Signal-Conditioned Bidding and Curtailment for Wind Power

Under the Nordic Imbalance Settlement Model

Harald Caap

Master's thesis
2026:E74

Lund Institute of Technology
Lund University

Faculty of Engineering
Centre for Mathematical Sciences
Mathematical Statistics

Centrum Scientiarum Mathematicarum

Abstract

Wind generation's growing share of Nordic electricity supply has made imbalance settlement an increasingly costly exposure for producers. Under the Nordic single-price mechanism, all residual imbalance is settled at a single imbalance price, regardless of direction, determined by the mFRR activation market. Before that price is known, the producer fixes a day-ahead position, adjusts on the continuous intraday market, and may curtail output via a setpoint. Whatever gap remains between the committed position and actual output is settled at the imbalance price, with a symmetric eSett fee on every megawatt-hour of imbalance. The incremental revenue from each commitment is determined by the spread between the day-ahead or intraday price and the imbalance price.

The central finding is that exploiting the stage-specific spread signal simultaneously raises mean daily revenue and reduces tail loss, with no Pareto trade-off at any stage where a directional spread signal is present. Across three Swedish parks, the proposed strategy raises mean daily revenue by approximately 9–18 % and compresses $\mathrm{CVaR}_{0.9}$ tail loss, by roughly 60 % at the most exposed park,[1] with the worst-decile expected revenue at another flipping from a small loss to net positive. At the day-ahead stage, the spread carries no interval-level directional information and no rule improves over the forecast baseline.

When the intraday price is high relative to the expected imbalance price, locking in volume at the intraday market pays; when the settlement price is expected to be higher, holding back volume to settle as imbalance pays. The governing principle is to tilt the bid in proportion to confidence in this spread signal, and to curtail output when negative settlement prices are expected. Empirically, signal quality grows monotonically from the day-ahead to the setpoint stage: at chance at the day-ahead stage, moderate accuracy at the intraday stage, and strong enough at the setpoint stage to support a discrete fee-band classifier.

Keywords: Wind Power; Nordic Electricity Market; Single-Price Imbalance Settlement; Quantile Bidding; Newsvendor Problem; Mean–CVaR Optimisation; Probabilistic Forecasting; Bootstrap Scenario Generation.

[1] All headline magnitudes in this thesis (per-park objective lifts, revenue and tail-loss percentages) are sourced from Table 8.1.

Preface

This thesis was carried out at the Centre for Mathematical Sciences, Division of Mathematical Statistics, Faculty of Engineering, Lund University, in collaboration with Holmen Energi AB.

I would like to thank my supervisor Erik Lindström and doctoral student Sofia Kockum at the Division of Mathematical Statistics for their guidance and support throughout this work.

A special thank you to Jakob Mitani at Holmen for giving me this opportunity and for making the collaboration both stimulating and enjoyable.

I would also like to thank Elias Kull and Erik Myklebust at Optimeering, who generously supported this thesis by supplying the imbalance price forecasts that underpin much of the empirical analysis.

Lund, May 2026

Harald Caap

Contents

List of Abbreviations

This list collects the abbreviations used repeatedly in the thesis. Thesis-specific named rules and forecasting models (BidF, BFS, BCW, ThreeRegion, ID-Linear, SP-Persistence, etc.) are defined where they are introduced: see Table 7.1 for the candidate bidding rules and Appendix A for the in-house production forecasts.

aFRR Automatic Frequency Restoration Reserve.

BRP Balance responsible party.

ACT mFRR activation price (the up- or down-regulation energy price that sets the imbalance price).

CCF Cross-correlation function.

CDF Cumulative distribution function.

CET Central European Time.

CRPS Continuous ranked probability score.

CVaR Conditional Value-at-Risk.

DA Day-ahead.

EAM Energy Activation Market.

ECMWF European Centre for Medium-Range Weather Forecasts.

ESS Effective sample size ($1/\sum_d w_d^2$).

EUR Euro.

ID Intraday.

IM Imbalance.

ISP Imbalance Settlement Period.

KDE Kernel density estimation.

kNN k-nearest neighbours.

KS Kolmogorov–Smirnov.

LP Linear program.

MAE Mean absolute error.

MAI Mean absolute imbalance.

MBA Market Balance Area.

mFRR Manual Frequency Restoration Reserve.

MI Mutual information (used only in the kNN feature screening of Chapter 6; the mean-imbalance metric is written $\overline{\nu}$).

MTU Market time unit.

MWh Megawatt-hour.

NWP Numerical weather prediction.

PIT Probability integral transform.

SAA Sample-average approximation.

SAA LP Sample-Average Approximation Linear Program: the scenario mean–CVaR program (Equation (3.11)).

SCADA Supervisory control and data acquisition.

SE1, SE3 Swedish bidding zones used in the empirical analysis.

SIDC Single Intraday Coupling.

SNR Signal-to-noise ratio.

SP Setpoint stage (thesis convention for the producer's pre-delivery curtailment decision; not an industry-standard trading stage).

TSO Transmission system operator.

UTC Coordinated Universal Time.

VaR Value-at-Risk.

VRE Variable renewable energy.

VWAP Volume-weighted average price.

List of Notation

The list below gives the main symbols used across the thesis. Local variables introduced only for a single derivation are defined where they appear.

Sets and Indices

$\mathcal{T}$	Set of settlement intervals in the evaluation horizon.
$\mathcal{T}_d$	Subset of intervals belonging to trading day d.
t	Settlement interval index.
d	Trading day index.
D	Number of trading days in the evaluation window.
N	Number of evaluation intervals, $N = \|\mathcal{T}\|$.
S	Number of scenarios in a sample-average approximation.
s	Scenario index, $s = 1, \ldots, S$.
Δt	Settlement interval duration, equal to 0.25 hours for 15-minute data.

Production, Forecasts, and Schedules

y_t	Realised production in interval t.
a_t	Available production before setpoint curtailment.
$\hat{a}_t^{\mathrm{DA}}$	Day-ahead forecast of available production.
$\hat{a}_t^{\mathrm{ID}}$	Intraday forecast of available production.
$\hat{a}_t^{\mathrm{SP}}$	Setpoint-stage forecast of available production.
s_t	Setpoint or production cap.
$P^{\max}$	Installed capacity of the wind park.
b_t^{DA}	Day-ahead cleared position.
Δb_t^{ID}	Accepted intraday adjustment.
b_t	Final schedule, $b_t = b_t^{\mathrm{DA}} + \Delta b_t^{\mathrm{ID}}$.

Prices and Spreads

π_t^{DA}	Day-ahead market price.
π_t^{ID}	Intraday price, represented as the effective price of the accepted ID adjustment.
π_t^{IM}	Imbalance price.
$\hat{\pi}_t^{\mathrm{DA}}$	Forecast or decision-time signal for the day-ahead price.
$\hat{\pi}_t^{\mathrm{IM,ID}}$	Imbalance-price forecast available at intraday decision time.
$\hat{\pi}_t^{\mathrm{IM,SP}}$	Imbalance-price forecast available at setpoint decision time.
δ_t^{DA}	Day-ahead to imbalance spread, $\delta_t^{\mathrm{DA}} = \pi_t^{\mathrm{DA}} - \pi_t^{\mathrm{IM}}$.
δ_t^{ID}	Intraday to imbalance spread, $\delta_t^{\mathrm{ID}} = \pi_t^{\mathrm{ID}} - \pi_t^{\mathrm{IM}}$.
$\hat{\delta}_t^{\mathrm{DA}}$	Point forecast of the DA-stage imbalance spread $\delta_t^{\mathrm{DA}} = \pi_t^{\mathrm{DA}} - \pi_t^{\mathrm{IM}}$ at $\mathcal{F}^{\mathrm{DA}}$ (so that $\hat{\pi}_t^{\mathrm{IM,DA}} = \pi_t^{\mathrm{DA}} - \hat{\delta}_t^{\mathrm{DA}}$). Derived from the Optimeering raw output via the sign flip in Remark 4.1.
$\hat{\delta}_t^{\mathrm{ID}}$	Point forecast of the ID-stage imbalance spread $\delta_t^{\mathrm{ID}} = \pi_t^{\mathrm{ID}} - \pi_t^{\mathrm{IM}}$ at $\mathcal{F}^{\mathrm{ID}}$, defined as $\pi_t^{\mathrm{ID}} - \hat{\pi}_t^{\mathrm{IM,ID}}$ (Equation (7.4)).
$\hat{\sigma}_t$, $\hat{\sigma}_{\delta,t}^{\mathrm{ID}}$	Predictive standard deviation of the spread forecast at the relevant stage; the ID-stage scale is read from the 10/90 fan as $\hat{\sigma}_{\delta,t}^{\mathrm{ID}} = (Q_{0.90}^{\mathrm{IM,ID}} - Q_{0.10}^{\mathrm{IM,ID}})/(2 \cdot 1.28)$.

Imbalance, Fees, and Revenue

ν_t	Signed imbalance, $\nu_t = y_t - b_t$.
f^{imb}	Symmetric imbalance fee per MWh.
F_t	Total fee term in interval t.
R_t	Total interval revenue.
R_t^{DA}	Day-ahead revenue component.
R_t^{ID}	Intraday revenue component.
R_t^{IM}	Imbalance revenue component.
$\bar{R}_d$	Daily revenue for day d.
$\bar{R}$	Average daily revenue over the evaluation window.

Information Sets and Forecast Distributions

$\mathcal{F}$, $\mathcal{F}_t$	Generic information set used in stage-agnostic decision-theoretic derivations; the stage-specific instances are $\mathcal{F}^{\mathrm{DA}}$, $\mathcal{F}^{\mathrm{ID}}$, $\mathcal{F}^{\mathrm{SP}}$.
$\mathcal{F}^{\mathrm{DA}}$	Information available at day-ahead decision time.
$\mathcal{F}^{\mathrm{ID}}$	Information available at intraday decision time.
$\mathcal{F}^{\mathrm{SP}}$	Information available at setpoint decision time.

Objective, Risk, and Metrics

$\mathbb{E}[\cdot]$	Expectation operator.
λ	Risk-aversion parameter.
λ_U, λ_L	Upper- and lower-tail dependence coefficients (Appendix C.9).
α	CVaR confidence level when used in CVaR_α; the standard value is 0.9.
CVaR_α	Conditional Value-at-Risk at confidence level α.
VaR_α	Value-at-Risk at confidence level α.
η	Rockafellar–Uryasev auxiliary variable for VaR.
z_s	Non-negative auxiliary variable for scenario s in the CVaR epigraph form.
$\overline{\nu}$	Mean imbalance.
MAI	Mean absolute imbalance.
nMAI	Capacity-normalised mean absolute imbalance; $\mathrm{nMAI} = \mathrm{MAI}/P^{\max}$.
CRPS	Continuous ranked probability score.
MAE	Mean absolute error.
RMSE	Root mean squared error.
ME	Mean error.
Cov_τ	Empirical coverage at quantile level τ: fraction of realised values falling within the central τ-prediction interval.

Policy and Regularization Parameters

γ_t	Signal-quality attenuation factor, $\gamma_t \in [0, 1]$; scales the risk-neutral bid tilt by the per-interval SNR of the spread forecast ($\gamma_t \to 0$: production-fan median; $\gamma_t \equiv 1$: full tilt).
θ	Generic scalar strategy parameter in adaptive bidding rules.
τ	Quantile level (when time-indexed as τ_t, the per-interval bid quantile); τ^* denotes the optimal bid quantile.
Q_p^{fc}	p-quantile of the predictive production fan; a central interval of half-width w is $[Q^{\mathrm{fc}}_{0.5-w}, Q^{\mathrm{fc}}_{0.5+w}]$, with the SAA LP using a per-day calibrated half-width w_d (see Appendix C).
W_t	Predictive fan width at interval t, $W_t = Q^{\mathrm{fc}}_{0.99} - Q^{\mathrm{fc}}_{0.01}$.
$\tau_{\mathrm{lo}}, \tau_{\mathrm{hi}}$	Lower and upper probability levels for the forecast-quantile bid bounds ($\tau_{\mathrm{lo}} < \tau_{\mathrm{hi}}$, both in $(0, 1)$).

Chapter 1

Introduction

1.1 Motivation

Holmen Energi operates three wind parks in Sweden: Blabergsliden and Blisterliden in bidding zone SE1 (Luleå) and Varsvik in SE3 (Stockholm). All three participate in the Nordic day-ahead (DA) market and can adjust their positions in the intraday (ID) continuous market. Any residual deviation between the final schedule and realised production is settled at the imbalance (IM) price.

Nordic imbalance prices have grown increasingly volatile, driven by rapid renewable expansion and tighter balancing conditions. This volatility has reshaped the economics of balance-responsibility outsourcing across the industry. Many wind producers have responded by taking direct imbalance exposure rather than transferring it to specialised third parties. Holmen Energi exemplifies this broader transition: since 2025, it carries the imbalance exposure directly for all three parks. The prior strategy was to bid full expected production on the DA market and let an outsourcing contract absorb any deviation. With direct exposure, the question becomes how to manage risk and revenue systematically under the Nordic single-price settlement rules.

The situation Holmen faces, a renewable producer operating under single-price imbalance settlement with volatile imbalance prices and sequential trading opportunities, is structurally identical to the situation facing many wind and solar producers across Europe. The shift from outsourced balance responsibility toward direct imbalance exposure is a Europe-wide trend as markets tighten and imbalance volatilities grow. The methodological framework developed in this thesis is therefore generalisable even though it is calibrated on specific parks and a specific market.

The thesis is written for a reader with a graduate-level quantitative background. Prior familiarity with electricity markets is helpful but not required. The market mechanics needed to follow the analysis are summarised in Chapter 2, and a consolidated notation and abbreviation list precedes the main text for reference.

1.2 Problem Statement

The thesis studies bidding by a price-taking wind producer as a sequential decision problem under uncertainty. A *price-taker* is a producer whose own traded volume is too small to move market prices, so it optimises against prices it takes as given (the assumption is stated in full in Section 2.5). For each settlement interval t, the producer faces up to three decisions at successive timepoints. First, on day $D-1$ before 12:00, the producer chooses a day-ahead position b_t^{DA}. Second, at least one hour before delivery, the producer may revise that position by an intraday adjustment Δb_t^{ID}. Third, shortly before delivery, the producer may choose a setpoint s_t that caps physical output below available production a_t. Any residual deviation between realised production y_t and the final schedule $b_t = b_t^{\mathrm{DA}} + \Delta b_t^{\mathrm{ID}}$ is settled at the imbalance price π_t^{IM}, observable only 2–13 days after delivery.

The three stages are referred to throughout as DA, ID, and SP. The *setpoint stage* (SP) is not a market-defined trading stage in the sense of the day-ahead and intraday markets; it is a thesis convention for the producer's last opportunity to reduce committed output before physical delivery, used here to give the curtailment decision a label symmetric with DA and ID. No new market clearing happens at SP — the DA and ID positions remain committed — but the producer can still influence imbalance volume by curtailing $s_t < a_t$.

All three stages are governed by the same underlying economics, derived formally in Chapter 2. At the DA and ID stages, the relevant object is the spread between the day-ahead/intraday price and the imbalance price: its sign determines whether committing more or less volume to the market is preferable. At the setpoint stage, the committed position is already fixed; the decision reduces to whether the imbalance price itself is high enough to make production worthwhile. What changes across stages is how much of that price signal is predictable in advance.

At the DA stage, 12–36 hours before delivery, no day-ahead or imbalance price has yet cleared, and only general meteorological and system-state information is available. At the ID stage, 60–105 minutes before delivery, the producer has substantially more information, including an updated production forecast, intraday bids, and market state information impacting the imbalance price. At the setpoint stage, approximately 20 minutes before delivery, the production forecast has tightened considerably and near-real-time system signals can be observed. Whether any of these information sets contains a statistically defensible directional signal — and if so, how strongly that signal should impact the producer's decisions — is the central empirical question of the thesis.

The organising conjecture is that the bid should tilt away from the production forecast in the direction the imbalance-price forecast points — to get the best price for the producer's volume and to stay off the wrong side of sharp price moves — but only as far as the signal-to-noise ratio of that forecast permits. A weak or absent signal leaves the production forecast as the natural bid; a strong signal warrants a tilt. Strategies are scored on a revenue–risk objective that penalises tail losses on bad trading days; the formal criterion is defined in Chapter 3.

1.3 Research Questions

The overarching question is whether stage-appropriate price signals can improve the revenue–risk trade-off of a Nordic wind producer relative to simply bidding the production forecast. Each stage is evaluated against its own baseline — bid the forecast at DA, do not adjust the bid on the ID market, do not curtail at SP — with an oracle providing the upper reference. This is broken into three sub-questions:

1. What revenue and risk does bidding the production forecast produce under current Nordic single-price rules, and where do the losses concentrate?
2. At each of the three stages, does the imbalance-price forecast carry directional information, and if so, how should it be translated into a bid?
3. How does the mean–CVaR improvement accumulate as each stage is added to the strategy, and where does the residual exposure concentrate after all signal-based rules are applied?

1.4 Contributions and Scope

The core contribution is *signal-quality-aware quantile bidding under Nordic single-price settlement*. The principle is to bid the τ_t^*-quantile of the predictive production distribution, with τ_t^* scaled by the signal-to-noise ratio of the available imbalance-spread forecast, applied stage-specifically as signal quality permits. Four contributions back the claim. Theoretically, Propositions 3.1–3.3 show that the per-MWh eSett settlement fee creates a closed-form optimal bid at every stage, resolving the all-or-nothing degeneracy that prior single-price work patches externally with hard position limits. As a framework, a stage-wise signal-credibility audit run before any bidding experiment establishes that directional content in the imbalance-spread forecast is at chance at the day-ahead stage, genuine but noisy at intraday, and strong at the setpoint stage. The principle therefore prescribes a plain forecast bid at DA, a bounded tilt at ID, and a discrete classifier at SP. Methodologically, the intraday tilt scaled by signal-to-noise raises mean revenue and reduces tail loss simultaneously at all three parks, and an unweighted variant is statistically indistinguishable on the headline objective, so the scaling is adopted for a materially lower imbalance footprint rather than for revenue. The setpoint classifier routes production away from the negative-price wrong-side tail that an unguided forecast bid enters. Empirically, the combined rules increase mean daily revenue by approximately 9–18 % across three Swedish parks while reducing tail loss simultaneously, with no Pareto trade-off and the improvement preserved under timing- and liquidity-honest execution. These four contributions are restated in full in Section 10.2.

The analysis is positive rather than normative: it evaluates what a single price-taking BRP can achieve under current Nordic rules, without modelling equilibrium feedback or claiming the results generalise to an aggregate of producers running the same strategy. The single-price settlement structure is load-bearing throughout — the optimality propositions and revenue decomposition do not carry over directly to dual-price markets. It is also a study of production bidding rather than financial trading, so the producer sizes only its own volume, and the thesis does not pursue a position in the day-ahead-to-intraday price spread. That bet is decoupled from production and would require forecasting the intraday price at the day-ahead gate, which the thesis does not do.

1.5 Relation to Prior Bidding Literature

Bidding by variable renewable energy producers under single-price imbalance settlement has become an active research area. The thesis draws on two traditions: the optimal-bid form from the newsvendor and quantile-bidding literature (Pinson, Chevallier, and Kariniotakis 2007; Zugno, Jónsson, and Pinson 2013) and the all-or-nothing diagnosis from recent single-price work (Bruninx et al. 2025; Heiser, Pourahmadi, and Kazempour 2024).

The closest comparators are Bruninx et al. (2025) on Belgian one-price balancing and Heiser, Pourahmadi, and Kazempour (2024) on a hybrid wind–electrolyzer plant: both diagnose that the unconstrained profit-maximising bid degenerates to all-or-nothing under single imbalance pricing. Neither models a per-MWh imbalance fee, so the pathology is patched from outside the model (hard position constraints; learned linear policies). In the Nordic eSett framework, the fee f^{imb} creates an interior newsvendor regime with a closed-form optimal quantile, derived here and absent from those papers.

Feature-driven and online prescriptive bidding (Muñoz, Morales, and Pineda 2020; Muñoz, Pinson, and Kazempour 2023) share the objective of adapting the decision to observed signals but learn the rule end-to-end from data rather than deriving it from the settlement structure. Strategic-bidding work (Zugno, Morales, et al. 2013; Sousa and Algarvio 2025) optimises the same revenue objective for producers with market power; this thesis occupies the complementary price-taking setting.

1.6 Thesis Structure

Figure 1.1 illustrates the logical dependencies between the thesis's core chapters. The structure follows directly from the core contribution: signal-quality-aware quantile bidding requires a calibrated production fan and a stage-specific spread signal as inputs, and a per-stage rule that maps the signal's quality into a quantile tilt. Each ingredient is developed in a dedicated chapter and the three are combined in the final methods chapter.

Chapter 2 introduces the Nordic single-price settlement mechanism, the three sequential trading stages, and the revenue decomposition that underlies the theoretical analysis. Chapter 3 derives the optimal bid for a price-taking producer, establishes the structure underlying the DA and ID stages, and derives the SP curtailment decision. Chapter 4 describes the three parks, data sources, evaluation design, and the forecast infrastructure used throughout.

Chapter 5 evaluates the forecast signals available at each stage: the production point forecasts that anchor the bid, and the stage-specific imbalance-spread forecast $\hat{\delta}_t$ with its predictive scale $\hat{\sigma}_t$, which together set the quantile level τ_t at which to bid. It measures how much

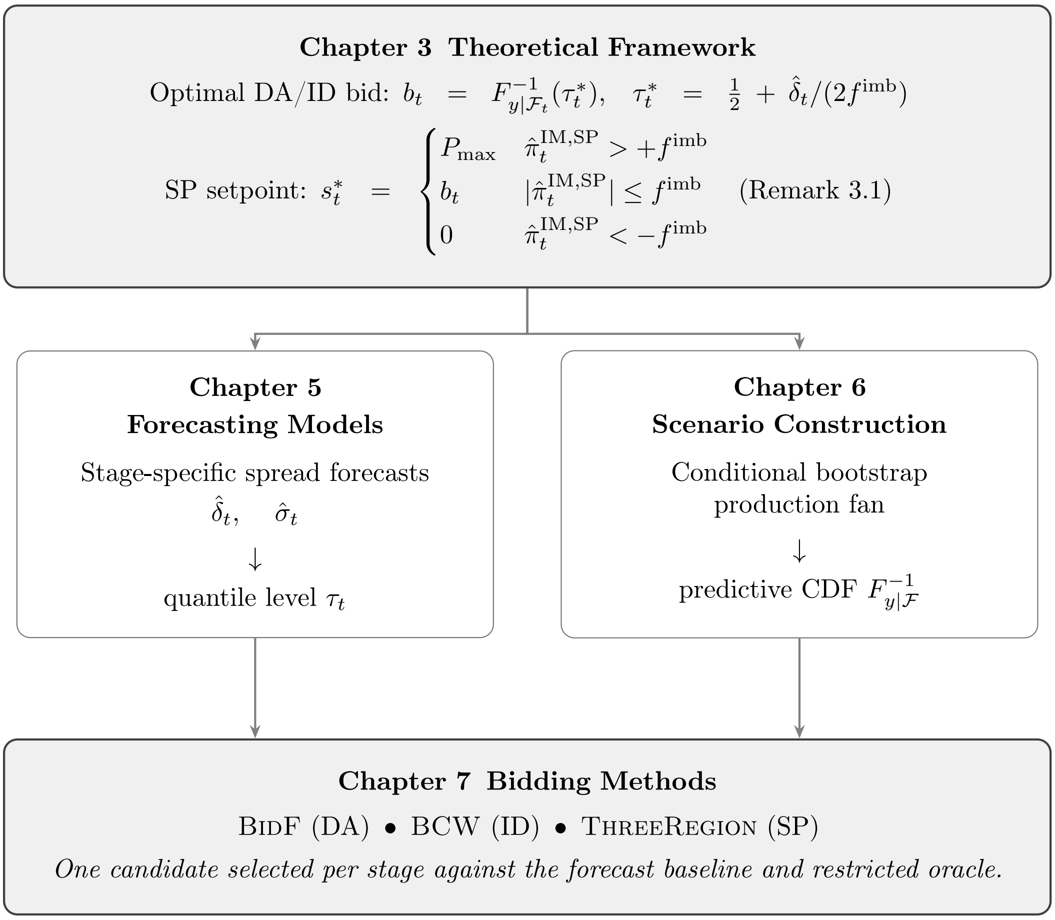

Figure 1.1: Logical structure of the thesis and dependencies between core chapters.

directional content each signal carries, stage by stage. Chapter 6 builds the second ingredient, the production fan: a conditional bootstrap that turns the production forecast into a calibrated predictive distribution, from which the τ_t-quantile is extracted as the bid. Chapter 7 assembles these two inputs into per-stage DA, ID, and SP rules and selects one candidate per stage against the forecast baseline and per-stage perfect-foresight oracle bounds on park data. Chapter 8 composes the per-stage rules into the combined strategy, reports combined revenue and tail-risk performance, and tests robustness under realistic execution conditions. Chapter 9 interprets the main findings. Chapter 10 summarises the findings and outlines directions for future work.

Chapter 2

Market Background and Problem Setting

This chapter describes the institutional setting that defines the producer's decision problem.

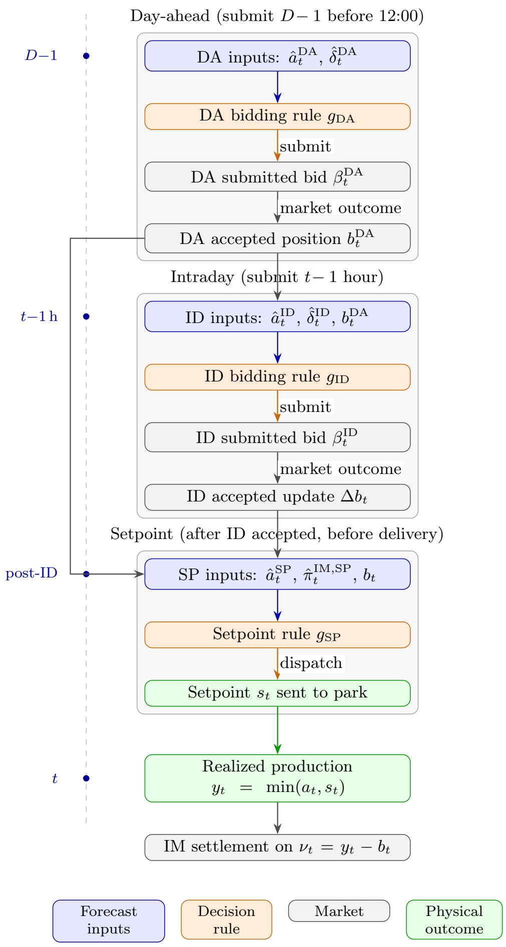

Figure 2.1: Information flow and decision timeline in the three stages. In the settlement box, a_t is the realised available production (the unhatted, stage-free counterpart of the stage forecasts $\hat{a}_t^{\mathrm{DA}}$, $\hat{a}_t^{\mathrm{ID}}$, $\hat{a}_t^{\mathrm{SP}}$) and s_t the setpoint cap, so realised output is $y_t = \min(a_t, s_t)$.

Figure 2.1 shows the general three-stage structure. Under the price-taker assumption (Section 2.5), submitted bids are accepted at their stated volume, so the thesis works directly with accepted positions: $\beta_t^{\mathrm{DA}} = b_t^{\mathrm{DA}}$ and $\beta_t^{\mathrm{ID}} = \Delta b_t^{\mathrm{ID}}$ throughout. The rule functions g_{DA}, g_{ID}, g_{SP} map the available information at each stage into a bid or setpoint; specifying these rules is the central task of the thesis.

2.1 Day-Ahead Market

Physical energy for delivery on day D is first traded in the day-ahead market on day $D-1$. Producers submit supply bids before 12:00 on day $D-1$, and the market clears to a zonal day-ahead price π_t^{DA} for each settlement interval.

Since 1 October 2025, the market time unit (MTU) in the DA market is 15 minutes, producing 96 clearing prices per bidding zone per day. This is also the settlement interval used throughout the thesis. Sweden's four bidding zones produce zone-specific clearing prices when transmission is constrained (Nord Pool AS 2025a).

2.2 Intraday Market

The Nordic intraday market consists of two parallel mechanisms: continuous trading and intraday auctions. This thesis uses only the continuous market; intraday auctions are not considered.

Continuous trading operates on a pan-European order book (the Single Intraday Coupling, SIDC) where orders are matched on a first-come-first-served basis (Nord Pool AS 2025b). The order book opens at $D-1$, 15:00, once the DA results are known. For all four 15-minute quarters within a given delivery hour, the continuous gate closes 60 minutes before the start of that hour. For example, continuous trading for all quarters of the 14:00 hour (14:00–14:15 through 14:45–15:00) closes at 13:00. Pricing is pay-as-bid: each trade clears at the price agreed between the two parties, so there is no single ID clearing price per MTU. Since the SIDC produces no official clearing price, the ID price π_t^{ID} used in the thesis is a modelling convention: it denotes the volume-weighted average price of all trades executed for that interval. The implications of this convention for counterfactual ID analysis are discussed in Section 2.5.

2.3 Imbalance Settlement

2.3.1 Settlement Mechanism

The Nordic imbalance settlement, operated by eSett, uses a single-price, single-position model (eSett Oy 2025). A single imbalance price π_t^{IM} applies per Market Balance Area (MBA) per 15-minute Imbalance Settlement Period (ISP), regardless of whether the balance responsible party is long or short. In Sweden, each MBA corresponds to a bidding zone (SE1–SE4). Unlike two-price settlement systems, there is no decomposition into helpful and harmful imbalance with different prices; a BRP's net imbalance is settled at a single price.

The imbalance price is derived from the mFRR Energy Activation Market (EAM) and depends on the dominating regulation direction in the MBA for that ISP:

- **Up-regulation ISP** (the system is short, net upward mFRR activated): $\pi_t^{\mathrm{IM}} = \pi_t^{\mathrm{ACT},\uparrow} \geq \pi_t^{\mathrm{DA}}$, where $\pi_t^{\mathrm{ACT},\uparrow}$ is the up-regulation mFRR activation price (ACT).
- **Down-regulation ISP** (the system is long, net downward mFRR activated): $\pi_t^{\mathrm{IM}} = \pi_t^{\mathrm{ACT},\downarrow} \leq \pi_t^{\mathrm{DA}}$.
- **No dominating direction** (no mFRR activated, or equal volumes in both directions): $\pi_t^{\mathrm{IM}} = \pi_t^{\mathrm{DA}}$. Deviations are settled at the DA price and produce no imbalance cost beyond the eSett fee.

An important country-specific rule applies: in Sweden and Norway, only the mFRR energy price is price-setting for the imbalance price; automatic frequency restoration reserve (aFRR) does not enter the formula (eSett Handbook §7.1.1).

In addition to the imbalance price, eSett levies several fees on BRPs. The most relevant is the *imbalance fee*: a symmetric charge of 1.15 EUR/MWh on absolute imbalance $|\nu_t|$, where ν_t is the signed deviation between delivered volume and the committed schedule (formalised in Equation (3.1)). It is applied regardless of whether the BRP is long or short. This means imbalance is never entirely free. Even in no-regulation ISPs where $\pi_t^{\mathrm{IM}} = \pi_t^{\mathrm{DA}}$, a BRP incurs a cost proportional to the size of its deviation. All fees are small relative to typical imbalance prices; their combined effect is captured by the fee term F_t in the revenue equation introduced in Chapter 3.

2.3.2 Imbalance Price Dynamics

Figure 2.2 compares the price distributions of the three markets using the available multi-year market-price history for SE1 and SE3. The means of the three markets are nearly indistinguishable in each zone, so settling at the imbalance price yields, on average, roughly the same revenue as selling on the DA market. The imbalance distribution is, however, substantially wider. Its standard deviation is more than twice that of the DA price in both zones.

In the tails, the imbalance price diverges sharply from the DA price; the producer's incentive to trade on the DA market is therefore about locking in a price, not capturing a premium. See Figure 2.3.

The economic environment facing wind producers in Sweden has changed materially since the early 2010s. Two structural developments account for most of the increase in imbalance price volatility. The first is the rapid expansion of variable renewable energy. Swedish installed wind capacity grew from roughly 6 GW in 2015 to over 16 GW by end-2023. As wind penetration rises, aggregate forecast errors increase in both frequency and magnitude, placing greater strain on the balancing market. Klyve et al. (2023) show that growing imbalance price variability, rather than larger physical deviations, is the primary driver of rising imbalance

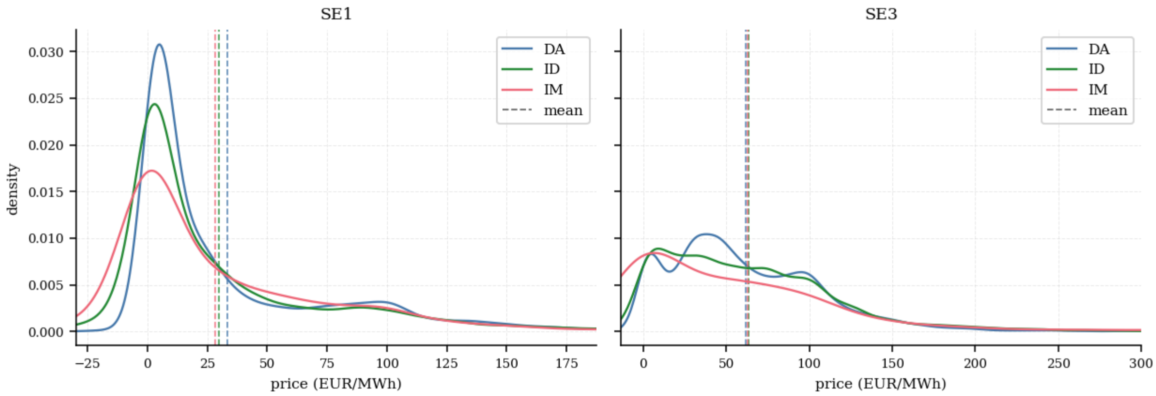

Figure 2.2: Kernel density estimates of day-ahead (π_t^{DA}), intraday (π_t^{ID}), and imbalance (π_t^{IM}) prices for SE1 and SE3 (multi-year history); dashed lines mark means.

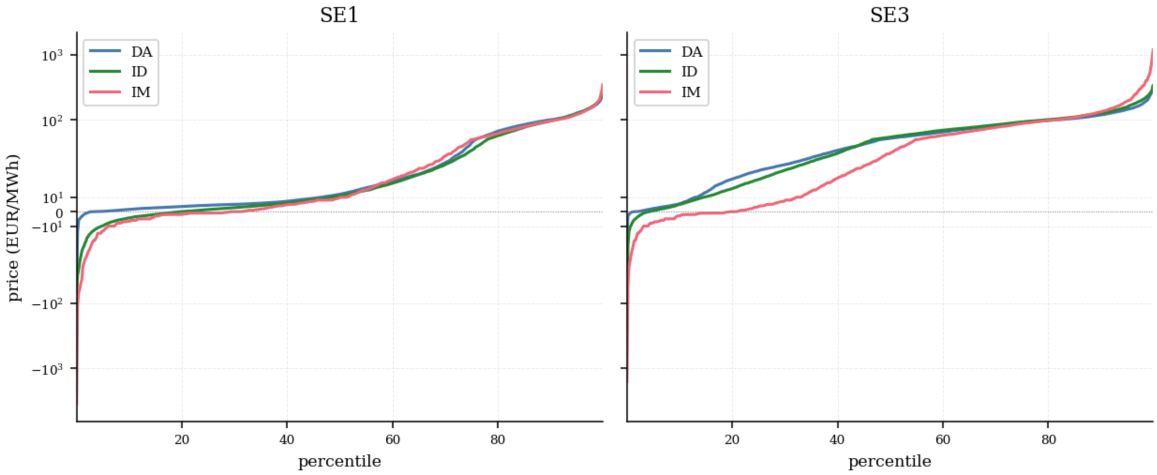

Figure 2.3: Percentile comparison of day-ahead, intraday, and imbalance prices for SE1 and SE3 (0.1 % to 99.9 %).

settlement costs for Nordic variable renewable energy (VRE) producers.

The second is a pair of market design changes implemented in Sweden in 2023–2025. Shorter imbalance settlement windows amplify the financial consequence of forecast errors: deviations that would have partially netted out within an hourly period are now settled individually each quarter-hour (eSett Oy 2025). On 4 March 2025, the Nordic transmission system operators (TSOs) activated the automated mFRR Energy Activation Market (mFRR EAM), replacing manual bid selection by TSO operators with an algorithm-based Activation Optimisation Function (Nordic TSOs 2025). The combination produced immediate and extreme price outcomes: imbalance prices in individual 15-minute periods exceeded EUR 2,000/MWh within the first week of operation. Nordenergi's post-implementation review explicitly identifies the mFRR EAM go-live as producing "unprecedented volatility in balancing prices" (Nordenergi 2025).

Figure 2.4 shows the imbalance price for both bidding zones throughout 2023–2026. The vertical dotted line marks the mFRR EAM activation on 4 March 2025. Rare price spikes cluster visibly after this date, with magnitudes that dwarf the DA price range in both zones. The faint yellow trace shows the DA price for reference. It stays within a narrow band while the imbalance price oscillates wildly around it.

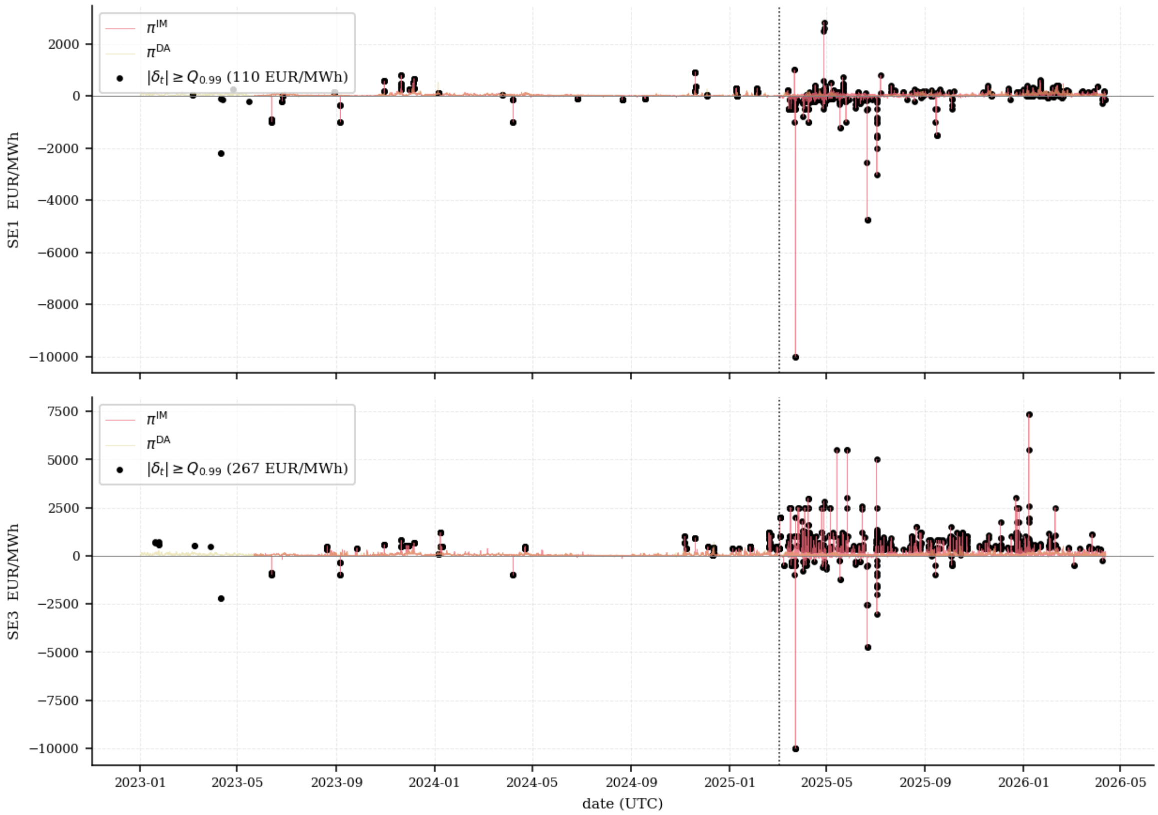

Figure 2.4: Imbalance price π_t^{IM} (SE1 and SE3, 2023–2026) with day-ahead price π_t^{DA} (faint yellow). Black dots mark the top 1 % of spread-magnitude intervals; the dotted line marks mFRR EAM activation on 4 March 2025.

2.4 Wind Production and Setpoint Control

Wind power production is uncertain at the time trading decisions must be made. Let a_t denote the available production in interval t, that is, the physical output that could be produced under the prevailing wind and equipment conditions. This is a realised quantity, observed only after delivery. At each decision stage, the producer works with a forecast of available production: $\hat{a}_t^{\mathrm{DA}}$ at the DA stage, $\hat{a}_t^{\mathrm{ID}}$ at the ID stage, and $\hat{a}_t^{\mathrm{SP}}$ at the setpoint stage. Without curtailment, realised production equals available production. The notation is therefore intentionally asymmetric: a_t is the *pre-curtailment* feasible output, while y_t is the *post-curtailment* realised output that is actually delivered and settled. When a setpoint is active, the realised production is

$$y_t = \min(a_t, s_t), \tag{2.1}$$

so y_t should not be interpreted as a direct observation of a_t in curtailed intervals.

Setpoint control is treated as a *third-stage economic decision*, on equal footing with the DA and ID decisions. Most stochastic bidding formulations treat realised production as exogenous (Pinson, Chevallier, and Kariniotakis 2007; Zugno, Morales, et al. 2013); by contrast, this thesis includes the setpoint as a recourse variable. By choosing s_t after the committed schedule $b_t^{\mathrm{DA}} + \Delta b_t^{\mathrm{ID}}$ is known and with the tightest available forecast, the producer can reduce imbalance exposure through curtailment when the expected imbalance price makes additional production unprofitable. The imbalance price π_t^{IM} is unobservable until $D+2$ to $D+13$, when eSett completes settlement, so it must be treated as a forecast input at every decision stage.

2.5 Thesis-Wide Modelling Assumptions

The empirical replay, benchmark models, and stochastic bidding models in the thesis share a small set of modelling assumptions. These assumptions define the scope of interpretation for the results. More specialised assumptions introduced later are *model-specific* and should be read in addition to the thesis-wide assumptions listed here.

1. **Price-taker behaviour.** Each park is treated as too small to affect the DA clearing price or the imbalance price. For the intraday market, trades are modelled at a volume-weighted representative price at the full desired volume, abstracting from price impact, order-book depth, and partial fills.

2. **Zero marginal cost of production.** Wind production is treated as having zero variable cost. The producer therefore offers every MWh available to the market, and the bidding problem reduces to allocation across the DA, ID, and imbalance markets. Revenue depends on the price *spreads* between markets rather than on absolute price levels, which is what allows Chapter 3 to reduce each stage's optimal bid to an expected-spread estimate.

3. **Intraday price observability.** The spread signal $\hat{\delta}_t^{\mathrm{ID}}$ is formed from π_t^{ID}, the full-MTU volume-weighted average intraday price, which is only known after gate closure. In practice, the imbalance-price forecast arrives before gate closure, so the most recent observable ID price at forecast arrival time is the last-printed trade.

4. **Idealised intraday execution.** Counterfactual intraday adjustments are assumed to execute at π_t^{ID} at the full desired volume. The SIDC produces no single official clearing price per MTU; this VWAP proxy abstracts from bid–ask spreads, order-book depth, and partial fills.

5. **Setpoint abstraction.** Setpoint control is modelled at settlement-interval level as a production cap, as in Equation (2.1). The thesis does not model turbine-level control, ramping limits, startup effects, or other detailed plant-operational dynamics.

The sensitivity of results to assumptions 1, 3, and 4 is quantified in Section 8.2.3.

Chapter 3

Theoretical Framework

3.1 Notation and Revenue Model

For each settlement interval t, the DA position is b_t^{DA}, the ID adjustment is Δb_t^{ID}, and the final schedule is $b_t = b_t^{\mathrm{DA}} + \Delta b_t^{\mathrm{ID}}$. Realised production is denoted by y_t, available production by a_t, and the setpoint cap by s_t. The signed imbalance is

$$\nu_t = y_t - b_t. \tag{3.1}$$

Forecasts target available production a_t; settlement and replay equations use realised production y_t (Section 2.4).

Under the price-taker assumption, interval revenue is linear in the scheduled volumes:

$$R_t = \pi_t^{\mathrm{DA}}\, b_t^{\mathrm{DA}} + \pi_t^{\mathrm{ID}}\, \Delta b_t^{\mathrm{ID}} + \pi_t^{\mathrm{IM}}\, \nu_t - F_t. \tag{3.2}$$

Here $F_t = f^{\mathrm{imb}}\,|\nu_t|$ is the eSett imbalance fee, levied on the absolute imbalance volume at the symmetric per-MWh rate f^{imb} (Section 3.4). Substituting $\nu_t = y_t - b_t^{\mathrm{DA}} - \Delta b_t^{\mathrm{ID}}$ yields the equivalent spread form

$$R_t = \pi_t^{\mathrm{IM}}\, y_t + \delta_t^{\mathrm{DA}}\, b_t^{\mathrm{DA}} + \delta_t^{\mathrm{ID}}\, \Delta b_t^{\mathrm{ID}} - F_t, \tag{3.3}$$

where $\delta_t^{\mathrm{DA}} = \pi_t^{\mathrm{DA}} - \pi_t^{\mathrm{IM}}$ and $\delta_t^{\mathrm{ID}} = \pi_t^{\mathrm{ID}} - \pi_t^{\mathrm{IM}}$. This representation separates the market-dependent baseline $\pi_t^{\mathrm{IM}}\, y_t$, identical across strategies on the same historical path, from the two spread terms the producer controls. A positive δ_t^{DA} means the DA price beat the imbalance price, so committing more volume in the DA market was profitable in hindsight; the bidding problem therefore reduces to forecasting the sign and magnitude of δ_t^{DA} and δ_t^{ID} before the respective gates close. The setpoint s_t enters through realised production: when $s_t < a_t$ the cap binds and $y_t = s_t$, changing the baseline term $\pi_t^{\mathrm{IM}}\, y_t$; unlike DA and ID, setpoint value derives directly from the imbalance price level (Section 3.4).

A natural question is whether the producer should oversell in the DA market and buy the excess back at ID, banking the difference between the two cleared prices. That round trip leaves the net committed position untouched and so changes neither realised production nor the imbalance fee. Its only payoff is the DA-to-ID price difference, a bet unrelated to the production problem studied here. Acting on it would require forecasting the intraday price at the DA gate, but the thesis uses forecasts of the imbalance price there, not the intraday price (Chapter 5). The thesis therefore takes the net committed position at each gate as the decision and does not pursue cross-market price arbitrage.

The objective is defined over daily revenue rather than per-interval revenue, so let the evaluation window contain D trading days and let $\mathcal{T}_d$ denote the set of settlement intervals belonging to day d. Daily and average daily revenue are

$$\bar{R}_d = \sum_{t \in \mathcal{T}_d} R_t, \qquad \bar{R} = \frac{1}{D} \sum_{d=1}^{D} \bar{R}_d. \tag{3.4}$$

3.2 Objective and Daily Risk Normalisation

The canonical objective is to maximise expected average daily revenue minus a penalty on adverse tails:

$$\max \ \mathbb{E}[\bar{R}] - \lambda \, \mathrm{CVaR}_\alpha(-\bar{R}_d), \tag{3.5}$$

with $\lambda \geq 0$ the risk-aversion parameter and $\alpha = 0.9$. The loss variable is $-\bar{R}_d$, so $\mathrm{CVaR}_\alpha(-\bar{R}_d)$ is the expected loss in the worst 10 % of trading days (Figure 3.1). Using *daily* revenue makes both the expectation and the risk penalty interpretable. This mean–CVaR objective follows (Artzner et al. 1999; Rockafellar and Uryasev 2000) and penalises the tail of the *outcome* distribution; the distinct problem of acting on an uncertain *estimate* is *parameter* (estimation) risk, which Bannör and Scherer (2013) capture by applying a convex risk measure to the parameter distribution; here it is handled by the signal-quality weighting of Section 3.5.

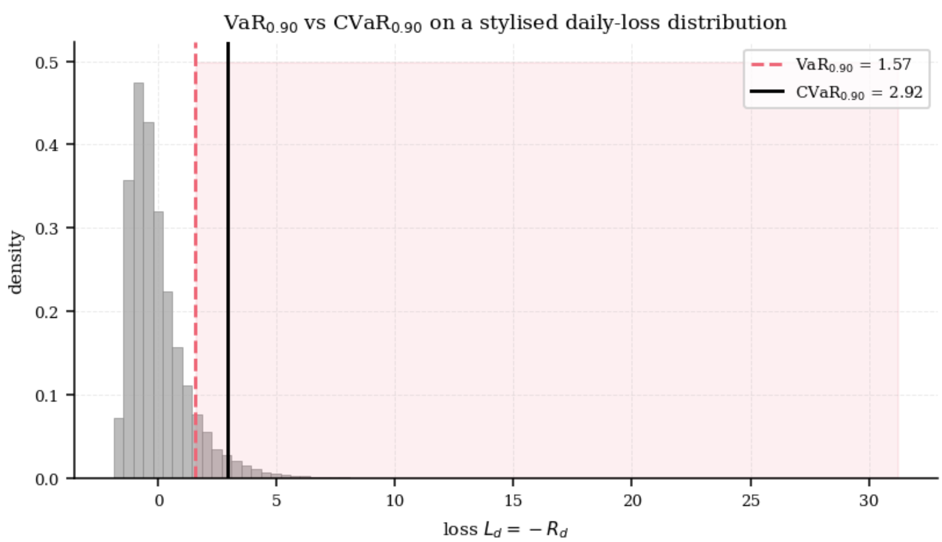

Figure 3.1: $\mathrm{VaR}_{0.9}$ and $\mathrm{CVaR}_{0.9}$ for a stylised daily loss distribution $L_d = -R_d$.

Empirical estimator. Over an evaluation window of D days with daily revenues sorted worst-to-best $\bar{R}_d^{(1)} \leq \cdots \leq \bar{R}_d^{(D)}$, the sample counterpart of (3.5)'s risk term is

$$\widehat{\mathrm{CVaR}}_{0.9}(-\bar{R}_d) = -\frac{1}{k}\sum_{i=1}^{k} \bar{R}_d^{(i)}, \qquad k = \lceil 0.1D \rceil. \tag{3.6}$$

This is the estimator used in all empirical comparisons in Chapter 7.

3.3 The Quantile Bid Framework

All bidding strategies in the thesis can be expressed within a single framework: the bid is a quantile of the conditional production distribution,

$$b_t = F^{-1}_{a_t|\mathcal{F}}(\tau), \tag{3.7}$$

where $\mathcal{F}$ denotes the information set (sigma-algebra of observable events) available at the relevant decision stage — $\mathcal{F}^{\mathrm{DA}}$, $\mathcal{F}^{\mathrm{ID}}$, or $\mathcal{F}^{\mathrm{SP}}$ for the DA, ID, and SP stages respectively — $F^{-1}_{a_t|\mathcal{F}}$ is the corresponding conditional quantile function of available production a_t, and $\tau \in [0, 1]$ is the quantile level. The framework decomposes every bid into two choices: the quantile function determines the shape of the production profile, while τ determines the *level* relative to the median forecast. Strategies differ in how F^{-1} is estimated — the subject of Chapter 6 — and how τ is set — the subject of Chapter 7.

Bid rules that tilt τ away from the median are truncated to the admissible quantile band $[\tau_{\mathrm{lo}}, \tau_{\mathrm{hi}}]$, parameterised by the *band half-width* $w \in (0, 0.5]$ as $\tau_{\mathrm{lo}} = 0.5 - w$ and $\tau_{\mathrm{hi}} = 0.5 + w$. The half-width w measures the distance from the median to each band edge: $w = 0.10$ gives a band $[Q_{0.40}, Q_{0.60}]$, the central 20 % of the production fan, while $w = 0.5$ admits the full span. A conservative band prevents the strategy from bidding on the far tails of the scenario fan and limits loss from noisy spread-sign misclassification while still allowing meaningful tilts above or below the median. The half-width w is a free parameter calibrated by walk-forward optimisation of the mean–CVaR objective; the empirical sensitivity to this choice is evaluated jointly with the SNR threshold z in Section 7.3.

3.4 Risk-Neutral Optimal Decisions under Single-Price Settlement

Under single-price settlement, the imbalance fee $f^{\mathrm{imb}} = 1.15$ EUR/MWh is symmetric in the deviation sign. The DA and ID decisions are then symmetric special cases of the asymmetric dual-price newsvendor solved in Pinson, Chevallier, and Kariniotakis (2007) and revisited in Zugno, Jónsson, and Pinson (2013); the SP decision shares the same three-region structure but is derived separately because realised production $y_t = \min(a_t, s_t)$ depends on the decision. Under single-price settlement, the three decisions respond to complementary price signals: DA and ID to the spreads δ_t^{DA} and δ_t^{ID}, and the setpoint to π_t^{IM} itself — when $\pi_t^{\mathrm{IM}} < 0$ curtailment avoids cost, when $\pi_t^{\mathrm{IM}} > 0$ curtailment sacrifices revenue. Throughout this section, the risk-aversion parameter is set to $\lambda = 0$, giving a risk-neutral objective; the extension to $\lambda > 0$ is treated in Section 3.5.

At each stage, the producer chooses a volume within a bounded feasible set; the revenue contribution is linear in that volume, and the only source of curvature is the eSett imbalance fee for DA and ID and additionally the curtailment kink $y_t = \min(a_t, s_t)$ for SP. This section derives the optimal decision *in isolation* at each stage. The isolation is a deliberate analytical choice: the spread-form revenue is additively separable across $\delta_t^{\mathrm{DA}} b_t^{\mathrm{DA}}$ and $\delta_t^{\mathrm{ID}} \Delta b_t^{\mathrm{ID}}$, so each stage's marginal value depends only on its own spread when the others are held fixed. The propositions characterise the optimal bid given a reliable spread forecast; Chapter 5 assesses whether the available forecasts deliver reliable information, and Chapter 7 tests whether acting on them improves the revenue–risk objective.

Proposition 3.1 (Optimal DA bid). *Under $\mathcal{F}^{\mathrm{DA}}$, the isolated optimal DA bid is*

$$\tau_t^{\mathrm{DA},0} = \frac{1}{2} + \frac{\mathbb{E}[\delta_t^{\mathrm{DA}} \mid \mathcal{F}^{\mathrm{DA}}]}{2f^{\mathrm{imb}}}, \qquad b_t^{\mathrm{DA}} = F^{-1}_{a_t \mid \mathcal{F}^{\mathrm{DA}}}\Big(\mathrm{clip}(\tau_t^{\mathrm{DA},0},\, \tau_{\mathrm{lo}},\, \tau_{\mathrm{hi}})\Big). \tag{3.8}$$

Proof. With $\Delta b_t^{\mathrm{ID}} = 0$, the b_t^{DA}-dependent expected contribution to (3.3) is

$$\mathbb{E}[\delta_t^{\mathrm{DA}} \mid \mathcal{F}^{\mathrm{DA}}]\, b_t^{\mathrm{DA}} - f^{\mathrm{imb}}\, \mathbb{E}[|a_t - b_t^{\mathrm{DA}}| \mid \mathcal{F}^{\mathrm{DA}}].$$

Writing $|a_t - b_t^{\mathrm{DA}}| = (a_t - b_t^{\mathrm{DA}})^+ + (b_t^{\mathrm{DA}} - a_t)^+$ and collecting terms gives the newsvendor form of Pinson, Chevallier, and Kariniotakis (2007, eq. 16) with overage and underage costs

$$c_o = f^{\mathrm{imb}} - \mathbb{E}[\delta_t^{\mathrm{DA}} \mid \mathcal{F}^{\mathrm{DA}}], \qquad c_u = f^{\mathrm{imb}} + \mathbb{E}[\delta_t^{\mathrm{DA}} \mid \mathcal{F}^{\mathrm{DA}}].$$

The difference from the dual-price case is that both costs are driven by the same fee f^{imb} rather than separate over- and under-production prices. When $|\mathbb{E}[\delta_t^{\mathrm{DA}} \mid \mathcal{F}^{\mathrm{DA}}]| \leq f^{\mathrm{imb}}$ both costs are non-negative and the critical-fractile optimum is

$$\frac{c_u}{c_o + c_u} = \frac{1}{2} + \frac{\mathbb{E}[\delta_t^{\mathrm{DA}} \mid \mathcal{F}^{\mathrm{DA}}]}{2f^{\mathrm{imb}}} = \tau_t^{\mathrm{DA},0}.$$

When $|\mathbb{E}[\delta_t^{\mathrm{DA}} \mid \mathcal{F}^{\mathrm{DA}}]| > f^{\mathrm{imb}}$ one cost is negative and the objective is monotone in b_t^{DA}, so the maximum on the admissible band is at the corresponding edge, which is what clipping $\tau_t^{\mathrm{DA},0}$ to $[\tau_{\mathrm{lo}}, \tau_{\mathrm{hi}}]$ produces. □

The optimal DA bid therefore shifts the median forecast upward when the DA price is expected to beat the imbalance price, and downward when the imbalance price is expected to beat the DA price, saturating at the quantile-band edge when the spread signal exceeds f^{imb}.

Remark 3.1. *Because $f^{\mathrm{imb}} = 1.15$ EUR/MWh is small relative to typical DA–IM spreads, the interior newsvendor regime $|\mathbb{E}[\delta_t^{\mathrm{DA}} \mid \mathcal{F}^{\mathrm{DA}}]| \leq f^{\mathrm{imb}}$ is structurally rare: the isolated optimum lies at a band edge for nearly all regulated intervals.*

Proposition 3.2 (Optimal ID adjustment). *Given b_t^{DA}, the isolated optimal ID adjustment* *under $\mathcal{F}^{\mathrm{ID}}$ is*

$$\tau_t^{\mathrm{ID},0} = \frac{1}{2} + \frac{\mathbb{E}[\delta_t^{\mathrm{ID}} \mid \mathcal{F}^{\mathrm{ID}}]}{2f^{\mathrm{imb}}}, \qquad b_t^{\mathrm{DA}} + \Delta b_t^{\mathrm{ID}} = F^{-1}_{a_t|\mathcal{F}^{\mathrm{ID}}}\Big(\mathrm{clip}(\tau_t^{\mathrm{ID},0},\, \tau_{\mathrm{lo}},\, \tau_{\mathrm{hi}})\Big). \tag{3.9}$$

Proof. With b_t^{DA} fixed, reparametrising $x = b_t^{\mathrm{DA}} + \Delta b_t^{\mathrm{ID}}$ casts the only Δb_t^{ID}-dependent contribution as the same newsvendor as in the proof of Proposition 3.1 with $\delta_t^{\mathrm{DA}} \mapsto \delta_t^{\mathrm{ID}}$ and $\mathcal{F}^{\mathrm{DA}} \mapsto \mathcal{F}^{\mathrm{ID}}$; (3.9) follows. □

Proposition 3.3 (Optimal setpoint)**.** *Assume π_t^{IM} and a_t are conditionally independent given $\mathcal{F}^{\mathrm{SP}}$. Given b_t, the isolated optimal setpoint under $\mathcal{F}^{\mathrm{SP}}$ is the three-region rule*

$$s_t = \begin{cases} P^{\max} & \text{if } \mathbb{E}[\pi_t^{\mathrm{IM}} \mid \mathcal{F}^{\mathrm{SP}}] > +f^{\mathrm{imb}}, \\ b_t & \text{if } |\mathbb{E}[\pi_t^{\mathrm{IM}} \mid \mathcal{F}^{\mathrm{SP}}]| \le f^{\mathrm{imb}}, \\ 0 & \text{if } \mathbb{E}[\pi_t^{\mathrm{IM}} \mid \mathcal{F}^{\mathrm{SP}}] < -f^{\mathrm{imb}}. \end{cases} \tag{3.10}$$

Remark 3.2. *The conditional-independence hypothesis $\pi_t^{\mathrm{IM}} \perp a_t \mid \mathcal{F}^{\mathrm{SP}}$ is what allows the marginal effect to factor into $\mathbb{E}[\pi_t^{\mathrm{IM}} \mid \mathcal{F}^{\mathrm{SP}}] \cdot \Pr(a_t > s_t \mid \mathcal{F}^{\mathrm{SP}})$ and yield the clean threshold rule. The hypothesis is tested in Appendix C.9: a rule free to exploit the joint structure of the scenario fan is empirically dominated by the three-region rule at every park, so the assumption is a robust inductive bias rather than a costly simplification.*

Proof. With b_t fixed, the s_t-dependent contribution is $\pi_t^{\mathrm{IM}} y_t - f^{\mathrm{imb}}|y_t - b_t|$ with $y_t = \min(a_t, s_t)$. The cap binds only when $a_t > s_t$, so the setpoint has no effect otherwise. The fee term shrinks as s_t moves toward b_t and grows beyond it. Under $\pi_t^{\mathrm{IM}} \perp a_t \mid \mathcal{F}^{\mathrm{SP}}$ with $\bar{\pi} := \mathbb{E}[\pi_t^{\mathrm{IM}} \mid \mathcal{F}^{\mathrm{SP}}]$, the expected slope is

$$\frac{d}{ds_t}\, \mathbb{E}[\pi_t^{\mathrm{IM}} y_t - f^{\mathrm{imb}}|y_t - b_t| \mid \mathcal{F}^{\mathrm{SP}}] = \begin{cases} (\bar{\pi} + f^{\mathrm{imb}})\, \Pr(a_t > s_t \mid \mathcal{F}^{\mathrm{SP}}), & 0 \le s_t < b_t, \\ (\bar{\pi} - f^{\mathrm{imb}})\, \Pr(a_t > s_t \mid \mathcal{F}^{\mathrm{SP}}), & b_t < s_t \le P^{\max}. \end{cases}$$

Since $\Pr(a_t > s_t \mid \mathcal{F}^{\mathrm{SP}}) \ge 0$, the sign of each slope is determined by the bracketed factor, and the maximum on $[0, P^{\max}]$ lies in $\{0, b_t, P^{\max}\}$:

- if $\bar{\pi} > +f^{\mathrm{imb}}$, both slopes are positive: $s_t^\star = P^{\max}$;
- if $\bar{\pi} < -f^{\mathrm{imb}}$, both slopes are negative: $s_t^\star = 0$;
- if $|\bar{\pi}| \le f^{\mathrm{imb}}$, slope ≥ 0 on $[0, b_t]$ and ≤ 0 on $[b_t, P^{\max}]$: $s_t^\star = b_t$.

This is (3.10). □

When the expected imbalance price is positive and large, producing more is profitable: set the turbine unconstrained. When it is negative and large, producing at all creates a loss that the fee makes worse: curtail to zero. In the neutral band, the typical case on quiet balancing days, match the committed schedule exactly to minimise the fee on any residual deviation.

All three propositions reduce to the same operational instruction: estimate the relevant expected price signal and decide whether it clears the fee threshold $\pm f^{\text{imb}}$. The signals differ (δ_t^{DA}, δ_t^{ID}, π_t^{IM}), the feasible sets differ (quantile band vs. physical envelope), but the decision rule is structurally identical. Table 3.1 lists the three resulting threshold problems side by side.

Table 3.1: Decision variables, revenue drivers, and signal rules at each bidding stage.

Stage	Decision variable	Revenue driver	Signal rule
DA	$b_t^{\text{DA}} \in [F^{-1}_{a_t\mid\mathcal{F}^{\text{DA}}}(\tau_{\text{lo}}),\, F^{-1}_{a_t\mid\mathcal{F}^{\text{DA}}}(\tau_{\text{hi}})]$	$\delta_t^{\text{DA}} \cdot b_t^{\text{DA}}$	$\mathbb{E}[\delta_t^{\text{DA}} \mid \mathcal{F}^{\text{DA}}]$ vs. $\pm f^{\text{imb}}$
ID	$b_t \in [F^{-1}_{a_t\mid\mathcal{F}^{\text{ID}}}(\tau_{\text{lo}}),\, F^{-1}_{a_t\mid\mathcal{F}^{\text{ID}}}(\tau_{\text{hi}})]$	$\delta_t^{\text{ID}} \cdot \Delta b_t^{\text{ID}}$	$\mathbb{E}[\delta_t^{\text{ID}} \mid \mathcal{F}^{\text{ID}}]$ vs. $\pm f^{\text{imb}}$
SP	$s_t \in [0,\, P^{\max}]$	$\pi_t^{\text{IM}} \cdot y_t$	$\mathbb{E}[\pi_t^{\text{IM}} \mid \mathcal{F}^{\text{SP}}]$ vs. $\pm f^{\text{imb}}$

3.5 Risk-Averse Extension

Propositions 3.1–3.3 give the risk-neutral optimal bid under perfect knowledge of the conditional spread mean $\mathbb{E}[\delta \mid \mathcal{F}]$. In practice, that conditional mean is estimated and may be noisy, and imbalance prices carry heavy tails that the risk-neutral analysis ignores. The mean–CVaR objective (3.5) addresses both concerns; this section introduces two implementations that target different sources of risk. The first takes the scenario fan as truth and penalises tail outcomes in the scenario distribution. The second takes the risk-neutral direction as truth and shrinks commitment by signal credibility, discounting how much the conditional spread mean is itself an estimate worth acting on.

Direct scenario optimisation. Using the Rockafellar–Uryasev representation (Rockafellar and Uryasev 2000) $\text{CVaR}_\alpha(L) = \min_\eta\{\eta + (1-\alpha)^{-1}\mathbb{E}[(L-\eta)_+]\}$, the scenario problem becomes a convex program: the *Sample-Average Approximation Linear Program* (SAA LP). It is written on the decision-controllable component of daily revenue $\bar{R}_d^{\text{ctrl}}$ — the bid-independent baseline $\pi_t^{\text{IM}}\, y_t$ removed for DA and ID, setpoint terms that change realised production retained.[1] On a scenario set $\{(\pi_t^{\text{DA},s}, \pi_t^{\text{ID},s}, \pi_t^{\text{IM},s}, a_t^s)\}_{s=1}^S$ the SAA LP is

$$\max_{\substack{b_t^{\text{DA}},\Delta b_t^{\text{ID}},s_t,\\ \eta,\, z\geq 0}} \quad \tfrac{1}{S}\sum_{s=1}^{S} \bar{R}_d^{s,\text{ctrl}} \;-\; \lambda\left[\eta + \tfrac{1}{(1-\alpha)S}\sum_{s=1}^{S} z_s\right]$$

$$\text{s.t.} \quad z_s \geq -\bar{R}_d^{s,\text{ctrl}} - \eta, \qquad F^{-1}_{a_t\mid\mathcal{F}}(\tau_{\text{lo}}) \leq b_t \leq F^{-1}_{a_t\mid\mathcal{F}}(\tau_{\text{hi}}). \tag{3.11}$$

Here η is the in-sample VaR_α of $-\bar{R}_d^{\text{ctrl}}$, each z_s is the per-scenario controllable loss above η, and the quantile band $[\tau_{\text{lo}}, \tau_{\text{hi}}] = [0.5 - w,\, 0.5 + w]$ caps exposure on every DA and ID bid.

Signal-quality-aware quantile family. The risk-neutral propositions deliver $\tau_t^\star = \frac{1}{2} + \hat{\delta}_t/(2f^{\text{imb}})$. A natural one-parameter generalisation attenuates the tilt by a signal-quality

[1] Removing the bid-independent baseline reduces sampling noise in the scenario objective without changing the maximiser; full realised revenue is restored for the final comparison in Chapter 7.

factor $\gamma_t \in [0, 1]$ and keeps the same band cap:

$$\tau_t = \mathrm{clip}\left(\frac{1}{2} + \gamma_t \frac{\hat{\delta}_t}{2 f^{\mathrm{imb}}}, \tau_{\mathrm{lo}}, \tau_{\mathrm{hi}}\right). \tag{3.12}$$

At $\gamma_t \equiv 1$, the rule commits to the full risk-neutral tilt; as $\gamma_t \to 0$, it collapses to the median forecast. Chapter 7 introduces specific forms of γ_t as candidates, including a per-interval signal-to-noise shrinkage of the spread forecast. The signal-quality factor is itself a form of risk control: acting on noise is risky, and shrinking commitment when the signal is unreliable attenuates that exposure directly. The band cap $[\tau_{\mathrm{lo}}, \tau_{\mathrm{hi}}]$ provides a hard backstop in both cases. In the terms of Section 3.2, the signal-quality factor controls *parameter* risk, the estimate's own uncertainty, as distinct from the outcome-tail risk penalised by the mean–CVaR objective.

3.6 Perfect-Forecast Counterfactuals

A natural diagnostic question, once a bidding rule is in place, is *how much of the remaining revenue gap is forecast noise the rule cannot remove?* A *perfect-forecast counterfactual* answers that question for one input at a time: replace the noisy estimate with the realised value at decision time, hold everything else equal, and record the mean–CVaR lift. The lift is the maximum value a better forecast of that input could deliver under the same rule; the gap between the lift and the deployed rule is the headroom a sharper forecast would unlock.

The deployed pipeline (assembled in Chapter 7 and run in Chapter 8) reads four uncertain inputs — the day-ahead production forecast, the intraday production forecast, the setpoint-stage imbalance-price forecast, and the intraday imbalance-spread forecast — and the counterfactuals are constructed input by input rather than aggregated. The reason is asymmetric: three of the four lifts are bounded by the problem, while the fourth is not.

The three bounded counterfactuals follow directly from the problem structure. The two production counterfactuals replace the forecast with realised availability a_t at the relevant gate; the resulting schedule incurs no imbalance, so the lift is capped by physical production. The setpoint counterfactual routes the Proposition 3.3 three-region rule through the realised π_t^{IM}; the lift is capped because the rule has only three discrete actions. Each lift is therefore a property of the input alone, independent of any tuning choice.

The intraday spread counterfactual is different in nature: with perfect foresight of the realised spread, the best response is to bid the largest position the band permits in its direction, so the lift mechanically scales with whatever band cap is allowed. Quoting it against an arbitrary cap reports the cap rather than the value of the information. The honest comparison instead inherits the band from the rule the counterfactual is being measured against: the bid is placed at the production-fan quantile $0.5 \pm w_t$, where w_t is the same per-day half-width the adopted ID rule selected for that day. The resulting lift is the upper bound on what an ideal ID-spread forecast could deliver *at the deployed rule's own risk budget*, which is the quantity a real-world replacement forecast would compete for.

The four counterfactuals are reported alongside each stage's candidate selection in Tables 7.2, 7.3 and 7.6. They are diagnostics for stage-level headroom, not aggregated into a single capture ratio.

Chapter 4

Empirical Setup

4.1 Thesis Parks and Study Periods

The thesis studies three wind parks: Blabergsliden and Blisterliden, both located in price area SE1, and Varsvik in SE3. Figure 4.1 shows their positions within the Swedish bidding zone structure.

The parks differ not only in size but also in the local imbalance-price regime shaped by the Swedish zonal market structure. For that reason, results are reported park by park rather than pooled into a single aggregate estimate.

The evaluation periods differ across parks depending on data availability. A central design choice for the thesis is therefore to make all park-level comparisons on the valid historical window for that specific park, while explicitly documenting any exclusions caused by data quality or missing signals. Table 4.1 summarises the parks and their data periods.

Table 4.1: Thesis parks and data coverage periods.

Park	Price area	Start (UTC)	End (UTC)	Duration
Blabergsliden	SE1	2025-06-01	2026-03-13	∼9.5 months
Varsvik	SE3	2025-07-01	2026-03-13	∼8.5 months
Blisterliden	SE1	2025-11-01	2026-03-13	∼4.5 months

In addition to the park-specific evaluation data, the thesis uses a separate multi-year price dataset for SE1 and SE3. This dataset contains DA and imbalance settlement prices from January 2023 to April 2026 and is used only for market background diagnostics, including seasonal regime structure, imbalance-price tails, and volatility trends. It is not used to extend the park-level backtests: baseline replay, scenario generation, and strategy evaluation are always restricted to the valid park-specific windows in Table 4.1.

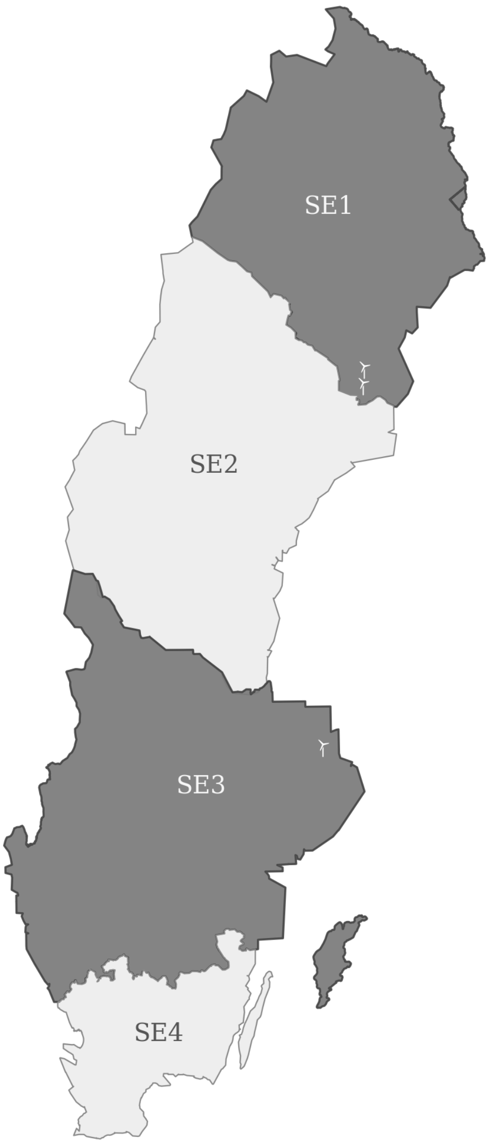

Figure 4.1: Swedish bidding zones and the thesis wind parks. Blabergsliden and Blisterliden are in SE1; Varsvik is in SE3.

4.2 Data Sources and Core Variables

Internal data platform. Realised production y_t, available production a_t, the historical DA position b_t^{DA}, and settlement prices $(\pi_t^{\mathrm{DA}}, \pi_t^{\mathrm{ID}}, \pi_t^{\mathrm{IM}})$ are sourced from the company's internal data platform at 15-minute resolution.

Day-ahead price forecasts. DA price forecasts are sourced from an external provider using the ECMWF *ec00* product, the operational medium-range forecast issued from the 00:00 UTC model run, published at approximately 06:00 CET on D-1, six hours before the Nordic gate closure at 12:00 CET. Forecasts are provided at hourly resolution and repeated across the four 15-minute intervals of each hour. In thesis notation: $\hat{\pi}_t^{\mathrm{DA}}$.

Optimeering imbalance spread forecasts. Optimeering is a Nordic power-market forecasting company whose platform provides calibrated probabilistic imbalance spread forecasts derived from publicly available TSO, exchange, and weather data. For this thesis, three stage-specific forecasts are used at 15-minute resolution for SE1 and SE3, each comprising a point estimate $\hat{\delta}_t$ and nine quantiles ($\tau \in \{0.01, 0.05, \ldots, 0.99\}$): a day-ahead batch (D-1, 11:55 CET), an intraday update available at the ID gate, and a setpoint product at approximately H-20 minutes lead time.

Two derived quantities are used throughout the rest of the thesis: the predictive fan width $W_t = Q^{\text{fc}}_{0.99} - Q^{\text{fc}}_{0.01}$, which measures forecast uncertainty, and the signal-to-noise ratio $\text{SNR}_t = |\hat{\delta}_t|/W_t$, which combines forecast magnitude with width into a single confidence measure.

Remark 4.1 (Optimeering sign convention)**.** *Optimeering defines the spread as* $\sigma^{\text{Opti}}_t = \pi^{\text{IM}}_t - \pi^{\text{DA}}_t$*, opposite to the thesis convention* $\delta^{\text{DA}}_t := \pi^{\text{DA}}_t - \pi^{\text{IM}}_t$*. The stored columns* `spread_{da, sp}_{quantile}` *retain the Optimeering sign. To recover an implied imbalance price level:*

$$\hat{\pi}^{\text{IM}}_t = \pi^{\text{DA}}_t + \hat{\sigma}^{\text{Opti}}_t. \tag{4.1}$$

4.3 Forecast and Data Pipeline Overview

Table 4.2 summarises every series used downstream, grouped into observed inputs, external forecasts, and in-house forecasts developed in this thesis (described in Chapter 5 and evaluated in Section 5.4). The external forecasts represent what a wind producer would have access to operationally; the in-house models are the thesis's own contributions.

Table 4.2: Data pipeline: observed series, external forecasts, and in-house forecasts.

Component	Type	Source	Resolution	Role
y_t, a_t	Observed	Internal platform	15 min	Target, error library
b^{DA}_t	Observed	Internal platform	15 min	Baseline replay
$\pi^{\text{DA}}_t, \pi^{\text{ID}}_t, \pi^{\text{IM}}_t$	Observed	Internal platform	15 min	Revenue, error library
$\hat{\pi}^{\text{DA}}_t$	External fc.	DA price fc. (ec00)	Hourly→15 min	DA price centre
$\hat{a}^{\text{DA}}_t, \hat{a}^{\text{ID,op}}_t$	External fc.	Park operator	15 min	DA / ID prod. baselines
$\hat{\delta}^{\text{DA}}_t, \hat{\delta}^{\text{ID}}_t, \hat{\delta}^{\text{SP}}_t$	External fc.	Optimeering	15 min	Stage-wise $\pi^{\text{IM}} - \pi^{\text{DA}}$
ID-Linear ($\hat{a}^{\text{ID,Lin}}_t$)	In-house fc.	This thesis	15 min	ID production centre
SP-Persistence ($\hat{a}^{\text{SP}}_t$)	In-house fc.	This thesis	15 min	SP production centre

Figure 4.2 places the components of Table 4.2 on a common time axis. The sources become available at different dates: zone prices have years of history, the DA price forecast begins in late 2024, and the Optimeering ID and SP spread fans only from 2025-03-19, just after the mFRR EAM go-live on 4 March 2025 (Section 2.3.2). Each park's evaluation window (coloured outlines) is therefore the intersection of the park's own series with these sources, minus a 30-day warm-up (lightly shaded) and capped at the 2026-03-13 16:00 UTC truncation. The figure also makes the residual data gaps visible; their handling is described in Section 4.4.

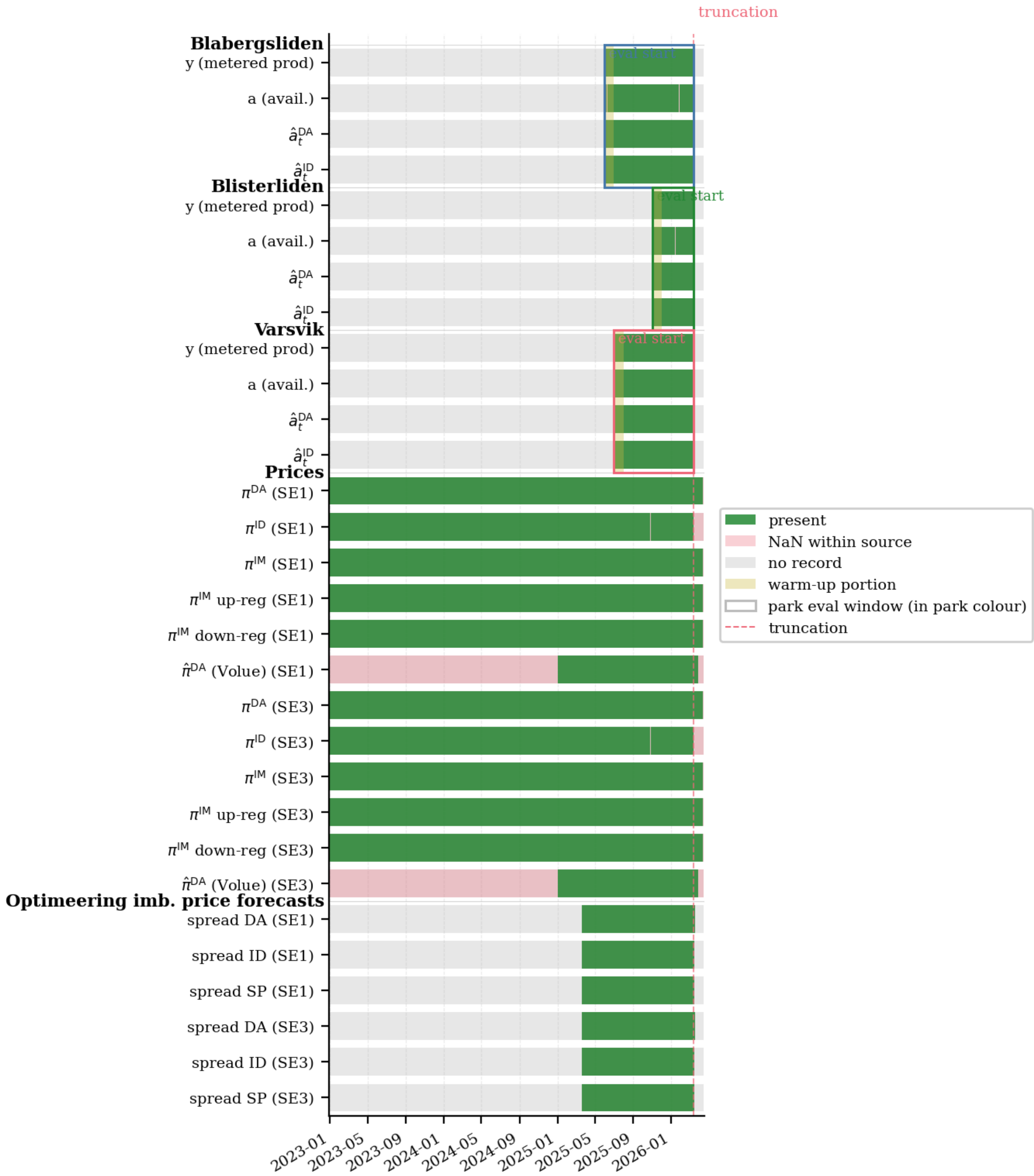

Figure 4.2: Variable completeness across the full dataset, grouped by source (parks, zone prices, Optimeering imbalance-spread forecasts). Green = present, faded red = NaN, light grey = no record. Each park's evaluation window is outlined in its canonical colour with the 30-day warm-up shaded; the dashed red line marks the 2026-03-13 16:00 UTC truncation.

4.4 Dataset Construction and Cleaning

The three sources are merged on UTC timestamps at the common 15-minute resolution. Rows of the merged 15-minute panel are not blanket-excluded: each pipeline stage drops only the rows where the columns it specifically requires are null. The revenue replay needs y_t, b_t^{DA}, Δb_t^{ID}, and the settlement prices; the scenario error libraries (the per-stage collections of historical forecast-error day vectors from which production scenarios are bootstrapped, Section 6.2) and the SP backtest additionally need the uncurtailed availability a_t along with the stage's forecast centre and Optimeering spread fan. The practical consequence is that rows with null a_t but otherwise complete data contribute to baseline and DA strategy evaluation, but are excluded from anything that requires a_t as the uncurtailed counterfactual.

Quality flagging. Availability a_t is additionally flagged where it exceeds the grid-connection $P^{\max}$, on static runs ($\geq$ 2 h of identical non-zero values), and on out-of-bounds prices. Flagged slots — 960, 404, and 660 at Blabergsliden, Blisterliden, and Varsvik — are dropped wherever a model is fitted or replayed, and days with more than 25 % flagged slots are dropped from the scenario error library.

All parks are truncated at 2026-03-13 16:00 UTC; subsequent data fall outside the evaluation window.

4.5 Expanding-Window Evaluation Protocol

All backtests in the thesis follow an expanding-window protocol, illustrated in Figure 4.3. This design ensures that every decision — scenario generation, forecast model retraining, parameter calibration, and bidding optimisation — uses only information that would have been available in real time.

For each evaluation day d^*, the eligible historical pool is

$$\mathcal{D}(d^*) = \big\{d : d < d^*,\ d \text{ complete}\big\}, \tag{4.2}$$

the set of all complete days strictly before d^*. As d^* advances, the pool grows monotonically; no data are ever discarded. A minimum of 30 complete days is required before the first scoring date; this warm-up ensures that day-level calibrators (the SP threshold rules, the band-width calibrator (Section 7.3), and the bootstrap error library used for scenario generation) have a non-degenerate history pool. The subset of each park's evaluation period following this initialisation phase is referred to throughout as the *post-warmup evaluation window*.

The expanding window governs three distinct processes:

1. **Forecast model retraining.** The in-house production model (ID-Linear) is retrained weekly using all data in $\mathcal{D}(d^*)$ up to the retraining date. Predictions from the retrained model are used for the following week.

2. **Error library construction.** The error library for scenario generation is drawn from $\mathcal{D}(d^*)$ at each evaluation date. Earlier evaluation dates draw from a smaller pool of historical forecast errors; later dates benefit from a richer pool.

3. **Strategy calibration.** Rule parameters and optimisation models are calibrated on $\mathcal{D}(d^*)$, then applied to day d^*.

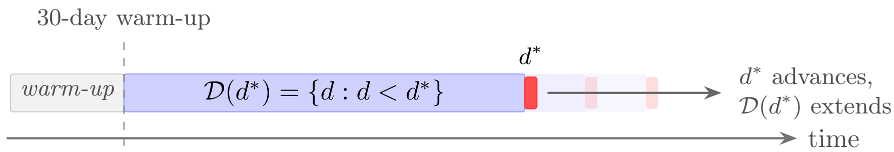

Figure 4.3: Expanding-window protocol used throughout the thesis. At each evaluation day d^*, scenarios are generated and parameters calibrated using only the eligible pool $\mathcal{D}(d^*) = \{d : d < d^*,\ d \text{ complete}\}$. The faint markers to the right indicate that d^* advances over time and the pool extends with it. The dashed line marks the 30-day minimum-history guard required before the first scoring date. No future information enters the pipeline at any stage. Best viewed in the digital version.

4.6 Diagnostic Metrics

4.6.1 Imbalance Metrics

Let $\mathcal{T}$ denote the set of evaluated intervals and $N = |\mathcal{T}|$ the count. The mean imbalance is

$$\overline{\nu} = \frac{1}{N} \sum_{t \in \mathcal{T}} \nu_t \tag{4.3}$$

and captures directional bias. The mean absolute imbalance is

$$\mathrm{MAI} = \frac{1}{N} \sum_{t \in \mathcal{T}} |\nu_t| \tag{4.4}$$

Capacity-normalised versions are defined as

$$\mathrm{nMAI} = \mathrm{MAI}/P^{\max} \tag{4.5}$$

expressing imbalance relative to installed capacity. These metrics summarise the operational consequences of a bidding strategy even before prices are considered.

4.6.2 Scenario and Forecast Evaluation Metrics

The operational metrics above evaluate the quality of a *decision* in physical terms. The metrics in this subsection evaluate the quality of the *scenarios* and probabilistic forecasts that feed into those decisions. A scenario set can only be trusted as input to the stochastic optimiser if it is well-calibrated, sufficiently sharp, and captures the relevant dependence structure. The metrics below formalise these properties and are reported in the scenario diagnostics of Section 6.5.

Marginal calibration and coverage For a set of N evaluation intervals with predictive quantile $\hat{y}_t^{(\tau)}$ and realisation y_t, the empirical coverage at level τ is

$$\mathrm{Cov}(\tau) = \frac{1}{N} \sum_{t=1}^{N} \mathbf{1}\left\{ y_t \le \hat{y}_t^{(\tau)} \right\}. \tag{4.6}$$

Ideal calibration requires $\mathrm{Cov}(\tau) \approx \tau$ for all $\tau \in (0, 1)$. When scenarios rather than explicit quantiles are available, the empirical quantiles of the scenario fan at each interval replace $\hat{y}_t^{(\tau)}$. A reliability diagram plots $\mathrm{Cov}(\tau)$ against τ and makes systematic over- or underdispersion visible at a glance.

Continuous ranked probability score The CRPS evaluates the full predictive distribution in a single number:

$$\mathrm{CRPS}(F_t, y_t) = \int_{-\infty}^{\infty} \big(F_t(x) - \mathbf{1}\{y_t \le x\}\big)^2 \, \mathrm{d}x \tag{4.7}$$

where F_t is the predictive CDF at interval t. The CRPS has the same unit as the predicted variable and can be decomposed into a reliability term and a sharpness term (Hersbach 2000), which makes it possible to diagnose whether a poor score is caused by miscalibration or by an overly wide distribution.

For an ensemble of S scenarios, the CRPS admits the energy-distance form

$$\mathrm{CRPS}(\{y_t^{(s)}\}_{s=1}^{S}, y_t) = \frac{1}{S} \sum_{s=1}^{S} \big|y_t^{(s)} - y_t\big| - \frac{1}{2S^2} \sum_{s=1}^{S} \sum_{r=1}^{S} \big|y_t^{(s)} - y_t^{(r)}\big| \tag{4.8}$$

where $y_t^{(s)}$ denotes the value of the forecast variable in scenario s (production or spread, as applicable).

The average CRPS over the evaluation window is

$$\overline{\mathrm{CRPS}} = \frac{1}{N} \sum_{t=1}^{N} \mathrm{CRPS}(F_t, y_t). \tag{4.9}$$

4.6.3 Forecast Accuracy Metrics

MAE captures average error magnitude; Pearson correlation captures the linear association between forecast and realisation. For a point forecast $\hat{y}_t$ and realisation y_t over N intervals, the mean absolute error is

$$\mathrm{MAE} = \frac{1}{N} \sum_{t=1}^{N} |y_t - \hat{y}_t| \tag{4.10}$$

and the Pearson correlation coefficient is

$$r = \frac{\sum_t (y_t - \bar{y})(\hat{y}_t - \bar{\hat{y}})}{\sqrt{\sum_t (y_t - \bar{y})^2}\,\sqrt{\sum_t (\hat{y}_t - \bar{\hat{y}})^2}}, \tag{4.11}$$

where $\bar{y}$ and $\bar{\hat{y}}$ are the sample means. Both metrics are reported in Chapter 5 for production and spread forecasts.

4.6.4 Statistical Significance of the Revenue Lift

Every revenue-comparison claim in this thesis is assessed with a single significance instrument, defined here and referenced wherever it is used. The instrument acts on the quantity the pipeline optimises, not on a proxy for it. The statistic is the lift in the mean–CVaR objective (3.5) of the proposed strategy over the baseline, evaluated on the post-warm-up trading days where both strategies are defined,

$$\Lambda \;=\; \big(\bar{R} - \lambda\,\widehat{\mathrm{CVaR}}_{0.9}(-\bar{R}_d)\big)_{\mathrm{prop}} \;-\; \big(\bar{R} - \lambda\,\widehat{\mathrm{CVaR}}_{0.9}(-\bar{R}_d)\big)_{\mathrm{base}}, \tag{4.12}$$

with the risk weight λ stated explicitly at each use.

The objective is a non-linear composite of a mean and a tail functional, so it has no closed-form standard error and its sampling distribution is not normal. Significance is therefore assessed by a paired case bootstrap rather than a parametric test. The procedure draws $B = 1000$ resamples. Each resample draws one set of post-warm-up days uniformly with replacement, applies that same set to both strategies so that the within-day market conditions the two share are preserved, and recomputes the lift on the resampled days. Resampling whole days uniformly treats the post-warm-up days as exchangeable, the assumption on which the interval rests. The reported point estimate is the full-sample Λ. The 95 % interval is the $[2.5, 97.5]$ percentile of the resampled lifts, and a lift is called significant when that interval excludes zero.

The risk weight is the only setting that varies between uses. The risk-aware headline of Chapter 8 sets $\lambda = 1$, a unit-consistency choice: the risk penalty is then in the same units as mean revenue, giving a direct euro-for-euro tradeoff. The headline ID-stage conclusion is robust to this choice; see Appendix C.5. Setting $\lambda = 0$ collapses each term of (4.12) to mean daily revenue, so the same instrument returns a distribution-free interval on the mean-revenue lift, reported by the per-stage lift tests of Appendix C.6, without the normality or symmetry assumptions of classical paired tests. The permutation null of Section C.7.2 is a distinct object, a diagnostic of per-slot signal content rather than a significance test of a revenue lift, and is labelled as such throughout.

Each park is tested on its own. All three independently show a positive lift whose 95 % interval lies entirely above zero. The conclusion therefore rests on three separate confirmations rather than one combined test, and no multiple-comparison adjustment is applied.

4.7 Baseline Exposure

The forecast baseline ($b^{\mathrm{DA}} = \hat{a}^{\mathrm{DA}}$, $\Delta b_t^{\mathrm{ID}} = 0$, $s_t = P^{\max}$) is profitable on average at every park. Worst-decile losses live in the imbalance-market settlement: the day-ahead market is reliably a revenue source, the fee is small, and almost all worst-decile losses come from $\pi_t^{\mathrm{IM}} \cdot \nu$, the imbalance settlement on the residual volume between bid and realised production. Figure 4.4 gives the per-park magnitude split into the market-form components of Equation (3.2);

Figure 4.5 places each day in the imbalance-price quadrants.[1]

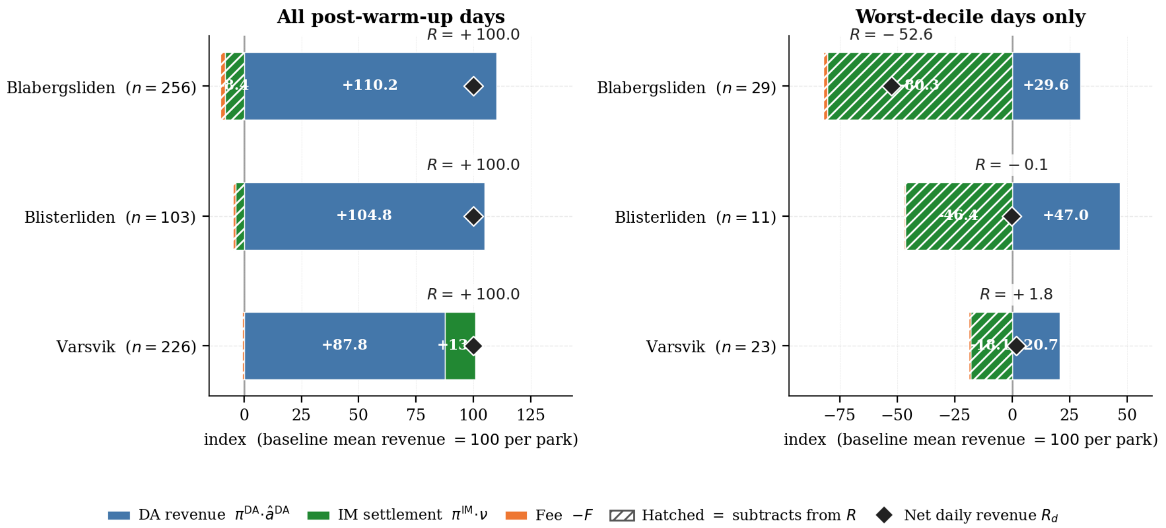

Figure 4.4: Forecast-baseline daily revenue decomposed into DA revenue (blue), IM settlement (green), and fee (orange); positives right, negatives left (hatched); diamond $= R_d$. *Left:* all days, *right:* worst decile.

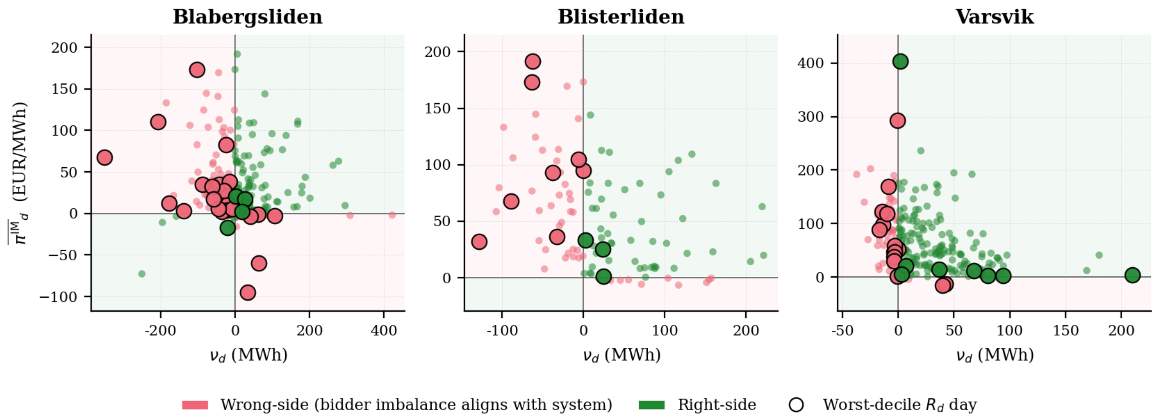

Figure 4.5: Wrong-side imbalance concentration. One dot per post-warm-up day in $(\nu_d, \overline{\pi_t^{\mathrm{IM}}}_d)$ space; red = wrong-side ($\mathrm{sign}(\nu) \neq \mathrm{sign}(\pi_t^{\mathrm{IM}})$, i.e. producer imbalance aligns with the system: long when the system is also long, short when the system is also short), green = right-side, worst-decile days outlined black.

Two wrong-side imbalance modes drive the worst-decile losses, both characterised by the producer's imbalance sharing its sign with the prevailing system imbalance. The dominant mode at every park is *short into positive IM*: the operator production forecast over-predicts wind, the producer is left short at delivery, and the shortfall is bought back at an up-regulating imbalance-price spike. Blabergsliden is additionally exposed to the inverse mode, *long into negative IM* on summer high-wind days when the imbalance price crashes below zero on volume the producer is actively generating. This mode is essentially absent at Blisterliden and Varsvik. The explanation differs by park. Blisterliden's data window covers only the

[1]Several figures in this thesis use colour to encode meaning (e.g. wrong-side / right-side outcomes, per-park markers). The colour distinctions are clearest in the digital PDF.

winter months 2025-11 through 2026-03, missing the summer negative-price episodes entirely. Varsvik sits in SE3, where negative imbalance prices are rarer than in the SE1 zone. The two modes are visible as red-region concentration in Figure 4.5 (bottom-right for the long mode, upper-left for the short mode). Table 4.3 quantifies the per-market attribution and the long/short split inside the IM-loss days.

Table 4.3: Forecast-baseline worst-decile days classified by loss type (DA-loss, IM-loss by position sign, fee-only), per park. $\sum F$ and $\sum R_d$ are summed over the worst decile.

		Worst-decile day count			IM-loss split (n / index)		Window total (index)	
Park	n	DA-loss	IM-loss	Fee-only	long	short	$\sum F$	$\sum R_d$
Blabergsliden	29	3	24	4	7 / −174	17 / −416	−13.5	**−384.5**
Blisterliden	11	0	10	1	2 / −6	8 / −122	−2.1	**−0.3**
Varsvik	23	0	16	7	4 / −48	12 / −68	−4.8	**+11.1**

The day-count totals make the per-park asymmetry quantitative. At Blabergsliden, the long-into-negative-IM mode is a minority of worst days ($n = 7$) but per-day damaging (~ 25 index points each), accounting for just under a third of the IM-cash loss. The rest is short-into-positive-IM ($n = 17$, $\sum = -416$ index). At Blisterliden and Varsvik, the long mode is near-absent and short-into-positive-IM dominates by both day-count and magnitude. Across all three parks, fees are negligible relative to the IM-cash loss per park over the entire window.

The setpoint-stage classifier (Section 7.4) targets the long-into-negative-IM mode by curtailing when the imbalance-price forecast crosses the negative fee band. The credibility-weighted ID-stage rule (Section 7.3) targets the wrong-side mechanism in general by tilting the bid so the producer's imbalance lands on the *right* side of the expected system imbalance: bid below the forecast when up-regulation is expected (so the producer ends up long into the high IM), bid above the forecast when down-regulation is expected (so the producer ends up short into the low IM). The first-order condition $\tau^* = \frac{1}{2} + \delta_t^{\mathrm{DA}}/(2f^{\mathrm{imb}})$ from Equation (7.1) encodes exactly that direction. Figure 8.3 returns to the same wrong-side encoding on full-pipeline outcomes and shows directly which mode each rule resolves.

Chapter 5

Market Signals and Forecast Evaluation

Chapter 3 reduced each stage's optimal bidding decision to an estimate of the expected spread signal at that stage (Propositions 3.1–3.3); the rule appropriate to each stage therefore depends on how much of that signal is empirically available. This chapter measures that availability.

The load-bearing finding is that the directional content of the spread forecast grows monotonically across the three decision stages: at DA it is a coin flip, at ID it is moderate, and at SP it is substantially better than chance. The heterogeneity is large enough that Chapter 7 will need a different rule at each stage.

These measurements also serve as the two empirical inputs that Chapter 7 consumes: stage-specific point forecasts $\hat{\delta}_t$ of the imbalance spread and the predictive scale $\hat{\sigma}_t$.[1]

5.1 Information at the Day-Ahead Stage

At the DA stage (D-1 12:00 CET), the producer must commit a bid volume without knowing the realised production or any of the prices in the spread. The decision-relevant inputs are therefore the production forecast, which anchors the bid centre, and the DA and imbalance-price forecasts that feed scenario construction.

5.1.1 Production Signal

The operator's DA production forecast anchors the DA bid. Per-park accuracy is reported in Table 5.4: correlations exceed 0.8 across all parks, but the residual uncertainty remains substantial. The forecast is also *biased low* at every park. The Bias column (mean signed error, forecast − realised) is negative throughout, growing with wind level and most severe at Varsvik, where it reaches roughly one-fifth of mean available production. The operator point forecast is therefore *not* the conditional mean of available production; this under-forecast, and

[1] The bidding rule sets the per-interval quantile level $\tau_t = \frac{1}{2} + \hat{\delta}_t/(2f^{\mathrm{imb}}) \cdot \gamma_t$; the construction of $\hat{\sigma}_t$ from the 10/90 fan width of the imbalance-price forecast is detailed where it is consumed, in Chapter 7.

what it implies for the choice of DA bid centre, is taken up in Chapters 6 and 7.

Figure 5.1 shows that wind production negatively correlates with the DA *price levels* through the merit-order effect, but has essentially no linear relationship with the imbalance *spread* δ_t^{DA} (numerical ranges in caption). Wind level does carry nonlinear regime information that differs sharply by zone (Appendix B.3): in SE1 down-regulation rises monotonically with wind quintile, whereas in SE3 there is no such gradient.

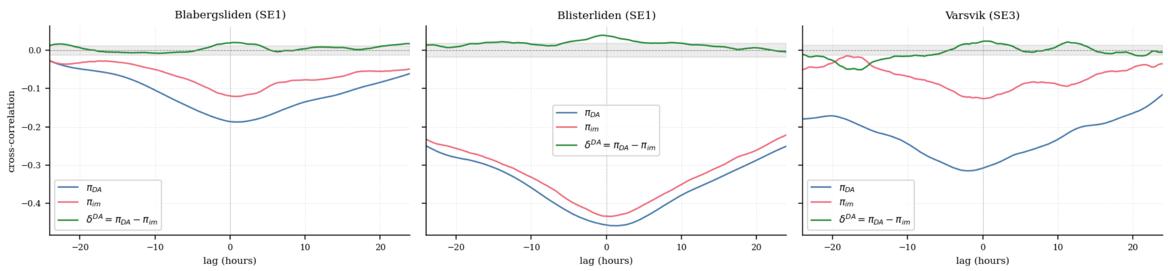

Figure 5.1: Cross-correlation between available production a_t and the three price series (π_t^{DA}, π_t^{IM}, δ_t^{DA}) at lags ±24 h, per park. Shaded band: 95 % CI under zero correlation.

5.1.2 Spread Signal

Production is treated as “free” (zero marginal cost), so the DA bid direction depends only on the expected spread $E[\delta_t^{\mathrm{DA}}]$, not on the level of π_t^{DA}, which is unknown at bid time anyway. The ec00 DA price forecast therefore appears in the thesis only as one low-ranked feature in the scenario kernel (Chapter 6); the DA-stage signal that matters is the imbalance-price forecast.

The Optimeering spread forecast at the DA stage carries essentially no directional information. Table 5.1 benchmarks it against four naïve alternatives. Two findings stand out. First, it has higher MAE than the zero forecast in both zones and lower sign accuracy than the majority-direction baseline. It is beaten by simply predicting the dominant regulation direction. Second, in SE3 extreme up-regulation spikes pull the unconditional mean positive despite a down-regulation majority; winsorising at $\pm 3\hat{\sigma}$ restores sign accuracy above 50 %, confirming the failure is tail-driven.

Table 5.1: DA-stage spread forecast accuracy. MAE in EUR/MWh; sign accuracy is the fraction of non-neutral intervals (those where $\delta_t^{\mathrm{DA}} \neq 0$) on which the forecast sign matches the realised sign.

	SE1		SE3	
Forecast	MAE	Sign acc.	MAE	Sign acc.
Zero (no spread)	22.8	—	41.9	—
Unconditional mean	22.8	63%	42.3	37%
Winsorized $\pm 3\hat{\sigma}$	22.7	63%	41.8	63%
Majority direction	—	63%	—	63%
Optimeering DA point	23.7	58%	49.5	51%

The DA-stage entry of Table 5.3 confirms this: Pearson correlation $r \approx 0$ and sign accuracy near 50 % on non-neutral intervals. The null result does not contradict the fact that distributional regularities exist across many intervals (e.g. extremes tend to be down-regulation in winter mornings): such regularities form a prior over the *distribution*, not a forecast of any specific interval.

A 180-day rolling winsorised mean ($\pm 3\hat{\sigma}$) of δ_t^{DA} therefore serves as the DA-stage point forecast $\hat{\delta}_t^{\mathrm{DA}}$ used throughout the thesis. It captures the seasonal/calendar prior, enough to clear 50 % sign accuracy by a small margin, but no more: this is the practical ceiling at the DA stage.

5.2 What the Intraday Stage Adds

At the ID stage (H-60 min before delivery), both signals improve in decision-relevant ways: a sharper in-house production forecast becomes available, and the imbalance-price forecast transitions from noise to a moderate, confidence-graded signal.

5.2.1 Production Forecast

ID-Linear (Appendix A.1) is a walk-forward linear autoregressive model that predicts the four 15-minute slots of the coming hour at a common origin 90 minutes before delivery. The appendix gives the full specification (lag structure, calendar features, training protocol).

Table 5.2 compares ID-Linear against the operator's ID forecast supplied at gate closure. ID-Linear is more accurate at every park. The gain is largest at Varsvik (SE3), where the operator's ID update actually *degrades* accuracy relative to its own DA forecast; ID-Linear nearly halves the operator's ID error there. The Bias column shows the same dominance on calibration: the operator ID forecast is biased low at every park, severely so at Varsvik (≈ -1.7 MWh per slot), whereas ID-Linear is essentially unbiased everywhere. Unlike the operator DA centre of Section 5.1.1, the production centre the fan is built on at the ID stage is therefore both sharper *and* essentially unbiased. The residual ID limitation is a conditional, per-regime one (Chapter 6), not a systematic low bias.

The accuracy gap translates directly into revenue, as documented in Chapter 7: swapping the operator forecast for ID-Linear lifts mean revenue at every park at the forecast-chase level and lowers tail risk at Blisterliden and Varsvik (Table 7.4); only at Blabergsliden, whose tail is set by sudden production collapses a lagging centre cannot anticipate, does the swap leave the tail slightly heavier. Error distributions across stages and parks are visualised in Figure A.1 (Appendix A).

Table 5.2: ID-stage production forecast accuracy on the post-warm-up evaluation window. The operator forecast ($\hat{a}_t^{\mathrm{ID,op}}$) is the vendor-supplied ID-stage forecast available at gate closure. The ID-Linear forecast ($\hat{a}_t^{\mathrm{ID,Lin}}$) is the in-house linear autoregressive centre (Appendix A.1). Errors are in MWh per 15-minute interval; Bias is the mean signed error (forecast − realised; negative ⇒ forecast biased low); r is the Pearson correlation between forecast and realised available production. Both forecasts are scored on the common set of intervals where the ID-Linear centre is defined (an inner join), so the operator figures here are on a slightly smaller sample than the full-window operator accuracy in Table 5.4.

Park	Forecast	MAE	RMSE	Bias	r
Blabergsliden (SE1)	Operator	2.94	4.54	−0.37	0.900
	ID-Linear	2.49	3.68	−0.14	0.935
Blisterliden (SE1)	Operator	3.19	4.58	−0.31	0.834
	ID-Linear	2.00	2.94	−0.06	0.935
Varsvik (SE3)	Operator	2.09	3.22	−1.61	0.710
	ID-Linear	1.24	1.74	0.09	0.898

5.2.2 Imbalance-Price Signal

At the ID stage, the imbalance-price forecast transitions from noise to a genuine signal. Table 5.3 documents the progression: at ID, sign accuracy on non-neutral intervals reaches roughly two-thirds and correlation with the realised spread is modest but unambiguously positive, far above the DA stage.

Binning sign accuracy by the SNR (Section 4.2) exposes the heterogeneity within each stage (Figure 5.3, consolidated in Section 5.4). At ID, the top SNR quintile reaches roughly 80–87 % sign accuracy across zones, and 90–94 % at SP. At DA the relationship is flat in SE1 and weakly anti-predictive at high SNR in SE3. The ID–SP heterogeneity — aggressive when SNR is high, conservative when it is low — motivates the credibility-weighted ID rule of Chapter 7, where the bid tilt scales with each interval's SNR.

The trading-relevant ID-stage spread estimate is

$$\hat{\delta}_t^{\mathrm{ID}} = \pi_t^{\mathrm{ID}} - \hat{\pi}_t^{\mathrm{IM,ID}},$$

where $\hat{\pi}_t^{\mathrm{IM,ID}}$ is the ID-stage imbalance-price forecast; its sign is positive when selling more on ID is profitable and negative when buying back is. This is the quantity the credibility-weighted ID rule of Chapter 7 conditions on.

5.3 What the Setpoint Stage Adds

At the SP stage (H-20 min before delivery), production uncertainty is essentially resolved: available production is highly autocorrelated at the 30-minute horizon, and the in-house **SP-Persistence** centre (Appendix A.2) is nearly exact (Table 5.4). Production no longer drives the SP decision; this section therefore addresses only the imbalance-price signal, which is now the sole driver.

The imbalance-price forecast reaches its strongest signal at SP — sign accuracy and correlation both peak across stages (Table 5.3) — enough to support a discrete curtailment decision.

The fan width W_t is at its most predictive here. Figure 5.2 sorts intervals by descending fan width: the top quintile concentrates the majority of tail events (defined as $|\delta_t| >$ 100 EUR/MWh), while narrow-fan intervals are nearly tail-free. Wide-fan intervals therefore concentrate the tail exposure that the SP-stage curtailment rules of Chapter 7 target.

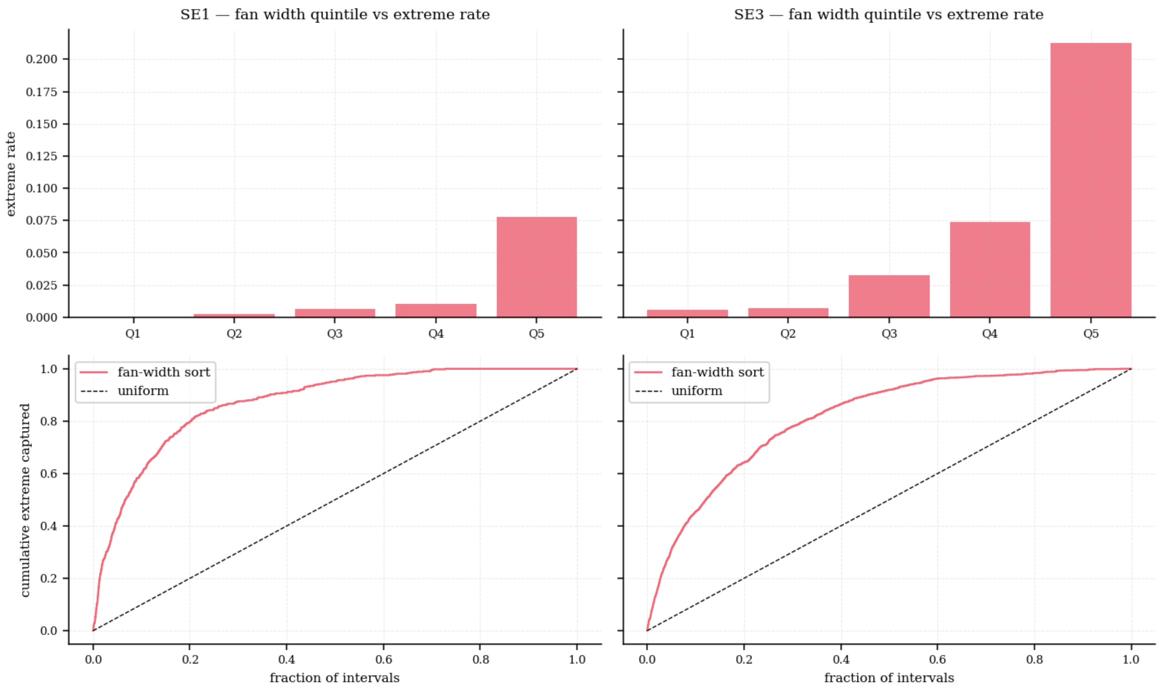

Figure 5.2: Fan width $W_t = Q^{\text{fc}}_{0.99} - Q^{\text{fc}}_{0.01}$ as a tail-intensity signal for SE1 (left) and SE3 (right). Tail events are defined here as $|\delta_t| > 100$ EUR/MWh. *Top:* tail-event rate by fan-width quintile; Q5 reaches 8 % in SE1 and 21 % in SE3, against essentially zero in Q1. *Bottom:* cumulative capture as intervals are flagged in descending fan-width order, versus the uniform baseline.

5.4 Summary: Signals Across Stages

Table 5.3 consolidates spread forecast quality across stages and zones. The DA $\rightarrow$ SP progression is not just monotonic but quantitatively large: top-quintile sign accuracy roughly doubles from DA to SP. This is the empirical condition under which the propositions of Chapter 3 prescribe distinct rules. Downstream chapters cite these tables rather than restating the numbers.

Table 5.3: Spread forecast quality by zone and stage: Pearson correlation r (Equation (4.11)), sign accuracy (fraction of non-neutral intervals correctly signed), and MAE (Equation (4.10)) in EUR/MWh.

Zone	Stage	r	Sign acc.	MAE (EUR/MWh)
SE1	DA	-0.006	58.2%	23.7
SE1	ID	0.230	65.9%	21.2
SE1	SP	0.334	73.7%	18.3
SE3	DA	0.031	50.8%	49.5
SE3	ID	0.289	68.8%	42.6
SE3	SP	0.418	76.9%	36.7

Figure 5.3 breaks the same progression down by SNR (Section 4.2), the quantity the credibility-weighted ID rule of Chapter 7 gates on: at ID, the top SNR quintile reaches 80–87 % sign accuracy across zones; at SP it reaches 90–94 %. The DA stage is flat in SE1 and weakly anti-predictive at high SNR in SE3, a reminder that the DA spread carries no exploitable directional content even after conditioning on forecast confidence.

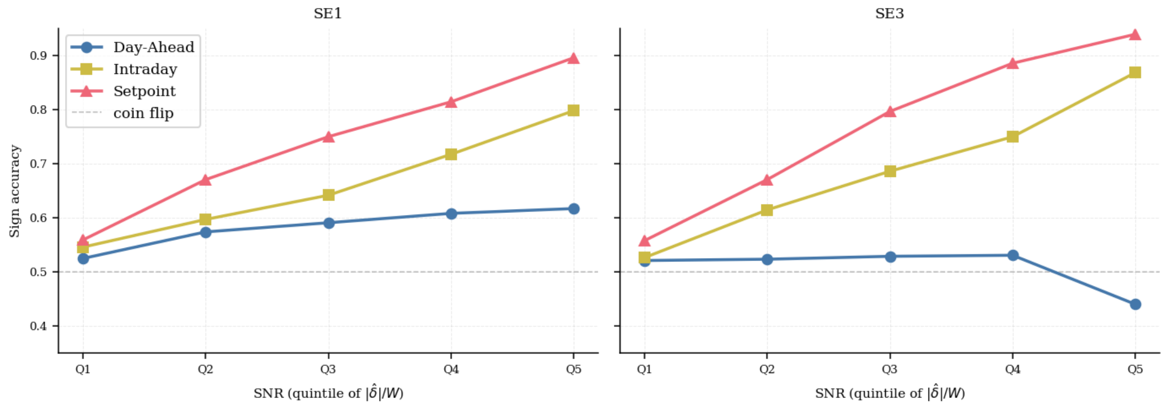

Figure 5.3: Sign accuracy on non-neutral intervals by SNR quintile ($|\hat{\delta}|/W$, $W = Q_{0.99} - Q_{0.01}$; Q5 = highest), per stage, SE1 (left) and SE3 (right).

Table 5.4 consolidates production forecast accuracy. Accuracy improves steadily from DA to SP at every park; spread sign accuracy improves sharply from ID onward.

Table 5.4: Production forecast accuracy per park and stage; spread forecast accuracy per zone. Each production row is the centre actually used by the pipeline at that stage: the operator (vendor) forecast at DA, the in-house ID-Linear centre at ID (Table 5.2), and the in-house SP-Persistence forecast at SP. r: Pearson correlation. MAE and Bias (mean signed error, forecast $-$ realised) in MWh/slot for production rows and EUR/MWh for spread rows.

Source	Stage	r	MAE	Bias
Blabergsliden (y)	DA	0.862	3.29 MWh	-0.45 MWh
Blabergsliden (y)	ID	0.935	2.49 MWh	-0.14 MWh
Blabergsliden (y)	SP	0.991	0.64 MWh	-0.06 MWh
Blisterliden (y)	DA	0.863	3.21 MWh	-1.18 MWh
Blisterliden (y)	ID	0.935	2.00 MWh	-0.06 MWh
Blisterliden (y)	SP	0.993	0.49 MWh	+0.00 MWh
Varsvik (y)	DA	0.824	1.72 MWh	-1.03 MWh
Varsvik (y)	ID	0.898	1.24 MWh	+0.09 MWh
Varsvik (y)	SP	0.983	0.45 MWh	-0.01 MWh
SE1 (δ)	DA	-0.006	23.7 EUR/MWh	-0.3 EUR/MWh
SE1 (δ)	ID	0.230	21.2 EUR/MWh	-1.0 EUR/MWh
SE1 (δ)	SP	0.334	18.3 EUR/MWh	-0.4 EUR/MWh
SE3 (δ)	DA	0.031	49.5 EUR/MWh	+8.2 EUR/MWh
SE3 (δ)	ID	0.289	42.6 EUR/MWh	-0.7 EUR/MWh
SE3 (δ)	SP	0.418	36.7 EUR/MWh	-1.8 EUR/MWh

Chapter 6

Production Scenario Construction

This chapter builds the conditional distribution of available production a_t that the bidding rules of Chapter 7 read quantiles from. It is a residual day-vector bootstrap plus two corrections that together turn it into a calibrated, conditional fan. *Variance rescaling* sets the fan's spread to the target day's regime; a *conditional kernel* builds it from history that resembles the target (selected per stage by one rule: screen by MI, confirm on CRPS, adopt).[1]

6.1 Why a Distribution

Propositions 3.1 and 3.2 establish that the revenue-optimal bid at every stage is a quantile of $F^{-1}_{a_t|\mathcal{F}}$, with the quantile level τ set by the spread signal. At DA the spread signal is close to uninformative, so $\tau \approx \frac{1}{2}$ and the optimal bid is the conditional *median* of a_t. At ID the signal carries genuine directional content and τ_t swings well away from $\frac{1}{2}$. In both cases the bid needs the full conditional distribution.

That distinction is not academic, for two compounding reasons. First, the operator DA point forecast is not even a calibrated centre: it is biased low at every park (Table 5.4), so it is neither the conditional mean nor the conditional median of a_t. Second, the production-residual distribution is markedly skewed. Figure 6.1 groups residuals by daily wind level: at low wind the zero floor truncates negative errors, producing strong right skew; at high wind the $P^{\max}$ ceiling flips the asymmetry to left skew. Under skew, the median sits a long way from the mean, so even a perfectly centred forecast would be the wrong object to bid at $\tau = \frac{1}{2}$: the newsvendor optimum there is the conditional *median*, not the mean. The theoretical optimum is therefore a quantile of the conditional distribution, not the operator point forecast: locating it — the conditional median at DA, a signal-shifted quantile off it — requires constructing that distribution. Whether *adopting* the DA correction it implies pays once single-price imbalance risk is priced is an empirical question, settled in Chapter 7.

[1] Feature-group selection was confirmed on CRPS over the full evaluation window rather than a held-out set. Given that the kernel's contribution to overall performance is small (Section 6.4), this has negligible practical effect on the reported bidding results.

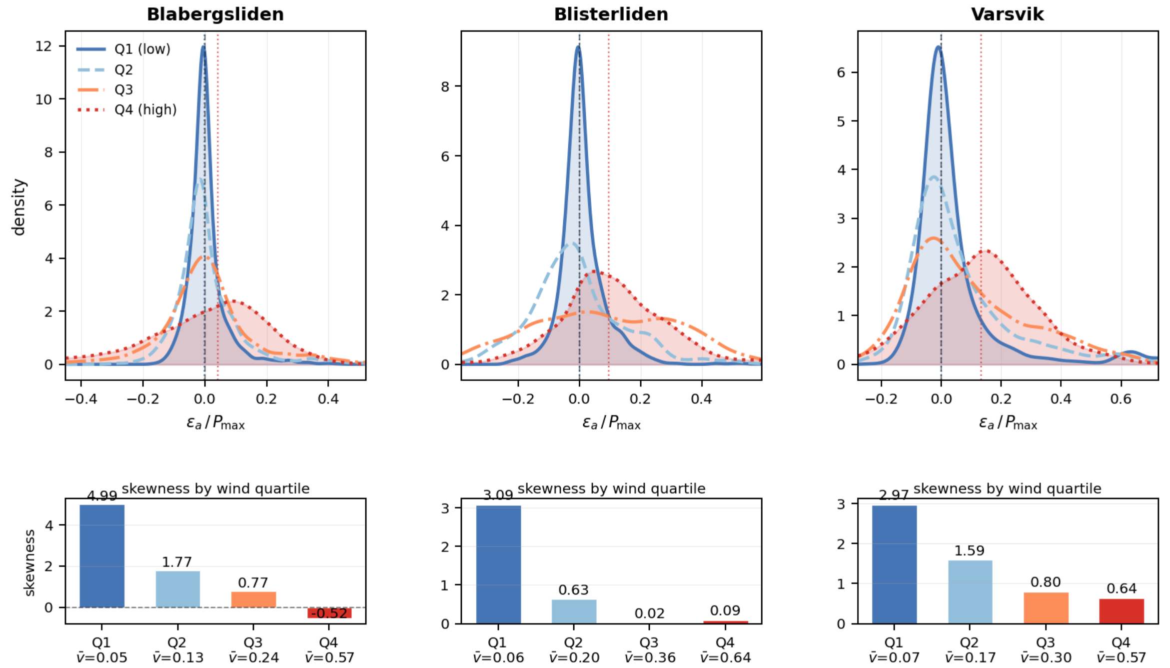

Figure 6.1: Production residuals $\epsilon_t^a = a_t - \hat{a}_t^{\mathrm{DA}}$ (normalised by $P^{\max}$) by daily wind-forecast quartile. *Top:* per-quartile KDE; dotted line = quartile median. *Bottom:* Pearson skewness by quartile.

6.2 Residual Day-Vector Bootstrap

Error library. A separate error library is built for each decision stage. For each historical day d and slot t, the production residual against the stage's point forecast $\hat{a}_t$ is

$$\epsilon_t^{a,d} \;=\; a_t - \hat{a}_t, \tag{6.1}$$

where $\hat{a}_t$ is the operator DA forecast at the DA stage, the ID-Linear intraday forecast at the ID stage (Appendix A.1), and the SP-Persistence setpoint forecast at the SP stage (Appendix A.2). The collection of all eligible day vectors is the error library $\mathcal{L}$; a library day is included only if all 96 quarter-hourly production observations are present, and the kernel excludes any day not strictly before the target day (no look-ahead). Table 6.1 lists the three production-residual libraries. The ID and SP forecasts are tighter than the DA forecast, so their residual libraries are narrower and the fan contracts as delivery approaches.

Table 6.1: Point forecast each stage's production residuals are centred on.

Stage	Production centre
DA	operator DA forecast
ID	ID-Linear
SP	SP-Persistence

Scenario assembly. Let d^* be the target day and d_s the sampled library day for scenario s. Each scenario is assembled by adding the stored day-vector of residuals to the target day's forecast:

$$a_t^{(s)} \;=\; \hat{a}_t \;+\; \epsilon_t^{a,\,d_s}, \tag{6.2}$$

clipped to $[0, P^{\max}]$. Because all 96 production residuals travel together as a single day vector, each scenario inherits the sampled day's within-day dependence structure: ramp timing and morning/evening asymmetry. The assembly step is identical at every stage; what changes is which error library is used, and therefore which point forecast $\hat{a}_t$ the residuals are centred on.

The bootstrap re-centres the DA forecast bias. Assembly (6.2) has a property that is decisive at the DA stage. The DA error library is *not* mean-zero: since $\epsilon_t^a = a_t - \hat{a}_t^{\mathrm{DA}}$ and the operator DA forecast is biased low (Table 5.4), the library's mean *is* that bias. Adding a sampled residual vector back to the forecast therefore restores the missing level. The fan's centre tracks the conditional mean of a_t even though the raw point forecast does not. This is a second, independent reason, beyond skew, that the bid must be read from the fan, not the forecast. The ID and SP stages differ in kind: their libraries are centred on the fitted ID-Linear/SP-Persistence models, whose residuals are mean-zero by construction. There the bootstrap does *not* self-correct. Whatever *conditional*, per-regime bias those models carry on a given day passes through. This is the forecast-quality bound returned to in Sections 6.4 and 6.5.

At the ID stage, each slot's bid is submitted approximately one hour before delivery, with the ID-Linear forecast evaluated at that rolling horizon; the residual library stores the per-slot, one-hour-ahead ID-Linear errors. A full 96-slot day vector is therefore valid in the backtest: the 96 decisions are independent at their respective one-hour horizons, and a library day's residual vector stores all 96 errors.

This skeleton — resample a whole historical error-day, add it to the target forecast — is not yet calibrated. An unconditional draw from it covers the realised production only 73 % of the time at the nominal 80 % level (Table 6.2, "Uniform").

6.3 Variance Rescaling

Wind-power forecast error is strongly heteroskedastic in the wind level: residuals on a windy day are roughly twice as wide as on a calm day, and roughly twice as wide in winter as in summer. Resampling a windy library day's residual vector for a calm target day therefore over-disperses the fan, and the reverse under-disperses it, the main reason the unconditional bootstrap is miscalibrated.

To correct it, every sampled day-vector is multiplied by one scalar

$$r_s \;=\; \frac{\hat{\sigma}(v^*)}{\sigma_{d_s}}, \qquad r_s \in [0.5,\, 3], \tag{6.3}$$

where σ_{d_s} is the standard deviation of the sampled day's own 96 residuals, v is the daily-mean normalised wind level $\overline{\hat{a}^{\mathrm{DA}}}/P^{\max}$ used as a regime proxy, and $\hat{\sigma}(v^*)$ is a kernel estimate of the residual standard deviation at the target day's wind level, computed only from library days strictly before the target (causality-safe).[2] The clip in (6.3) bounds the correction on poorly-populated wind levels. Because r_s multiplies the *entire* day-vector, the within-day dependence structure — ramp timing, morning/evening asymmetry — is preserved; only the overall spread is brought into line with the target day's regime. Rescaling acts on the residual *spread*, not its centre: at ID and SP the mean-zero libraries are unaffected, while at DA the bias scales with r_s by a per-day factor but is not undone (Section 6.2).

Applied to the plain unconditional bootstrap, rescaling moves 80 % coverage from 73 % to 78 %, within roughly two points of nominal, and cuts CRPS by about 2 % (Table 6.2): the spread correction doing its job, the fan's width now tracking the target day's regime. The second tool, the conditional kernel of the next section, builds the fan from days that resemble the target; it helps too, and because both respond to the same heteroskedastic structure their coverage gains overlap rather than add.

6.4 Conditional Day-Selection

The second correction weights historical days by similarity to the target, so a windy target day draws preferentially from windy history. Because rescaling has already corrected the shared miscalibration, conditioning adds only a further 0.6 % of CRPS and essentially no coverage gain. That small *incremental* number is not an argument against conditioning: Section 6.1 fixed the bid as a *conditional* quantile, so the fan must condition on the target day regardless. What carries weight is then the *selection rule*, applied identically at both stages: screen candidate features by mutual information, confirm the winner on CRPS, adopt it. Days are weighted

$$\Pr(d_s = d) \;\propto\; \exp\!\Big(-\tfrac{1}{2}\,\Delta(x_d, x_{d^*})^2/h^2\Big), \tag{6.4}$$

with x_d the day's feature vector (standardised over the library), Δ weighted Euclidean distance, and h the median target-to-library distance. A uniform blend stabilises the weights at low effective sample size, and recency enters as a down-weighted feature; both parameters are fixed across stages (Table 6.3).

DA stage. Five groups are screened beyond the base set by kNN mutual information (Kraskov, Stögbauer, and Grassberger 2004) against next-day mean absolute production error $\bar{e}_d = \overline{|a_t - \hat{a}_t^{\mathrm{DA}}|}$ (a per-day proxy; the bake-off tests the full distribution directly): **wind-shape**, **price-shape**, **spread-fan**, **lagged outcomes**, **calendar**. Wind-shape leads; lagged outcomes and calendar are negligible and dropped (Figure 6.2, left; per-feature scores in Appendix B.1). Scored directly on CRPS — the same groups, in the *Day-ahead* block of Table 6.2 — the feature groups are interchangeable: wind-shape, spread-fan and price-shape are indistinguishable on the proper score, and none shifts coverage. Within conditioning, the coverage gain comes from recency: the *Recency only* row lifts coverage by about three points at no CRPS cost. The MI winner is no worse than any alternative, so wind-shape is adopted; with rescaling applied, the adopted kernel lands within two points of nominal coverage (Table 6.2).

[2] Gaussian kernel, bandwidth 0.15 on the normalised wind level $v \in [0, 1]$.

Table 6.2: Per-stage conditioning ablation. Each block scores, directly on CRPS, the candidate groups that stage's own MI screen ranks (Figure 6.2): wind-/price-shape and spread-fan at DA; IDLinear-profile, cleared DA price and the ID-stage spread fan at ID. $S = 200$, expanding-window walk-forward, causality-safe (30-day warm-up at DA, 90-day at ID), 3-park average, production a. CRPS energy-form (4.8); Cov_{80} = slot fraction in $[Q_{10}, Q_{90}]$. † marks the adopted kernel (Table 6.3). Rescaling and conditioning are two corrections on the same fan; the adopted kernels sit ≈ 2.5% (DA) and ≈ 2.4% (ID) below their uniform baselines, the feature groups interchangeable on CRPS.

Configuration	CRPS (MWh)	Cov_{80} (%)
Day-ahead stage		
Uniform	2.108	73.7
Recency only	2.073	76.6
Base + wind-shape	2.090	73.8
Base + price-shape	2.096	73.8
Base + spread-fan	2.093	73.7
Uniform + rescaling	2.064	78.2
Base + wind-shape + recency + rescale	2.056†	78.3
Intraday stage		
Uniform	1.470	71.5
Recency only	1.459	73.0
Base + wind-shape	1.466	71.4
Base + IDLin-profile	1.460	71.3
Base + actual-DA-price	1.468	71.3
Base + ID-spread-fan	1.465	71.3
Uniform + rescaling	1.439	77.1
Base + wind-shape + IDLin-prof. + recency + rescale	1.436†	76.9

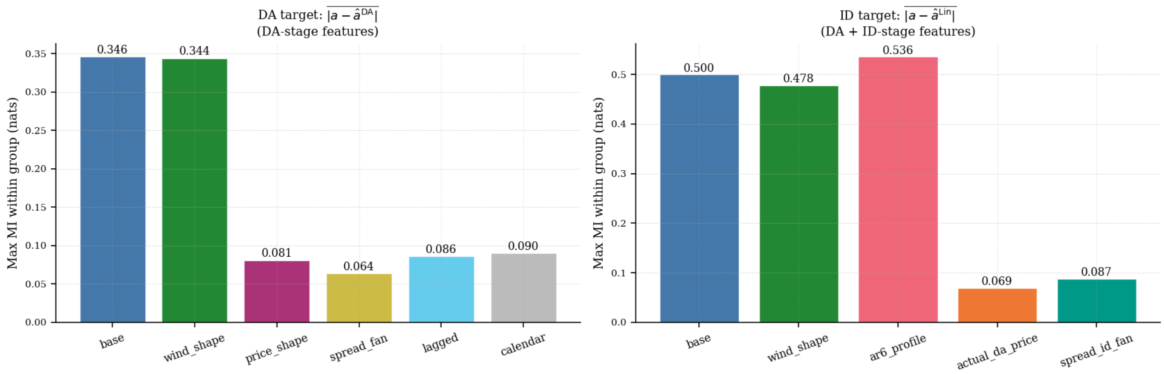

Figure 6.2: Feature-group kNN MI (Kraskov, Stögbauer, and Grassberger 2004), 3-park average ($k = 5$). *Left:* DA stage vs. $\overline{|a - \hat{a}^{\text{DA}}|}$ (DA groups). *Right:* ID stage vs. $\overline{|a - \hat{a}^{\text{Lin}}|}$, incl. three ID-only groups.

ID stage. The same rule runs at ID against the ID-Linear-centred error $\bar{e}_d^{\text{Lin}} = \overline{|a_t - \hat{a}_t^{\text{Lin}}|}$, now screening the three groups admissible only after DA gate closure: **ID-Linear-profile** (block/shape of the ID-Linear forecast), **actual DA price**, and **ID spread fan**. ID-Linear-profile leads the screen by a clear margin (Figure 6.2, right) — the forecast's own shape encodes the regime driving its residuals — with base and wind-shape close behind and the other two an order of magnitude lower. Scored directly on CRPS — the same groups, in the *Intraday* block of Table 6.2 — the MI winner is no worse than any ID alternative on the proper score and does not degrade coverage. ID-Linear-profile is therefore adopted on top of base + wind-shape, with rescaling applied as well — the same rule as at DA, a different winner because a different forecast is now available.

The kernel's gain at ID is small for a structural reason that bounds any kernel here: rescaling moves spread, not centre (Section 6.2), so whatever bias ID-Linear carries on a day passes through regardless of which days are drawn. ID underdispersion is forecast-bias-bound; the binding constraint is forecast quality, the highest-leverage extension (Chapter 9).

Adopted vector. DA uses *base + wind-shape + recency* (22 features); ID extends it with ID-Linear-profile (35). Table 6.3 defines it.

Table 6.3: Conditioning vector x_d. Base, wind-shape and recency use DA-gate data ($\hat{a}_t^{\mathrm{DA}}$, EC00 DA price) and apply at both stages; recency is a feature down-weighted to $\omega_{\mathrm{age}} = 0.5$, and the kernel weights are blended with a uniform component ($\beta = 0.2$). ID extends x_d with the ID-Linear-profile block (13 statistics from the intraday forecast), selected by the ID-stage MI screen (Figure 6.2, right).

Group	Symbol	Definition	Dim
Base	$\bar{v}_k,\ k = 1\ldots4$	Mean $\hat{a}_t^{\mathrm{DA}}/P^{\max}$ over 6-h block k (00–06h, 06–12h, 12–18h, 18–24h UTC)	4
	$\bar{v},\ \sigma_v$	Daily mean and std of $\hat{a}_t^{\mathrm{DA}}/P^{\max}$	2
	$\bar{\pi}_k^{\mathrm{DA}},\ k = 1\ldots4$	Mean EC00 DA price forecast per 6-h block (EUR/MWh)	4
	$\bar{\pi}^{\mathrm{DA}},\ \sigma_\pi$	Daily mean and std of DA price forecast	2
	$\sin\theta_d,\ \cos\theta_d$	Day-of-year encoding: $\theta_d = 2\pi \cdot \mathrm{day}_d/365.25$	2
Wind-shape	$v_{\max},\ v_{\min},\ v_{\mathrm{ptp}}$	Daily max, min, and range of $\hat{a}_t^{\mathrm{DA}}/P^{\max}$	3
	$r_\uparrow,\ r_\downarrow$	Largest 15-min upward / downward ramp in $\hat{a}_t^{\mathrm{DA}}/P^{\max}$	2
	$h_d^{\sin},\ h_d^{\cos}$	Cyclical encoding of peak-production slot: $2\pi\, t^*/96$	2
Recency	age_d	Calendar days before target; standardised feature, kernel weight $\omega_{\mathrm{age}} = 0.5$	1
ID-Linear-prof. (ID only)	$\hat{a}_t^{\mathrm{Lin}}$ block/shape	The 13 base+wind-shape statistics, recomputed from the ID-Linear forecast $\hat{a}_t^{\mathrm{Lin}}/P^{\max}$	13
		Total in x_d — DA / ID	**22 / 35**

Figure 6.3 shows the adopted fan on one representative day: the DA fan is wider (larger forecast uncertainty at the earlier gate); the ID fan collapses around the ID-Linear forecast as delivery approaches.

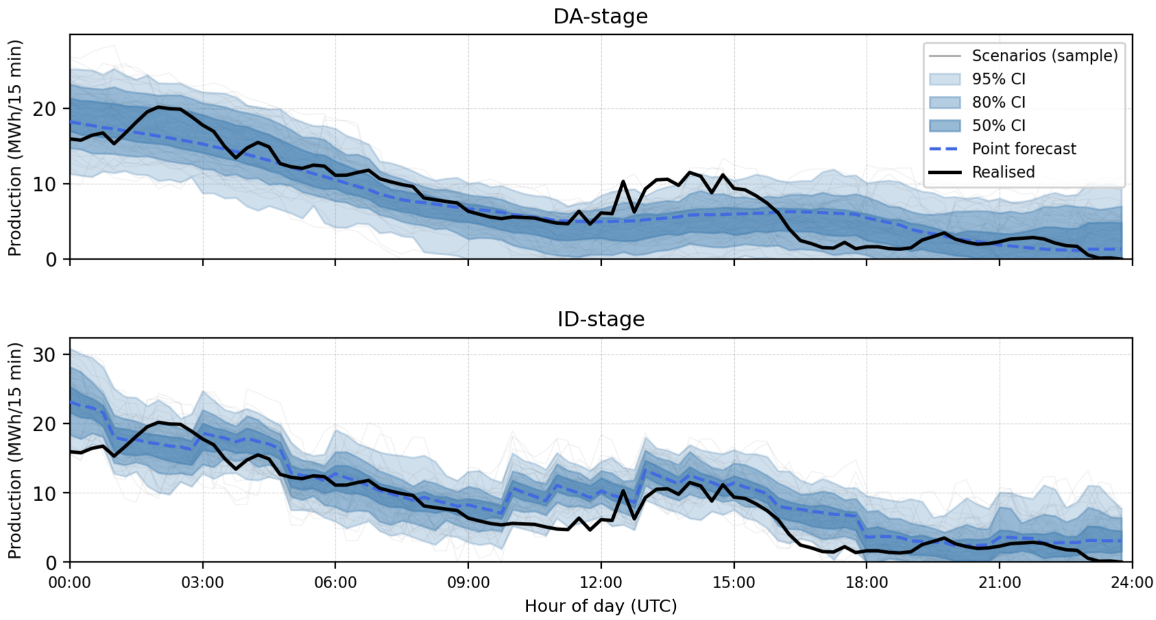

Figure 6.3: Production fan, Blabergsliden 2025-11-07 ($S = 200$). *Top:* DA stage. *Bottom:* ID stage (ID-Linear centre, kernel extended with ID-Linear-profile; Table 6.3). Bands: 50/80/95 % PIs; grey: 60 sample paths; dashed blue: point forecast; black: realised.

6.5 Verification

The adopted fan is calibrated if the probability integral transform (PIT) of realised a_t (each a_t through its own predictive CDF) is Uniform$(0, 1)$ and the 80 % central interval holds a_t about 80 % of the time, evaluated out-of-sample over all day-slot pairs in an expanding-window walk-forward. CRPS is reported alongside as a proper score.

At the DA stage, the PIT is near-uniform across all three parks (top row of Figure 6.4) and 80 % coverage lands within a few points of nominal (Table B.1, App. B): the bid quantiles Chapter 7 reads at DA are sound.

At the ID stage coverage spreads more across parks — the forecast-bias bound of Section 6.4 realised per park, not a kernel failure. The ID quantiles are sound where ID-Linear is well-centred and mildly optimistic where it is biased; closing that gap is a forecasting problem, taken up as future work (Chapter 9). The price marginal is not separately PIT-validated. The adaptive-quantile rules read the production marginal, and the joint (a, π^{IM}) construction is consumed only by the SAA LP of Appendix C.7.

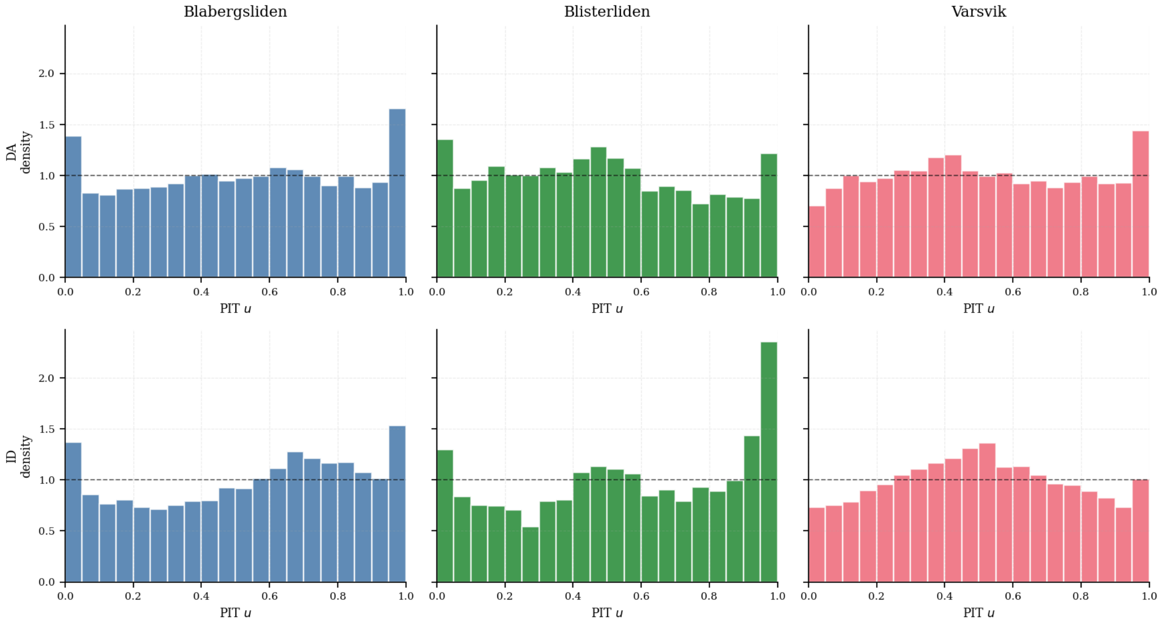

Figure 6.4: PIT of realised a_t under the adopted per-stage kernels, per park, full expanding-window. *Top:* DA stage (30-day warm-up). *Bottom:* ID stage (90-day warm-up, ID-Linear centre). Flat = calibrated.

Chapter 7

Signal-Conditioned Bidding and Curtailment

This chapter assembles the spread-driven quantile level τ_t and the predictive production fan $F_{y|\mathcal{F}_t}$ developed in Chapters 5 and 6 into the bid $b_t = F^{-1}_{y|\mathcal{F}_t}(\tau_t)$. Chapter 3 already fixes the bid form — a signal-quality-weighted production-fan quantile at DA and ID, a discrete setpoint rule at SP. That form does not change across the three gates; only the quality of the directional signal feeding it does (Table 5.3), and the rule adopted at each gate is the one matched to that quality. Section 7.1 states the shared rule form. The three stage sections then take the gates in order of sharpening signal: Section 7.2 the day-ahead *control*, where the signal is at chance; Section 7.3 the methodological core, where an informative-but-noisy signal earns the credibility-weighted tilt; Section 7.4 the setpoint stage, where a strong signal supports a discrete classifier. Chapter 8 then composes the three adopted rules into the proposed pipeline.

7.1 The Shared Rule Form

The risk-neutral per-interval optimum of Propositions 3.1 and 3.2 (at $\lambda = 0$ in (3.5), spread-form revenue (3.3)) is the chapter's starting point:

$$x^* \;=\; F^{-1}_{y|\mathcal{F}}(\tau^*), \qquad \tau^* \;=\; \tfrac{1}{2} + \frac{\mathbb{E}[\delta \mid \mathcal{F}]}{2f^{\text{imb}}}. \tag{7.1}$$

The bid is a quantile of the predictive production fan (Chapter 6) at a level τ^* set by the expected imbalance spread and the eSett fee; it needs only the conditional mean of δ and the marginal CDF of y, so substituting the point estimate $\hat{\delta}_t$ for $\mathbb{E}[\delta \mid \mathcal{F}]$ makes it directly usable. This form is not the chapter's contribution; what is open is everything it leaves unspecified, which the rest of this section fixes.

Specific signal-quality forms. Section 3.5 established the signal-quality-aware quantile family: by (3.12) the bid is the risk-neutral tilt attenuated by a factor $\gamma_t \in [0,1]$ ($\gamma_t \to 0$: production-fan median; $\gamma_t \equiv 1$: full tilt), clipped to the band. The two specific forms

(Table 7.1) are

$$\tau_t = \mathrm{clip}\left(\tfrac{1}{2} + \gamma_t \,\frac{\hat{\delta}_t}{2 f^{\mathrm{imb}}},\ \tau_{\mathrm{lo}},\ \tau_{\mathrm{hi}}\right), \qquad \gamma_t = \begin{cases} 1 & \text{(BFS)},\\ \min\left(1,\ \dfrac{|\hat{\delta}_t|}{z\,\hat{\sigma}_t}\right) & \text{(BCW)}. \end{cases} \tag{7.2}$$

BFS commits fully; BCW shrinks the tilt by the per-interval signal-to-noise ratio $|\hat{\delta}_t|/(z\,\hat{\sigma}_t)$, so any BCW–BFS difference is the SNR shrinkage alone. The SAA LP (Section 3.5, (3.11)) solves the scenario program directly; it serves only as an in-sample upper bound the family is judged against and is never a candidate for adoption. The two free parameters, the band half-width w and BCW's SNR threshold z, are re-selected per delivery day by walk-forward mean–CVaR empirical-risk minimisation over a fixed 8×8 grid ($w \in \{0.10, \ldots, 0.45\}$, $z \in \{0.5, \ldots, 4\}$),

$$(\hat{w}_d,\ \hat{z}_d) \;=\; \arg\max_{w',\,z'} \left[\overline{R}^{(w',z')}_{<d} \;-\; \lambda\, \widehat{\mathrm{CVaR}}_{0.9}\big(-\overline{R}^{(w',z')}_{<d}\big) \right], \tag{7.3}$$

on history strictly before day d (the expanding-window protocol of Section 4.5). BFS and the SAA LP calibrate w only, since $\gamma_t \equiv 1$ removes any role for z; per-day $(\hat{w}_d, \hat{z}_d)$ trajectories are in Appendix C.4. Each stage applies these forms to its own signal and adds at most one stage-specific baseline.

Table 7.1: Shared signal-responsive rule family. BidF: no-signal endpoint. BFS and BCW: two parametrisations of (3.12) differing in γ_t. SAA LP: direct mean–CVaR optimisation on the scenario fan.

Candidate	Abbrev.	Mechanism	Band	Free params
BidForecast	BidF	production forecast as bid; no signal-conditioned adjustment	—	—
BoundedFullSignal	BFS	$\gamma_t \equiv 1$ on (3.12), band cap	$[\tau_{\mathrm{lo}}, \tau_{\mathrm{hi}}]$	w
BoundedCredWeighted	BCW	$\gamma_t = \min(1, \lvert\hat{\delta}_t\rvert/(z\hat{\sigma}_t))$, band cap	$[\tau_{\mathrm{lo}}, \tau_{\mathrm{hi}}]$	w, z
Sample-Average Approximation Linear Program	SAA LP	direct scenario optimisation (3.11); in-sample upper bound	$[\tau_{\mathrm{lo}}, \tau_{\mathrm{hi}}]$	w

7.2 Day-Ahead

The DA-stage estimate of the imbalance spread has sign accuracy at chance (Table 5.3), so the theoretical tilt is zero and the principle prescribes a bid at the conditional median of the production fan. The only open question is whether any rule acting on this chance-level signal, or re-centring the conditionally low operator forecast (Section 5.1.1), can beat the operator bid in practice. Of the candidates in Table 7.2, the one adopted in the proposed pipeline is BidF itself. None of the alternatives beats it robustly: the per-park objective winner differs by park, no candidate wins at more than one of the three sites, and the lift over BidF never exceeds roughly +10 index points (Section 9.1). The null result is the expected one. The pipeline adds no value where the signal carries none.

The candidate set applies the Section 7.1 family to the per-slot DA-stage spread estimate (BFS/BCW/SAA LP; Table 7.1) and adds four DA baselines that use no per-interval signal: the operator bid BIDF, its fan-median and fan-mean re-centrings (FANMEDIANDA/FANMEANDA), and a 180-day rolling winsorised-mean rule (RollingMean-DA; defined in Appendix C.1). Table 7.2 reports mean revenue, $\text{CVaR}_{0.9}$, and the mean–CVaR objective for the DA candidates at every park. The fan-mean re-centring (FANMEANDA) and the full bias-correction decomposition are deferred to Appendix C.3.

Table 7.2: DA-stage candidates: mean daily revenue $\bar{R}$, $\text{CVaR}_{0.9}$ of daily losses, and mean–CVaR objective ($\lambda = 1$), all in index units (baseline = 100 per park), on the post-warm-up evaluation window. A star ($\star$) marks the per-park objective winner; the dagger (†) marks the candidate adopted in the proposed pipeline.

	Blabergsliden			Blisterliden			Varsvik		
Candidate	$\bar{R}$	CVaR	Obj	$\bar{R}$	CVaR	Obj	$\bar{R}$	CVaR	Obj
BIDFORECAST †	100.0	58.5	41.5	100.0	−0.0	100.0 *	100.0	−1.5	101.5
FANMEDIANDA	99.9	55.1	44.9	101.9	2.9	99.0	100.9	−0.2	101.1
BOUNDEDFULLSIGNAL	102.8	71.6	31.2	103.3	7.3	96.0	101.1	9.8	91.3
BOUNDEDCREDWEIGHTED	100.8	53.0	47.8 *	103.1	6.8	96.3	102.5	8.0	94.6
SAA LP	98.7	54.7	44.1	97.9	2.4	95.5	100.9	−3.3	104.2 *
BIDTRUEPROD (DA)	113.0	0.4	112.6	112.8	−10.3	123.1	106.6	−9.2	115.8

The per-park detail behind the null: the objective winner is BCW at Blabergsliden, BIDF at Blisterliden, and the SAA LP at Varsvik (Table 7.2); BFS ($\gamma_t \equiv 1$) loses at all three. The cross-park inconsistency is the binding observation — no candidate generalises — which is why BIDF is the robust choice.

Bias correction. The operator forecast is biased low and the fan re-centres that bias by construction (Section 6.2), so a natural question is whether bidding the fan's median or mean beats the biased operator bid. It does not robustly. Re-centring lifts mean revenue at nearly every park, but on the mean–CVaR objective FANMEDIANDA is point-positive at only one of three parks and within the DA-stage noise. The fuller conditional-mean re-centring is the *worst* choice at the most right-skewed park, where the extra over-bid widens the imbalance-fee tail. Correcting a genuine forecast bias still manufactures no DA edge *within the tested re-centrings*; what an ideal DA production forecast *could* deliver is reported in the BIDTRUEPROD (DA) row of Table 7.2 — the counterfactual of Section 3.6, which bids the realised availability at the gate and therefore carries no imbalance. The gap to BIDF sits well above any tested DA candidate at the SE1 parks, locating the dominant residual exposure in DA-stage production-forecast error rather than in the absence of a spread tilt; this is taken up in Sections 9.4 and 10.4. The per-park decomposition and supporting table for the re-centrings themselves are in Appendix C.3.

7.3 Intraday

This is the chapter's methodological core. At the ID gate, the DA price is realised, so the forecast is an imbalance-price forecast rather than the spread of two unknowns it was at DA, and it is genuinely informative: sign accuracy and correlation rise sharply over the DA stage (Table 5.3). It also arrives with a 10/90 fan, hence a per-interval uncertainty scale. Of the candidates in Table 7.3, the one adopted is the credibility-weighted tilt BCW, which turns this informative-but-noisy signal into an objective lift an order of magnitude larger than any DA-stage lift of Section 7.2 by shrinking the bid toward the production-fan median in proportion to the per-interval SNR (Figure 7.1). Why this rule shape emerges from this signal quality is interpreted in Section 9.2.

The signal is the gap between the observed ID price and the imbalance-price forecast, with a one-sigma uncertainty read from the 10/90 fan:

$$\hat{\delta}_t^{\mathrm{ID}} = \pi_t^{\mathrm{ID}} - \hat{\pi}_t^{\mathrm{IM,ID}}, \qquad \hat{\sigma}_{\delta,t}^{\mathrm{ID}} = \frac{Q_{0.90}^{\mathrm{IM,ID}}(t) - Q_{0.10}^{\mathrm{IM,ID}}(t)}{2 \cdot 1.28}. \tag{7.4}$$

Substituting into (7.2) gives BFS and BCW directly, with the τ_t-quantile read off the production fan centred on the ID-Linear forecast (Section 5.1) and $(\hat{w}_d, \hat{z}_d)$ walk-forward calibrated by (7.3). The derivation targets the risk-neutral marginal optimum: a marginal production quantile is enough because the spread multiplies the deterministic bid in (3.3). The mean–CVaR criterion the rule is *evaluated* on is a joint-law functional this marginal form does not model; the residual production–spread tail coupling is quantified in Appendix C.8.

Three baselines isolate the source of the gain by tracking only a production forecast. BidF keeps the DA bid, IDF bids the operator's ID-stage forecast, and ID-Linear bids the in-house centre (Appendix A.1). The gaps measure the value of any ID update, of a better update, and of the imbalance-price forecast itself. The SAA LP of Section 7.1 serves as the scenario comparator: it ignores estimator uncertainty and bids off the scenario mean as though it were the truth (Appendix C.7). Table 7.3 reports mean revenue, $\mathrm{CVaR}_{0.9}$, and the mean–CVaR objective for every ID candidate at every park; Figure 7.1 traces cumulative ΔR and cumulative Δtail loss vs BidF.

Table 7.3: ID-stage candidates: mean daily revenue $\bar{R}$, $\mathrm{CVaR}_{0.9}$, and mean–CVaR objective ($\lambda = 1$), all in index units (baseline = 100 per park). A star ($\star$) marks the per-park objective winner; the dagger (†) marks the candidate adopted in the proposed pipeline.

	Blabergsliden			Blisterliden			Varsvik		
Candidate	$\bar{R}$	CVaR	Obj	$\bar{R}$	CVaR	Obj	$\bar{R}$	CVaR	Obj
BidForecast	100.0	58.5	41.5	100.0	−0.0	100.0	100.0	−1.5	101.5
IDForecast	100.7	52.9	47.8	101.0	−2.1	103.2	99.7	−2.8	102.4
ID-Linear Correction	100.9	54.0	46.9	102.5	−2.7	105.2	102.8	−4.7	107.5
BoundedFullSignal	112.2	50.2	62.0	108.3	−3.1	111.4 *	116.5	−13.0	129.5 *
BoundedCredWeighted †	111.5	49.5	62.1 *	107.9	−2.8	110.7	115.6	−12.9	128.5
SAA LP	106.0	56.6	49.4	103.0	−0.2	103.2	107.3	−7.9	115.2
BidTrueProd (ID)	106.7	25.8	80.9	105.6	−3.6	109.2	105.2	−8.3	113.5
BCW on δ_t^{ID}	132.5	37.5	95.0	118.1	−7.8	125.9	139.8	−19.6	159.4

Forecast quality is a smaller, secondary lever. The in-house ID-Linear forecast does lift mean revenue at every park relative to the operator forecast, but the magnitude is small relative to the parametric-rule lift over BidF, and the tail effect is mixed across parks (Ta-

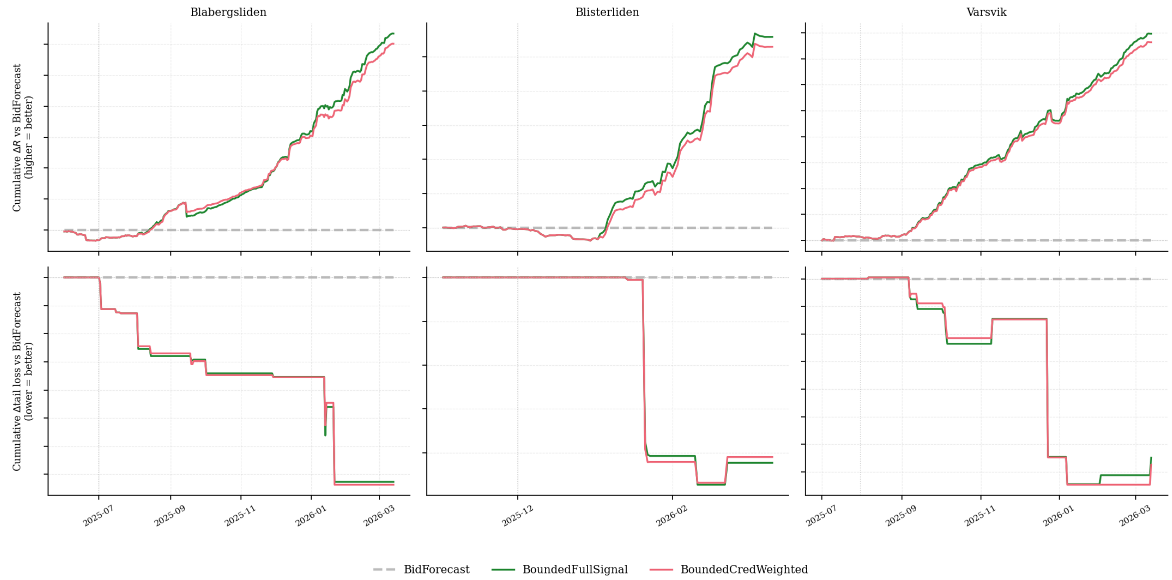

Figure 7.1: ID-stage cumulative revenue and tail-loss trajectories of BCW and BFS relative to the BIDF baseline. *Top*: cumulative ΔR (positive = earns more). *Bottom*: cumulative Δtail loss (lower = less tail damage). Dashed line: warm-up boundary.

ble 7.4). Better forecasting contributes; it does not explain the headline result. The BIDTRUE-PROD (ID) row in Table 7.3 sharpens this: even an idealised intraday update that bids realised availability, the upper bound on any ID-stage production-forecast improvement (Section 3.6), earns less mean revenue than BCW at every park, so the deliberate spread-driven position is a separate source of revenue that no production forecast alone replicates. On the mean–CVaR objective BCW still wins at Blisterliden and Varsvik, but not at Blabergsliden, where perfect production foresight instead collapses the production-forecast tail the spread tilt leaves intact (Section 9.4).

Table 7.4: ID-stage forecast benchmark at $\gamma = 1$: operator ID forecast vs. in-house ID-Linear, per park. MAI: mean absolute imbalance (MWh). $\Delta\bar{R}$: mean daily revenue change vs. the no-ID baseline (% of park baseline). $\mathrm{CVaR}_{0.9}$: CVaR of daily losses (index, baseline = 100 per park).

Park	Forecast	MAI (MWh)	$\bar{R}$	$\Delta\bar{R}$	$\mathrm{CVaR}_{0.9}$
Blabergsliden	Operator	2.786	100.7	0.7	52.9
	ID-Linear	2.414	100.9	0.9	54.0
Blisterliden	Operator	3.072	101.0	1.0	−2.1
	ID-Linear	2.247	102.5	2.5	−2.7
Varsvik	Operator	2.209	99.7	−0.3	−2.8
	ID-Linear	1.342	102.8	2.8	−4.7

The per-interval imbalance-price forecast is the dominant lever. Both parametric candidates climb above every forecast-chase variant from the warm-up boundary onward, and the objective lift over BIDF is far larger than any forecast-quality improvement alone can deliver (Table 7.3). Two checks confirm the lift is a genuine signal effect. First, the SAA LP underperforms both parametric candidates at every park: it carries no representation of *parameter* *risk* — the uncertainty in the spread *estimate* itself (Section 3.5) — treats the scenario mean as truth and over-tilts accordingly, while the parametric band cap and SNR shrinkage regularise against that overcommitment (Appendix C.7). Second, the significance instrument of Section 4.6.4, applied at $\lambda = 0$, returns a 95 % CI excluding zero at every park, including the demanding BCW vs ID-Linear comparison (Appendix C.6). The BCW-on-δ_t^{ID} row of Table 7.3 bounds how much room a sharper spread forecast itself leaves: the counterfactual inherits BCW's walk-forward w_d and bids the production-fan band edge in the realised-spread direction (Section 3.6), so the gap to BCW measures the residual ID-spread headroom at the same risk budget. The residual is narrow at Blabergsliden and material at the two parks with the strongest signal quality (Table 5.3), pointing to ID-spread forecast sharpening as a stage-specific extension where the deployed signal is already strongest.

Why BCW and not BFS. BFS and BCW are statistically indistinguishable on objective. The gap is at most 1 index point and not separable from zero (Table 7.5), consistent with the second-order role of tilt magnitude isolated in Section 8.1.1. BCW is adopted for two reasons. First, it is the rule the Section 3.5 derivation produces from the predictive uncertainty in $\hat{\delta}_t$: the shrinkage γ_t is the explicit parameter-risk control, collapsing the tilt toward the median when the signal is weak. BFS is the $\gamma_t \equiv 1$ ablation that ignores parameter risk. Second, BCW produces up to ∼18 % lower mean absolute imbalance, with the shrinkage active on a substantial fraction of intervals (Table 7.5), a tangible operational benefit at no statistically significant objective cost.

Table 7.5: Head-to-head comparison of BFS and BCW at the ID stage, both with walk-forward calibrated bands. "mean $|\tau - \frac{1}{2}|$" measures bid aggressiveness (larger = bid further from the production-fan median). "frac $\gamma_t < 0.5$" is the share of intervals on which BCW's SNR shrinkage actively suppresses the tilt below half-strength; "frac $\gamma_t > 0.9$" is the share where the signal is strong enough to trigger near-full commitment.

	Blabergsliden		Blisterliden		Varsvik	
Metric	BFS	BCW	BFS	BCW	BFS	BCW
$\mathbb{E}[\bar{R}]$ (index)	112.2	111.5	108.0	107.9	116.2	115.6
$\mathrm{CVaR}_{0.9}$ (index)	+50.2	+49.5	−3.1	−2.8	−12.9	−12.9
Objective (index)	62.0	62.1	111.1	110.7	129.2	128.5
MAI (MWh / 15-min slot)	4.302	3.531	3.275	3.163	2.622	2.374
nMAI (% of $P^{\max}$)	12.0	9.9	13.4	12.9	20.6	18.6
mean $\lvert\tau - \frac{1}{2}\rvert$	0.372	0.258	0.304	0.273	0.369	0.308
mean γ_t (BCW only)	1.000	0.219	1.000	0.610	1.000	0.328
frac $\gamma_t < 0.5$ (%)	0.0	89.2	0.0	43.2	0.0	78.4
frac $\gamma_t > 0.9$ (%)	100.0	5.4	100.0	38.9	100.0	11.7

7.4 Setpoint

The setpoint signal is the strongest in the pipeline: sign accuracy > 70 % and peak correlation across the three stages (Table 5.3), strong enough to act on directly. The decision space also differs from DA and ID: the producer selects a *setpoint* $s_t \in [0, P^{\max}]$ to which the asset is curtailed, not a quantile of the production fan. Of the candidates in Table 7.6, the one adopted in the proposed pipeline is the discrete classifier ThreeRegion. It wins at every park on the mean–CVaR objective and has the lowest $\mathrm{CVaR}_{0.9}$ where SP candidates carry material tail exposure: the fee-band threshold $\pm f^{\mathrm{imb}}$ is the correct decision boundary, and holding the committed bid in the interior band is empirically valuable.

The baseline is NoCurtailment ($s_t = P^{\max}$ unconditionally); every other candidate acts on the SP-stage imbalance-price forecast $\hat{\pi}_t^{\mathrm{IM,SP}}$ — the realised DA price plus the published $\hat{s}_t^{\mathrm{IM,SP}}$, which forecasts $\pi_t^{\mathrm{IM}} - \pi_t^{\mathrm{DA}}$. The adopted ThreeRegion rule is the direct instantiation of Proposition 3.3,

$$s_t = \begin{cases} P^{\max} & \text{if } \hat{\pi}_t^{\mathrm{IM,SP}} > +f^{\mathrm{imb}}, \\ b_t & \text{if } |\hat{\pi}_t^{\mathrm{IM,SP}}| \le f^{\mathrm{imb}}, \\ 0 & \text{if } \hat{\pi}_t^{\mathrm{IM,SP}} < -f^{\mathrm{imb}}. \end{cases} \tag{7.5}$$

The fee thresholds $\pm f^{\mathrm{imb}}$ are break-even boundaries: produce at capacity when the expected price clears the fee, and hold the committed bid in the interior band where any deviation costs the fee in expectation. Below $-f^{\mathrm{imb}}$, curtail to zero. A deliberate shortfall earns positive imbalance revenue net of the fee.

Four non-adopted variants probe the rule's structure. CurtailToBid and CurtailToZero trigger at $\hat{\pi}_t < 0$ rather than at $-f^{\mathrm{imb}}$ (routing to b_t and 0 respectively), and SoftAnchored replaces the hard switch with a calibrated tanh ramp without interior hold band; it collapses to CurtailToZero rather than (7.5) even as $k \to \infty$ (Appendix C.2). The fourth, SAA LP, drops the conditional-independence hypothesis $\pi_t^{\mathrm{IM}} \perp a_t \mid \mathcal{F}^{\mathrm{SP}}$ behind Proposition 3.3: it picks s_t per slot by mean–CVaR grid search on the joint $(a_t, \pi_t^{\mathrm{IM}})$ scenario fan of Chapter 6, with per-slot price scenarios re-centred on $\hat{\pi}_t^{\mathrm{IM,SP}}$ so the price input matches the one ThreeRegion reads. It beats (7.5) only if the joint structure the fan carries is decision-relevant; whether that holds is the empirical question the SAA LP row in Table 7.6 answers, with two further tests in Appendix C.9.

Table 7.6: SP-stage candidates: mean daily revenue $\bar{R}$, $\mathrm{CVaR}_{0.9}$, and mean–CVaR objective ($\lambda = 1$), all in index units (baseline = 100 per park). A star ($\star$) marks the per-park objective winner; the dagger (†) marks the candidate adopted in the proposed pipeline.

	Blabergsliden			Blisterliden			Varsvik		
Candidate	$\bar{R}$	CVaR	Obj	$\bar{R}$	CVaR	Obj	$\bar{R}$	CVaR	Obj
NoCurtailment	100.0	58.5	41.5	100.0	−0.0	100.0	100.0	−1.5	101.5
CurtailToBid	103.7	33.6	70.1	100.7	−0.1	100.8	100.7	−3.6	104.3
CurtailToZero	106.5	32.3	74.2	101.3	0.6	100.7	101.3	−3.4	104.7
ThreeRegion†	106.7	32.1	74.6*	101.3	0.4	100.9*	101.4	−3.4	104.8*
SoftAnchored	106.6	32.3	74.4	101.4	0.5	100.9	101.2	−3.5	104.7
SAA LP	86.5	75.5	11.1	99.5	−0.0	99.6	86.4	8.4	77.9
3R on π_t^{IM}	111.7	29.7	81.9	103.9	−0.8	104.6	103.9	−4.9	108.9

CurtailToZero is essentially tied with ThreeRegion: the two diverge only in the negative sub-band $-f^{\mathrm{imb}} < \hat{\pi}_t < 0$, where CurtailToZero curtails (incurring the fee) while ThreeRegion holds (no fee). The near-tie says these intervals are rare: when negative-price events occur, the SP forecast clears $-f^{\mathrm{imb}}$ comfortably. CurtailToBid loses because

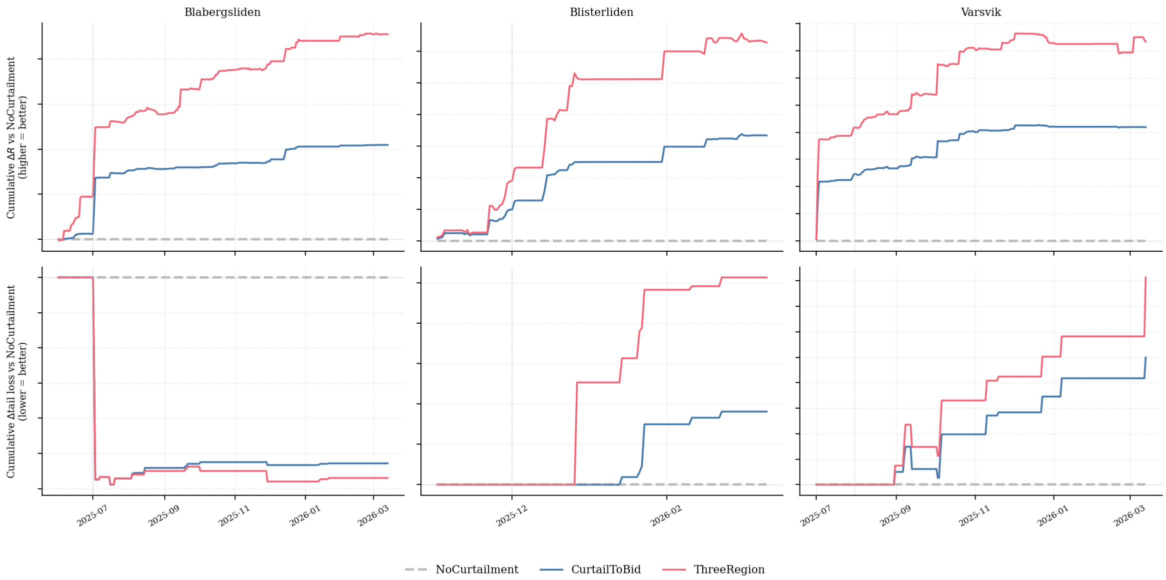

Figure 7.2: SP-stage cumulative revenue and tail-loss trajectories of ThreeRegion and CurtailToBid relative to the NoCurtailment baseline. *Top*: cumulative ΔR (positive = earns more). *Bottom*: cumulative Δtail loss (lower = less tail damage). Dashed line: warm-up boundary.

it stops overproduction without claiming the deliberate-shortfall revenue. SoftAnchored loses everywhere because it lacks the interior hold band: the ramp deviates from b_t inside $\pm f^{\text{imb}}$ and pays the fee where ThreeRegion holds fee-free (Appendix C.2). SAA LP also underperforms ThreeRegion at every park: the joint (a_t, π_t^{IM}) structure the fan carries is too noisy at the SP horizon to beat the three-corner discrete rule, which serves as a regulariser against scenario-estimation noise, the same mechanism as the SAA LP-versus-BCW contrast at ID (Section 9.2). The bulk of the gap is the empirical $\text{CVaR}_{0.9}$ estimated on only 25 tail scenarios at $S = 250$; a risk-neutral run shrinks the gap to within 3.4 index points at every park (Section C.9.1). The 3R-on-π_t^{IM} row of Table 7.6, ThreeRegion routed through the realised π_t^{IM} rather than the forecast (Section 3.6), sits only marginally above the deployed rule at every park: with three discrete actions and decision thresholds at $\pm f^{\text{imb}}$, the rule already captures most of what perfect SP-price knowledge could give, so SP-side forecast sharpening is not a high-leverage extension.

The value of curtailment is concentrated at Blabergsliden, whose SP-stage lift over NoCurtailment is roughly an order of magnitude larger than at the other two parks (Table 7.6). Its NoCurtailment curve diverges sharply on several negative-price events that ThreeRegion avoids, whereas the other parks' tail damage is small even without curtailment (Figure 7.2). The concentration tracks each park's evaluation window, not its bidding zone. ThreeRegion is identical across parks; its value is bounded by how often the trigger fires. Negative-price intervals concentrate almost entirely in the SE1 summer-2025 episodes covered by Blabergsliden's window. Blisterliden is also in SE1 but its post-warm-up window starts later and misses most of those episodes. Varsvik is in SE3, where negative-price intervals are markedly rarer.

7.5 The Composed Pipeline on a Single Day

The three sections above each isolated one stage against its own baseline. Figure 7.3 shows what the composed BidF–BCW–ThreeRegion pipeline looks like in practice, tracing all three rules across a single day at Blabergsliden (17 Feb 2026, $\Delta R_d = +86$ index points over baseline).

The four callouts in panel (a) cover every operating regime within this one day. The bid tilts above the ID-Linear fan median when $\hat{\delta}_t^{\mathrm{ID}} > 0$, below it when $\hat{\delta}_t^{\mathrm{ID}} < 0$, holds at the median when SNR shrinkage collapses the tilt to zero, and is forced to zero during the late-afternoon $\hat{\pi}_t^{\mathrm{IM,SP}} < -f^{\mathrm{imb}}$ curtailment episode. Panel (b) shows the ID spread forecast tracking the realised spread in sign on most intervals; the SP imbalance-price forecast falls well below $-f^{\mathrm{imb}}$ in the curtailment window, confirming the curtailment fired on a true negative-price event. Panel (d) attributes the +92 index-point lift at the interval level: the curtailment window contributes the bulk, with a long tail of small ID-tilt gains accumulating across the remainder of the day. Panels (e) and (f) tie every visible bid feature back to its rule — τ_t tracking within the calibrated band for BCW, the three-mode state strip for ThreeRegion.

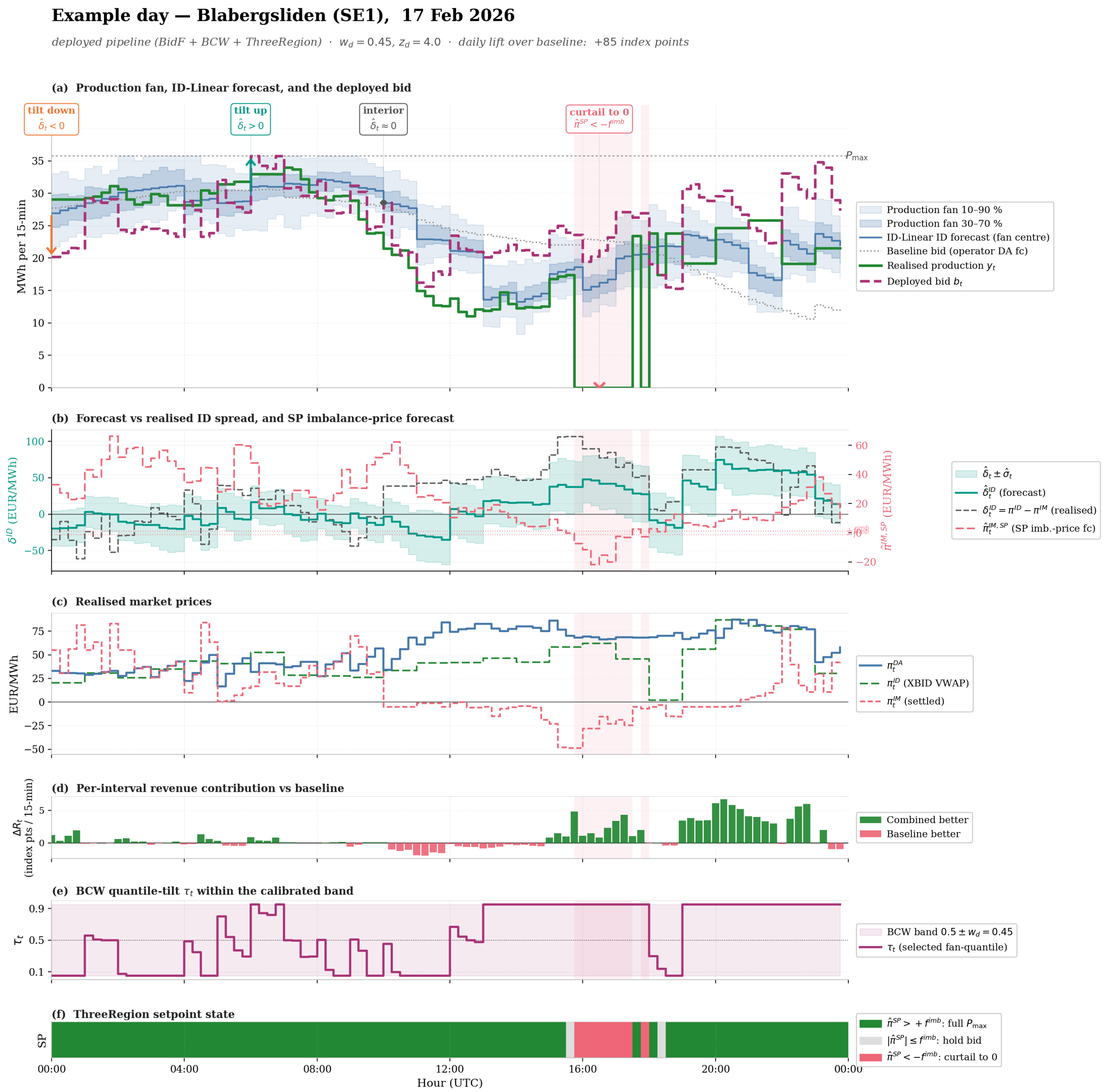

Figure 7.3: All three pipeline stages on a single representative day at Blabergsliden (17 Feb 2026). Panel **(a)**: production-fan bands (10–90 % and 30–70 %) around the ID-Linear ID forecast, realised production y_t, baseline bid, and the final pipeline bid b_t; callouts mark one interval each of *tilt up*, *tilt down*, *interior* (SNR shrinkage to zero), and *curtail to 0*. Panel **(b)**: ID spread forecast $\hat{\delta}_t^{\text{ID}}$ vs realised δ_t^{ID} (left axis) and SP imbalance-price forecast $\hat{\pi}_t^{\text{IM,SP}}$ (right axis). Panel **(c)**: realised market prices. Panel **(d)**: per-interval revenue difference ΔR_t vs baseline (green = pipeline better). Panel **(e)**: BCW quantile τ_t inside the walk-forward calibrated band $0.5 \pm w_d$. Panel **(f)**: ThreeRegion setpoint mode. A faint red band spans the curtailment window across all panels. Best viewed in the digital version.

Chapter 8

Results

Chapter 7 selected one rule per stage by isolating each stage against its own baseline. This chapter composes the per-stage winners into the proposed BIDF–BCW–THREEREGION pipeline and reports its empirical performance (Section 8.1), including the only remaining operational lever, the DA bid size itself (Section 8.2.1). Section 8.2 then checks whether the headline lift survives sampling uncertainty and realistic execution conditions.

8.1 Combined Signal-Based Pipeline

The proposed pipeline (BIDF–BCW–THREEREGION) lifts the mean–CVaR objective at all three parks and turns $\text{CVaR}_{0.9}$ negative at two of them. Table 8.1 reports indexed metrics for the three stages of the build-up (Baseline → +ID → +ID+SP), configuration held constant across all three.

Table 8.1: Combined pipeline results on the post-warm-up evaluation window. All monetary metrics in index units (baseline $\bar{R} = 100$ per park). $\bar{R}$: mean daily revenue index. CVaR: $\text{CVaR}_{0.9}$ of daily losses (positive = worst-decile loss; negative = profitable worst decile). Obj: $\bar{R} - \text{CVaR}_{0.9}$.

Park	Pipeline	$\bar{R}$	CVaR	Obj
Blabergsliden	Baseline	100.0	58.5	41.5
	+ ID	111.5	49.5	62.1
	+ ID + SP	118.2	23.8	94.4
Blisterliden	Baseline	100.0	−0.0	100.0
	+ ID	107.9	−2.8	110.7
	+ ID + SP	109.3	−2.4	111.6
Varsvik	Baseline	100.0	−1.5	101.5
	+ ID	115.6	−12.9	128.5
	+ ID + SP	117.0	−13.6	130.5

Table 8.1 shows the ID stage carrying the lift at every park; at Blisterliden and Varsvik the SP contribution falls within sampling uncertainty (Section 4.6.4). THREEREGION is identical across parks, so its value tracks how often the negative-price trigger fires in each evaluation window, not any park-specific property of the rule.

8.1.1 Band Half-Width Sensitivity

The headline numbers in Table 8.1 report the proposed pipeline at the walk-forward calibrated $(\hat{w}_d, \hat{z}_d)$ point. Sweeping the BCW band half-width w over $[0.05, 0.45]$ at fixed $z = 2$ (the calibration default of Equation (7.3)) shows both that the lift does not hinge on that calibration and, more pointedly, where the lift comes from. Figure 8.1 plots the resulting $(\mathrm{CVaR}_{0.9}, \mathbb{E}[\bar{R}_d])$ frontier per park, on the same total-revenue mean–CVaR axes used elsewhere in the chapter; Table 8.2 reports, at every w, the objective lift over the BIDF baseline alongside the deliberate-imbalance footprint (nMAI, defined in Equation (4.5)).

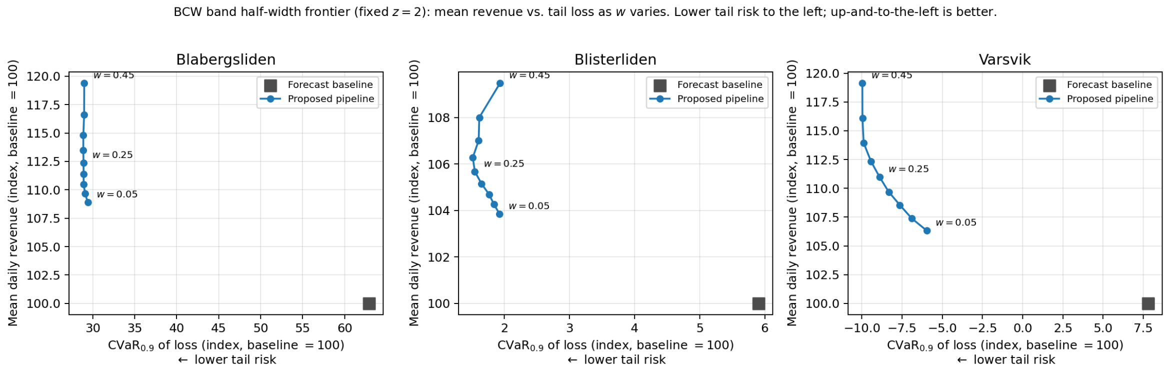

Figure 8.1: Revenue–risk frontier of the proposed pipeline as the BCW band half-width w varies, at fixed $z = 2$. Each park is one panel; each blue marker is one $w \in \{0.05, 0.10, \ldots, 0.45\}$; the grey square is the forecast baseline. Lower tail risk is to the left, so up-and-to-the-left is strictly better.

Table 8.2: Objective lift over the BIDF baseline and deliberate-imbalance footprint (nMAI) at each BCW band half-width w, fixed $z = 2$, post-warm-up window.

	Blabergsliden		Blisterliden		Varsvik	
w	ΔObj (index)	nMAI (%)	ΔObj (index)	nMAI (%)	ΔObj (index)	nMAI (%)
baseline	—	9.4	—	13.8	—	14.0
0.05	+42.4	9.6	+7.8	12.2	+20.1	13.5
0.1	+43.5	9.7	+8.3	12.3	+22.1	13.7
0.15	+44.5	9.8	+8.8	12.4	+24.0	14.1
0.2	+45.4	10.0	+9.4	12.5	+25.8	14.6
0.25	+46.4	10.3	+10.0	12.8	+27.7	15.2
0.3	+47.5	10.6	+10.7	13.0	+29.6	16.0
0.35	+48.8	11.1	+11.3	13.4	+31.6	17.2
0.4	+50.5	11.7	+12.3	14.1	+33.9	18.9
$0.45^{\dagger}$	$+53.3^{\star}$	12.8	$+13.5^{\star}$	15.2	$+36.9^{\star}$	21.6

Baseline = BIDF (DA forecast bid, no ID/SP action); $\Delta\mathrm{Obj} = \bar{R} - \mathrm{CVaR}_{0.9}$ lift over baseline (index, baseline = 100 per park). nMAI is post-warm-up, % of $P^{\max}$, on the identical scale as Table 8.3. $z = 2$ fixed. † calibrated operating band; * best objective per park.

The pipeline beats the forecast baseline on *both* revenue and tail loss at every w and every park; no band width reverses the result. The substantive finding is that the benefit is bought almost entirely by smaller tilts. At the narrowest band, $w = 0.05$, the pipeline already captures the majority of the maximum mean–CVaR lift and essentially the entire tail-risk improvement (Figure 8.1). This comes at *no net cost in deliberate imbalance*: at $w = 0.05$ the nMAI is essentially unchanged from the BIDF baseline at Blabergsliden and *below* it at Blisterliden and Varsvik, because the ID-Linear centre removes more forecast-error imbalance than the small tilt adds (Table 8.2). Widening the band buys only the remaining objective increment, and pays for it in deliberate imbalance volume that rises monotonically with w. The band half-width is thus a *revenue dial bought with imbalance volume*; the tail-risk benefit is free, secured at the narrowest setting.

The separation is structural: w caps only the *magnitude* of the quantile departure from the fan median, while the *sign* of each tilt, and the decision to tilt at all, is governed by $\hat{\delta}_t^{\mathrm{ID}}$ and the SNR shrinkage (7.2), neither of which depends on w. Capturing the correct *side* of imbalance settlement is the first-order effect; band width only scales second-order volume around it.[1]

The criterion that matters for the ID stage is therefore the sign of the *realised* spread, not the sign of π^{IM}: an adjustment earns precisely when $\operatorname{sign} \Delta b_t = \operatorname{sign} \delta_t^{\mathrm{ID}}$, with per-MWh P&L $\delta_t^{\mathrm{ID}}\, \Delta b_t$. By that measure, roughly two-thirds of the intraday-adjustment volume lands on the right side of the spread and the gross spread P&L is positive at every park, consistent with the ID sign accuracy of Chapter 5 established *before* any bidding. The tilt positions the producer on the right side of the *spread* it trades, not the imbalance settlement it deliberately accepts.

8.1.2 Imbalance Footprint and Economics

The BIDF-only baseline conflates forecast improvement with deliberate signal trading; Table 8.3 separates the two by comparing three pipelines: BIDF-only, an ID-Linear-corrected forecast chase without any tilt, and the full proposed pipeline.

Table 8.3: Imbalance footprint and revenue across three pipelines. BIDF: DA forecast bid, no ID or SP adjustment. ID-Linear-corrected: BIDF + ID-Linear forecast chase at ID, no spread tilt or curtailment. Full: the proposed BIDF–BCW–THREEREGION pipeline. nMAI: mean absolute imbalance as a percentage of $P^{\max}$ (Equation (4.5)).

	nMAI (% of $P^{\max}$)			Daily revenue $\bar{R}$ (index)		
Park	BidF	ID-Lin-corr.	Full	BidF	ID-Lin-corr.	Full
Blabergsliden	9.4	6.8	12.4	100.0	101.0	117.9
Blisterliden	13.8	9.2	15.6	100.0	102.8	109.5
Varsvik	14.0	10.5	21.1	100.0	102.9	116.9
Decomposition of the change (index, MWh per 15-min slot):						
	Better-forecasting (BidF → ID-Lin-corr.)			Deliberate (ID-Lin-corr. → Full)		
Park	ΔnMAI	ΔMAI	$\Delta\bar{R}$	ΔnMAI	ΔMAI	$\Delta\bar{R}$
Blabergsliden	−2.7	−0.95	+0.9	+5.6	+2.02	+17.2
Blisterliden	−4.6	−1.13	+2.5	+6.4	+1.56	+6.7
Varsvik	−3.4	−0.44	+2.8	+10.6	+1.35	+14.3

Updating to the ID-Linear forecast lowers nMAI and lifts revenue: better forecasting alone reduces imbalance exposure. Layering BCW and THREEREGION on top then raises nMAI deliberately, at a net revenue per MWh of added imbalance that is order tens of EUR/MWh against the $f^{\mathrm{imb}} = 1.15$ EUR/MWh settlement fee (Table 8.3).

[1]The robustness of this result to the z parameter is documented in Appendix C.4.

That added volume is controlled: the adverse footprint is essentially unchanged from the BIDF baseline at both SE1 parks and grows only at Varsvik. Figure 8.2 breaks the intraday spread leg into its two directions. Both earn a positive net, with the short direction carrying the larger payoff because the negative-imbalance-price tail makes the realised spread largest exactly when the rule is short.

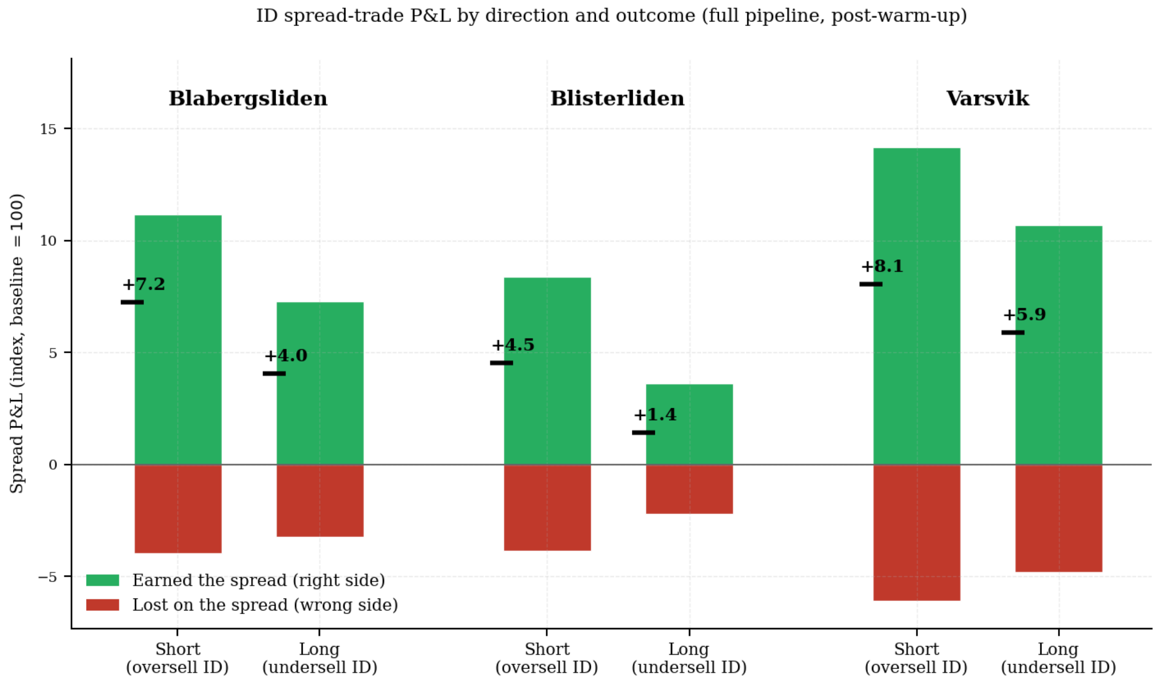

Figure 8.2: ID spread-trade P&L by direction and outcome, full pipeline, post-warm-up. *Short*: the rule bid above its IDLinear production forecast (a deliberate over-commit, settled back at π^{IM}); *long*: bid below it. Green is P&L on intervals where the realised spread $\delta^{\mathrm{ID}} = \pi^{\mathrm{ID}} - \pi^{\mathrm{IM}}$ paid the direction (right side), red where it did not; the black tick marks each direction's net. The reference is the IDLinear-corrected bid, so forecast accuracy is netted out and short + long + re-centring equals the deliberate-trade ID-leg in the imbalance footprint table (residual zero).

8.1.3 Worst-Day Recovery and Residual Exposure

On the three worst baseline days per park, the proposed pipeline recovers value almost everywhere. Blisterliden's two largest losses reduce to near-zero, and all three of Varsvik's loss days flip to gains. At Blabergsliden, two of the three losses are recovered substantially, but on one day in January 2026 the pipeline amplifies the loss: the intraday spread signal misfired ahead of a production collapse (Section 9.4 examines this residual mode).

One structural failure mode persists. Figure 8.3 uses the same wrong-side (red) / right-side (green) encoding as Figure 4.5, applied to the full-pipeline daily outcomes, so the resolution of the two modes identified in Section 4.7 reads directly off the figure. The long-into-negative-IM wrong-side region (lower right, $\nu_d > 0$, $\bar{\pi}_d^{\mathrm{IM}} < 0$) is empty at all three parks: the SP-stage classifier curtails on exactly those intervals and removes that exposure by construction. The residual worst-decile loss is then essentially all short-into-positive-IM (upper left, $\nu_d < 0$,

$\bar{\pi}_d^{\mathrm{IM}} > 0$): underproducing into a positive (typically elevated) imbalance price. Every day the pipeline cannot fully recover sits in the same wrong-side cell, by the same mechanism.

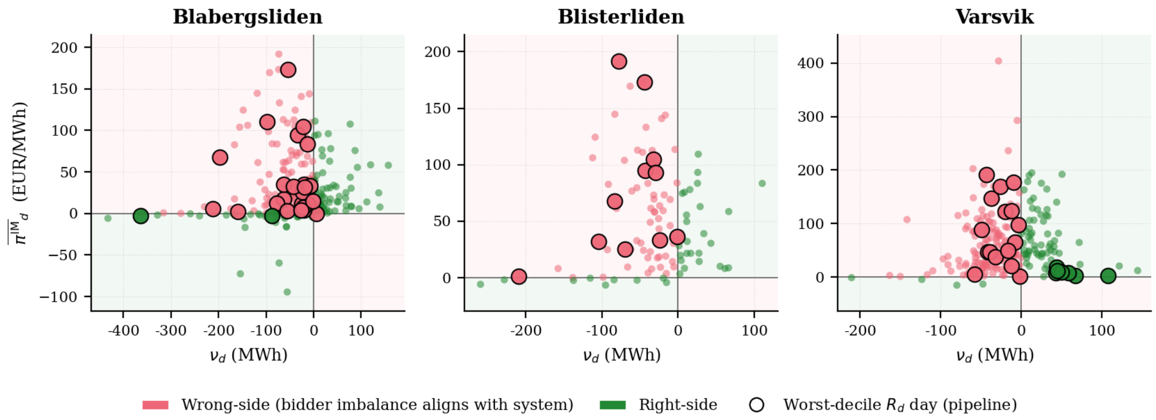

Figure 8.3: Full-pipeline (BIDF–BCW–THREEREGION) daily outcomes in $(\nu_d, \bar{\pi}_d^{\mathrm{IM}})$-space; same wrong-side (red) / right-side (green) encoding as Figure 4.5. Outlined markers: worst-decile R_d days (pipeline).

Table 8.4 decomposes those short-into-positive-IM worst-decile days. The only real residual loss is at Blabergsliden. On the large majority of those days, the production forecast was too high even with the in-house centre and no tilt, accounting for essentially the whole loss. The spread tilt was on average mildly protective rather than the cause, and the case where forecast and tilt were both wrong is a minority. At Blisterliden and Varsvik, the short-into-positive-IM worst-decile days are already net positive. The pipeline has closed those tails. The residual is thus a production-forecast problem, concentrated on the days the centre over-predicts.

Table 8.4: Short-into-positive-IM worst-decile days (the $\mathrm{CVaR}_{0.9}$ tail of Table 8.1) decomposed by origin, per park, post-warm-up window. $\bar{R}$: mean daily revenue on those days, signed (negative = loss), with $\bar{R}$ = fc-short + tilt; fc-short is the ID-Linear-corrected revenue (better forecast, no tilt or curtailment) and tilt is the spread-tilt increment. Positive $\bar{R}$ at Blisterliden and Varsvik means the tail is already profitable. The share columns give the percentage of those days that were forecast-high (fc-high), tilt-losing (tilt-lost), or both.

		Share of those days (%)			Mean (index)		
Park	Days	fc-high	tilt-lost	both	$\bar{R}$	fc-short	tilt
Blabergsliden	23	83	39	26	−26.1	−28.5	+2.4
Blisterliden	11	73	64	36	+2.1	+2.4	−0.0
Varsvik	16	100	44	44	+14.0	+12.6	+1.9

8.2 Robustness Checks

8.2.1 DA Bid Size

The pipeline of Section 8.1 is reported at full DA commitment, $b_t^{\mathrm{DA}} = \hat{a}_t^{\mathrm{DA}}$. A producer with stronger tail aversion reaches first for the lever the signal rules do not touch, namely committing less at the day-ahead stage. Conventional risk management treats this as a frontier to ride, shaving the day-ahead position to trade mean revenue for a smaller tail.

Sweeping a uniform DA multiplier $\alpha \in [0,1]$ with $b_t^{\mathrm{DA}} = \alpha\, \hat{a}_t^{\mathrm{DA}}$, while holding BCW at ID and ThreeRegion at SP fixed, draws that frontier explicitly (Figure 8.4).[2]

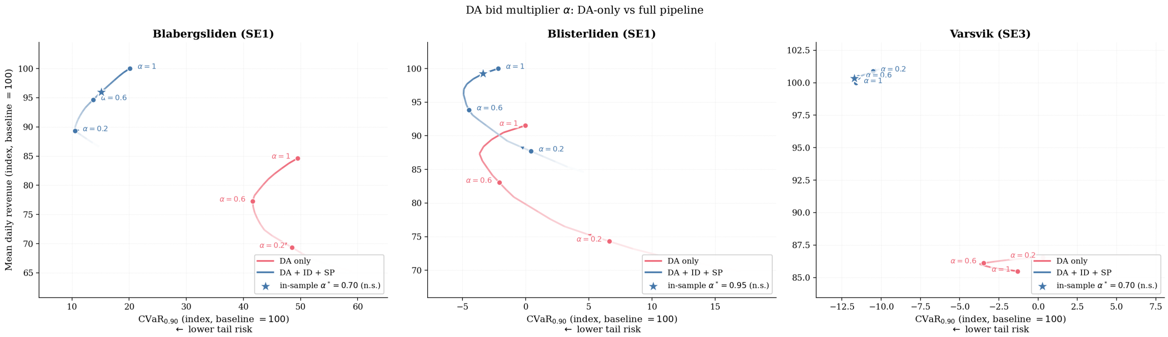

Figure 8.4: Revenue–risk frontier as the DA bid multiplier α sweeps $[0,1]$, per park. *Blue*: full pipeline (DA + ID + SP) at varying α. *Red*: DA-only (no ID or SP rules). Lower tail risk lies to the left; tick annotations mark $\alpha \in \{0.2, 0.4, 0.6, 0.8, 1\}$. The blue star marks the in-sample α^* maximising the mean–CVaR objective on the evaluation window. Viewed best in digital version.

The frontier appears at every park, bending as expected when α falls, except at Varsvik, where the full-pipeline tail is already near zero at $\alpha = 1$ and little is left to compress. The curvature is sharpest where the residual left tail is heaviest, which tracks the window-coverage of negative-price intervals rather than the bidding zone. The frontier is empirically real, but no rule can act on it. A walk-forward replay over $\alpha \in \{0.70, 0.80, 0.85, 0.90, 0.95, 1.00\}$, re-calibrating BCW's (w, z) at each evaluation day, finds no $\alpha < 1$ whose objective gain over $\alpha = 1$ has a 95 % paired-day bootstrap CI excluding zero.[3] Trimming the day-ahead commitment is the one genuine revenue–risk trade-off this thesis exposes: it appears on every panel and is statistically insignificant on all of them.

8.2.2 Statistical Significance of Results

The significance instrument of Section 4.6.4, applied at $\lambda = 1$ to the proposed-vs-baseline mean–CVaR objective lift, gives 95 % CIs whose lower bounds exclude zero at all three parks (Table 8.5), so the consistent-outperformance ranking in Table 8.1 is statistically significant under the exchangeability assumption of Section 4.6.4.

Decomposed by stage, the ID increment is similarly distinguishable from zero at all three parks. The SP increment is not: its CI straddles zero at Blisterliden and Varsvik, consistent with the attribution in Section 8.1 that SP neither helps nor hurts those parks materially in this window.

[2] Here $\alpha \in [0,1]$ is the DA bid multiplier; it is distinct from the CVaR confidence level $\alpha = 0.90$ used throughout the thesis.

[3] The largest in-sample gain over $\alpha = 1$ anywhere on the grid is under 1.5 index points, and the 95 % paired-day bootstrap CI straddles zero at every α and every park; see the starred points in Figure 8.4.

Table 8.5: 95 % paired-day bootstrap CIs (index points) for the mean–CVaR objective lift, post-warm-up evaluation window. Full: proposed pipeline vs. BidF baseline. ID: ID-stage increment. SP: SP-stage increment. Intervals whose lower bound excludes zero are marked *.

Park	Full	ID	SP
Blabergsliden	[+22.2, +92.3]*	[+13.0, +29.0]*	[+4.8, +71.0]*
Blisterliden	[+6.4, +18.3]*	[+5.5, +17.1]*	[−0.6, +3.1]
Varsvik	[+19.3, +40.1]*	[+17.9, +37.2]*	[−0.0, +4.3]

8.2.3 Realistic-Conditions ID Lift

The headline lift figures assume idealised ID execution, the natural ceiling. Two corrections establish a tighter, more defensible bracket. First, a *causally honest* signal: the canonical pipeline forms $\hat{\delta}_t^{\mathrm{ID}}$ from the full-MTU VWAP π_t^{ID}, which is only known after gate closure, so the pipeline substitutes the last-printed trade $\pi_t^{\mathrm{ID,last}}$, observable at forecast arrival time. Second, a *liquidity-honest* cap: hourly absolute adjustments are capped at 20 % of zone-matched SIDC volume, following the sweep in Section 2.5 (tighter caps cost lift; looser caps buy none). Settlement remains at the full-MTU VWAP π_t^{ID} as the price-taker proxy for the marginal counterfactual order. Table 8.6 reports the four configurations $\{\mathrm{VWAP}, \mathrm{Last}\} \times \{\text{no cap}, 20\ \%\ \text{cap}\}$ alongside the BidF anchor, isolating each correction independently. The realistic configuration (*Last + 20 % cap*) outperformed BidF on both axes at all three parks, but the lift bracket is materially tighter than the proposed-pipeline ceiling, with Blisterliden the most affected.

Table 8.6: ID-stage lift under timing- and liquidity-honest corrections. {VWAP, Last} × {no cap, 20 % cap} configurations plus the BidF anchor, post-warm-up evaluation window. Mean revenue ($\bar{R}$), $\mathrm{CVaR}_{0.9}$, and mean–CVaR objective per park.

Park	Configuration	$\mathbb{E}[\bar{R}]$ (index)	$\mathrm{CVaR}_{0.9}$ (index)	ΔObj vs BidF (index)
Blabergsliden	BidF	100.0	58.4	—
	BCW, VWAP signal, no cap (proposed)	111.4	49.4	+20.5
	BCW, VWAP signal, 20% cap	109.4	50.8	+17.1
	BCW, Last signal, no cap	108.6	49.3	+17.8
	BCW, Last signal, 20% cap (*realistic*)	107.3	50.8	+15.0
Blisterliden	BidF	100.0	0.0	—
	BCW, VWAP signal, no cap (proposed)	108.0	−2.8	+10.7
	BCW, VWAP signal, 20% cap	106.0	−2.1	+8.1
	BCW, Last signal, no cap	105.9	−1.9	+7.7
	BCW, Last signal, 20% cap (*realistic*)	104.6	−1.4	+5.9
Varsvik	BidF	100.0	−1.5	—
	BCW, VWAP signal, no cap (proposed)	115.7	−12.9	+27.0
	BCW, VWAP signal, 20% cap	114.7	−12.7	+25.9
	BCW, Last signal, no cap	115.2	−12.9	+26.5
	BCW, Last signal, 20% cap (*realistic*)	114.3	−12.7	+25.5

The realistic configuration retains the majority of the proposed-pipeline objective lift across all three parks, with Varsvik and Blabergsliden robust on both mean revenue and tail compression. Blisterliden is the most sensitive, with a noticeably tighter bracket on both axes.

Chapter 9

Discussion

The central finding of this thesis is affirmative. Acting on the stage-appropriate imbalance-price signal raises mean daily revenue and lowers tail-loss risk at the same time, at every evaluated park, with no trade-off between the two wherever a usable directional signal exists (Table 8.1). This is achieved by adjusting the traded position on the intraday market and curtailing physical output at the setpoint stage when the imbalance price is expected to turn negative. This chapter interprets that result.

9.1 Where the Improvement Comes From

On the intervals where the signal is informative, meaning the intraday spread signal flags that the imbalance price is about to turn against full commitment, bidding the full production forecast is not a cautious choice that occasionally goes wrong. It is the worse choice on both revenue and risk axes at once. When the market is about to pay the producer to hold back, committing full output earns less and, on the same intervals, does the most damage to the tail. Reducing the commitment there does not trade mean revenue for a smaller tail. It removes a recurring wrong-side commitment. This is why no meaningful trade-off appears. The one lever that would be a genuine trade-off, committing less at the day-ahead stage regardless of conditions, delivers no risk-adjusted gain statistically distinguishable from full commitment on this window (Section 8.2.1). It is therefore not the explanation either.

The same reasoning explains why a directional rule at the day-ahead stage adds nothing. A day ahead of delivery, the costly intervals cannot be told apart from ordinary ones, so there is no specific wrong-side commitment to cease. This is not specific to the rules evaluated here. Even the strongest rule in the published one-price literature has, at the day-ahead stage, only the sign of the spread to act on. When that sign is no better than chance, as it is at this horizon, that rule too collapses to bidding the forecast. The upper bound reported for it in (Bruninx et al. 2025) conditions on the realised spread sign, so the DA null result here is consistent with it. A day-ahead estimate of the imbalance spread with sign accuracy materially above one half would change this, but no commercial product supplies one at that horizon. The signal-quality evidence of Chapter 5 sets the bar for any rule that would replace the forecast bid.

9.2 Acting in Proportion to Confidence

The three stage rules look different on paper, but they follow one principle. The form of the rule should track the credibility of the directional signal at that stage, where credibility means the size of the signal measured against the uncertainty of the estimate, not the size alone. This is the principle the thesis calls *signal-quality-aware quantile bidding.* The producer bids the production quantile the signal warrants, scaled by how far that signal can be trusted. It is also the distinction between the two ways of acting on a signal examined earlier. The scenario-based optimiser improves the objective one scenario at a time. Because the imbalance fee is small, the case for sitting anywhere in the interior of the band almost never arises for a risk-neutral optimiser, which therefore pushes to one band edge on virtually every interval, committing fully in the direction of the estimated spread (Appendix C.7). The decisive point is not that this optimiser is risk-neutral. It carries no representation of how uncertain the spread estimate itself is, and it treats the scenario draws as the truth rather than as a noisy estimate of a conditional spread mean. Risk aversion does not repair this. Penalising the revenue tail penalises bad revenue outcomes, not the estimator's uncertainty, and turning up the weight does pull the intraday optimiser off the edge, to cap the revenue tail, not because the spread estimate is unreliable (Appendix C.7). This is the distinction between risk in the realised outcome and risk in the estimate itself (Bannör and Scherer 2013). The credibility-weighted rule reads the same point estimate but also consults how noisy that estimate is, the parameter risk the optimiser ignores, abstaining on weak slots and committing only on confident ones.

Empirically, the flexible optimiser loses at both gates where it has a structured rival, at every park: the scenario optimiser is beaten by the credibility-weighted rule at the intraday stage, and a scenario rule that drops the independence assumption and optimises on the joint fan is beaten by the discrete classifier at the setpoint stage (Appendix C.9). Neither shape was imposed in advance; each followed from the signal's credibility at its gate. Both defeats reflect one principle. The settlement structure is known with certainty, its spread sign, fee threshold $\pm f^{\text{imb}}$, and newsvendor quantile all fixed in advance, so building it in regularises the decision against the estimation noise that dominates a near-chance signal and a short, heavy-tailed record, where a flexible optimiser instead overfits. This is regime-specific: with a strong signal and deep data those same constraints would bind, and a flexible rule could recover what they discard.

The intraday lift does not hinge on the exact weighting. The adopted rule and its no-shrinkage corner case are tied within the daily-revenue standard error at every park (Section 7.3), and even the narrowest band recovers most of it at no extra imbalance (Section 8.1.1). The gain is therefore the disciplined-band principle itself, not a finely tuned rule. Credibility weighting earns its place not on headline revenue but as the rule the derivation produces, with the lighter imbalance footprint (Section 7.3).

The practical implication follows directly. A staged bidding pipeline should be designed by first auditing signal quality at each stage, and only then choosing how to act on it. The audit decides which stages can carry a directional rule at all. The decision rules are closed-form and auditable rather than learned, fixed by the propositions from the spread signal and the fee with only two scalars calibrated walk-forward.

A complementary isolation makes the deliberate position concrete from the opposite direction. An idealised counterfactual with perfect production foresight at the intraday gate but no spread position (BidTrueProd (ID) row in Table 7.3) still earns less mean revenue than the proposed pipeline at every park: the spread tilt is a separate source of revenue that even a perfect production forecast does not capture, not a substitute for forecast quality.

9.3 Why the In-House Forecast Wins

One input to the pipeline is worth a remark of its own, because it is a change with the large return for a low effort. The production forecast used at the intraday stage is a deliberately simple in-house model: a few autoregressive lags of recent available production, short rolling statistics, and the day-ahead forecast as a level anchor. Despite that simplicity it is not merely competitive with the commercial product it replaces but strictly more accurate, at every evaluated park and on every conventional score, with lower MAE and RMSE, smaller bias, and higher correlation with realised production (Table 5.2). The gap is largest at Varsvik, where the in-house centre nearly halves the operator error. That accuracy advantage carries into the bidding outcome, lifting mean revenue at every park (Table 7.4).

The win is not uniform on risk, and the exception is informative. At the two parks whose tails are mild the sharper centre also lowers $\mathrm{CVaR}_{0.9}$, but at Blabergsliden it does not (Table 7.4). The worst-day decomposition locates why: that park's tail is set by days on which the centre over-predicts and production then collapses, and a forecast swap that does not catch those days leaves the exposure in place (Section 8.1.3). The in-house model is essentially unbiased on average (Table 5.2); its limitation is this conditional, per-regime miss on rare collapses, not a systematic bias (Section 6.4). Closing that conditional miss, the pipeline's main residual exposure, is the highest-leverage extension a future iteration has (Sections 9.4 and 10.4).

9.4 What the Pipeline Does Not Model

The pipeline never targets the tail directly. It is a chain of directional rules, and whatever tail control it achieves is a by-product of getting the direction right, not something the objective sets out to optimise. Two things follow, both modelling omissions rather than tuning failures: the pipeline does not model uncertainty in the production centre, and it does not model the decision as a multistage problem with recourse.

The first omission is where the loss concentrates. The optimal bid reads two inputs (Propositions 3.1 and 3.2), a production centre and a spread tilt, but only the tilt is risk-weighted: the credibility weight pulls the tilt back on a weak signal, while nothing pulls back the commitment when the centre is wrong. The model has no representation that the centre itself can fail. And it does, in one specific way. The residual worst-decile days are almost all the producer short into a positive imbalance price, the opposite of the mode the setpoint rule removes, concentrated at the most exposed park in a handful of sudden production collapses (Section 8.1.3). There the centre, not the tilt, is the culprit: calibrated on average but heavy-tailed, it over-predicts sharply on the collapses it cannot see coming, and the bid sits above output that then falls away (Section 6.4). The failure is in the tail, not the mean, so a re-centred mean cannot help; nor would a richer dependence model, since that dependence is weak and the rule built to exploit it loses to the one that ignores it (Appendix C.9). What would help is a centre that widens its own fan when it is unreliable, letting the rule pull the commitment back as it already does the tilt (Section 10.4). This omission is the benign one: it shrinks as the production forecast sharpens.

The second omission is structural. The stages are solved one at a time, here and now, exact at the day-ahead and intraday gates, but not at the setpoint, where curtailment changes the very production the earlier stage treated as fixed (Remark 3.2). The setpoint is really a recourse option, the right to curtail after seeing more, and an option has value before it is exercised: a bid that priced it could lean further than a rule reading the spread alone would dare. The pipeline never prices it. Nothing in the thesis is recourse-aware. Closing this means a genuine multistage formulation with the setpoint as recourse (Section 10.4); unlike the forecast omission, no improvement in forecasting removes it.

9.5 Limitations and External Validity

The results are a snapshot, a short, multi-month window at three parks under the post-go-live mFRR EAM regime. The setpoint value is the most regime-contingent component, its payoff riding on the frequency of extreme negative-price events; the intraday rule earns from the spread present on every interval rather than from the frequency of rare events, so it should be the more transferable part, and its lift survives timing- and liquidity-honest execution (Section 8.2.3). The methodology carries; the magnitudes do not.

The analysis is also of a single price-taker: were many producers to act on the same forecast, their common response would move the prices the rule reads. The erosion is gradual. Each contributes only its deliberate imbalance, smallest at the conservative band the pipeline favours (Section 8.1.1), and the aggregate equilibrium is the natural next question (Section 10.4).

Chapter 10

Conclusion

This chapter answers the research questions of Section 1.3 from the structural and empirical findings, and then sets out the open problems that remain.

10.1 Answers to the Research Questions

The overarching answer is affirmative: acting on the stage-appropriate imbalance-price signal improves the revenue–risk trade-off at every evaluated park, raising mean revenue and lowering tail loss together wherever a usable signal exists.

RQ1, forecast-bid revenue and risk, and where losses concentrate. Bidding the operator's day-ahead production forecast with no intraday correction or setpoint curtailment is profitable on average at all three parks but carries a sharply asymmetric tail (Section 4.7). The worst-decile losses concentrate at Blabergsliden and are wrong-side settlements of two kinds: the dominant one, common to every park, is underproducing into a positive imbalance-price spike when the forecast over-predicts wind; the second, which the other two parks largely escape, is overproducing into a negative imbalance price on high-wind hours, rarer but concentrated where negative prices are common.

RQ2, directional content of the imbalance-price forecast at each stage, and how to use it. The directional content rises monotonically across the stages (Table 5.3). At day-ahead the spread forecast carries none: no per-interval rule on it beats the forecast bid (Table 5.1). At intraday it is a genuine but noisy signal, sign-accurate on roughly two-thirds of non-neutral intervals, and the per-interval scale read from the 10/90 fan lets the bid tilt in proportion to its signal-to-noise ratio (BCW, Equation (7.2)). At the setpoint stage it is strong enough for a discrete fee-band classifier (ThreeRegion), which curtails when the predicted setpoint imbalance price falls below the negative fee. The translation follows the credibility: no rule, a shrunk tilt, a hard classifier.

RQ3, how mean–CVaR improvement accumulates, and where residual exposure lives. Each added stage improves the mean–CVaR objective at every park, and the improvement is statistically significant throughout (Table 8.1 and Section 8.2): mean revenue rises and worst-decile loss compresses at the same time, with no Pareto trade-off anywhere a directional signal is available. The one genuine frontier, how much of the forecast to commit at day-ahead, yields no $\alpha < 1$ whose objective gain is distinguishable from full commitment on this window (Section 8.2.1), so it is theoretical only here. The residual exposure after the full pipeline lives almost entirely in worst-decile days that underproduce into a positive imbalance-price spike, where the production centre over-predicted while the spread tilt was small or protective (Table 8.4). It is a production-forecast residual, not a joint-signal one: no causal intraday rule corrects a centre that over-predicts a collapse, so further reduction must come from a sharper, uncertainty-quantified production forecast (Sections 9.4 and 10.4).

10.2 Contributions

The thesis makes four main contributions, stated in full in Section 1.4 and instantiated by the answers above.

- **Theoretical.** A unified single-price newsvendor structure that resolves the all-or-nothing degeneracy from inside the model. The closest single-price comparators find the profit-maximising bid collapses to a bang-bang corner and patch it externally, with hard position limits or learned policies. Propositions 3.1 to 3.3 show that the per-MWh eSett fee instead creates an interior regime with a closed-form optimum (Chapter 3). Every stage reduces to one comparison of an expected price signal against the fee, a τ_t^*-quantile of the predictive production distribution at DA and ID and the same three-region form at SP. The eSett fee, not the dual-price spread asymmetry, is the structural boundary of that regime.
- **Framework.** A stage-wise signal-credibility audit that turns the propositions into a rule-selection procedure. Credibility is the size of the directional signal measured against the uncertainty of its estimate, not the size alone, and the audit is run per stage before any bidding experiment. It predicts rather than imposes the rule shape, the plain forecast bid where the signal is noise, a bounded tilt where it is moderate, and a discrete classifier where it is strong.
- **Methodological.** A bounded credibility-weighted tilt on the ID spread signal (Equation (7.2)) that raises mean revenue and reduces $\mathrm{CVaR}_{0.9}$ simultaneously at all three parks, robust rather than tuned. Its no-shrinkage ablation BFS is statistically indistinguishable on the objective, so BCW is adopted for a materially lower imbalance footprint and graceful degradation as the signal weakens, not for headline revenue.
- **Empirical.** A stage-by-stage validation on data from three Swedish parks. At *Intraday* a simple in-house linear centre (ID-Linear, Appendix A.1) beats the operator's commercial ID forecast on accuracy and mean revenue at every park (Table 7.4). The credibility-weighted tilt raises mean revenue and compresses $\mathrm{CVaR}_{0.9}$ simultaneously, significant at every park and robust in that it stays positive even at the narrowest band, where it adds no imbalance beyond the forecast baseline (Section 8.1.1). At the *Set-*

point stage a discrete fee-band classifier closes the negative-price wrong-side tail where the evaluation window exposes it. In *combination* they simultaneously cause revenue and tail-risk improvement with no Pareto trade-off, significant at all three parks and preserved under timing- and liquidity-honest execution (Section 8.2.3).

10.3 Practical Takeaway

For a Nordic wind producer carrying direct imbalance exposure, the result is rather practical. Three changes, applied in order and none needing a stochastic solver, raise mean daily revenue and compress the worst-case loss at the same time at every park studied. The single highest-leverage lever is participating in the Intraday market with a reliable production forecast. Moreover, this thesis showed that even a deliberately simple in-house linear model, refitted weekly on a walk-forward loop, improved on objective metrics over the commercial intraday forecast Holmen was using. Bringing that forecast in-house is the first and cheapest move, and sharpening it further is where the largest remaining gain lies. On top of it the credibility-weighted intraday tilt (BCW) improves the economic lift further by leaning the bid toward the imbalance-price signal. Last is the fee-band curtailment rule at the setpoint stage. It closes the negative-price tail and pays off where such events occur. The day-ahead bid is best left at the production forecast, because no signal was found in this work worth acting on.

Under timing- and liquidity-honest execution, every park keeps the majority of the headline lift and the proposed pipeline still beats the forecast bid on both revenue and tail risk (Section 8.2.3). A producer unwilling to carry much deliberate imbalance need not give up most of the gain. The narrowest intraday band, a clip of a few percentiles around the forecast, already secures most of the revenue lift and essentially the entire tail-risk reduction at no net imbalance cost (Section 8.1.1). Widening the band is therefore a revenue choice and not a safety one.

10.4 Further Work

Several families of questions remain open.

- **Market realism.** The thesis prices the intraday adjustment at the volume-weighted average of executed trades and assumes the producer is a price-taker. A timing- and liquidity-honest check already brackets the cost of these assumptions (Section 8.2.3). The open extension is a full replay of BCW against historical SIDC order books, which would price queue position and true slippage rather than a volume cap. The complementary normative question, raised in Section 9.5, is the horizon at which the price-taker assumption breaks. If enough producers adopt the rules, they shift the aggregate day-ahead position and the spread distribution itself.

- **Portfolio-level bidding.** The thesis evaluates each park independently, which is methodologically clean but differs from the operational setting where a single BRP account covers all three parks. Optimising the intraday tilt at the aggregate portfolio level would net partially uncorrelated production errors across parks in SE1 and SE3. It could achieve comparable ID-spread capture with a smaller deliberate-imbalance footprint (Section 8.1.1) than the per-park sum implies.

- **DA-stage extensions.** The zero-tilt day-ahead result has two natural openings. A DA-stage estimate of the imbalance spread with sign accuracy materially above 50 %, at the ID-stage level say, would in principle change the verdict. The low signal-to-noise ratio at the DA horizon sets the bar for replacing BidF. Separately, the worst-decile residual lives in the underproduction mode where day-ahead exposure modulation provides no measurable relief (Section 8.2.1). Any further reduction has to condition on something the DA stage can observe, for example a production-side risk constraint with a defensible per-producer floor.

- **An uncertainty-quantified ID-stage production forecast.** The ID-Linear centre beat the operator's commercial product but is deliberately simple, and that simplicity has two binding costs (Appendix A.1). The linear conditional mean lags sudden production changes that variance rescaling cannot absorb. The day-ahead anchor carries no within-hour shape, so short-horizon shape errors fall on the AR lags alone. A gradient-boosted regressor per horizon, or NWP-augmented features at the ID gate, could cut how often the centre misses a weather event. Accuracy alone, however, does not reach the residual that dominates the post-pipeline tail (Section 9.4). That residual is a production-forecast one: on the rare collapse days that set the tail, the centre over-predicts. Nothing shrinks the commitment when the centre is wrong the way the credibility weight shrinks the tilt when the spread signal is weak. The structurally relevant extension is therefore an uncertainty-quantified ID forecast whose fan widens on the days the centre is unreliable, paired with a production-side credibility weight that shrinks the commitment when that fan is wide. This mirrors the spread-side weight the pipeline already carries. The centre enters through a single interface, so a richer model is a drop-in replacement and also tightens the residual library every BCW bid is drawn from. By contrast, perfect foresight of the SP-stage imbalance price lifts the mean–CVaR objective by only +4–+8 index pts over the deployed SP rule (3R-on-π_t^{IM} row in Table 7.6), so SP-side forecast sharpening is not a comparably high-leverage extension. The production-side forecast is the high-leverage axis (Section 9.4, BidTrueProd (DA) row in Table 7.2), and this is the highest-leverage single extension to the proposed pipeline.

- **The mFRR Capacity Market.** The thesis sees the mFRR Energy Activation Market only through its imprint on the imbalance price (Section 2.3.1). The same probabilistic-forecasting approach could be applied to the upstream mFRR Capacity Market, where the TSO procures balancing-reserve availability in a day-ahead auction gating at 07:30 CET on $D-1$. The similarity is structural but partial. The offered reserve is again a quantile of the predictive available-production distribution, so the same forecasting machinery applies. The quantile is conservative and low, governed by a capacity-price-to-non-delivery-penalty ratio rather than the symmetric eSett-fee boundary of this thesis.

The obstacle was that no in-house production forecast was available at that early gate, the earliest centre being the operator day-ahead forecast at $D - 1$ 12:00.

- **Methodological refinements.** The residual day-vector bootstrap was selected for transparency and for preserving within-day structure. Generative approaches such as a Gaussian copula or deep generative models might deliver tighter calibration. The relevant test is whether any coverage improvement propagates to the downstream mean–CVaR objective rather than only to the scoring rules. On the SP side, the three-region rule fixes its thresholds at $\pm f^{\text{imb}}$ from theory. Sweeping them at each park would identify where the SP rule's revenue cost ceases to be justified by its $\text{CVaR}_{0.9}$ compression for producers with a different risk profile.
- **Recourse-aware bidding.** The adopted pipeline chains stage-separable rules, and the scenario optimiser, though it co-optimises the bid and the setpoint, chooses both as here-and-now quantities. Neither values the downstream curtailment option when the earlier bid is set (Section 9.4). A genuine multistage stochastic program with nonanticipativity, or a dynamic program over the three gates with the setpoint as recourse, would optimise the earlier bid against the distribution of optimal future curtailment. It would let an intraday under-commitment lean on the insurance the setpoint option provides. The expected gain is bounded, because curtailment fires on a small number of negative-price events, but the limitation is structural and would persist under perfect forecasts.
- **Cross-zone and within-zone replication.** Three SE1/SE3 parks constrain how strongly the per-park heterogeneity findings generalise. Two replication directions are complementary. Cross-zone replication (SE2, SE4, the Norwegian zones, or non-Nordic single-price markets) tests which conclusions are zone-specific versus market-structural. Within-zone, longer-window replication is the more direct test of the headline SP and DA-multiplier results. The dispositive within-SE1 contrast (Blabergsliden vs Blisterliden, Section 8.1) is window-coverage-driven, so a longer Blisterliden history spanning a comparable volatility regime would test whether the difference closes.

Bibliography

Artzner, Philippe et al. (1999). "Coherent Measures of Risk". In: *Mathematical Finance* 9.3, pp. 203–228. DOI: 10.1111/1467-9965.00068.

Bannör, Karl F. and Matthias Scherer (2013). "Capturing Parameter Risk with Convex Risk Measures". In: *European Actuarial Journal* 3.1, pp. 97–132. DOI: 10.1007/s13385-013-0070-z.

Bruninx, Max et al. (2025). *Day-Ahead Bidding Strategies for Wind Farm Operators under a One-Price Balancing Scheme*. Accepted at ACM e-Energy 2025 workshop; arXiv:2505.05153v3, 27 June 2025. arXiv: 2505.05153. URL: https://arxiv.org/abs/2505.05153 (visited on 05/20/2026).

eSett Oy (2025). *Nordic Imbalance Settlement Handbook: Instructions and Rules for Market Participants*. Version 5.2, 29 October 2025. eSett Oy. URL: https://www.esett.com/app/uploads/2025/10/NBS-Handbook-v5.2.pdf (visited on 05/20/2026).

Heiser, Yannick, Farzaneh Pourahmadi, and Jalal Kazempour (2024). *Betting vs. Trading: Learning a Linear Decision Policy for Selling Wind Power and Hydrogen*. Also published in Electric Power Systems Research; arXiv:2412.18479v2, 26 March 2025. arXiv: 2412.18479. URL: https://arxiv.org/abs/2412.18479 (visited on 05/20/2026).

Hersbach, Hans (2000). "Decomposition of the Continuous Ranked Probability Score for Ensemble Prediction Systems". In: *Weather and Forecasting* 15.5, pp. 559–570. DOI: 10.1175/1520-0434(2000)015<0559:DOTCRP>2.0.CO;2.

Klyve, Øyvind Sommer et al. (2023). "Limiting imbalance settlement costs from variable renewable energy sources in the Nordics: Internal balancing vs. balancing market participation". In: *Applied Energy* 350, p. 121696. DOI: 10.1016/j.apenergy.2023.121696.

Kraskov, Alexander, Harald Stögbauer, and Peter Grassberger (2004). "Estimating Mutual Information". In: *Physical Review E* 69.6, p. 066138. DOI: 10.1103/PhysRevE.69.066138.

Muñoz, Miguel Angel, Juan M. Morales, and Salvador Pineda (2020). "Feature-Driven Improvement of Renewable Energy Forecasting and Trading". In: *IEEE Transactions on Power Systems* 35.5, pp. 3494–3504. DOI: 10.1109/TPWRS.2020.2975246. arXiv: 1907.07580.

Muñoz, Miguel Angel, Pierre Pinson, and Jalal Kazempour (2023). "Online Decision Making for Trading Wind Energy". In: *Computational Management Science* 20, p. 33. DOI: 10.1007/s10287-023-00462-2.

Nord Pool AS (2025a). *Bidding areas*. URL: https://www.nordpoolgroup.com/en/the-power-market/Bidding-areas/ (visited on 05/12/2026).

Nord Pool AS (Sept. 2025b). *Intraday Market Regulations.* URL: https://www.nordpoolgroup.com/49efe0/globalassets/download-center/rules-and-regulations/intraday-market-regulations-september-2025.pdf (visited on 05/12/2026).

Nordenergi (June 2025). *Impact of Recent Market Changes on the Nordic Balancing Market.* Tech. rep. Nordenergi. URL: https://www.nordenergi.eu/wp-content/uploads/2025/06/mFRR-EAM-and-consequences-for-the-balancing-market_2025.pdf (visited on 05/20/2026).

Nordic TSOs (2025). *Automated Nordic mFRR Energy Activation Market.* Nordic Balancing Model roadmap page. URL: https://nordicbalancingmodel.net/roadmap-and-projects/automated-nordic-mfrr-energy-activation-market/ (visited on 04/2026).

Pinson, Pierre, Christophe Chevallier, and George N. Kariniotakis (2007). "Trading Wind Generation From Short-Term Probabilistic Forecasts of Wind Power". In: *IEEE Transactions on Power Systems* 22.3, pp. 1148–1156. DOI: 10.1109/TPWRS.2007.901117.

Rockafellar, R. Tyrrell and Stanislav Uryasev (2000). "Optimization of Conditional Value-at-Risk". In: *Journal of Risk* 2.3, pp. 21–41.

Sousa, Vivian and Hugo Algarvio (2025). "Strategic Bidding to Increase the Market Value of Variable Renewable Generators in Electricity Markets". In: *Energies* 18, p. 1586. DOI: 10.3390/en18071586.

Zugno, Marco, Tryggvi Jónsson, and Pierre Pinson (2013). "Trading Wind Energy on the Basis of Probabilistic Forecasts Both of Wind Generation and of Market Quantities". In: *Wind Energy* 16.6, pp. 909–926. DOI: 10.1002/we.1531.

Zugno, Marco, Juan M. Morales, et al. (2013). "Pool Strategy of a Price-Maker Wind Power Producer". In: *IEEE Transactions on Power Systems* 28.3, pp. 3440–3450. DOI: 10.1109/TPWRS.2012.2235078.

Appendix A

In-House Forecast Model Specifications

This appendix documents the feature specifications and training procedures for the two in-house production forecast centres introduced in Section 5.1. The main text summarises forecast accuracy in Tables 5.2 and 5.4; Figure A.1 shows the same accuracy visually as error distributions across stages. These models are not the central contribution of the thesis; they serve as more accurate forecast centres for the error libraries used in scenario generation.

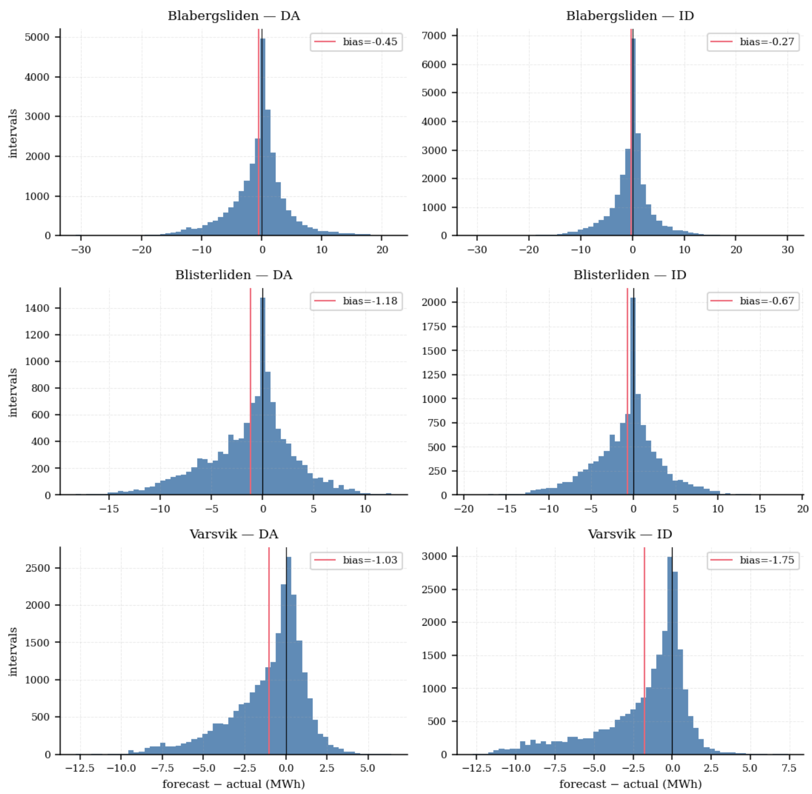

Figure A.1: Production forecast error distributions (forecast minus available production) by stage and park; the red line marks the bias.

A.1 ID-Linear

ID-Linear is the in-house ID-stage production centre. The ID stage is represented as an hour-block decision: at a common origin 90 minutes before the delivery hour, the model predicts the four 15-minute target slots using horizons 6, 7, 8, and 9 (counted in 15-minute MTUs from the origin). For example, for delivery hour 14:00 the origin is 12:30, and the four target slots 14:00, 14:15, 14:30, 14:45 take horizons $h = 6, 7, 8, 9$ respectively. The cached series `a_ID_fc` is built by selecting, for each target slot, the horizon prediction that matches the slot's position within its delivery hour.

Feature Specification

One linear model is fitted per horizon (four models in total, one each for $h = 6, 7, 8, 9$). Features are computed using only information available at the origin time $t - h$, with one exception: the DA production forecast, issued the day before delivery, is known for the target interval at origin time and is therefore entered at target time. The features are:

- lagged available production a_{t-h-k} for $k \in \{1, \ldots, 6\}$ (six features);
- rolling origin-time means of available production over the most recent 4 and 8 intervals, plus a 4-step ramp term $a_{t-h} - a_{t-h-4}$ (three features);
- cyclical hour-of-day and day-of-week encodings $(\sin, \cos)$ for the target interval (four features);
- the DA production forecast $\hat{a}_t^{\mathrm{DA}}$ as an exogenous anchor, entered at target time with no shift (one feature).

The feature vector has dimension 14 per horizon. Each horizon model is a `StandardScaler` followed by an unregularised `LinearRegression`. The fitted prediction for target t at horizon h is

$$\hat{a}_t^{(h)} = \beta_0^{(h)} + \sum_{k=1}^{6} \beta_k^{(h)}\, a_{t-h-k} + \beta_{\bar{4}}^{(h)}\, \tfrac{1}{4}\sum_{k=0}^{3} a_{t-h-k} + \beta_{\bar{8}}^{(h)}\, \tfrac{1}{8}\sum_{k=0}^{7} a_{t-h-k} + \beta_{\mathrm{r}}^{(h)}\, \big(a_{t-h} - a_{t-h-4}\big) + \beta_{\mathrm{cal}}^{(h)\top} \mathbf{c}_t + \beta_{\mathrm{DA}}^{(h)}\, \hat{a}_t^{\mathrm{DA}}, \quad \text{(A.1)}$$

where $\mathbf{c}_t = (\sin \frac{2\pi h_t}{24}, \cos \frac{2\pi h_t}{24}, \sin \frac{2\pi w_t}{7}, \cos \frac{2\pi w_t}{7})$ encodes the target slot's hour-of-day h_t and day-of-week w_t. All predictors are standardised to zero mean and unit variance on the training window before the regression is fit. The standardisation parameters are retrained alongside the coefficients at each walk-forward refit.

Why both persistence and a DA anchor. The two feature families capture complementary information. The AR lags, rolling means, and ramp term encode the recent production trajectory at origin time, short-term wind variability the day-old DA forecast cannot see. The DA forecast $\hat{a}_t^{\mathrm{DA}}$ carries the day-scale meteorological regime that the 90-minute history alone cannot reconstruct. The linear regression learns the relative weight: where the DA forecast is locally biased, the lag features dampen it; where production is on a persistent trajectory the DA forecast missed, the lag features pull the prediction back toward the realised path. This fusion is the main reason ID-Linear outperforms the operator's commercial ID-stage forecast on the evaluated parks (Table 5.2).

From four horizon outputs to one per-slot forecast. Each target slot reads the prediction from its matching horizon model: a target at minute 0 of the delivery hour takes the $h = 6$ prediction, minute 15 takes $h = 7$, minute 30 takes $h = 8$, and minute 45 takes $h = 9$. The resulting per-slot series, named `a_ID_fc` and indexed by target time at 15-min resolution, feeds the downstream scenario fan (Chapter 6) and bidding rules (Chapter 7).

Accuracy

Table 5.2 (in Section 5.2.1) reports accuracy on the evaluation window alongside the operator's ID forecast. The ID-Linear centre achieves Pearson correlation $r = 0.937$, 0.936, and 0.900 at Blabergsliden, Blisterliden, and Varsvik respectively, against $r = 0.897$, 0.837, and 0.706 for the operator forecast. The largest gain is at Varsvik, where the operator's ID update is less accurate than its own DA forecast at the 90-minute horizon. The ID-Linear centre recovers meaningful predictive skill there by combining production persistence with the DA forecast anchor.

Design Choices and Limitations

Three non-obvious design choices are worth flagging, because they bound what ID-Linear can deliver and motivate the future-work direction in Section 10.4.

1. **Linear OLS on the centre; non-parametric bootstrap on the spread.** Each horizon is a plain `StandardScaler` → `LinearRegression` without regularisation, interactions, or nonlinear bases, fit by ordinary least squares. The predictive distribution of a_t is supplied separately, by the residual day-vector bootstrap of Chapter 6: ID-Linear produces the centre, and the bootstrap produces the spread around it. The decoupling is deliberate and lets each component be audited independently. The transparency cost of OLS is paid in two directions: every coefficient in (A.1) is interpretable, and the unregularised mean estimator carries no shrinkage bias the downstream fan would need to correct. The structural cost is that the bootstrap assumes residuals are mean-zero conditional on the kernel features (Table 6.3); where ID-Linear's linearity assumption is violated in a regime outside the kernel's feature support, the residuals carry a conditional bias the bootstrap inherits silently. Variance rescaling can correct the fan's width but not its position (Section 6.5).

2. **Same feature set across horizons.** Under the hour-block convention all four horizons share the same origin ($t_0 = H - 90$ min for delivery hour H), and the lag-ℓ feature at target t is read from $t - h - \ell$. With $h = 6, 7, 8, 9$ for the four target slots within the hour, the offsets cancel and all four horizon models receive the same lag, rolling, and ramp features — literally the same physical timestamps, not just the same recipe.

The DA anchor is also the same across slots within the hour (the DA price forecast is at hourly resolution, repeated across the four MTUs). Only the cyclical hour-of-day encoding differs, and only by the 45-minute spread between the four target slots. The four horizon models therefore differ in the target they regress against, not in the information they receive. Any accuracy gap between $h = 9$ and $h = 6$ at the same delivery hour is bounded by the inherent difficulty of predicting further into the hour from a fixed information set.

3. **DA anchor at target time, no within-hour shape signal.** $\hat{a}_t^{\mathrm{DA}}$ is the only non-AR exogenous feature, entered at target time. It anchors the daily *level* of the forecast but provides no information about *within-hour shape*: ramps, gusts, and minute-scale variability are entirely covered by the AR lags, which see production only up to the 90-minute-old origin slot and carry no information about what happens between origin and target. Improving the centre on shape-driven errors therefore requires a different exogenous input (e.g. NWP features), not a richer functional form on the existing one.

A more capable ID centre — gradient-boosted regression per horizon, or NWP-augmented features — would close the bias gap that the linearity assumption (item 1) and the level-only DA anchor (item 3) leave open, and is the highest-leverage extension to the pipeline (Section 10.4).

A.2 SP-Persistence

SP-Persistence is the SP-stage production centre, providing the forecast anchor for the SP-stage scenario fan used in the independence test of appendix C.9. The SP gate sits at approximately $H - 20$ minutes for delivery hour H. Feature construction mirrors ID-Linear (appendix A.1) but uses only lags $k \in \{1, 2\}$ of a_{t-h-k} and omits the DA forecast anchor: at this horizon, lag-1 of available production is so informative that the day-old forecast adds negligible incremental signal. The rolling-mean, ramp, and cyclical calendar features remain, giving a feature vector of dimension 9. Training follows the same 60-day warm-up and weekly expanding-window refit. The resulting cached series `a_SP_fc` is near-exact at this horizon; Table 5.4 reports the SP-stage MAE across parks.

Appendix B

Scenario Construction and Market-Regime Diagnostics

This appendix provides supplementary material for Chapter 6: the per-feature mutual-information scores omitted from the main chapter, calibration diagnostics for the production bootstrap fan (PIT–KS distance and 80 % coverage), and a DA-stage market-regime finding that motivates the scenario kernel design, the nonlinear relationship between wind level and regulation state. The regime finding does not constitute an interval-level DA directional signal; it enters the scenario kernel as a conditioning variable shaping the production fan.

B.1 Per-feature mutual information scores

Figure B.1 shows the full per-feature MI breakdown, supplementing the group-level summary in the left panel of Figure 6.2.

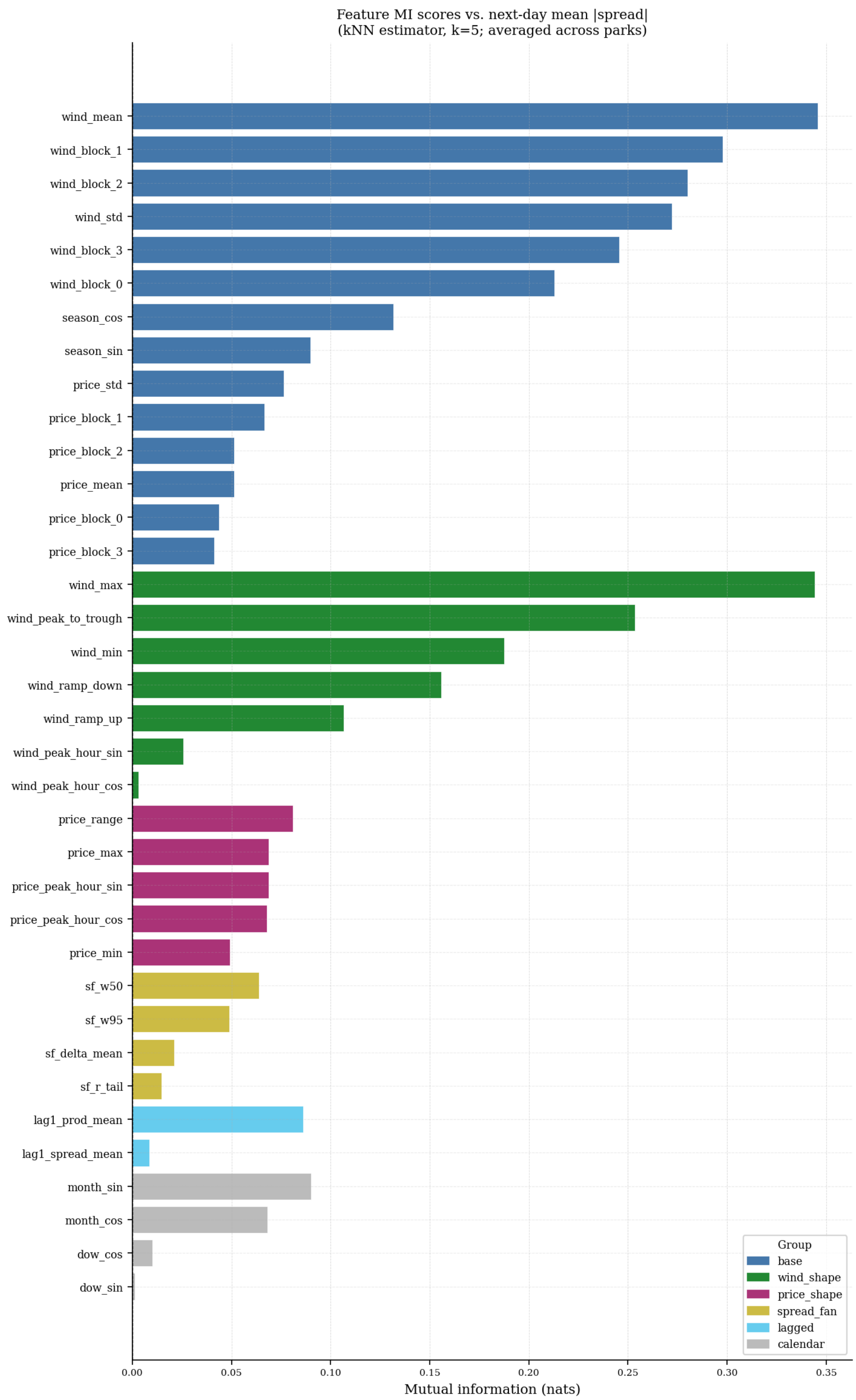

Figure B.1: Per-feature kNN MI scores against next-day mean absolute production error $\bar{e}_d = \overline{|a_t - \hat{a}_t^{\mathrm{DA}}|}$, averaged across the three parks. Features coloured by group.

B.2 Production fan reliability and shape fidelity

Table B.1 reports the PIT–KS statistic and empirical 80 % coverage of the adopted bootstrap fan per park and stage, complementing the PIT histogram of Figure 6.4.

Table B.1: PIT–KS distance and empirical 80 % coverage of the production bootstrap fan, by park and stage.

Park	Stage	PIT-KS ↓	80% cov.	gap
Blabergssliden	DA	0.036	76.1%	-3.9 pp
Blabergssliden	ID	0.079	76.6%	-3.4 pp
Blisterliden	DA	0.055	78.3%	-1.7 pp
Blisterliden	ID	0.091	69.5%	-10.5 pp
Varsvik	DA	0.026	81.1%	+1.1 pp
Varsvik	ID	0.045	83.7%	+3.7 pp

B.3 Regulation state by wind quintile

Figure B.2 partitions the capacity-normalised DA wind forecast into five quintiles and shows the resulting regulation-state frequencies. In SE1, down-regulation is most pronounced at high wind (Q5), consistent with excess supply pushing the system long. In SE3 there is no such gradient — up-regulation is actually *highest* at Q5, the opposite of the physical intuition — suggesting that SE3 regulation is driven by broader market factors beyond park-level wind.

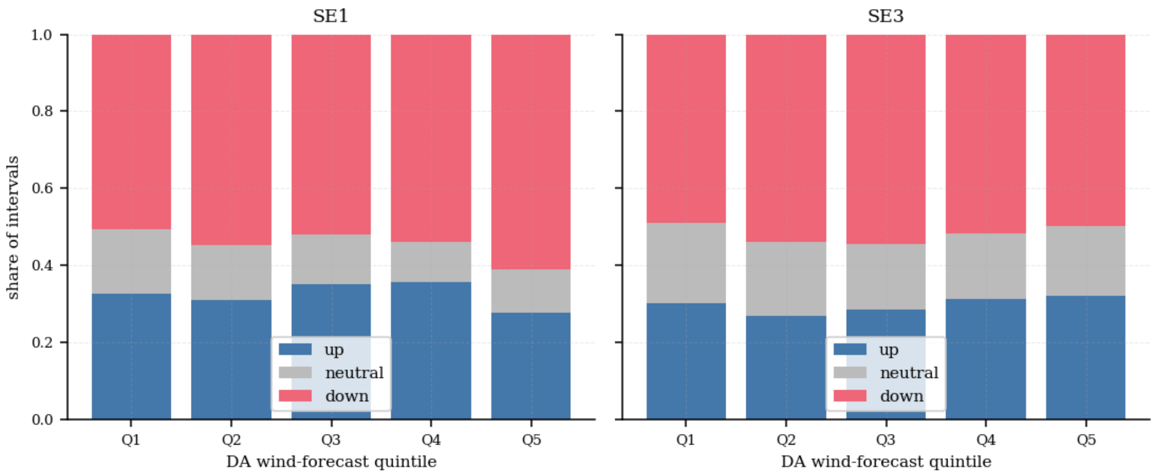

Figure B.2: Regulation-state frequency by DA wind forecast quintile for SE1 and SE3. Down-regulation ($\delta_t^{\mathrm{DA}} > 0$, red), up-regulation ($\delta_t^{\mathrm{DA}} < 0$, blue), neutral (grey).

This contrast between SE1 and SE3 motivates including DA price-block features alongside wind features in the scenario kernel (the *base* feature group of Table 6.3): for SE3, the park-level wind forecast alone is insufficient to characterise the regulation regime, so the kernel needs price-level signals to identify analogous historical days.

Appendix C

Bidding-Method Diagnostics

This appendix collects diagnostic evidence supporting the bidding-method comparisons introduced in Chapter 7 and the empirical results reported in Chapter 8. The material covers the DA-stage rule specification, the day-ahead bias-correction decomposition, the walk-forward stability and grid adequacy of the proposed BCW rule's calibrated parameters, paired significance tests for the ID-stage lift, the scenario LP's bang-bang behaviour and per-slot signal quality, the ID-stage production–spread tail coupling, and an empirical test of the SP conditional-independence assumption.

C.1 DA-stage rule specification

The only DA-stage directional candidate evaluated in the bake-off is RollingMean-DA. It takes the last 180 daily-mean realised spreads observable at bid time (the eSett publication delay of 2–13 days, Section 2.3, cuts the window at $d-2$), winsorises them at ± 3 standard deviations to get $\bar{\delta}_W^{\mathrm{DA}}$ and $\hat{\sigma}_\delta^{\mathrm{DA}}$, and applies

$$\tau = \mathrm{clip}\left(\tfrac{1}{2} + \frac{\bar{\delta}_W^{\mathrm{DA}}}{2f^{\mathrm{imb}}}\,\gamma,\ \tau_{\mathrm{lo}},\ \tau_{\mathrm{hi}}\right), \qquad \gamma = \min\left(1,\ \frac{|\bar{\delta}_W^{\mathrm{DA}}|}{3\,\hat{\sigma}_\delta^{\mathrm{DA}}}\right), \tag{C.1}$$

as a single per-day quantile level on the production fan. The 180-day history is available from the first evaluation day at every park because zone-level price data extend back to January 2022.

C.2 SP-stage SoftAnchored: form and calibration

The non-adopted SoftAnchored candidate at the setpoint stage (Section 7.4) replaces the hard switch of (7.5) with a calibrated tanh ramp of the fee-normalised forecast,

$$s_t = b_t + \max(\alpha_t, 0)\,(P^{\max} - b_t) + \min(\alpha_t, 0)\,b_t, \qquad \alpha_t = \tanh\left(k\,\frac{\hat{\pi}_t^{\mathrm{IM,SP}}}{f^{\mathrm{imb}}}\right), \tag{C.2}$$

with $k \geq 0$ walk-forward calibrated on the mean–CVaR objective. SoftAnchored has *no* interior hold band — it switches around $\hat{\pi}_t = 0$, not at $\pm f^{\text{imb}}$ — so even as $k \to \infty$ it collapses to CurtailToZero, *not* to (7.5). This is why the candidate loses everywhere in Table 7.6: the ramp pays the fee over the interior band where ThreeRegion holds fee-free.

C.3 Day-ahead bias-correction diagnostic

FanMedianDA and FanMeanDA differ from BidF only in the bid centre: they read the fan's median and mean per slot instead of the operator point forecast. Since the fan corrects the operator's low bias (Section 6.2), the comparison isolates the value of that re-centring — to the conditional median, or the fuller conditional mean — from any spread input (Table C.1, post-warm-up window).

Re-centring does lift mean revenue at nearly every park: both variants recover part of the operator's systematic underprediction. On the mean–CVaR objective, however, no lift survives uncertainty quantification: every 95 % paired bootstrap CI spans zero (Table C.1), confirming the cross-park inconsistency already visible in the point estimates.

The median and mean re-centrings differ in aggressiveness. The fan median is a conservative correction — it closes the operator's bias without overshooting — whereas the fan mean shifts the bid further toward the distributional centre, creating more deliberate long imbalance. The $\text{CVaR}_{0.9}$ column tracks the cost directly: at Varsvik, the most right-skewed park, FanMeanDA widens $\text{CVaR}_{0.9}$ from -1.5 to $+6.7$ index points while FanMedianDA leaves it almost unchanged (-0.2). The more aggressive the re-centre, the worse the tail, and neither escapes the statistical noise at this stage. The operator forecast, despite its low bias, is therefore not dominated at the DA stage.

Table C.1: DA bias-correction diagnostic: BidF, FanMedianDA, and FanMeanDA. Mean daily revenue, $\text{CVaR}_{0.9}$, and mean–CVaR objective ($\lambda = 1$) in index units (BidF mean revenue = 100 per park). CI: 95 % paired case-bootstrap interval on objective lift vs BidF ($B = 1000$); all CIs span zero.

Park	DA bid centre	$\bar{R}$	$\text{CVaR}_{0.9}$	Obj	95 % CI vs BidF
Blabergsliden	BidF† (operator)	100.0	58.5	41.5	—
	FanMedianDA (fan median)	99.9	55.1	44.9	[-0.4, +8.7]
	FanMeanDA (fan mean)	101.2	55.2	46.0	[-1.1, +11.7]
Blisterliden	BidF† (operator)	100.0	0.0	100.0	—
	FanMedianDA (fan median)	101.9	2.9	99.0	[-3.8, +2.6]
	FanMeanDA (fan mean)	103.6	5.2	98.4	[-7.5, +4.0]
Varsvik	BidF† (operator)	100.0	-1.5	101.5	—
	FanMedianDA (fan median)	100.9	-0.2	101.1	[-5.5, +4.0]
	FanMeanDA (fan mean)	100.5	6.7	93.8	[-18.9, +0.6]

C.4 BCW parameter sensitivity: grid adequacy and walk-forward

The proposed BCW rule re-selects $(\hat{w}_d, \hat{z}_d)$ every delivery day on the joint 8×8 grid $\mathcal{W} \times \mathcal{Z}$ via the calibrator of (7.3). This section reports the per-day trajectory stability and whether the grid is wide enough along z (the w axis is covered by Section 8.1.1).

Walk-forward stability. Figure C.1 shows the per-day $(\hat{w}_d, \hat{z}_d)$ trajectories. A stable mode (the desirable pattern) indicates the calibrator has settled rather than chasing recent noise inside the warm-up grid; a drifting trace would warn that the 30-day warm-up is too short to identify (w, z) on the BCW objective surface. The annotated mode and its frequency in each panel summarise this stability at a glance.

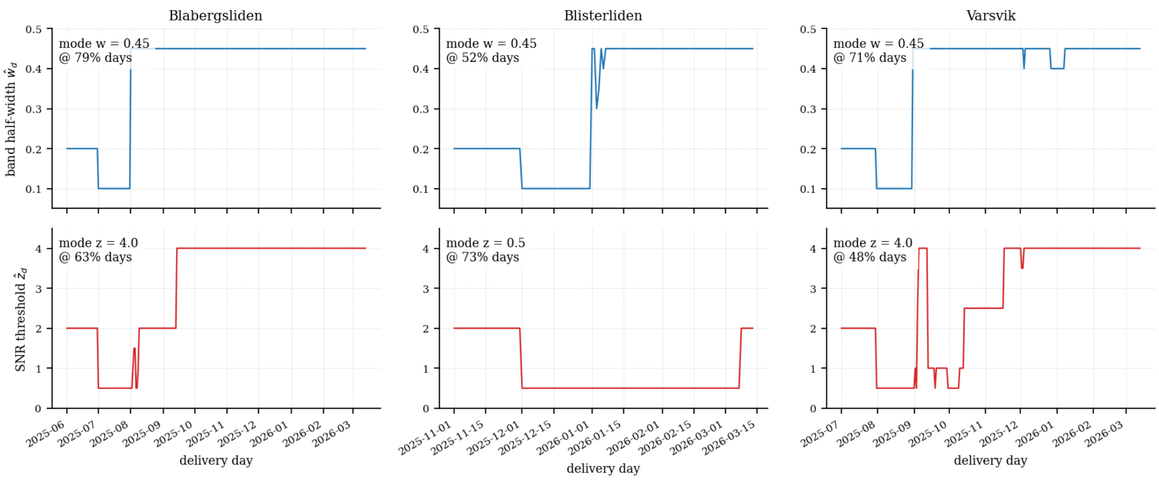

Figure C.1: Walk-forward calibration of BCW: per-day band half-width $\hat{w}_d$ (top) and SNR threshold $\hat{z}_d$ (bottom), per park. Annotated boxes show the post-warm-up mode and its frequency; dashed line marks the first calibrated day.

Grid adequacy along z. To test whether the production grid's $z \in [0.5, 4]$ is wide enough, hold w at its calibrated mode ($w = 0.45$, the band $[Q_{0.05}, Q_{0.95}]$) and sweep the SNR threshold z far past both grid edges, from $z = 0.1$ (near the $\gamma_t \equiv 1$ BFS limit, no shrinkage) to $z = 24$ (shrinkage so aggressive the tilt is all but suppressed and the rule collapses toward BIDF). Table C.2 reports mean daily revenue, $\text{CVaR}_{0.9}$, and the mean–CVaR objective on the post-warm-up availability window (the same Panel.A basis as Table 7.3), per park.

Table C.2: BCW objective sensitivity to the SNR threshold z at $w = 0.45$, post-warm-up window. $\bar{R}$, $\text{CVaR}_{0.9}$, and mean–CVaR objective ($\lambda = 1$) in index units (baseline = 100 per park). Rows marked † are the grid edges $\mathcal{Z} = \{0.5, \ldots, 4\}$.

	Blabergsliden			Blisterliden			Varsvik		
z	$\bar{R}$	CVaR	Obj	$\bar{R}$	CVaR	Obj	$\bar{R}$	CVaR	Obj
0.10	113.4	50.6	62.9	109.2	−3.5	112.7	122.0	−13.7	135.7
0.50†	113.3	50.6	62.8	109.1	−3.1	112.3	121.7	−13.7	135.4
1.00	113.3	50.2	63.0	109.1	−3.1	112.1	121.7	−13.6	135.3
2.00	113.1	50.0	63.1	108.9	−3.0	111.9	121.6	−13.8	135.4
4.00†	112.8	49.7	63.1	108.7	−3.0	111.7	121.4	−14.0	135.3
8.00	111.8	49.6	62.3	108.2	−3.1	111.3	120.7	−13.9	134.6
16.00	110.5	49.1	61.4	107.6	−3.8	111.4	119.6	−14.0	133.6
24.00	109.3	49.1	60.1	107.1	−4.0	111.1	118.8	−13.9	132.7

The picture is the same at all three parks. Inside the production grid the objective is essentially flat: it spans under 0.7 index points at every park over the full $[0.5, 4]$, a fraction of the DA-stage noise the bake-off itself exhibits. The argmax is at the upper grid edge for Blabergsliden ($z = 4$) and at the low-z BFS corner for Blisterliden and Varsvik, but the gain from pushing z below the lower grid edge is small (+0.4 index points at Blisterliden, +0.3 at Varsvik, +0.1 at Blabergsliden) and *monotone* toward the $\gamma_t \equiv 1$ limit: the surface is a flat plateau that slopes gently into the BFS corner, not a basin with a missed optimum outside the grid. Above the upper edge the objective declines monotonically as over-shrinkage suppresses the tilt and the rule degenerates toward BIDF (at $z = 24$, −3.0, −0.6, and −2.6 index points relative to $z = 4$); there is no hidden peak. The calibrated parameters therefore sit at grid edges because the objective is flat there, not because the optimum lies beyond the search: the headline ID results are grid-independent.

C.5 Sensitivity of the BCW conclusion to the risk-aversion weight

The mean–CVaR objective (3.5) is reported at $\lambda = 1$ throughout the thesis: the risk penalty is then in the same units as mean revenue, giving a direct euro-for-euro tradeoff. That choice does, however, enter the BCW pipeline twice over: the per-day calibrator (7.3) reads λ when selecting $(\hat{w}_d, \hat{z}_d)$, and the headline objective uses it again to score the realised revenue series. This appendix checks whether the ID-stage BCW-vs-BIDF conclusion is load-bearing on $\lambda = 1$ by re-running the same pipeline at $\lambda \in \{0, 0.5, 1, 2\}$, per park. The baseline is λ-independent and run once.

Table C.3 reports the results. The mean-revenue lift is positive at every park for every λ, including $\lambda = 0$ (no risk term), so BCW beats the baseline on pure mean revenue before the risk penalty enters. BCW also reduces $\text{CVaR}_{0.9}$ at every park and every λ: the gain is not a mean-vs-tail trade-off but a strict dominance on both axes. The objective lift is therefore positive at every cell and grows monotonically in λ, as the heavier risk weight rewards the existing tail advantage more strongly. The calibrator shifts $\hat{z}$ exactly as theory predicts: at $\lambda = 0$ it settles at the low end (minimal SNR shrinkage, maximise mean) at every park; at

$\lambda \geq 0.5$ the two tail-prone parks jump to the high end (maximal shrinkage), while Blisterliden, whose baseline $\mathrm{CVaR}_{0.9}$ is already small, never sees enough tail pressure to invoke shrinkage. The band half-width sits at the grid edge $\hat{w} = 0.45$ throughout, consistent with the grid-adequacy analysis of appendix C.4. The headline conclusion is therefore not sensitive to the specific choice $\lambda = 1$.

Table C.3: ID-stage BCW-vs-BIDF lift across $\lambda \in \{0, 0.5, 1, 2\}$, per park, post-warm-up window. $\Delta\bar{R}$: mean daily revenue lift. $\mathrm{CVaR}^{\mathrm{base}} - \mathrm{CVaR}^{\mathrm{BCW}}$: tail reduction (positive = BCW tighter). Obj lift: $\Delta\bar{R} - \lambda\,\Delta\mathrm{CVaR}$. All in index units (BIDF mean revenue = 100 per park). $\hat{w}, \hat{z}$: post-warm-up modes of the per-day calibrated band half-width and SNR threshold on the grids of appendix C.4.

Park	λ	$\Delta\bar{R}$	$\mathrm{CVaR}_{0.9}^{\mathrm{base}} - \mathrm{CVaR}_{0.9}^{\mathrm{BCW}}$	Obj lift	$\hat{w}$	$\hat{z}$
Blabergsliden	0.0	+11.15	+7.18	+11.15	0.45	0.5
	0.5	+10.76	+7.40	+14.46	0.45	4.0
	1.0	+10.78	+7.57	+18.36	0.45	4.0
	2.0	+10.20	+7.36	+24.92	0.45	4.0
Blisterliden	0.0	+7.67	+1.87	+7.67	0.45	0.5
	0.5	+7.57	+1.87	+8.50	0.45	0.5
	1.0	+7.40	+1.87	+9.27	0.45	0.5
	2.0	+7.29	+1.87	+11.03	0.45	0.5
Varsvik	0.0	+16.03	+8.52	+16.03	0.45	0.5
	0.5	+15.85	+8.33	+20.02	0.45	4.0
	1.0	+15.58	+8.33	+23.92	0.45	4.0
	2.0	+13.24	+8.26	+29.76	0.45	4.0

C.6 Significance tests for the ID-stage lift

The ID-stage lifts of Table 7.3 might reflect favourable market conditions rather than a genuine signal. Each is tested with the paired case-bootstrap interval on mean revenue lift at $\lambda = 0$ (Section 4.6.4); no additional methodology is introduced here. Table C.4 reports two paired comparisons: BCW vs ID-Linear isolates the value of price conditioning given that forecast quality is already optimised (the more demanding test); BCW vs BIDF measures the total ID-stage gain.

Table C.4: ID-stage lift significance at $\lambda = 0$ (paired case-bootstrap). $\overline{\Delta R}$: mean per-day revenue difference; 95 % CI: paired case-bootstrap interval ($B = 1000$). n varies by park due to differences in the post-warm-up window over which all required price and stage forecasts are simultaneously available.

Park	Comparison	n	$\overline{\Delta R}$ (index)	95% CI (index)
Blabergsliden	BCW-ID vs BIDFORECAST	255	+11.5	[+8.1, +14.8]
Blabergsliden	BCW-ID vs ID-LINEAR CORRECTION	255	+10.6	[+7.9, +12.9]
Blisterliden	BCW-ID vs BIDFORECAST	103	+7.9	[+4.4, +11.7]
Blisterliden	BCW-ID vs ID-LINEAR CORRECTION	103	+5.4	[+3.3, +7.5]
Varsvik	BCW-ID vs BIDFORECAST	226	+15.6	[+12.0, +19.3]
Varsvik	BCW-ID vs ID-LINEAR CORRECTION	226	+12.9	[+10.2, +15.6]

C.7 LP solver diagnostics: bang-bang behaviour and per-slot signal quality

Joint scenario apparatus. The SAA policy consumes joint (a, π^{IM}) day vectors drawn from the same residual day-vector bootstrap of Section 6.2. Price residuals are stored alongside the production residuals on the same day vector: at the DA stage both $\epsilon_t^{\pi^{\mathrm{DA}},d}$ and $\epsilon_t^{\pi^{\mathrm{IM}},d}$ are stored, while at the ID and SP stages the DA price is already observed and only $\epsilon_t^{\pi^{\mathrm{IM}},d}$ is stored. The joint scenario is

$$\pi_t^{\mathrm{IM},(s)} \;=\; \hat{\pi}_t^{\mathrm{IM}} \;+\; \epsilon_t^{\pi^{\mathrm{IM}},\,d_s}, \tag{C.3}$$

paired with the production scenario of Equation (6.2) from the same sampled day d_s (and analogously for π_t^{DA} at the DA stage). Because both residuals travel on the same day vector, the empirical production–price correlation is preserved by construction.

Let $\delta_t^{\mathrm{ID},(s)}$ denote the realised ID-stage spread in scenario s at slot t, and write the per-slot scenario mean as

$$\bar{\delta}_t^{\mathrm{ID}} \;=\; \frac{1}{S}\sum_{s=1}^{S} \delta_t^{\mathrm{ID},(s)}. \tag{C.4}$$

The LP's first-order condition uses $\bar{\delta}_t^{\mathrm{ID}}$ in place of the unobservable conditional expectation $\mathbb{E}[\delta_t^{\mathrm{ID}} \mid \mathcal{F}^{\mathrm{ID}}]$.

The main-body argument of Section 8.1 combines two theoretical claims. First, the scenario LP is structurally prone to bang-bang: risk-neutrally, because the fee is small, its interior first-order condition pushes the position to a band edge on almost every slot. Risk aversion tempers this unevenly — it survives at the day-ahead stage and weakens at intraday (Table C.5) — but at either stage the LP commits on the scenario mean as though it were the truth. Second, the per-slot scenario mean $\bar{\delta}_t^{\mathrm{ID}}$ that drives the LP's directional decision, a sample average over bootstrap-resampled historical days, carries only weak directional information about the realised δ_t^{ID} on the target day, so the LP's 96 independent slot-level bets are largely uninformed. This section provides empirical support for both claims on the post-warm-up evaluation window. Coverage tracks ID-Linear ID-forecast cache availability, since each replay requires the ID-Linear forecast, all realised prices, and a solvable LP simultaneously.

The LP clamps the position to the per-slot production-quantile band $[Q_{0.5-w}, Q_{0.5+w}]$ at the modal calibrated half-width $w = 0.45$ (the band $[Q_{0.05}, Q_{0.95}]$, essentially the full production fan), the value walk-forward calibration selects on most days at all three parks; a slot counts as bang-bang when the position sits within 2% of $P_{\max}$ of a band edge. We report both the risk-neutral weight $\lambda = 0$ and the deployed mean–CVaR weight $\lambda = 1$, at the day-ahead and intraday stages.

C.7.1 Empirical On-Off Frequencies

Table C.5 reports, per park and stage, the fraction of 15-minute slots on which the LP's total position $q_t = b_t^{\mathrm{DA}} + \Delta b_t^{\mathrm{ID}}$ sits within 2 % of a band edge, at the risk-neutral weight $\lambda = 0$ and the deployed mean–CVaR weight $\lambda = 1$.

Table C.5: Bang-bang behaviour of the scenario LP, by stage and risk-aversion weight, on the post-warm-up evaluation window. For each park and stage, "bang-bang" is the fraction of 15-minute slots on which the LP places its position within 2% of $P_{\max}$ of a band edge; the columns give this fraction at the risk-neutral weight $\lambda = 0$ and the deployed mean–CVaR weight $\lambda = 1$, with the interior fraction at $\lambda = 1$. The position is clamped to the production-quantile band $[Q_{0.05}, Q_{0.95}]$ at the modal calibrated half-width $w = 0.45$; the risk-neutral fractions are insensitive to w.

		Bang-bang [%]		Interior [%]
Park	Stage	$\lambda = 0$	$\lambda = 1$	$\lambda = 1$
Blabergsliden	DA	94.1	94.6	5.4
	ID	97.8	61.2	38.8
Blisterliden	DA	97.7	93.5	6.5
	ID	98.4	78.8	21.2
Varsvik	DA	96.9	94.7	5.3
	ID	98.6	49.4	50.6

Risk-neutrally the LP is bang-bang at both stages and every park (94–99 %): with the fee small, sitting interior is almost never optimal. The deployed mean–CVaR weight changes this only at the intraday stage. The day-ahead LP stays bang-bang (93–95 %) — with no directional signal to weigh against the revenue tail, it still commits fully — while the intraday LP is pulled off the edge, sitting interior on 21 to 51 % of slots (most at Varsvik). Those intraday interior moves are not a response to spread-estimate uncertainty: the LP still bids off the per-slot scenario mean regardless of how reliable that estimate is, which is why BCW, whose shrinkage tracks the per-interval signal-to-noise ratio, beats it on the objective (Table 7.3).

C.7.2 Per-slot scenario mean $\bar{\delta}^{\mathrm{ID}}$ vs realised δ^{ID}

The LP converts the scenario set into a bid via the per-slot scenario mean $\bar{\delta}_t^{\mathrm{ID}}$ defined in (C.4). Figure C.2 plots this quantity against the realised δ_t^{ID} on the same slot, pooled across all evaluation days for each park. The Pearson correlations and per-slot sign-accuracy statistics printed in each panel title quantify the directional information content of the scenario-mean signal at the slot level.

This is a signal-content diagnostic, not a significance test of a revenue lift in the sense of Section 4.6.4: it asks whether the per-slot signal carries any directional information, not whether a revenue difference is non-zero, and the two should not be conflated. The realised Pearson correlations and sign-accuracies sit well above the permutation null: each panel title in Figure C.2 reports the 95th-percentile $|r|$ and sign accuracy from 1,000 within-park shuffles of δ_t^{ID}, with two-sided permutation $p \approx 0$ at all three parks. The per-slot signal therefore carries genuine directional information. But the small magnitude of r (≤ 0.23) and the sub-60 % sign accuracy mean the signal is far too weak to drive aggressive directional bets at the slot level.

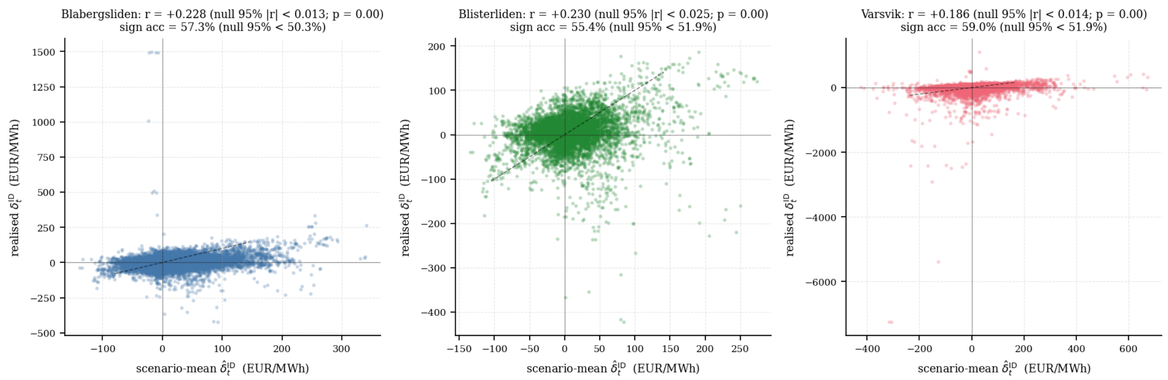

Figure C.2: Per-slot scenario mean $\bar{\delta}_t^{\mathrm{ID}}$ versus the realised δ_t^{ID}, pooled across all evaluation days per park. Pearson correlation r and slot-level sign accuracy are printed in each panel title.

C.8 ID-stage production–spread tail coupling

This is a scope diagnostic on the ID mean–CVaR, not an independence defence: the DA/ID rule is joint-free at the risk-neutral optimum by construction (Section 7.3). The rule is *evaluated* on mean–CVaR, however, and the $\mathrm{CVaR}_{0.9}$ term is a functional of the joint law of (a, δ) that the marginal-quantile form does not model. A near-zero pooled cross-correlation (Figure 5.1, $|\mathrm{CCF}| \leq 0.05$) does not settle this, because sign-asymmetric coupling cancels in any linear or rank correlation.

The decision-relevant cut is therefore signed. With a_t and the signed ID spread $\delta_t^{\mathrm{ID}} = \pi_t^{\mathrm{ID}} - \pi_t^{\mathrm{IM}}$ over the post-warm-up window, Table C.6 reports the decile co-exceedance probability $\Pr(\delta^{\mathrm{ID}}$ top decile $\mid a_t$ bottom decile$)$ against the unconditional reference 0.10, with a 95 % day-block bootstrap interval and a copula-free permutation null. Production-collapse intervals carry the δ^{ID}-top tail at 1.2 to 1.6 times the unconditional rate by point estimate, though the effect is statistically distinguishable from chance only at Blabergsliden — the Blisterliden and Varsvik intervals include 0.10 — while the mirror corner (production high, δ^{ID} low) sits at or below chance. This sign asymmetry is why the pooled CCF reads ≈ 0 on a non-symmetric conditional law.

The corner identified here is the intraday-rich one (π_t^{ID} high relative to π_t^{IM}), distinct from the underproduce-into-π_t^{IM}-spike loss mode of Section 9.4 (π_t^{IM} high $\Rightarrow \delta^{\mathrm{ID}}$ in its *bottom* decile).

Internal validity is unaffected: the paired-bootstrap lift is computed on realised revenues, so any joint coupling sits inside both revenue vectors and biases neither the point estimate nor the no-trade-off conclusion. External validity is bounded by the modest asymmetry and wide day-block intervals. The marginal-quantile rule is sound and the residual joint tail is a named, quantified caveat.

Table C.6: ID-stage signed production–spread decile co-exceedance, post-warm-up window. $\Pr(\delta^{\text{ID}} \uparrow \mid a \downarrow)$ is $\Pr(\delta^{\text{ID}}$ in its top decile $\mid a$ in its bottom decile), unconditional reference 0.10; "×ref" is the lift over 0.10; "perm" is the mean of a 1000-replicate copula-free day-block permutation null; the 95 % CI is the day-block bootstrap interval. $\Pr(\delta^{\text{ID}} \downarrow \mid a \uparrow)$ is the mirror corner. Signed throughout; scope diagnostic on the ID mean–CVaR.

		$\Pr(\delta^{\text{ID}} \uparrow \mid a \downarrow)$			null	$\Pr(\delta^{\text{ID}} \downarrow \mid a \uparrow)$	
Park	n	$\hat{p}$	95% CI	×ref	perm	$\hat{p}$	95% CI
Blabergsliden	23654	0.141	[0.109, 0.180]	1.41	0.101	0.040	[0.020, 0.070]
Blisterliden	9666	0.162	[0.095, 0.219]	1.62	0.102	0.055	[0.011, 0.107]
Varsvik	21005	0.116	[0.081, 0.154]	1.16	0.099	0.075	[0.052, 0.105]

C.9 Empirical test of the SP conditional-independence assumption

Proposition 3.3 derives the three-region setpoint rule under the conditional-independence hypothesis $\pi_t^{\text{IM}} \perp a_t \mid \mathcal{F}^{\text{SP}}$, flagged in Remark 3.2 as an empirical simplification. The assumption is load-bearing only at the setpoint stage. At DA and ID the objective separates algebraically (Propositions 3.1 and 3.2), because the decision is a financial position. At SP, however, the setpoint physically truncates production via $y_t = \min(a_t, s_t)$, so the factorisation of the marginal effect into $\mathbb{E}[\pi_t^{\text{IM}} \mid \mathcal{F}^{\text{SP}}] \cdot \Pr(a_t > s_t \mid \mathcal{F}^{\text{SP}})$ requires the hypothesis. The SAA LP row in Table 7.6 already addresses this at the bake-off level: under the deployed regime ($\lambda = 1$, post-warm-up window) the dependence-aware rule underperforms ThreeRegion at every park. The mechanism is the same as at the intraday stage (Section 9.2): empirical mean–CVaR optimisation on $S = 250$ scenarios estimates the tail noisily, so the grid search picks suboptimal corners. This appendix adds two further tests. The first is a risk-neutral diagnostic ($\lambda = 0$, b_t at the deployed DA bid), which isolates whether the joint (a_t, π_t^{IM}) structure carries any mean-revenue signal beyond what the marginal forecast already gives ThreeRegion. The second is a model-free tail-dependence check (Section C.9.2) against the one objection a scenario-based optimiser cannot answer about itself.

C.9.1 Decision-Level Test: The Setpoint Rule Without the Assumption

If the joint (a_t, π_t^{IM}) structure carried decision-relevant signal, a rule free to exploit it would beat one that assumes it away on *mean* revenue alone, before any risk-aversion enters. The SAA LP policy maximises the empirical mean revenue $\frac{1}{S}\sum_s \pi_t^{\text{IM},(s)} \min(a_t^{(s)}, s_t) - f^{\text{imb}} | \min(a_t^{(s)}, s_t) - b_t|$ on the joint (a, π_t^{IM}) scenario fan of Appendix C.7 by a 1-D grid search over $s_t \in [0, P^{\max}]$, so it can exploit any conditional-on-curtailment structure the bootstrap fan carries. We run it at the same b_t the deployed bake-off uses (the DA bid) under risk-neutrality ($\lambda = 0$); per-slot price scenarios are re-centred on the marginal SP forecast so the price input matches the one ThreeRegion reads.

Table C.7 reports both policies and the no-curtailment baseline on the full post-warm-up window. SAA LP loses to ThreeRegion by only 0.5 to 3.4 index points and wins on 17 to 25 % of days, with the largest gap at Blabergsliden, the park with the strongest production–price coupling (Section C.9.2). Even with the freedom to exploit any conditional-on-curtailment structure the joint fan carries, the dependence-aware rule cannot beat the discrete classifier on mean revenue. The deployed-regime gap in Table 7.6 is therefore not what the joint structure could have bought; it is the noisy-empirical-CVaR mechanism of Section 9.2 reproduced at the setpoint stage.

Table C.7: SAA LP versus ThreeRegion (TR) and the no-curtailment baseline, mean daily revenue and gaps in index units (baseline = 100 per park), per park, full post-warm-up window, b_t at the deployed DA bid, risk-neutral ($\lambda = 0$), $S = 250$ scenarios with per-slot price re-centring on the marginal SP forecast.

		Mean daily revenue (index)			Gap (index)	
Park	n	NoCurt	TR	SAA	TR−NoCurt	SAA−TR
Blabergsliden	255	100.0	106.7	103.2	+6.7	−3.4
Blisterliden	103	100.0	101.3	100.4	+1.3	−1.0
Varsvik	226	100.0	101.4	100.9	+1.4	−0.5

C.9.2 Model-Free Backstop: The Decision-Relevant Tail Cell

A scenario-based optimiser that finds no signal is ambiguous on its own: the null could mean there is no exploitable structure, or that the bootstrap fan represents the joint structure poorly. A model-free check resolves the ambiguity by measuring production–price dependence directly in the data, with no fan and no optimiser. The check must be aimed correctly. The setpoint can only curtail, $y_t = \min(a_t, s_t)$, never add output, so the only consequential SP action is *downward* curtailment when the imbalance price is expected negative. The factorisation error of Proposition 3.3 is $\mathrm{Cov}(\pi_t^{\mathrm{IM}}, \mathbf{1}\{a_t > s_t\} \mid \mathcal{F}^{\mathrm{SP}})$, so the assumption can only change the curtailment decision or its payoff in one quadrant: high available production coinciding with a strongly negative imbalance price. That is the cell the system-balance channel of Appendix B.3 would populate. The opposite cell, low production into a positive price spike, is a bid-side forecast concern the setpoint cannot act on, and is the separate production-forecast-bias residual of Section 9.4, not this assumption.

With F_a and F_π the empirical marginals of a_t and the *signed* imbalance price π_t^{IM} over the post-warm-up window, the relevant signed co-exceedance coefficient at threshold u is

$$\lambda_{\mathrm{LU}}(u) = \Pr\big(F_a(a_t) > u \;\big|\; F_\pi(\pi_t^{\mathrm{IM}}) < 1 - u\big), \tag{C.5}$$

the probability that available production is in its upper tail given the imbalance price in its lower (strongly negative) tail. The subscript reads in conditioning order: L is the price in its *lower* tail, the conditioning event where the rule curtails, and U is production in its *upper* tail, the outcome that makes curtailment bite. Figure C.3 traces it over $u \in [0.80, 0.99]$ against the exact independence reference $1 - u$ (by construction of the rank transform, 0.05 at $u = 0.95$) and a copula-free day-block permutation null that breaks the production–price pairing while preserving each marginal and within-day structure. At $u = 0.95$ the coefficient sits at or within sampling error of the independence reference at every park: 0.04 at Blabergsliden (below 0.05), 0.06 at Varsvik, and 0.08 at Blisterliden with a 95 % interval $[0.03, 0.17]$ that spans the reference and stays inside the permutation-null envelope. The system-balance channel is real at the imbalance *price-level* scale, visible in the cross-correlation (Figure 5.1) and the wind-quintile gradient (appendix B.3), but it does not survive into tail co-exceedance in the per-slot residual the curtailment rule reads. The one coupling that could break the rule is empirically at the independence level, so the decision-level null is not an artefact of a poor scenario model.

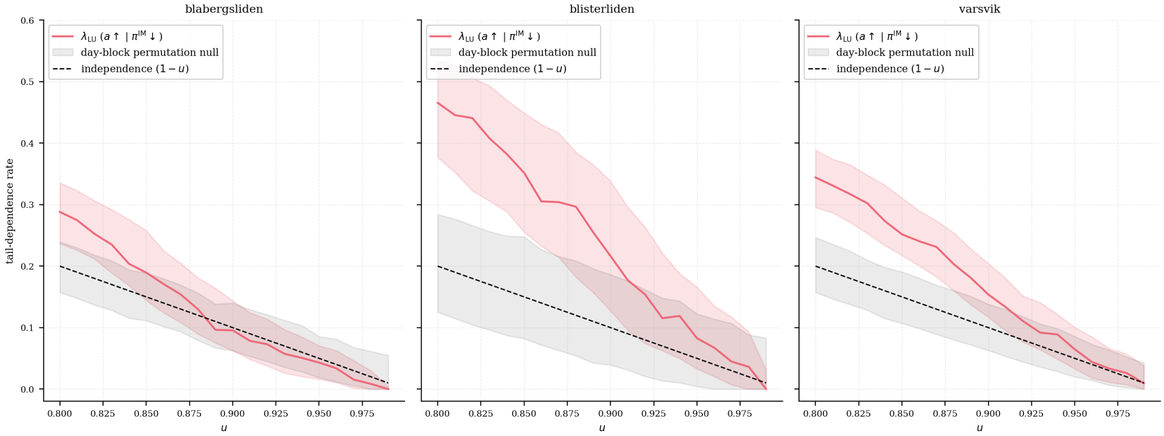

Figure C.3: Signed tail coefficient λ_{LU} (probability that production is in its upper tail given imbalance price in its lower tail) vs. threshold u, per park. Shaded band: 95 % bootstrap CI. Dashed line: exact independence reference $1-u$; grey band: permutation null.

C.9.3 Verdict and Scope

Three lines of evidence converge. The bake-off row (Table 7.6, $\lambda = 1$, deployed b_t) shows SAA LP losing to ThreeRegion at every park; the risk-neutral diagnostic (Table C.7, $\lambda = 0$, deployed b_t) shows the loss persists even before any risk-aversion enters, so the joint structure carries no decision-relevant mean-revenue signal to exploit; and the tail-coexceedance check (Figure C.3) measures the one cell that could break the curtailment rule at the independence level, so the decision-level null is not an artefact of a poor scenario model. The conditional-independence hypothesis of Proposition 3.3 is therefore not a costly simplification: the three-region rule is the empirically optimal rule, not merely the rule that is optimal once independence is assumed. This is distinct from the production-forecast-bias residual of Section 9.4, which remains open.

Populärvetenskaplig sammanfattning

Smarta bud på elmarknaden: hur vindkraftägare kan använda marknadsinformation för att tjäna mer och riskera mindre

Vindkraft är en av hörnstenarna i dagens och framtidens elsystem, men vinden är svår att förutsäga. Det skapar ett ekonomiskt problem: på elmarknaden måste producenter lägga bud på en viss mängd el redan dagen innan, utan att med säkerhet veta hur mycket vinden faktiskt kommer att producera. Om produktionen sedan avviker från det utlovade budet uppstår en *obalans*, som regleras till ett särskilt obalanspris. Ibland är det priset fördelaktigt, ibland medför det en kostnad. Det är under den osäkerheten som det lönar sig att vara smart.

Det här examensarbetet är utfört på uppdrag av Holmen Energi, som äger och driver tre vindkraftparker i Sverige. I takt med stigande och mer volatila obalanspriser i Norden har allt fler vindkraftproducenter börjat bära obalansrisken direkt, i stället för att lägga ut den på en extern balansansvarig part. Holmen Energi är ett exempel på den här utvecklingen och har sedan 2025 det direkta ekonomiska ansvaret för obalanserna i alla tre parker. Det ställde frågan på sin spets: hur ska man systematiskt hantera den osäkerhet som följer med att producera vind i ett volatilt elsystem?

Att justera sina bud under dagen. Mellan budet dagen innan och den faktiska leveransen finns en *intradagsmarknad* där producenter kan köpa och sälja el för att korrigera sina positioner. Ska man utnyttja den? Och i så fall, hur mycket ska man ändra sitt bud?

Svaret beror på vad marknaden just nu signalerar om obalansen förväntas bli lönsam eller kostsam. Examensarbetet visar att man kan använda obalansprisprognosen för att löpande avgöra huruvida, och hur kraftigt, budet ska justeras. Nyckeln är att vikta signalens trovärdighet: om prognosen är osäker är det klokt att hålla sig nära ursprungsbudet; om prognosen är tydlig och stark är det motiverat att agera mer beslutamt. Den regeln, en adaptiv budkorrigering skalad efter signalens tillförlitlighet, visade sig vara en regel som *både* höjde snittintäkten *och* minskade risken för stora förlustdagar vid samtliga tre parker. Skillnaden jämfört med att inte alls handla intradagsmarknaden var statistiskt säkerställd.

Att begränsa parkens produktion när priset är negativt. Ibland produceras det mer el i elnätet än vad som förbrukas, till exempel en blåsig helgnatt med låg efterfrågan. Då kan obalanspriset bli negativt, vilket innebär att producenten faktiskt *betalar* för att leverera el ingen vill ha. I sådana stunder kan det vara lönsamt att stänga ner turbinerna och begränsa parkens produktion.

Examensarbetet undersöker hur en kortsiktig prisprognos, tillgänglig ungefär tjugo minuter innan leveransen, kan användas för att fatta just det beslutet. Resultaten visar att ett enkelt beslutskriterium baserat på det prognosticerade obalanspriset fungerar väl: begränsa parkens produktion om prognosen signalerar ett tydligt negativt pris, håll dig till det inlämnade budet annars. Regeln aktiverades sällan, eftersom negativa priser fortfarande är ovanliga, men när den väl aktiverades var den i de flesta fall korrekt, och intäktsvinsten var meningsfull.

Vad det sammantaget innebär. Kombinationen av de två reglerna, adaptiv intradagsbudgivning och nedstyrning vid negativa priser, förbättrade det ekonomiska utfallet vid alla tre parker. Vid två av parkerna var risken för riktigt dåliga dagar liten redan från början, och reglerna tog i praktiken bort den helt. Vid den tredje, som var den enda med verkliga förlustdagar, minskade svansrisken tydligt men en del kvarstod, framför allt på enstaka dagar då produktionsprognosen var för hög och produktionen sedan föll oväntat.

Den övergripande slutsatsen är att en vindkraftägare med direkt balansansvar kan förbättra sin intjäning och minska sin risk med hjälp av enkla beslutsregler baserade på marknadens prisprognoser.

Master's Theses in Mathematical Sciences 2026:E74
ISSN 1404-6342

LUTFMS-3563-2026

Mathematical Statistics
Centre for Mathematical Sciences
Lund University
Box 118, SE-221 00 Lund, Sweden

http://www.maths.lu.se/