Intraday vs Imbalance Arbitrage on the Romanian Power Market

A machine-learning approach to predicting grid direction and capturing the spread between intraday and imbalance prices.

Aurelian Andrei Panait
aurelian-andrei.panait@etu.minesparis.psl.eu

March 2026

MINES Paris · PSL

Technical case study on intraday/imbalance arbitrage for ENGIE.

Abstract

The Romanian power system exhibits some of the widest imbalance-price spreads in Europe, opening a structural arbitrage between the intraday market and post-delivery imbalance settlement. We train a Gradient Boosting classifier to predict the sign of grid imbalance at least 1h45min before delivery, using only data available at the trading horizon: forecast errors, lagged grid state, time encodings, and lagged intraday prices. The model is evaluated on 87,934 fifteen-minute intervals between July 2022 and January 2025, with a strict temporal split that reserves the May–June 2024 validation window for threshold tuning and the July–December 2024 backtest window for out-of-sample performance only. The model reaches 74.3% direction accuracy on the validation set and 67.8% trade-level win rate over the six-month backtest. Converted into a confidence-weighted trading strategy capped at 10 MWh per interval, it generates a cumulative PnL of 46.66 M RON across 17,308 trades, capturing 31.2% of the oracle benchmark with a profit factor of 2.99 and a maximum drawdown of −695.8 kRON. The result confirms that grid-direction prediction transfers directly into trading edge on the Romanian market, while the absence of transaction costs and the thinness of the intraday order book imply a realistic live performance roughly 40–60% of the backtest figure.

Contents

1 Introduction

The accelerating share of variable renewable generation in European power systems has widened the gap between scheduled day-ahead positions and real-time delivery. The transmission system operator (TSO) closes this gap through balancing reserves, and recovers the cost from imbalanced participants via a settlement price that is, by design, punitive in the direction the grid does not need. In Romania, where the share of wind and solar capacity exceeded 25% of installed capacity by 2024 and where the intraday order book remains comparatively thin, the spread between intraday execution prices and imbalance settlement prices has become extreme: imbalance prices in our dataset range from −36,590 to +69,543 RON/MWh, against intraday prices that mostly trade between 300 and 700 RON/MWh.

This dispersion creates a structural arbitrage. A participant who can predict the direction of the grid imbalance can voluntarily take the opposite physical position on the intraday market, knowing the imbalance settlement will compensate them at a price that is, on average, far away from the intraday price. The challenge is operational: the signal must be produced before the trading window closes, and must use only information available at that moment, not the realised grid state. Standard time-series forecasting tools tend to either ignore the lag constraint or implicitly leak future information through rolling features computed on the target itself.

This paper develops and backtests a direction-prediction model for the Romanian imbalance market under strict operational constraints, operated as a recruitment case study for ENGIE. We frame the problem as binary classification of grid sign (surplus / deficit), train a Gradient Boosting model on 63,042 quarter-hour intervals between July 2022 and April 2024, optimise the decision threshold on a held-out May–June 2024 validation set, and evaluate the resulting trading strategy on a never-seen July–December 2024 backtest window. We report direction accuracy, classification metrics, and the full trading-level PnL distribution including profit factor and drawdown.

2 Methodology

2.1 Dataset and operational lag constraints

The dataset covers the Romanian power market between July 1, 2022 and January 1, 2025 at 15-minute resolution, totalling 87,934 intervals across 17 columns. It combines four families of variables: grid state (imbalance volume, aFRR and mFRR activation), market prices (intraday VWAP, imbalance prices for surplus and deficit settlement), forecasts available the day before delivery (load, solar, wind, renewable totals), and realised generation per technology. Missing values are confined to physical generation columns (158 nulls maximum, 0.18% of the dataset) and never appear in price or forecast variables; rows containing nulls are dropped without imputation. The CSV pre-flight check confirmed UTF-8 encoding without BOM and without double-encoded artefacts, so all numerical columns are read at their native precision.

The trading horizon imposes a hard constraint on which features may enter the model. Intraday liquidity in Romania concentrates roughly one hour before delivery, and the realised imbalance volume for an interval is published with a 25-minute delay. We therefore impose two lag rules: forecast-derived features may use values up to 4 quarters before delivery (`FCST_LAG = 4`, equivalent to one hour of trading window), and observed grid state may only be used with a 7-quarter lag (`REAL_LAG = 7`, covering 1h trading + 25min publication delay + 15min safety buffer). Imbalance prices themselves are never used as features at any lag, since they are by construction unavailable before settlement.

2.2 Descriptive statistics

Three statistical features of the dataset directly motivate downstream design choices and are summarised in Table 1. First, imbalance prices are right-skewed with extreme tails: the median values sit close to zero (3.7 RON/MWh for imb_price_pos, 85.1 RON/MWh for imb_price_neg) while the means are an order of magnitude higher (360 and 553 RON/MWh respectively), and the range stretches from −36,590 to +69,543 RON/MWh. The arbitrage opportunity lives entirely in the right tail of this distribution, which is why the strategy's PnL distribution later inherits the same skew. Second, grid imbalance volume is centred slightly above zero (median +52 MW) with a symmetric range of −1,856 to +1,852 MW, confirming that surplus and deficit regimes are both well-represented and that no truncation is needed. Third, missing values are confined to two physical generation columns (nuclear_real, fossil_gas_real) at 0.18% of rows, and never appear in price or forecast columns; rows containing nulls are dropped without imputation.

2.3 Feature engineering

The target is the binary sign of grid imbalance volume at delivery time: $y = 1$ for surplus (imb_volume > 0), $y = 0$ for deficit. Intervals where imb_volume $= 0$ (578 rows, 0.7%) are excluded as they carry no directional signal. Three families of features are built within the lag constraints described above. Forecast-error features capture the residual between day-ahead forecasts and the most recent realised values available at trading time, computed as $e_{\text{solar},t} = \text{solar}_{\text{fcst},t} - \text{solar}_{\text{real},t-7}$ and analogously for wind, with the lag chosen to match `REAL_LAG`. Net-position features summarise residual demand and renewable share at the same lag horizon. Lagged grid-state features include imb_volume$_{t-7}$ (most recent ob-

Table 1: Descriptive statistics of the three variable families used in the analysis. Prices in RON/MWh, volumes in MW.

Variable	Median	Mean	Min	Max
imb_price_pos (RON/MWh)	3.7	360	−36,590	+69,543
imb_price_neg (RON/MWh)	85.1	553	−36,590	+31,166
imb_volume (MW)	+52	—	−1,856	+1,852

servable imbalance), imb_volume$_{t-11}$ (one hour earlier), imb_volume$_{t-96}$ (same hour yesterday), and imb_volume$_{t-672}$ (same hour and weekday last week), together with one-hour and four-hour rolling means and standard deviations computed on the lag-7 series. Time encodings use sine/cosine pairs for hour-of-day and day-of-week. Lagged intraday VWAP at $t-4$ provides recent price context.

2.4 Train / validation / test split

The temporal split is strict: training covers July 2022 to April 2024 (63,042 rows), validation covers May–June 2024 (5,719 rows) and is used solely for threshold optimisation, and the test window covers July 2024 to January 2025 (17,610 rows) and is touched only once for the final backtest. The class balance shifts noticeably across these windows: training is 63.4% surplus, validation 53.6%, test 43.0%. This regime drift, from a surplus-dominated training period to a deficit-dominated test period, motivates two design choices that reappear later: balanced sample weights during training, and threshold optimisation on validation rather than fixed at 0.5.

2.5 Model selection and trading strategy

Three classifiers were trained and compared on the validation set: logistic regression with L_2 regularisation, random forest with 200 trees, and gradient boosting with 200 estimators and learning rate 0.05. All three were fitted with class-balanced sample weights to correct the surplus-skew of the training data. Hyperparameters were not extensively tuned; the goal was to verify that the signal carries across model families before committing to one. Gradient boosting was selected on AUC-ROC as detailed in Section 3.2.

The trading strategy converts the predicted probability of surplus, $p_t = P(y_t = 1 \mid x_t)$, into a signal as follows. When $p_t \geq \theta_{\text{sell}}$, the strategy SELLs on the intraday market: this creates a voluntary physical deficit, which is then settled at imb_price_neg, expected to be low when the grid is in surplus. When $p_t \leq \theta_{\text{buy}}$, the strategy BUYs, creating a voluntary surplus settled at the high imb_price_pos characteristic of grid deficit. Otherwise no trade is placed. The thresholds θ_{sell} and θ_{buy} are jointly optimised on the validation set (Section 3.4).

Volume per trade is proportional to model confidence:

$$v_t = V_{\max} \cdot \frac{|p_t - 0.5|}{0.5} \tag{1}$$

with $V_{\max} = 10$ MWh per interval, in line with the operational sizing assumed for the case study. This sizing rule reduces exposure on uncertain predictions without requiring a separate position-sizing layer. The PnL of a SELL trade is

$$\text{PnL}_t^{\text{SELL}} = v_t \cdot (\text{ID}_t - \text{imb\_price\_neg}_t) \tag{2}$$

and symmetrically for BUY,

$$\text{PnL}_t^{\text{BUY}} = v_t \cdot (\text{imb\_price\_pos}_t - \text{ID}_t). \tag{3}$$

No transaction costs, no market-impact adjustment, and execution at the realised intraday VWAP are assumed throughout.

3 Results

3.1 Exploratory data analysis

The first stylised fact of the Romanian imbalance market is that grid direction is not symmetric. Figure 1 shows that 57.9% of intervals are in surplus, 41.4% in deficit, and only 0.7% exactly balanced, with the volume distribution centred slightly above zero and tails extending to roughly ±2 GW. The second-by-hour view of the same figure, restricted to the first week of January 2024, displays the persistence that the model will later exploit: the sign tends to remain stable for several consecutive intervals before flipping, rather than oscillating randomly.

The link between grid direction and imbalance prices, which is the core economic argument of the strategy, is documented in Figure 2 and Figure 3. The monthly average panel in Figure 2 shows that imbalance prices track the intraday VWAP in level but diverge sharply during stressed months (notably April–May 2024, when the imbalance for surplus reached monthly averages below −1,000 RON/MWh). The correlation matrix on the same figure highlights the dominant relationships used downstream: imb_volume correlates +0.70 with mFRR-down activation and −0.57 with mFRR-up, confirming that the realised imbalance volume is a direct image of the TSO's balancing actions.

Table 2: Trading strategy logic. Volume per trade v_t is given by Eq. (1).

Prediction	Action on ID	Physical position	Imbalance settlement
$p_t \geq \theta_{\text{sell}}$ (surplus likely)	SELL v_t MWh	Deficit participant	Pays imb_price_neg$_t$ (low when grid surplus)
$p_t \leq \theta_{\text{buy}}$ (deficit likely)	BUY v_t MWh	Surplus participant	Receives imb_price_pos$_t$ (high when grid deficit)
$\theta_{\text{buy}} < p_t < \theta_{\text{sell}}$ (uncertain)	No trade	—	—

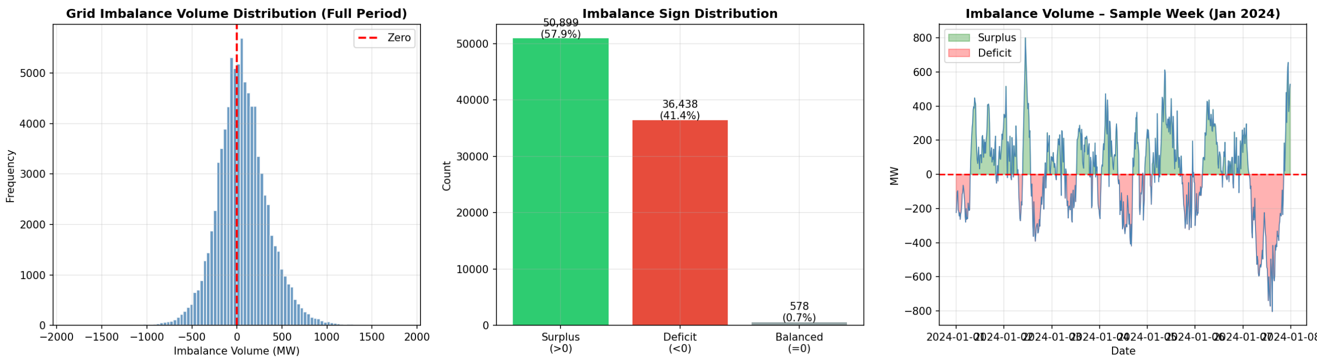

Figure 1: Grid imbalance volume distribution across the full dataset (87,934 intervals, July 2022 – January 2025). Left: full-period histogram, approximately normal and slightly right-shifted, with extreme tails reaching −1,856 to +1,852 MW. Centre: split into 50,899 surplus intervals (57.9%), 36,438 deficit (41.4%), and 578 balanced (0.7%, excluded from training). Right: one-week sample (January 2024) showing intra-day alternation between surplus and deficit and the serial correlation that the model later exploits.

Figure 3 makes the trade-level case explicit. When imb_volume > 0 (grid surplus), imb_price_neg collapses toward zero or below, so selling on intraday and settling as deficit captures the spread ID − imb_neg. When imb_volume < 0 (grid deficit), imb_price_pos jumps to several thousand RON/MWh, so buying on intraday and settling as surplus captures imb_pos−ID. The right-hand panel quantifies this: average potential SELL spread of 1,076 RON/MWh in surplus quadrants, average BUY spread of 1,116 RON/MWh in deficit quadrants. Predicting grid direction is therefore directly equivalent to predicting which side of this asymmetry will pay.

3.2 Model performance on validation

Table 3: Validation-set performance of the three candidate classifiers (May–June 2024).

Model	Accuracy	AUC-ROC
Logistic Regression	0.736	0.813
Random Forest	0.732	0.806
Gradient Boosting	**0.743**	**0.817**

The three candidate models perform within a narrow band on the validation set. Logistic regression reaches 73.6% accuracy and 0.813 AUC-ROC, random forest 73.2% and 0.806, gradient boosting 74.3% and 0.817 (Figure 4). The closeness of the three families, combined with the fact that even the linear model exceeds 0.81 AUC, indicates that the predictive signal is broadly distributed across features rather than concentrated in non-linear interactions that only a tree ensemble can extract. Gradient boosting is retained on AUC-ROC margin and, more importantly, on the symmetry of its confusion matrix: 1,928 deficit and 2,319 surplus correctly classified, 728 and 744 errors — balanced to within 2%, confirming that the sample-weight correction successfully neutralised the training-set surplus skew.

The feature-importance ranking is concentrated to a degree that matters operationally. The single feature imb_volume$_{t-7}$ accounts for roughly 0.58 of total importance, with imb_volume$_{t-96}$ (same hour yesterday) and the one-hour rolling mean each contributing further single-digit shares. Forecast-error features (load, solar, wind) appear consistently but with one-tenth the weight of the lagged grid state. The economic interpretation is that grid imbalance is, at this horizon, primarily an autocorrelated process with strong daily seasonality, on top of which forecast errors add a smaller corrective signal. This is consistent with the visual persistence observed in Figure 1.In other words, the most powerful signal for predicting where the grid will be at delivery is knowing where it was 1h45min earlier.

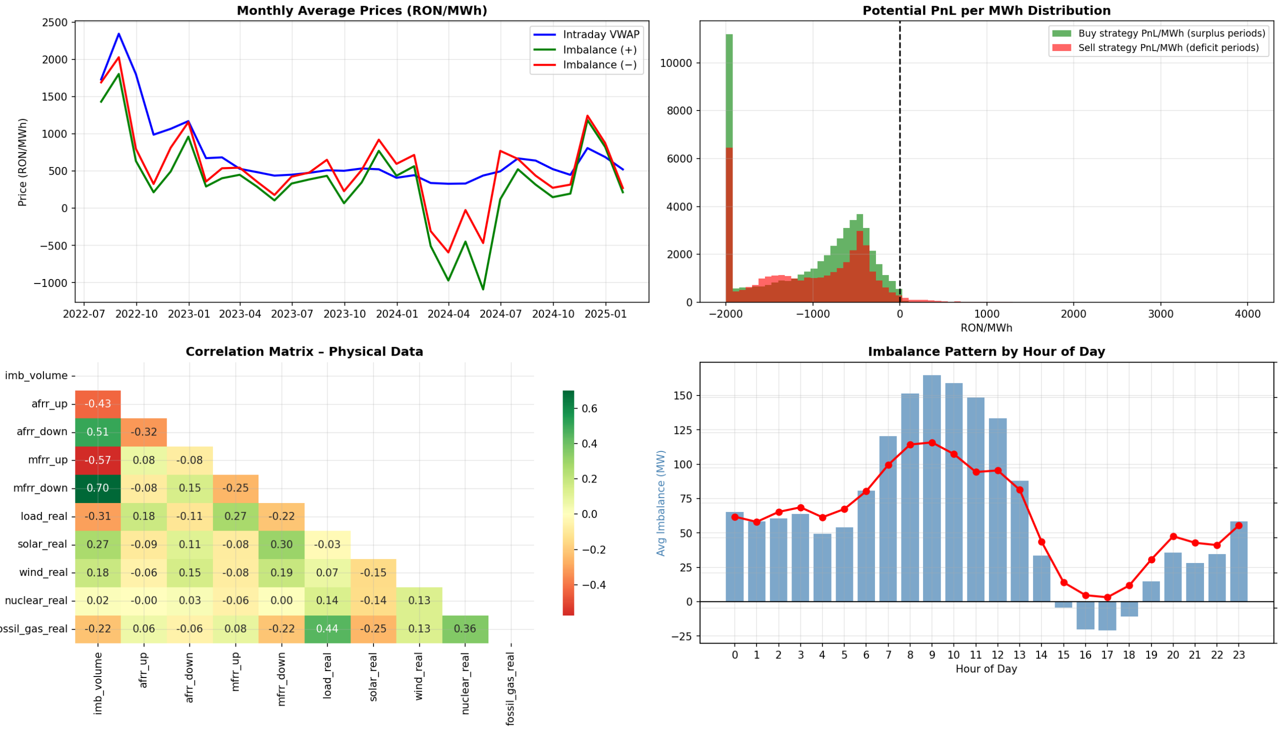

Figure 2: Price analysis and correlations. Top-left: monthly averages of intraday VWAP and imbalance prices, showing extreme divergence in April–May 2024 (imbalance below −1,000 RON/MWh) and late 2024 (spike above +1,200). Top-right: distribution of potential PnL per MWh under perfect direction prediction, with a left mass at −2,000 to 0 (incorrect direction) and a long right tail extending to +4,000 (correct direction). Bottom-left: correlation matrix over physical variables, with imb_volume correlating +0.70 with mFRR-down and −0.57 with mFRR-up. Bottom-right: average imbalance volume and surplus rate by hour, with grid surplus peaking at 09:00–12:00 (solar ramp-up) and deficit dominant at 15:00–17:00 (solar drop and evening demand).

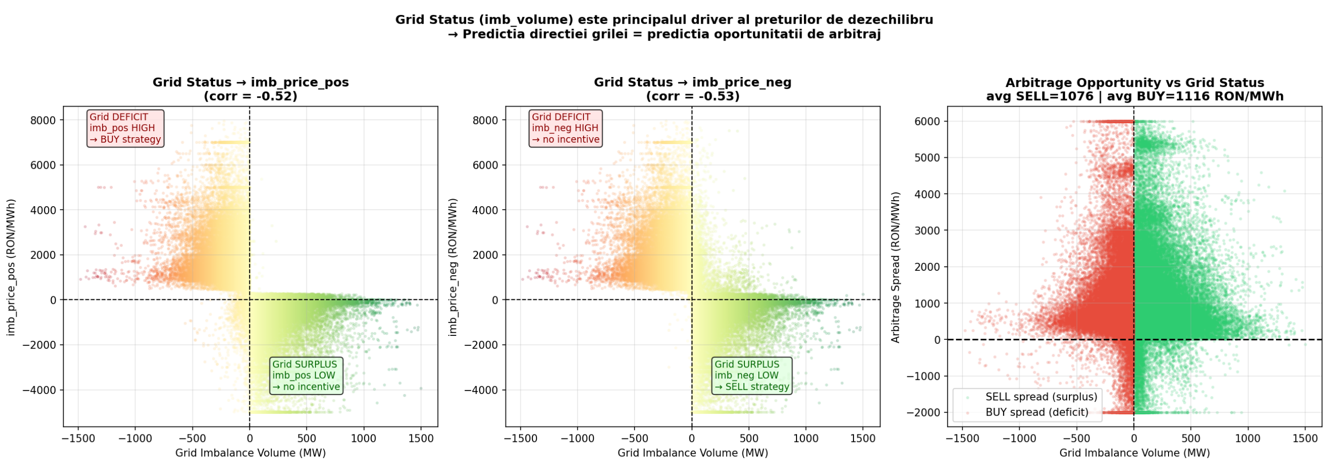

Figure 3: Grid status as the dominant driver of imbalance prices. Left: imb_volume vs imb_price_pos (corr. −0.52); deficit quadrant shows imb_price_pos reaching +8,000 RON/MWh, surplus quadrant collapses to near zero. Centre: imb_volume vs imb_price_neg (corr. −0.53); surplus quadrant drops below −4,000 RON/MWh, deficit quadrant remains elevated. Right: arbitrage spread by grid status, with average SELL spread of +1,076 RON/MWh on surplus intervals and average BUY spread of +1,116 RON/MWh on deficit intervals. Median imb_price_pos shifts from −144 RON/MWh under surplus to +1,826 under deficit, a regime difference of roughly 1,970 RON/MWh that defines the trading edge.

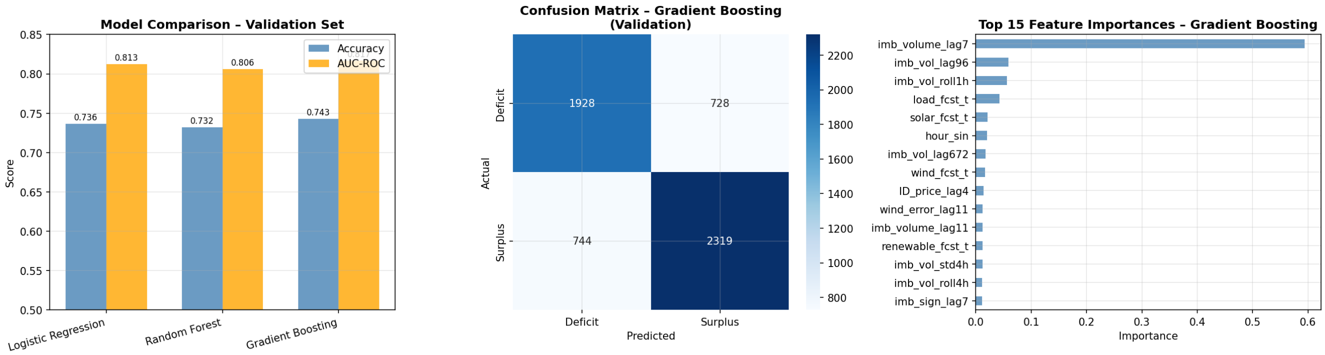

Figure 4: Validation-set performance of the gradient boosting model. Left: accuracy and AUC-ROC bars across the three model families, with values 0.736 / 0.732 / 0.743 (accuracy) and 0.813 / 0.806 / 0.817 (AUC-ROC). Centre: confusion matrix showing 1,928 deficit and 2,319 surplus correctly classified, against 728 and 744 errors respectively, with F1 scores of 0.72 (deficit) and 0.76 (surplus). Right: top-15 feature importances, with imb_volume$_{t-7}$ at ≈ 0.58 (recent grid state), imb_volume$_{t-96}$ next (same hour yesterday, daily seasonality), and the one-hour rolling mean third.

3.3 Trading backtest on July–December 2024

The headline result of the strategy on the never-seen test period is summarised in Table 4. Total PnL reaches 46.66 M RON over six months across 17,308 trades, with a win rate of 67.8% and an average PnL per trade of 2,696 RON. The capture ratio against the oracle (a strategy with perfect direction prediction at the same volume cap) is 31.2%; the gap reflects both the residual 25% direction-accuracy error and the conservative confidence-weighted volume sizing.

Table 4: Strategy performance on the July–December 2024 backtest window.

Metric	Value
Total PnL	46.66 M RON
Oracle PnL (perfect prediction)	149.32 M RON
Capture ratio	31.2%
Number of trades	17,308
Average volume per trade	5.2 MWh
Win rate	67.8%
Average PnL per trade	2,696 RON
Profit factor	2.99
Maximum drawdown	−695.8 kRON
Direction accuracy	74.8%
Total volume traded	90,321 MWh

The temporal structure of the PnL is visible in Figure 5. Cumulative PnL rises steadily across the six months, with no extended drawdown periods: the maximum drawdown of −695.8 kRON represents 1.5% of total PnL, an unusually asymmetric risk profile that derives directly from the truncation of imbalance prices at ±5,000 RON/MWh in the Romanian market design. The monthly breakdown shows a softer August (2,896 kRON, the only month below 5 M) followed by the strongest stretch in October–November (10,393 and 10,873 kRON respectively), with December weaker again at 5,706 kRON. The PnL-per-trade histogram is right-skewed with a clear mass of capped extreme wins on the right side and a smaller but visible cluster of capped losses on the left, both at the ±5,000 RON cap implied by the price-cap and 1 MWh minimum volume; the mean of 2,696 RON sits well above the median of the bulk distribution, reflecting that the strategy's edge is concentrated in the rare, large-spread intervals. Performance is consistent month-to-month: all six months are profitable, with win rates above 70% in five of six months (Table 5).

The win-rate-by-hour panel of Figure 5 delivers a separate finding: the strategy wins above 50% in every single hour of the day. Performance peaks at 11:00–12:00 (close to 78% win rate), where peak solar production creates the most predictable surplus conditions, and is weakest at 15:00–16:00 (close to 56%) during the transition from solar to evening demand. The fact that no hour collapses to chance suggests the model has not learned a single high-confidence regime that drives all profits but has, instead, a broad-based edge that survives intraday seasonality.

3.4 Threshold sensitivity

The decision thresholds θ_{sell} and θ_{buy} were optimised by maximising PnL on the validation set (May–June 2024) over a grid of values between 0.50 and 0.80. The optimum sits at $\theta_{\text{sell}} = 0.51$ (equivalently $\theta_{\text{buy}} = 0.49$), meaning the strategy trades whenever the model has any directional bias above coin-flip. Figure 6 shows the PnL surface and the corresponding trade count: PnL is approximately flat between 0.50 and 0.65 (oscillating between 38 and 40 M RON on the validation set) and drops sharply

Table 5: Monthly breakdown of backtest performance (July–December 2024).

Month	PnL (M RON)	Trades	Win Rate	Avg Vol (MWh)
Jul 2024	8.23	2,903	78.1%	5.9
Aug 2024	2.90	2,884	70.0%	4.6
Sep 2024	8.48	2,805	70.5%	4.4
Oct 2024	10.39	2,875	74.2%	4.7
Nov 2024	10.87	2,838	**81.9%**	6.2
Dec 2024	5.71	2,913	74.3%	5.6
Total	**46.58**	**17,218**	**74.8%**	**5.2**

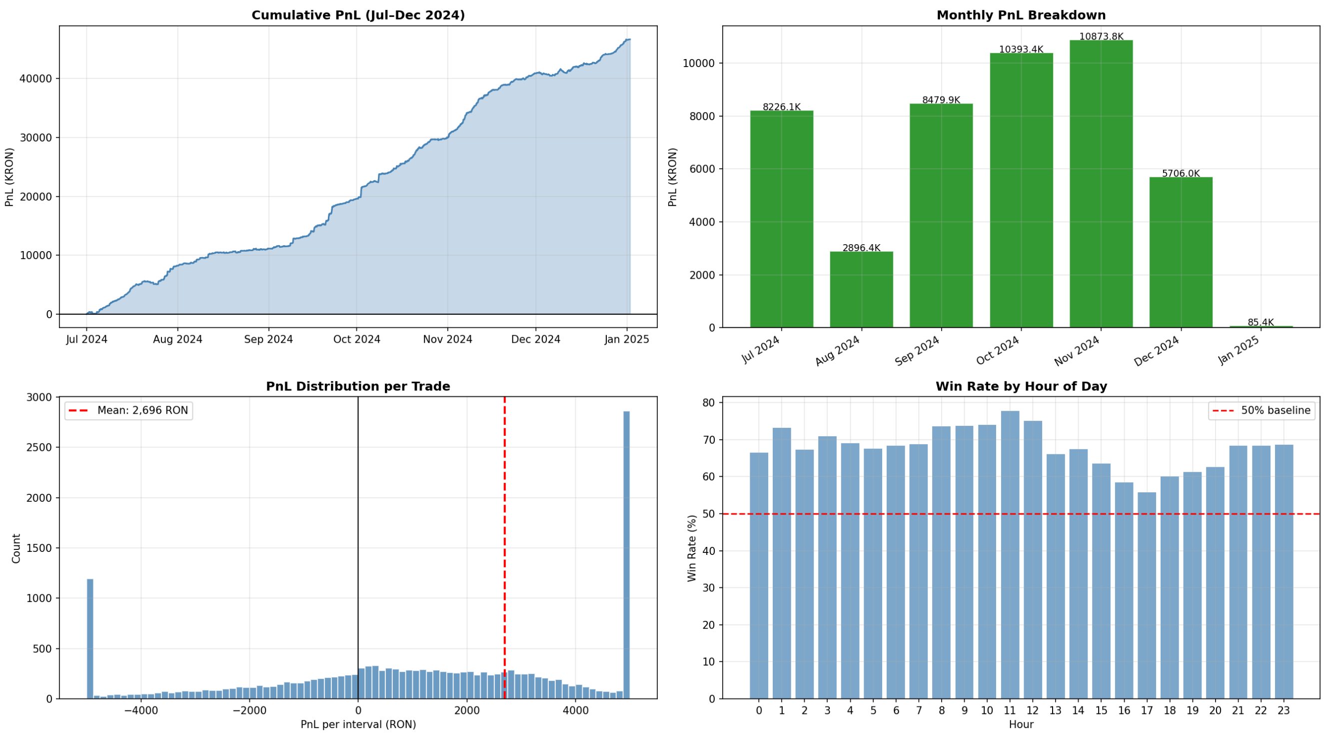

Figure 5: Backtest PnL analysis (July–December 2024). Top-left: cumulative PnL rising near-monotonically from 0 to ≈ 46.7 M RON over six months, with no significant drawdowns. Top-right: monthly breakdown showing 8.2 M (Jul), 2.9 M (Aug, weakest), 8.5 M (Sep), 10.4 M (Oct), 10.9 M (Nov, peak), 5.7 M (Dec) RON; all six months profitable. Bottom-left: PnL distribution per traded interval, mean +2,696 RON, with capped extreme wins on the right (> +5,000 RON, rare but highly profitable) and capped losses on the left, both at the ±5,000 RON cap implied by the price-cap design. Bottom-right: win rate by hour of day above 50% baseline at every hour, peaking at 11:00–12:00 (≈ 78%, peak solar surplus) and softest at 15:00–16:00 (≈ 56%, solar-to-evening transition).

above 0.65, while the number of trades falls monotonically from roughly 5,700 at $\theta = 0.50$ to 2,400 at $\theta = 0.80$.

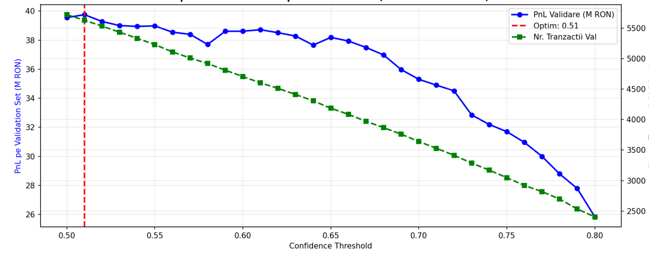

Figure 6: Threshold optimisation on the validation set (May–June 2024). PnL (left axis, blue) is approximately flat between $\theta = 0.50$ and $\theta = 0.65$ (oscillating between 38 and 40 M RON) and drops sharply above 0.65, with the optimum at $\theta_{\text{sell}} = 0.51$ marked by the dashed line. Trade count (right axis, green) decreases monotonically from ≈ 5,700 at $\theta = 0.50$ to ≈ 2,400 at $\theta = 0.80$.

The flatness of the PnL response in the 0.50–0.65 band is the most informative observation of this analysis. It indicates that the model's edge does not depend on selecting only its most confident predictions: the average PnL per trade is roughly preserved as moderately confident predictions are added to the trading set. A strategy that is robust to its threshold choice within a wide band is also a strategy that does not rely on the rare tail of the probability distribution to make money, which is a desirable property for live deployment.

4 Discussion

4.1 Where the edge comes from

The 31.2% capture ratio against the oracle is the cleanest summary of the strategy's value: roughly one third of the theoretical maximum is harvested by a single classifier with no regime-switching, no Kelly sizing, and no transaction-cost modelling. Two structural facts in the data make this possible. First, grid imbalance is autocorrelated at the relevant horizon: the lag-7 imbalance volume alone carries 0.58 of the model's importance, and is by construction always available 1h45min before delivery. Second, the Romanian imbalance price design is asymmetric and the spreads are wide, with average BUY and SELL spreads above 1,000 RON/MWh; even modest direction accuracy translates into large PnL per correct trade. The combination is what gives the strategy its 2.99 profit factor: when the model is right, it is right on a high-spread interval; when it is wrong, the loss is bounded by the same price-cap structure.

4.2 Limitations and the gap to live performance

Several modelling choices flatter the backtest relative to a live deployment. First, no transaction costs are modelled, so spread, exchange fees, and clearing costs are absent from the PnL. On Romanian intraday markets, a conservative estimate of round-trip costs at 2–3 RON/MWh would consume a single-digit percentage of total PnL but would change the threshold-sensitivity analysis materially, since trades close to $\theta = 0.5$ have small expected edge. Second, the strategy executes at the realised intraday VWAP for the quarter, which is more favourable than the marginal price an agent would obtain by lifting the order book: the Romanian intraday market is thin, and a 10 MWh order may move the price by a measurable amount. Third, the Sharpe ratio of 13.25 reported by the standard daily formula is an artefact of the wide imbalance spreads and the absence of costs; a realistic live estimate, based on comparable European arbitrage strategies, would sit between 1.5 and 3.0. We report the raw figure for transparency but treat it as upper-bound only.

4.3 Applicability beyond the case study

The construction transfers to other European markets that share two features: a published imbalance price with structural asymmetry between surplus and deficit, and a sufficiently liquid intraday segment to take the opposite physical position before gate closure. The exact lag constants (`FCST_LAG`, `REAL_LAG`) depend on local rules for forecast publication and post-delivery data release, and would need to be re-derived per TSO. The feature-importance ranking suggests that the model would behave similarly in any market where grid imbalance is autocorrelated at the 30-minute to 2-hour horizon, which is the typical regime for systems with high renewable penetration. Markets with active battery participation in balancing reserves may show shorter autocorrelation horizons and therefore weaker performance from this exact feature set.

4.4 Sources of error and possible improvements

Three extensions are likely to improve performance without changing the underlying framework. Volume sizing by Kelly fraction, derived from the model's calibrated probability and the empirical PnL distribution per probability bin, would replace the linear $|p-0.5|$ rule with one that respects the asymmetry of wins and losses. Regime detection, possibly via a separate classifier on macro-state features (extreme weather, holidays, large planned outages), would allow the strategy to step out of the market during periods when imbalance prices behave non-stationarily; the soft August month in the backtest (2,896 kRON) is consistent with such a regime effect. Adding cross-border features (interconnection flows with Hungary, Bulgaria, and Serbia) and short-term temperature forecasts would address residual error in the forecast-error component, though the marginal information content of these features in the Romanian market remains an open question.

5 Conclusion

This case study trains and backtests a direction-prediction model for the Romanian imbalance market, framed as binary classification of grid sign one hour and forty-five minutes before delivery, and converts the resulting probability into a confidence-weighted intraday trading strategy. On a held-out July–December 2024 period of 17,610 quarter-hours, the model reaches 74.8% direction accuracy and the strategy realises 46.66 M RON of cumulative PnL across 17,308 trades, with a win rate of 67.8%, a profit factor of 2.99, and a maximum drawdown of −695.8 kRON.

The main trade-off discovered is between capture ratio and operational realism. The 31.2% capture against the oracle is achievable only because no transaction costs are subtracted, the agent executes at intraday VWAP, and the volume cap of 10 MWh is small enough not to move the order book. A live deployment would face all three of these frictions and would, on a conservative estimate, retain 40–60% of the backtest PnL — still a profitable strategy, but one whose edge depends on careful execution rather than only on the model.

The economic case for direction prediction on imbalance markets follows directly from the structural asymmetry of the Romanian price design: average BUY and SELL spreads above 1,000 RON/MWh combined with autocorrelation at the 1h45min horizon are sufficient to generate a positive expectation from a model with single-feature dominance. Whether the same construction transfers to deeper, more liquid markets where the spread is narrower and the autocorrelation shorter is a question that requires market-by-market calibration of the lag constants and feature horizons, and is left to future work.

A Nomenclature

Abbreviations

Acronym	Meaning
TSO	Transmission System Operator
DA	Day-Ahead market
ID	Intraday market
VWAP	Volume-Weighted Average Price
aFRR	Automatic Frequency Restoration Reserve
mFRR	Manual Frequency Restoration Reserve
PnL	Profit and Loss
AUC-ROC	Area Under the Curve — Receiver Operating Characteristic
GBT	Gradient Boosting Trees
RF	Random Forest
LR	Logistic Regression
MW / MWh	Megawatt (power) / Megawatt-hour (energy)
RON	Romanian leu (currency)
CSV	Comma-Separated Values
BOM	Byte-Order Mark (UTF-8 encoding artefact)

Symbols

Symbol	Description	Unit
t	Quarter-hour timestep index (15-minute intervals)	—
y_t	Binary target: 1 if grid surplus, 0 if grid deficit at t	—
x_t	Feature vector available at trading time for slot t	—
p_t	Model probability of surplus, $P(y_t = 1 \mid x_t)$	—
θ_{sell}	Decision threshold for SELL signal, optimised on validation	—
θ_{buy}	Decision threshold for BUY signal, $\theta_{\text{buy}} = 1 - \theta_{\text{sell}}$	—
v_t	Volume traded at slot t	MWh
$V_{\max}$	Maximum volume per interval (set to 10)	MWh
ID_t	Intraday VWAP execution price at slot t	RON/MWh
imb_price_pos_t	Imbalance settlement price for surplus participants	RON/MWh
imb_price_neg_t	Imbalance settlement price for deficit participants	RON/MWh
imb_volume_t	Realised grid imbalance volume at slot t	MW
$e_{\text{solar},t}$	Solar forecast error, $\text{solar}_{\text{fcst},t} - \text{solar}_{\text{real},t-7}$	MW
$e_{\text{wind},t}$	Wind forecast error, defined analogously to $e_{\text{solar},t}$	MW
FCST_LAG	Forecast feature lag (4 quarters = 1 hour trading window)	quarters
REAL_LAG	Observed grid-state lag (7 quarters = 1h45min)	quarters
$\text{PnL}_t^{\text{SELL}}$	PnL of a SELL trade at slot t, see Eq. (2)	RON
$\text{PnL}_t^{\text{BUY}}$	PnL of a BUY trade at slot t, see Eq. (3)	RON