Received: 25 February 2021 | Revised: 24 June 2021 | Accepted: 19 August 2021

DOI: 10.1002/asmb.2645

RESEARCH ARTICLE

WILEY

A self-exciting modeling framework for forward prices in power markets

Giorgia Callegaro[1] | **Andrea Mazzoran**[1] | **Carlo Sgarra**[2]

[1]Dipartimento di Matematica Applicata, Universitá di Padova, Padua, Italy

[2]Dipartimento di Matematica, Politecnico di Milano, Milan, Italy

Correspondence
Carlo Sgarra, Dipartimento di Matematica, Politecnico di Milano, Milan, Italy.
Email: carlo.sgarra@polimi.it

Abstract

We propose and investigate two model classes for forward power price dynamics, based on continuous branching processes with immigration, and on Hawkes processes with exponential kernel, respectively. The models proposed exhibit jumps clustering features. Models of this kind have been already proposed for the spot price dynamics, but the main purpose of the present work is to investigate the performances of such models in describing the forward dynamics. We adopt a Heath–Jarrow–Morton approach in order to capture the whole forward curve evolution. By examining daily data in the French power market, we perform a goodness-of-fit test and we present our conclusions about the adequacy of these models in describing the forward prices evolution.

KEYWORDS

branching processes, forward prices, Hawkes processes, Heath–Jarrow–Morton model, jumps clustering, power markets, self-exciting processes

1 | INTRODUCTION

Energy markets, and in particular, electricity markets, exhibit very peculiar features. The historical series of both futures and spot prices include seasonality, mean-reversion, spikes, and small fluctuations. One can alternatively describe the power price dynamics by modeling the spot or the forward price. In the former case, the spot price can be obtained as a limit of the forward price when the maturity is close to the current time, in the latter case it is possible to derive the forward price from the spot by computing the conditional expectation with respect to a suitable risk-neutral measure of the spot price at the maturity. Both spot and forward electricity prices models are extremely relevant for investors, consumers,[1] and regulators in order to define optimal strategies both for risk management and portfolio optimization.[2,3] Some derivatives contracts moreover are peculiar of energy markets.[4]

After the pioneering paper by Schwartz,[5] where an Ornstein–Uhlenbeck dynamics is assumed to describe the spot price behavior, several different approaches have been investigated in order to describe the power price evolution. A comprehensive literature review until 2008 is offered in the book by Benth et al..[6] A similar effort has been devoted to identify reliable models for the forward price dynamics, and a huge amount of literature is available focusing on Heath–Jarrow–Morton type models as in Benth et al.[7] and in Filimonov et al.,[8] in the attempt to provide a description of the whole forward curves dynamics, in analogy with forward interest rates in fixed-income markets as in Heath et al..[9] Some of the classical models proposed include jumps and/or stochastic volatility. Benth and Paraschiv[10] propose a random field approach based on Gaussian random fields by adopting the Musiela parameterization in order to describe

the forward curve dynamics. Empirical evidence suggests that in many assets prices often jumps appear in cluster, thus requiring the introduction of jump processes exhibiting a clustering or self-exciting behavior.

Kiesel and Paraschiv[11] recently presented a systematic empirical investigation of electricity intraday prices and suggested an approach based on Hawkes processes in order to describe the power price dynamics with jump clustering features. Self-exciting features in electricity prices attracted already some attention by several authors: Herrera and Gonzalez[12] proposed a self-excited model for electricity spot prices, while Christensen et al.[13] and Clements et al.[14] pointed out that time between spikes has a significant impact on the likelihood of future occurrences, thus providing a strong support to models including self-exciting properties.

The large class of models available in the literature describing the power price dynamics is then widening in order to include models exhibiting self-exciting features.

We also mention the paper by Jiao et al.,[15] where a model based on continuous branching processes with immigration (CBI) for power spot prices was proposed, and the forward prices computed with respect to a suitable structure preserving equivalent martingale measure.

Eyjolfsson and Tjøshteim[16] describe a class of Hawkes processes and present an empirical investigation based on data from UK power market supporting Hawkes-type models for spot prices.

The purpose of the present article is to investigate if self-exciting features can arise in the power forward prices evolution as well, and in order to perform this investigation we shall focus on two different model classes: the CBI and the Hawkes processes. While CBI processes are always affine, Hawkes processes in general are not, but when the kernel describing the intensity dynamics is of exponential type they are, and this feature makes the Hawkes processes with exponential kernel appealing from the modeling point of view. By considering the two model classes mentioned before, that is, CBI and Hawkes processes, we then want to provide the description of the full term structure of power forward prices, following a Heath–Jarrow–Morton approach.

Power is a flow commodity, this meaning that instantaneous forward contracts are not directly traded on the market, but futures (sometimes called flow forward) are. So, in order to perform any kind of inference on the model proposed, it is necessary to extract the relevant information on the forward dynamics included in the futures prices. This can be done by applying suitable optimization procedures proposed in the literature, eventually modified in order to provide the best performances in the case under examination. These procedures are far from trivial from the computational point of view and require a careful implementation of the optimization step. We deliberately chose to work on daily data, in order to show how self-exciting effects can arise not only on a small time scale, but also at a coarser level.

The article is organized as follows: in Section 2, we introduce the processes on which our models are based and in Section 3, we present and discuss the models proposed for the forward power price dynamics. In Section 4, we discuss the dynamics of futures contracts when the forward dynamics is assumed to be given by the models introduced. From Sections 5 to 8, we provide the theoretical background and numerical results relative to the calibration/parameters' estimation for the model proposed. We conclude by providing some remarks and discussing future extensions of the present work.

2 | THE MODELING FRAMEWORK

2.1 | Continuous branching processes with immigration

We now introduce our modeling framework for the electricity price, which is based on stochastic differential equations driven by Lévy random fields. We consider a Lévy random field, which is a combination of a Gaussian random measure W and a compensated Poisson random measure N independent of W. For a background on such general stochastic equations with jumps, we refer the readers, for example, to Dawson and Li,[17] Li and Ma,[18] and Walsh.[19]

Let us now briefly introduce all the relevant ingredients of our work and recall some preliminary results. We fix a probability space $(\Omega, \mathcal{A}, \mathbb{P})$. A *white noise* W on $\mathbb{R}^2_+$ is a Gaussian random measure such that, for any Borel set $A \in \mathcal{B}(\mathbb{R}^2_+)$ with finite Lebesgue measure $|A|$, $W(A)$ is a normal random variable of mean zero and variance $|A|$ and if $A_1, \dots, A_n$ are disjoint Borel sets in $\mathcal{B}(\mathbb{R}^2_+)$, then $W(A_1), \dots, W(A_n)$ are mutually independent. We denote by N the Poisson random measure on $\mathbb{R}^3_+$ with *intensity* λ which is a Borel measure on $\mathbb{R}^3_+$ defined as the product of the Lebesgue measure on $\mathbb{R}_+ \times \mathbb{R}_+$ with a Borel measure μ on $\mathbb{R}_+$ such that $\int_0^\infty (z \wedge z^2)\mu(dz) < +\infty$. Note that μ is a Lévy measure since $\int_0^\infty (1 \wedge z^2)\mu(dz) < +\infty$. Recall that for each Borel set $B \in \mathcal{B}(\mathbb{R}^3_+)$ with $\lambda(B) < +\infty$, the random variable $N(B)$ has the Poisson

distribution with parameter $\lambda(B)$. Moreover, if $B_i, i = 1, \dots, n$ are disjoint Borel sets in $\mathcal{B}(\mathbb{R}^3_+)$, then $N(B_1), \dots, N(B_n)$ are mutually independent. We let $\sim N = N - \lambda$ be the compensated Poisson random measure on $\mathbb{R}^3_+$ associated to N.

We introduce the filtration $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$ as the natural filtration generated by the Lévy random field (see Dawson and Li[17]) and satisfying the usual conditions. Namely, for any Borel subset $A \in \mathcal{B}(\mathbb{R}_+)$ and $B \in \mathcal{B}(\mathbb{R}^2_+)$ of finite Lebesgue measure, the processes $(W([0,t]\times A), t \geq 0)$ and $(\sim N([0,t]\times B), t \geq 0)$ are $\mathbb{F}$-martingales.

We consider the following stochastic differential equation in the integral form. Let $a, b, \sigma, \gamma \in \mathbb{R}_+$ be constant parameters. Consider the equation:

$$Y(t) = Y(0) + \int_0^t a\,(b - Y(s))\,ds + \sigma \int_0^t \int_0^{Y(s)} W(ds, du) + \gamma \int_0^t \int_0^{Y(s-)} \int_{\mathbb{R}^+} z\tilde{N}(ds, du, dz), \tag{1}$$

where $W(ds, du)$ is a white noise on $\mathbb{R}^2_+$ with unit covariance, $\sim N(ds, du, dz)$ is an independent compensated Poisson random measure on $\mathbb{R}^3_+$ with intensity $ds\,du\,\mu(dz)$ with $\mu(dz)$ being a Lévy measure on $\mathbb{R}_+$ and satisfying $\int_0^\infty (z \wedge z^2)\mu(dz) < \infty$.

The integrals appearing in Equation (1) (and in the following) are both in the sense of Walsh.[19] It follows from Dawson and Li[17(Theorem3.1)] or Li and Ma[18(Theorem2.1)] that Equation (1) has a unique strong solution.

Our model actually belongs to the family of CBI processes, a class of stochastic processes commonly used in modeling population dynamics, as in Pardoux.[20] The self-exciting features, arising from the integrals in Equation (1) extended on the domain $[0, Y(s))$ with respect to the integration variable u, describe the growth of the population due to the reproduction of the previous generations. In the present modeling framework, they are used to describe jumps generated by previous jumps. We briefly recall the definition by Kawazu and Watanabe.[21(Def.1.1)]

Definition 1. A Markov process Y with state space $\mathbb{R}_+$ is called a CBI process characterized by branching mechanism $\Psi(\cdot)$ and immigration rate $\Phi(\cdot)$, if its characteristic representation is given, for $p \geq 0$, by:

$$\mathbb{E}_y\left[e^{-pY(t)}\right] = \exp\left(-yv(t,p) - \int_0^t \Phi\,(v(s,p))\,ds\right),$$

where $\mathbb{E}_y$ denotes the conditional expectation with respect to the initial value $Y(0) = y$. The function $v : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}_+$ satisfies the following differential equation:

$$\frac{\partial v(t,p)}{\partial t} = -\Psi(v(t,p)), \quad v(0,p) = p$$

and Ψ and Φ are functions of the variable $q \geq 0$ given by

$$\Psi(q) = aq + \frac{1}{2}\sigma^2 q^2 + \gamma \int_0^\infty (e^{-qu} - 1 + qu)\pi(du),$$
$$\Phi(q) = abq + \int_0^\infty (1 - e^{-qu})\nu(du),$$

with $\sigma, \gamma \geq 0$, $\beta \in \mathbb{R}$, and π, ν being two Lévy measures such that

$$\int_0^\infty (u \wedge u^2)\pi(du) < \infty, \quad \int_0^\infty (1 \wedge u)\nu(du) < \infty.$$

It is proved in Reference 17 (Theorem 3.1) that the process in Equation (1) is a CBI process with the branching mechanism Ψ given by:

$$\Psi(q) = aq + \frac{1}{2}\sigma^2 q^2 + \int_0^\infty (e^{-q\gamma z} - 1 + q\gamma z)\mu(dz) \tag{2}$$

and the immigration rate $\Phi(q) = abq$.

The link between CBI processes and the affine term structure models has been established by Filipović[22]. If the process Y takes values in $\mathbb{R}_+$ he proved equivalence between the two classes. We recall that the joint Laplace transform of a CBI process Y and its integrated process, which is given in Filipović,[22(Theorem 5.3)] is defined as follows: for non-negative real numbers ξ and θ, we have:

$$\mathbb{E}_y\left[e^{-\xi Y(t)-\theta\int_0^t Y(s)ds}\right]=\exp\left\{-yv(t,\xi,\theta)-\int_0^t\Phi\left(v(s,\xi,\theta)\right)ds\right\}, \tag{3}$$

where $v(t,\xi,\theta)$ is the unique solution of

$$\frac{\partial v(t,\xi,\theta)}{\partial t}=-\Psi(v(t,\xi,\theta))+\theta,\quad v(0,\xi,\theta)=\xi. \tag{4}$$

2.2 | Hawkes processes

A Hawkes process is a special counting process with a random intensity function. We introduce now the Hawkes processes with exponential kernel. They can be written as follows:

$$Y(t)=Y(0)+\sum_{i=1}^{N_t}Z_i=Y(0)+\int_0^t\int_0^\infty zJ(ds,dz), \tag{5}$$

where the last term is an Itô integral, N_t is the number of jumps in the interval between 0 and t, and $J(dz,ds)$ is a Poisson random measure with intensity $\lambda(t)$, satisfying the SDE:

$$\begin{aligned}\lambda(t)&=\lambda(0)-\beta\int_0^t\lambda(s)ds+\alpha\int_0^t\int_0^\infty zJ(ds,dz)\\&=\exp(-\beta t)\lambda(0)+\alpha\sum_{i=1}^{N_t}\exp[-\beta(t-t_i)]Z_i.\end{aligned} \tag{6}$$

Here $\beta>0$ is the rate of exponential decay of the influence of previous jumps on the intensity level and α the amplitude of the memory kernel, t_i are the jumps times and Z_i the jump sizes, which we shall assume distributed according to an exponential density with parameter δ, so that only positive jumps appear in both Equations (5) and (6), and we can write $\tilde{J}(ds,dz)=J(ds,dz)-\lambda(s)\mu(dz)ds$ and $\mu(dz)=\delta\exp(-\delta z)dz$, where $\tilde{J}(ds,dz)$ denotes the compensated version of the Poisson measure $J(ds,dz)$. We assume the following condition holds: $\beta-\alpha/\delta>0$, granting the non-explosiveness of the Hawkes process (see, e.g., Bernis et al.[23]).

Hawkes processes with exponential kernel are the only class of Hawkes processes exhibiting both the Markov property and an affine structure (see, e.g., Errais et al.[24]). They have been extensively used in order to describe the dynamics of several asset classes, including equities as in Hainaut and Moraux,[25] commodities as in Eyjolfsson and Tjøsteim,[16] exchange rates as in Rambaldi et al.,[26] and credit risk as in Errais et al..[24]

The left-hand side plot in Figure 1 illustrates three trajectories of a Hawkes process with exponential kernel (without diffusion and deterministic component and integrated with respect to the non-compensated jump measure). It is obtained by a MATLAB simulation algorithm with parameters $\lambda(0)=2.9,\beta=2.27,\alpha=1.5$. The right-hand side graph in Figure 1 shows a trajectory of a CBI process without diffusion and deterministic components and with intensity measure $ds\mu(dz)du$ with $\mu(dz)$ in (1) chosen as a spectrally positive α-stable Lévy measure (for the relation between α-stable processes and CBI see Reference 27) with parameter $\alpha=0.7$, obtained with a MATLAB simulation algorithm as well. A simulation scheme for CBI based on the Lamperti transform can be found in the paper by Caballero et al.,[28] while Dassios and Zhao[29] propose an efficient simulation scheme for Hawkes processes.

3 | FORWARD PRICES MODELING

In this section, we are going to introduce the two alternative models for the forward prices, that we are going to test against electricity market data. In both cases, the price at time t of a forward contract with maturity $T\geq t$ is additive:

$$f(t,T)=\Lambda(t)-\Lambda(0)+\sum_{i=1}^{n}X_i(t,T), \tag{7}$$

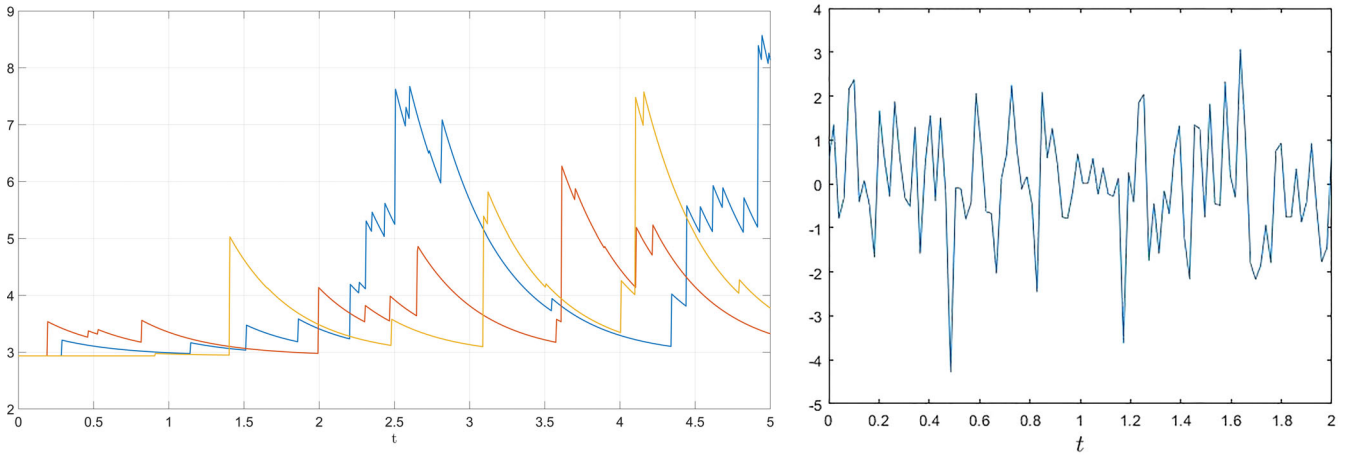


FIGURE 1 Three trajectories of a Hawkes process with exponential kernel (on the left) and a trajectory of a CBI process with α-stable Lévy measure (on the right)

where $\Lambda(t)$ is a deterministic seasonality function that will be made precise later on, n is the number of factors used and each of the terms X_i is an underlying factor, whose dynamics will be specified in Sections 3.1 and 3.2.

3.1 | The forward model based on CBI

Let n be the number of factors in model (7) and let $a_i, \sigma_i, \gamma_i \in \mathbb{R}_+,\ i = 1, \dots, n$ be constant parameters. Our first model assumes the following dynamics for the factors $X_i, i = 1, \dots, n$:

$$X_i(t,T) = X_i(0,T) - \int_0^t a_i X_i(s,T)ds + \sigma_i \int_0^t \int_0^{X_i(s,T)} W_i(ds,du) + \gamma_i \int_0^t \int_0^{X_i(s-,T)} \int_{\mathbb{R}_+} z\tilde{N}_i(ds,du,dz). \tag{8}$$

Assuming the dynamics for the factors X_i as in (8), it is possible to rewrite Equation (7) as follows:

$$f(t,T) = \Lambda(t) - \Lambda(0) + f(0,T) - \sum_{i=1}^n \int_0^t a_i X_i(s,T)ds + \sum_{i=1}^n \sigma_i \int_0^t \int_0^{X_i(s,T)} W_i(ds,du) + \sum_{i=1}^n \gamma_i \int_0^t \int_0^{X_i(s-,T)} \int_{\mathbb{R}_+} z\tilde{N}_i(ds,du,dz), \tag{9}$$

where $f(0,T) = \sum_{i=1}^n X_i(0,T)$.

The relation between the dynamics of the forward price with respect to the historical measure $\mathbb{P}$ and the risk-neutral dynamics, written with respect to $\mathbb{Q}$, can be easily obtained by applying the following result, proved in the paper by Jiao et al.[15(Proposition4.1)]

Proposition 1. *Let $X_1, X_2, \dots, X_n$ be independent CBI processes where for each $i \in \{1, \dots, n\}$, X_i is a CBI process under the probability measure $\mathbb{P}$, with dynamics given by Equation (8). Assume that the filtration $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$ is generated by the random fields $W_1, W_2, \dots, W_n$ and by the compensated Poisson random measures $\sim N_1, \sim N_2, \dots, \sim N_n$. For each i, fix $\eta_i \in \mathbb{R}$ and $\xi_i \in \mathbb{R}_+$ and define*

$$U_t := \sum_{i=1}^n \eta_i \int_0^t \int_0^{X_i(s)} W_i(ds,du) + \sum_{i=1}^n \int_0^t \int_0^{X_i(s-)} \int_0^\infty (e^{-\xi_i z} - 1)\tilde{N}_i(ds,du,dz). \tag{10}$$

Then the Doléans–Dade exponential $\mathcal{E}(U)$ is a martingale under $\mathbb{P}$ and the probability measure $\mathbb{Q}$ defined by

$$\left.\frac{d\mathbb{Q}}{d\mathbb{P}}\right|_{\mathcal{F}_t} = \mathcal{E}(U)_t, \tag{11}$$

is equivalent to $\mathbb{P}$. Moreover, under $\mathbb{Q}$, X_i is a CBI process with parameters $(a_i^{\mathbb{Q}}, b_i^{\mathbb{Q}}, \sigma_i^{\mathbb{Q}}, \gamma_i^{\mathbb{Q}}, \mu_i^{\mathbb{Q}})$, where:

$$a_i^{\mathbb{Q}} = a_i^{\mathbb{P}} - \sigma_i^{\mathbb{P}}\eta_i - \int_0^\infty z(e^{-\theta_i z} - 1)\mu_i^{\mathbb{P}}(dz), \tag{12}$$

$$b_i^{\mathbb{Q}} = a_i^{\mathbb{P}} b_i^{\mathbb{P}} / a_i^{\mathbb{Q}}, \quad \sigma_i^{\mathbb{Q}} = \sigma_i^{\mathbb{P}}, \quad \gamma_i^{\mathbb{Q}} = \gamma_i^{\mathbb{P}}, \tag{13}$$

$$\mu_i^{\mathbb{Q}}(dz) = e^{-\theta_i z}\mu_i^{\mathbb{P}}(dz), \quad \delta_i^{\mathbb{Q}} = \delta_i^{\mathbb{P}}. \tag{14}$$

Remark 1. In this context, the parameters η_i, ξ_i can be interpreted as the market price of risk associated with the diffusion/jump part of X_i, $i = 1, \dots, n$, respectively.

Remark 2. In order to avoid arbitrage opportunities, we shall assume that the de-seasonalized dynamics of every factor X_i is a local martingale under $\mathbb{Q}$ and this will automatically imply that $a_i = 0$ under $\mathbb{Q}$. Since the first integral is defined with respect to the Gaussian white noise $W_i(ds, du)$ and the second integral is defined with respect to the compensated Poisson random measure $\tilde{N}_i(ds, du, dz)$, each process $X_i(t, T)$ is in fact a local martingale with respect to $\mathbb{Q}$.

Remark 3. From (14), specifying the relations between the model parameters under the risk-neutral measure $\mathbb{Q}$ and the historical measure $\mathbb{P}$, it is clear that in the present modeling framework, for each factor X_i, a mean reversion speed coefficient a_i can be non-null under $\mathbb{P}$ and zero under $\mathbb{Q}$. As far as the immigration term b_i is concerned, if it vanishes under $\mathbb{Q}$, it will be zero under any equivalent probability measure.

3.2 | The forward model based on Hawkes processes

Let n be the number of factor in model (7) and let $c_i, \sigma_i \in \mathbb{R}_+$, $i = 1, \dots, n$ be constant parameters. As alternative to the model proposed in the previous subsection, we consider, under ¶, the dynamics (8), where now each X_i, $i = 1, \dots, n$, satisfies a SDE of the following form:

$$X_i(t, T) = X_i(0, T) - \int_0^t c_i X_i(s, T)ds + \sigma_i \int_0^t \sqrt{X_i(s, T)}dW_i(s) + \int_0^t \int_0^\infty z\tilde{J}_i(dz, ds), \tag{15}$$

where $\tilde{J}_i(dz, ds)$ are compensated marked point process with intensity $\lambda_i(t)$, satisfying the SDE:

$$\lambda_i(t) = \lambda_i(0) - \beta_i \int_0^t \lambda_i(s)ds + \alpha_i \int_0^t \int_0^\infty z J_i(ds, dz), \tag{16}$$

where we recall that $\beta_i > 0$ is the rate of exponential decay of the influence of previous jumps on the intensity level and α_i the amplitude of the memory kernel of each factor X_i. We assume the jump size distributed according to an exponential density with parameter δ_i for each (λ_i, X_i), so we can write:

$$\tilde{J}_i(ds, dz) = J_i(ds, dz) - \lambda_i(s)\mu(dz)ds = J_i(ds, dz) - \lambda_i(s)\delta_i \exp(-\delta_i z)(dz)ds. \tag{17}$$

Remark 4. The choice of a square-root process for the diffusion part of the forward curves dynamics is motivated by the positivity requirement as well as the choice of the exponential distribution for the size of the jumps.

Therefore, in this case, Equation (7) takes the following form:

$$\begin{aligned} f(t, T) = \Lambda(t) - \Lambda(0) + f(0, T) - \sum_{i=1}^n \int_0^t c_i X_i(s, T)ds + \sum_{i=1}^n \int_0^t \sigma_i \sqrt{X_i(s, T)}dW_i(s) \\ + \sum_{i=1}^n \int_0^t \int_0^\infty z\tilde{J}_i(dz, ds), \end{aligned} \tag{18}$$

where $f(0,T) = \sum_{i=1}^{n} X_i(0,T)$.

In order to make the presentation of the two model classes more homogeneous, we can introduce the Dawson–Li representation for the Hawkes-type dynamics as well and write the SDE governing the dynamics of forward prices under the historical measure $\mathbb{P}$ as follows:

$$X_i(t,T) = X_i(0,T) - \int_0^t c_i X_i(s,T)ds + \int_0^t \int_0^{X_i(s,T)} \sigma_i W_i(du,ds) + \int_0^t \int_0^{X_i(s_-,T)} \int_{\mathbb{R}_+} z\tilde{N}_i(dz,du,ds), \tag{19}$$

where the definition of the integrals and the notations are the same as in Section 2.1 and the $\lambda_i(t)$ evolve according to Equation (16).

Remark 5. We point out the main difference between the stochastic differential equations (8) and (15) consists in the type of stochastic integrals used: in the first case Walsh type integral are in use and in the second Ito type integrals appear, by following the usual representations adopted in the literature for CBI and Hawkes processes, respectively.

It is immediate to remark that the dynamics described by the two model classes look almost identical when written in the Dawson–Li representation, the main difference being the specification of the equation governing the evolution of the intensity processes. This is one of the reasons behind the choice of these two alternative models to describe the forward prices' evolution.

The dynamics just described is given with respect to the historical probability measure $\mathbb{P}$. In order to obtain a description with respect to the risk-neutral measure $\mathbb{Q}$ we need to introduce a measure change. The following proposition provides a measure change preserving the Hawkes-type dynamics. A proof can be found in Bernis et al..[30]

Proposition 2. *Let (λ_i, X_i) be described by Equations (6) and (19) under the historical probability $\mathbb{P}$. Fix $(\eta,\xi) \in \mathbb{R} \times (-\delta_i, \infty)$ and define:*

$$U_t := \sum_{i=1}^{n} \eta_i \sigma_i \int_0^t \int_0^{X_i(s)} W_i(ds,du) + \sum_{i=1}^{n} \int_0^t \int_0^{\lambda_i(s-)} \int_{\mathbb{R}^+} \left(e^{-\xi_i z} - 1\right) \tilde{J}_i(ds,du,dz).$$

Then the Doléans–Dade exponential $\mathcal{E}(U)$ is a martingale under $\mathbb{P}$ and the probability measure $\mathbb{Q}$ defined by $\frac{d\mathbb{Q}}{d\mathbb{P}}\Big|_{\mathcal{F}_t} := \mathcal{E}(U)_t$ is equivalent to $\mathbb{P}$. The dynamics with respect to $\mathbb{Q}$ takes the following form:

$$X_i(t,T) = X_i(0,T) + \int_0^t \int_0^{X_i(s,T)} \sigma_i^{\mathbb{Q}} W_i(du,ds) + \int_0^t \int_0^{X_i(s_-,T)} z\tilde{J}_i^{\mathbb{Q}}(dz,du,ds),$$
$$\lambda_i(t) = \lambda_i(0) - \int_0^t \beta_i^{\mathbb{Q}} \lambda(s)ds + \alpha_i^{\mathbb{Q}} \int_0^t \int_0^{\infty} \exp[-\beta^{\mathbb{Q}}(t-s)] J_i^{\mathbb{Q}}(dz,ds),$$

where

$$c_i^{\mathbb{Q}} = c_i^{\mathbb{P}} - \sigma_i^{\mathbb{P}} \eta_i - \int_0^{\infty} z(e^{-\theta_i z} - 1)\mu_i^{\mathbb{P}}(dz), \quad \sigma_i^{\mathbb{Q}} = \sigma_i^{\mathbb{P}},$$
$$\alpha_i^{\mathbb{Q}} = \alpha_i^{\mathbb{P}}, \quad \beta_i^{\mathbb{Q}} = \beta_i^{\mathbb{P}}, \quad \mu_i^{\mathbb{Q}}(dz) = e^{-\theta_i z}\mu_i^{\mathbb{P}}(dz).$$

Remark 6. In this context, the parameters η_i, ξ_i can be interpreted as the market price of risk associated with the diffusion/jump part of the i th factor X_i, respectively.

Remark 7. We shall assume, as for the previous model, that the de-seasonalized dynamics of X_i is a local martingale under $\mathbb{Q}$ and this will automatically imply that the mean reversion speed c_i of any X_i must vanish under $\mathbb{Q}$. Both the diffusion and the jump terms are in fact local martingales with respect to $\mathbb{Q}$.

Remark 8. From the formulas in the previous lines, specifying the relations between the model parameters under the risk-neutral measure $\mathbb{Q}$ and the historical measure $\mathbb{P}$, it is clear that in the Hawkes modeling framework, for each factor X_i, a mean reversion speed coefficient c_i can be nonzero under $\mathbb{P}$ and zero under $\mathbb{Q}$. A nonzero mean-reverting

term can then appear in the dynamics written with respect to the historical measure $\mathbb{P}$, although this term vanishes under $\mathbb{Q}$.

4 | THE FUTURES DYNAMICS

We focus here on forward contracts delivering a quantity of energy over a finite period of time. We shall refer to them as futures, even if in the literature they are sometimes called swaps or flow forwards.

Definition 2. The price at time $t \geq 0$ of a futures contract with delivery period $[T_1, T_2]$ with $t \leq T_1 \leq T_2$ is given by

$$F(t, T_1, T_2) = \frac{1}{T_2 - T_1} \int_{T_1}^{T_2} f(t, x)\, dx, \tag{20}$$

where $f(t, \cdot)$ is the price at time t of the forward contract to be paid upon delivery.

Remark 9. From Definition 2, it is clear why futures are sometimes called flow forwards: the owner of a futures with delivery period over $[T_1, T_2]$ would substantially receive a constant flow of the commodity over this period. Notice also that a futures contract delivering the commodity over a time period which collapses into a single point coincides with a forward.

The value at time t of a futures contract with delivery period $[T_1, T_2]$ is given, in our modeling framework, by (recall Equation 7):

$$F(t, T_1, T_2) = \frac{1}{T_2 - T_1} \int_{T_1}^{T_2} f(t, x) dx = \Lambda(t) - \Lambda(0) + \frac{1}{T_2 - T_1} \left(\sum_{i=1}^{n} \int_{T_1}^{T_2} X_i(t, x) dx \right). \tag{21}$$

By introducing the dynamics of the factors X_i into the above equation, we get the following equation describing the futures' dynamics under the risk-neutral probability $\mathbb{Q}$ both in the CBI framework (recall Equation 8):

$$\begin{aligned} F(t, T_1, T_2) = \Lambda(t) - \Lambda(0) &+ \frac{1}{T_2 - T_1} \int_{T_1}^{T_2} f(0, x) dx \\ &+ \frac{1}{T_2 - T_1} \sum_{i=1}^{n} \sigma_i \int_{T_1}^{T_2} \int_0^t \int_0^{X_i(s,x)} W_i(ds, dy) dx \\ &+ \frac{1}{T_2 - T_1} \sum_{i=1}^{n} \gamma_i \int_{T_1}^{T_2} \int_0^t \int_0^{X_i(s-,x)} \int_{\mathbb{R}_+} z \tilde{N}_i(ds, dy, dz) dx \end{aligned}$$

and in the Hawkes setting (recall Equation 15):

$$\begin{aligned} F(t, T_1, T_2) = \Lambda(t) - \Lambda(0) &+ \frac{1}{T_2 - T_1} \int_{T_1}^{T_2} f(0, x) dx \\ &+ \frac{1}{T_2 - T_1} \sum_{i=1}^{n} \sigma_i \int_{T_1}^{T_2} \int_0^t \sqrt{X_i(s, x)} dW_i(s) dx \\ &+ \frac{1}{T_2 - T_1} \sum_{i=1}^{n} \int_{T_1}^{T_2} \int_0^t \int_{\mathbb{R}_+} z \tilde{J}_i(dz, ds) dx. \end{aligned}$$

Assumption 1. From now on, in view of our numerical analysis, we will assume that one driving factor is sufficient. Namely, we will consider the case $n = 1$.

In order to rule out arbitrage opportunities the prices of futures with different delivery periods must satisfy specific time-consistency relations. In particular, the value of a futures contract with delivery period $[T_1, T_n]$ is linked to the values of the contracts with delivery on intervals $[T_i, T_{i+1}]$, $i = 1, \ldots, n-1$, where $[T_i, T_{i+1}]$ represents a partition of the interval $[T_1, T_n]$, by the following relation:

$$F(t, T_1, T_n) = \frac{1}{T_n - T_1} \sum_{i=1}^{n-1} (T_{i+1} - T_i) F(t, T_i, T_{i+1}). \tag{22}$$

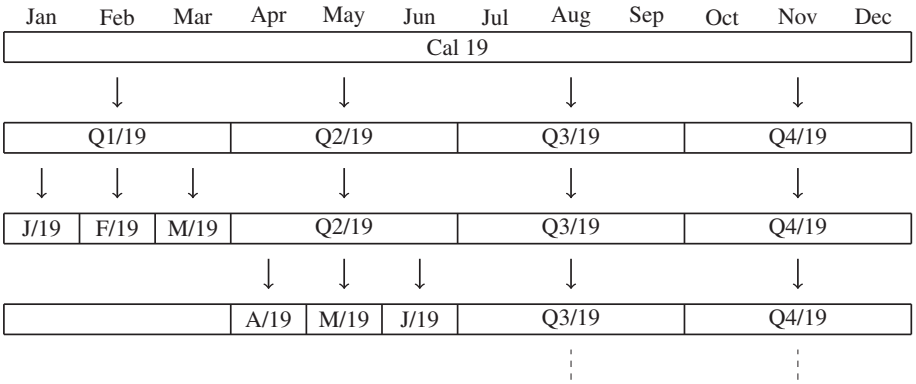


FIGURE 2 Cascade unpacking mechanism of futures contracts. For each given calendar year, as time passes by, futures are unpacked first in quarters, then in the corresponding months. It may happen that the same delivery period is covered by different contracts, for example, one simultaneously finds quotes for the monthly contracts J/19, F/19, M/19 and for the quarterly Q1/19

In Figure 2, we show the so called "cascade unpacking mechanism."

5 | DATA ANALYSIS: FROM FUTURES PRICES TO FORWARD CURVES

5.1 | Data set description

From a theoretical point of view, the contracts are settled continuously over the delivery period, as you can see from Equation (20), but in practice they are settled at discrete times. Assuming settlement at N points in time $u_1 < u_2 < \ldots < u_N$, with $u_1 = T_1$, $u_N = T_2$, then the discrete version of Equation (20) becomes:

$$F(t, T_1, T_2) = \frac{1}{T_2 - T_1} \sum_{i=1}^{N} f(t, u_i)\ \Delta_i,$$

where $\Delta_i = u_{i+1} - u_i$

The main goal of what follows is to provide a forward dynamics formulation starting from the futures prices that we observe in the market. We will build a smooth curve describing today's forward prices from quoted futures prices, according to the Heath–Jarrow–Morton framework outlined before.

This is a well-studied problem in literature and there are basically two ways to proceed: either fitting a parametric function to the entire yield curve by regression, or fitting all observed yields with a spline (see, e.g., Anderson and Deacon[31] for a survey on different methods for constructing yield curves). Here we opt for the latter approach, mainly because of its flexibility and easiness to implement. In addition, it can be used both for interpolation and smoothing, as we shall see in Section 5.2.

Throughout the article, we will also use the notation T_i^s, T_i^e to denote the first (start) and the last (end) day of the delivery period of the ith contract, and T^s, T^e in case of no ambiguity.

We are working on financially settled base load daily futures contract based upon the daily prices of the French power market. The term *base load* indicates that the delivery of electricity takes place for each hour of the day, in contrast to *peak load* contracts, that instead settle the delivery only for specified hours (from 8 to 20 typically). These contracts prescribe the delivery of 1 MW per day in the contract period (week, month, quarter, or year) and their price is in Euros and Euro cents per MWh.

More precisely, our data set consists of French futures closing prices downloaded from Thomson Reuters, that span over a period of 17 years, from 2002 to 2019. These contracts are divided with respect to the length of the delivery period into: weekly (tickers F7B1-B5), monthly (ticker F7BM), quarterly (ticker F7BQ), and yearly (ticker F7BY) contracts. Each of them presents four typologies of rolling contracts, namely, $c1$, $c2$, $c3$, and $c4$, where $c1$ and $c4$ are the ones with the closest and the farthest delivery period, respectively (Figure 3 in Section 5.3 shows an example to see how these rolling contracts work).

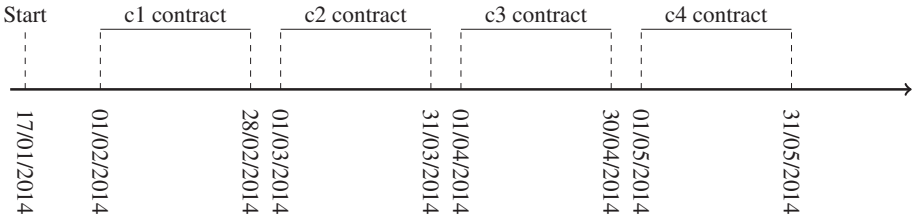


FIGURE 3 An example of how rolling contracts work, in the case when today, the "0" date, is January 17, 2014

On the market, we observe the quantity $F(0, T^s, T^e)$ for every contract, for different choices of T^s, T^e with $T^e - T^s =$ 7, 30, 90, and 365 days according to the weekly, monthly, quarterly, and yearly contracts, respectively, where "0" is the current date, the first available being July 1, 2002, while the last available being March 15, 2019, for a total number of 4234 trading days. More precisely, on each day, that we label as the "0" day, we have different number of contracts with different delivery periods, depending on the market data availability of that day. Our goal is to extract the curve $f(0, u)$ for all the different choices of the "0" date, as we are going to see in Section 5.2.

Notation 1. When possible from now on we will write $f(u)$ instead of $f(0, u)$ and $F(T_1, T_2)$ instead of $F(0, T_1, T_2)$ to shorten the notation.

5.2 | Extracting smooth forward curves from market data

Obtaining a smooth curve of forward prices from futures prices is a well-studied problem in the literature, see, for example, Fleten and Lemming.[32] The initial condition for using a Heath–Jarrow–Morton approach when modeling forwards is a smooth curve describing today's forward prices, which must be extracted from the futures prices observed in the market. We will follow the approach of Benth et al.[6(Chapter7)] by imposing the following

Assumption 2. The forward curve can be represented as the sum of two continuous functions $\Lambda(u)$ and $\varepsilon(u)$:

$$f(u) = \Lambda(u) - \Lambda(0) + \varepsilon(u), \quad u \in [T^s, T^e], \tag{23}$$

where T^s is the starting day of the settlement period for the contract with the closest delivery period and T^e is the last day of the settlement period for the contract with the farthest delivery period. We interpret $\Lambda(u)$ as a seasonality function and $\varepsilon(u)$ to be an adjustment function that captures the forward curve's deviation from the seasonality.

For the specification of the seasonality function, we follow the approach of Benth et al.,[6] namely, we define

$$\Lambda(u) = a \ \cos\left((u - b) \cdot \frac{2\pi}{365}\right). \tag{24}$$

The parameter $a \in \mathbb{R}_+$ is obtained by finding the minimum of the prices over all the contracts, while b is the normalized distance between the end of the last day of the year from the day when the minimum occurs.

This procedure leads to:

$$a = 13.600, \quad b = 1358.038. \tag{25}$$

There are several other methods for extracting the seasonality function from the data (see, e.g., Paraschiv[33] and Kiesel and Parashiv[11] for an application to hourly data), but since this topic is not the main focus of our study, we prefer to stick to the well-known method proposed by Benth et al.[34] and systematically described in Benth et al.[6(Chapter7)]

We shall see now how the adjustment function ε is obtained. We follow the approach proposed by Benth et al.,[6] in that we apply a maximum smoothness criterion to the adjustment function ε.

Remark 10. One may ask why the maximum smoothness criterion is applied only to the adjustment function ε and not to the entire forward function f . This ensures the presence of a seasonality pattern that, otherwise, would have possibly been smoothed out.

The properties we require for the adjustment function are that it is twice continuously differentiable and horizontal at time T_e, that is,

$$\varepsilon'(T^e) = 0. \tag{26}$$

This flatness condition is due to the fact that the long end of the curve may be several years ahead, and obviously the market's view on risk become less and less sensitive as time goes by.

Let us denote by $C_0^2([T^s, T^e])$ the set of real-valued functions on the interval $[T^s, T^e]$ which are twice continuously differentiable with zero derivative in T_e. We consider $\mathcal{C}$ as the set of polynomial spline functions of order four which belong to $C_0^2([T^s, T^e])$.

Definition 3. We define the smoothest possible forward curve on an interval $[T^s, T^e]$ as the function which minimizes, over $\mathcal{C}$, the integral

$$\int_{T^s}^{T^e} [\varepsilon''(u)]^2 du$$

and such that the closing prices matching condition holds (this is made precise in Equation A6).

We interpret the smoothest forward curve (23) to be the one for which ε solves the minimization problem above, with Λ chosen as in (24) and a, b as in (25).

In Appendix A, we provide a detailed description of the algorithm used for construction of forward curves from futures prices. In Benth et al.,[6] a comprehensive presentation of this and other approaches can be found.

5.3 | Numerical results

Recall that we are working with closing prices of French futures from 2002 to 2019. The yearly contracts span from 2002 to 2019, the quarterly from 2011 to 2019, the monthly from 2011 to 2019, and the weekly from 2010 to 2019. We are working with *rolling contracts*, called $c1$, $c2$, $c3$, and $c4$. In Figure 3, an example which shows how these contracts roll for a monthly contract is provided.

As you can see from Figure 3, the $c1$ contract is the closer one to the current date and its delivery period spans from 01/02/2014 to 28/02/2014. After 28/02/2014, there is a rollover from the $c1$ contract to the $c2$ contract and so on.

Figure 4 shows the plot of the futures closing prices for the $c1$ weekly, monthly, quarterly and yearly contracts. In the x-axis there are the different dates, while in the y-axis there is the price. As you can see the presence of seasonality is pretty strong, especially in the monthly and quarterly contracts.

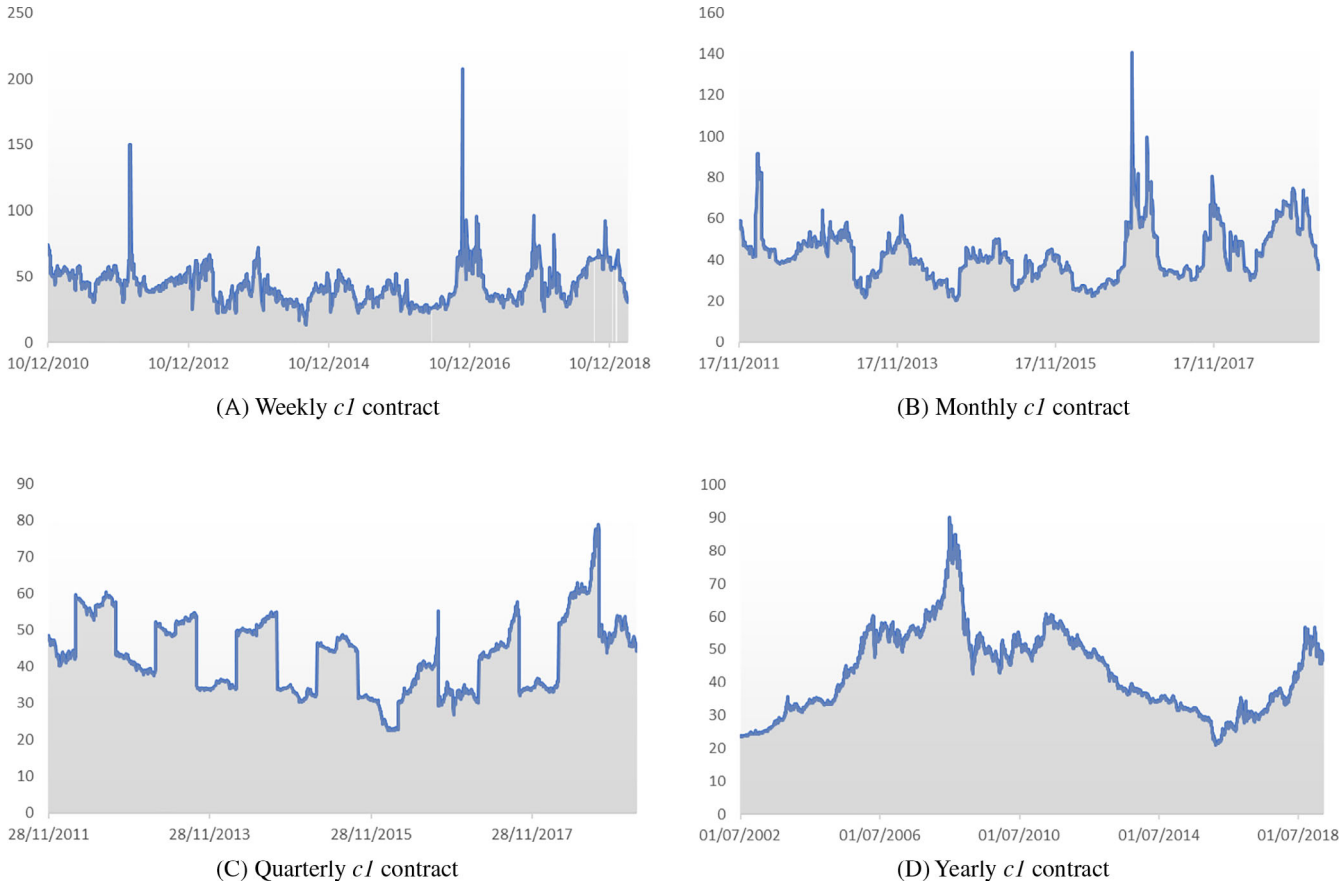
Since we were worried that our analysis could have been affected by the presence of the quarterly contracts being more sensible to the seasonality pattern, we performed the analysis in both cases, with and without the quarterly contracts. We noticed that the results were coherent in both cases, so from now on we will focus only on contracts different from quarterly.

The algorithm presented takes approximately 40 s to extract the 4234 different curves, that is, the curves $f(0, u)$, for all the 4234 different values of "0."

Figure 5 shows, for every day from 2002 to 2019, a subset of 599 de-seasonalized forward curves. In the x-axis we have the time to maturity while in the y-axis we have the prices. Note that different colors in the curves mean different type of contracts. Note also that the further we move on the x-axis, the flatter the curves become. This is in line with the flatness constraint in Equation (26). The reason behind the difference in the shapes of the curves is to be investigated in the price constraint but also in the nature of the contracts, since for each day the number and the type of available contracts were different, having to deal also with overlapping settlement periods.

6 | JUMP DETECTION

We now want to detect the jumps. We will be only dealing with positive jumps since we have supposed that the jump size is distributed according to an exponential density, as already described in Section 3. In the next subsection, we will describe an algorithm that allows to detect jumps, and, as a by-product, which also gives their size.



(A) Weekly *c1* contract

(B) Monthly *c1* contract

(C) Quarterly *c1* contract

(D) Yearly *c1* contract

FIGURE 4 Prices in Euros of the contracts *c*1, starting from December 10, 2010 and ending on March 14, 2019

6.1 | Description of the algorithm

In order to detect jumps, we proceed in the following way: for a fixed maturity T, we define

$$V_t = f(t, T),$$

the vertical section at maturity T, where the parameter t ranges through all the curves, that is, $t = 1, \dots, 4234$.

Roughly speaking, looking at Figure 5, this is nothing but the intersection between the vertical line $x = T$ and the curves. There are several ways to detect jumps from the data, the more natural one consisting in fixing a threshold $\Theta \in \mathbb{R}_+$ and saying that a jump occurs at time $\tilde{t}$ if $|V_{\tilde{t}+1} - V_{\tilde{t}}| \geq \Theta$. We follow here an iterative weighted least square approach. Define $n = 4234$ (the total number of curves) and $\mathcal{N} = \{1, 2, \dots, n-1\}$. The algorithm to detect jumps reads as follows:

Clearly, as the number of iterations increases, σ_i^2 decreases and so does the number of jumps detected. After several tests on the data, we noticed that stopping at $k \in \mathbb{N}$ such that $m_k = 0$ would lead to too many jumps, the last detected being much smaller compared to the ones discovered at the first iterations. Indeed, as you can see from Table 1, σ_2 is about three times σ_3 and this leads to an unreasonable number of small size's jumps. Therefore, we chose to stop the algorithm after the first two iterations.

6.2 | Jumps analysis

We selected different values of T, namely, $T = 200$, $T = 400$, and $T = 700$ days. This covers all the different shapes of the forward curves and so it represents a good sampling of our data. After applying the algorithm described in Section 6.1, we end up with Figure 6 showing the size and distribution of the jumps detected at $T = 200$, $T = 400$, and $T = 700$. Jumps are highlighted by a vertical orange line, the corresponding price being represented by the continuous blue line.

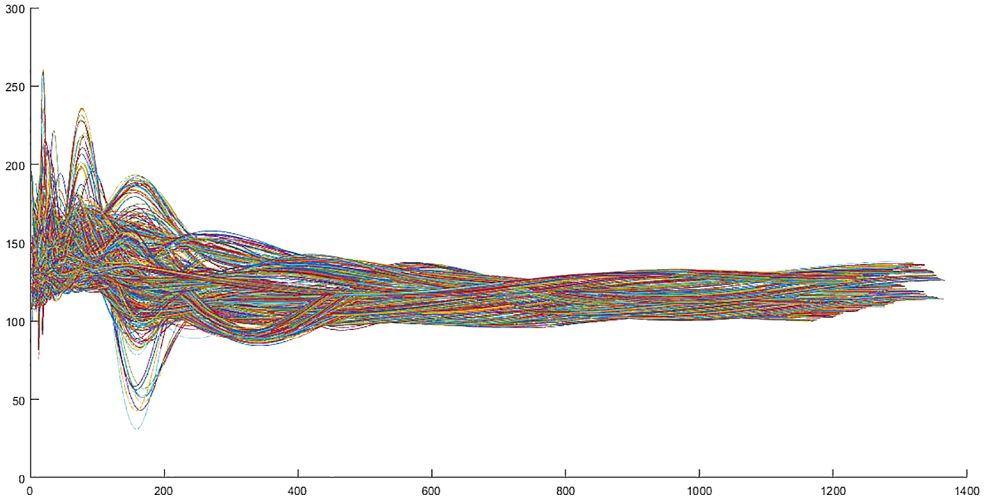


FIGURE 5 A subset of 599 de-seasonalized forward curves extracted without considering quarterly contracts. In the x-axis there is the time to maturity while in the y-axis there are the prices in Euros

Algorithm 1. Iterative weighted least square algorithm for jumps detection

Initialization
Set $\sigma_1^2 = \frac{1}{n-2}\sum_{t\in\mathcal{N}} \frac{(V_{t+1}-V_t)^2}{V_t}$;
Identify all $t \in \mathcal{M}_1 \subseteq \mathcal{N}$ such that $\frac{V_{t+1}-V_t}{\sqrt{V_t}} \geq 3\sigma_1$;
Set $m_1 = |\mathcal{M}_1|$;
Set $i = 1$;
while $m_i = |\mathcal{M}_i| \neq 0$ **do**
  $i \leftarrow i + 1$;
  $\sigma_i^2 = \frac{1}{n-(\sum_{j=1}^{i-1} m_j)-1}\sum_{t\in\mathcal{N}\setminus(\bigcup_{j=1}^{i-1}\mathcal{M}_j)} \frac{(V_{t+1}-V_t)^2}{V_t}$;
  Identify all $t \in \mathcal{M}_i \subseteq \mathcal{N} \setminus (\bigcup_{j=1}^{i-1}\mathcal{M}_j)$ such that $\frac{V_{t+1}-V_t}{\sqrt{V_t}} \geq 3\sigma_i$;
  Set $m_i = |\mathcal{M}_i|$.
end

TABLE 1 Values of $\sigma_i, i = 1, 2, 3$ at different maturities

T	200	400	700
σ_1	0.4581	0.2743	0.0909
σ_2	0.2187	0.1295	0.0664
σ_3	0.0712	0.0212	∼0

Table 2 shows the number of jumps detected at each iteration, for different value of T.

As one can see from Figure 6, at $T = 200$ the jumps detected are the bigger ones with respect to their amplitude, and this is not surprising looking at Figure 5, where one can clearly see that the price movements are significant at $T = 200$. On the other hand, when $T = 700$, the jumps detected are quite small and this is due to the fact that at $T = 700$ the curves are flattened, leading to prices which are close to each other.

7 | PARAMETERS ESTIMATION

Before starting with the statistical tests on the two models, we still have to estimate: the size of the jumps and the parameters characterizing the drift and the volatility coefficients. We start with δ, the size of the jumps. Recall that in

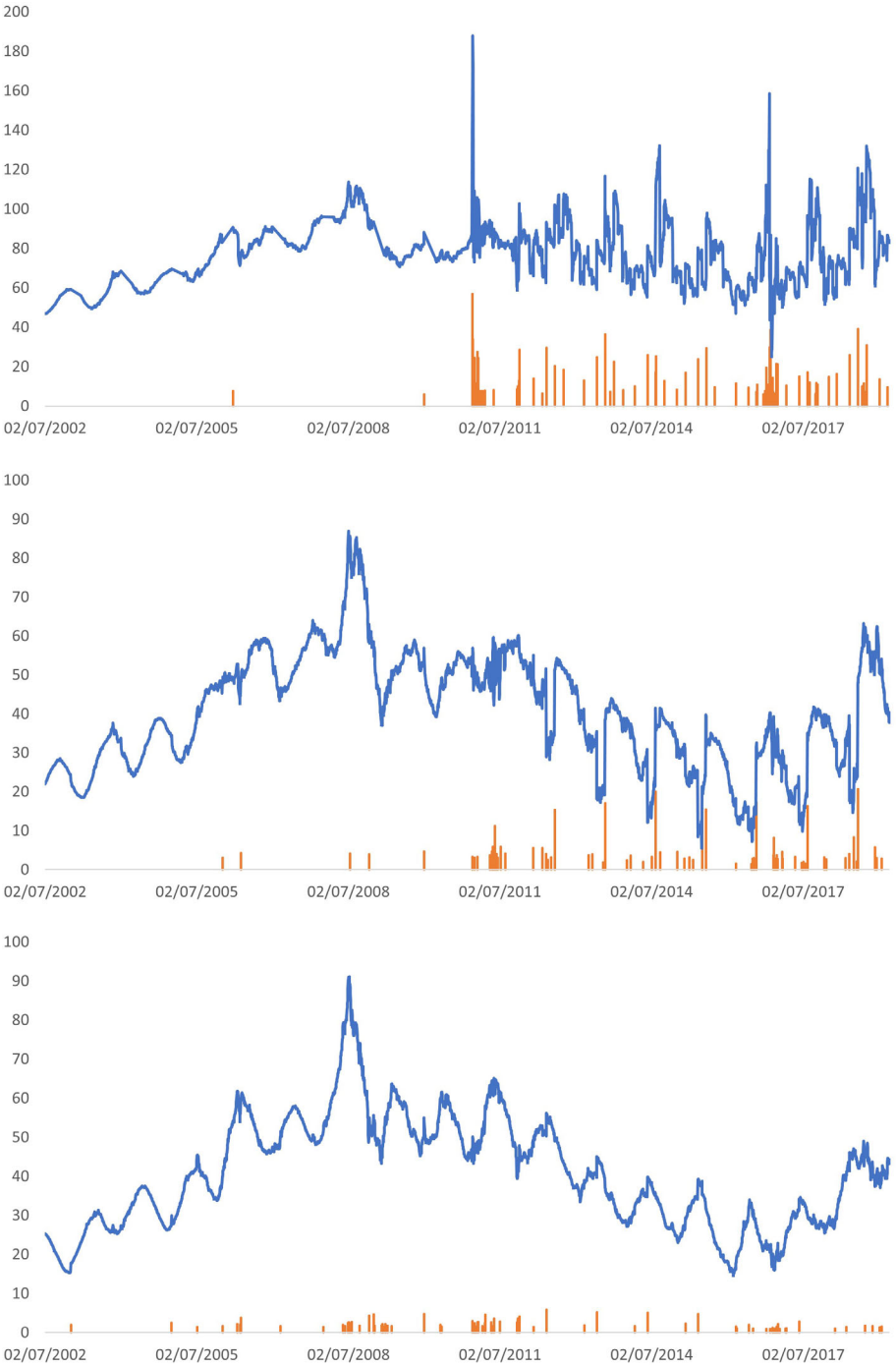


FIGURE 6 Jumps detected (vertical orange lines) and price plot (blue line) at $T = 200$, $T = 400$, and $T = 700$, from the top to the bottom, respectively

Sections 3.2 and 3.1, we assumed for both models the size of the jumps to be distributed like an exponential random variable with parameter $\delta > 0$. Let z_i be the size of the ith jump, where $i = 1, \ldots, L$ and L is the number of jumps at the chosen maturity T (see Table 2). Then δ can be estimated, for example, via its maximum likelihood estimator:

$$\hat{\delta} = \frac{L}{\sum_{i=1}^{n} z_i}. \tag{27}$$

We obtain what follows (the fact that we have the smallest values of $\hat{\delta}$ at $T = 200$ is not surprising at all, because at the beginning jumps are bigger, as said before).

TABLE 2 Number of jumps detected at different maturities and at different iterations

T	200	400	700
First iteration	38	19	43
Second iteration	48	55	36
Total	86	74	79

TABLE 3 Parameter estimation for δ

T	200	400	700
$\hat{\delta}$	0.064282682	0.194300598	0.430239733

TABLE 4 Parameters estimation for $\tilde{a}$

T	200	400	700
$\hat{\tilde{a}}$	−0.001387342	−0.001771845	−0.000237952

TABLE 5 Parameter estimation for σ

T	200	400	700
$\hat{\sigma}$	0.218667945	0.129560813	0.066361067

We now need to estimate the parameter appearing in the drift coefficient of our forward dynamics, namely, the parameter a_1 in Equation (7) and c_1 in Equation (15). For simplicity, in this section, we denote both a_1 and c_a by $\tilde{a}$. This drift term $\tilde{a}$ can be seen as a mean reversion speed of the stochastic factor toward the null mean reversion level (see Table 3 below).

Remark 11. The dynamics under the historical measure $\mathbb{P}$, both for the CBI (see Equation 9) and the Hawkes settings (see Equation 15), in discrete form, reads, for a fixed T:

$$\sim X(t+1,T) = \sim X(t,T) - \int_t^{t+1} \sim a \sim X(s,T)ds,$$

where $\sim X$ is the unique factor in the forward dynamics without the seasonality and with no jumps.

In order to estimate $\sim a$, we simply use Remark 11 and we obtain the following estimation

$$\hat{\tilde{a}} := 1 - \frac{\sim X(t+1,T)}{\sim X(t,T)}. \tag{28}$$

We find the following estimates:

As you can see from Table 4, the estimated value of $\sim a$ in all the three cases is really small.

Now it remains to estimate the volatility parameter σ, appearing in both Equations (9) and (15). By recalling the iterative algorithm presented in Section 6.1 and taking as $\hat{\sigma}$ the value of σ_2 (viz., the estimation after the second iteration), we get Table 5.

8 | TESTING THE MODELS

In this section, we want to perform statistical tests concerning the intensity of the jumps. We want to check what is the best process modeling the jumps we have detected before. We test the two models based on Hawkes and branching processes, plus the Poisson, which is a toy-model:

(0) Poisson process;
(1) Hawkes process;
(2) Self exciting branching process.

TABLE 6 Parameters estimation for the three models at $T = 200$

	Parameters				
Model	$\lambda(0)$	α	β	λ^P	γ
Poisson	–	–	–	0.023	–
Hawkes	0.017	0.074	0.094	–	–
Branching	–	–	–	–	0.00028

We will mainly rely on the Kolmogorov–Smirnov (KS) test, namely, we will test the null hypothesis H_0, stating that the data have the same cumulative distribution function as the one coming from one of the above models, against the alternative H_1. We fix a significance level equal to 0.05.

8.1 | Jump intensity estimation

Before using the KS test to check whether the jumps distribution comes from one of the three models, we need to estimate the intensity from our data. The input in all the cases will be the time occurrences of the jumps over $[0, T]$ (for the three different values of T), $0 < \tau_1 < \tau_2 < \ldots < \tau_N = T$, where N can take the values 86, 74, and 79 depending on the chosen maturity T, as you can check from Table 2.

(0) [Poisson] The (constant) intensity, $\lambda^P > 0$, is estimated as the ratio between the total number of (positive) jumps and the sum of the inter-times between two consecutive jumps.
(1) [Hawkes] Here the intensity is given by Equation (16), so this case will be treated in a separate subsection.
(2) [Branching] In this case, the stochastic intensity $\lambda^B(t) \propto X(t, T)$ Jiao et al.[15(Section5.1,p152)] and the constant of proportionality γ is estimated, for a fixed T, as the ratio between the total number of (positive) jumps and the cumulative (de-seasonalized) forward prices.

8.1.1 | The Hawkes setting: Estimating λ

Recalling Equation (16), it is clear that we have to estimate three parameters: $\lambda(0)$, α, and β. We mainly rely on the paper by Ozaki[35] and we will find a maximum likelihood estimation (MLE).

The log-likelihood of a Hawkes process whose response function is of the form $\alpha e^{-\beta t}$, is given by

$$\log L(\tau_1, \ldots, \tau_N) = -\lambda(0)\tau_N + \sum_{i=1}^{N} \frac{\alpha}{\beta}\left(e^{-\beta(\tau_N - \tau_i)} - 1\right) + \sum_{i=1}^{N} \log\left(\lambda(0) + \alpha A(i)\right), \tag{29}$$

where $A(i) = \sum_{\tau_j < \tau_i} e^{-\beta(\tau_i - \tau_j)}$ for $i \geq 2$ and $A(1) = 0$.

In order to estimate the parameters $\lambda(0), \alpha, \beta$, we need to find the maximum of the function in (29), which is a real value function of three variables. The maximum was found using the command *fminsearch* of MATLAB.

Tables 6–8 show the parameters estimated for the jump intensity of the three different models at the maturities $T = 200$, $T = 400$, and $T = 700$. As far as the branching model is concerned, $\lambda^B(t)$ is proportional to the process $X(t, T)$ (for a fixed T), and the parameter to be estimated is the γ, which is the constant ratio between the two processes.

8.2 | KS test for the models

We perform a KS test in order to check which of the proposed distributions best models the jumps in our data. At the end of the subsection, we will provide the p-values to conclude.

TABLE 7 Parameters estimation for the three models at $T = 400$

	Parameters				
Model	$\lambda(0)$	α	β	λ^P	γ
Poisson	–	–	–	0.018	–
Hawkes	0.0026	0.012	0.016	–	–
Branching	–	–	–	–	0.00021

TABLE 8 Parameters estimation for the three models at $T = 700$

	Parameters				
Model	$\lambda(0)$	α	β	λ^P	γ
Poisson	–	–	–	0.0019	–
Hawkes	0.040	0.059	0.085	–	–
Branching	–	–	–	–	0.00032

(0) [Poisson] We check whether the jumps inter-times are drawn from an exponential distribution with parameter λ^P, where λ^P is the one given in Tables 6–8. The p-value is automatically given by MATLAB via the function *kstest*.
(1) [Hawkes] We mainly adapt the methods in Lallouache and Challet[36] to our purpose. In particular, we check if the time-deformed series of durations $\{\theta_i\}_{i=1,\dots,N}$, defined by

$$\theta_i = \int_{\tau_{i-1}}^{\tau_i} \widehat{\lambda}_t dt, \tag{30}$$

has an exponential distribution of parameter 1, where $\widehat{\lambda}_t$ is the intensity estimated before (the estimated parameters can be found in Tables 6–8), and where recall that the τ_i's are the jumps arrival times. The p-value is automatically given by MATLAB via the function *kstest*.
(2) [Branching] The procedure here is quite different from the ones adopted before and it is the object of the following subsection.

8.2.1 | Setting the KS test for the branching model

The KS test we will perform in this case was constructed based on the following classical result (see, e.g., Cox and Lewis[37(Section3.3)] or, for the proof, Vedyuschenko[38(Thm5)]).

Proposition 3. *Let $(N_t)_{t\geq 0}$ be a nonhomogeneous Poisson process with continuous expectation function. If n events have occurred in $(0, T]$, then the arrival times $\tau_1, \dots, \tau_n$ are distributed as the order statistics from a sample with cumulative distribution function*

$$F(t) = \frac{\int_0^t \lambda^B(s)ds}{\int_0^T \lambda^B(s)ds}, \quad 0 \leq t \leq T, \tag{31}$$

where in our case $\lambda^B(s) \propto f(s, T)$, for a fixed T.

Remark 12. Recall that $\lambda^B(s) \propto f(s, T)$, for any fixed T. It is crucial to notice, from Proposition 3, that the distribution of the arrival times is independent of the factor of proportionality connecting λ^B and $f(\cdot, T)$.

So in this case we perform a KS test, comparing the cumulative distribution function $F(t)$ in Equation (31) with the empirical one relative to the jump times.

TABLE 9 Maximum distance between the empirical distribution function and the theoretical one for the branching case

T	200	400	700
D_n	0.2151	0.2276	0.2513

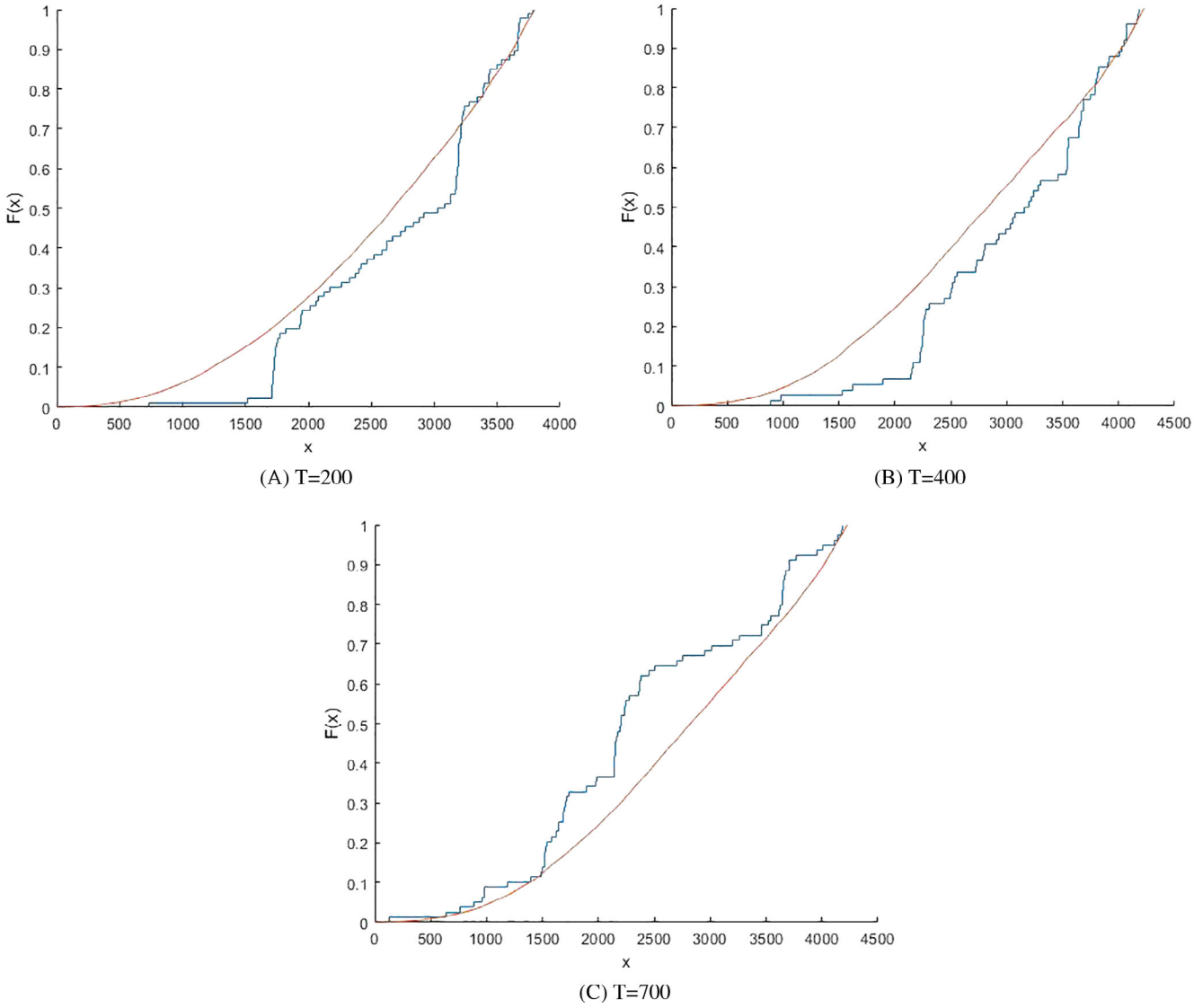


FIGURE 7 Comparison between the cumulative distribution function of F and the empirical cumulative distribution function of the jumps arrival times

Since $F(t)$ given in Proposition (3) is not a priori associated to a known distribution, we cannot use the MATLAB command kstest and we have to rely on the classical theory on the KS test. The KS statistics for the test is

$$D_n = \sup_{x \in \mathbb{R}} |S_n(x) - F(x)|, \tag{32}$$

where n is the number of our data (recall Table 2), F is the cumulative distribution function in Equation (31) and $S_n(x)$ is the empirical cumulative distribution function of the jump arrival times. We find in Table 9.

In Figure 7, we graphically compare the two cumulative distribution functions at the three different maturities.

TABLE 10 *p*-value test for the three models for different *T*

Values of *T*	200	400	700
Poisson	∼0	∼0	∼0
Branching	0.041	0.018	0.042
Hawkes	0.23	0.31	0.13

In order to obtain the *p*-value in the branching case, we apply the asymptotic results in Facchinetti[39] in the case when the dataset is greater than 35. We provide the *p*-values in Table 10.

As one can see from Table 10, the hypothesis that the intensity follows a Poisson process is highly rejected, as we expected. In the branching case, the hypothesis is also rejected, even if the *p*-value in this case was much closer to the acceptance level of 0.05. For the Hawkes case, the test fails to reject the hypothesis since all the three values are above our level of acceptance.

As a conclusion we can resume the main achievement presented in the present article. We proposed two alternative models for power forward prices evolution, based on a HJM approach, extending to forward prices dynamics two models already proposed for the spot price dynamics in Eyjolfsson and Tjøshteim[16] and in Jiao et al..[15] After extracting forward curves from quoted futures prices, we proposed a parameters estimation method for both models and then we performed a test on the adequacy of the two models in describing the observed forward prices evolution. The final conclusion of our test is that the hypothesis that forward prices follow a CBI-type dynamics is rejected, while the hypothesis of a Hawkes type dynamics is not. This conclusion suggests that self-exciting effects can arise in power forward dynamics as well as in the spot dynamics, and that an approach based on Hawkes processes can capture these effects in a natural and parsimonious way.

Futures perspectives of the present work include the investigation, from a numerical point of view, of the multivariate setting. Numerical extensions and calibration in a multivariate framework can be quite complicated: a two-dimensional case might be a reasonable starting point. Simulations and Monte Carlo schemes are mandatory to price complex derivatives in power markets. So, a deeper analysis of Monte Carlo simulations techniques and relative pricing schemes in these complex settings is the subject of a future research program.

ACKNOWLEDGMENTS

We thank S. Scotti and R. Brignone for fruitful discussions in preparing this article. We are grateful to the associate editor and two anonymous referees for helpful remarks that allowed to improve the original version of the article. Open Access Funding provided by Politecnico di Milano within the CRUI–CARE Agreement.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

ORCID

Carlo Sgarra https://orcid.org/0000-0001-9790-5292

REFERENCES

1. Kovacevic RM. Valuation and pricing of electricity delivery contracts: the producer's view. *Ann Oper Res*. 2019;275:421-460.
2. Alves MJ, Antunes CH. A semivectorial bilevel programming approach to optimize electricity dynamic time-of-use retail pricing. *Comput Oper Res*. 2018;92:130-144.
3. Lucheroni C, Mari C. Risk-shaping of optimal electricity portfolios in the stochastic LCOE theory. *Comput Oper Res*. 2018;96:374-385.
4. Wahab MIM, Lee C-G. Pricing swing options with regime switching. *Ann Oper Res*. 2011;185:139-160.
5. Schwarz ES. The stochastic behaviour of commodity prices: implications for valuation and hedging. *J Financ*. 1997;52(3):923-973.
6. Benth FE, Salthyte-Benth J, Koekebakker S. *Stochastic Modelling of Electricity and Related Markets*. World Scientific; 2008.
7. Benth FE, Piccirilli M, Vargiolu T. Mean-reverting additive energy forward curves in a Heath-Jarrow-Morton framework. *Math Finan Econ*. 2019;13(4):543-577.
8. Filimonov V, Bicchetti D, Maystre N, Sornette D. Quantification of the high level of endogeneity and structural regime shifts in commodity markets. *J Int Money Financ*. 2014;42(C):174-192.

9. Heath D, Jarrow R, Morton A. Bond pricing and the term structure of interest rates: a new methodology for contingent claim valuation. *Econometrica*. 1992;60(1):77-105.
10. Benth FE, Paraschiv F. A space-time random field model for electricity forward prices. *J Bank Financ*. 2018;95:203-216.
11. Kiesel R, Paraschiv F. Econometric analysis of 15-minutes intraday electricity prices. *Energy Econ*. 2017;64:77-90.
12. Herrera R, Gonzalez N. The modeling and forecasting of extreme events in electricity spot markets. *Int J Forecast*. 2014;30(3):477-490.
13. Christensen TM, Hurn AS, Lindsay KA. It never rains, but it pours: modelling the persistence of spikes in electricity markets. *Energy J*. 2009;30(1):25-48.
14. Clements A, Fuller J, Hurn AS. Semi-parametric forecasting of spikes in electricity prices. *Econ Rec*. 2013;89(287):508-521.
15. Jiao Y, Ma C, Scotti S, Sgarra C. A branching process approach to power markets. *Energy Econ*. 2019;79:144-156.
16. Eyjolfsson H, Tjøshteim D. Self-exciting jump processes with applications to energy markets. *Ann Inst Stat Math*. 2018;70(2):373-393.
17. Dawson A, Li Z. Stochastic equations, flows and measure-valued processes. *Ann Probab*. 2012;40(2):813-857.
18. Li Z, Ma C. Asymptotic properties of estimators in a stable Cox-Ingersoll-Ross model. *Stoch Process Appl*. 2015;125(8):3196-3233.
19. Walsh J. *An Introduction to Stochastic Partial Differential Equations*. Ecole d'été de Probabilités de Saint-Flour XIV-1984, Lecture Notes in Mathematics. Vol 1180. Springer; 1980:265-430.
20. Pardoux E. *Probabilistic Models of Population Evolution*. Springer; 2016.
21. Kawazu K, Watanabe S. Branching processes with immigration and related limit theorems. *Theory Probab Appl Ther*. 1971;16(1):36-54.
22. Filipović D. A general characterization of one factor affine term structure models. *Finance Stochast*. 2001;5(3):389-412.
23. Bernis G, Salhi K, Scotti S. Sensitivity analysis for marked Hawkes processes: application to CLO pricing. *Math Financ Econ*. 2018;12(4):541-559.
24. Errais E, Giesecke K, Goldberg LR. Affine point processes and portfolio credit risk. *SIAM J Financ Math*. 2010;1(1):642-665.
25. Hainaut D, Moraux F. A switching self-exciting jump diffusion process for stock prices. *Ann Finance*. 2019;15(2):267-306.
26. Rambaldi Q, Pennesi X, Lillo F. Modeling foreign exchange market activity around macroeconomic news: Hawkes process approach. *Phys Rev E*. 2015;91(1):012819.
27. Jiao Y, Ma C, Scotti S. Alpha-CIR model with branching processes in sovereign interest rate modeling. *Finance Stochast*. 2017;71:789-813.
28. Caballero ME, Garmendia JLP, Uribe Bravo G. A Lamperti-type representation of continuous-state branching processes with immigration. *Ann Probab*. 2013;41(3A):1585-1627.
29. Dassios A, Zhao H. Exact simulation of Hawkes process with exponentially decaying intensity. *Electron Commun Probab*. 2012;18:1-13.
30. Bernis G, Scotti S, Sgarra C. A Gamma Ornstein-Uhlenbeck model driven by a Hawkes process. *Math Finan Econ*. 2021;15:747-773.
31. Anderson FB, Deacon M. *Estimating and Interpreting the Yield Curve*. John Wiley & Sons; 1996.
32. Fleten SE, Lemming J. Constructing forward price curves in electricity markets. *Energy Econ*. 2003;25(5):409-424.
33. Paraschiv F. Price dynamics in electricity markets. *Handbook of Risk Management in Energy Production and Trading*. Boston, MA: Springer; 2013:47-69.
34. Benth FE, Koekebakker S, Ollmar F. Extracting and applying smooth forward curves from average-based commodity contracts with seasonal variation. *J Deriv*. 2007;15(1):52-66.
35. Ozaki T. Maximum likelihood estimation of hawkes self-exciting point processes. *Ann Inst Stat Math*. 1979;31(1):145-155.
36. Lallouache M, Challet D. The limits of statistical significance of Hawkes processes fitted to financial data. *Quant Finance*. 2016;16(1):1-11.
37. Cox DR, Lewis PAW. *The Statistical Analysis of Series of Events, Methuen's Monographs on Applied Probability and Statistics*. M.S. Bartlett F.R.S; 1966.
38. Vedyushenko A. *Non-homogeneous Poisson Process - Estimation and Simulation*. Master thesis. Charles University; 2018.
39. Facchinetti S. A procedure to find exact critical values of Kolmogorov-Smirnov test. *Italian J Appl Stat*. 2009;21(3-4):337-359.

How to cite this article: Callegaro G, Mazzoran A, Sgarra C. A self-exciting modeling framework for forward prices in power markets. *Appl Stochastic Models Bus Ind*. 2021;1–22. https://doi.org/10.1002/asmb.2645

APPENDIX A. SMOOTH FORWARD CURVE CONSTRAINED BY CLOSING PRICES

In this section, we present the general procedure to extract the forward dynamics in a general situation with a fixed number of contracts from the market, but we will often make references to our own case. Before presenting the algorithm we need to introduce a procedure in order to deal with overlapping periods. Let

$$\mathcal{T} = \left\{(T_1^b, T_1^e), \ \dots \ , (T_m^b, T_m^e)\right\}$$

be a list of start and end dates for the settlement periods of m different futures contracts for a given day (in our case, $m = 16$). We need to be able to handle the problem of overlapping settlement periods to rule out arbitrage opportunities. This was a concrete issue working with our data because it happens that, in a given day, two or more contracts have

delivery periods that intersect. To overcome this, we construct a new list of dates $\tilde{\mathcal{T}}$, namely

$$\tilde{\mathcal{T}} = \{T_0, T_1, \dots, T_n\},$$

where overlapping contracts are split into sub-periods. In our case n is typically 24, T_0 denotes the starting day of the contract with the closest delivery period, while T_n denotes the last day of the contract with the farthest delivery period. The procedure is illustrated in Figure A1.

As we can see from Figure A1, the elements of this new list are basically the elements in $\mathcal{T}$ sorted in ascending order, with duplicate dates removed. The futures prices could be taken into account either by exact matching or by a constraint on the bid-ask spread prices. Dealing with closing prices, here we impose an exact matching prices on closing prices (see Equation A6). From now on we denote by F_i^C the closing price for the future $i, i \in \{1, \dots, m\}$.

The adjustment functions ε is chosen in the class C, namely (with a slight abuse of notation we use $\varepsilon(u; \boldsymbol{x})$ instead of $\varepsilon(u)$ to stress the dependence on $\boldsymbol{x}$)

$$\varepsilon(u; \boldsymbol{x}) = \begin{cases} a_1 u^4 + b_1 u^3 + c_1 u^2 + d_1 u + e_1, & u \in [T_0, T_1], \\ a_2 u^4 + b_2 u^3 + c_2 u^2 + d_2 u + e_2, & u \in [T_1, T_2], \\ \vdots & \\ a_n u^4 + b_n u^3 + c_n u^2 + d_n u + e_n, & u \in [T_{n-1}, T_n], \end{cases}$$

where $\boldsymbol{x}' = [a_1, b_1, c_1, d_1, e_1, \dots, a_n, b_n, c_n, d_n, e_n]$ is the row vector of the coefficients of the splines that we want to find.

In this way we have, roughly speaking, a spline for every settlement period. To find the unknown parameters $\boldsymbol{x}' = [a_1, b_1, c_1, d_1, e_1, \dots, a_n, b_n, c_n, d_n, e_n]$ in order to fully recover the adjustment function, we need to solve the following equality constrained convex quadratic programming problem:

$$\min_{\boldsymbol{x} \in \mathbb{R}^{5n}} \int_{T_0}^{T_n} [\varepsilon''(u; \boldsymbol{x})]^2 du, \tag{A1}$$

subject to the following constraints:

(i) continuity of the derivatives up to second order at the knots, for $j = 1, \dots, n-1$,

$$a_{j+1} T_j^4 + b_{j+1} T_j^3 + c_{j+1} T_j^2 + d_{j+1} T_j + e_{j+1} = a_j T_j^4 + b_j T_j^3 + c_j T_j^2 + d_j T_j + e_j, \tag{A2}$$

$$4a_{j+1} T_j^3 + 3b_{j+1} T_j^2 + 2c_{j+1} T_j + d_{j+1} = 4a_j T_j^3 + 3b_j T_j^2 + 2c_j T_j + d_j, \tag{A3}$$

$$12a_{j+1} T_j^2 + 6b_{j+1} T_j + 2c_{j+1} = 12a_j T_j^2 + 6b_j T_j + 2c_j, \tag{A4}$$

(ii) flatness at the end (see Equation 26)

$$\varepsilon'(T_n; \boldsymbol{x}) = 0, \tag{A5}$$

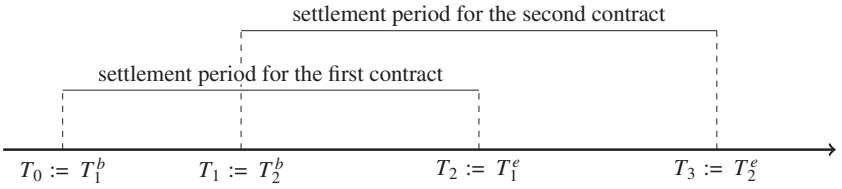


FIGURE A1 Dealing with overlapping delivery time windows

(iii) matching of the closing prices (see Equation 20), for $i = 1, \dots, m$,

$$F_i^C = \frac{1}{T_i^e - T_i^b} \int_{T_i^b}^{T_i^e} [\Lambda(u) + \varepsilon(u; \boldsymbol{x})]\, du. \tag{A6}$$

By this way, the minimization problem (A1) has a total of $3n + m - 2$ constraints (i.e., $3(n-1)$ constraints from (A2)–(A4), one constraint from (A5), and m constraints from (A6)). By computing the second derivative of ε and inserting it in Equation (A1) and integrating for every delivery period, we can rewrite the minimization problem (A1) as

$$\min_{\boldsymbol{x} \in \mathbb{R}^{5n}} \boldsymbol{x}' \mathbf{H} \boldsymbol{x}, \tag{A7}$$

where

$$\mathbf{H} = \begin{bmatrix} h_1 & \dots & 0 \\ & \ddots & \\ 0 & \dots & h_n \end{bmatrix} \quad \text{with} \quad h_j = \begin{bmatrix} \frac{144}{5}\Delta_j^5 & 18\Delta_j^4 & 8\Delta_j^3 & 0 & 0 \\ 18\Delta_j^4 & 12\Delta_j^3 & 6\Delta_j^2 & 0 & 0 \\ 8\Delta_j^3 & 6\Delta_j^2 & 4\Delta_j^1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{A8}$$

and

$$\Delta_j^l = T_j^l - T_{j-1}^l, \tag{A9}$$

for $j = 1, \dots, n$, and $l = 1, \dots, 5$.

We clearly see that the constraints (A2)–(A6) are linear w.r.t. $\boldsymbol{x}$, and so they can be formulated in a matrix form as $\mathbf{A}\boldsymbol{x} = \mathbf{b}$, where $\mathbf{A}$ is a $(3n + m - 2) \times 5n$-dimensional matrix, and $\mathbf{b}$ is a $(3n + m - 2)$-dimensional vector. Solving the problem (A1) with the constraints (A2)–(A6) is equivalent to solving (A7) with the constraints written in the form $\mathbf{A}\boldsymbol{x} = \mathbf{b}$. Let $\lambda' = [\lambda_1, \lambda_2, \dots, \lambda_{3n+m-2}]$ be the corresponding Lagrange multiplier vector to the constraints (A2)–(A6). So, we can now express (A1) as the following unconstrained minimization problem

$$\min_{\boldsymbol{x} \in \mathbb{R}^{5n}, \lambda \in \mathbb{R}^{3n+m-2}} \boldsymbol{x}' \mathbf{H} \boldsymbol{x} + \lambda' (\mathbf{A}\boldsymbol{x} - \mathbf{b}). \tag{A10}$$

Remark 13. The advantage of dealing with problem (A10), instead of (A1) with the constraints (A2)–(A6), is that (A10) is a unconstrained problem that can be simply solved. Indeed the solution $[\overline{\mathbf{x}}, \overline{\lambda}]$ is obtained just solving the linear system

$$\begin{bmatrix} 2\mathbf{H} & \mathbf{A}' \\ \mathbf{A} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{x} \\ \lambda \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{b} \end{bmatrix}. \tag{A11}$$

The dimension of the left matrix is $(8n + m - 2) \times (8n + m - 2)$. Solving (A11) numerically is standard, and can be done using various techniques (e.g., QR or LU factorization).