Electricity Derivatives Summary

Gabriele Calandrino

February 2026

1 Introduction and Roadmap

This monograph situates *electricity derivatives* at the intersection of (i) market design, (ii) stochastic price modelling, and (iii) valuation of *real* (asset-based) derivatives. The starting observation is that the deregulation of electricity markets and the emergence of spot and futures quotations have generated a large quantitative-finance literature, yet electricity remains an outlier among tradable commodities.

1.1 Why electricity is a special underlying

The author highlights three structural reasons motivating dedicated models and methods:

1. **Non-storability and extreme spot behaviour.** Electricity cannot be economically stored, which fosters pronounced spikes; in addition, negative spot prices have been observed in European markets due to operational constraints of generation units combined with non-storability.

2. **Complex market microstructure.** System security requirements and limited flexibility of generation assets lead to multi-layered trading and balancing architectures; the market is engineered to allow trading while preserving physical feasibility and reliability.

3. **Path-dependent derivatives and optimal control.** Many relevant contracts correspond to sequences of exercise decisions reflecting operational constraints (e.g., start-up, ramping, inventory constraints). Their valuation naturally leads to stochastic control and related numerical methods rather than closed-form formulas.

1.2 Aim and scope of the monograph

The stated purpose is twofold:

- to provide a concise state-of-the-art overview of the *pricing and hedging* issues raised by electricity derivatives;

- to present modelling and methodological approaches that can be used to tackle them in practice.

The book is *not* intended as an exhaustive treatise; instead, it offers a broad, structured landscape of the electricity business through the market features, the price models, and the derivatives that are most specific to electricity (e.g., power plants, tolling, swings, storage).

A recurring motivation is that, despite two decades of progress, the field still lacks a broadly accepted *reference model* comparable to Black–Scholes for equity derivatives or HJM-type frameworks for interest rates; this gap contributes to heterogeneous valuations and highlights the need for further innovation.

1.3 What the reader will find in the next chapters

The monograph is structured around three successive layers:

$$\textbf{Markets} \rightarrow \textbf{Price models} \rightarrow \textbf{Derivatives}$$

Chapter 2: Electricity markets. The reader is introduced to the electricity commodity and the implications of (i) non-storability and (ii) locality (network constraints), with a brief engineering-oriented discussion aimed at explaining how these features shape the market microstructure.

Chapter 3: Price models. The focus shifts to stochastic models for spot and futures prices designed for *hedging and pricing* applications. The emphasis is on modelling the *futures dynamics* (rather than pure forecasting), and on assessing models through desiderata relevant to implementation (e.g., realism, internal consistency, computational efficiency, robustness, and generality).

Chapter 4: Electricity-specific derivatives. The book then discusses derivatives that are central to the electricity industry. Spread options are presented as fundamental building blocks, and more complex contracts (power plants/tolling, storage, swing) are framed as path-dependent options and control/switching problems, typically requiring numerical schemes. The chapter also touches on retail contracts (embedded optionality in customer load) and weather derivatives as risk-transfer instruments for climate-sensitive exposures.

2 Electricity Features

Electricity exhibits two technical characteristics that have first-order consequences for pricing, hedging, and market design: *(i) non-storability* and *(ii) network-constrained transport*. The author focuses on these features because they directly shape the behaviour of spot prices, the structure of forward contracts, and the nature of electricity-specific derivatives.

2.1 Storage

The statement "electricity cannot be stored" should be understood as "cannot be stored at reasonable cost"; yet, even with storage technologies, electricity systems face simultaneously an *energy* problem (how much energy can be stored) and a *capacity* problem (how fast it can be released, i.e. the power constraint).

Hydroelectric reservoirs as the benchmark. The most economical large-scale storage for electricity generation remains hydroelectric reservoirs, but this solution is geographically constrained and depends on hydrological inflows. The book illustrates the scale mismatch with a French example: hydro capacity is about 25 GW out of 110 GW installed capacity; hydro generation is about 15% of total generation; and the associated energy storage is about 10 TWh versus roughly 500 TWh annual consumption.

System security and the reserve stack. Because power cannot be stored, a persistent imbalance between demand and generation jeopardizes system frequency and can lead to blackouts. This implies (i) real-time monitoring of the balance on very short horizons and (ii) the need for *operating reserves* held by the transmission system operator (TSO) to cope with forecast errors and contingencies.

Reserves are organised by response time:

- **Primary and secondary reserves** (automatic, mobilised within less than 15 minutes),
- **Tertiary reserves** (manual activation beyond 15 minutes; in practice split into rapid and additional tiers),
- **Deferred reserves** (mobilised after more than half an hour).

Orders of magnitude reported for France (where hourly load ranges roughly from 50 GW in summer off-peak to 100 GW in winter peak) include minimum levels of approximately 700 MW (primary), 500 MW (secondary), and 1,500 MW (tertiary).

A useful operational concept is the TSO *operating margin* at horizon $t + h$, defined as the difference between the demand forecast for $t + h$ and the sum of the reserves available over that horizon. The margin requirement depends on the contingency set (e.g. loss of the largest unit on short horizons) and on a reliability criterion (e.g. probability of exceptional measures such as load shedding); the text provides an illustrative French value of about 2.3 GW to cover the next two hours during peak periods.

2.2 Transport

Electricity transport obeys Kirchhoff's laws: current balance at each node and zero net voltage drop along loops. A key implication is that, in a meshed network, power flows from an injection point to a withdrawal point through *all* available paths, not along a single contractual line.

Loop flows and congestion externalities. The book illustrates loop-flow effects using a three-node network where line capacities differ (e.g. one direct line with higher capacity and two indirect lines with lower capacities). Even if each market participant intends to use distinct lines "like trucks", Kirchhoff flows split the injections across paths and can produce unintended congestion on a line shared indirectly by several trades, potentially violating its thermal limit.

From network physics to traded transfer capacity. Because feasible cross-border exchanges depend on the whole network state and on generation patterns, transfer capacities available for trading require explicit modelling assumptions before they can be computed. In Europe, *net transfer capacities* (NTC) are centrally managed and published.

Allocation mechanisms and inefficiencies. To handle scarce interconnection capacity, early European markets relied on *explicit auctions*: participants trading power across borders had to separately purchase transmission rights at TSO-run auctions. While simple, this design can induce inefficiencies and even "adverse flows" (flows from a high-price area to a low-price area) due to timing and coordination frictions between the energy market and the capacity market.

2.3 My considerations 1

It is important to keep in mind that $P_{load} = P_{gen}$ no matter what, if the relation does not hold then we have a deviation in frequency (how fast the current oscillates) from 50 Hz, hurting the systems. If frec. $\neq 50$ we activate the reserves if the Operating Margin isn't enough to absorb the deviation.
That being said the *loading* term includes not only the demand but also the losses and the storage charging.

3 Market Microstructure

Electricity markets are organised in multiple sequential layers whose purpose is to (i) ensure physical feasibility of the system and (ii) allow economic price discovery. Because electricity cannot be stored economically and must be balanced in real time, the market structure is fundamentally different from that of storable commodities.

The trading architecture is typically decomposed into:

$$\text{Forward Market} \rightarrow \text{Day-Ahead Market} \rightarrow \text{Intraday Market} \rightarrow \text{Balancing Mechanism}$$

Each layer reduces uncertainty and corrects previous forecasts.

3.1 Day-Ahead Market

The day-ahead market is the central price formation mechanism.

Participants submit bids for each hour of the following day:

- Producers submit supply curves,
- Consumers (or retailers) submit demand curves.

A centralized auction determines the clearing price for each delivery hour. The result is:

- A scheduled generation plan,
- A scheduled consumption plan,
- A day-ahead spot price for each hour.

The clearing price corresponds to the marginal unit needed to satisfy demand. This is a uniform-price auction.

At this stage, the system is balanced in expectation, but uncertainty remains due to forecast errors in:

- Demand,
- Renewable generation,
- Plant availability.

3.2 Intraday Market

The intraday market allows participants to adjust positions between the day-ahead auction and real-time delivery.

As new information arrives (weather updates, outages, revised demand forecasts), market participants can:

- Buy additional energy,

- Sell excess contracted energy.

The intraday market reduces forecast errors but does not eliminate them. Residual imbalances are handled by the balancing mechanism.

3.3 Balancing Mechanism

The balancing mechanism is operated by the Transmission System Operator (TSO) and ensures real-time system security.

Its objectives are:

1. Maintain system frequency at its nominal value,
2. Correct deviations between scheduled and actual generation/consumption.

Producers owning flexible generation units must submit reserve bids, specifying:

- A quantity,
- A price,
- Activation constraints.

If the system is short (generation $<$ demand), the TSO activates upward reserves.

If the system is long (generation $>$ demand), the TSO activates downward reserves.

Participants whose actual position deviates from their scheduled position are settled at an imbalance price, which can differ significantly from the day-ahead price.

This introduces volume risk and operational risk into the market.

3.4 Forward Market

Forward and futures contracts allow participants to hedge price risk for delivery periods extending from weeks to years.

Unlike storable commodities, electricity forwards are not linked to the spot price via a cost-of-carry relation. Instead, forward prices reflect:

- Expectations of future supply-demand conditions,
- Risk premia,
- Fuel prices,
- Regulatory and structural constraints.

As delivery approaches, forward contracts are progressively replaced by contracts with shorter delivery horizons, a mechanism often referred to as cascading.

The convergence property

$$\lim_{t \to T} F(t, T) = S_T$$

is approximately observed but must be interpreted carefully, since many contracts settle on averages over delivery periods.

3.5 Diversity of Electricity Markets

Electricity markets differ across countries due to:

- Generation mix (nuclear, hydro, renewables, thermal),
- Interconnection capacity,
- Market design rules,
- Regulatory frameworks.

Consequently, price dynamics are not universal. Models must reflect the physical and institutional characteristics of the specific market under consideration.

4 Real Derivatives

Like any other commodity market, electricity has standard options written on quoted futures. However, the most challenging problems for practitioners are not the valuation of such standard products, but rather their *pricing, hedging, and structuring* into exotic tradable products. These products are called *real derivatives.* The key point is that electricity operators (generators, utilities) and retailers typically hold *embedded optionalities* in their physical portfolios. For instance, a power plant can be interpreted as a strip of call options, while a retail contract with a curtailment clause (penalising the customer when a curtailment signal is ignored) contains an embedded put-type optionality. By analogy with real options in investment decisions, these rights can be viewed as real derivatives.

Moreover, since electricity cannot be stored, the Transmission System Operator (TSO) must procure *flexibility* to cope both with uncertainty and with the dynamic constraints of the generation system. In some situations it is more economical to reduce (or defer) consumption than to satisfy it immediately; in others it may be preferable to *increase* consumption to avoid shutting down an inflexible unit. A canonical example is the remote control of household water heaters in France: they can be automatically started at night when system consumption is lowest. These flexibilities can be seen as *real options* bought by the TSO from users, or by market participants from customers.

The following three real derivatives are the daily concerns of operators, together with their financial representations: (i) power plants and tolling agreements, (ii) energy storage and swing contracts, (iii) retail contracts.

4.1 Power plant and tolling agreement

The owner of a power plant creates value by *selling electricity* and *buying fuel.* Estimating the value of the generation asset is required for investment decisions, long-term contract negotiations, and risk management. A power plant is subject to many technical, environmental, and legal constraints; nevertheless, a first crude approximation represents a thermal plant as a strip of calls on a spread. When carbon emission prices are included, one obtains *clean* spreads.

Let S_t denote the electricity spot price (power) at time t, S_t^f the spot price of the plant's fuel (e.g. gas or coal), and h the plant's *heat rate.* The heat rate h converts one unit of electricity output into required thermal fuel input (so that hS_t^f is the instantaneous fuel cost per unit of electricity produced). Let S_t^c denote the spot price of an emission permit (carbon), and let g be the plant's *emission factor* (so that gS_t^c is the instantaneous carbon cost per unit of electricity produced). Then the payoff *per MW* over a time period $[0, T]$ is approximated by

$$\int_0^T \left(S_t - hS_t^f - gS_t^c\right)^+ dt. \tag{1}$$

Expression (1) makes explicit why a plant can be identified with a strip of call options. At the same time, the instantaneous payoff depends on *three* underlying prices (power, fuel, carbon), which already raises technical difficulties.

More importantly, operational constraints drastically reduce the possibility of capturing all successive positive spreads. Start-up costs, minimum running times, ramp-up and ramp-down constraints, and limited numbers of cycles per day all limit flexibility; hence (1) typically *overestimates* the achievable profit. Incorporating such constraints naturally leads to an optimal control formulation. One possible representation is

$$\sup_{q_t \in \mathcal{A}} E\left[\int_0^T q_t \left(S_t - hS_t^f - gS_t^c - \kappa\right)^+ dt\right], \tag{2}$$

where q_t is the generation rate (power output) at time t, constrained to belong to an admissible set $\mathcal{A}$ encoding the operational constraints, and κ is a start-up cost (included here as a simplified cost term inside the spread). Problems of the form (2) are closely related to classical *unit commitment* and *mid-term generation management*: scheduling groups of plants over hours/days, or assessing generation levels over months. Numerical optimization approaches (e.g. mixed-integer programming or Lagrangian relaxation) were developed long before deregulation, in settings where monopolies aimed to meet random demand at minimal cost. Spot markets modify the context by adding a bidding phase, but the underlying optimization techniques remain relevant.

A major novelty introduced by financial markets is the hedging question for (2). Natural hedging instruments include futures on fuels, carbon emissions, and power. Yet, power futures do not match the fine granularity of spot exposure: if T is one year, the payoff is exposed to 8,760 hourly risk factors, while only a handful of futures contracts are liquidly available. In this sense, the electricity market is incomplete; consequently, even the notion of "the" value of (1) is ambiguous and lacks a consensual answer. This directly explains why exchange prices for tolling contracts are difficult to establish.

A *tolling contract* is the financial counterpart of a power plant: the owner concedes an exploitation right in exchange for a fixed premium. Depending on the price model, the operational constraints accounted for, and the hedging capacities, large discrepancies between quoted prices may arise.

4.2 Energy storage and swing contract

Hydroelectric plants are highly flexible and can provide electricity on short notice. While valuation issues similar to thermal plants arise, hydro introduces a fundamental additional ingredient: the "fuel" (water) is *limited*, so one faces a storage management problem.

The simplest hydro storage management problem considers a single reservoir. The valuation problem is written as

$$\sup_{q_t \in [0,\overline{q}],\, \delta_t \geq 0} E\left[\int_0^T q_t\, S_t\, dt + g(S_T, X_T)\right], \tag{3}$$

where S_t is the electricity spot price and X_t is the reservoir level at time t. The dynamics of the reservoir level can be represented (with explicit dependence on initial conditions (t,x)) by

$$dX_s^{t,x} = \left(a_s^{t,a} - q_s - \delta_s\right) ds, \tag{4}$$

where $(a_s^{t,a})_{s \geq t}$ denotes random inflows (with initial level a at time t). The reservoir is subject to level constraints $X_s \in [\underline{x}, \overline{x}]$. The control δ_s is a *spilling* variable: when the reservoir is full ($X_s = \overline{x}$) and inflows exceed the generation capacity $\overline{q}$, the operator must spill the excess water. The terminal function g represents a final value (reward) attached to having a certain water level at the terminal time T.

In more general settings, a network of reservoirs and hydro plants interact, constraints become more complex, and the state dimension increases dramatically. Such hydro management problems were historically a driving force behind the development of dynamic programming. The main difficulty is the curse of dimensionality as the number of reservoirs grows; classical approaches include stochastic dynamic programming methods, such as dual dynamic programming and decomposition techniques. The advent of electricity markets revived these questions due to the availability of spot prices, and alternative approaches based on optimal switching or stochastic control have also been proposed.

As for power plants and tolling, a financial representation of storage facilities was developed: *swing options*. The aim is to provide the ability to store power without the full technical constraints of a real hydro system. An important class of swing contracts arises from demand-side management: customers benefit from a lower tariff most of the year, except for a few (unknown in advance) days where the price becomes dissuasive; the utility announces such days one day ahead. Hence the utility holds an option on the customer's consumption. For an industrial customer, the response is essentially deterministic (the firm cancels consumption), whereas for a large pool of households, the resulting avoided consumption is random.

4.3 Retail contract

A retailer's pricing policy consists in choosing the form of contracts and the prices to charge. In analogy with insurance, retailers first design contract menus and prices; customers then select a contract and consume. When setting this policy, a retailer does not know its future sourcing cost, nor the exact future consumption of its customers.

Consumption differs substantially across customer types, and customers are typically grouped into three classes: industrial, professional, and household. These classes differ in number, volume and patterns of consumption, economic behaviour, needs, and the information available to the retailer. Industrial customers are fewer but represent a significant share of aggregate consumption. Their loads are often precisely metered, they may be less weather-sensitive than households, and they can switch retailer at very low per-MWh savings because even small price differences amount to large yearly sums. Household switching is typically more viscous, and household consumption is often still measured infrequently in many countries.

These differences imply very different contract structures. Industrial contracts can be tailored to operational constraints, while household contracts are typically standardised mass-market products. Industrial retail contracts often include embedded options (e.g. rights to resell delivered power to the market at a specified price). Household contracts commonly involve fixed premiums proportional to subscribed capacity plus a price per kWh, sometimes with two-level tariffs depending on the hour of consumption (peak/off-peak).

This topic has received less attention in quantitative energy finance than in economics. Historically, pricing policy during monopoly periods was framed through marginal cost pricing theory (notably illustrated by Boiteux), which aims at maximising social welfare under regulated monopoly. Under competition and with actual market prices, selling at marginal generation cost is not guaranteed to be value-maximising for the firm; moreover, the marginal cost pricing rule relies on long-term equilibrium conditions on generation assets that are unlikely to hold in real market conditions.

5 Models

5.1 Preliminary Remarks: spot–forward relations in electricity

This section recalls the classical no-arbitrage links between spot and forward prices for *storable* assets, and explains why these links largely fail in electricity markets. It then isolates two relations that remain central for modelling purposes: a (near) convergence property and a representation of forwards as expectations under a pricing measure, together with the practical caveats induced by market incompleteness.

Spot and forward prices. Let S_t denote the spot price at time t (in electricity, the relevant reference for delivery is typically the *day-ahead* price), and let $F(t,T)$ denote the forward (or futures) price observed at time t for delivery at future time $T > t$. The interest rate is denoted by r and is taken constant in the benchmark relations below. In this chapter the terms *forward* and *future* are used interchangeably, since their difference is not essential for the modelling objectives pursued here.

Benchmark: freely storable assets. If the underlying can be stored without frictions, standard no-arbitrage arguments imply

$$F(t,T) = e^{r(T-t)} S_t.$$

For storable commodities, storage costs matter. A first correction replaces the above by

$$F(t,T) = e^{r(T-t)+c(T-t)} S_t,$$

where $c(T-t)$ is a (time-to-maturity dependent) storage cost. Empirically, even this relation is insufficient because forward prices may lie below spot prices; a classical explanation introduces a *convenience yield* $y(t,T)$ (the benefit of physically holding inventory), leading to

$$F(t,T) = e^{r(T-t)+c(T-t)-y(t,T)} S_t.$$

In electricity, such inventory-based reasoning essentially breaks down: storage is not available at scale and the convenience-yield concept becomes hard to justify in its usual form.

Convergence and basis risk. A second modelling hypothesis is that, as delivery approaches, the forward should converge to the spot:

$$\lim_{t \to T} F(t,T) = S_T.$$

If this relation holds, the basis risk between forward and spot becomes negligible near delivery, and an agent can in principle unwind a hedge by switching between the two markets without incurring a systematic mismatch. In electricity, however, the situation is more delicate because standard contracts are written on *delivery periods* (month, quarter, year) and settle against *average* day-ahead prices over the delivery period. This aggregation induces a cascading hedging practice: as shorter delivery-period contracts become available, operators roll their hedges from longer to shorter maturities.

A concrete test case is provided by *day-ahead futures* whose underlying is the average spot price over the delivery day. These contracts are traded only on business days: for a Monday delivery, the last quote occurs on the preceding Friday. The discrepancy between the last futures quote and the realised average spot over Monday therefore measures the quality of convergence (and, in practice, the residual weekend basis). Empirical evidence reported in the section suggests that this discrepancy is not statistically different from zero on average across hours, although it can reach a few percent for specific peak hours. Hence, for modelling purposes, convergence can be accepted as an approximation, possibly up to a small premium reflecting residual basis effects.

Risk-neutral representation and incompleteness. A third classical cornerstone for storable assets is the existence of a risk-neutral (pricing) measure Q such that

$$F(t,T) = E_t^Q[S_T]\,.$$

For valuation, it is desirable that Q be unique. Electricity models, however, typically lead to *incomplete* markets: the set of traded instruments is too small to span all relevant risks (in particular, the hourly granularity of spot exposure versus the coarse set of liquid forwards). As a consequence, the risk-neutral measure is generally not unique, and one must select a pricing measure according to modelling choices. The section emphasises that, beyond theoretical existence and uniqueness, a key practical criterion is the model's ability to fit the observed forward curve.

Risk premium. A common object of study in the electricity literature is the forward risk premium,

$$R(t,T) = F(t,T) - E_t^P[S_T]\,,$$

where P denotes the historical (physical) probability measure. The sign and term-structure of $R(t,T)$ are economically important, but a detailed discussion is not required for the core objective of this text, namely derivative pricing based on models delivering consistent forward dynamics.

5.2 HJM-Style Forward Curve Models

This section presents the class of models inspired by the Heath–Jarrow–Morton (HJM) methodology originally developed for interest-rate term structures. The guiding idea is to start from the *observed* forward curve, model its dynamics directly, and (only if needed) define the spot as a zero-maturity limit. This "forward-first" viewpoint is attractive in practice because it enforces an exact fit to market quotes at each calibration date, which is a primary requirement on trading desks.

5.2.1 Principle

Let $F(t,T)$ denote the forward (or futures) price observed at time t for delivery at the future instant $T > t$. Throughout this section we assume the existence of a pricing measure Q under which traded forward prices are (local) martingales. The HJM specification is then given by prescribing volatility functions $(\sigma_i(t,T))_{1\leq i\leq n}$ and writing

$$\frac{dF(t,T)}{F(t,T)} = \sum_{i=1}^{n} \sigma_i(t,T)\, dW_t^i, \tag{5}$$

where $(W^1,\ldots,W^n)$ is an n-dimensional Brownian motion under Q (components possibly correlated). The family $\sigma_i(t,T)$ controls the evolution of the entire term structure $T \mapsto F(t,T)$.

The model is initialized by setting $F(0,T)$ equal to the observed forward curve at time 0. Hence, calibration at the initial date is tautological: the model reproduces the market curve by construction, and only the subsequent dynamics are modelled.

If one wishes to define a spot price process, a natural definition is the immediate-delivery limit

$$S_t := F(t,t), \tag{6}$$

so that the spot inherits its dynamics from (5). This definition is standard in HJM-type term-structure models: the spot is not modelled independently but extracted from the forward curve.

A practical issue concerns the number of factors n. In principle, one could associate one factor per maturity, yielding a very high-dimensional system. Empirically, however, principal component analysis (PCA) in many markets shows that a small number of factors captures most of the variance of forward returns, motivating low-dimensional specifications of (5).

5.2.2 The Case of Electricity

Applying an HJM methodology to electricity forward curves raises additional issues that are far less pronounced in interest rates or in storable commodities.

Seasonality and the need for a deterministic trend. A first technical difficulty is the strong seasonal pattern of electricity forward prices. This seasonality reflects well-understood seasonal variations of demand and of the marginal cost of production, hence it is natural to introduce a time-dependent deterministic component (a "trend") that captures the seasonal structure of the curve. Several estimation procedures have been proposed. One approach is to rely on a fundamental forecasting model (built for spot prices) to produce a seasonal profile, and then fit this profile to the observed forward curve using constrained least squares. A key modelling requirement is to impose smoothness with respect to maturity in order to avoid artificial discontinuities between neighbouring contracts (e.g. between the last daily contract of summer and the first daily contract of autumn). A simpler alternative is to fit seasonality using a truncated Fourier series; in practice, a forecast-based seasonal trend can provide a slight improvement.

Number of factors and distributional shape. A second (and more important) difficulty concerns the factor dimension needed to represent the dynamics of electricity forwards. Empirical PCA studies (notably on Nord Pool data) indicate that electricity forward curves exhibit a much more volatile and less "compressible" behaviour than typical commodity or bond markets. For instance, one study reports that, while three factors may explain essentially all the variance in copper, crude oil, or bonds, the same number of factors explains only a limited portion of the variance in electricity forwards. Another study finds that one factor explains about 68% of the variance, two factors about 75%, three factors about 80%, and four factors about 83%, and that more than seven factors are required to exceed 90%. A key economic interpretation is that short-term products (especially weekly contracts) may move quite independently from seasonal or yearly contracts. Moreover, forward returns may deviate from Gaussianity even for relatively long maturities; this motivates extensions beyond Brownian-driven HJM, including Lévy-driven infinite-dimensional models (e.g. ambit-field approaches).

Delivery periods and the unobserved hourly forward curve. Further difficulties arise from a structural feature of electricity: the spot price is fundamentally *hourly*, while exchange-traded forward products typically involve *delivery periods*. In the HJM notation $F(t,T)$, it must therefore be clarified whether T denotes a quoted delivery period (month, quarter, year) or an *hour* in the future.

To value real derivatives (plants, storage, swing), one would like access to an *instantaneous-delivery* (or "hourly") forward curve. Denote by $f(t,T)$ the (generally unobserved) price at time t for delivery of 1 MW during the hour T. A traded forward contract delivering 1 MW continuously over the period $[T_1,T_2]$ has an observed price $F(t,T_1,T_2)$, which is naturally linked to f by an averaging identity:

$$F(t,T_1,T_2) \;=\; \frac{1}{T_2-T_1}\int_{T_1}^{T_2} f(t,s)\,ds. \tag{7}$$

Even at the initial date $t=0$, reconstructing $T\mapsto f(0,T)$ from a finite set of period quotes $\{F(0,T_1,T_2)\}$ is an ill-posed inverse problem: in general, there is no unique hourly curve consistent with the same set of averages. For modelling purposes, one assumes that a reasonable construction of $f(0,\cdot)$ is available and uses it as the initial condition of the HJM dynamics for $f(t,T)$.

Admissible volatility functions and the Markov constraint. If one models the unobserved $f(t,T)$ in HJM form and then defines the spot by the zero-maturity limit $S_t := f(t,t)$, one must confront a strong restriction: not all volatility specifications are compatible with a Markov spot. In particular, when the delivery-period relation (7) is imposed and one requires the Markov property for the derived spot, it can be shown (under mild regularity conditions) that the volatility of $f(t,T)$ must be *independent of the maturity* T. Hence $\sigma(t,T)$ must reduce to a function of t only. This leaves few practical alternatives, and it explains why Black-type specifications (constant volatility) are frequently used in practice for such forward-curve dynamics.

Modelling period contracts directly and a two-factor specification used in practice. An alternative is to specify dynamics directly for observed delivery-period contracts:

$$dF(t,T_1,T_2) \;=\; \Sigma(t,T_1,T_2)\,F(t,T_1,T_2)\,dW_t,$$

for any pair (T_1,T_2), and to infer (approximately) an instantaneous curve and a spot price from these. A concrete example developed in the literature considers monthly futures $F(t,T)$ (where T denotes the start of the delivery month) and treats longer products as portfolios of monthly contracts. A commonly used two-factor model under the pricing measure is:

$$dF(t,T) \;=\; e^{-\kappa(T-t)}\sigma_1\,F(t,T)\,dW_t^{(1)} \;+\; \sigma_2\,F(t,T)\,dW_t^{(2)}, \tag{8}$$

where $\kappa>0$ is a mean-reversion speed shaping how the first factor loads across maturities, $\sigma_1,\sigma_2>0$ are volatilities, and $W^{(1)},W^{(2)}$ are (often taken) independent Brownian motions. For this model, one obtains an explicit expression for the variance of $\ln F(t,T)$, which enables calibration of $(\kappa,\sigma_1,\sigma_2)$ from at-the-money option-implied variances across maturities. The benefit of (8) is simplicity and daily calibration feasibility; the limitation is that it may not reproduce all features required for complex derivative valuation.

Remark: valuation beyond no-arbitrage HJM. Because of the restrictive admissible volatility classes when insisting on a Markov spot together with delivery-period constraints, an alternative line of work formulates the forward dynamics under the historical probability and performs valuation via a quadratic hedging criterion, rather than enforcing a risk-neutral SDE for forward prices.

Joint modelling with fuels. Finally, for many real derivatives (e.g. thermal plants and tolling) the relevant state is not electricity alone but the joint dynamics of electricity forwards and fuel (and possibly carbon) prices. This joint modelling aspect is essential for spread-based payoffs, yet it is less frequently addressed than the single-commodity forward-curve modelling problem.

5.3 One-Factor Spot Models

The goal of *one-factor spot models* is to define an electricity spot-price dynamics that (i) preserves a Markov structure for the spot and (ii) allows one to compute the whole forward curve as an expectation of the spot under a suitable pricing measure. In the text, four representative specifications are discussed, with increasing complexity (Table 1).

Model	Seasonality	Process	Jumps	Change of measure
Lucia–Schwartz	Deterministic	Gaussian MR	–	Constant
Cartea–Figueroa	Deterministic	Gaussian MR	Poisson	Constant
Benth et al.	Deterministic	Gaussian MR	Lévy	Time-dependent
Benth–Saltyte-Benth	Deterministic	Non-Gaussian MR	–	(Esscher)

Table 1: Summary of one-factor spot models discussed in the book.

5.3.1 Mean-Reversion Process (Lucia–Schwartz benchmark)

The benchmark model uses the *daily average* spot price, denoted by P_t (not the hourly spot). It is specified on the log-price by a deterministic seasonal component plus a Gaussian mean-reverting factor:

$$\ln P_t = f(t) + Y_t, \tag{3.21}$$

$$dY_t = -\kappa Y_t \, dt + \sigma \, dW_t, \tag{3.22}$$

where

- $f(t)$ is a known deterministic seasonal function;
- Y_t is an Ornstein–Uhlenbeck (OU) process with mean-reversion speed $\kappa > 0$ and volatility $\sigma > 0$;
- W_t is a standard Brownian motion (under the historical measure).

The distribution of P_t is lognormal, and its first two moments are explicit:

$$E_0[P_t] = \exp\left(f(t) + (\ln P_0 - f(0))e^{-\kappa t} + \frac{\sigma^2}{4\kappa}\left(1 - e^{-2\kappa t}\right) \right), \tag{9}$$

$$\mathrm{Var}_0[P_t] = E_0[P_t]^2 \left[\exp\left(\frac{\sigma^2}{2\kappa}\left(1 - e^{-2\kappa t}\right) \right) - 1 \right]. \tag{10}$$

For valuation, the book assumes the existence of a (chosen) pricing measure under which futures equal conditional expectations of the spot. A constant market price of risk λ is introduced, and the mean-reversion structure is preserved under the pricing measure. Under this assumption, the forward price has the closed form:

$$\ln F(t,T) = f(T) + e^{-\kappa(T-t)}\big(\ln P_t - f(t)\big) + \alpha^* \big(1 - e^{-\kappa(T-t)}\big) + \frac{\sigma^2}{4\kappa}\big(1 - e^{-2\kappa(T-t)}\big), \tag{3.23}$$

where α^* is the (constant) drift adjustment induced by the constant market price of risk (in the book: $\alpha^* = -\lambda\sigma/\kappa$).

Calibration in the benchmark. A typical specification for the deterministic component is sinusoidal with an additional dummy for holidays/weekends:

$$f(t) = a + bD_t + c\cos\big((t+\tau)2\pi/365\big),$$

where $D_t \in \{0,1\}$ indicates whether day t is a holiday/weekend, and (a,b,c,τ) are parameters. In the benchmark study reported in the book, parameters $(\kappa,\sigma,a,b,c,\tau)$ are estimated jointly via a nonlinear least-squares procedure on several years of NordPool daily data. The key qualitative outcome is that the model can fit the *smooth* daily dynamics reasonably well, but it produces no spikes and, with a constant market price of risk, it cannot perfectly match the observed forward curve; moreover, strong mean reversion dampens the dynamics of longer maturities.

5.3.2 Mean-Reverting Jump–Diffusion Models (Cartea–Figueroa)

A first improvement is to add jumps to generate spikes. The spot is specified by a deterministic log-seasonality $g(t)$ and a stochastic factor Y_t:

$$\ln S_t = g(t) + Y_t, \tag{11}$$

where g is deterministic and differentiable. The stochastic component follows a mean-reverting jump–diffusion:

$$dY_t = -\alpha Y_t\,dt + \sigma(t)\,dW_t + (\ln J)\,dq_t. \tag{12}$$

The model ingredients are:

- $\alpha > 0$: mean-reversion speed (fast mean reversion makes spikes short-lived);
- $\sigma(t) \geq 0$: (possibly time-dependent) diffusion volatility;
- W_t: Brownian motion;
- q_t: Poisson process with intensity $l > 0$ (jump arrival rate);
- J: proportional jump multiplier, assumed lognormal: $\ln J \sim \mathcal{N}(\mu_J, \sigma_J^2)$.

A modelling restriction used in the book is $E[J] = 1$, so that jumps do not introduce an additional drift on average; for lognormal J, this implies

$$\mu_J = -\frac{\sigma_J^2}{2}.$$

Under this specification, the spot-price SDE can be written (with S_{t-} the left limit at time t):

$$dS_t = \alpha\big(\rho(t) - \ln S_t\big)S_t\,dt + \sigma(t)S_t\,dW_t + S_{t-}(J-1)\,dq_t, \tag{13}$$

where the deterministic function $\rho(t)$ is

$$\rho(t) = \frac{1}{\alpha} g'(t) + \frac{1}{2}\sigma^2(t) + g(t).$$

Pricing with a constant market price of risk. As in the benchmark, a constant market price of risk λ is used. Under the pricing measure, if $x_t = \ln S_t$, the dynamics takes the form

$$dx_t = \alpha\big(\mu^*(t) - x_t\big)\, dt + \sigma(t)\, dW_t^* + (\ln J)\, dq_t, \qquad \mu^*(t) = \frac{1}{\alpha} g'(t) + g(t) - \frac{\lambda}{\alpha}\sigma(t).$$

In this setting, the book highlights that futures prices admit a closed-form decomposition:

$$F(t,T) = G(T) \left(\frac{S_t}{G(t)}\right)^{e^{-\alpha(T-t)}} D(t,T)\, J(t,T), \tag{14}$$

where $G(t) = \exp(g(t))$ is the seasonal component, and the remaining multiplicative terms separate diffusion and jumps:

$$D(t,T) = \exp\left(\int_t^T \left(\frac{1}{2}\sigma^2(s) e^{-2\alpha(T-s)} - \lambda\sigma(s) e^{-\alpha(T-s)}\right) ds\right), \tag{15}$$

$$J(t,T) = \exp\left(l \int_t^T (\xi(T,s) - 1)\, ds\right), \qquad \xi(T,s) = \exp\left(-\frac{\sigma_J^2}{2}\left(e^{-\alpha(T-s)} - e^{-2\alpha(T-s)}\right)\right). \tag{16}$$

Interpretation and limitations. The constant λ makes it difficult to perfectly fit the initial forward curve. The jump intensity l and jump-size parameter σ_J reduce the forward price through the jump term. Finally, when α is large (fast mean reversion), jumps have a strongly dampened effect on long maturities, which tends to generate an unrealistically flat long-maturity forward dynamics.

5.3.3 Non-Gaussian Mean-Reversion Models (Benth–Saltyte-Benth; NIG example)

An alternative to adding explicit Poisson jumps is to represent the stochastic driver by a *non-Gaussian* Lévy process. The spot is modelled as

$$S_t = \Lambda_t e^{X_t}, \tag{3.24}$$

where Λ_t is a deterministic seasonal function and X_t follows an OU-type dynamics driven by a Lévy process L_t:

$$dX_t = a(m - X_t)\, dt + dL_t. \tag{3.25}$$

Here $a > 0$ is the mean-reversion speed, m is the long-run mean, and $(L_t)_{t\geq 0}$ has stationary independent increments. The book emphasises the Normal Inverse Gaussian (NIG) family as a convenient choice (explicit density with four parameters: location μ, tail heaviness α, skewness β, and scale δ), enabling efficient maximum likelihood estimation.

Forward prices via an Esscher transform. In an incomplete market there are infinitely many equivalent martingale measures. The book presents a change of measure (Esscher transform) parameterised by a deterministic function $\theta(\cdot)$ that can be used to fit the initial forward curve. In the NIG-driven OU setting, forward prices are given by:

$$F^\theta(t,T) = \Lambda(T)\exp\Big(\frac{\mu}{a}\big(1-e^{-a(T-t)}\big)\Big)\left(\frac{S_t}{\Lambda(t)}\right)^{e^{-a(T-t)}} \tag{17}$$

$$\times\exp\left(\delta\int_t^T\Big[\sqrt{\alpha^2-(\theta(s)+\beta)^2}-\sqrt{\alpha^2-(\theta(s)+\beta+e^{-a(T-s)})^2}\Big]\,ds\right). \tag{18}$$

Fitting the initial forward curve with piecewise-constant θ. Assume that at time 0 one observes n forward prices $(f_i)_{i=1}^n$ for non-overlapping maturities $0 < T_1 < \cdots < T_n$. Choose $\theta(t) = \theta_i$ on each interval (T_{i-1}, T_i). Then the book writes the system

$$f_i = q_i\,\Theta_i, \tag{3.27}$$

with

$$q_i = \Lambda(T_i)\exp\Big(\frac{\mu}{a}\big(1-e^{-aT_i}\big)\Big)\left(\frac{S_0}{\Lambda(0)}\right)^{e^{-aT_i}}, \tag{19}$$

$$\Theta_i = \exp\left(\delta\int_0^{T_i} h(s,T_i,\theta(s))\,ds\right), \tag{20}$$

$$h(s,T_i,\theta(s)) = \sqrt{\alpha^2-(\theta(s)+\beta)^2}-\sqrt{\alpha^2-(\theta(s)+\beta+e^{-a(T_i-s)})^2}. \tag{21}$$

The book explains that one can solve recursively for $\theta_1, \theta_2, \ldots$ by progressively using the maturities $T_1, T_2, \ldots$ (the recursive construction is discussed in the broader Lévy setting in the references cited there).

Main qualitative limitation. Even with a richer (non-Gaussian) marginal distribution and an admissible change of measure, strong mean reversion still dampens the forward dynamics at long maturities: short maturities remain volatile, while long maturities rapidly collapse toward the seasonal component.

5.3.4 Conclusion of the One-Factor Approach

One-factor spot models can reproduce key stylised facts of electricity spot prices (seasonality, mean reversion, spikes). However, the book stresses that they typically fail to reproduce a realistic *term-structure dynamics* for long maturities, precisely because mean reversion erases the influence of spot shocks over time. This motivates moving to multi-factor spot models, where distinct factors can separately control spikes and longer-term forward-curve movements.

5.4 Multi-factor Spot Models

Motivation. Single-factor spot models typically imply that the forward curve becomes *too stable* for medium/long maturities: the impact of spot shocks decays quickly with $T-t$, hence forward volatilities collapse as maturity increases. Multi-factor specifications introduce at least two mean-reversion speeds so that (i) short-term spikes dissipate fast, while (ii) a slower component keeps the whole curve genuinely stochastic at longer horizons.

5.4.1 An illustrative two-factor model: diffusion + mean-reverting spikes

Specification. Let Λ_t be a deterministic seasonality function and model the spot multiplicatively as

$$S_t = \Lambda_t \exp(X_t + Y_t). \tag{3.28}$$

A typical choice is:

- a *slow* Gaussian factor X_t (Ornstein–Uhlenbeck),

$$dX_t = \lambda_X(\mu_X - X_t)\,dt + \sigma_X\,dW_t;$$

- a *fast* spike factor Y_t, mean-reverting with upward jumps,

$$dY_t = -\lambda_Y Y_t\,dt + h\,dN_t,$$

  where N_t is Poisson with intensity l and the jump size h is lognormal, $\ln h \sim \mathcal{N}(\mu_J, \sigma_J)$.

Implication for forward dynamics. With $F(t,T) = E[S_T \mid \mathcal{F}_t]$, the two distinct speeds $\lambda_Y \gg \lambda_X$ generate a *two-scale* term structure:

- near maturities: volatility dominated by Y (spikes), but with quick decay because of strong mean reversion;
- far maturities: X remains influential, preventing the forward curve from becoming almost deterministic.

This is precisely what single-factor models struggle to deliver.

Practical difficulty. Only S_t is observed, while (X_t, Y_t) are latent. Calibration therefore becomes a *filtering* problem; increasing the number of factors improves fit but raises statistical and computational complexity.

5.4.2 Review of multi-factor approaches: sorting by the hidden state

Sorting principle. A workable (though not unique) taxonomy is: (i) *continuous* hidden state vs *discrete* Markov chain, and (ii) *Gaussian* vs *non-Gaussian* conditional dynamics.

(A) Continuous hidden state, Gaussian (Kalman-filter friendly). A representative family uses two Gaussian factors to drive a log-spot (or spot) with distinct economic interpretations ("level" vs "equilibrium" component). Typical specifications include:

- $S_t = e^{X_t}$ with X_t mean-reverting towards another (possibly drifting) factor Y_t;
- Y_t either Brownian with drift or mean-reverting to a long-run level;
- alternatively, multiplicative diffusions for (S_t, Y_t).

Advantage: linear-Gaussian state-space structure allows Kalman filtering. *Issue:* parameters of the latent dynamics can be weakly identified when spot noise is large; stability and robustness of estimation become central concerns.

(B) Continuous hidden state, non-Gaussian (Lévy/OU superpositions). A widely used arithmetic class is

$$S_t = \mu(t) + \sum_{i=1}^{n} w_i X_t^{(i)}, \qquad dX_t^{(i)} = -\lambda_i X_t^{(i)}\, dt + \sigma_i(t)\, dL_t^{(i)}, \tag{3.29}$$

where $L^{(i)}$ may include pure-jump components (optionally with seasonal intensity). *Key benefit:* tractable (often explicit) forward prices even with *delivery periods* (weeks/months/quarters), a crucial electricity feature. *Trade-off:* non-Gaussianity complicates filtering and calibration relative to Kalman-based setups.

(C) Hidden Markov models (regime switching). Here spikes are captured by a latent discrete regime with transition probabilities. A common calibration idea is to enforce that "spike states" have *low persistence*; then the model rapidly "forgets" spikes, which helps reconstruct forward prices at multiple granularities (week, month, quarter, year). *Limitation:* ensuring consistency with a full set of forward products and achieving stable estimation remains delicate.

Takeaway. Multi-factor models address forward-curve dynamics more credibly than one-factor models, but the modeling gain is paid in filtering complexity and calibration fragility; in practice, parsimonious specifications are often preferred.

5.5 Structural Models

Definition. Structural models build the electricity price as a (mostly) deterministic function of *observable* fundamentals (demand, available capacity, fuel prices, etc.). Randomness is then pushed into the dynamics of those fundamentals. This differs from reduced-form spot models where the state is largely latent.

5.5.1 The "mother" structural model: inverse demand with a capacity cap

Core idea. Price is obtained from an inverse demand curve $g(\cdot)$ with a price cap when demand approaches (or exceeds) maximal capacity. A tractable non-linear form uses $g(x) = a_0 - b_0 x^\alpha$ with $\alpha < 0$ and a cap mechanism. The resulting spot specification can be written as

$$S_t = \begin{cases} \left(\dfrac{a_0 - D_t}{b_0}\right)^{1/\alpha}, & \text{for moderate demand,} \\ A_0, & \text{near/above capacity (cap),} \end{cases} \tag{3.31}$$

with the cap level linked to auxiliary parameters (via ε_0). Under suitable transformations, the induced dynamics can be expressed as a non-linear OU-type process,

$$S_t = \begin{cases} (1+\alpha X_t)^{1/\alpha}, & 1+\alpha X_t \le \varepsilon_0, \\ A_0, & 1+\alpha X_t \ge \varepsilon_0, \end{cases} \tag{3.32}$$

where the driver X_t follows

$$dX_t = -\lambda(X_t - a)\,dt + \sigma\,dW_t. \tag{3.33}$$

Pros/cons. *Pro:* spikes arise endogenously when demand approaches capacity. *Con:* calibration is sensitive (strong non-linearity and cap effects); moreover, the mapping from fundamentals to traded forwards can be challenging in realistic market settings.

5.5.2 Auxiliary-variable models: explicit link to demand, capacity, and fuels

Demand & capacity as observable OU processes. A common brick is to set the spot (or a proxy) as an explicit function of demand D_t and capacity C_t, e.g.

$$P_t = \beta \exp(\gamma C_t + \alpha D_t), \qquad \alpha, \beta > 0,\ \gamma < 0, \tag{3.34}$$

with

$$D_t = g_t^D + X_t^D, \qquad dX_t^D = -\kappa_D X_t^D \, dt + \sigma_D(t) \, dW_t^D, \tag{3.35}$$

$$C_t = g_t^C + X_t^C, \qquad dX_t^C = -\kappa_C X_t^C \, dt + \sigma_C(t) \, dW_t^C, \tag{3.36}$$

where W^D and W^C are independent and g^D, g^C encode seasonality.

Analytical forwards. These choices yield closed-form conditional expectations for forward prices (exponentials of Gaussian terms), leading to explicit forward expressions (see the exponential-Gaussian structure summarized by (3.37)).

Fuel-based structural decomposition. Let S_t^i be the fuel price for technology i with heat rate h_i and available capacity C_t^i. The system marginal fuel price proxy is

$$\widehat{P}_t = \sum_{i=1}^{n} h_i S_t^i \, \mathbf{1}_{\{D_t \in I_t^i\}}, \tag{3.38}$$

where I_t^i is the demand interval covered by technology i given the merit order. To capture scarcity, define total capacity $C_t = \sum_i C_t^i$, reserve margin $R_t = C_t - D_t$, and set

$$P_t = g(R_t) \, \widehat{P}_t, \tag{3.39}$$

with scarcity function

$$g(x) := \min\Big(\frac{\gamma}{x^\nu}, \, M\Big) \, \mathbf{1}_{\{x>0\}} + M \, \mathbf{1}_{\{x \le 0\}}. \tag{3.40}$$

This extends the previous inverse-demand logic by modulating the marginal fuel price by a scarcity multiplier.

Hedgeable vs unhedgeable components and forward formula. A key structural insight is the separation between: (i) hedgeable fuel-price risk (via traded fuel futures), and (ii) unhedgeable fundamental risk (demand/capacity). For an electricity forward delivering 1 MWh at hour T,

$$F^e(t,T) = \sum_{i=1}^{n} h_i \, G_i^T(t, C_t, D_t) \, F_i(t,T), \tag{3.41}$$

where $F_i(t,T)$ is the *quoted* fuel-i futures price and the weights satisfy

$$G_i^T(t, C_t, D_t) = E_t^P\Big[g(R_T) \, \mathbf{1}_{\{D_T \in I_T^i\}}\Big]. \tag{3.42}$$

Thus, no stochastic model is required for $F_i(t,T)$ (they are observed), whereas G_i^T depends on the (historical) dynamics of demand/capacity and scarcity.

5.5.3 Stack curve models (bid/offer curves)

From fundamentals to bids. Rather than relying on a merit-order proxy only, stack-curve models represent the market clearing mechanism through bid curves $b_i(\cdot)$ for each technology/fuel. A generic formulation is

$$P_t = \max\left(\min_i b_i(S_t^i),\ \sup\left\{p : \sum_i \widehat{b}_i^{-1}(p, S_t^i) < D_t\right\}\right), \tag{3.43}$$

where $\widehat{b}_i^{-1}(p, S_t^i)$ is the offered quantity from unit i at price p (given fuel price S_t^i), and D_t is demand.

Interpretation and complexity. With few fuels (e.g. gas and coal), the clearing price can be described by a finite set of cases depending on which fuel is marginal and whether scarcity binds; this enables semi-explicit reasoning. However, realistic bid curves introduce:

- non-linearities and discontinuities (switches in merit order, capacity constraints),
- a strong dependence on market design and bidding behavior,
- substantial estimation burden (bid curve calibration and validation).

Takeaway. Structural models provide economic interpretability and allow decompositions (notably via fuel forwards and weights depending on scarcity and marginality). Their main limitations are the complexity of market bidding mechanisms and the calibration effort needed to turn observable fundamentals into reliable derivative prices.

5.6 Derivatives

General perspective. Electricity derivatives are shaped by three structural features: (i) *non-storability* of the underlying spot commodity, (ii) *delivery-period payoffs* (week/month/quarter baseload or peakload rather than a single instant), and (iii) *operational constraints* of generation and consumption. Consequently, many valuation problems are not plain "European option pricing" but either (a) *spread-option pricing* for correlated underlyings (power vs fuel / zones / carbon), or (b) *stochastic control / optimal switching* where the holder chooses how and when to exercise repeatedly.

5.6.1 Spreads

Definition and canonical payoff. A spread option is the right to receive the positive part of the difference between two assets (or two prices) net of a strike. In its simplest form, with maturity T,

$$p(T; S^a, S^b, K) = \left(S_T^a - S_T^b - K\right)^+. \tag{4.1}$$

In electricity, this structure appears naturally:

- **Locational spread:** S^a and S^b are prices in two zones/hubs; K represents (possibly simplified) transmission or congestion costs.
- **Fuel spread (spark/dark):** S^a is electricity, S^b is a fuel cost converted into /MWh through a heat rate (and possibly other add-ons).

Black–Scholes setting and the exchange case ($K = 0$). Assume two correlated lognormal assets under a risk-neutral measure:

$$dS_t^i = rS_t^i\,dt + \sigma_i S_t^i\,dW_t^i, \qquad i \in \{a, b\}, \qquad dW_t^a\,dW_t^b = \rho\,dt.$$

When $K = 0$, the spread option becomes an *option to exchange* one asset for another. Its value admits a closed form:

$$p_0 = e^{-rT}\Big(S_0^a N(d_1) - S_0^b N(d_2)\Big), \tag{4.2}$$

with the standard normal CDF $N(\cdot)$ and

$$d_1 = \frac{\ln(S_0^a/S_0^b) + \frac{1}{2}\sigma^2 T}{\sigma\sqrt{T}}, \qquad d_2 = d_1 - \sigma\sqrt{T}, \qquad \sigma^2 = \sigma_a^2 + \sigma_b^2 - 2\rho\sigma_a\sigma_b. \tag{4.3}$$

This is the key benchmark: for $K = 0$ one can price exactly via an equivalent volatility on the ratio S^a/S^b.

The general case ($K > 0$) and why approximations are used. For $K > 0$, no general closed form exists even in the lognormal framework. In practice, one relies on accurate and fast approximations whose quality depends on correlation, relative volatilities, and moneyness.

5.6.2 Kirk-type approximation

Idea. Replace the spread $(S_T^a - S_T^b - K)^+$ by an exchange-type structure where S_T^b is shifted by K and the ratio volatility is adjusted. The approximation has a Black–Scholes form:

$$\widehat{p}_K = e^{-rT}\Big(S_0^a N(d_1^K) - (S_0^b + K)N(d_2^K)\Big), \tag{4.4}$$

where

$$d_1^K = \frac{\ln\big(S_0^a/(S_0^b + K)\big) + \frac{1}{2}\sigma_K^2 T}{\sigma_K\sqrt{T}}, \qquad d_2^K = d_1^K - \sigma_K\sqrt{T},$$

and the *effective volatility* is

$$\sigma_K^2 = \sigma_a^2 + \left(\tfrac{S_0^b}{S_0^b+K}\right)^2 \sigma_b^2 - 2\rho\sigma_a\sigma_b\left(\tfrac{S_0^b}{S_0^b+K}\right). \tag{4.5}$$

Remark (asymmetry). The factor $\frac{S_0^b}{S_0^b+K}$ gives S^b and K a special role; the approximation is not symmetric in (a, b).

5.6.3 Eydeland–Wolyniec variant

Idea. Use the same shifted Black–Scholes structure as in Kirk, but with a different (more symmetric) volatility adjustment:

$$\widehat{p}_E = e^{-rT}\Big(S_0^a N(d_1^E) - (S_0^b + K)N(d_2^E)\Big), \tag{4.6}$$

with d_1^E, d_2^E defined as usual from an effective σ_E (constructed so that the adjustment does not privilege one asset as strongly as Kirk's).

$$\sigma_E^2 = \sigma_a^2 - 2\rho\sigma_a\sigma_b\frac{S_0^a}{S_0^b + K} + \sigma_b^2\left(\frac{S_0^a}{S_0^b + K}\right)^2$$

5.6.4 Moment-matching approximation

Idea. Approximate the distribution of the terminal spread by a Gaussian distribution with matching first two moments, then price as if the spread were normal. Denote by m the (discounted) mean of $(S_T^a - S_T^b)$ and by σ_M^2 the (discounted) variance. The approximation reads

$$\widehat{p}_M = (m - Ke^{-rT})\, N(d_M) - \sigma_M\, \varphi(d_M), \tag{4.7}$$

with φ the standard normal density and

$$d_M = \frac{m - Ke^{-rT}}{\sigma_M}.$$

Under the lognormal assumptions for (S^a, S^b), σ_M^2 can be written explicitly from $(\sigma_a, \sigma_b, \rho)$; in particular it includes the cross-term driven by $e^{\rho\sigma_a\sigma_b T}$:

$$\sigma_M^2 = e^{-2rT}\Big({S_0^a}^2(e^{\sigma_a^2 T} - 1) + {S_0^b}^2(e^{\sigma_b^2 T} - 1) - 2S_0^a S_0^b(e^{\rho\sigma_a\sigma_b T} - 1)\Big). \tag{4.8}$$

Practical interpretation. Moment-matching often behaves well when the spread is not too skewed and when the normal proxy captures the central mass of the distribution. In strongly skewed regimes, a direct Monte Carlo benchmark is typically used to validate approximation errors.

5.6.5 Multi-asset spread: electricity vs fuel and carbon

Clean spread structure. For thermal generation, the relevant margin often includes fuel and carbon costs. A common practical approximation extends Kirk's idea to three assets (electricity S^a, fuel S^b, carbon S^c) by shifting the denominator:

$$\widehat{p}_L = e^{-rT}\Big(S_0^a N(d_1^L) - (S_0^b + S_0^c + K)N(d_2^L)\Big), \tag{4.9}$$

with d_1^L, d_2^L obtained from an effective volatility σ_L built by weighting the fuel and carbon components through

$$\pi_b = \frac{S_0^b}{S_0^b + S_0^c + K}, \qquad \pi_c = \frac{S_0^c}{S_0^b + S_0^c + K},$$

and combining the volatilities and correlations accordingly. This provides a fast proxy for *clean* spreads when a full multi-dimensional numerical price is too heavy for daily risk usage.

What matters operationally. In electricity practice, spread options are not merely "exotic" instruments: they are the basic building block behind more complex contracts (tollings, swing, plant valuation). The key modeling inputs are joint dynamics and correlations (power–fuel–carbon, or zone–zone), and the key numerical requirement is *robustness* across market regimes (spikes, stress correlations, changing vol levels).

5.6.6 Power Plants and Tollings

From a single spread to a sequence of decisions. A tolling contract (or, more generally, the operational value of a thermal power plant) can be viewed as the right to repeatedly exploit a spread margin over time *subject to operating constraints*. Unlike a European spread option, the holder can decide when to run the plant and at which regime (off / on / partial load), incurring start-up and switching costs.

5.6.7 Optimal switching formulation

State process and regimes. Let X_t be a Markov process collecting the relevant market factors (electricity price, fuel price, possibly CO_2, possibly other variables), and let u_t be a *regime* taking values in a finite set $\mathcal{I}$ (e.g. off/on, or multiple production levels). For each regime $i \in \mathcal{I}$, define a running payoff rate $\phi(t, X_t, i)$ (net margin, including fixed costs). Switching from regime i to j costs $C_{i,j} \geq 0$.

Objective functional. Starting at time t with state $X_t = x$ and regime i, the operational value is

$$J(t,x,i) = \sup_{u \in \mathcal{U}(t)} E\Big[\int_t^T \phi(s, X_s, u_s)\, ds - \sum_{\tau_k \le T} C_{u_{\tau_k^-}, u_{\tau_k}}\Big]. \tag{4.10}$$

Here $(\tau_k)_k$ are switching times and $u_{\tau_k^-}$ denotes the regime just before switching. In this risk-neutral operational formulation, the main difficulty is not the choice of measure but the *control* embedded in the contract.

5.6.8 Dynamic programming with a bounded number of switches

Tractable recursion. Introduce $J_k(t,x,i)$ as the value when the number of switches is limited to k. Then:

$$J_0(t,x,i) = E\Big[\int_t^T \phi(s, X_s, i)\, ds \,\Big|\, X_t = x\Big], \tag{4.11}$$

and for $k \ge 1$,

$$J_k(t,x,i) = \sup_{\tau \in [t,T]} E\Big[\int_t^\tau \phi(s, X_s, i)\, ds + M_{k,i}(\tau, X_\tau) \,\Big|\, X_t = x\Big], \tag{4.12}$$

where the *intervention value* is

$$M_{k,i}(t,x) = \max_{j \in \mathcal{I},\, j \ne i} \Big(-C_{i,j} + J_{k-1}(t,x,j)\Big). \tag{4.13}$$

This recursion encodes the core economics: *wait* and accrue ϕ in the current regime until a stopping time τ, then *switch* to the best alternative net of cost.

5.6.9 Quasi-variational inequality characterization

Generator form. If X_t satisfies an Itô diffusion $dX_t = \mu(X_t)dt + \sigma(X_t)dW_t$, denote by $\mathcal{L}_X$ its infinitesimal generator. Under regularity conditions, the value function solves a QVI of the form

$$\max\Big\{\mathcal{L}_X v + \phi,\ -v + \mathcal{M}v\Big\} = 0, \tag{4.14}$$

where $(\mathcal{M}v)(t,x,i) = \max_{j \ne i}\{-C_{i,j} + v(t,x,j)\}$. The two terms correspond respectively to continuation (do not switch) and intervention (switch now). The QVI expresses the partition of the state space into continuation and switching regions.

5.6.10 BSDE representation (simulation-oriented)

Reflected system. A further representation (useful for numerical schemes) uses a backward stochastic differential equation system. With $Y_t^{k,i} = J_k(t, X_t, i)$, one can formulate

$$Y_t^{k,i} = \int_t^T \phi(s, X_s, i)\, ds + A_T^{k,i} - A_t^{k,i} - \int_t^T Z_s^{k,i}\, dW_s, \tag{4.15}$$

subject to the reflection constraint

$$M_{k,i}(t, X_t) \le Y_t^{k,i}, \tag{4.16}$$

and the Skorokhod condition (minimal push)

$$\int_0^T \big(Y_t^{k,i} - M_{k,i}(t, X_t)\big)\, dA_t^{k,i} = 0, \qquad A_0^{k,i} = 0. \tag{4.17}$$

This formulation emphasizes that the value must stay above the switching value; A enforces the constraint, and Z acts as the hedge-like sensitivity in the Brownian directions.

5.6.11 Hedging, incompleteness, and indifference pricing

Utility-based valuation. When the owner can trade some instruments (e.g. liquid forwards/futures) but cannot hedge all sources of risk, the market is incomplete. A common approach is to maximize expected utility of terminal wealth under admissible operating (*switching*) and hedging strategies:

$$V(t, y, x, i) = \sup_{\pi,\xi} E_{t,x,y,i}\big[U(X_T^{t,x,\pi,\xi})\big], \tag{4.18}$$

where π denotes the hedging strategy and ξ the operating control; U is often exponential utility $U(w) = -e^{-\gamma w}$.

Indifference price. Let $U_0(t, x)$ denote the maximal utility without owning/operating the plant (only trading). The *indifference price* $p_{t,T}(x, y, i)$ is defined by

$$V(t, y, x, i) = U_0\big(t, x + p_{t,T}(x, y, i)\big). \tag{4.19}$$

It is the cash amount that makes the owner indifferent between (a) having the plant and lower initial cash, and (b) having more cash but no plant. This definition yields a valuation in currency even though V itself is in utility units.

Optimal hedge (qualitative structure). In exponential-utility settings, the optimal hedge has a delta-like form modified by a factor that accounts for the effectiveness of the traded hedge instrument; schematically,

$$\pi_t^* \propto -\rho\, \frac{\partial p(t, \cdot)}{\partial(\text{hedge factor})}, \tag{4.20}$$

where the proportionality involves volatility scaling and reflects that hedging is imperfect when the hedge instrument is not perfectly correlated with the plant value driver.

Key takeaway. Power plant/tolling valuation is fundamentally a *multi-exercise* problem with switching costs and constraints. Closed forms are the exception; dynamic programming, QVIs, or BSDE-based simulation schemes are the standard tools, and incompleteness naturally motivates indifference pricing.

5.6.12 Storage and Swings

Why "storage" exists in electricity. While electricity cannot be stored economically as a commodity, *flexibility* exists through hydro reservoirs (stored potential energy), pumped storage, or contractual flexibility (swing). Such products embed intertemporal constraints, turning valuation into control.

5.6.13 Hydro storage as a continuous-time control problem

State variables and controls. Consider a single-reservoir hydro system. Let S_t be the (Markov) spot price, A_t a (Markov) inflow process, and X_t the reservoir level. Controls are the production rate $q_t \in [0, \overline{q}]$ and a spillage rate $\delta_t \geq 0$.

Value function. A canonical formulation is

$$V(t, s, a, x) = \sup_{q_u \in [0,\overline{q}],\ \delta_u \geq 0} E\Big[\int_t^T q_u\, S_u\, du + g(S_T, X_T)\Big], \tag{4.21}$$

where g is a terminal reward (or salvage value) preventing the degenerate policy of emptying the reservoir at T.

Reservoir dynamics and constraints. The reservoir evolves as

$$dX_u = (A_u - q_u - \delta_u)\, du, \tag{4.22}$$

with hard bounds $X_u \in [\underline{x}, \overline{x}]$. The spillage δ is needed to enforce the upper bound when inflows are high.

Interpretation. The control trades off current revenue $q_u S_u$ against the option value of water kept for future high-price periods, under physical bounds. Numerically, this is a high-dimensional DP/PDE/Monte-Carlo control problem once multiple reservoirs or stochastic inflows are included.

5.6.14 Swing options as discrete-time constrained control

Contract structure. A standard swing contract grants the right to choose purchase/sale quantities at a finite set of exercise dates $t_0, \ldots, t_{N-1}$, each quantity bounded:

$$q(t_i) \in [\underline{q}, \overline{q}].$$

Define cumulative consumption $Q(t_i) = \sum_{j \leq i} q(t_j)$, typically constrained at maturity by

$$Q_T \in [Q_m, Q_M].$$

Valuation problem. A representative formulation is

$$V(t, s, Q) = \sup_{q(t_i) \in [\underline{q}, \overline{q}]} E\Big[\sum_{i=0}^{N-1} \psi\big(t_i, q(t_i), S(t_i), Q(t_i)\big) + P(T, S_T, Q_T)\Big], \quad (4.23)$$

where ψ encodes the instantaneous cashflow at each nomination date and P is a terminal penalty/reward enforcing global constraints.

Central difficulty: intertemporal coupling. Unlike a strip of independent options, the constraint on Q couples all decisions across time. The optimal policy is typically characterized by threshold regions in (S, Q) and solved via backward induction (on a discretized state space) or via simulation-based methods.

Key takeaway. Storage and swing valuation is *control with constraints*. The essential modeling choices are the Markov structure for S (and possibly A), the discretization of the control set, and the numerical method used to preserve monotonicity and constraint feasibility.

5.6.15 Retail Contracts

Economic setting. Retail electricity contracts (selling to end customers) often involve a fixed price while the utility faces:

- **volume risk:** stochastic customer load,
- **price risk:** stochastic spot/imbalance prices for the unhedged residual,
- **basis risk:** imperfect match between hedge instruments and realized consumption profile.

Even with forward hedging, the residual risk is material because consumption is not perfectly tradable.

5.6.16 Risk-adjusted pricing via capital and target return

Principle. A practical approach sets the retail price p to achieve a target return μ on a risk-capital measure. Denote by $\Pi(p)$ the expected cashflow (profit) associated with retail price p, and by $R(p)$ the random cashflow itself (or a loss variable derived from it). Using a quantile-based risk measure (cashflow-at-risk), one seeks p satisfying

$$\mu = \frac{\Pi(p)}{\Pi(p) - q_\alpha(R(p))}, \quad (4.24)$$

where $q_\alpha(\cdot)$ is the α-quantile.

Interpretation. The denominator $\Pi(p) - q_\alpha(R(p))$ plays the role of economic capital: it measures the buffer needed so that, with confidence level α, cashflows do not fall below an acceptable threshold.

Refinements commonly required. Implementations typically extend this idea to incorporate:

- correlation between customer load and system load (hence correlation with spot prices),
- the fraction of future load that is hedgeable with standard forward products,
- spike-sensitive spot models consistent with the utility's exposure (especially imbalance settlement).

Key takeaway. Retail pricing is less about replication and more about *risk budgeting*: quantify residual exposure after hedging and translate it into a price through a return-on-capital constraint.

5.6.17 Weather Derivatives

Underlying and motivation. Weather derivatives are written on non-tradable, non-storable underlyings (temperature, precipitation, wind). Utilities are highly sensitive to temperature (heating/cooling demand) and increasingly to wind (renewables generation variability). These products sit at the boundary between insurance and finance.

5.6.18 Contract standardization: four ingredients

Ingredients. A standard contract is defined by:

- a **weather index** (e.g. degree days),
- a **delivery period** (month/season),
- a **location** (weather station),
- a **tick size** converting index units to currency.

Degree-day indices. For day t, let T_t^M and T_t^m be the maximum and minimum daily temperatures, and define the daily mean $\frac{T_t^M + T_t^m}{2}$. With threshold θ (often 18°C), the *heating degree days* (HDD) are

$$\eta_t = \left(\theta - \tfrac{T_t^M + T_t^m}{2}\right)^+,$$

and similarly cooling degree days (CDD) are defined with the sign reversed. Indices are typically accumulated over a month or season: $\sum_{t \in \text{period}} \eta_t$.

Tick size and payoff scale. If the tick size is A (e.g. 100 per degree-day unit), a forward on the accumulated HDD index pays $A \sum_t \eta_t$ at maturity. Vanilla options can be written as well, for instance a call on the accumulated index with strike K:

$$A\Big(\sum_t \eta_t - K\Big)^+.$$

5.6.19 Pricing in an incomplete market

Why classical replication fails. Neither actuarial diversification (across many independent policyholders) nor dynamic hedging is fully effective for weather claims:

- weather shocks can affect many utilities simultaneously (limited cross-sectional diversification),
- the underlying is not directly tradable (no perfect temporal diversification).

Main approaches in the literature. Three families dominate:

- **Utility-based pricing:** representative-agent equilibrium ideas, or agent-specific marginal utility valuation, or indifference pricing.
- **Equilibrium with two agents:** characterize conditions under which a buyer and a seller agree on an admissible price for an unhedgeable claim.
- **Market price of risk calibration:** infer a risk premium by fitting models to traded weather derivatives quotes (when such quotes are available), analogously to fitting an electricity forward curve via a market price of risk.

5.6.20 Market-development constraint

Liquidity and customization. Despite long-standing efforts and exchange listings (notably in the U.S.), standardized weather-derivative markets have remained relatively thin. A significant fraction of transactions is conducted OTC via customized contracts, reflecting the difficulty of designing standardized indices with sufficiently direct balance-sheet impact for heterogeneous industrial exposures.

Key takeaway. Weather derivatives combine (i) careful index engineering and (ii) incomplete-market valuation. Practically, the modeling challenge reduces to describing the underlying meteorological factor (temperature, wind, precipitation) and choosing a consistent risk-premium or utility-based valuation principle.