Tertiary Reserve Spread & SPOT Bidding-Curve Sensitivity

Anticipating Tertiary-Reserve Price Spikes from the Shape of the Aggregated SPOT Bidding Curves

Aurelian Andrei Panait

MINES Paris · PSL

aurelian-andrei.panait@etu.minesparis.psl.eu

June 2026

A technical case study on the French balancing market for a renewable energy aggregator

Abstract

On the French balancing market, energy procured through upward tertiary reserve is settled at a marginal price that is normally above the day-ahead SPOT price but occasionally spikes far above it. This study measures whether the *shape* of the aggregated SPOT bidding curves can anticipate those spikes. Three quarter-hourly datasets covering 1 March to 27 May 2026 — the aggregated SPOT bidding curves, the system imbalance with tertiary activation, and the tertiary marginal prices — are aligned to a common CET clock and joined into 8348 common steps. Each supply curve is reduced to two interpretable features: the cheap volume offered just above clearing (headroom_{100}) and the curve steepness (slope_{2000}). The upward tertiary spread is almost always positive ($\sim$91% of steps) with a fat right tail reaching the 4000 €/MWh cap, so a spike is defined empirically as the top decile (q90 $\approx$ 270 €/MWh). The relationship between curve shape and the spread is not linear ($|\text{Pearson}| \lesssim 0.1$) but exhibits a clear threshold effect: the spike rate jumps from $\sim$3% to $\sim$28% in the top activation quintile and rises monotonically as headroom shrinks, reaching $\sim$23% (about 4× the base rate) when a tight curve meets a large activation. A transparent, unfitted ex-ante risk indicator combining low headroom, high |imbalance| and high price flags risky hours with a $\sim$2.2× lift over the 10% base rate — a usable early-warning filter and a baseline for a fitted model.

Contents

1 Introduction

On the SPOT day-ahead market, generation and consumption are planned a day in advance, but real time never matches the plan exactly. The transmission system operator (RTE in France) corrects the residual gap by activating reserves in successive layers: FCR (instant, automatic), aFRR (fast, automatic), and finally tertiary / mFRR & RR (manual, last resort, most expensive). Tertiary energy is bought at a marginal price that is normally a little above the SPOT clearing price but occasionally spikes far above it. The difference between the two is the object of this study.

The difficulty of anticipating these spikes is one of timing. The single strongest explanatory variable — how much reserve the TSO ends up activating — is only known *after* the activation, so it cannot be used to flag risk in advance. A useful indicator must instead rely on information available once the day-ahead auction clears, in particular the shape of the aggregated supply curve. The central question is whether that shape, measured just above the clearing point, carries enough information about spike risk to be operationally useful.

This paper reproduces the full chain on the provided data: aligning the three sources to a common clock, building the tertiary spread, reducing each bidding curve to two interpretable shape features, quantifying their relationship with the spread, and assembling a simple ex-ante risk indicator. The analysis follows the three tasks of the exercise — exploration, correlation, and a risk indicator — and pays particular attention to the non-linearity of the relationship, the threshold at which a tight curve becomes dangerous, and the trade-off between predictive power and ex-ante availability.

2 Methodology

2.1 Data

The dataset comprises three files at 15 min resolution covering 1 March to 27 May 2026. The bidding-curve file (`aggregated_ccv_reconstructible.parquet`) contains, for each timestamp, the points of the aggregated supply and demand curves together with the SPOT clearing price and volume. The imbalance file (`imbalance_and_marginal_data.xlsx`) contains the system imbalance, the system trend, and the tertiary active power activated upward and downward; empty activation rows indicate that the TSO balanced the grid with other reserves (aFRR/FCR) rather than tertiary, and are kept as "not activated" rather than zero. The marginal-price file (`marginal_prices_energy_reserve.xlsx`) contains the marginal price of the tertiary offer activated by the TSO in each direction.

A subtlety that matters for the join is that the three files store time differently — one in naive local time, one with an explicit `+02:00` offset, one in UTC. The ingestion pipeline places all three on a common CET clock, localising naive timestamps and converting offset-aware ones, which handles the late-March daylight-saving transition correctly. All three sources are quarter-hourly and align directly on the timestamp; the inner join retains 8348 common steps.

2.2 The tertiary spread

The quantity studied is the upward tertiary spread, defined per timestep as

$$\text{spread\_up} = \text{mp\_up} - \text{clearing\_price}, \tag{1}$$

where mp_up is the upward tertiary marginal price and clearing_price is the SPOT clearing price. It is the extra cost per MWh of procuring energy through upward tertiary reserve instead of having bought it a day ahead on SPOT. The study focuses on the upward direction, where the financially material spikes occur; the downward spread is computed for completeness but not analysed.

A spike is defined empirically rather than by a fixed price level. As shown in Section 3, the spread distribution is far from Gaussian, so a fixed or normal-model threshold would misrepresent tail risk. A spike is therefore any step whose spread exceeds the empirical 90th percentile,

$$\text{spread\_up} > q_{90} \approx 270\,€/\text{MWh}, \tag{2}$$

which sets the base spike rate to 10% by construction.

2.3 Bidding-curve features

Each aggregated supply curve is summarised by two features that measure its shape just above the clearing point, where any upward reserve need would be served. The headroom is the cumulative supply volume offered within a 100 €/MWh band above the clearing price,

$$\text{headroom}_{100} = V(\text{cp} + 100) - V(\text{cp})\,, \tag{3}$$

where cp is the clearing price and $V(p)$ is the cumulative supply volume up to price p; a large value means abundant cheap slack (a relaxed curve), a small value a tight one. The slope is the price increment needed to add 2000 MW of supply on top of the cleared volume,

$$\text{slope}_{2000} = p(V_c + 2000) - \text{cp}, \tag{4}$$

where V_c is the cleared volume; a large value means a steep curve. Both bands are fixed reference windows; since the features are used only in a relative sense, the exact band widths affect the absolute numbers but not the tight-versus-relaxed ranking.

2.4 Ex-ante risk indicator

Because activated power is known only *ex post*, the risk indicator is built from three quantities available before activation: the curve tightness headroom_{100}, the imbalance magnitude $|\text{imbalance}|$ (magnitude rather than sign, since the imbalance distribution is symmetric), and the price level clearing_price. They are combined into a transparent, unfitted standardised sum

$$\begin{aligned} \text{RISK} = z(-\text{headroom}_{100}) + z(|\text{imbalance}|) \\ + z(\text{clearing\_price}) \,, \end{aligned} \tag{5}$$

where the z-score $z(x) = (x - \mu)/\sigma$ standardises each variable to zero mean and unit variance. Standardisation places the three quantities, which carry different units, on a common dimensionless scale so they can be summed; equal weights and the absence of fitting keep the indicator interpretable and free of over-fitting.

3 Exploration

The aggregated bidding curves are the raw material of the study. For a single timestamp, stacking all sell offers cheapest-first gives the supply curve and stacking all buy offers gives the demand curve; their intersection is the clearing price and volume (Figure 1). The steepness of the supply curve just above this point governs how expensive any additional reserve need will be.

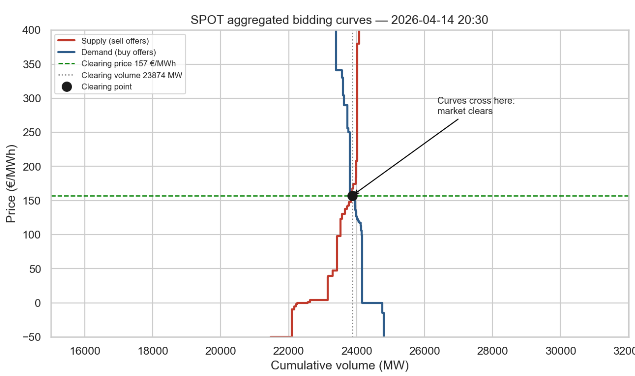

Figure 1: Aggregated SPOT bidding curves for one timestamp. Supply (red) stacks sell offers cheapest-first, demand (blue) stacks buy offers; the two cross at the clearing point.

The tertiary spread of Equation (1) has a distinctive distribution (Figure 2). It is almost always positive ($\sim$ 91% of steps), centred near a median of 65 €/MWh, but with a fat right tail: the mean ($\approx$ 100 €/MWh) exceeds the median, the standard deviation ($\approx$ 169 €/MWh) exceeds the mean, and the maximum reaches the 4000 €/MWh technical cap. The middle 50% of values sits between roughly 27 and 117 €/MWh, yet the top decile stretches all the way to the cap.

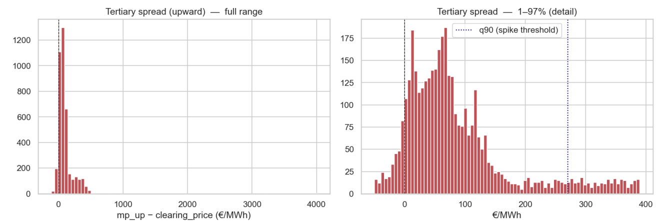

Figure 2: Distribution of the upward tertiary spread. Left: full range, dominated by the tail to the 4000 €/MWh cap. Right: the 1–97% detail, with the q90 spike threshold marked.

Both the mean exceeding the median and the standard deviation exceeding the mean are signatures of a heavy right tail — the reason a normal model would badly under-estimate tail risk, and the justification for the empirical spike definition of Equation (2). The distributions of the four core variables (clearing price, marginal price, imbalance, activated power) are collected in Appendix Figure 9, and the activation structure — direction following the sign of imbalance — in Appendix Figure 8.

4 Correlation and threshold effect

4.1 Reading the two curve features

The two features of Equations (3)–(4) are best understood on two real hours (Figure 3). In the relaxed hour the supply curve is nearly flat above clearing, so the 100 € band holds a large cheap volume ($\sim$6500 MW) and any extra reserve need is cheap. In the tense hour the curve rises almost vertically right after clearing, so the same band holds almost nothing ($\sim$145 MW); the spread that hour was $\sim$790 €/MWh. This contrast is exactly what headroom_{100} captures. The distributions of both features appear in Appendix Figure 10; notably, slope_{2000} saturates near the 4000 €/MWh cap, which limits its discriminating power.

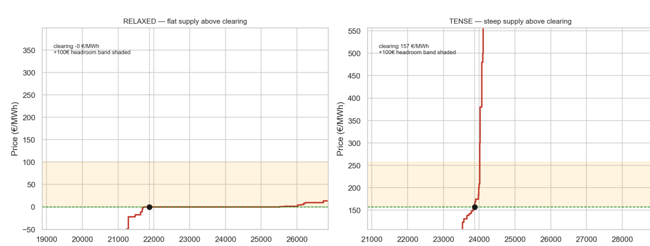

Figure 3: The 100 €/MWh band (shaded) above clearing in a relaxed hour (left) and a tense hour (right). A flat supply curve holds abundant cheap volume; a steep one holds almost none, making any reserve need expensive.

4.2 The relationship is not linear

The link between curve shape and the spread is weak under linear correlation. Figure 4 compares the absolute Pearson (linear) and Spearman (monotonic, rank) correlations of each candidate factor with the spread. All linear correlations are weak ($|\text{Pearson}| \lesssim 0.1$ for the curve features), but for imbalance and activated power the Spearman correlation is markedly larger than Pearson (-0.18 vs -0.08; 0.40 vs 0.23), the signature of a monotonic but non-linear relationship rather than a proportional one.

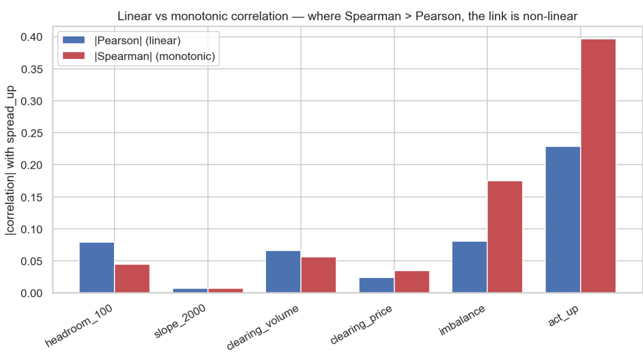

Figure 4: Absolute Pearson (linear) versus Spearman (monotonic) correlation with the spread. Where the Spearman bar clearly exceeds the Pearson bar (imbalance, activated power), the relationship is monotonic but non-linear.

Two readings follow. First, activated power is the strongest signal, but it is known only *ex post* and so cannot serve as a predictor — a point that motivates the ex-ante indicator of Section 5. Second, headroom looks weak in this table because a simple pairwise correlation cannot capture an interaction: its effect on the spread appears only when combined with a large activation, as the next subsection shows. The raw scatter of spread against activated power is given in Appendix Figure 11, where heavy overlap hides the pattern that the quintile analysis recovers.

4.3 A threshold effect

The weak linear correlations motivate testing for a threshold (regime) effect, in which a variable barely matters until it crosses a critical level and then its impact jumps. Bucketing each factor into quintiles and measuring the spike rate per bucket reveals exactly this (Figure 5). For activated power the spike rate is flat at $\sim$3–4% across the bottom three quintiles, then jumps to $\sim$12% (Q4) and $\sim$28% (Q5), with the break around 250 MWh per step. For headroom the spike rate falls monotonically with supply slack, from $\sim$17% for the tightest curves to $\sim$3% for the most relaxed. The slope feature is non-monotonic and noisier, because it saturates near the price cap and so rank-orders spikes less reliably than headroom.

The two factors matter together rather than separately. Combining low headroom *and* high activation lifts the spike rate to $\sim 23\%$, about $4\times$ the base rate, whereas neither factor alone is sufficient. The SPOT bidding curve does carry spike information, but only non-linearly: it matters most when a tight curve meets a large reserve need.

5 Ex-ante risk indicator

The indicator of Equation (5) uses only quantities available before activation, and is assessed with precision (of flagged hours, the share that are real spikes), recall (of all spikes, the share caught), and lift (precision divided by the base rate). The results are consolidated in Table 1. Flagging the top 10% of hours by RISK gives $\sim$22% precision against a 10% base rate — a $\sim$2.2$\times$ lift — while catching about a fifth of all spikes; relaxing to the top 30% raises recall to $\sim$49% at lower precision. There is a clear precision–recall trade-off set by how many hours are flagged.

Flag rule	Precision	Recall	Lift
Top 30%	16.5%	49.4%	1.65×
Top 20%	19.0%	37.9%	1.90×
Top 10%	22.4%	22.4%	2.24×

Table 1: Ex-ante indicator performance at three flagging thresholds. The base spike rate is 10% by construction.

The indicator rank-orders risk monotonically (Figure 6): the spike rate rises from $\sim$ 2% in the lowest RISK decile to $\sim$22% in the highest. Discrimination is sharpest at the extremes; the middle deciles are flat near the base rate, so the indicator separates clearly-safe and clearly-risky hours better than it orders the ambiguous middle.

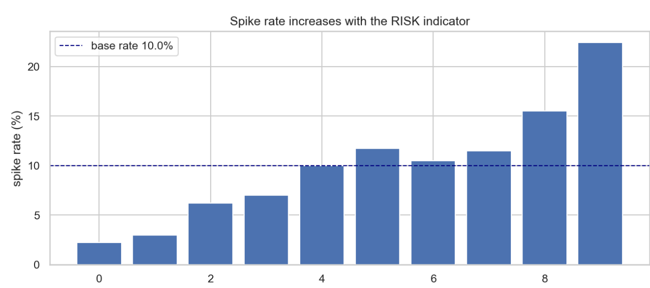

Figure 6: Spike rate by RISK decile (base rate 10%, dashed line). The monotonic increase confirms that the indicator rank-orders risk; discrimination is sharpest at the extremes.

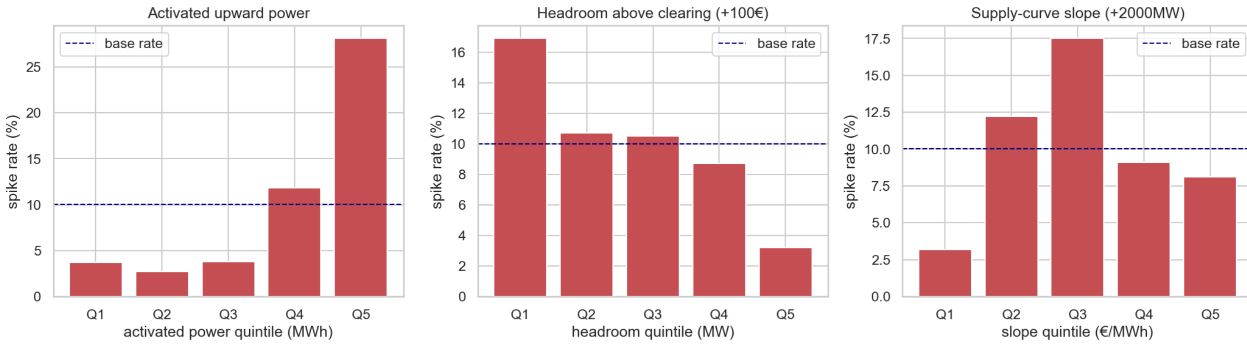

Figure 5: Spike rate by quintile (base rate 10%, dashed line). Activated power shows a sharp threshold in the top quintile; headroom declines monotonically as the curve relaxes; slope is noisier due to saturation at the price cap.

6 Discussion

6.1 Why the relationship is non-linear

The threshold effect of Section 4.3 has a physical reading. The spread does not respond proportionally to any single factor because a tertiary spike is the result of a coincidence: a reserve need that lands on a part of the supply curve with no cheap volume left to serve it. When the curve is relaxed, even a large activation is absorbed cheaply; when it is tight, even a moderate activation climbs the steep part of the curve. Neither tightness nor activation alone determines the outcome — it is their conjunction, which is precisely why a linear correlation on either one is weak and a quintile interaction is strong.

6.2 The ex-ante / ex-post trade-off

The central operational tension is between predictive power and timing. Activated power is the best single predictor of a spike, but it is realised only after the TSO has acted, so a desk cannot use it to reposition in advance. The indicator of Equation (5) deliberately discards this strongest signal in exchange for being computable once the day-ahead auction clears. The resulting $\sim$2.2$\times$ lift is modest but real: it is enough to make the indicator useful as an early-warning filter that concentrates attention on the riskiest hours, while being far from a precise forecaster. The flat middle deciles of Figure 6 make this explicit — the indicator is informative at the extremes and uninformative in the ambiguous middle.

6.3 Limitations

Three limitations temper the result. First, the study covers a single three-month window; spike behaviour may differ across seasons, and a longer sample would test the stability of both the q90 threshold and the quintile breaks. Second, the curve features use fixed bands (100 €/MWh, 2000 MW); a feature measured at the *expected* activation volume would tie the curve shape more directly to the reserve need actually faced. Third, the indicator is unfitted and equal-weighted by design; a fitted model would almost certainly improve the lift, at the cost of the transparency that makes the present indicator easy to audit and reason about.

6.4 Extensions

Four extensions follow directly. Replacing the equal-weight composite with a fitted logistic regression or gradient-boosting model on the same ex-ante features should yield a clear gain while keeping the inputs unchanged. Adding temporal features — hour-of-day, recent imbalance ramp, spike persistence — would exploit the fact that spikes cluster in time, visible in the price time series of Appendix Figure 12 and the hour-of-week structure of Appendix Figure 13. Using a directional curve feature measured at the expected activation volume would sharpen the link between curve shape and reserve need. Finally, calibrating the spike threshold to the business cost of a missed spike versus a false alarm, rather than a fixed q90, would align the indicator with the desk's actual asymmetric loss.

7 Conclusion

This study measures whether the shape of the aggregated SPOT bidding curves can anticipate tertiary-reserve price spikes on the French balancing market. The upward tertiary spread is almost always positive with a fat tail to the 4000 €/MWh cap, so spikes are best defined empirically as the top decile. The relationship between curve shape and the spread is not linear, but a threshold effect is clear: the spike rate jumps once activation is large and once the curve is tight, reaching about 4$\times$ the base rate when both conditions hold. A transparent ex-ante indicator built from curve tightness, imbalance magnitude and price level achieves a modest but real $\sim$2.2$\times$ lift, usable as an early-warning filter rather than a precise forecaster.

Beyond the headline result, the analysis isolates the structure that any improved model would build on: the non-linearity is an interaction, not a missing transformation; the strongest signal is unavailable ex ante by construction; and the useful information lives at the extremes of the risk distribution. The computational backbone — alignment, spread construction, curve featurisation, and evaluation — is settled; the interesting remaining work is in fitting the interaction and calibrating the threshold to the operational cost of error.

A Nomenclature

Abbreviations

Acronym	Meaning
RTE	*Réseau de Transport d'Électricité* — French Transmission System Operator
TSO	Transmission System Operator
SPOT	Day-ahead wholesale electricity market
FCR	Frequency Containment Reserve (instant, automatic)
aFRR	automatic Frequency Restoration Reserve (fast, automatic)
mFRR	manual Frequency Restoration Reserve
RR	Replacement Reserve
Tertiary	The mFRR & RR layer — manual, last-resort balancing reserve
CET	Central European Time (CET/CEST, with daylight saving)
q90	90th percentile (empirical spike threshold)

Symbols

Symbol	Description	Unit
clearing_price	SPOT day-ahead clearing price	€/MWh
clearing_volume	SPOT cleared volume	MW
mp_up	Upward tertiary marginal price	€/MWh
mp_down	Downward tertiary marginal price	€/MWh
spread_up	Upward tertiary spread, mp_up − clearing_price	€/MWh
imbalance	System imbalance (negative = deficit)	MWh
act_up	Upward tertiary power activated (*ex post*)	MWh
act_down	Downward tertiary power activated (*ex post*)	MWh
headroom_{100}	Supply volume offered within +100 €/MWh above clearing	MW
slope_{2000}	Price rise to add 2000 MW above clearing	€/MWh
$V(p)$	Cumulative supply volume up to price p	MW
$z(\cdot)$	Standardisation (z-score), $(x-\mu)/\sigma$	—
q_{90}	90th percentile of the spread (spike threshold)	€/MWh
RISK	Ex-ante composite, Equation (5)	—

B Supporting figures

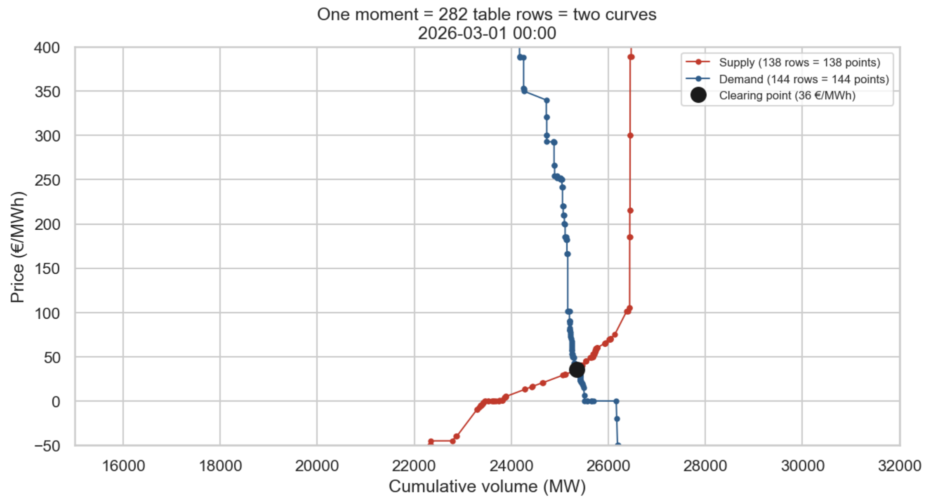

Figure 7: Anatomy of one moment: the ∼282 table rows of a single timestamp are the points of its supply and demand curves, which cross at the clearing point.

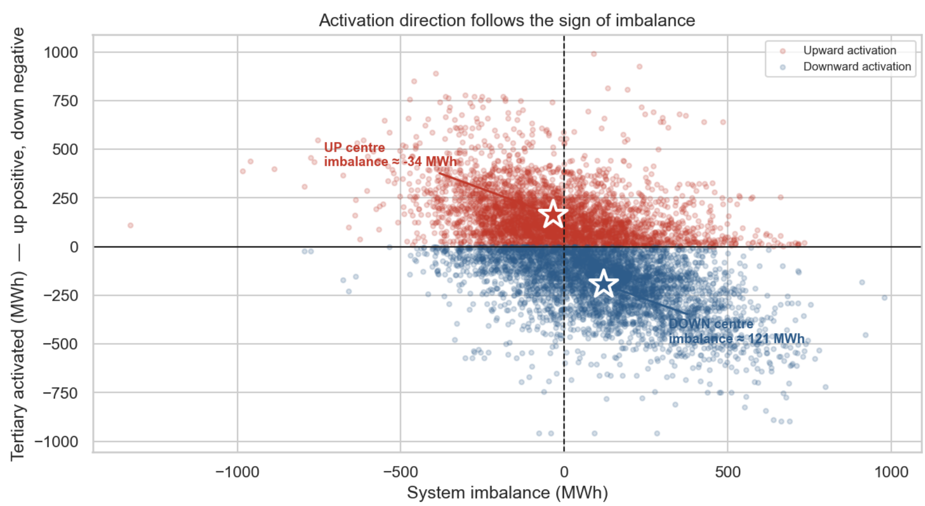

Figure 8: Activation direction versus the sign of system imbalance. Upward activations (red) cluster on the negative-imbalance side (deficit), downward (blue) on the positive side (surplus); the cloud centres are marked.

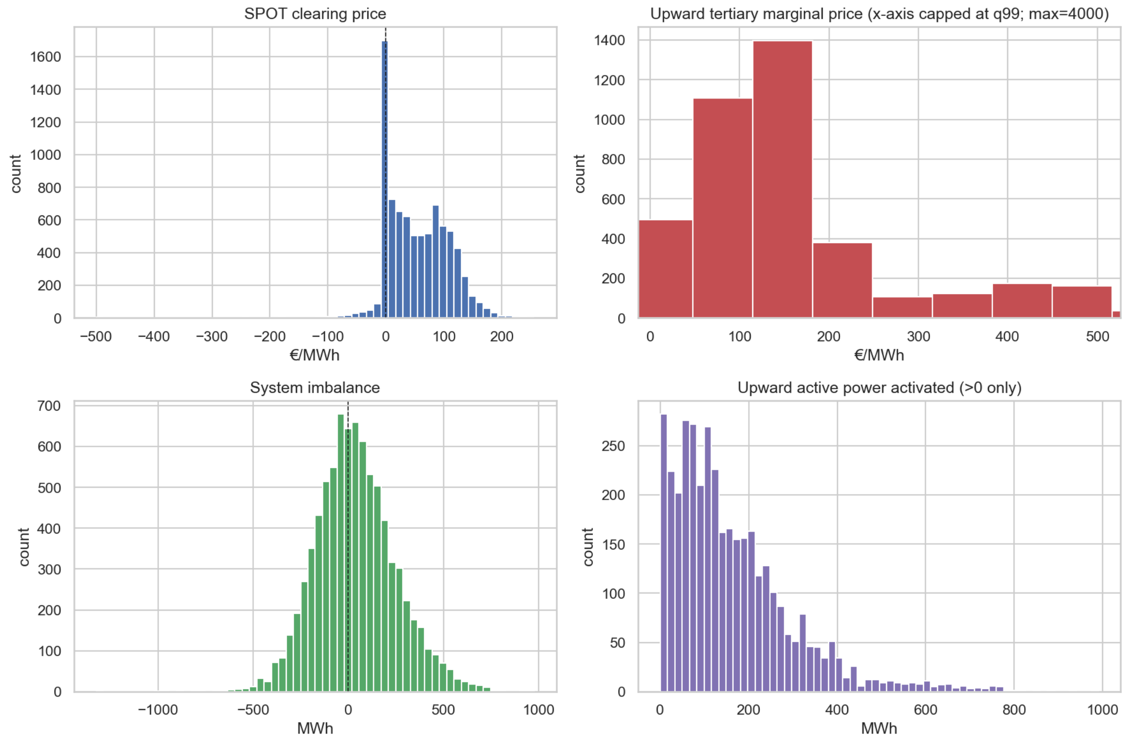

Figure 9: Distributions of the four core variables: SPOT clearing price, upward tertiary marginal price (x-axis capped at q99), system imbalance, and activated upward power (> 0 only).

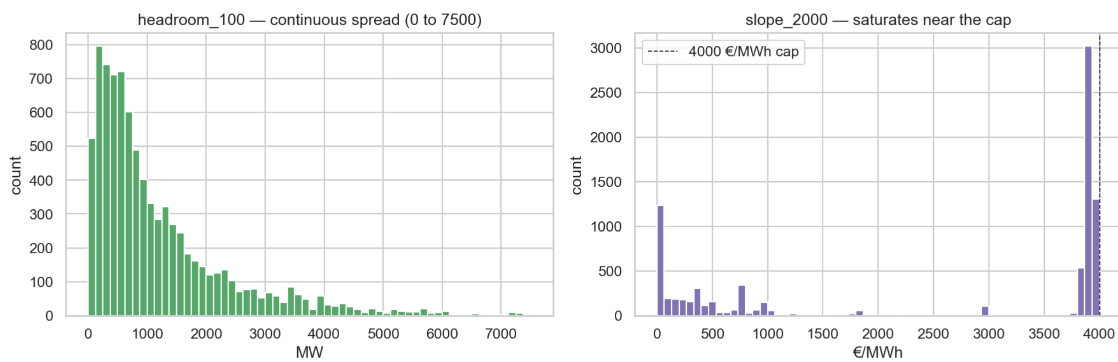

Figure 10: Distributions of the two curve features. headroom_{100} is continuously distributed; slope_{2000} saturates near the 4000 €/MWh cap, which is why it rank-orders spikes less reliably.

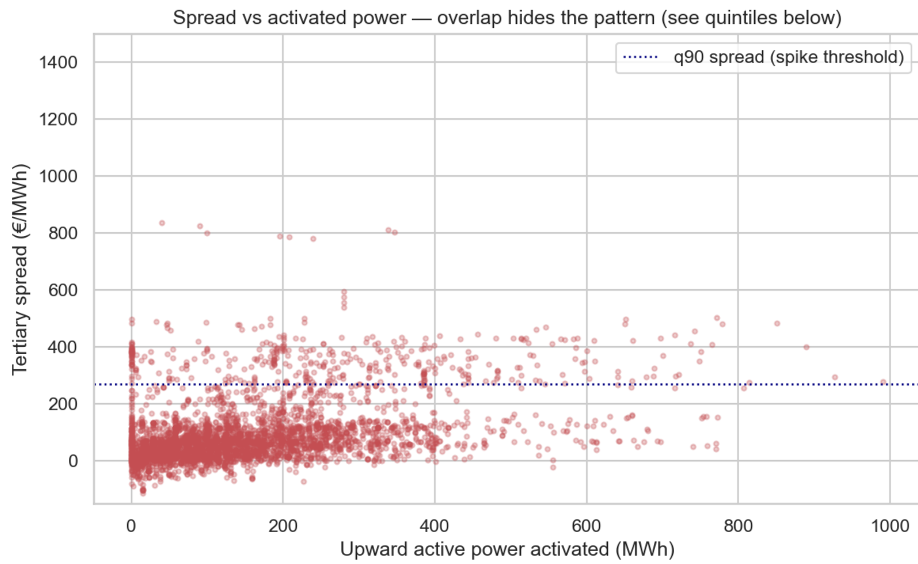

Figure 11: Spread versus activated power. Most points sit in a low-spread band across the whole range; the spike share rises with activation volume, but heavy overlap hides it — the quintile analysis quantifies it more reliably.

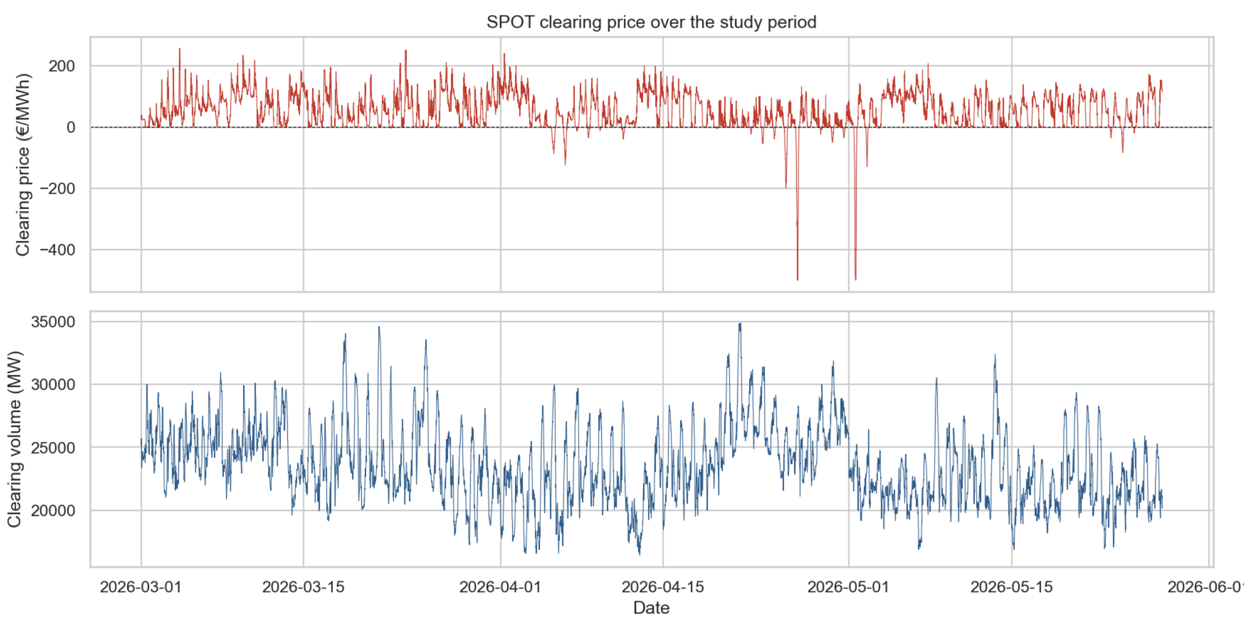

Figure 12: SPOT clearing price (top) and volume (bottom) over the study period. The price shows a daily cycle with occasional deep negative-price episodes; volume cycles daily with no trend.

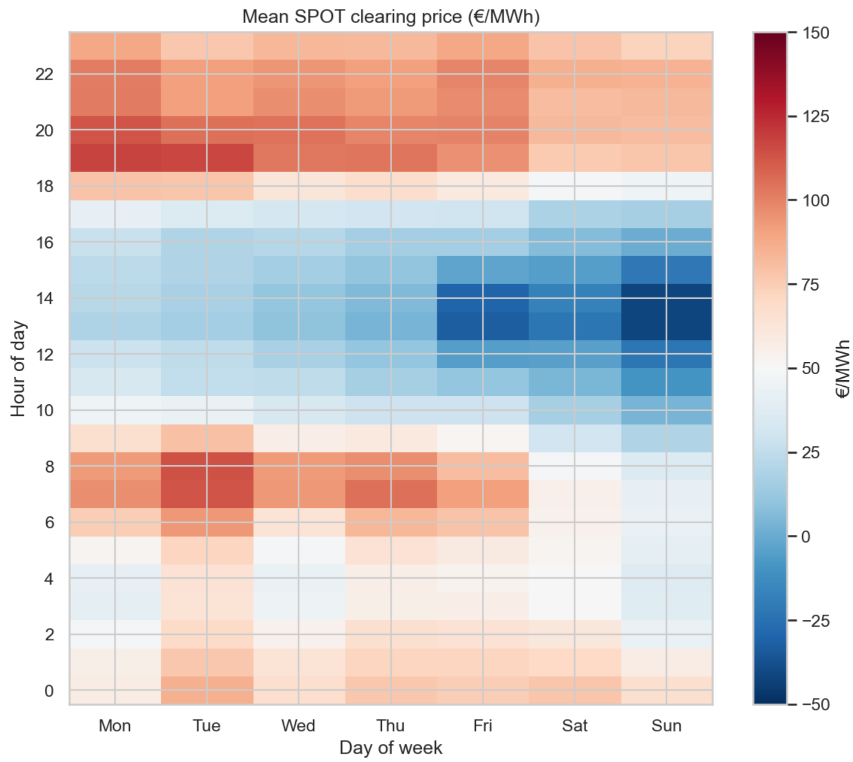

Figure 13: Mean SPOT clearing price by hour-of-day and day-of-week. The classic load shape appears: morning and evening peaks, midday and overnight dips, lower weekend prices.