual market conditions and activation requirements are fully known. These decisions must be made while considering both market-side uncertainty and plant-side technical constraints.

The first source of difficulty lies in uncertain reserve prices. Although reserve markets offer potential additional revenue, the value of upward and downward reserve varies over time and is difficult to predict perfectly. A strategy that is attractive under one price realization may become suboptimal under another. As a result, reserve allocation cannot be based solely on deterministic expectations without potentially exposing the unit to economic inefficiency.

The second source of uncertainty is activation. Even if reserve capacity is offered and accepted, the actual fraction of reserve that will be called is uncertain and may differ substantially across time periods and operating conditions. This means that the operator must not only decide how much reserve to provide, but also consider how potential activation affects the unit's realized dispatch and cost.

A further challenge comes from the interaction between energy and reserve markets. If the gas turbine allocates more capacity to energy production, it may earn higher day-ahead electricity revenue, but it may lose the flexibility required to provide upward reserve. Conversely, maintaining significant reserve capability may constrain the preferred energy-market operating point. Therefore, reserve participation should not be modeled independently of energy-market incentives.

Finally, all decisions must remain technically feasible. Gas turbines are subject to minimum and maximum output limits, ramping constraints, and reserve capability limits. These technical restrictions directly affect how much reserve can be offered and how the unit can respond under activation. The practical problem is therefore to find a strategy that is both economically sound and operationally feasible across uncertain market scenarios.

1.3 Research Objectives

The overall objective of this thesis is to develop and evaluate a decision-support framework for gas turbine participation in the aFRR market. The framework should combine market-related forecasting inputs with optimization-based operational planning in order to determine profitable and technically feasible reserve strategies.

More specifically, the thesis has the following objectives:

- to develop forecasting-based inputs for the optimization model, including reserve market signals, activation-related quantities, and day-ahead electricity price information,
- to formulate both deterministic and stochastic optimization models for joint base-generation scheduling and reserve allocation,
- to represent activation uncertainty in a tractable way that is suitable for single-unit optimization,

- to analyze the trade-off between energy-market participation and reserve provision,
- to evaluate the resulting strategies through backtesting, model comparison, benchmark comparison, and sensitivity analysis.

By achieving these objectives, the thesis aims to provide both methodological insight and practical operational interpretation for gas turbine reserve participation.

1.4 Research Questions

To guide the analysis, the thesis is structured around the following research questions:

- How should a gas turbine allocate base generation and reserve capacity in the aFRR market under uncertain prices and activation conditions?
- How can activation uncertainty be represented in a tractable way for single-unit reserve optimization?
- Does scenario-aware scheduling improve the commonly evaluated objective relative to an EV policy, and how should that difference be interpreted?
- How does the inclusion of day-ahead electricity market value affect the optimal reserve scheduling strategy?
- How sensitive are the optimization results to market assumptions, technical constraints, and activation-ratio calibration?

These questions are intended to connect the technical modeling work with broader operational and economic insights.

1.5 Scope and Limitations

This thesis focuses on the reserve scheduling problem of a single gas turbine participating in the aFRR market. The purpose is not to build a full market-clearing simulator, but rather to construct a tractable and interpretable optimization model that supports single-unit operational decision making.

Several scope limitations are introduced deliberately. First, the model focuses on single-unit optimization rather than portfolio-level coordination across multiple units. Second, activation is represented through a simplified activation-ratio approach, where the realized activation is modeled as a fraction of the offered reserve rather than as a full system-level dispatch assignment. This formulation is intended to provide a practical approximation of activation intensity while avoiding unrealistic assumptions that a single turbine must meet total system balancing demand.

Third, the base version of the model does not include startup and shutdown decisions. The gas turbine is assumed to be available during the optimization horizon, which keeps the formulation continuous and avoids the additional complexity of unit commitment variables and binary commitment logic. Fourth, market settlement is represented in simplified form. The model captures day-ahead electricity value and reserve-related revenue, but does not fully replicate all real-world market-clearing and balancing settlement mechanisms.

These limitations mean that the model should be interpreted as a research-oriented operational prototype rather than a complete industrial bidding engine. Nevertheless, the framework is sufficiently rich to analyze the interaction between uncertainty, technical feasibility, and economic decision making.

1.6 Thesis Structure

The remainder of this thesis is organized as follows.

Chapter 2 reviews the relevant background and literature on ancillary service markets, gas turbine flexibility, reserve participation, and optimization under uncertainty. It positions the thesis within the existing research field and identifies the main research gap addressed in this work.

Chapter 3 presents the data and problem setup. It describes the data sources, preprocessing steps, key variables, and the representation of gas turbine technical parameters and market inputs.

Chapter 4 introduces the methodology. It explains the forecasting and scenario-generation logic, formulates the deterministic and stochastic optimization models, and describes the key modeling assumptions.

Chapter 5 describes the implementation of the proposed framework, including solver setup, benchmark strategies, and the evaluation design used for backtesting and model comparison.

Chapter 6 presents the results. It includes representative single-day illustrations, multi-day backtesting results, deterministic versus stochastic comparisons, benchmark strategy comparisons, and sensitivity analysis.

Chapter 7 discusses the implications of the findings, the strengths and limitations of the current modeling framework, and possible future improvements.

Finally, Chapter 8 concludes the thesis by summarizing the main contributions and outlining potential directions for future work.

Chapter 2

Background and Literature Review

2.1 Forecasting in Electricity and Reserve Markets

Forecasting plays a central role in electricity-market participation because operational and bidding decisions must typically be made before market outcomes are fully realized. For power producers, profitability depends not only on plant efficiency and technical flexibility, but also on the ability to anticipate price signals and system conditions with reasonable accuracy. In this sense, forecasting is a necessary upstream component of optimization-based decision making.

In electricity markets, a large body of literature has studied price forecasting methods. Early work often relied on classical time-series models such as ARIMA, which provided transparent baselines for short-term electricity price prediction [5]. As market behavior became increasingly nonlinear and volatile, more advanced machine learning and deep learning methods gained prominence. Architectures such as LSTM, GRU, and hybrid CNN-LSTM models have been shown to improve predictive performance by capturing temporal dependencies and nonlinear price dynamics [10]. More recently, Transformer-based approaches have attracted substantial attention because their self-attention mechanism allows them to model long-range dependencies more effectively in complex energy-market time series [3].

Electricity price forecasting is a mature research field with a long tradition of statistical, machine-learning, and deep-learning methods. Reviews by [16] and [10] show that electricity prices are difficult to forecast because of seasonality, volatility, spikes, and the non-storable nature of electricity. Classical time-series and regression-based models remain important baselines, while modern machine-learning methods can capture nonlinear and high-dimensional dependencies [9, 17].

However, forecasting in balancing and reserve markets is less mature than day-ahead electricity price forecasting. Balancing prices are closer to real time, more volatile, and more directly linked to system imbalance and activation needs. Recent work on balancing market forecasting emphasizes that models successful in day-ahead markets do not necessarily transfer directly to balancing markets [13]. This motivates treating aFRR price forecasting as a separate prediction problem rather than as a simple extension of day-ahead price forecasting.

One common way to handle this challenge is through uncertainty representation and scenario generation. Rather than relying only on point forecasts, scenario-based approaches preserve multiple possible future realizations of prices and activation-related conditions. These scenarios can then be used as inputs to stochastic optimization mod-

els. From a methodological perspective, scenario generation is therefore the link between predictive modeling and operational decision making under uncertainty.

In the context of this thesis, forecasting is not treated as an isolated prediction task, but as a functional input to optimization. The purpose of reserve-price forecasts, day-ahead price inputs, and activation-related scenarios is not only to reduce forecast error, but to enable better operational decisions for a gas turbine participating in the aFRR market.

2.2 Optimization under Uncertainty

Because reserve-market decisions must be made before the actual market and activation conditions are known, optimization under uncertainty is a natural framework for this problem. Stochastic programming provides a standard modelling approach for sequential decision problems in which first-stage decisions are made before uncertainty is revealed and second-stage decisions adapt after scenario realization [1, 4]. In electricity markets, this structure is particularly relevant because bidding, scheduling, and reserve-allocation decisions are often made before the actual prices, renewable output, imbalance state, or activation volumes are known [12].

A deterministic optimization model replaces uncertain inputs by point forecasts or expected values. This is computationally attractive and easy to interpret, but it does not explicitly represent the distribution of possible outcomes. A stochastic formulation, instead, optimizes expected performance over a set of scenarios and can therefore capture the effect of uncertainty on reserve and dispatch decisions. Similar deterministic, stochastic, and robust comparisons have been studied for reserve-providing flexible resources, where activation uncertainty is a central modelling challenge [15].

A stochastic optimization approach addresses this limitation by representing uncertainty through multiple scenarios and optimizing decisions across their joint probability distribution. In stochastic programming, the objective is typically to maximize expected value or minimize expected cost while ensuring that decisions remain feasible across a range of possible future outcomes. This framework is particularly suitable for reserve scheduling because both reserve prices and activation intensity are uncertain at the time reserve offers are made.

Among stochastic optimization methods, two-stage stochastic programming is especially relevant. In a two-stage setting, first-stage decisions are made before uncertainty is realized, while second-stage decisions adapt after the scenario becomes known. In power-system applications, this structure is widely used because it reflects the real sequence of operational decisions: producers first commit or schedule capacity and only later observe the actual market realization. A standard reference on uncertainty-aware decision making in power systems is provided by [4], who discuss the role of stochastic programming and related methods in energy applications.

In balancing-market participation, stochastic programming has been used to support bidding decisions under uncertain prices and demand conditions. For example, risk-aware stochastic bidding formulations often combine expected-profit maximization with downside-risk measures such as Conditional Value-at-Risk (CVaR) [11]. Such extensions are valuable when decision makers care not only about average profitability but also about avoiding poor outcomes under adverse market conditions.

Nevertheless, stochastic optimization also has practical challenges. Its quality depends

strongly on scenario construction, probability calibration, and the realism of the uncertainty representation. If the scenario set is poorly calibrated or overly conservative, the resulting strategy may sacrifice profitability without delivering a meaningful robustness advantage. This issue is particularly relevant for single-unit reserve scheduling, where activation uncertainty must often be approximated rather than directly observed.

For this reason, the current thesis adopts a pragmatic approach. It compares a deterministic baseline with a two-stage stochastic model in order to evaluate whether explicit uncertainty modeling improves reserve scheduling performance. At the same time, the thesis recognizes that the value of stochastic optimization depends not only on the mathematical framework itself, but also on how well the underlying uncertainty inputs are constructed.

2.3 Research Gap

Existing literature provides strong foundations in electricity price forecasting, reserve-market analysis, and optimization under uncertainty. However, an important gap remains at the intersection of these fields. Much of the forecasting literature focuses on prediction accuracy alone, while many optimization studies assume simplified or exogenously given market inputs. As a result, the practical connection between forecasting outputs and plant-level reserve scheduling decisions is often underdeveloped.

A first gap concerns the limited integration of reserve-market forecasting with operational optimization. Compared with the large literature on day-ahead electricity price forecasting, fewer studies explicitly address how reserve price uncertainty and activation uncertainty should be translated into decision-ready inputs for unit-level scheduling. In particular, for balancing products such as aFRR, the challenge is not only to predict prices, but also to represent how likely and how strongly reserve is expected to be activated.

A second gap concerns the representation of activation at the single-unit level. In many practical settings, historical datasets provide system-level balancing signals rather than exact unit-level activation assignments. This makes it difficult to directly embed activation uncertainty into gas turbine scheduling models. Existing work often either ignores this issue or assumes simplified balancing behavior without clearly connecting system-level uncertainty to unit-level dispatch feasibility.

A third gap concerns the evaluation of deterministic versus stochastic formulations in a practically interpretable setting. While stochastic optimization is widely recognized as a theoretically appealing framework, its practical value depends heavily on the quality of the underlying scenarios and calibration. There is therefore a need for empirical comparison frameworks that test whether stochastic reserve scheduling actually outperforms deterministic approaches and heuristic benchmark strategies under realistic market input assumptions.

The literature therefore provides strong foundations in electricity price forecasting, stochastic programming, and reserve-market modelling. Nevertheless, a gap remains at the intersection of these fields. Much of the electricity-price forecasting literature focuses on prediction accuracy, while many optimization studies assume that uncertainty scenarios are already available. For practical aFRR participation, however, the full decision pipeline matters: market data must be transformed into forecasts, forecasts must be converted into decision-ready scenarios, and the resulting schedules must be evaluated

economically and computationally.

This thesis addresses this gap by linking forecasting-derived reserve-price and activation inputs with deterministic and stochastic gas-turbine scheduling models. The emphasis is not only on whether stochastic optimization can improve expected performance, but also on whether the additional modelling and computational effort is justified in a practical single-unit aFRR scheduling setting.

First, it combines day-ahead electricity value, reserve prices, and activation-related uncertainty in a single operational scheduling framework. Second, it introduces an activation-ratio formulation that provides a tractable approximation of reserve activation intensity for single-unit optimization. Third, it evaluates deterministic optimization, stochastic optimization, and benchmark heuristics in a unified multi-day backtesting setup.

In this way, the thesis moves beyond isolated forecasting or isolated optimization and instead focuses on the full decision pipeline relevant for gas turbine participation in balancing markets. While hydrogen-related operational effects motivated part of the broader project context, the present thesis concentrates on the market-oriented scheduling problem itself, with the aim of providing a tractable and empirically grounded decision-support model for aFRR participation.

Chapter 3

Data and Problem Setup

3.1 Data Sources

The empirical dataset is mainly based on publicly available European electricity-market and balancing-market data from the ENTSO-E Transparency Platform [6]. The platform provides market, load, generation, transmission, and balancing-related data that are commonly used in European power-market studies. Weather variables are obtained from Open-Meteo [14], while commodity and carbon-price series are used as additional exogenous market indicators.

The day-ahead electricity price for SE3 is used as the economic signal for planned energy production. In addition to the SE3 price, day-ahead prices from neighbouring Swedish bidding zones SE1, SE2, and SE4 are included as explanatory variables in the forecasting stage. These neighbouring prices are not direct optimization targets, but they help represent spatial market conditions in the Nordic electricity system.

In this thesis, the *expected-value problem* (EV) denotes the deterministic optimization problem obtained by replacing the scenario-dependent reserve-price and activation-ratio inputs with their probability-weighted mean trajectories. Solving EV yields the optimal first-stage decision vector $x^{\text{EV}} = (p^{\text{base}}, r^{\text{up}}, r^{\text{down}})$. Its internal objective value, denoted by z^{EV} , describes performance under the mean inputs and is not used directly as the evaluated performance of the EV policy. The expected result of using the EV solution (EEV) is obtained by fixing x^{EV} and evaluating it with the same full set of 50 scenarios, scenario probabilities, recourse equations, and feasibility conditions used for the stochastic recourse problem (RP). The value of the stochastic solution is then $\text{VSS} = \text{RP} - \text{EEV}$. Section 4.5 gives the formal definitions.

The aFRR market data include upward and downward contracted reserve prices and corresponding reserve quantities for SE3. The main reserve-price targets are the upward and downward aFRR capacity prices. These time series are used first as historical targets in the aFRR price forecasting model and then to generate scenario realizations. The RP model uses those scenarios directly; the EV model uses their probability-weighted trajectories so that both formulations refer to the same discrete distribution.

Activation-related uncertainty is represented using system imbalance and balancing-market information. Exact unit-level activation instructions for the gas turbine are not available in the dataset. Therefore, the thesis uses system-level imbalance and activated balancing indicators as proxies for activation conditions. These data are transformed into activation states and activation-ratio inputs that approximate the fraction of offered reserve that may be activated in each interval.

Additional explanatory variables are included to improve forecasting performance. These include actual load, load forecasts, wind and solar forecasts, actual generation by technology, cross-border flow indicators, weather observations for northern and southern Sweden, natural gas prices, and EU carbon allowance prices. The weather variables include temperature and 100-meter wind speed from Open-Meteo. Gas price data are obtained through financial market data, while carbon allowance prices are based on historical EUA price data.

Table 3.1 summarizes the main data categories used in the thesis.

Table 3.1: Main data sources and their role in the modelling framework.

Data category	Main variables	Role in thesis	Resolution
Day-ahead market	SE3 day-ahead price; SE1, SE2, SE4 neighbouring prices	Energy value signal and forecasting features	15 min aligned
aFRR capacity market	Upward and downward reserve prices and quantities	Forecasting targets and optimization revenue inputs	15 min
Balancing and imbalance data	Imbalance indicators, activated balancing price, imbalance volume	Proxy information for activation uncertainty	15 min aligned
System data	Actual load, load forecast, generation mix, wind and solar forecasts	Explanatory variables for forecasting	15 min aligned
Weather data	Temperature and 100 m wind speed for northern and southern Sweden	Exogenous forecasting features	15 min aligned
Commodity and carbon data	Natural gas price and EUA carbon price	Cost- and market-related explanatory variables	15 min aligned
Gas turbine data	Generation limits, ramping limits, reserve capability, variable cost	Technical and economic optimization parameters	Static parameters

3.2 Data Preparation

The raw datasets are first converted to a common timestamp format, sorted chronologically, deduplicated, and aligned to a 15-minute UTC index. Hourly or daily data are forward-filled to the 15-minute grid when this is economically reasonable, for example for day-ahead prices, commodity prices, and carbon prices. This alignment is necessary because the optimization model makes one decision for each 15-minute delivery interval, resulting in 96 intervals per delivery day.

Missing values are handled differently depending on the data type. Critical target variables, such as day-ahead prices and aFRR reserve prices, are required to be available for the forecasting and optimization stages. For explanatory variables, limited imputation is used where appropriate. Commodity and carbon prices are forward-filled or backfilled

over short gaps. Weather observations and renewable forecasts are imputed over short windows when the missingness is limited. Sparse balancing-market features are retained only when they provide useful information and are not used in a way that would introduce leakage.

The final unified modelling table, `df_model_unified.parquet`, contains 44 columns and 111,840 observations. Table 3.2 reports the coverage of selected key variables after harmonization.

Table 3.2: Coverage and summary statistics of selected variables in the unified dataset.

Variable	Unit	Coverage	Min.	Mean	Max.
Day-ahead price, SE3	EUR/MWh	100.0%	-60.04	47.70	707.52
aFRR upward reserve price	EUR/MW	100.0%	0.00	21.50	575.78
aFRR downward reserve price	EUR/MW	100.0%	0.00	27.07	5389.86
aFRR upward reserve amount	MW	100.0%	0.00	0.29	30.00
aFRR downward reserve amount	MW	100.0%	0.00	3.67	60.00
Actual load	MW	100.0%	5198	9655	16747
Load forecast	MW	100.0%	5516	9636	15827
Wind onshore forecast	MW	99.996%	19.30	1137.69	3637.20
Solar forecast	MW	99.996%	0.00	114.31	1180.43
Natural gas price	EUR/MWh	99.74%	22.93	37.30	74.30
Carbon EUA price	EUR/tCO ₂	99.91%	52.51	75.35	97.47

The forecasting models use a chronological split to avoid information leakage. The first 70% of the dataset is used for training, the next 15% for validation, and the final 15% for testing. No random shuffling is applied, since the problem is time-series based.

Table 3.3: Chronological split of the unified modelling dataset.

Subset	Intervals	Start	End
Training	78,288	2023-01-01 00:00	2025-03-26 11:45
Validation	16,776	2025-03-26 12:00	2025-09-17 05:45
Test	16,776	2025-09-17 06:00	2026-03-10 23:45

3.3 Definition of Targets and Optimization Inputs

The main day-ahead forecasting target is the SE3 day-ahead electricity price. This price enters the optimization model as the revenue signal for planned base generation. Forecasting features include calendar variables, lagged prices, rolling statistics, neighbouring bidding-zone prices, load and renewable forecasts, weather variables, gas prices, carbon prices, and selected balancing-market indicators.

For aFRR price forecasting, the two main regression targets are the upward and downward reserve prices:

$$\lambda_t^{\text{up}} \quad \text{and} \quad \lambda_t^{\text{down}}.$$

These are first predicted as point forecasts and then used to generate scenario-dependent price trajectories. The RP model uses the daily scenario bundle, whereas the EV model uses the probability-weighted mean price and activation trajectories derived from that same bundle.

Activation uncertainty is represented through a three-state proxy: upward activation, neutral state, and downward activation. The state labels are constructed from imbalance-related proxy variables rather than exact unit-level activation instructions. The resulting daily prediction output contains, for each 15-minute interval, predicted upward and downward reserve prices and probabilities for the three activation states:

$$p_t^{\text{up}}, \quad p_t^{\text{neutral}}, \quad p_t^{\text{down}}.$$

For the stochastic model, these probabilities are used to construct daily scenario bundles. Each scenario uses one identifier and contains 96 interval-level realizations of upward reserve price, downward reserve price, activation state, and activated volume. This identifier groups the sampled variables but does not imply that their full temporal or cross-variable dependence is preserved. The final scenario bundle used in the optimization experiments covers 42 delivery days from 28 January 2026 to 10 March 2026. Each day contains 96 intervals and 50 scenarios, giving 201,600 scenario-interval observations in total.

The activation-ratio formulation translates scenario-level activation information into a fraction of offered reserve that is activated. For scenario s and interval t , the realized activated reserve is modelled as

$$a_{t,s}^{\text{up}} = \alpha_{t,s}^{\text{up}} r_t^{\text{up}}, \quad a_{t,s}^{\text{down}} = \alpha_{t,s}^{\text{down}} r_t^{\text{down}},$$

where r_t^{up} and r_t^{down} are first-stage reserve decisions, while $\alpha_{t,s}^{\text{up}}$ and $\alpha_{t,s}^{\text{down}}$ are scenario-dependent activation ratios. This formulation avoids assuming that the single gas turbine must follow total system-level activation volumes. Instead, activation is interpreted as a proportional call on the reserve capacity offered by the unit.

3.4 Gas Turbine Parameterization

The gas turbine is represented by a simplified set of technical and economic parameters. These parameters are based on information provided by the industrial supervisor and are reported in anonymized form due to confidentiality considerations. The exact values are not disclosed; instead, rounded ranges and normalized quantities are used. This preserves the operational meaning of the model while avoiding publication of plant-specific data.

The minimum and maximum generation limits define the feasible operating range of the unit. The minimum stable generation level is approximately one third of the maximum output. The ramping capability is high relative to the unit size, meaning that the turbine can change output substantially within one 15-minute interval. This is consistent with the role of gas turbines as flexible assets in balancing-market applications.

Reserve capability is represented through upward and downward reserve limits. The upward reserve capability is of the same order as, but smaller than, the total output range of the unit. Downward reserve capability is treated analogously in the optimization model. Variable operating cost is represented by a short-run marginal cost term, while startup and shutdown costs are not included in the base optimization model.

The present thesis does not include startup and shutdown decisions in the optimization model. Therefore, C^{start} is reported for completeness but is not active in the base formulation. Including startup and shutdown would require binary commitment variables, minimum up-time and down-time constraints, and startup cost terms. Such an exten-

Table 3.4: Anonymized gas turbine parameterization used to define the optimization model.

Parameter	Description	Reported form	Unit
$P^{\min}$	Minimum stable generation	approximately 25–35% of $P^{\max}$	MW
$P^{\max}$	Maximum generation capacity	anonymized reference capacity	MW
RU	Maximum upward ramp within one interval	approximately 0.9–1.1 $P^{\max}$	MW/15 min
RD	Maximum downward ramp within one interval	approximately 0.8–1.0 $P^{\max}$	MW/15 min
$R^{\max, \text{up}}$	Maximum upward reserve capacity	approximately 35–45% of $P^{\max}$	MW
$R^{\max, \text{down}}$	Maximum downward reserve capacity	assumed comparable order unless otherwise stated	MW
c^{var}	Variable operating cost component	low single-digit value, adjusted by a calibration factor	EUR/MWh
C^{start}	Startup cost	order of 10^3	EUR/start

sion would change the model from a linear programming problem into a mixed-integer optimization problem and would likely increase computational complexity substantially.

In the computational experiments, the plant is assumed to be available throughout each daily optimization horizon. The model therefore studies the reserve scheduling problem conditional on the gas turbine being online. This assumption keeps the focus on the interaction between base generation, reserve provision, activation uncertainty, and market prices.

Chapter 4

Methodology

This chapter presents the modelling framework used to connect forecasting outputs with optimization-based aFRR scheduling decisions. Section 4.1 gives an overview of the end-to-end framework, while Section 4.2 describes the forecasting and scenario generation layer. Sections 4.3 and 4.4 formulate the deterministic and stochastic optimization models, respectively. Section 4.5 compares the two formulations, and Section 4.6 summarizes the main modelling assumptions.

4.1 Overall Framework

The methodology used in this thesis consists of four main components: forecasting, scenario generation, optimization, and evaluation. These components form an end-to-end workflow that converts historical and exogenous market information into operational reserve scheduling decisions for a gas turbine.

First, forecasting models are used to estimate reserve-related market signals and activation-related quantities. Second, these outputs are transformed into stochastic scenario bundles and probability-weighted EV inputs. Third, optimization models use these inputs to determine base generation and reserve allocation decisions. Finally, the resulting strategies are evaluated using backtesting, model comparison, benchmark comparison, and sensitivity analysis.

This workflow makes it possible to connect uncertain market conditions with operational decision making in a structured and transparent way.

4.2 Forecasting and Scenario Generation

Reserve Price Forecasts

The forecasting layer produces 96 quarter-hourly point forecasts for upward and downward aFRR reserve prices for each delivery day. Price uncertainty is represented by a residual bootstrap, a standard way of converting a point forecast into an empirical predictive distribution when the residual sample is informative [16, 12]. For direction $d \in \{\text{up}, \text{down}\}$, interval t , and scenario s , the generated price is

$$\lambda_{t,s}^d = \hat{\lambda}_t^d + \gamma_d \varepsilon_{t,s}^d, \quad (4.1)$$

where residuals are sampled with replacement from a validation residual pool conditioned on hour of day (or from the full pool when an hourly pool is too small). The scale factor γ_d is selected on the validation sample to target 80% coverage between the empirical 0.1 and 0.9 scenario quantiles. Each scenario combines sampled reserve-price residuals, activation states, and conditional activation volumes under a common scenario identifier. The current implementation mainly preserves hour-conditioned residual distributions, interval-specific activation-state probabilities, and state-conditional activation-volume distributions. Full temporal dependence and cross-variable correlation are not explicitly modelled and remain subjects for future validation.

Activation-Ratio Construction

Because exact unit-level activation instructions are unavailable, the three-state activation process (up, neutral, or down) is sampled from the interval-specific class probabilities. Conditional activation volumes are then converted to bounded directional ratios using the 95th percentile of the corresponding daily scenario volumes as the normalization scale:

$$\alpha_{t,s}^d = \mathbf{1}\{z_{t,s} = d\} \min\left\{\max\left(\frac{V_{t,s}}{q_{0.95}^d}, 0\right), 1\right\}. \quad (4.2)$$

The requested activation is linked to offered reserve through

$$a_{t,s}^d + u_{t,s}^d = \alpha_{t,s}^d r_t^d, \quad d \in \{\text{up}, \text{down}\}, \quad (4.3)$$

where $u_{t,s}^d \geq 0$ is a penalized delivery shortfall. The same shortfall definition and penalty are used in the stochastic recourse problem and when evaluating the expected-value solution. This provides feasible recourse without changing the activation representation for one model only.

This formulation provides a tractable approximation of reserve activation while avoiding unrealistic assumptions that a single unit must satisfy the full system-level balancing demand.

Day-Ahead Price Integration

The DA price trajectory λ_t^{DA} is treated as a common, non-random input in all scenarios. It is taken from the same day-ahead point-forecast file for EV, EEV, and RP and is aligned by delivery day and quarter-hour slot. It is therefore not averaged over the aFRR scenarios and enters every formulation identically.

Scenario Generation Method

Each daily bundle contains 50 scenario trajectories and 96 intervals per trajectory. The Monte Carlo trajectories are equiprobable, $\pi_s = 1/50$, so that

$$\pi_s \geq 0, \quad \sum_{s \in S} \pi_s = 1. \quad (4.4)$$

The bootstrap residuals are not forced to have zero sample mean after drawing. The implementation therefore adopts the probability-consistent alternative of defining the

deterministic expected-value trajectory directly from the realized scenario bundle:

$$\lambda_t^{d,\text{EV}} = \sum_{s \in S} \pi_s \lambda_{t,s}^d, \quad (4.5)$$

$$\alpha_t^{d,\text{EV}} = \sum_{s \in S} \pi_s \alpha_{t,s}^d, \quad (4.6)$$

for both directions. This is preferable to combining the scenarios with an independently stored point forecast, because the EV and recourse problems then refer to exactly the same discrete probability distribution [1]. Automated checks verify the day and slot keys, complete scenario coverage, non-negative normalized probabilities, and equality of Equations (4.5)–(4.6) to numerical precision.

4.3 Deterministic Optimization Model

The deterministic model is the expected-value (EV) formulation of the daily reserve scheduling problem. It replaces uncertain reserve prices and activation ratios by the probability-weighted trajectories of the scenario bundle and therefore supplies the first-stage policy used in the EEV evaluation. The formulation is developed within this MSc project, using standard linear programming and stochastic programming principles as modelling foundations [1, 4].

4.3.1 Sets, Parameters, and Variables

The notation used in the deterministic model is summarized in Table 4.1. Mnemonic superscripts such as up, down, base, and real are written in roman font because they represent labels rather than mathematical indices.

Table 4.1: Notation for the deterministic optimization model.

Symbol	Description
Sets and indices	
T	Set of 15-minute intervals in the daily optimization horizon.
t	Time-interval index, $t \in T$.
Parameters	
$P^{\min}, P^{\max}$	Minimum and maximum generation output of the gas turbine.
RU, RD	Maximum upward and downward ramp between two consecutive intervals.
$R^{\max, \text{up}}, R^{\max, \text{down}}$	Maximum upward and downward reserve capacity.
λ_t^{DA}	Day-ahead electricity price in interval t .
$\lambda_t^{\text{up}}, \lambda_t^{\text{down}}$	Probability-weighted EV trajectories of upward and downward reserve prices.
$\alpha_t^{\text{up}}, \alpha_t^{\text{down}}$	Probability-weighted EV trajectories of upward and downward activation ratios.
$c^{\text{fuel}}, c^{\text{var}}$	Fuel and variable operating cost coefficients.
$c^{\text{pen, up}}, c^{\text{pen, down}}$	Penalty coefficients for undelivered requested activation.
Δ	Length of one time interval in hours.
Decision variables	
p_t^{base}	Planned base generation in interval t .
$r_t^{\text{up}}, r_t^{\text{down}}$	Upward and downward reserve capacity offered in interval t .
$a_t^{\text{up}}, a_t^{\text{down}}$	Realized upward and downward activation in interval t .
$u_t^{\text{up}}, u_t^{\text{down}}$	Non-negative upward and downward activation shortfall.
p_t^{real}	Realized generation after activation in interval t .

4.3.2 Objective Function

The deterministic objective maximizes the net economic value of energy and reserve participation:

$$\max \sum_{t \in T} \left(\lambda_t^{\text{DA}} p_t^{\text{base}} \Delta + \lambda_t^{\text{up}} r_t^{\text{up}} + \lambda_t^{\text{down}} r_t^{\text{down}} - (c^{\text{fuel}} + c^{\text{var}}) p_t^{\text{real}} \Delta - c^{\text{pen, up}} u_t^{\text{up}} - c^{\text{pen, down}} u_t^{\text{down}} \right). \quad (4.7)$$

The first term represents day-ahead energy revenue, the second and third terms represent reserve-related revenue, and the fourth term represents the operating cost of realized generation. The final two terms are penalties for undelivered upward and downward activation.

4.3.3 Constraints

The base generation bounds ensure that the planned operating point remains within the technical generation range of the turbine:

$$P^{\min} \leq p_t^{\text{base}} \leq P^{\max}, \quad \forall t \in T. \quad (4.8)$$

Reserve capacity is limited both by product-specific reserve limits and by the physical headroom around the base generation schedule:

$$0 \leq r_t^{\text{up}} \leq R^{\max, \text{up}}, \quad 0 \leq r_t^{\text{down}} \leq R^{\max, \text{down}}, \quad \forall t \in T, \quad (4.9)$$

$$p_t^{\text{base}} + r_t^{\text{up}} \leq P^{\max}, \quad p_t^{\text{base}} - r_t^{\text{down}} \geq P^{\min}, \quad \forall t \in T. \quad (4.10)$$

These constraints prevent the model from offering reserve that cannot be delivered without violating the minimum or maximum output limits.

Ramping constraints restrict changes in the planned base generation between consecutive intervals:

$$p_t^{\text{base}} - p_{t-1}^{\text{base}} \leq RU, \quad p_{t-1}^{\text{base}} - p_t^{\text{base}} \leq RD, \quad \forall t \in T \setminus \{1\}. \quad (4.11)$$

The first interval is initialized relative to the assumed initial operating point.

Activation is modelled as a fraction of the offered reserve:

$$a_t^{\text{up}} + u_t^{\text{up}} = \alpha_t^{\text{up}} r_t^{\text{up}}, \quad a_t^{\text{down}} + u_t^{\text{down}} = \alpha_t^{\text{down}} r_t^{\text{down}}, \quad \forall t \in T. \quad (4.12)$$

Delivered activation cannot exceed the corresponding reserve offer:

$$0 \leq a_t^{\text{up}} \leq r_t^{\text{up}}, \quad 0 \leq a_t^{\text{down}} \leq r_t^{\text{down}}, \quad \forall t \in T. \quad (4.13)$$

The realized dispatch is then determined by the planned base generation and the activation of upward and downward reserve:

$$p_t^{\text{real}} = p_t^{\text{base}} + a_t^{\text{up}} - a_t^{\text{down}}, \quad \forall t \in T. \quad (4.14)$$

Finally, the realized dispatch must remain within the generation limits,

$$P^{\min} \leq p_t^{\text{real}} \leq P^{\max}, \quad \forall t \in T, \quad (4.15)$$

and satisfy the corresponding ramping restrictions,

$$p_t^{\text{real}} - p_{t-1}^{\text{real}} \leq RU, \quad p_{t-1}^{\text{real}} - p_t^{\text{real}} \leq RD, \quad \forall t \in T \setminus \{1\}. \quad (4.16)$$

This ensures that the deterministic schedule is feasible both before and after the activation approximation is applied.

4.4 Stochastic Optimization Model

The stochastic model extends the deterministic baseline by explicitly representing uncertainty in reserve prices and activation ratios. It follows a two-stage structure. The base generation and reserve offers are first-stage decisions, which must be fixed before the

scenario is known. Realized activation and realized dispatch are second-stage quantities, evaluated separately for each scenario.

4.4.1 Sets, Parameters, and Variables

The notation specific to the stochastic formulation is summarized in Table 4.2. Notation already introduced in Table 4.1 keeps the same interpretation unless a scenario index is added.

Table 4.2: Additional notation for the stochastic optimization model.

Symbol	Description
Sets and indices	
S	Set of scenarios in the daily scenario bundle.
s	Scenario index, $s \in S$.
Parameters	
π_s	Probability of scenario s .
$\lambda_{t,s}^{\text{up}}, \lambda_{t,s}^{\text{down}}$	Scenario-dependent upward and downward reserve prices.
$\alpha_{t,s}^{\text{up}}, \alpha_{t,s}^{\text{down}}$	Scenario-dependent upward and downward activation ratios.
Decision variables	
$a_{t,s}^{\text{up}}, a_{t,s}^{\text{down}}$	Scenario-dependent realized upward and downward activation.
$u_{t,s}^{\text{up}}, u_{t,s}^{\text{down}}$	Scenario-dependent non-negative activation shortfall.
$p_{t,s}^{\text{real}}$	Scenario-dependent realized generation after activation.

4.4.2 Objective Function

The stochastic objective maximizes expected net economic value across the scenario set:

$$\max \sum_{t \in T} \lambda_t^{\text{DA}} p_t^{\text{base}} \Delta + \sum_{s \in S} \pi_s \sum_{t \in T} \left(\lambda_{t,s}^{\text{up}} r_t^{\text{up}} + \lambda_{t,s}^{\text{down}} r_t^{\text{down}} - (c^{\text{fuel}} + c^{\text{var}}) p_{t,s}^{\text{real}} \Delta - c^{\text{pen,up}} u_{t,s}^{\text{up}} - c^{\text{pen,down}} u_{t,s}^{\text{down}} \right). \quad (4.17)$$

The day-ahead revenue term is outside the scenario summation because p_t^{base} and λ_t^{DA} are common to all scenarios. Reserve revenue and realized operating cost are scenario-dependent and are therefore weighted by the scenario probabilities.

4.4.3 Constraints

The stochastic formulation retains the first-stage generation bounds, reserve-capacity bounds, reserve headroom, and base-ramping constraints defined in Equations (4.8), (4.9), (4.10), and (4.11), respectively. These constraints are imposed only on p_t^{base} , r_t^{up} , and r_t^{down} . Consequently, all scenarios share the same first-stage schedule before the uncertain scenario is observed, which provides the non-anticipativity structure of the two-stage formulation.

For each scenario, activation is linked to the first-stage reserve offer by

$$a_{t,s}^{\text{up}} + u_{t,s}^{\text{up}} = \alpha_{t,s}^{\text{up}} r_t^{\text{up}}, \quad a_{t,s}^{\text{down}} + u_{t,s}^{\text{down}} = \alpha_{t,s}^{\text{down}} r_t^{\text{down}}, \quad \forall t \in T, \forall s \in S. \quad (4.18)$$

These constraints express that the activation ratio is uncertain, but the reserve offer itself is fixed before the scenario is observed.

The delivered activation in every scenario is also bounded by the corresponding first-stage reserve offer:

$$0 \leq a_{t,s}^{\text{up}} \leq r_t^{\text{up}}, \quad 0 \leq a_{t,s}^{\text{down}} \leq r_t^{\text{down}}, \quad \forall t \in T, \forall s \in S. \quad (4.19)$$

The realized dispatch in each scenario is given by

$$p_{t,s}^{\text{real}} = p_t^{\text{base}} + a_{t,s}^{\text{up}} - a_{t,s}^{\text{down}}, \quad \forall t \in T, \forall s \in S. \quad (4.20)$$

Scenario feasibility is enforced by the scenario-indexed generation bounds

$$P^{\min} \leq p_{t,s}^{\text{real}} \leq P^{\max}, \quad \forall t \in T, \forall s \in S, \quad (4.21)$$

and ramping constraints

$$p_{t,s}^{\text{real}} - p_{t-1,s}^{\text{real}} \leq RU, \quad p_{t-1,s}^{\text{real}} - p_{t,s}^{\text{real}} \leq RD, \quad \forall t \in T \setminus \{1\}, \forall s \in S. \quad (4.22)$$

Equations (4.18)–(4.22) are the scenario-indexed counterparts of the deterministic activation and realized-dispatch constraints (4.12)–(4.16). Thus, the model optimizes expected value while checking physical feasibility for every represented scenario realization.

4.4.4 Interpretation of the Activation-Ratio Formulation

A central modelling choice in this thesis is the activation-ratio formulation. Instead of assuming that a single gas turbine must satisfy the full system-level balancing demand, the model assumes that only a fraction of the offered reserve is activated in each scenario. This avoids the unrealistic interpretation that one unit absorbs the full activation volume of the balancing market and provides a tractable link between system-level balancing signals and single-unit scheduling.

The formulation should nevertheless be interpreted as an approximation of reserve-use intensity rather than as an exact reconstruction of unit-level dispatch instructions. It may lead to optimistic results if reserve capacity is assumed to be fully accessible whenever it is offered, or if additional activation-related wear, opportunity costs, or market acceptance effects are not fully represented. These limitations are revisited in Section 7.3.

4.5 Comparison of the Deterministic and Stochastic Formulations

The comparison follows the standard expected-value terminology of stochastic programming [1]. Let $x = (p^{\text{base}}, r^{\text{up}}, r^{\text{down}})$ denote the first-stage schedule and let $Q(x, \xi_s)$ denote the optimal second-stage value under scenario s , including the same realized-dispatch constraints and shortfall penalties in every evaluation.

The *expected-value problem* (EV) replaces each uncertain coefficient by the probability-weighted trajectory defined in (4.5) and (4.6). The optimal values of the first-stage decision variables are collected in the vector x^{EV} , and the EV model’s internal objective value is denoted by z^{EV} . This internal value is calculated under the mean inputs and is distinct from the scenario-evaluated performance of the EV policy. The expected result of using the EV solution is

$$\text{EEV} = \sum_{s \in S} \pi_s Q(x^{\text{EV}}, \xi_s), \quad (4.23)$$

which is obtained by fixing every component of x^{EV} in the same scenario model used for the stochastic solution and re-optimizing only second-stage variables.

The recourse problem (RP) chooses the first-stage schedule inside that scenario model:

$$\text{RP} = \max_x \sum_{s \in S} \pi_s Q(x, \xi_s). \quad (4.24)$$

For this maximization problem, the value of the stochastic solution is

$$\text{VSS} = \text{RP} - \text{EEV} \geq 0 \quad (4.25)$$

up to solver tolerance. A relative diagnostic is reported as $\text{VSS}/|\text{RP}|$. The direct difference $\text{RP} - z^{\text{EV}}$ is retained only as a model-internal diagnostic: the two terms optimize different information structures, so that difference is not an out-of-sample performance measure and is not VSS.

EV, EEV, and RP use the same 96 intervals, 50 scenarios, probabilities, common DA price, activation-ratio construction, plant parameters, objective components, technical constraints, and shortfall penalties. Automated checks confirm that all first-stage EV variables remain fixed in EEV. Computationally, EV scales mainly with $|T|$, whereas EEV and RP contain scenario-indexed recourse variables and constraints of order $|T||S|$.

4.6 Model Assumptions

The current optimization framework relies on several simplifying assumptions. First, the gas turbine is treated as a *single independent unit*. The model therefore does not consider portfolio-level coordination, interactions with other units, or system-level dispatch decisions. Second, the base formulation assumes that the unit is available throughout the optimization horizon. The model consequently focuses on reserve scheduling conditional on availability rather than on the broader unit commitment problem.

Startup and shutdown decisions are not included in the present formulation. This means that the model does not represent binary on/off status, startup costs, shutdown costs, or minimum up- and down-time constraints. Excluding these decisions keeps the model linear and computationally tractable, but it also limits the operational realism of the formulation.

Activation is represented through deterministic or scenario-dependent *activation ratios* rather than exact unit-level dispatch instructions. This is a necessary approximation because exact activation instructions for the individual gas turbine are not available in the dataset. Similarly, reserve acceptance and awarded reserve volumes are not modelled explicitly. The model therefore assumes that reserve decisions can be interpreted as economically relevant offered or scheduled reserve quantities.

The market settlement representation is also simplified. The objective captures day-ahead energy value, reserve-related revenues, operating cost, and selected penalty terms, but it does not reproduce the full settlement mechanism of real balancing markets. Finally, risk aversion is not included in the baseline formulation. The stochastic model maximizes expected value across scenarios, while downside-risk measures such as CVaR are left for future extensions.

Taken together, these assumptions keep the model transparent and suitable for repeated multi-day testing. At the same time, they define the interpretation of the results: the model should be viewed as a tractable decision-support framework for studying the energy-reserve trade-off, rather than as a complete industrial bidding or unit commitment tool.

Chapter 5

Implementation

5.1 Model Implementation

This chapter describes how the proposed optimization framework is implemented in practice. The implementation is designed to support a complete workflow from processed forecasting outputs to optimization-based reserve scheduling and result evaluation.

Software Environment

The implementation is developed in Python, which provides a flexible and well-established environment for data handling, time-series preprocessing, optimization model construction, solver interaction, result aggregation, and visualization. The main scientific and modeling libraries used in the project include `pandas` for tabular data handling and preprocessing, `numpy` for numerical operations, `matplotlib` for visualization, and `Pyomo` for optimization model formulation. In addition, standard Python utilities are used for file management and experiment execution.

The overall implementation is modular, meaning that forecasting, optimization, and evaluation are kept as separate components as far as possible. This improves transparency and makes the framework easier to debug, extend, and interpret.

Pyomo and Solver Setup

Both the deterministic and stochastic optimization models are formulated in `Pyomo`. `Pyomo` is used because it provides a clear and transparent way to translate mathematical formulations into executable optimization code, while also maintaining close consistency between the implementation and the equations presented in the methodology chapter.

The solver is selected dynamically from a list of available candidates. In the current implementation, the candidate solvers include Gurobi, CBC, GLPK, and HiGHS. This approach allows the code to be executed on different systems depending on local solver availability. Since the current optimization formulations are linear and do not yet include binary commitment variables or nonlinear constraints, standard linear programming solvers are sufficient for the implementation.

After each model run, the implementation reports the solver status and termination condition. This serves as a basic verification step to ensure that the problem is feasible, that an optimal solution has been found, and that the resulting objective value can be interpreted consistently.

Data Input and Output Structure

The optimization scripts are designed to consume processed input files rather than raw datasets. This allows the forecasting stage and the optimization stage to remain loosely coupled, which improves modularity and reduces unnecessary complexity in the implementation.

The main optimization inputs consist of a daily forecast table, a scenario bundle table, and a set of gas turbine technical and economic parameters. The forecast table contains point estimates such as reserve prices, day-ahead prices, and optional deterministic activation-ratio signals. The scenario bundle table contains scenario-specific reserve-price realizations and activation-ratio inputs. The plant parameter set defines the technical and cost characteristics of the gas turbine used in the scheduling problem.

In the current implementation, these inputs are typically stored in **parquet** or **csv** format. During execution, the optimization script loads the relevant files, filters them by delivery day where necessary, and maps them into model-ready data structures indexed by time interval and scenario.

The model outputs are stored in two main forms. First-stage output tables contain variables such as **p_base**, **r_up**, and **r_down**, while second-stage output tables contain scenario-dependent quantities such as realized activation and realized dispatch. In addition, the implementation generates summary outputs for later evaluation, including daily objective values, economic decomposition terms, average reserve allocations, and comparison-ready result tables.

Scripts and Workflow

The implementation follows a structured workflow that moves from forecasting outputs to evaluation-ready optimization results. In broad terms, the process begins with preparation and cleaning of the unified dataset, followed by the generation of forecasting outputs and scenario bundles. These processed inputs are then loaded for either deterministic or stochastic optimization over a selected delivery day or range of days. After solving the optimization problem, the implementation extracts the first-stage and second-stage outputs, aggregates the results across multiple days, and produces plots and tables for model comparison and case-study analysis.

In practice, this workflow is supported by several dedicated scripts and notebooks. Some are used for repeated optimization over multiple days, others for deterministic versus stochastic comparison, benchmark evaluation, and result visualization. This overall design makes it possible to move beyond isolated single-day experiments and instead evaluate the framework systematically across a larger backtesting horizon and multiple strategy settings.

5.2 Benchmark Strategies

To assess whether the optimization model provides value beyond simple heuristic rules, a set of benchmark strategies is implemented. These strategies are intentionally much simpler than the proposed optimization framework and serve as reference points for comparison.

The **fixed_reserve_20_20** strategy represents a fixed-rule reserve policy. In each interval, it aims to provide 20 MW of upward reserve and 20 MW of downward reserve.

The base generation level is set at a feasible intermediate operating point and is then adjusted as necessary to respect ramping and technical constraints. This strategy does not respond to market prices or scenario information and is intended to represent a simple mechanical reserve rule. If the target reserve level is not technically feasible, the implementation automatically reduces the reserve level to restore feasibility.

The `max_reserve` strategy prioritizes reserve provision. In each interval, it attempts to provide the largest feasible symmetric reserve level subject to upward reserve capability limits, downward reserve capability limits, generation feasibility, and an additional implementation-level safety cap. This strategy represents an aggressive reserve-oriented operating philosophy and serves as a benchmark for what happens when reserve participation is prioritized over a coordinated energy-reserve trade-off.

The `energy_focused` strategy is intended to represent an energy-market-oriented operating rule. In this benchmark, no reserve is intentionally provided and the turbine is operated close to a high output level, typically near a specified fraction of maximum output. Conceptually, this benchmark represents a strategy that focuses primarily on capturing day-ahead electricity revenue rather than balancing-market flexibility.

The `no_optimization_rule` strategy represents a simple fixed dispatch heuristic. It does not optimize against market prices or uncertainty, but instead sets zero reserve and fixes the base generation schedule at a medium-to-high operating level based on a predetermined fraction of the feasible output range. Compared with the energy-focused strategy, this rule is less aggressive and more closely resembles a practical fixed-load operating heuristic.

Taken together, these benchmark strategies represent different simplified operating philosophies, including fixed reserve provision, maximum reserve participation, energy-priority operation, and fixed non-optimized dispatch. Comparing them under the common scenario evaluation shows how their modelled objective values and operating decisions differ from RP.

5.3 Evaluation Design

The evaluation framework is designed to test the optimization model systematically. The purpose is not only to show that the model can be solved successfully, but also to assess the consistency, feasibility, and interpretability of its modelled outcomes across the tested market conditions.

Multi-Day Backtesting

A central component of the evaluation is multi-day backtesting. Instead of analyzing only a single representative day, the optimization framework is run repeatedly over a larger set of delivery days. This makes it possible to observe how the strategy behaves under varying market conditions and whether its performance is stable over time.

For each day, the evaluation stores summary quantities such as the objective value, the economic decomposition, the average base generation level, the average upward and downward reserve allocations, and selected activation-related indicators. This allows the backtesting exercise to answer practical questions about stability, consistency, and sensitivity to changing market structure.

Deterministic vs Stochastic Comparison

The EV–stochastic comparison is based on EV, EEV, RP, and VSS as defined in Section 4.5. For each of the 42 delivery days, the EV trajectory is reconstructed from that day’s scenario bundle; x^{EV} is then solved, fixed in the stochastic model to obtain EEV, and compared with the optimized RP value. Consequently, RP and EEV are evaluated under the same 50-scenario distribution rather than by comparing their separate in-sample objective values.

The stored diagnostics include solver status, termination condition, maximum constraint violation, the maximum deviation of fixed first-stage variables, scenario coverage and probability checks, revenue and cost decomposition, average first-stage decisions, VSS, and relative VSS. A negative VSS beyond 10^{-5} is treated as an implementation failure, not as evidence that stochastic optimization performs worse. The direct EV objective remains available only as a diagnostic.

Benchmark Comparison

The benchmark comparison evaluates the proposed optimization framework against the heuristic strategies described above. This part of the analysis addresses whether the RP schedule has a higher mean daily objective value than simple operating rules under the same scenario evaluation.

The comparison is conducted over the same set of test days, and the resulting strategy performance is summarized through mean daily objective values in EUR/day, daily distributions, reserve-allocation statistics, and simple feasibility indicators. These are model-evaluation quantities under the stated scenario and penalty assumptions, rather than realized industrial profits.

Sensitivity Analysis

Sensitivity analysis is used to examine how the stochastic model responds to selected operational and market-input parameters. Rather than treating one parameterization as fixed and definitive, the thesis explores how the model objective and decisions change when these quantities are varied.

The tested dimensions include day-ahead electricity prices, fuel cost, reserve capability limits, and activation-ratio scaling assumptions. These experiments describe general parameter response; they do not constitute a sensitivity analysis of VSS and do not validate the monetary magnitude of VSS with respect to the shortfall-penalty coefficient.

Case Studies

In addition to aggregate statistical evaluation, the framework includes selected case studies. The purpose of the case-study component is to complement summary statistics with a more detailed operational interpretation of individual days.

A case-study day is selected because it is representative of RP outcomes, exhibits high price/activation stress, or has the largest absolute VSS. The comparison case is thus chosen by a common-distribution performance measure rather than by the incomparable difference between the EV and RP internal objectives. For each case, the EV policy is reported through EEV, alongside RP and a rule-based policy evaluated in the same scenario model.

Case studies are valuable because they illustrate not only whether the model performs well on average, but also how its internal decision logic responds to concrete market conditions.

Overall Role of the Evaluation Design

Taken together, the evaluation design combines multi-day testing, an EV–EEV–RP comparison, common-scenario benchmark evaluation, selected parameter-response checks, and interpretable case studies. This combination provides a structured assessment of the current modelling framework while retaining the stated calibration limitations.

5.4 Computational Setup and Instance Size

The optimization models were implemented in Python using Pyomo. The computational experiments were carried out on a daily rolling basis, where each optimization instance corresponds to one delivery day. Each day is represented by 96 time intervals of 15 minutes. The stochastic model uses 50 scenarios per delivery day, while the deterministic EV model uses one probability-weighted trajectory derived from those scenarios.

The evaluation horizon used in the optimization experiments covers 42 delivery days, from 28 January 2026 to 10 March 2026. This gives 4032 optimized time intervals in the multi-day backtest. The stochastic scenario bundle contains 201,600 scenario-interval observations, corresponding to

$$42 \times 96 \times 50 = 201,600.$$

Both optimization formulations are linear programming models in the current version. The deterministic model contains only continuous variables indexed by time. The stochastic model has the same first-stage variables as the deterministic model, but its activation and realized dispatch variables are indexed by both time and scenario. As a result, the stochastic model is substantially larger even though it remains linear.

Table 5.1 reports the size of a representative daily optimization instance. The counts are obtained directly from the Pyomo model objects after model construction.

Table 5.1: Size of representative daily optimization instances.

Model	Time intervals	Scenarios	Variables	Constraints
Deterministic LP	96	1	576	1,632
Stochastic LP	96	50	14,688	43,968

The difference in model size is mainly caused by the second-stage variables and constraints. In the deterministic formulation, activation and realized dispatch are indexed only by time. In the stochastic formulation, the corresponding variables are indexed by both time and scenario. For example, the stochastic model includes 96×50 realized dispatch variables, whereas the deterministic model includes only 96. The same scenario expansion applies to activated upward reserve, activated downward reserve, and the associated feasibility constraints.

Table 5.2 gives a more detailed breakdown of the variables in the two formulations.

The backtesting scripts recorded wall-clock solution time for each daily instance using the same solver, hardware, and input days. The reported times cover the optimization

Table 5.2: Variable breakdown for one daily optimization instance.

Variable	Deterministic model	Stochastic model
Base generation p_t^{base}	96	96
Upward reserve r_t^{up}	96	96
Downward reserve r_t^{down}	96	96
Activated upward reserve $a_{t,s}^{\text{up}}$	96	4,800
Activated downward reserve $a_{t,s}^{\text{down}}$	96	4,800
Realized dispatch $p_{t,s}^{\text{real}}$	96	4,800
Total	576	14,688

solve only and exclude data loading, model construction, scenario generation, EEV evaluation, result export, and visualization. Table 5.3 summarizes the results over the 42-day backtesting horizon.

Table 5.3: Wall-clock solution time over the 42-day backtesting horizon.

Model	Solved days	Mean time (s)	Median time (s)	Max time (s)
Deterministic LP	42	0.17	0.16	0.44
Stochastic LP	42	5.15	5.16	5.98

The stochastic LP required approximately 30.3 times as much solution time on average as the deterministic LP. This difference is consistent with the substantially larger scenario-indexed formulation. Nevertheless, all stochastic daily instances were solved in less than six seconds, indicating that both formulations remain computationally tractable for the tested daily scheduling setting.

Although the stochastic model is much larger than the deterministic model, both models remain linear in the current formulation. The additional computational burden therefore comes primarily from scenario expansion rather than from integer variables or nonlinear constraints. This is important for interpreting the deterministic–stochastic comparison: the stochastic model should be evaluated not only by its expected objective value, but also in relation to its additional model size and solution time.

In the present experiments, the stochastic model has approximately 25.5 times as many variables and 26.9 times as many constraints as the deterministic model. This increase is expected because most second-stage variables and constraints scale linearly with the number of scenarios. If the number of scenarios is denoted by $|S|$ and the number of time intervals by $|T|$, the deterministic model scales mainly with $|T|$, while the second-stage part of the stochastic model scales mainly with $|T||S|$.

This also has implications for future extensions. The current models do not include startup and shutdown decisions. If unit commitment decisions are added, binary variables would be required to represent on/off status, startups, shutdowns, and minimum up- and down-time constraints. The problem would then become a mixed-integer optimization model rather than a linear programming model. In that case, computational time could become a more important practical limitation, especially for the stochastic formulation.

Chapter 6

Results

6.1 Multi-Day Backtesting Results

To evaluate the consistency of the proposed optimization framework beyond a single representative day, the stochastic model was tested over a sequence of delivery days covering the late January to early March period. The objective of this multi-day backtest is to examine whether the model produces economically meaningful and technically interpretable reserve scheduling decisions under changing market conditions.

Daily Objective Evolution

Figure 6.1 shows the evolution of the daily objective value over the backtesting horizon. Several observations can be made. First, the model remains profitable throughout the full evaluation period, as all daily objective values are positive. This indicates that, under the current market assumptions and parameterization, the joint energy-and-reserve strategy remains economically viable across a wide range of days.

Second, the daily objective values vary substantially over time. The lowest daily values are close to the lower end of the range, around 10^5 , while the highest values approach 8×10^5 . This wide spread suggests that the profitability of the strategy is strongly influenced by day-to-day variation in market conditions, especially the interaction between day-ahead electricity value and reserve-related opportunities.

Third, the temporal pattern is not random. There are several local peaks, for example around mid-February, late February, and early March, as well as a few pronounced dips around the beginning of February and again before the late-February peak. This pattern suggests that the model is responsive to market structure rather than producing nearly constant dispatch and reserve decisions across days.

Overall, the daily objective evolution indicates that the model is both economically active and sufficiently flexible to adapt to changing market signals.

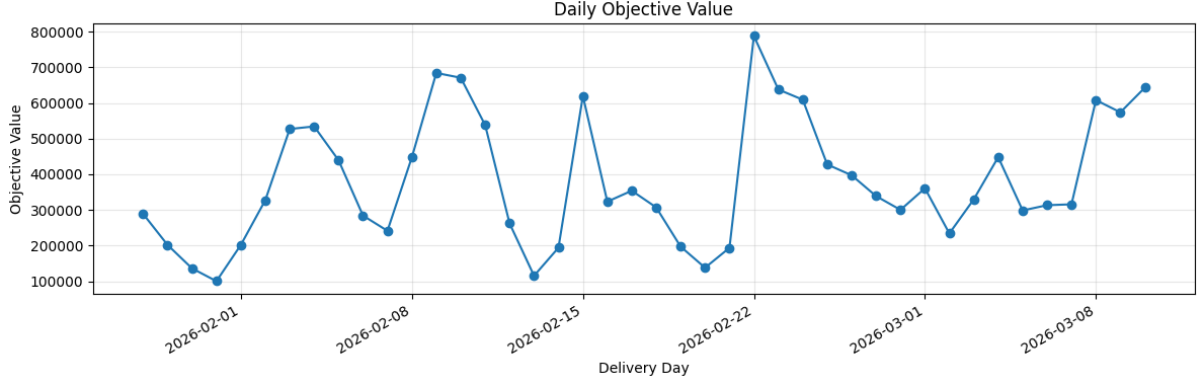


Figure 6.1: Daily objective value over the multi-day backtesting horizon.

Revenue Decomposition

Figure 6.2 presents the daily revenue decomposition into day-ahead revenue, upward reserve revenue, downward reserve revenue, operating cost, and penalty cost. The dominant positive contribution comes from the day-ahead energy component. On most days, day-ahead revenue forms the largest share of the objective value, which shows that the turbine is not acting as a pure reserve-only unit. Instead, the model chooses a joint energy-and-reserve participation strategy in which base generation remains economically important.

Reserve revenues also contribute to the objective, although their role is more complementary than dominant. Both upward and downward reserve revenues vary across the backtesting horizon, reflecting changes in reserve-market conditions and in the model's preferred allocation of flexibility. The relative importance of upward and downward reserve is therefore not constant over time, which indicates that the model adapts the reserve mix to daily market conditions.

Operating cost is a non-negative cost quantity that is subtracted from market revenue and therefore appears as a negative contribution in the objective-value decomposition. Its magnitude remains below the combined positive revenue contributions on all tested days. Shortfall-penalty cost follows the same sign convention: the underlying cost is non-negative, but its contribution is plotted below zero because it is deducted from the objective value. The penalty contribution is negligible in these reported RP results, indicating that the optimized schedules do not rely materially on activation shortfall under the current scenario design.

The bars therefore decompose the model objective into positive revenue components and subtractive cost components. Day-ahead and reserve-market revenues are shown above zero; operating-cost and activation-shortfall-penalty contributions are shown below zero. Adding all signed bars gives the daily net objective value, equivalently total market revenue minus operating cost and shortfall-penalty cost. Taken together, the decomposition shows that DA revenue is the main positive component, reserve products provide additional revenue, and operating cost remains controlled.

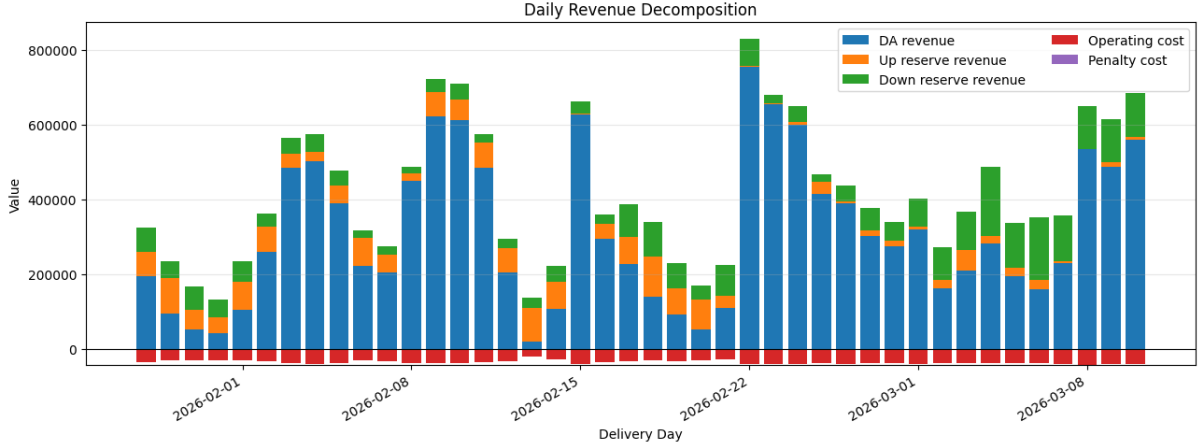


Figure 6.2: Daily objective-value decomposition during the multi-day backtesting period. Revenue components are plotted above zero; operating-cost and shortfall-penalty contributions are plotted below zero because they are subtracted from revenue.

Average Reserve Allocation and Dispatch

Figure 6.3 shows the evolution of average base generation, average upward reserve allocation, and average downward reserve allocation over the backtest period.

The average base generation remains relatively high across most days, typically in the range of roughly 200–300 MW, with several periods close to the upper end of the feasible operating range. This indicates that the model often prefers to keep the turbine at a relatively elevated operating point. Economically, this is consistent with the strong role of DA revenue observed in the decomposition results.

The average upward reserve allocation varies more strongly over time. On some days, it remains at a moderate level, while on others it drops close to zero. This behavior is economically intuitive. When the base output is pushed upward to capture more DA value, the upward headroom available for reserve provision decreases, which naturally reduces the feasible and attractive upward reserve offer.

The average downward reserve allocation is more stable and tends to remain positive throughout the backtest horizon, often increasing when the base generation is relatively high. This is also consistent with the physical structure of the problem: when the unit operates at a higher output level, it typically has more room to reduce output than to increase it further. Therefore, the model often combines high base generation with meaningful downward reserve capability.

These results indicate that the model captures a realistic energy-reserve trade-off. When base generation is high, the available upward headroom is reduced, which limits upward reserve provision. At the same time, higher base generation generally increases the room to reduce output, making downward reserve more attractive and technically feasible. The model therefore links market incentives to the physical flexibility of the gas turbine in an economically interpretable way.

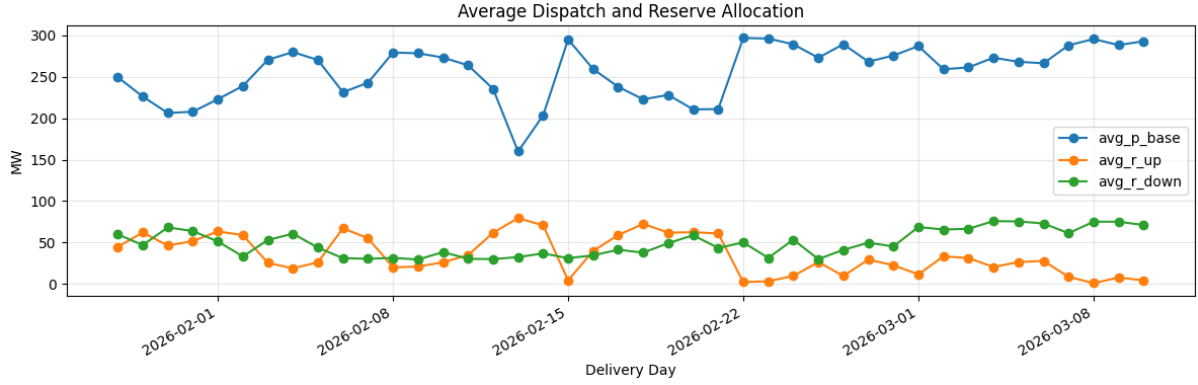


Figure 6.3: Average base generation and reserve allocation over the backtesting horizon.

Distribution of Daily Results

Figure 6.4 summarizes the distribution of daily objective values through a boxplot. The distribution confirms that the model produces consistently positive outcomes while still exhibiting substantial day-to-day variation.

The median daily objective lies in the mid-range of the observed values, while the interquartile range is relatively wide. This indicates that the strategy is not producing almost identical profits every day, but rather adapts to the economic conditions of individual delivery days. At the same time, the absence of negative outliers suggests that the model does not generate systematically poor performance under the tested conditions.

The upper whisker extends much further than the lower whisker, which indicates that some days offer exceptionally favorable market opportunities and that the model is able to exploit them. This asymmetric spread is economically meaningful: in balancing-related applications, a few high-value days may contribute disproportionately to overall profitability.

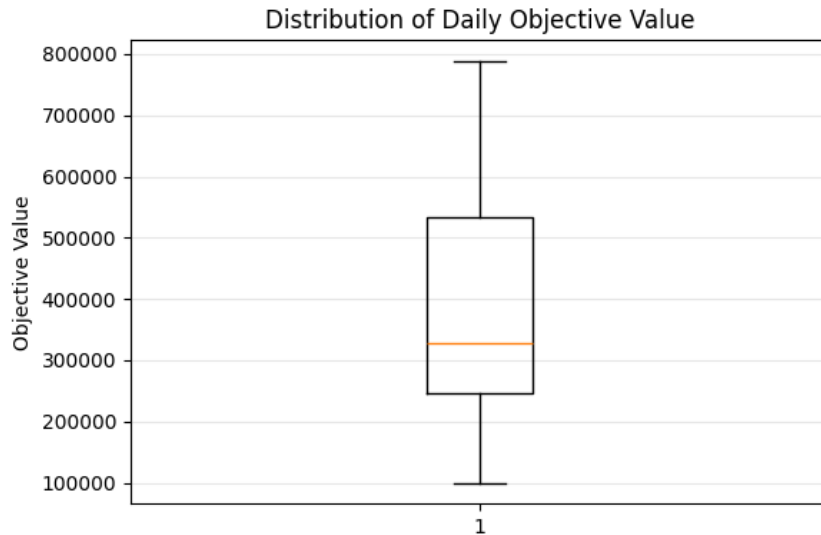


Figure 6.4: Distribution of daily objective values in the multi-day backtest.

Overall Interpretation

Overall, the multi-day backtesting results show that the model is not limited to isolated single-day success. It produces consistently positive objective values, economically interpretable revenue structures, technically plausible reserve allocations, and stable profitable operation over the tested delivery days. The results also suggest that profitability is driven primarily by day-ahead energy value, while reserve participation provides an important supplementary source of revenue.

6.2 Deterministic vs Stochastic Comparison

The comparison covers all 42 delivery days from 28 January to 10 March 2026. For each day, the EV policy is evaluated in exactly the same 50-scenario recourse model as the RP solution. Thus, the relevant comparison is RP versus EEV, while the EV objective is shown only as an internal diagnostic.

RP, EEV, and VSS

Figure 6.5 plots both comparable time series. Mean daily RP is EUR 380,267.95/day, whereas mean EEV is EUR 32,370.34/day; the related medians are EUR 329,164.58/day and EUR 422.24/day. EEV is negative on 21 of 42 days because the EV first-stage policy incurs large shortfall penalties when exposed to individual activation realizations. RP exceeds EEV on every day (Figure 6.6). Mean VSS is EUR 347,897.61/day and median VSS is EUR 363,192.73/day, with a range from EUR 29,776.84/day to EUR 527,793.11/day. There are no negative VSS observations, including within the 10^{-5} numerical tolerance.

The positive VSS demonstrates the value of scenario-aware scheduling within the specified evaluation model. However, its monetary magnitude is driven mainly by avoided activation-shortfall penalties and should not be interpreted directly as an industrial business-case value.

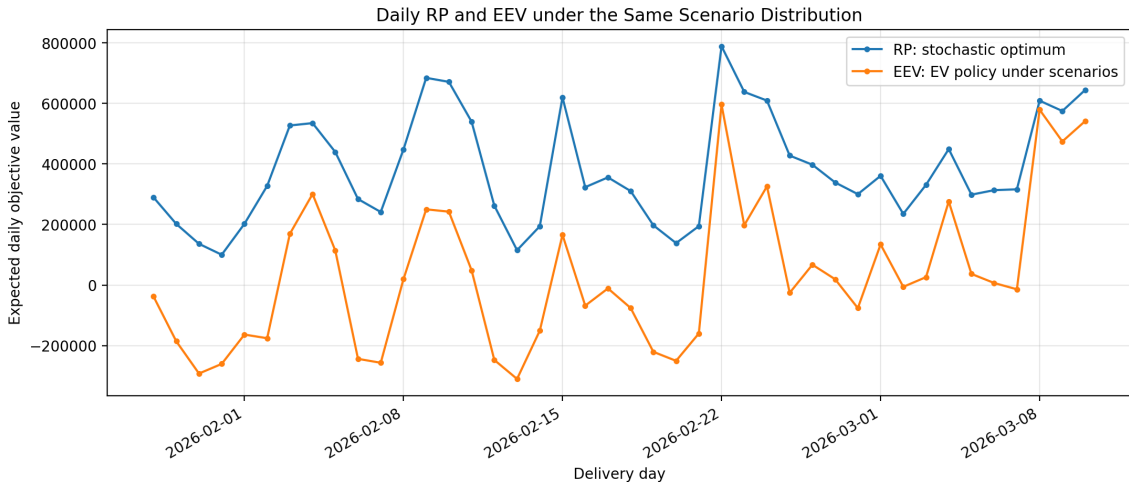


Figure 6.5: Daily RP and EEV objective values under the same scenario distribution (EUR/day).

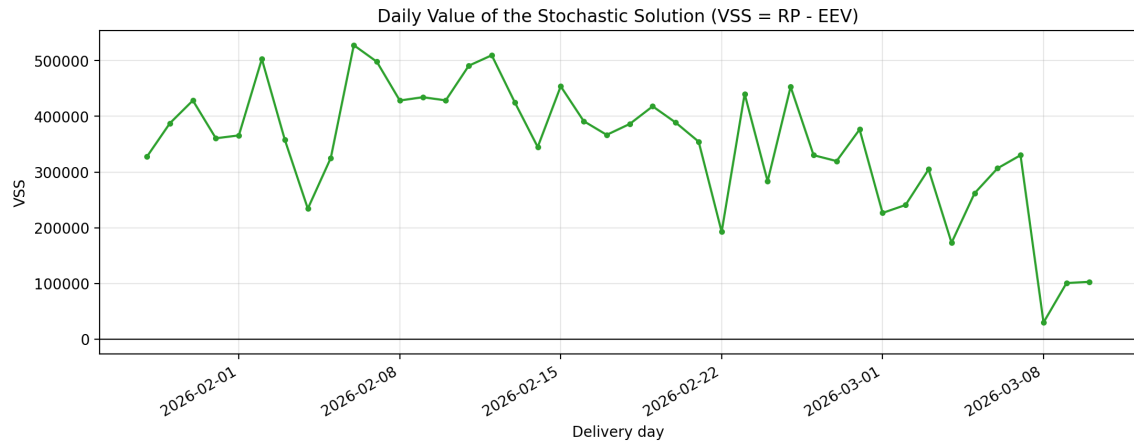


Figure 6.6: Daily value of the stochastic solution, $VSS = RP - EEV$, in EUR/day.

Relative VSS, defined as $VSS/|RP|$, has a mean of 124.22% and a median of 100.18%; it ranges from 4.89% to 369.21% (Figure 6.7). Values above 100% are possible because EEV is negative or small on many days. They should therefore be interpreted as evidence of large penalty exposure relative to the RP value, not as a conventional percentage return.

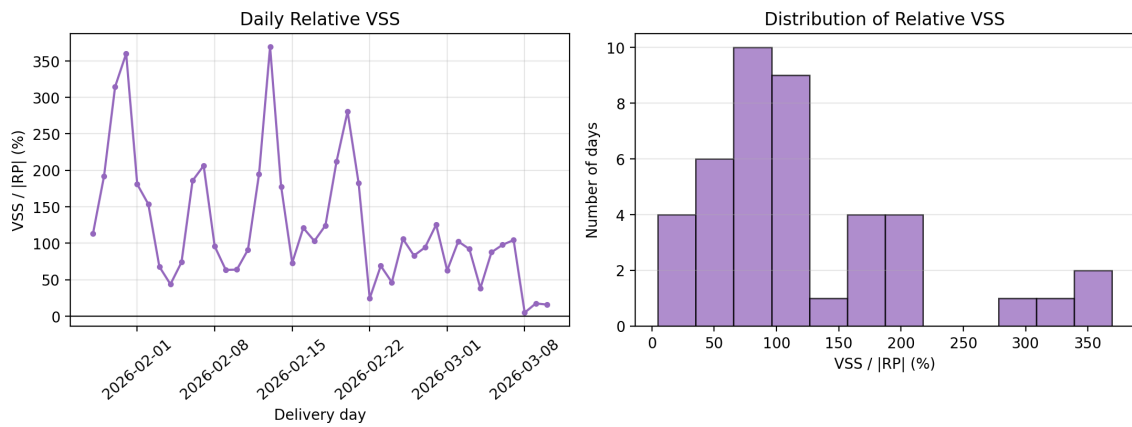


Figure 6.7: Time series and distribution of relative VSS, defined as $VSS/|RP|$.

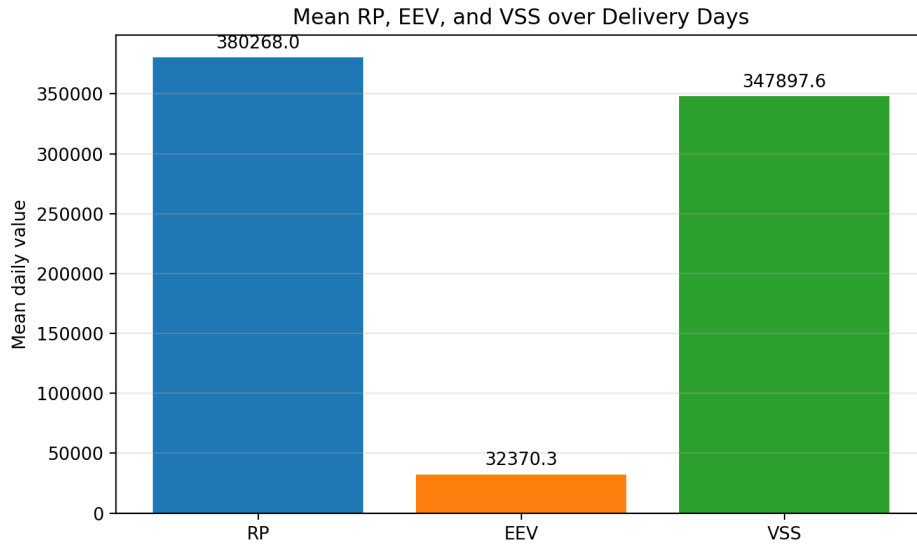


Figure 6.8: Mean daily RP, EEV, and VSS over the 42 delivery days (EUR/day).

With EEV on the horizontal axis and RP on the vertical axis, the scatter in Figure 6.9 places every observation above the 45-degree line. Thus RP exceeds EEV and VSS is positive for every plotted day. The boxplot in Figure 6.10 also demonstrates why the EV objective cannot be used as a performance proxy: its mean is EUR 419,740.81/day, but the same policy produces mean EEV of only EUR 32,370.34/day under the scenario distribution.

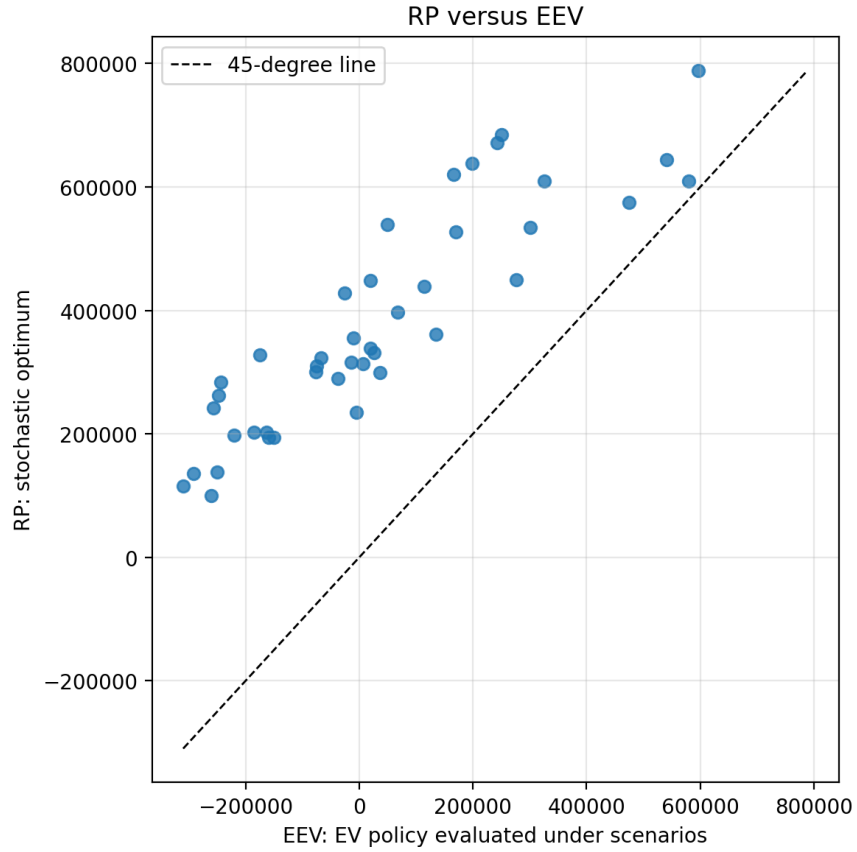


Figure 6.9: Daily RP against EEV (EUR/day). The dashed 45-degree line denotes equality; points above it have $RP > EEV$ and positive VSS.

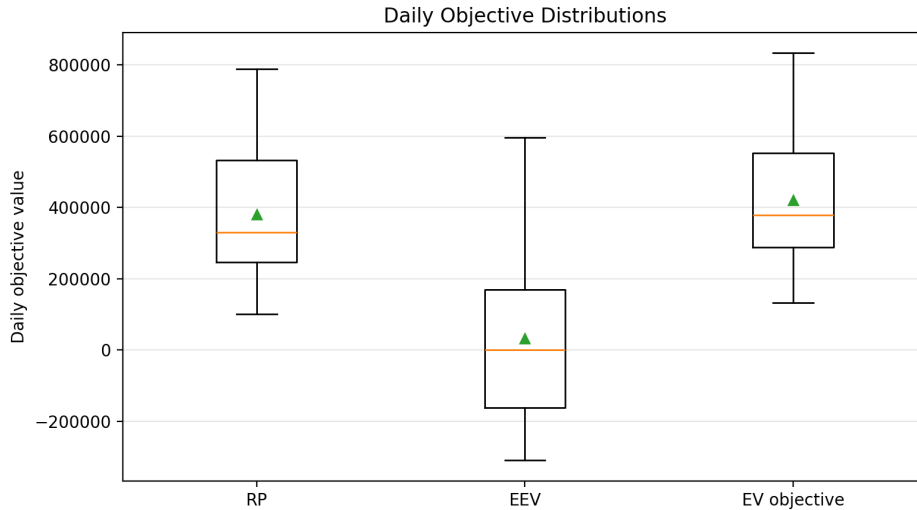


Figure 6.10: Daily distributions of RP, EEV, and the internal EV objective (EUR/day). Only RP and EEV are comparable performance measures; the EV objective is diagnostic. The triangle denotes the mean, the line inside each box the median, the box the interquartile range, and the whiskers the most extreme observations within 1.5 times that range.

Decision and Value Decomposition

The EV solution averages 253.52 MW of base output, 43.11 MW of upward reserve, and 79.93 MW of downward reserve. RP averages 256.45 MW, 35.25 MW, and 50.11 MW, respectively. Relative to EV, RP therefore increases mean base output by 2.93 MW and reduces upward and downward reserve by 7.86 MW and 29.82 MW. The largest systematic change is the reduction in downward reserve exposure; this statement follows from the observed decisions and does not imply a general risk-aversion property of stochastic solutions.

The objective decomposition identifies the source of VSS. EEV incurs a mean daily shortfall penalty of EUR 387,032.74/day, compared with EUR 1,144.58/day for RP. Moving from the EV policy to RP therefore avoids EUR 385,888.16/day of expected penalty cost. This benefit is partly offset by reductions of EUR 8,807.58/day and EUR 30,799.96/day in upward and downward reserve revenue, an increase of EUR 881.73/day in operating cost, and is supplemented by EUR 2,498.72/day of additional DA revenue. Hence the estimated VSS is primarily a shortfall-penalty effect, not an increase in reserve revenue.

The largest absolute VSS occurs on 6 February 2026 (EUR 527,793.11/day; relative VSS 185.81%). The largest relative VSS occurs on 13 February 2026 (369.21%). Across days, VSS has only weak correlation with the combined activation-and-price stress score 0.09 and with upward and downward reserve-price volatility (0.17 and -0.07). It is negatively correlated with mean upward activation ratio (-0.77) and positively correlated with mean downward activation ratio (0.79), consistent with the large downward-reserve exposure of the EV policy. The highest-VSS dates therefore cannot be characterized simply as the most volatile or most highly activated days.

Interpretation and Numerical Validation

The results support the stochastic formulation within the specified scenario-and-penalty model: RP is the optimum over the same feasible first-stage set in which the EV policy is evaluated, and its advantage is positive on all 42 days. They do not establish robustness outside that distribution. In particular, the magnitude of VSS is sensitive to the activation-ratio normalization, the sampled scenario realizations, and the shortfall penalty of 1,000. Because that penalty dominates the value decomposition, penalty calibration and out-of-sample scenario validation are essential before interpreting the monetary levels as industrial business-case estimates.

All EV, EEV, and RP solves reached optimal termination. The maximum recorded constraint violation is 4.89×10^{-9} , the maximum error in fixed EV first-stage variables is zero, every scenario covers all 96 intervals, and scenario probabilities sum to one. The expected-value price and activation trajectories equal the probability-weighted scenario means to machine precision.

6.3 Benchmark Strategy Comparison

All strategies in this section are evaluated in the same 50-scenario model, with the same common DA trajectory, activation-ratio construction, technical constraints, and shortfall penalties. The EV-based optimized policy is represented by EEV rather than by its internal EV objective. This change materially affects the interpretation of the ranking.

Figure 6.11 reports mean daily objective values in EUR/day. RP ranks first at EUR 380,267.95/day. Among the four rule-based policies, the energy-focused strategy ranks second overall at EUR 262,097.71/day, followed by max reserve at EUR 231,220.86/day, fixed reserve 20/20 at EUR 228,758.72/day, the non-optimized rule at EUR 224,443.95/day. The EV policy has mean EEV of EUR 32,370.34/day and ranks last because its relatively large reserve positions generate substantial activation-shortfall penalties under the scenario realizations.

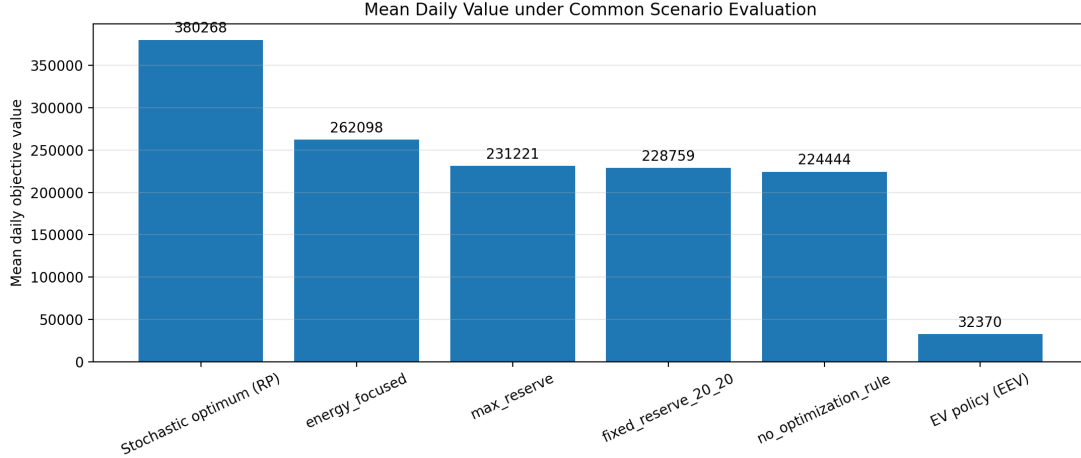


Figure 6.11: Mean daily objective value by strategy under the common scenario evaluation (EUR/day). The EV policy is represented by EEV, not by its internal EV objective.

The distribution in Figure 6.12 reinforces the distinction between an optimized RP schedule and the evaluated EV policy. RP achieves a median daily objective value of EUR 329,164.58/day, with positive objective values on all tested days. The energy-focused and non-optimized rules have negative minima, whereas the EV policy reaches EUR $-310,151.06$ /day. These values are scenario-evaluation outcomes, not direct comparisons of separate optimization objectives.

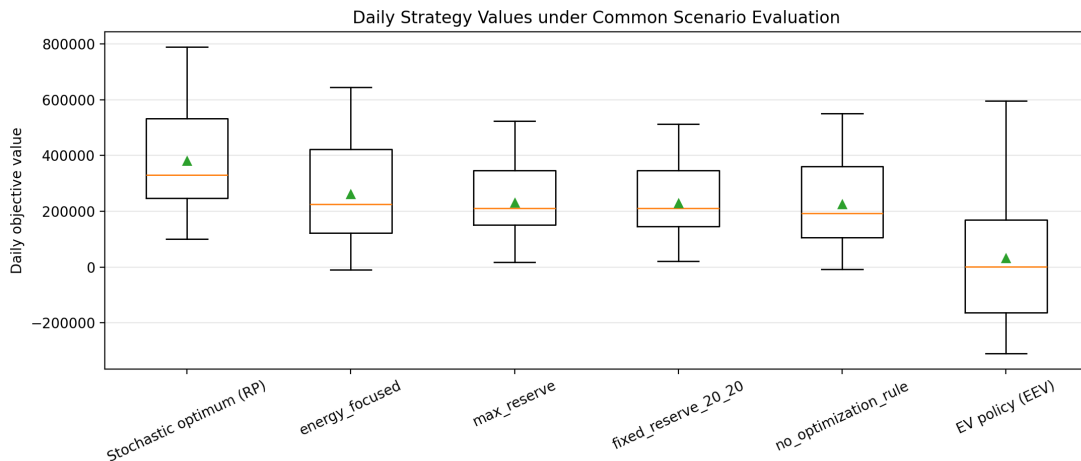


Figure 6.12: Distribution of daily strategy objective values under the common scenario evaluation (EUR/day). The EV policy is shown as EEV.

The benchmark result supports using optimization to coordinate energy and reserve

in the tested model, because RP outperforms each fixed rule on average. It does not support the earlier claim that both optimization-based formulations outperform every heuristic: once the EV policy is evaluated as EEV, its exposure to scenario-dependent shortfall makes it the weakest average performer. Nor is the ranking an out-of-sample guarantee, because RP is optimized on the same scenario bundle used for evaluation. A stricter future test should optimize each policy on a training scenario set and evaluate it on withheld market realizations or an independently generated scenario set.

6.4 Sensitivity Analysis

A one-factor-at-a-time (OFAT) experiment is conducted on the stochastic model over 42 delivery days. In each run, only one input is perturbed at levels $0.8\times$, $1.0\times$, and $1.2\times$ (or $\pm 20\%$ for prices and fuel cost), while all other parameters remain at baseline. The factors are: DA price, fuel cost c_{fuel} , symmetric reserve capability limits ($R_{\text{max}}^{\uparrow}$, $R_{\text{max}}^{\downarrow}$), and activation-ratio scaling. Reserve market prices enter through the scenario bundle and are not varied separately in this sweep. These tests describe the response of the stochastic model objective and decisions; they are not an RP–EEV sensitivity experiment.

Figure 6.13 shows that the mean daily model objective is highly sensitive to DA price, moderately sensitive to reserve limits, and only weakly sensitive to fuel cost and activation scaling. Relative to baseline, a $\pm 20\%$ DA price change shifts the mean objective by roughly $\pm 17\%$, whereas the same perturbation in fuel cost changes it by only about $\pm 2\%$. The tornado ranking in Figure 6.16 identifies DA price as the dominant tested driver (span ≈ 125 kEUR/day), followed by reserve capacity (≈ 25 kEUR/day), fuel cost (≈ 12 kEUR/day), and activation scale (≈ 3 kEUR/day).

The decision-response plots (Figure 6.14) show the expected energy–reserve trade-off: higher DA prices increase average p_{base} and reduce upward reserve bids, while higher reserve limits raise $r_{\uparrow}$ and $r_{\downarrow}$ at the expense of base dispatch. The decomposition in Figure 6.15 shows that DA revenue drives most of the objective change under price shocks, whereas reserve revenue responds mainly to capacity limits. Activation scaling has a small effect on the objective in this particular sweep. Daily boxplots (Figure 6.17) indicate that these directional responses are not driven by a single day.

Overall, the tested model responds consistently to these selected parameter changes. The current sensitivity analysis focuses primarily on operational and market-input parameters. Since avoided shortfall penalties account for most of the measured VSS, the robustness of the VSS magnitude with respect to the penalty coefficient has not yet been fully established and remains an important topic for further validation. The existing figures therefore support only the reported parameter-response statements and do not validate either the sign or monetary magnitude of VSS under alternative penalty settings.

Multi-day mean (successful runs): Objective sensitivity

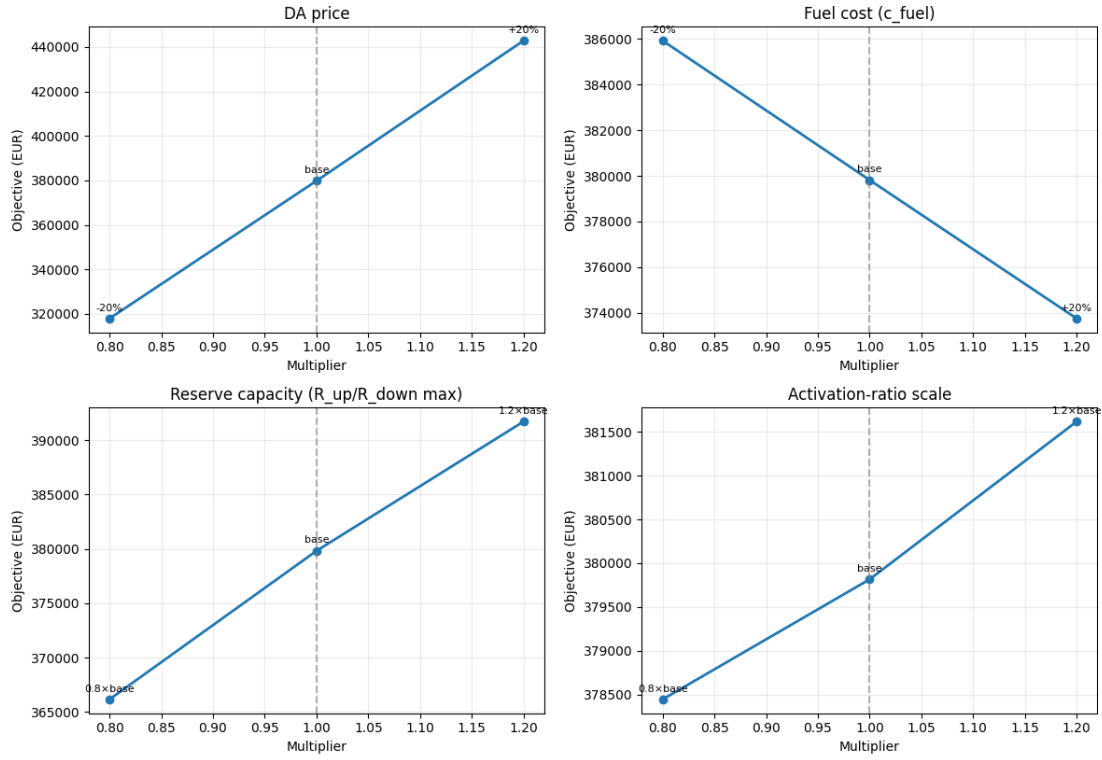


Figure 6.13: Mean daily stochastic-model objective under OFAT perturbations of DA price, fuel cost, reserve limits, and activation-ratio scale (EUR/day). This is a general parameter-response plot, not a VSS sensitivity test.

Multi-day mean (successful runs): First-stage decisions

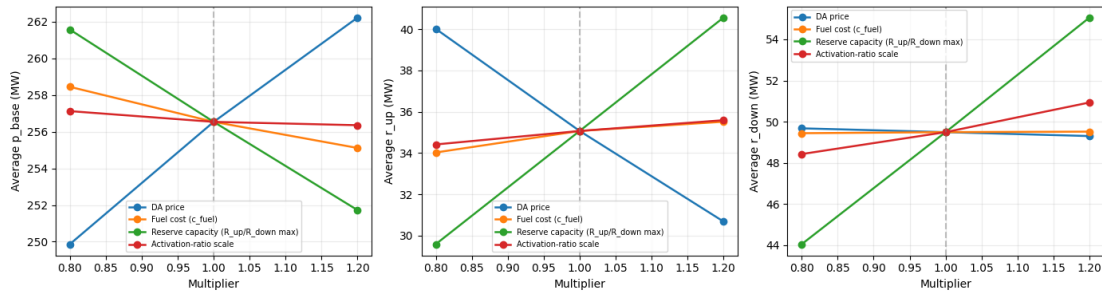


Figure 6.14: Multi-day mean first-stage decisions (p_{base} , $r_{\uparrow}$, $r_{\downarrow}$) across parameter levels.

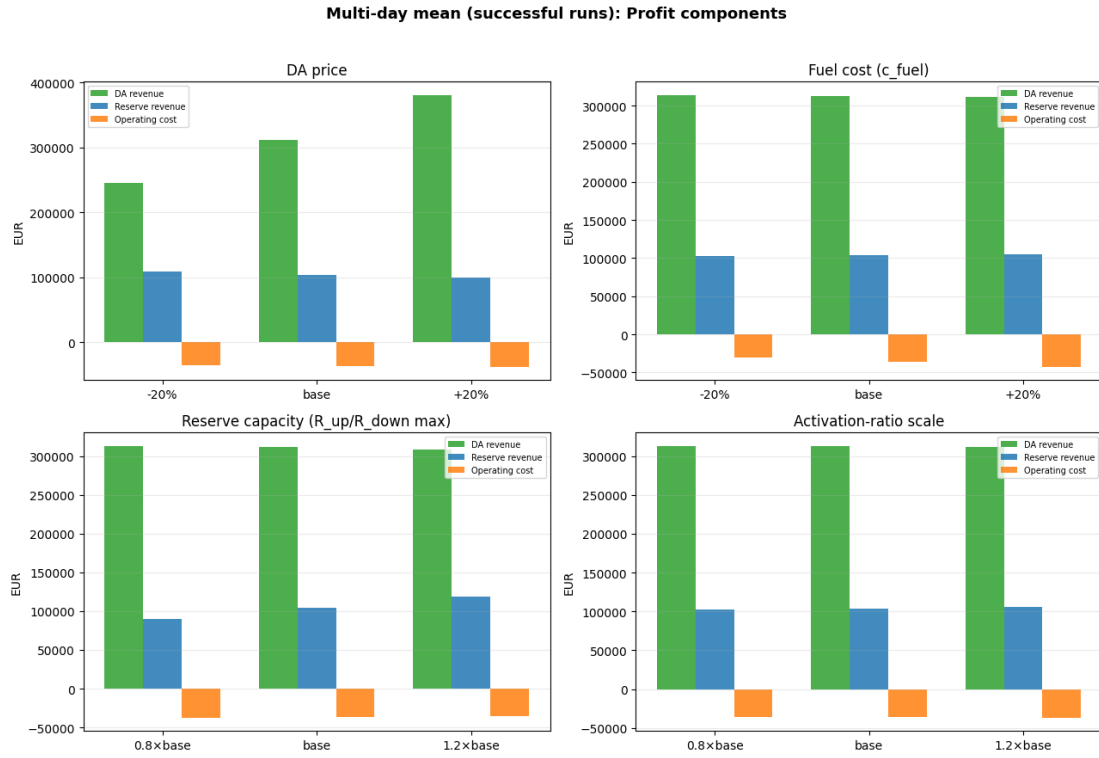


Figure 6.15: Mean daily objective components (DA revenue, reserve revenue, operating cost) by perturbed parameter (EUR/day). Shortfall-penalty sensitivity is not tested here.

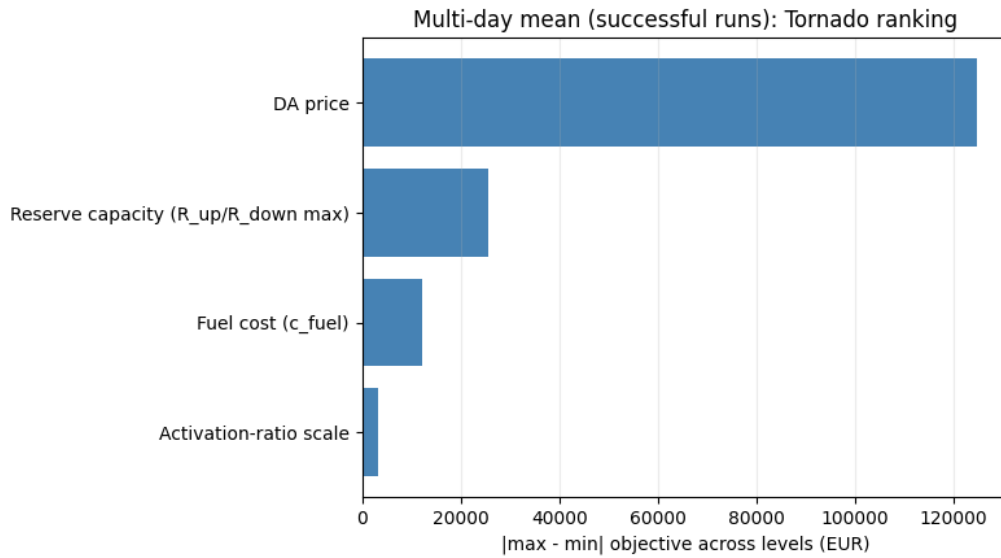


Figure 6.16: Tornado chart: range of the mean daily stochastic-model objective across $\pm 20\%$ (or 0.8–1.2 $\times$) levels for each tested factor (EUR/day); this does not establish VSS robustness.

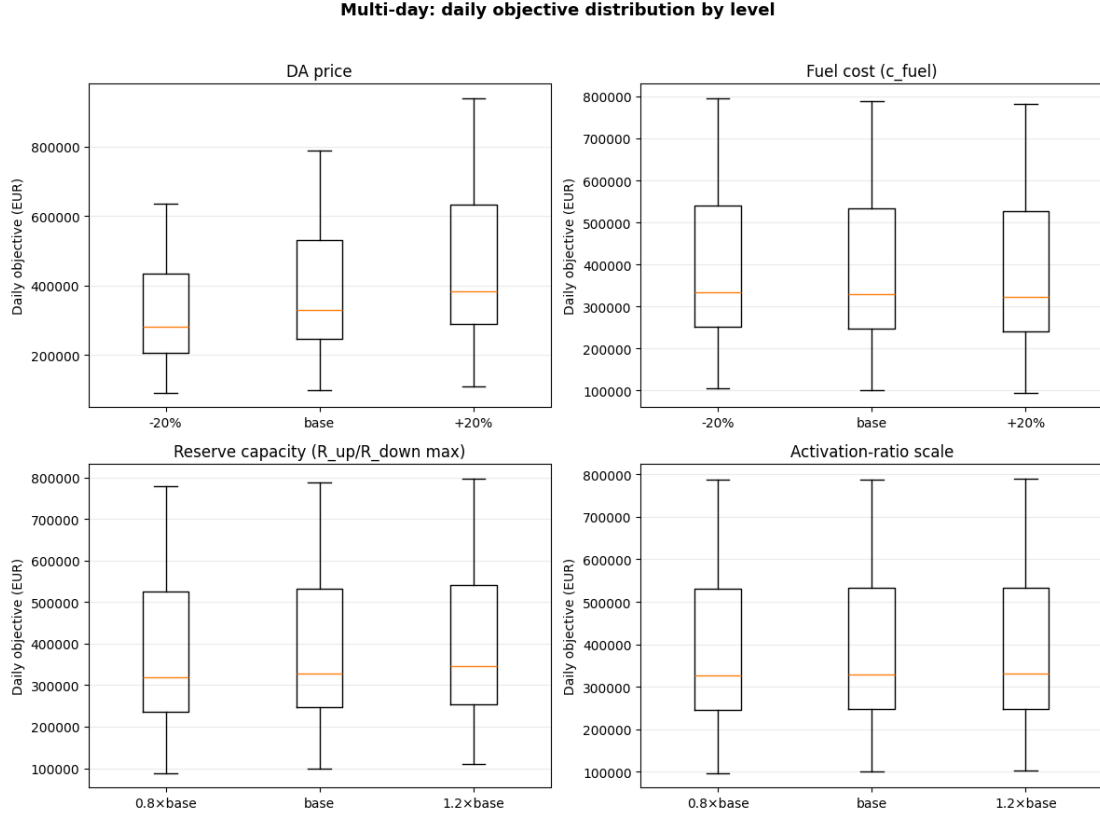


Figure 6.17: Distribution of daily stochastic-model objective values across delivery days for each tested parameter level (EUR/day).

6.5 Case Studies

Three representative delivery days are selected using reproducible rules. The *typical* day, 2026-03-03, has an RP objective close to the sample median. The *stress* day, 2026-03-09, has high combined activation and price stress. The *comparison* day, 2026-02-06, has the largest absolute VSS. For every day, the RP policy, the EV first-stage policy evaluated as EEV, and a fixed-reserve heuristic (20 MW upward and downward) are evaluated with the same scenario probabilities, recourse equations, operating costs, and shortfall penalties.

On the typical day (Figure 6.18), the RP objective reaches EUR 330,999.63/day, compared with an EEV of EUR 26,201.35/day. The resulting VSS is EUR 304,798.28/day. The result shows that a day near the RP median need not be benign for the EV policy: its reserve schedule can still create costly activation shortfalls when evaluated over the same 50-scenario distribution used for RP.

On the stress day (Figure 6.19), the RP objective reaches EUR 574,923.79/day, compared with an EEV of EUR 474,356.57/day. The resulting VSS is EUR 100,567.22/day. Both policies have high model-evaluated values under favourable market conditions, while RP changes the time-varying reserve exposure in response to the joint price and activation scenarios. This is a consequence of the stated risk-neutral expected-profit model and should not be interpreted as a claim of distributional robustness.

The comparison day provides the sharpest model-evaluation contrast. The RP objective reaches EUR 284,045.53/day. The EV policy has an EEV of EUR -243,747.58/day, while the fixed-reserve benchmark reaches EUR 179,829.42/day. This corresponds to a

VSS of EUR 527,793.11/day. As shown in Figures 6.20 and 6.21, the EV policy selects substantially different reserve quantities and then incurs a large expected shortfall penalty under the scenario evaluation. RP’s advantage on this day therefore comes mainly from avoiding that penalty, rather than from uniformly higher market revenue.

These cases reinforce the multi-day result: the correct comparison is RP versus EEV, not RP versus the EV model’s own objective. They also show why the penalty parameter and activation calibration are economically important; VSS in this experiment measures the value of scenario information under that specified evaluation model.

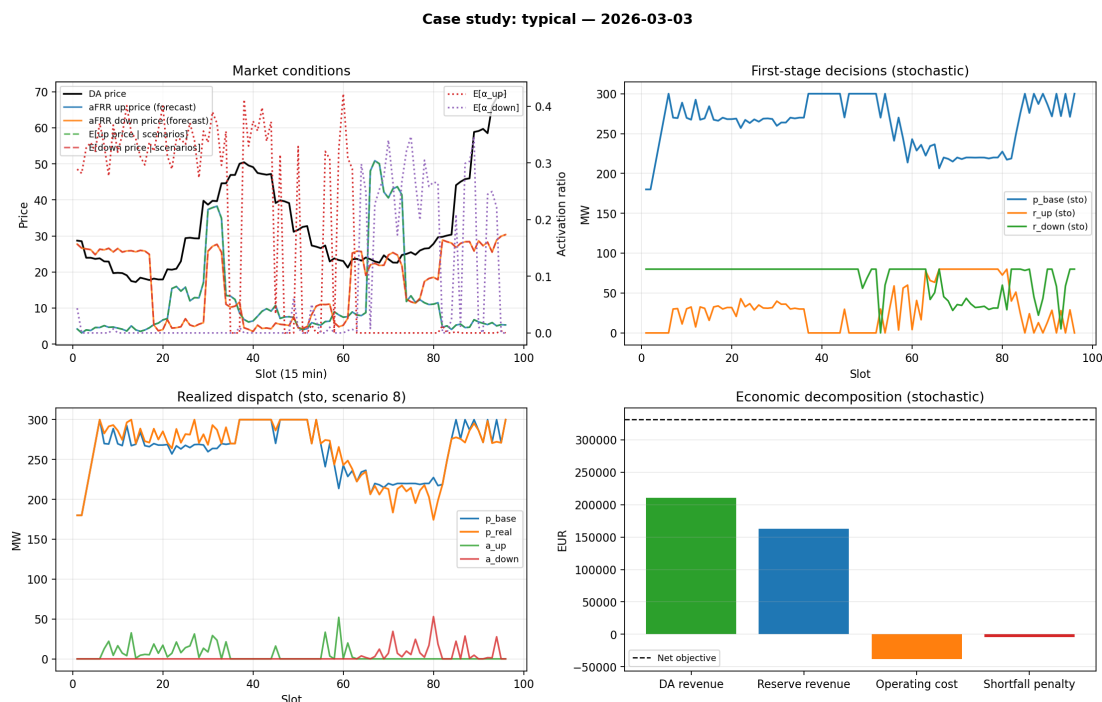


Figure 6.18: Typical day (2026-03-03): market inputs, RP first-stage decisions, representative-scenario dispatch, and objective decomposition in EUR/day including shortfall penalties. The figure’s “sto” and “stochastic” labels denote RP.

Case study: stress — 2026-03-09

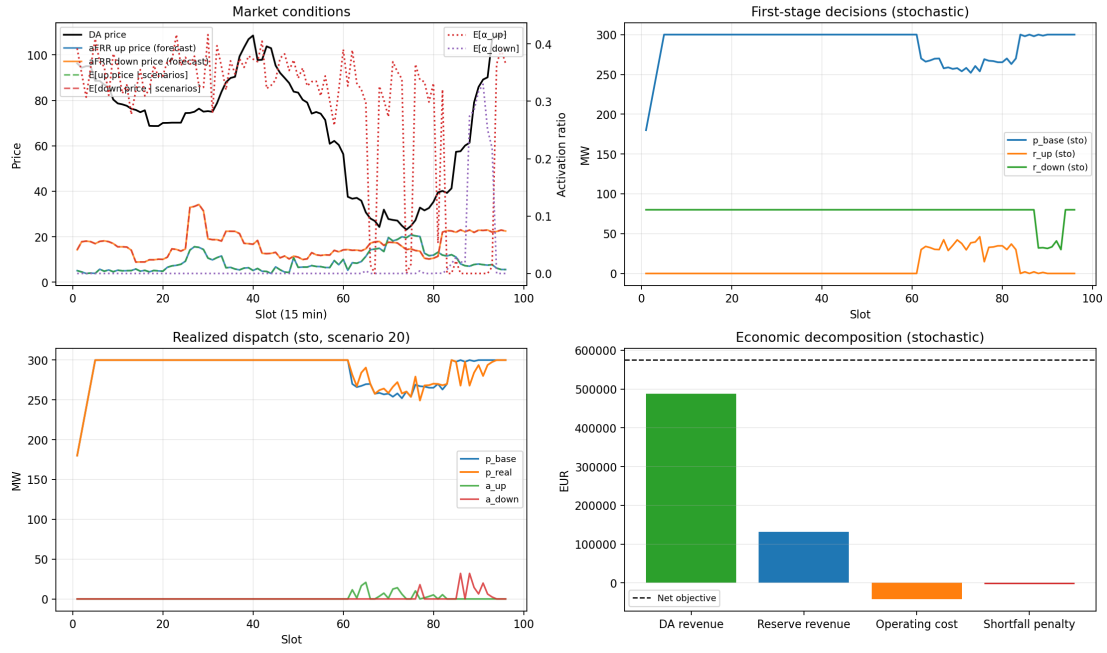


Figure 6.19: Stress day (2026-03-09): market inputs, RP decisions, representative-scenario dispatch, and objective decomposition in EUR/day. The figure’s “sto” and “stochastic” labels denote RP.

Case study: comparison — 2026-02-06

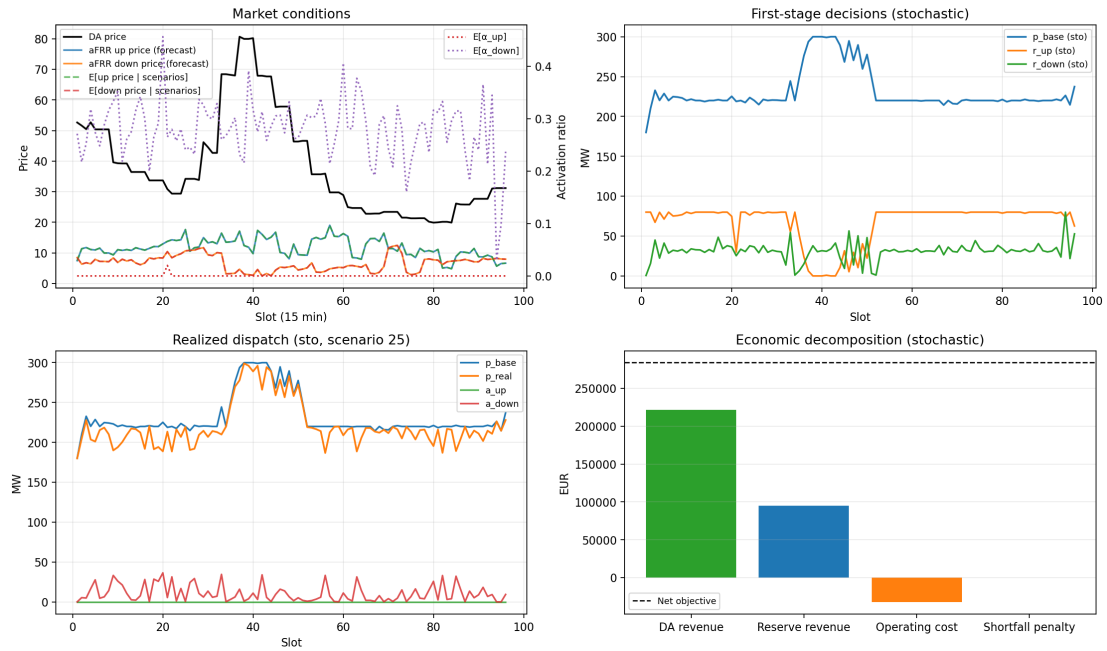


Figure 6.20: Comparison day (2026-02-06), selected by the largest absolute VSS; the objective decomposition is in EUR/day. The figure’s “sto” and “stochastic” labels denote RP.

Strategy comparison — comparison / 2026-02-06

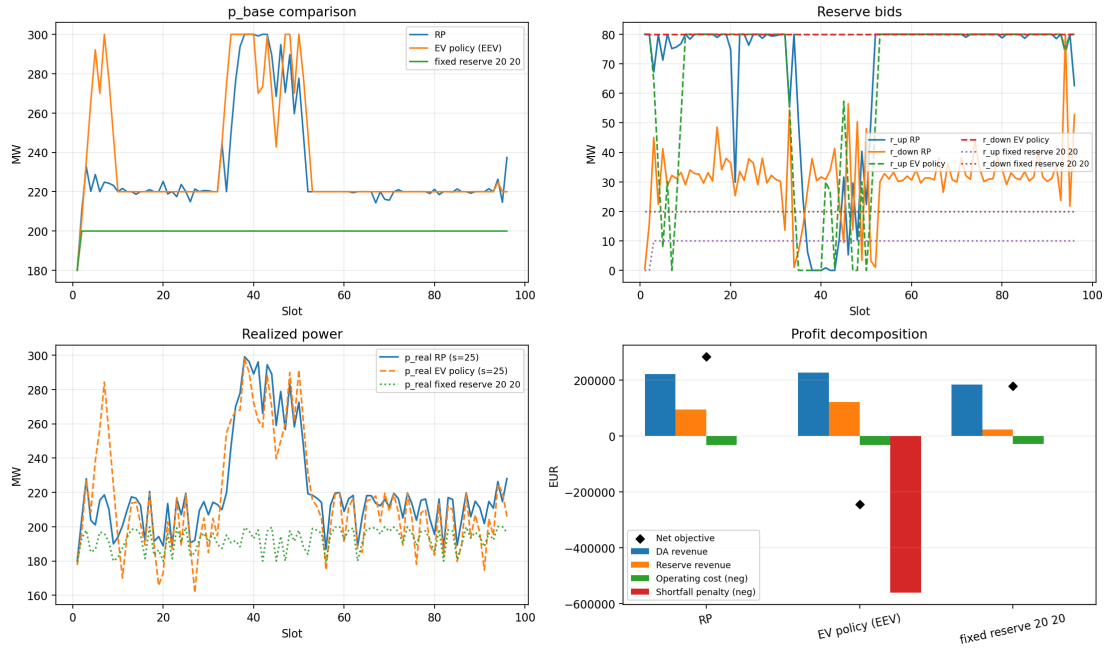


Figure 6.21: Strategy comparison on 2026-02-06: RP, the EV policy evaluated as EEV, and the fixed-reserve benchmark. Diamonds denote daily net objectives in EUR/day; penalty bars are shown explicitly.

Chapter 7

Discussion

7.1 Interpretation of Main Findings

The results of this thesis provide several important insights into the reserve scheduling problem of a gas turbine participating in the aFRR market. These findings can be interpreted from operational, economic, and methodological perspectives.

Operational Perspective

From an operational perspective, the results show that the gas turbine should not be treated as either a pure energy unit or a pure reserve unit. Instead, the most effective strategy is to jointly manage base generation and reserve allocation while respecting the technical flexibility limits of the plant.

The multi-day backtesting results indicate that the optimization model frequently selects a relatively high base generation level, especially when day-ahead electricity prices are favorable. At the same time, the model adjusts upward and downward reserve provision depending on the remaining technical headroom and downward flexibility. In particular, the model often reduces upward reserve when the unit operates close to its maximum output, while maintaining meaningful downward reserve capability at high operating levels. This reflects a realistic operational trade-off: a gas turbine cannot simultaneously maximize energy output and preserve full upward balancing flexibility.

The case-study analysis further confirms that the model behaves in an operationally interpretable way. On high-value energy days, the model pushes the base generation schedule upward, whereas on days with stronger balancing opportunities it adapts reserve allocation accordingly. The resulting dispatch patterns remain physically plausible and respect output and ramping constraints, which indicates that the framework successfully links market opportunities to feasible turbine operation.

Economic Perspective

From an economic perspective, the results suggest that day-ahead electricity revenue is the dominant value stream in the current setup, while reserve participation provides an important complementary source of income. This finding is visible in the revenue decomposition, where DA revenue contributes the largest share of total profit on most days, and reserve revenues add further value depending on market conditions.

The benchmark comparison is interpreted only after placing every policy in the same scenario-based evaluation model. This avoids awarding an optimized policy its own in-sample objective while evaluating another policy under uncertainty.

The EV–stochastic comparison has a different economic interpretation from a direct comparison of model objectives. Mean RP is EUR 380,267.95/day and mean EEV is EUR 32,370.34/day, yielding mean VSS of EUR 347,897.61/day. The advantage comes predominantly from avoiding activation shortfall penalties: RP gives up some reserve revenue and changes the reserve mix, especially by reducing downward reserve, but reduces mean penalty cost from EUR 387,032.74/day to EUR 1,144.58/day. Thus the tested scenario model assigns substantial value to reducing scenario-evaluated delivery shortfalls. This result is conditional on the scenario generator, activation-ratio normalization, and high shortfall penalty; it should not be generalized as evidence of an inherent risk-aversion or robustness property.

The existing OFAT sensitivity analysis does not close this calibration question. It examines selected market inputs, technical limits, fuel cost, and activation scaling, but does not systematically vary the shortfall-penalty coefficient in an RP–EEV comparison. Consequently, it supports statements about general parameter response only, not the robustness of the reported VSS magnitude.

Methodological Perspective

From a methodological perspective, the thesis demonstrates that it is possible to combine forecasting outputs, scenario-based uncertainty representation, and optimization-based reserve scheduling into one coherent framework. The implemented structure is sufficiently rich to capture the interaction between energy value, reserve provision, activation intensity, and gas turbine feasibility, while remaining tractable enough for repeated multi-day testing.

A particularly important methodological contribution is the activation-ratio formulation. Instead of forcing a single gas turbine to satisfy the full system-level balancing demand, the model interprets activation as a fraction of the offered reserve. This makes the uncertainty representation more suitable for single-unit scheduling and avoids unrealistic dispatch interpretations. Although this formulation is still an approximation, it provides a practical bridge between system-level balancing signals and unit-level reserve optimization.

At the same time, the results show that methodology and calibration quality are closely connected. VSS is mathematically well defined only after the deterministic trajectory is constructed from the same distribution and its policy is re-evaluated scenario by scenario. The large observed VSS is also a calibration warning: because avoided penalties dominate the decomposition, better validation of activation ratios, temporal dependence, and the economic shortfall coefficient is required before transferring the numerical magnitude to practice.

7.2 Strengths of the Current Model

The current framework has several clear strengths.

First, it captures the joint trade-off between energy-market participation and reserve provision. This is one of the central practical challenges for gas turbine operation in

balancing markets, and the model addresses it directly by combining day-ahead electricity value with reserve-market remuneration and operating cost.

Second, the model incorporates uncertainty in a structured way. Even though the current stochastic formulation can still be improved, the framework already allows uncertainty in reserve prices and activation intensity to be represented explicitly rather than ignored completely.

Third, the model remains technically interpretable. The inclusion of generation limits, ramping constraints, reserve-capability constraints, and scenario-based realized dispatch ensures that the resulting strategies are not only economically attractive but also physically meaningful.

Fourth, the framework is computationally tractable. The model can be solved repeatedly over many days, which makes systematic backtesting, comparison, and sensitivity analysis feasible. This is particularly important for thesis-level empirical work, where evidence across multiple test days is more informative than an isolated single-day example.

Fifth, the framework is modular and extensible. Forecasting, scenario generation, optimization, and evaluation are implemented as separate but connected components. This makes it possible to refine individual parts of the pipeline without having to rebuild the full structure from scratch.

7.3 Limitations

Despite its strengths, the current model has several limitations that should be acknowledged clearly.

Simplified Market Settlement

The current treatment of market settlement is simplified. The model includes day-ahead electricity value and reserve-related revenue, but it does not fully reproduce the detailed settlement logic of real balancing markets. In particular, the current framework does not explicitly model all interactions between planned energy schedules, activation energy, imbalance settlement, and possible market-specific payment structures. Therefore, the model should be interpreted as an economically meaningful approximation rather than a full market-clearing engine.

Activation-Ratio Approximation

The activation-ratio approach is one of the key simplifications of the thesis. While it provides a tractable way to link reserve offers to realized activation, it remains a proxy rather than a direct representation of unit-level dispatch instructions from the transmission system operator. As a result, the model captures activation intensity rather than exact activation assignment. This is a useful approximation for single-unit scheduling, but it inevitably limits realism.

This approximation may also make the results optimistic. The model assumes that the offered reserve capacity is accessible when needed and that activation can be represented as a proportional call on the offered reserve. In practice, unit-level activation may be affected by market acceptance, detailed dispatch instructions, availability restrictions, wear-related costs, and plant-specific operating conditions. These effects are only partly represented in the current formulation.

No Startup and Shutdown Decisions

The current base model assumes that the gas turbine is available throughout the optimization horizon and does not include startup or shutdown decisions. This avoids the complexity of binary commitment variables, startup costs, shutdown costs, and minimum up/down time constraints. However, it also means that the model cannot answer a broader operational question of whether the unit should be online at all. The present framework therefore focuses on reserve scheduling conditional on unit availability rather than full unit commitment.

No Reserve Acceptance Logic

Another limitation is that the model does not explicitly represent reserve acceptance or awarded reserve volumes. In the current formulation, reserve-related revenues are derived from modeled reserve decisions and market inputs without a separate award mechanism. This may overstate the degree to which all offered reserve can be considered fully market-relevant. A more detailed representation of market acceptance would improve realism.

Data Limitations

The quality of the final optimization results depends strongly on the quality of the forecasting and scenario inputs. Any limitations in reserve-price forecasts, activation proxies, scenario probabilities, or data coverage will directly affect model quality. In particular, the current results suggest that the stochastic model may be sensitive to scenario calibration, which reinforces the importance of high-quality uncertainty inputs.

Scenario Dependence and VSS Calibration

Each daily scenario uses a common identifier for its sampled reserve-price residuals, activation states, and conditional activation volumes. This construction represents differences among scenarios, but it does not explicitly model the full temporal dependence within a day or the complete cross-variable correlation among price, activation direction, and activation volume. The current implementation preserves hour-conditioned residual distributions, interval-specific state probabilities, and state-conditional volume distributions. Validation of temporal and cross-variable dependence remains future work.

The reported VSS is also conditional on the activation-ratio normalization and the shortfall-penalty coefficient. Because avoided shortfall penalties account for most of VSS, the present OFAT analysis cannot be used to claim that the absolute VSS amount is robust. Additional penalty calibration and RP-EEV sensitivity analysis are required before attaching an industrial monetary interpretation to that amount.

7.4 Implications for Practice

Although the framework is developed in a research setting, it already provides several useful practical insights for operators and industrial stakeholders such as Siemens Energy.

First, the model shows that reserve scheduling can be studied as a joint optimization problem rather than through fixed heuristics. Under the common in-sample scenario

evaluation, RP has a higher mean daily objective value than each tested rule. This ranking is conditional on the current scenario and penalty calibration.

Second, the results highlight the importance of energy-market conditions in reserve participation decisions. In the current setup, day-ahead electricity revenue is often the dominant driver of profitability. This means that reserve participation strategies should not be developed independently of energy-market expectations.

Third, the framework provides a structured way to assess flexibility value. By explicitly showing how reserve allocation changes with prices, activation conditions, and turbine constraints, the model helps quantify when flexibility is economically valuable and when it is better used for energy production.

Fourth, the VSS experiment provides a model diagnostic for uncertainty modelling. In the tested sample, the RP schedule substantially reduces activation shortfall relative to the EV policy. An industrial implementation could use the same fixed-policy evaluation to examine the modelled cost of ignoring uncertainty before changing operating practice. However, the present monetary magnitude must not be treated as realized savings or a business-case forecast: it is highly sensitive to the assumed shortfall penalty and to scenario calibration.

For Siemens Energy specifically, the framework is best interpreted as a prototype decision-support and model-validation tool. It can compare candidate first-stage turbine schedules on one controlled scenario distribution, expose whether value arises from DA revenue, reserve revenue, operating cost, or deliverability penalties, and identify days for detailed engineering review. Deployment would additionally require plant-specific activation instructions, validated penalty or settlement coefficients, reserve-award logic, and out-of-sample testing. The single-unit, continuous-dispatch formulation does not yet support commitment or portfolio decisions.

7.5 Future Work

Several directions for future work follow naturally from the results of this thesis.

A first priority is out-of-sample calibration of the activation ratios and the shortfall penalty. Avoided penalty cost accounts for most of the measured VSS, so the 95th-percentile volume normalization and the coefficient of 1,000 should be tested against unit-level delivery data or defensible settlement consequences. Reporting VSS over a penalty grid would distinguish structural scheduling value from penalty-driven value.

A second priority is scenario validation. Future work should test marginal calibration, cross-direction and temporal dependence, extreme-event coverage, and stability as the number of scenarios changes. A genuinely out-of-sample evaluation should fix policies using one scenario set and assess them on withheld realizations or an independently generated scenario set.

A further extension is the inclusion of reserve acceptance constraints or awarded reserve logic. This would make the market representation more realistic by distinguishing between reserve offers and reserve that is actually accepted in the market.

Another improvement is the explicit representation of activation energy remuneration. At present, the economic role of activation is captured only indirectly through realized dispatch and operating cost. A more detailed treatment of activation-related settlement would strengthen the market-side realism of the model.

Risk-aware optimization, for example through CVaR or other downside-oriented mea-

tures, could then be studied separately from the risk-neutral RP formulation. The present VSS result measures the value of using the distribution in expected-value optimization; it does not by itself establish a robustness benefit.

Finally, if the scope is expanded, startup and shutdown decisions could be incorporated through a unit commitment formulation. This would allow the model to address not only reserve scheduling conditional on availability, but also the more fundamental decision of whether the unit should be online in the first place.

For larger stochastic instances, decomposition methods could also be considered. Since the current stochastic model has a two-stage linear structure, Benders decomposition, also known as the L-shaped method in stochastic programming, is a natural candidate for separating first-stage scheduling decisions from scenario-dependent recourse calculations. This was not implemented in the thesis because the tested instances remain small enough to solve directly, but it would become more relevant if the scenario set or the plant representation were expanded.

Another possible extension is a system-level formulation based on column generation. In such a framework, a single-unit profit-maximization model similar to the one developed in this thesis could be used as a pricing problem that generates profitable dispatch and reserve schedules. A system-level master problem could then select schedules across multiple units to meet balancing requirements. This would move the model from single-unit decision support toward a broader market or system optimization framework.

Overall, the future development path is clear: the current framework provides a useful and interpretable core, and the next improvements should focus on better uncertainty representation and more realistic market modeling before adding the full complexity of commitment decisions.

Chapter 8

Conclusion

8.1 Summary of the Thesis

This thesis developed an end-to-end framework for scheduling the base generation and aFRR reserve capacity of a single gas turbine under uncertain reserve prices and activation. The framework combines next-day forecasting, residual-bootstrap scenarios, a two-stage linear program, multi-day evaluation, benchmark policies, sensitivity tests, and operational case studies. Its scope remains reserve scheduling conditional on unit availability: unit commitment, startup and shutdown, portfolio coordination, reserve acceptance, and full balancing-market settlement are outside the model.

A central correction in the final evaluation is the alignment of the deterministic and stochastic formulations. The deterministic model is defined as the expected-value problem of the actual 50-scenario daily distribution. Its reserve-price and activation-ratio inputs are the probability-weighted scenario means, while the DA price is a common non-random trajectory. The EV first-stage policy is then fixed and re-optimized only in recourse to obtain EEV. This permits the standard maximization measure $VSS = RP - EEV$; a direct comparison of the EV and RP internal objectives is not used as a performance claim.

Across 42 complete delivery days, mean RP is EUR 380,267.95/day and mean EEV is EUR 32,370.34/day. Mean VSS is EUR 347,897.61/day and median VSS is EUR 363,192.73/day. VSS is positive on all 42 days, with no negative observation beyond or within the 10^{-5} tolerance. RP reduces average upward reserve from 43.11 MW to 35.25 MW and downward reserve from 79.93 MW to 50.11 MW relative to the EV policy, while increasing average base output by 2.93 MW. The value does not arise from higher reserve revenue. Instead, mean activation-shortfall penalty falls from EUR 387,032.74/day under EEV to EUR 1,144.58/day under RP, more than offsetting lower reserve revenues and slightly higher operating cost.

The empirical conclusion is therefore conditional but clear. Within the specified scenario distribution and penalty model, using the stochastic formulation improves the expected performance of the first-stage schedule under the same evaluation environment. The positive VSS demonstrates the value of scenario-aware scheduling within the specified evaluation model. However, its monetary magnitude is driven mainly by avoided activation-shortfall penalties and should not be interpreted directly as an industrial business-case value. It remains sensitive to activation-ratio normalization, scenario calibration, and the penalty coefficient. The current sensitivity analysis does not fully test VSS against changes in that coefficient, so it neither establishes VSS robustness nor

an out-of-sample monetary business case.

The main practical contribution is a transparent validation framework: candidate turbine schedules can be evaluated on a common uncertainty distribution, their value decomposed, and scenario-delivery risks identified explicitly. Before industrial use, future work should validate activation and penalty assumptions against unit-level and settlement data, evaluate policies out of sample, and add reserve-award and commitment logic. Subject to these boundaries, the results support scenario-based optimization as a useful research and decision-support approach for coordinating energy and reserve scheduling.

Bibliography

- [1] John R. Birge and François Louveaux. *Introduction to Stochastic Programming*. Springer, 2 edition, 2011.
- [2] Theodor Borsche, Andreas Ulbig, and Göran Andersson. A new frequency control reserve framework based on energy-constrained units, 2013.
- [3] Sergio Cantillo-Luna, Ricardo Moreno-Chuquen, Jesus Lopez-Sotelo, and David Celeita. An intra-day electricity price forecasting based on a probabilistic Transformer neural network architecture. *Energies*, 16(19):6767, 2023.
- [4] Antonio J. Conejo, Miguel Carrión, and Juan M. Morales. *Decision Making Under Uncertainty in Electricity Markets*. Springer, 2010.
- [5] Javier Contreras, Rosario Espinola, Francisco J. Nogales, and Antonio J. Conejo. ARIMA models to predict next-day electricity prices. *IEEE Transactions on Power Systems*, 18(3):1014–1020, 2003.
- [6] ENTSO-E. ENTSO-E transparency platform. <https://transparency.entsoe.eu/>, 2026. Accessed: 2026-06-18.
- [7] ENTSO-E. PICASSO: Platform for the international coordination of automated frequency restoration and stable system operation. https://www.entsoe.eu/network_codes/eb/picasso/, 2026. Accessed: 2026-06-18.
- [8] European Commission. Commission regulation (EU) 2017/2195 establishing a guideline on electricity balancing. <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32017R2195>, 2017. Accessed: 2026-06-18.
- [9] Jesus Lago, Fjo De Ridder, and Bart De Schutter. Forecasting spot electricity prices: Deep learning approaches and empirical comparison of traditional algorithms. *Applied Energy*, 221:386–405, 2018.
- [10] Jesus Lago, Grzegorz Marcjasz, Bart De Schutter, and Rafał Weron. Forecasting day-ahead electricity prices: A review of state-of-the-art algorithms, best practices and an open-access benchmark. *Applied Energy*, 293:116983, 2021.
- [11] Elaheh Mashhour and Seyed Masoud Moghaddas-Tafreshi. Bidding strategy of virtual power plant for participating in energy and spinning reserve markets—Part I: Problem formulation. *IEEE Transactions on Power Systems*, 26(2):949–956, 2011.
- [12] Juan M. Morales, Antonio J. Conejo, Henrik Madsen, Pierre Pinson, and Marco Zugno. *Integrating Renewables in Electricity Markets: Operational Problems*. Springer, 2014.

- [13] Ciaran O'Connor, Joseph Collins, Steven Prestwich, and Andrea Visentin. Electricity price forecasting in the irish balancing market, 2024.
- [14] Open-Meteo. Open-meteo weather forecast API. <https://open-meteo.com/>, 2026. Accessed: 2026-06-18.
- [15] Ivan Pavić, Hrvoje Pandžić, and Tomislav Capuder. Electric vehicle aggregator as an automatic reserves provider in the european market setting, 2020.
- [16] Rafał Weron. Electricity price forecasting: A review of the state-of-the-art with a look into the future. *International Journal of Forecasting*, 30(4):1030–1081, 2014.
- [17] Florian Ziel and Rafał Weron. Day-ahead electricity price forecasting with high-dimensional structures: Univariate vs. multivariate modeling frameworks. *IEEE Transactions on Power Systems*, 33(3):3312–3321, 2018.