

A Heath-Jarrow-Morton framework for energy markets: review and applications for practitioners

Matteo Gardini* Edoardo Santilli†

June 26, 2025

Abstract

This article delves into the application of the Heath-Jarrow-Morton framework to energy markets. Its objective is to provide a comprehensive overview of this subject, emphasizing practical considerations over theoretical aspects, which have already been extensively explored in the literature. The aim of this work is to serve as a practical guide for industry professionals and anyone involved in navigating the practical challenges inherent in applying this approach to energy markets.

In particular, the article focuses on several key areas, including market structure analysis, model calibration using Principal Component Analysis, Monte Carlo simulations, and the pricing of derivatives. European power and gas markets are examined as primary applications. The model is calibrated using historical futures data, followed by simulations of futures and spot prices, with subsequent analysis of the outcomes.

Moreover, the article explores the pricing of complex financial derivatives within energy markets, such as virtual power plants, swing options and gas storage.

Keywords: Stochastic processes, Heath-Jarrow-Morton, Energy Markets, Monte Carlo, Principal Components Analysis, Calibration, Option Pricing.

1 Introduction

Electricity markets worldwide exhibit differences attributable to various factors, including fundamental aspects such as demand and generation mix, as well as regulatory frameworks. In many advanced economies, the deregulation of energy markets aims to enhance their ability to react efficiently to fluctuations in supply and demand.

Focusing on the United States, Park et al. [55] demonstrated that deregulation has led to the emergence of more competitive and interconnected environments in both electricity and natural gas markets. Building upon this, Emery and Liu [29] found that the daily settlement prices of New York Mercantile Exchange's (NYMEX) California-Oregon Border (COB) and Palo Verde (PV) electricity futures contracts exhibit cointegration with prices

*Eni Plenitude, Via Ripamonti 85, 20136, Milan, Italy, email matteo.gardini@eniplenitude.com,
Department of Mathematics, University of Genoa, Via Dodecaneso 45, 16146, Genoa, Italy, email gardini@dima.unige.it

†Eni Plenitude, Via Ripamonti 85, 20136, Milan, Italy, email edoardo.santilli@eniplenitude.com

of natural gas futures contracts. This finding was corroborated by Mjelde and Bessler [51], who highlighted that electricity prices primarily respond to shocks in the coal market. Additionally, Bachmeier and Griffin [3] extensively explored the degree of market integration between crude oil, coal, and natural gas prices.

Several authors have explored the impact of fuel price fluctuations on electricity prices in European countries. Asche et al. [2] conducted tests to determine market integration among natural gas, electricity, and oil prices in the UK during a period when the natural gas market was deregulated but not yet connected to the continental European gas market. Their findings highlighted evidence of an integrated energy market. Similarly, Panagiotidis and Rutledge [53] examined the relationship between UK wholesale gas prices and Brent oil prices, discovering co-integration throughout the entire sample period (1996–2003). Additionally, utilizing daily price data for Brent crude oil, NBP UK natural gas, and EEX electricity, Bencivenga et al. [7] demonstrated the integration of gas, oil, and electricity markets. Conversely, through robust multivariate long-run dynamics analysis, Bosco et al. [16] revealed the presence of four highly integrated central European electricity markets (France, Germany, the Netherlands, and Austria). The trends observed in these four electricity markets also appear to be reflected in gas prices, but not in oil prices.

The recent invasion of Ukraine by Russia, coupled with concerns over potential gas supply shortages for Europe, precipitated a surge in gas and electricity prices¹—a phenomenon unprecedented in scale, as illustrated in Figure 1. This can be readily understood from an economic standpoint, given that natural gas accounts for 19.2% of electricity production in Europe, with natural gas power plants typically serving as the marginal technology in the electricity supply curve.

By now, it should be clear that considering the integration between energy markets is essential for those interested in energy commodities modeling, risk management, or derivatives pricing.

Over the years, numerous approaches have been proposed to model energy markets within a univariate framework. Pioneering works in this field trace back to the turn of the century: Schwartz and Smith [63], Lucia and Schwartz [47], and Schwartz [62] primarily focused on Gaussian frameworks. Meanwhile, Cartea and Figueroa [23] introduced a mean-reverting model incorporating jumps and a deterministic seasonality component for electricity spot prices. Saifert and Uhrig-Homburg [58] conducted a comparison of different modeling approaches for power markets, while Benth et al. [10] offers a comprehensive overview of energy market modeling.

Many non-Gaussian models, based on Lévy processes originally proposed for equities, have also been applied to energy markets. These include the variance gamma model by Madan et al. [49], the jump-diffusion model by Merton [50] and the normal-inverse Gaussian process by Barndorff-Nielsen [5], with a two-factor version of the latter recently applied to the energy context by Piccirilli et al. [57]. Additionally, various stochastic volatility models, such as those introduced by Heston [37] and Bates [6], can be adapted to capture commodity price behavior. These sophisticated tools enable the consideration of key stylized facts, such as jumps in price trajectories, skewness and fat-tails in log-

¹It is worth noting that this increase began prior to the invasion, following the Covid-19 pandemic, with numerous complex factors contributing to it. Nonetheless, the impact of war on European energy commodity prices has been unmistakable.

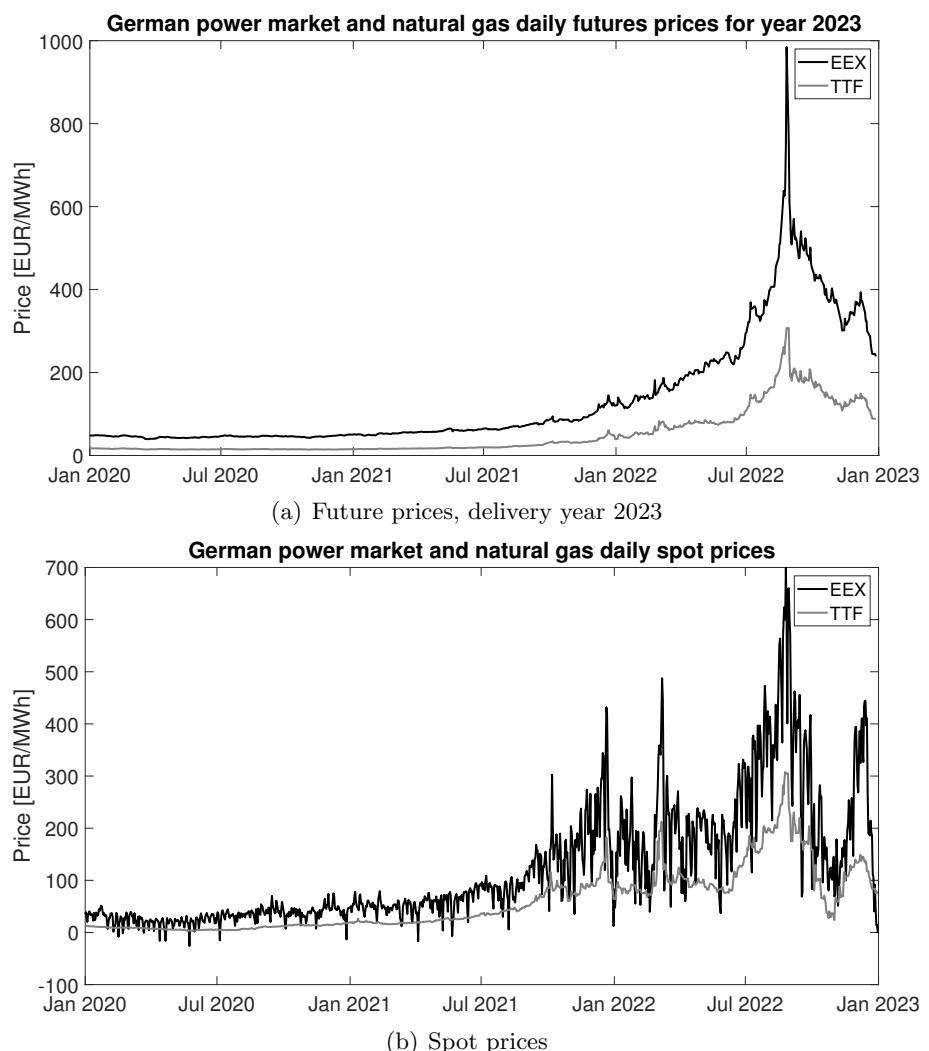


Figure 1: German power (EEX) and Dutch natural gas (TTF) daily prices from January 1, 2020 to the December 31, 2022.

returns distribution, and volatility smiles. A comprehensive review of financial modeling with jump processes can be found in Cont and Tankov [26].

While financial modeling in a univariate setting has been extensively explored, challenges arise when transitioning to a multi-commodity market. In this context, applying former modeling techniques becomes more complex in practice, and the literature is not as abundant as in the one-dimensional framework. Petroni and Sabino [56] demonstrated how certain standard models, such as those proposed by Black and Scholes [14], Schwartz and Smith [63], and Cartea and Figueroa [23], can be extended to a multivariate context by incorporating dependent jumps modeled using self-decomposable laws. Additionally, Kiesel and Kusterman [41] introduced a structural model to effectively account for the market coupling effect in electricity spot markets, drawing inspiration from the approach proposed by Carmona and Coulon [22].

An established approach for the modeling of energy markets was proposed by Benth and Saltyte-Benth [9], which adapted the framework introduced by Heath et al. [36] for energy futures markets. Numerous authors, including Sclavounos and Ellefsen [64], Hinderks et al. [38], Benth et al. [11], Broszkiewicz-Suwaj and Weron [19], Edoli et al. [28], and Feron and Gruet [30], have studied this modeling technique and its calibration. Despite their mathematical rigor, these articles are often challenging to apply in practice due to several practical considerations being overlooked. Common omissions include details on data preparation, step-by-step implementation guidelines, a practitioner-oriented interpretation of both theory and results and management strategies for typical implementation issues. This work seeks to fill these gaps by revisiting prior research on the application of the (Gaussian) Heath-Jarrow-Morton (HJM) framework in energy markets. Focusing on European power and gas futures markets, we present a comprehensive approach for data analysis and preparation, model calibration, simulations, and pricing.

The Gaussian hypothesis in log-returns can be relaxed by employing Lévy processes (as discussed by Benth et al. [10]) or Lévy copulas, as proposed by Panov and Samarin [54] and Cont and Tankov [26]. However, these approaches pose significant practical challenges, as acknowledged by the authors themselves. Indeed, the calibration step is particularly difficult to manage, even when dealing with a versatile multidimensional distribution like the Normal Inverse Gaussian. This difficulty is exacerbated when the number of underlying assets is large.

Many authors, including Luciano and Semeraro [48], Schoutens [61], Ballotta and Bonfiglioli [4], Buchmann et al. [20], and Gardini et al. [32], have explored alternative techniques to incorporate dependence in log-returns while remaining within a Lévy framework. While these approaches tend to perform well when the number of risky assets is small, complications arise when dealing with numerous underlying assets due to the rapid increase in the number of parameters. As a result, calibration becomes challenging in practice and the reliability of its results may be compromised.

In view of these comments, despite its limitations, the Gaussian framework remains a cornerstone for practitioners in multi-commodity energy markets. Conversely, when focusing on a single product, it may be better to consider a more evolved model from those listed above.

The article is structured as follows: Section 2 introduces the model, while Section 3 describes the market structure and the dataset. Section 4 covers the necessary data preparation steps required for fitting the model using the PCA technique. This task is addressed in Section 5, whereas, in Section 6, we discuss the simulations generated by the model and price some of the most common energy derivative contracts, such as virtual power plants, swing options and gas storage. Section 7 concludes the work.

2 The model

In this section, we delve into a detailed discussion of the HJM framework as applied to energy markets. In Section 2.1, we provide a brief overview of the mathematical foundations and review pertinent literature, directing interested readers to Benth et al. [10, Chapter 6.8]. In Sections 2.2 and 2.3, we outline the framework we used in this article to accurately model both forward and spot energy markets.

2.1 Modeling aspects and mathematical setting

In energy markets a variety of products are traded. In subsequent sections, we adhere to the notation introduced by Benth et al. [10, Chapter 1]. Specifically:

- $S(t)$ represents the spot prices of the commodity at time t .
- $F(t, \tau^s, \tau^e)$ denotes the price at time t of a contract delivering energy between τ^s and τ^e at a fixed price. Such contracts are commonly referred to as “swap contracts”, and the interval $[\tau^s, \tau^e]$ is designated as the ”delivery period.” These delivery periods typically span months, quarters and years, although shorter intervals such as weeks or days may also be encountered.
- $f(t, T)$ indicates the forward price at time t of a contract delivering energy at time T , where $t \leq T$.

The primary determinant of the swap price is the underlying spot price process $S = \{S(t); t \geq 0\}$. Consider the scenario where we have engaged in a forward contract to deliver the spot at time τ , assuming a constant interest rate $r \geq 0$. The payoff at time τ is given by:

$$S(\tau) - f(t, \tau).$$

Assuming the absence of arbitrage opportunities in the market, there exists a measure \mathbb{Q} , referred to as the “risk-neutral measure”: under this measure, the price of any derivative is determined by the discounted expectation of its payoff. Given that entering into a forward contract involves no initial cost, we have that:

$$0 = e^{-r(T-\tau)} \mathbb{E}^{\mathbb{Q}} [S(\tau) - f(t, \tau) | \mathcal{F}_t],$$

where $(\mathcal{F}(t); t \geq 0)$ represents the filtration encompassing all available market information up to time t . The forward price is determined at time t , rendering $f(t, \tau)$ adapted to $(\mathcal{F}(t); t \geq 0)$, thus yielding:

$$f(t, \tau) = \mathbb{E}^{\mathbb{Q}} [S(\tau) | \mathcal{F}_t].$$

Consequently, the process $f = \{f(t, \tau), t \leq \tau\}$ is a \mathbb{Q} -martingale.

With this in mind, consider T^* as the market's final time. Following the HJM approach, for $0 \leq t \leq \tau \leq T^*$, we model the dynamics of forward prices under \mathbb{Q} as:

$$f(t, \tau) = f(0, \tau) \exp \left\{ \int_0^t a(u, \tau) du + \sum_{k=1}^{\tilde{N}} \sigma_k(u, \tau) dW_k(u) \right\} \quad (1)$$

where $a(u, \tau)$ and $\sigma_k(u, \tau)$ are real-valued continuous functions on $[0, \tau] \times [0, T^*]$, with the additional conditions that σ_k must be positive functions. To ensure that $f(t, u)$ behaves as \mathbb{Q} -martingales, certain constraints on the drift $a(u, \tau)$ are necessary. As in Benth et al. [10, Proposition 6.1], we have the following drift condition:

$$\int_0^t a(u, \tau) du + \frac{1}{2} \sum_{k=1}^{\tilde{N}} \int_0^t \sigma_k^2(u, \tau) du = 0 \quad (2)$$

Failure to satisfy Equation (2) may lead to the emergence of arbitrage opportunities.

In energy markets, it is common to encounter contracts structured as $F(t, \tau^s, \tau^e)$ with overlapping delivery periods. For instance, one might encounter a contract delivering electricity for the entire calendar year alongside four contracts, each delivering electricity quarterly within the same year. To prevent arbitrage opportunities, it is imperative to ensure certain conditions are met among the prices of these swap contracts. In continuous time, we must adhere to “no-arbitrage conditions” of the following form:

$$F(t, \tau^s, \tau^e) = \int_{\tau^s}^{\tau^e} \hat{w}(u, \tau^s, \tau^e) f(t, u) du, \quad (3)$$

where $\hat{w}(u, \tau^s, \tau^e)$ is a weighting function of the form:

$$\hat{w}(u, s, t) = \frac{w(u)}{\int_t^s w(v) dv},$$

where $w(u)$ is a positive function and we can assume $w(u) = 1$ if the settlement of the swap takes place at the end of the delivery period (see Benth et al. [10, Chapter 4]).

Drawing an analogy to the modeling approach utilized for future contracts $f(t, u)$, we might impose the dynamics for $F(t, \tau^s, \tau^e)$ to be:

$$F(t, \tau^s, \tau^e) = F(0, \tau^s, \tau^e) \exp \left\{ \int_0^t A(u, \tau^s, \tau^e) du + \sum_{k=1}^{\tilde{N}} \int_0^t \Sigma_k(u, \tau^s, \tau^e) dW_k(u) \right\}. \quad (4)$$

Here, $(W_1, \dots, W_{\tilde{N}})$ are independent Brownian motions, $A(u, \tau^s, \tau^e)$ and $\Sigma_k(u, \tau^s, \tau^e)$ denote continuous real-valued functions and $F(0, \tau^s, \tau^e)$ represents the initial forward price of a contract. To ensure that swap prices behave as \mathbb{Q} -martingales, it is essential that:

$$\int_0^t A(u, \tau^s, \tau^e) du + \frac{1}{2} \sum_{k=1}^{\tilde{N}} \int_0^t \Sigma_k(u, \tau^s, \tau^e) du = 0.$$

Regrettably, the model proposed for the swap price $F(t, \tau^s, \tau^e)$ fails to adhere to the continuous-time no-arbitrage conditions outlined in Equation (3). Consequently, alternative models must be explored. One approach involves deriving models from forward contracts $f(t, \tau)$. Alternatively, inspiration can be drawn from LIBOR market models within fixed income theory (refer to Brigo and Mercurio [17] and Oosterlee and Grzelak [52]).

2.1.1 Swap models and market models

The initial methodology to capture the dynamics of swap contracts $F(t, \tau^s, \tau^e)$ is known as *swap model* and it involves commencing with the dynamics of forward prices $f(t, \tau)$ as described in Equation (1), alongside the drift condition outlined in Equation (2). The swap price is then derived as follows:

$$F(t, \tau^s, \tau^e) = \int_{\tau^s}^{\tau^e} \hat{w}(u, \tau^s, \tau^e) f(t, u) du,$$

Obviously, this formulation ensures compliance with the non-arbitrage relations specified in Equation (3). However, this approach often results in non-Markovian models (with exceptions being trivial cases), posing challenges in calibration and simulation, as extensively discussed in Benth et al. [10, Chapter 6].

An alternative approach relies on what are known as *market models*. Drawing inspiration from LIBOR models, these frameworks construct dynamics for traded contracts that align with the observed volatility term structure. Unlike the previous approach, market models exclusively consider products actively traded in the market, thereby bypassing the continuous-time no-arbitrage conditions.

Initially, contracts that cannot be decomposed into smaller delivery periods are singled out. For instance, we focus solely on monthly contracts and omit year and quarter products. These contracts, referred to as *basic contracts*, whose delivery periods are denoted by:

$$\left\{ [\tau_1^s, \tau_1^e], \dots, [\tau_{\tilde{N}}^s, \tau_{\tilde{N}}^e] \right\},$$

where \tilde{N} represents the total number of contracts considered. Consequently, we define the dynamics for each basic contract as:

$$F_m(t, \tau_m^s, \tau_m^e) = F_m(0, \tau_m^s, \tau_m^e) \exp \left\{ \int_0^t A_m(u, \tau_m^s, \tau_m^e) du + \sum_{k=1}^{\tilde{N}} \int_0^t \Sigma_{m,k}(u, \tau_m^s, \tau_m^e) dW_k(u) \right\},$$

for $m = 1, \dots, \tilde{N}$, where A_m and $\Sigma_{m,k}$ represent continuous real-valued functions defined on $[0, \tau_m^s]$. A condition analogous to the one stated in Equation (4) ensures that the dynamics of swap prices form a \mathbb{Q} -martingale.

In the subsequent section we show how the HJM approach can be used to properly model a multi-commodity market.

2.2 A model for a multi-commodity market

In this section, we initially concentrate on a single-factor model to gain a deeper understanding of how to accurately represent the volatility structure. Subsequently, we show how to extend the model to incorporate an arbitrary number of stochastic factors.

2.2.1 One factor model

Consider the price of the swap contract $F(t, \tau^s, \tau^e)$ with maturity² $T \leq \tau^s$, signed at t_0 , for $t \in [t_0, T]$. If we consider only a single Brownian motion, assuming we are working under the risk neutral measure \mathbb{Q} , we consider a dynamics of the following type:

$$\frac{dF(t, \tau^s, \tau^e)}{F(t, \tau^s, \tau^e)} = \sigma(t, T)dW(t), \quad F(t_0, \tau^s, \tau^e) = F(0, \tau^s, \tau^e) \quad a.s., \quad t \leq T, \quad (5)$$

where the volatility $\sigma(t, T)$ is assumed to be a deterministic time dependent function $\sigma(t, T) : [t_0, T] \rightarrow \mathbb{R}^+$ and $F(0, \tau^s, \tau^e)$ is the value of the future contract at time t_0 .

According to Samuelson [59], the term structure of commodity forward price volatility typically declines with contract horizon, a phenomenon commonly known as “Samuelson’s effect”. Therefore, it is customary to assume a volatility function which depends on the time to maturity ($T - t$), namely $\sigma(t, T) = \sigma(T - t)$. In particular, $\sigma(T - t)$ will be decreasing in $T - t$.

Despite any form can be assumed for $\sigma(T - t)$ in Equation (5), from a practical point of view a step-wise volatility structure is usually considered. If we take a monthly partition of the time interval $[t_0, T]$ with dates $t_0 = T_{M_0} < \dots < T_{M_{j-1}} < T_{M_j} < \dots < T_{M_M} = T$, the volatility function assumes the following form:

$$\sigma(T - t) = \sum_{j=1}^M \sigma_j \mathbb{1}_{I_j}(T - t),$$

where $\sigma_j \in \mathbb{R}^+$ and $I_j = (T_{M_{j-1}}, T_{M_j}]$. Hence volatility has a step-wise structure as the one shown in Figure 2 where parameters σ_j must be estimated.

To calibrate the parameters σ_j , the key approach involves constructing M synthetic fixed-delivery products, denoted as F_{M_j} , from \tilde{N} swap prices $F(r, \tau_n^s, \tau_n^e)$. Here, M_j indicates that the product delivery occurs j months after the current time. The dynamics of F_{M_j} is given by by:

$$\frac{dF_{M_j}(t, T_{M_j})}{F_{M_j}(t, T_{M_j})} = \sigma_j dW(t),$$

allowing us to compute σ_j by evaluating the standard deviation of log-returns.

We will revisit the data preparation methodology, and the calibration algorithm for the general case will be addressed in Section 5.

²In order to simplify the discussion, we usually assume that $T = \tau^s$, namely that the contract expires as soon as the delivery period starts. This assumption is no so far from what happens in real markets, where swaps contract are usually traded up to few days before the beginning of the delivery period.

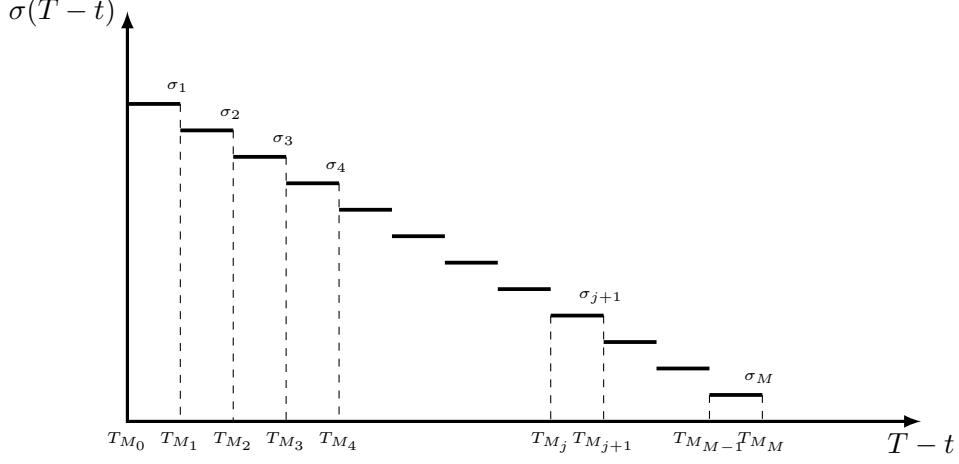


Figure 2: Volatility structure $\sigma(T - t)$.

2.2.2 The general multi-factor model

Extending the dynamics outlined in Equation (5) to a multi-commodity framework is a straightforward process. Consider a scenario where we have $k = 1, \dots, K$ markets with a designated final time for trading denoted as T^* . As before, we establish a monthly partition denoted by $t_0 = T_{M_0} < \dots < T_{M_{j-1}} < T_{M_j} < \dots < T_{M_M} = T^*$ over the interval $[t_0, T^*]$. Initially, in the market we have $F^k(t, \tau_n^s, \tau_n^e)$, where $n = 1, \dots, N_k$ and N_k signifies the total number of contracts in market k . For each market k , we can then generate an equivalent number $M_k = M$ of contracts with monthly delivery periods, expressed as $F_{M_j}^k(t, T_{M_j})$, where $j = 1, \dots, M$. A detailed explanation of this process can be found in Section 4.

We start by delineating the dynamics of the fixed-delivery futures contract $F_{M_j}^k(t, T_{M_j})$. We consider each of these contracts as a “random factor” capable of influencing the overall market dynamics. Consequently, we are dealing with a total of $\tilde{N} = M \cdot K$ random factors. The dynamics of a future product k with fixed-delivery T_{M_j} is given by:

$$\frac{dF_{M_j}^k(t, T_{M_j})}{F_{M_j}^k(t, T_{M_j})} = \sum_{i=1}^{\tilde{N}} \sigma_{ji}^k dW_i(t), \quad F_{M_j}^k(t_0, T_{M_j}) = F_{M_j}^k(0, T_{M_j}) \quad a.s., \quad t \in [t_0, T_{M_j}]. \quad (6)$$

where $W_i = \{W_i(t); t \geq 0\}$ $i = 1, \dots, \tilde{N}$ are independent Brownian motions. From Equation (6) observe that the dynamics of the single monthly product $F_{M_j}^k(t, T_{M_j})$ potentially depends on the those of all other monthly futures products. From empirical evidence, energy markets are strongly co-integrated and hence it is clear that considering \tilde{N} independent Brownian motions appears to be unreasonable. A possible methodology, based on the dimensional reduction inherited from the Principal Component Analysis (PCA), to select a lower number of stochastic factors, will be presented in Section 5.

Equation (6) can be written in a matrix form as:

$$\frac{d\mathbf{F}(t, T)}{\mathbf{F}(t, T)} = \boldsymbol{\sigma} \cdot d\mathbf{W}(t), \quad (7)$$

where $\mathbf{W} = (W_1, \dots, W_{\tilde{N}})$ is \tilde{N} -dimensional a standard Brownian motion with independent components, $\boldsymbol{\sigma}$ is a $\tilde{N} \times \tilde{N}$ and:

$$\mathbf{F}(t, T) = \begin{cases} \mathbf{F}^1(t, T) & \left[\begin{array}{c} F_{M_1}^1(t, T_{M_1}) \\ F_{M_2}^1(t, T_{M_2}) \\ \vdots \\ F_{M_M}^1(t, T_{M_M}) \end{array} \right] \\ \mathbf{F}^2(t, T) & \left[\begin{array}{c} F_{M_1}^2(t, T_{M_1}) \\ F_{M_2}^2(t, T_{M_2}) \\ \vdots \\ F_{M_M}^2(t, T_{M_M}) \end{array} \right] \\ \vdots & \vdots \\ \mathbf{F}^k(t, T) & \left[\begin{array}{c} \vdots \\ \vdots \\ \vdots \\ F_{M_1}^k(t, T_{M_1}) \\ F_{M_2}^k(t, T_{M_2}) \\ \vdots \\ F_{M_M}^k(t, T_{M_M}) \end{array} \right] \\ \vdots & \vdots \\ \mathbf{F}^K(t, T) & \left[\begin{array}{c} \vdots \\ \vdots \\ \vdots \\ F_{M_1}^K(t, T_{M_1}) \\ F_{M_2}^K(t, T_{M_2}) \\ \vdots \\ F_{M_M}^K(t, T_{M_M}) \end{array} \right] \end{cases}, \quad \boldsymbol{\sigma} = \begin{cases} \boldsymbol{\sigma}^1 & \left[\begin{array}{ccccc} \sigma_{1,1}^1 & \sigma_{1,2}^1 & \sigma_{1,3}^1 & \cdots & \sigma_{1,\tilde{N}}^1 \\ \sigma_{2,1}^1 & \sigma_{2,2}^1 & \sigma_{M+2,3}^1 & \cdots & \sigma_{2,\tilde{N}}^1 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ \sigma_{M,1}^1 & \sigma_{M,2}^1 & \sigma_{M,3}^1 & \cdots & \sigma_{2M,\tilde{N}}^1 \\ \sigma_{1,1}^2 & \sigma_{1,2}^2 & \sigma_{1,3}^2 & \cdots & \sigma_{1,\tilde{N}}^2 \\ \sigma_{2,1}^2 & \sigma_{2,2}^2 & \sigma_{2,3}^2 & \cdots & \sigma_{2,\tilde{N}}^2 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ \sigma_{M,1}^2 & \sigma_{M,2}^2 & \sigma_{M,3}^2 & \cdots & \sigma_{M,\tilde{N}}^2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \boldsymbol{\sigma}^K & \left[\begin{array}{ccccc} \sigma_{1,1}^K & \sigma_{1,2}^K & \sigma_{1,3}^K & \cdots & \sigma_{1,\tilde{N}}^K \\ \sigma_{2,1}^K & \sigma_{2,2}^K & \sigma_{2,3}^K & \cdots & \sigma_{2,\tilde{N}}^K \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ \sigma_{M,1}^K & \sigma_{M,2}^K & \sigma_{M,3}^K & \cdots & \sigma_{M,\tilde{N}}^K \end{array} \right] \end{cases} \quad (8)$$

where $\boldsymbol{\sigma}^k \in \mathbb{R}^{M \times \tilde{N}}$ denotes the matrix associated to random factors of k -th market product.

Once we have defined the dynamics for $F_{M_j}^k(t, T_{M_j})$ that of the $F_n^k(t, \tau_n^s, \tau_n^e)$ is a direct consequence. The dynamics of the n -th swap contract $F_n^k(t, \tau_n^s, \tau_n^e)$ for $t \in [t_0, \tau_n^s]$ is given by:

$$\frac{dF_n^k(t, \tau_n^s, \tau_n^e)}{F_n^k(t, \tau_n^s, \tau_n^e)} = \sum_{i=1}^{\tilde{N}} \sum_{j=1}^M \sigma_{ji}^k \mathbb{1}_{I_j}(T^* - t) dW_i(t), \quad (9)$$

and its explicit solution for $t \in [t_0, \tau_n^s]$ is given by:

$$F^k(t, \tau_n^s, \tau_n^e) = F^k(t_0, \tau_n^s, \tau_n^e) \exp \left\{ -\frac{1}{2} \sum_{i=1}^{\tilde{N}} \sum_{j=1}^M \left(\sigma_{ji}^k \right)^2 \int_{t_0}^t \mathbb{1}_{I_j}(T^* - s) ds + \sum_{i=1}^{\tilde{N}} \sum_{j=1}^M \sigma_{ji}^k \int_{t_0}^t \mathbb{1}_{I_j}(T^* - s) dW_i(s) \right\}. \quad (10)$$

The availability of a closed-form solution enables us to exactly simulate the paths of the process without resorting to any discretization methods, such as Euler's or Millstein's, as discussed in Seydel [65]. It is worth noting that when t is close to t_0 , the difference $T^* - t$ is relatively large, indicating we are considering the volatility of products with longer maturities. However, as time progresses and t increases, $T^* - t$ diminishes, leading us to focus on the volatility of fixed-delivery products with shorter maturities.

A sanity check of the implemented simulation algorithm can be conducted by simulating the process according to Equation (10) over a brief duration, specifically when $T^* - t$ is approximately equal to $T^* - t_0$. Consequently, Equation (10) simplifies to:

$$F_n^k(t, \tau_n^s, \tau_n^e) = F_n^k(t_0, \tau_n^s, \tau_n^e) \exp \left\{ -\frac{1}{2} \Delta t \sum_{i=1}^{\tilde{N}} \left(\sigma_{m^*i}^k \right)^2 + \sum_{i=1}^{\tilde{N}} \sigma_{m^*i}^k W_i(t) \right\}, \quad (11)$$

where $\Delta t = t - t_0$ and m^* is such that σ_{m^*i} , $i = 1, \dots, \tilde{N}$ are the volatilities associated to $F_{M_{m^*}}^k(t, T_{M_{m^*}})$ with $T_{M_{m^*}} = \tau_n^s$. This implies that if we aim to simulate the product $F^k(t, \tau_n^s, \tau_n^e)$ with a monthly delivery period $[\tau_n^s, \tau_n^e]$ three months from the present, with, for example, a Δt of one day, we only need to consider the volatility linked to the monthly product $F_{M_3}^k(t, T_{M_3})$, which represents the fixed-delivery product expiring three months from now. By computing the log-return $x^k(t) = \ln F_M(t, T_{M_{m^*}}) - \ln F(t_0, T_{M_{m^*}})$ we get:

$$x^k(t) = -\frac{1}{2} \Delta t \sum_{i=1}^{\tilde{N}} \left(\sigma_{m^*i}^k \right)^2 + \sum_{i=1}^{\tilde{N}} \sigma_{m^*i}^k dW_i(t), \quad (12)$$

and if we compute its variance we get:

$$\text{Var} \left[x^k(t) \right] = \Delta t \sum_{i=1}^{\tilde{N}} \left(\sigma_{m^*i}^k \right)^2.$$

Hence, we have a simple way to check that the variance of the simulated product is correct, by summing up the squares of the entries of the matrix $\boldsymbol{\sigma}^k$ and multiplying by Δt .

Clearly if the approximation $T^* - t \approx T^* - t_0$ does not hold, namely for larger t such expression is no longer valid. On the other hand, simple computations show that:

$$\text{Var} \left[\log F_n^k(t, \tau_n^s, \tau_n^e) \right] = \sum_{i=1}^{\tilde{N}} \sum_{j=1}^M \left(\sigma_{ji}^k \right)^2 \int_{t_0}^t \mathbb{1}_{I_j}(T^* - s) ds, \quad (13)$$

which can be easily used to check the correctness of the simulations.

We conclude by noting that the formulas for European vanilla options' styles option, as presented by Black [13], can be readily derived once the expression for $\text{Var} [\log F_n^k(t, \tau_n^s, \tau_n^e)]$ is known, as discussed in Clewlow and Strickland [25]. Therefore, such formulas can be easily derived within the framework we have proposed, ensuring a straightforward method to calibrate the volatility structure over quoted derivatives, if available, as discussed, for instance, in Cont and Tankov [26].

2.3 The spot dynamics

The dynamics of the spot prices $S^k(t)$ can be obtained by setting $\tau^s = \tau^e = t$ in the spirit of Clellow and Strickland [24]. From Equation (10) we get

$$S^k(t) = F^k(t_0, t) \exp \left\{ -\frac{1}{2} \sum_{j=1}^{\tilde{N}} \sum_{i=1}^M \left(\sigma_{ij}^k \right)^2 \int_{t_0}^t \mathbb{1}_{I_i}(t-s) ds + \sum_{j=1}^{\tilde{N}} \sum_{i=1}^M \sigma_{ij}^k \int_{t_0}^t \mathbb{1}_{I_i}(t-s) dW_j(s) \right\}. \quad (14)$$

Observe that it is important to understand how volatility behaves as time t increases. If $t \approx t_0$ then the volatility of the spot price $S^k(t)$ is the one of the fixed-delivery futures product with the shorter time to maturity. As time t increases, the spot price $S^k(t)$ somehow includes all the volatility effects from the products with shorter time to maturity to the ones with a longer one. This is reasonable from an economic point of view. Indeed, if we aim to simulate the process $S^k(t)$ over the course of a year, it would involve M_{12} , M_{11} , and so forth down to M_0 , necessitating the aggregation of all their volatility contributions. Similarly, in this scenario, an expression for $\text{Var}[S^k(t)]$ akin to the one provided in Equation (13) can be derived.

From a practical point of view, especially for power prices, it is customary to consider a hourly granularity. $F^k(t_0, t)$ represents the power forward price today for hour t and hence the function $F^k(t_0, t)$ for $t \in [t_0, T^*]$, represents the hourly forward curve for the spot market k . In simulation routines, to streamline the process, we assume that the “shocks” concerning the hourly forward curve occur on a daily basis. For commodities with daily granularity, like natural gas, all operations are conducted on a daily timeframe. The hourly or daily forward curve can be derived from futures market instruments represented as $F(t, \tau^s, \tau^e)$, and various methodologies, such as the approach suggested by Benth et al. [10, Chapter 7], can be employed to derive it.

3 Power and gas markets: data-set description

In this section we focus solely on the European power and gas futures markets, but the analysis can be easily adapted to any market. In particular, we consider data from the European Energy Exchange (EEX) for the power markets, whereas data regarding the natural gas markets comes from the Intercontinental Exchange (ICE).

We concentrate on the German (DE), Italian (IT), French (F7), and Swiss (CH) power futures markets, as well as the TTF and PSV, which represent the Dutch and Italian natural gas hubs, respectively. The dataset comprises all quotations from the year 2020. For each market, we analyze fixed delivery products $F_d(t, T_d)$ of Section 2 with delivery up to two years into the future, obtained by rearranging contracts $F(t, \tau^s, \tau^e)$ (see Section 4.2), with $d \in \{M_0, M_1, M_2, Q_1, Q_2, Q_3, Y_1, Y_2\}$ ³.

³Recall that, for example, a power futures (or swap, denoted by $F(t, \tau^s, \tau^e)$ in Section 2) 2020 calendar is a contract between two counterparts to buy or sell a specific volume of energy in MWh at fixed price, decided at time t , for all the hours of the year. In this case the delivery period lies between $\tau^s = 01/01/2020$

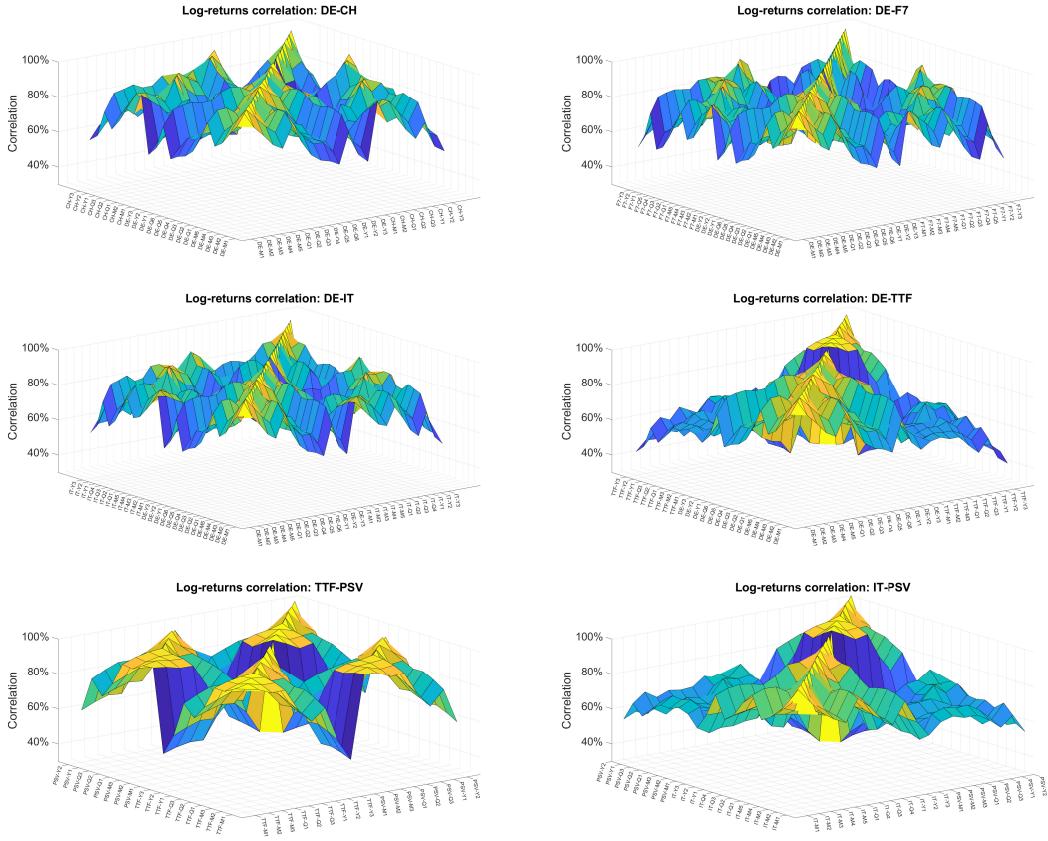


Figure 3: Log-returns correlation surfaces for different commodities.

Each traded contract contributes to the uncertainty, or “random factor”, shaping the forward curve dynamics. However, to model these curves effectively, the goal is to use a limited number Brownian motions, given the interconnectedness of these markets. In line with Sclavounos and Ellefsen [64], Figure 3 depicts correlation surfaces among daily log-returns of various futures products with fixed-delivery $F_d(t, T_d)$, with $d \in \{M_j, Q_j, Y_j\}$. Notably, a significant linear correlation coefficient is evident, suggesting that only a few stochastic factors seem to drive the overall market structure.

At this point of the discussion, a pertinent question arises: Is the model we have proposed suitable for the market we aim to represent? Our approach is built upon Brownian motion, a process that falls within the broader class of Lévy processes. Operating within a Lévy framework allows for models with a rich structure capable of capturing many stylized facts efficiently (refer to Sato [60], Applebaum [1], and Cont and Tankov [26] for more details). Lévy processes, by definition, have independent increments. Hence, before employing Lévy processes for modeling purposes, it is crucial to verify the independence

and $\tau^e = 31/12/2020$. τ^s plays also the role of maturity. As stated before, actually the product with delivery starting at τ^s is traded until the business day before the beginning of the delivery period, but in order to simplify the notation we consider τ^s as the maturity. See Benth et al. [10, Chapter 1] for a good description of power and gas markets.

of these increments. Following the methodology outlined in Brigo et al. [18], we compute the auto-correlation function (ACF) for six distinct time series, each corresponding to a specific market, focusing on calendar products with a delivery year of 2021. These series consist of daily log-returns x_1, x_2, \dots, x_n , with

$$x_i = \ln \frac{P_{t_{i+1}}}{P_{t_i}},$$

where P_{t_i} denotes the price of a (general) risky asset at time t_i and we compute the ACF with lag k

$$ACF(k) = \frac{1}{(n-k)\hat{v}} \sum_{i=1}^{n-k} (x_i - \hat{m})(x_{i+k} - \hat{m}), \quad k = 1, \dots, 20,$$

where \hat{m} and \hat{v} are the sample mean and variance. Roughly speaking, we can consider the ACF as an estimate of the correlation between the random variables $X(t_i)$ and $X(t_{i+k})$. The ACF plots for the six selected products mentioned above are displayed in the charts in Figure 4. Across all of these plots, we do not detect any significant lags in the historical return time series. This indicates that the assumption of independence is acceptable in this scenario. Moreover, similar results are obtained when altering the delivery period of the product. Consequently, Lévy processes can be deemed suitable for effectively modeling futures prices.

One fundamental consequence of our proposed modeling framework is the assumption that log-returns follow a normal distribution. However, as noted by several authors (e.g., Benth and Koekbakker [8], Frestad et al. [31], and Green [35]), log-returns in energy markets often deviate from normality. Instead, their distribution tends to exhibit skewness and heavy tails, indicative of fat tails effects. Additionally, log-return volatility is typically non-constant, often displaying clustering behavior. In Figure 5, we illustrate the empirical probability density function of log-returns compared to the normal distribution fitted on the same dataset. While the normal distribution may seem to provide a reasonable fit for this dataset, we observe distributions with peakedness and heavy tails, especially in power forward markets. Moreover, focusing on different time periods would reveal non-Gaussian log-returns, such as during 2022, when price spikes occurred due to the Russia-Ukraine conflict, resulting in a highly volatile market.

Considering these findings, adopting a normal distribution may oversimplify market characteristics. Nonetheless, practitioners often accept this assumption to establish a straightforward, efficient calibration methodology, albeit recognizing its limitations.

As a final note, it is worth mentioning that most power markets quote both peak-load and base-load products. The distinction between them lies in the delivery schedule: peak-load contracts supply electricity only during specific hours (from 8 AM to 8 PM) on working days, while base-load contracts provide power for all hours throughout the delivery periods. Traders and risk managers often consider both types of products. While we have not yet specified the distinction between base-load and peak-load contracts, we can easily incorporate them into our modeling approach. Specifically, in spot simulations, the volatility of both peak-load and base-load products can be accounted for. One approach involves separately simulating spot prices for peak-load and off-peak-load products

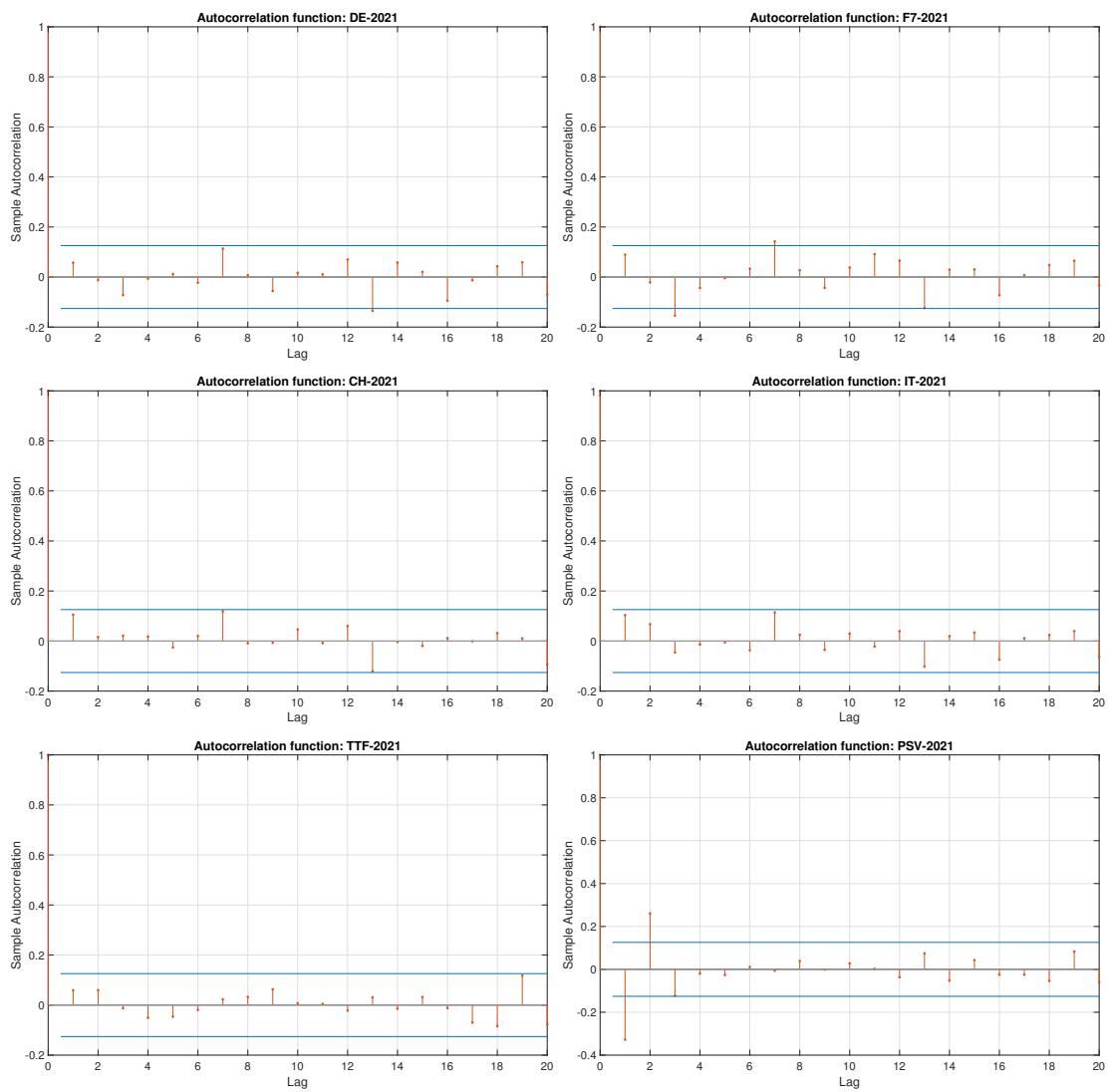


Figure 4: Sample ACF computed on log-returns.

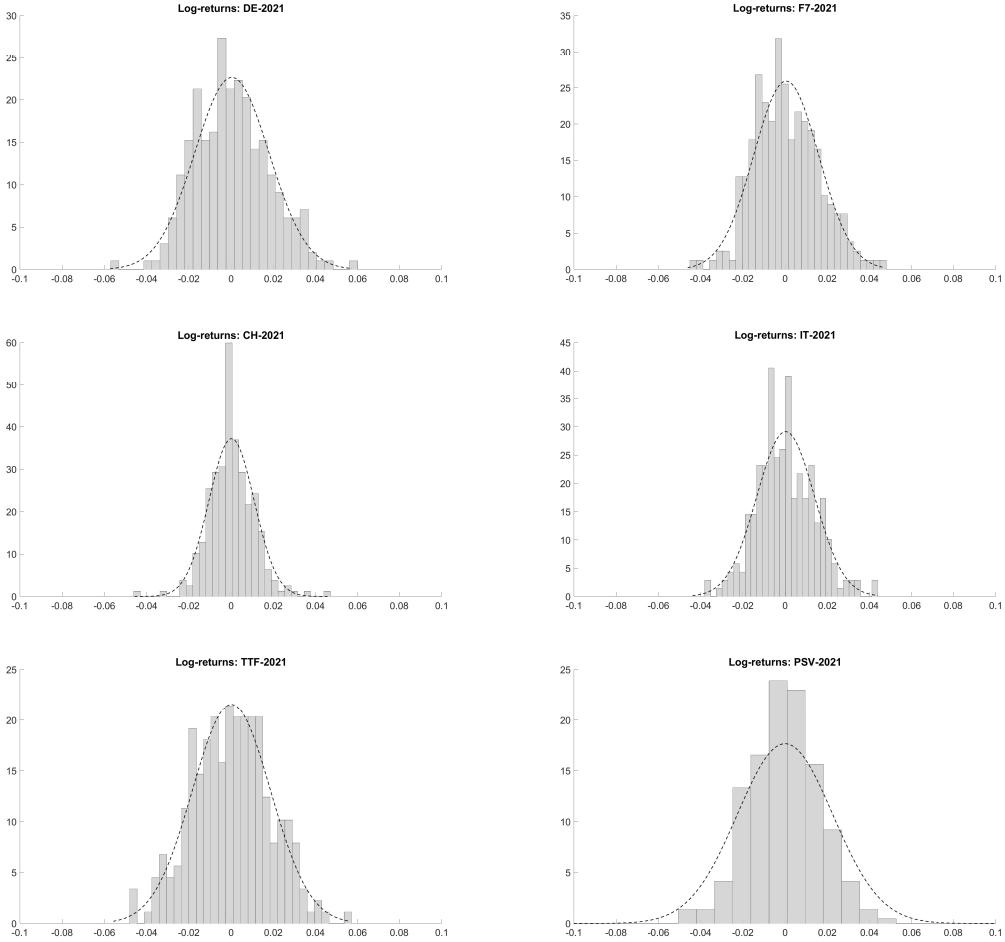


Figure 5: Daily futures log-returns densities for calendar 21 products.

and then combining them. However, in many power futures markets, peak-load contracts are less liquid than their corresponding base-load contracts. Consequently, an "expertise adjustment" is necessary to accurately estimate peak-load quotations when they are unavailable. This introduces subjective decisions, which could potentially influence spot simulation outcomes. Therefore, for the sake of simplicity in presentation, we focus only on base-load quotations.

In the following section, we discuss the methodology for data-set preparation, a crucial prerequisite for the accurate calibration of the model discussed in Section 2.2.1.

4 Data-set preparation

The data preparation stage is of paramount importance and requires meticulous attention. Inadequate preparation of the data can result in erroneous parameter estimations during calibration, leading to misleading output results. This section is divided into two parts:

Trading date	Jan-20	Feb-20	Mar-20	Q2-20	Q3-20	Q4-20	2021	2022
2020-01-02	36.05	39.76	37.15	35.50	39.05	45.30	43.85	46.55

Table 1: Some of the products $F(t, \tau^s, \tau^e)$ available on German futures market for the trading date 02/01/2020.

initially, we discuss the data preparation process, then we proceed by the utilization of the PCA technique for model calibration.

4.1 From $F(t, \tau^s, \tau^e)$ to $F_{M_j}(t, T_{M_j})$

As outlined in Section 2, we have assumed that the volatility $\sigma(t, T)$ is a step-wise function of the time to maturity $T - t$. As described in Section 2.2.1, volatility functions σ_j refer to products of the form $F_{M_j}(t, T_{M_j})$, which can be derived from $F(t, \tau^s, \tau^e)$ as follows. In Table 1, we focus on the German power futures market: however, the same rationale can be applied to any market. On January 2, 2020 we observe different products, such as the one that delivers on February 2020 (red): we can simply reconstruct the product $F_{M_1}(t, T_{M_1})$ delivering on the following month by setting: $F_{M_1}(t, T_{M_1}) = 39.76$. The same approach holds if we want to obtain the product $F_{Q_2}(t, T_{Q_2})$ from the swap contracts in Table 1 (blue) and so on. In this example, we have restricted the time-frame to two years from now, but longer delivery can be easily considered.

Attention should be directed towards the transition between months. For instance, as depicted in Table 2, on March 31, 2020, $F_{M_2}(t, T_{M_2})$ (highlighted in green) corresponds to April 2020; however, on April 1, 2020, the same product refers to May 2020. Thus, when referring to $F_{M_2}(t, T_{M_2})$, we are actually indicating the product delivering throughout May 2020 and not to April 2020. This market feature, known as a “rolling mechanism”, applies to products with more refined delivery schedules, such as quarters and months.

It is crucial to note that with each occurrence of a “rolling event”, the fixed-delivery product shifts to the left (as observed with the green products in Table 2). When computing the log-returns of a product like M_1 , for instance, a spike in log-returns coincides with the rolling event. It is clear that this spike is not a market-driven phenomenon but rather a consequence of a modeling assumption. Consequently, during the calibration process, log-returns spanning across months should be excluded. Otherwise, additional volatility, not reflective of the market, would be inadvertently incorporated.

Consider now $K = 6$ markets (DE, IT, F7, CH, TTF and PSV), each of them having the same number N_k , $k = 1, \dots, K$ of products $F^k(t, \tau_n^s, \tau_n^e)$, for $n = 1, \dots, N_k$, with monthly, quarterly and yearly delivery periods. Hence we can introduce the following set of fixed-delivery $F_d^k(t, T_d)$ with $d \in \{M_0, M_1, M_2, Q_1, Q_2, Q_3, Y_1, Y_2\}$.

By following the procedure outlined for the German power market, for all markets $k = 1, \dots, K$ we construct the fixed-delivery forward $F_d^k(t, T_d)$ and we obtain a data-set similar to the one depicted in Table 2.

At this stage, we have transitioned from swap contracts in the format $F^k(t, \tau^s, \tau^e)$ to those in the format $F_d^k(t, T_d)$. The subsequent step involves transforming contracts of the form $F_{Q_j}^k(t, T_{Q_j})$ and $F_{Y_j}^k(t, T_{Y_j})$ into contracts of the form $F_{M_j}^k(t, T_{M_j})$.

Trading date	M_0	M_1	M_2	Q_1	Q_2	Q_3	Y_1	Y_2
2020-01-02	36.05	39.76	37.15	35.50	39.05	45.30	43.85	46.55
2020-01-03	38.06	40.40	37.80	36.55	39.85	45.97	44.85	47.08
...
2020-03-31	15.74	17.06	19.79	26.74	27.63	34.45	35.65	39.05
2020-04-01	16.98	19.04	23.24	27.02	33.76	36.24	34.95	38.56
...
2020-12-30	43.52	54.53	53.47	51.72	43.72	46.33	48.28	49.80

Table 2: Fixed-delivery products $F_{M_j}(t, T_{M_j})$, $F_{Q_j}(t, T_{Q_j})$ and $F_{Y_j}(t, T_{Y_j})$ for the German forward market.

4.2 From $F_{Q_j}^k(t, T_{Q_j})$ and $F_{Y_j}^k(t, T_{Y_j})$ to $F_{M_j}^k(t, T_{M_j})$

To derive products in the form of $F_{M_j}(t, T_{M_j})$ from those with coarser granularity, we proceed as follows. Let T denote the upper bound of maturity time in the market. We establish a partition on $[t_0, T^*]$, $t_0 = T_{M_0} < \dots < T_{M_{i-1}} < T_{M_i} < \dots < T_{M_M} = T^*$, ensuring that each interval spans exactly one month. Consequently, starting from products with coarser granularity, namely quarters and years ($F_{Q_j}(t, T_{Q_j})$ and $F_{Y_j}(t, T_{Y_j})$), we generate monthly products $F_{M_j}(t, T_{M_j})$. To achieve this, we need to calculate a “flat monthly forward curve” by determining the value of each monthly futures contract, ensuring that no arbitrage opportunities arise. A widely recognized method for constructing a smooth forward curve with a seasonal effect, coherent with observed futures quotations in the market, has been proposed by Benth et al. [10, Chapter 7]. A similar procedure can be applied and simplified by replacing the smooth curve with a step-wise one.

Consider the example in Figure 6 where, for a given market k , overlapped products are allowed and suppose we want to define a step-wise forward curve on the intervals $[\tau_{i-1}, \tau_i]$, $i = 1, \dots, M$ with $M = 7$ of the form:

$$\epsilon(u) = \sum_{i=1}^M a_i \mathbb{1}_{[\tau_{i-1}, \tau_i]}(u),$$

where $\{a_i\}_{i=1}^M$ are the values we have to fit. When considering $F(t_0, \tau_i^s, \tau_i^e)$, it is imperative to ensure that arbitrages are avoided by satisfying the following relation for each quoted product $i = 1, \dots, n$:

$$F(t_0, \tau_i^s, \tau_i^e) = \frac{1}{\tau_i^e - \tau_i^s} \int_{\tau_i^s}^{\tau_i^e} \epsilon(u) du.$$

Finding the values $\{a_i\}_{i=1}^n$ reduces to solving a linear system, which can be efficiently accomplished numerically. Once the values $\{a_i\}_{i=1}^n$ are obtained, $\epsilon(u)$ can be determined. Subsequently, for each trading date and market, such a monthly forward curve must be computed. If we take the trading date January 4, 2020, the resulting monthly forward curve is depicted in Figure 7. Of course, if during the construction of the forward curve some products are completely overlapped they must be removed preserving those with the finer granularity. The final set of fixed-delivery products $F_{M_j}(t, T_{M_j})$, $j = 1, \dots, M$

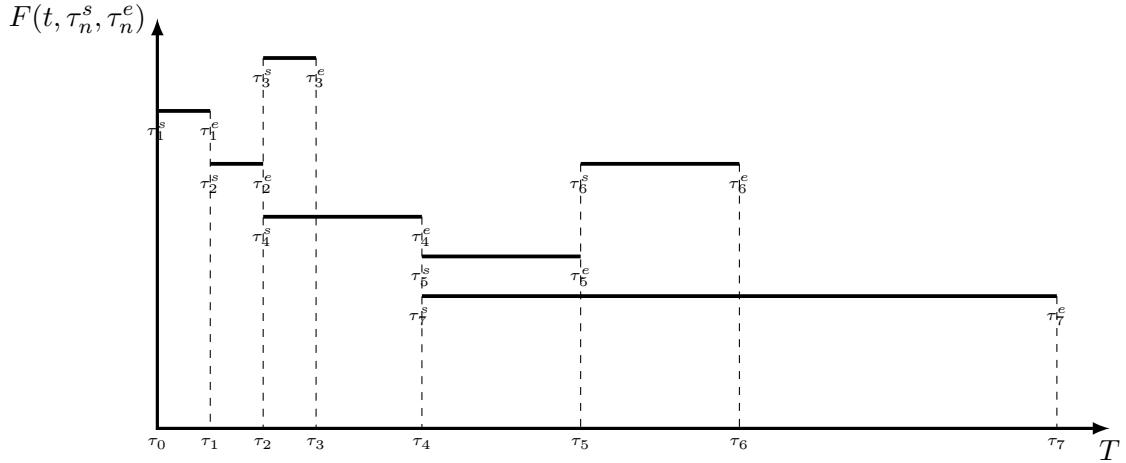


Figure 6: Quoted futures product $F(t, \tau_n^s, \tau_n^e)$.

has been obtained. The final data-set, for the German forward market is shown in Table 3.

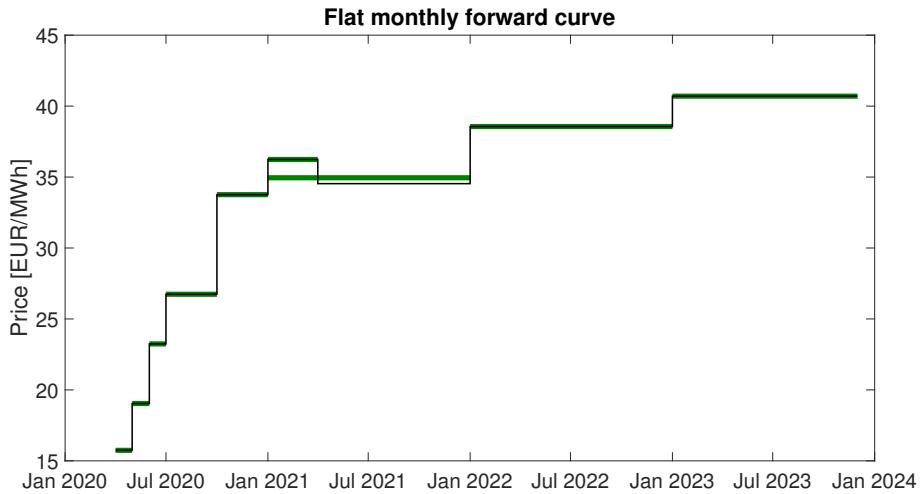


Figure 7: Reconstructed monthly forward curve (black) compared with the quoted forward price for the trading date January 4, 2020 (green). Observe that non arbitrage constraints are satisfied on the year 2021.

5 Calibration

Starting from the following products $M_0, M_1, M_2, Q_1, Q_2, Q_3, Y_1, Y_2$ for each market, by following the procedure depicted in the previous section, we create all the monthly products $F_{M_j}^k(t, T_{M_j})$ for $j = 1, \dots, M$ and $k = 1, \dots, K$, where $M = 24$ and $K = 6$, resulting in a total of $\tilde{N} = 144$ fixed-delivery products. As mentioned earlier, each of these can be viewed as a stochastic factor. Their dynamics in matrix form are outlined by Equation

Trading date	M_0	M_1	M_2	M_3	M_4	M_5	...	M_{22}	M_{23}
2020-01-02	36.05	39.76	37.15	35.50	35.50	35.50	...	43.85	43.85
2020-01-03	38.06	40.40	37.80	36.55	36.55	36.55	...	44.85	44.85
...
2020-03-31	22.49	17.06	19.79	24.14	27.63	27.63	...	39.05	39.05
2020-04-01	15.74	19.04	23.24	27.02	27.02	27.02	...	38.56	38.56
...
2020-12-31	43.52	54.53	53.47	46.59	43.72	43.72	...	49.80	49.80

Table 3: Fixed-delivery products $F_{M_j}^k(t, T_{M_j})$ for the German forward market, reconstructed from products in Table 2.

(7), as follows:

$$\frac{d\mathbf{F}(t, T)}{\mathbf{F}(t, T)} = \boldsymbol{\sigma} \cdot d\mathbf{W}(t),$$

where $\mathbf{W} = (W_1, \dots, W_{\tilde{N}})$ is a standard Brownian motion with $\tilde{N} = M \cdot K$ independent components, \mathbf{F} and $\boldsymbol{\sigma}$ has been defined in Equation (7) and with $T = \min_{m \in [1, M]} T_m$.

$\boldsymbol{\sigma}$ can be efficiently estimated using real market data. To accomplish this, we make the assumption that we have n_{obs} equally spaced daily observations of the process \mathbf{F} at times $t_0 \leq t_1 \leq \dots \leq t_{n_{obs}}$. We define Δt as the time difference between consecutive observations, expressed in fractions of years, for instance, $\Delta t = \frac{1}{260}$. Define the log-return for the product in market k with delivery T_m as:

$$X_i^{k,m} = \ln \frac{F_m^k(t_{i+1}, T_m)}{F_m^k(t_i, T_m)}$$

and assume that the vector $\mathbf{X} = [X^1, X^2, \dots, X^{\tilde{N}}]$ follows a normal distribution with a mean of $\boldsymbol{\mu}$ and a covariance of $\boldsymbol{\Sigma}$: this assumption aligns with our model. Since $\boldsymbol{\Sigma}$ is a covariance matrix it is a symmetric and positive-definite matrix and hence it factorizes as:

$$\boldsymbol{\Sigma} = \mathbf{C}\boldsymbol{\Gamma}\mathbf{C}^T = \mathbf{C}\boldsymbol{\Gamma}^{\frac{1}{2}} \left(\mathbf{C}\boldsymbol{\Gamma}^{\frac{1}{2}} \right)^T$$

for $\boldsymbol{\Gamma} \in \mathbb{R}^{\tilde{N} \times \tilde{N}}$ diagonal matrix and $\mathbf{C} \in \mathbb{R}^{\tilde{N} \times \tilde{N}}$ an orthogonal matrix such that $\mathbf{C}^T \mathbf{C} = \mathbf{I}$, where \mathbf{I} is the identity matrix. On the other hand, we have that $\boldsymbol{\Sigma} = \mathbf{X}^T \mathbf{X}$ and it can be proved that:

$$\boldsymbol{\Sigma} = \Delta t \boldsymbol{\sigma} \boldsymbol{\sigma}^T.$$

Therefore, $\boldsymbol{\sigma}$ can be estimated as:

$$\boldsymbol{\sigma} = \frac{\mathbf{C}\boldsymbol{\Gamma}^{\frac{1}{2}}}{\Delta t^{\frac{1}{2}}}.$$

This procedure is summarized in Algorithm 1.

Algorithm 1 Estimation of σ

- 1: Assume to observe n_{obs} realization of the random vector \mathbf{X} and list them into the matrix $\hat{\mathbf{X}}$.
 - 2: Compute the sample covariance matrix $\hat{\Sigma}$ as $\hat{\Sigma} = \hat{\mathbf{X}}^T \hat{\mathbf{X}}$.
 - 3: Diagonalize the matrix $\hat{\Sigma} = \mathbf{C}\Gamma\mathbf{C}^T$.
 - 4: $\hat{\sigma}$ is given by: $\hat{\sigma} = \mathbf{C}\Gamma^{1/2}\Delta t^{-1/2}$.
-

5.1 PCA for dimension reduction

In this section, we show how PCA can help identify a relatively small number of stochastic factors that drive the entire market. In our experiment, we have K markets, each with the same number of contracts with given maturities $F_{M_j}^k(t, T_{M_j})$, where $k = 1, \dots, K$ and $j = 1, \dots, M$. Consider a matrix with n_{obs} rows and $\tilde{N} = M \cdot K$ columns and compute the matrix of log-returns $\mathbf{X} \in \mathbb{R}^{(n_{obs}-1) \times \tilde{N}}$.

The following proposition is the essence of the PCA.

Theorem 5.1. [Johnson and Wichern [40, Result 8.2]] Let $\mathbf{X} = (X_1, \dots, X_d)$ be a random vector with covariance matrix Σ with eigenvalue-eigenvector pair $(\lambda_1, v_1), \dots, (\lambda_d, v_d)$ where $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_d \geq 0$. Let $Y_i = v_i \mathbf{X}$ for $i = 1, \dots, d$ be the principal components. Then:

$$\sigma_{11} + \dots + \sigma_{dd} = \sum_{i=1}^d \text{var}(X_i) = \lambda_1 + \dots + \lambda_d = \sum_{i=1}^d \text{var}(Y_i)$$

This result states that the total population variance is given by the sum of the eigenvalue λ_i , for $i = 1, \dots, d$. Hence the percentage of the variance explained by the principal component Y_k is given by:

$$\frac{\lambda_k}{\sum_{i=1}^d \lambda_i}, \quad k = 1, \dots, d.$$

The number of principal components k is chosen such that they are enough to explain a sufficiently large percentage of the variance. Usually this choice is made heuristically, namely there is not a rigorous way to chose the number k . See Johnson and Wichern [40] for details.

Once that we have chosen how many principal components to consider, say N , we can define the matrix σ^* as:

$$\sigma^* = \mathbf{C}^* (\Gamma^*)^{1/2} \Delta t^{-1/2},$$

where \mathbf{C}^* is the $\mathbb{R}^{\tilde{N} \times N}$ matrix consisting on the first N eigenvector associated with the eigenvalues $\lambda_1, \dots, \lambda_N$, Γ^* is the diagonal $\mathbb{R}^{N \times N}$ matrix containing N eigenvalues and Δt is the time interval between quotations (usually a day: $\Delta t = 1/260$). Once we have fitted the matrix σ^* we can use N independent stochastic factors to simulate both forward and spot pricing by using Equations (10) and (14).

5.2 PCA analysis

In this section, we use Principal Component Analysis (PCA) to accomplish two objectives: first, to estimate the matrix σ^* , and second, to determine the appropriate number of

stochastic factors needed to effectively capture a significant portion of market variance.

We begin by deriving the log-returns matrix \mathbf{X} from the reconstructed quotations denoted as $F_{M_j}(t, T_{M_j})$. As discussed in Section 4.2, it is important to address potential spurious fluctuations, termed as “fake spikes”, in the log-returns due to the rolling mechanism. Additionally, assuming Gaussian distribution of log-returns, we employ a filtering method, inspired by the approach proposed by Cartea and Figueroa [23], to remove outliers. Briefly, during the first iteration of the filtering procedure, returns with absolute values greater than three times the standard deviation of the returns of the series at that specific iteration are removed. On the second iteration, the standard deviation of the remaining series (stripped from the first filtered returns) is again calculated; those returns which are now greater than three times this last standard deviation are filtered again. The process is repeated until no further returns can be filtered.

After preparing the matrix \mathbf{X} , we conduct a PCA analysis to identify an adequate number of stochastic factors that sufficiently account for a significant portion of the overall market variance. Research by Feron and Gruet [30], Geman [33], and Koekebakker and Fridthjof [43] suggest that in power markets, 5-10 factors are typically sufficient to explain a substantial portion of the variance, while in other markets like metals, only 4 factors may be necessary. We use PCA on the matrix \mathbf{X} , which contains the log-returns of each $F_{M_j}^k(t, T_{M_j})$ for $k = 1, \dots, K$ and $j = 1, \dots, M$. In Figure 8, we depict the cumulative percentage variance against the number of components. Consistent with the previous discussion, it is evident that approximately ten principal components are sufficient to explain nearly 95% of the total variance.

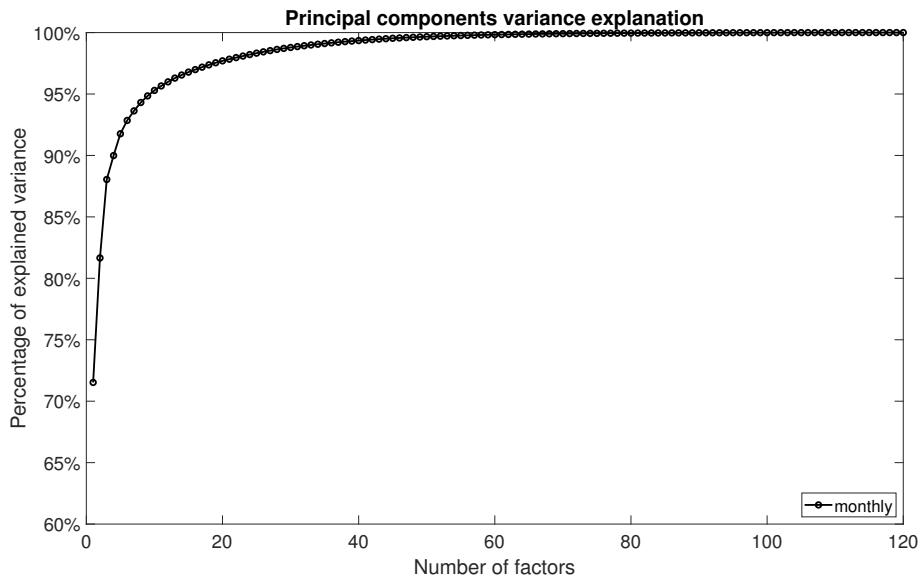


Figure 8: Variance explained from the principal components of the monthly data.

It is also worth considering whether the introduction of synthetic monthly products $F_{M_j}^k(t, T_{M_j})$ might influence the determination of the number of factors during PCA. Intuitively, as we only derived monthly products from those with a coarser granularity, we should not expect to introduce additional information and variability compared to the

scenario where we work with raw data from Table 2. To verify this assertion, we conduct PCA using both raw and monthly data. As depicted in Figure 9 (where we truncated the x axis at 48, as the details beyond are not pertinent), we observe that the same number of stochastic factors explains a similar percentage of variance in both cases, despite the monthly data containing a larger original total number of forward products than the raw data.

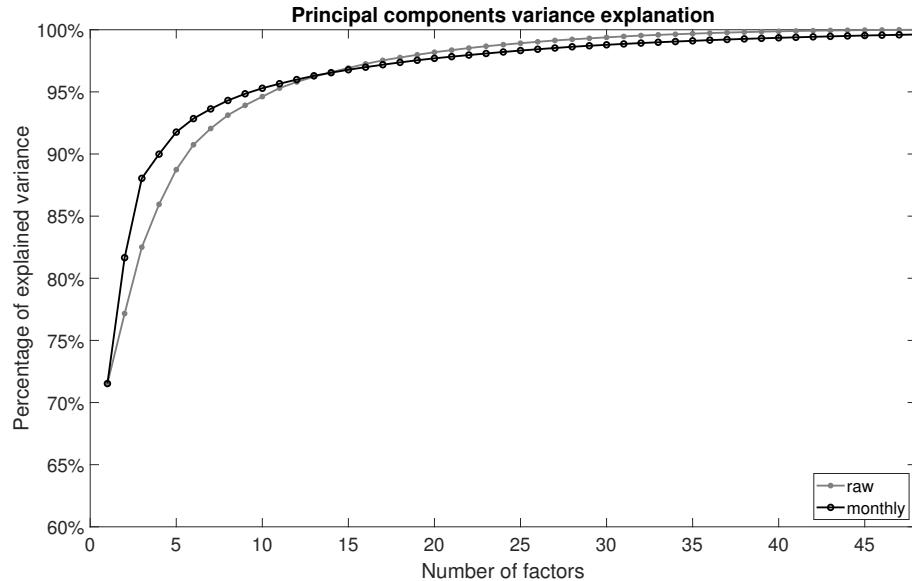


Figure 9: Percentage of the total variance explained by principal components on the raw and on the monthly data.

Given the preceding results, we can assert that the volatility parameter calibration using the outlined PCA procedure can be applied to both raw data ($F_{Y_j}^k(t, T_{Y_j})$, $F_{Q_j}^k(t, T_{Q_j})$, $F_{M_j}^k(t, T_{M_j})$) and reconstructed monthly forward contracts $F_{M_j}^k(t, T_{M_j})$. The latter approach is preferable in situations where commodities with varying granularity are present. In such cases, working with products of uniform granularity facilitates clearer analysis and simpler coding.

Through these simulations, we can compute correlation surfaces in simulated log-returns and compare them with historical data. A comparison between Figure 10 and Figure 11 reveals that the correlation structure is accurately replicated by the proposed model. However, due to the use of only 10 stochastic factors instead of the original 144, the correlation surface computed from the simulations appears smoother than the original. Specifically, the simulation-derived correlation surface may appear flat for long maturities. This occurs because PCA selects a single stochastic factor to influence the entire long-term volatility structure. This observation aligns with empirical evidence, as products with longer maturities tend to be strongly correlated and exhibit similar oscillatory behavior.

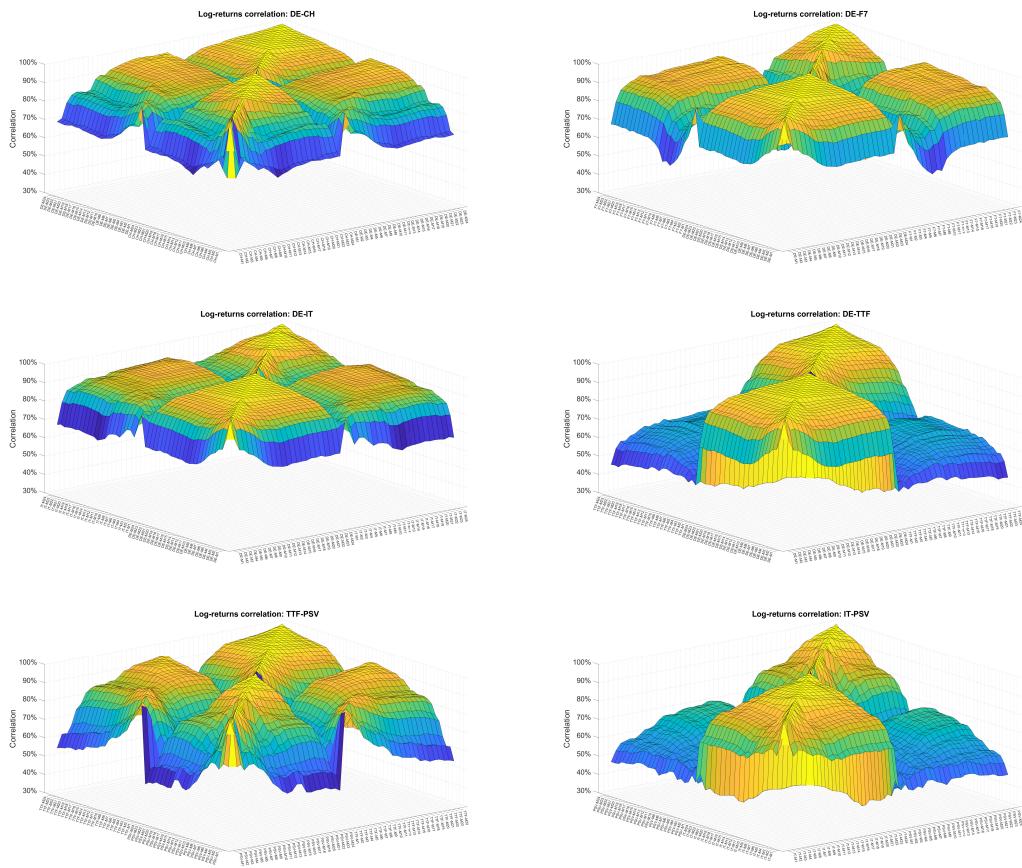


Figure 10: Historical daily log-returns correlation surfaces for different commodities.

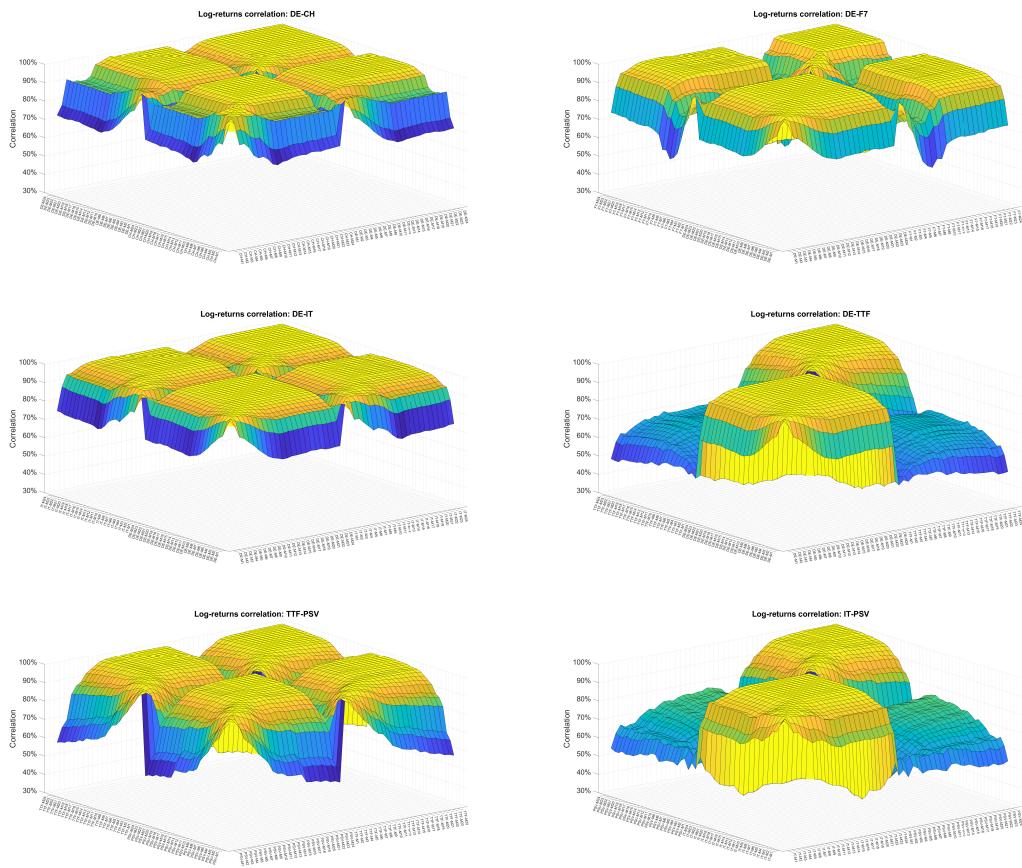


Figure 11: Simulated daily log-returns correlation surfaces for different commodities.

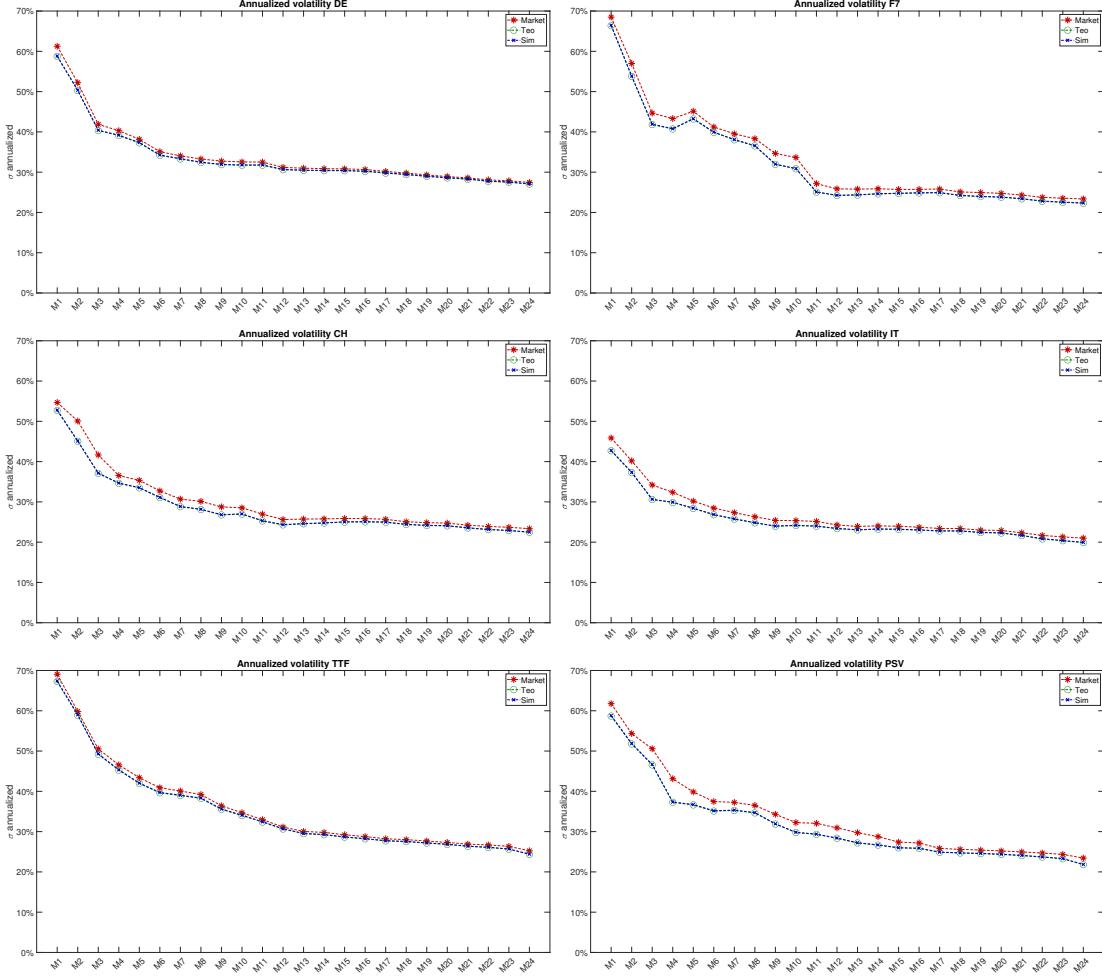


Figure 12: Simulated, market and theoretical annualized volatility.

In Figure 12, we compare the annualized historical market volatility with that of forward simulations. The theoretical volatility is computed using Equation (12), while the simulated volatility is determined by calculating the standard deviation of log-return simulations. A comparison between the theoretical and simulated volatilities indicates that the model has been accurately implemented. Additionally, by comparing the market volatility with that derived from the model, we can verify the model's effectiveness in capturing the market's volatility structure.

6 Numerical simulations and derivatives pricing

In this section, we apply the proposed model to address several significant challenges common in energy markets, namely price simulations, computation of risk metrics, and derivatives pricing. Initially, we present a series of paths generated by our model and explore how simulations of forward products can facilitate the computation of portfolio value at risk. Subsequently, we delve into derivatives pricing, employing Monte Carlo

simulations to price various exotic products commonly traded over-the-counter (OTC) in energy markets, such as virtual power plants, swing options, and gas storages.

6.1 Price simulations and risk metrics

Once the model has been calibrated using futures market prices $F(t, \tau^s, \tau^e)$ spanning from January 1, 2020, to December 31, 2020, we are prepared to simulate the forward curve structure. Assuming $t_0 = 04/01/2021$, according to Equation (10), we simulate futures prices $F^k(t, \tau^s, \tau^e)$ for all markets considered ($k = 1, \dots, K$) with delivery period $\tau^s = 01/01/2022$ and $\tau^e = 31/12/2022$ from t_0 up to maturity $T = 31/12/2021$. The results are depicted in Figure 13. The price dynamics exhibit a notably high level of interdependence. In particular, both power and natural gas prices demonstrate a tendency to move in tandem throughout the entire simulation period, as one would expect in reality.

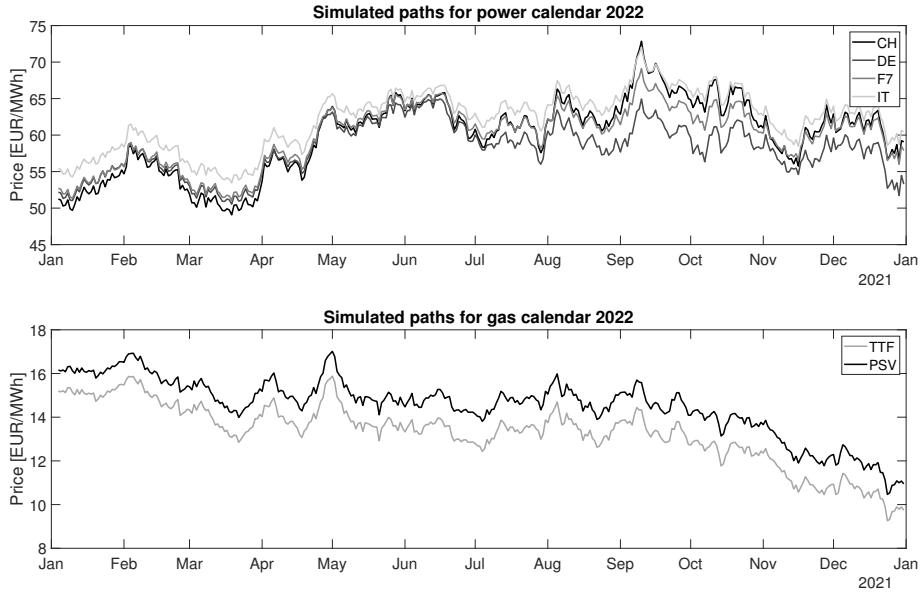


Figure 13: Sample paths for the product calendar 2022 with delivery date from 01/01/2022 to 31/12/2022.

In many cases, practitioners seek to compute risk metrics (such as VaR, TVaR, CVaR, and others) for a given portfolio. Various approaches, including historical, parametric, and Monte Carlo methods, among others, are available for this task. If a Monte Carlo approach is selected, simulations of the forward curve must be conducted. To achieve this, we use Equation (11) to simulate the forward curve after a few days, typically one or two. The results are depicted in Figure 14. In particular, we observe that the expected value of the simulated forward curves (μ) converges to today's forward curve $F(t_0, T)$, as expected. The fifth and ninety-fifth percentiles, highlighted in red, provide insight into the range of the simulations. Once simulations are available, the desired risk metric can be computed. If non-linear derivatives are present, the Delta-Gamma approximation can be employed, as proposed by Glasserman [34].

A careful reader may notice that we employ a model formulated under the risk-neutral probability measure \mathbb{Q} to compute risk metrics within the real-world probability measure \mathbb{P} .

Under \mathbb{P} , forward prices may exhibit seasonality, trends, risk premia, and mean-reverting dynamics. In such cases, one could develop the Heath-Jarrow-Morton (HJM) framework under \mathbb{P} and calibrate the drift under this measure, as suggested by Broszkiewicz-Suwa and Weron [19]. Alternatively, models with different price dynamics, including mean reversion, have been proposed by Latini et al. [45] and Benth et al. [11], among others.

Moreover, it is worth noting that the PCA method explains only a portion of the market variance. This could potentially result in an underestimation of risk metrics, which needs to be considered in practical applications. For instance, it may be prudent to adopt a conservative approach when setting the VaR limit for the company. However, despite these limitations, the framework presented here enables the evaluation of a multi-commodity portfolio using a comprehensive Monte Carlo approach, with reasonable computational effort.

Being able to compute risk-metrics on forward portfolio is crucial in practical application: indeed, even if many OTC contracts have as underlying asset the spot price they are very illiquid and can be hedged only by using forward products: hence selling a derivatives on spot prices automatically generates an exposition to forward prices which risk must be properly quantified.

On the other hand, derivatives on forward prices often encompass vanilla European options or exchange options, for which analytical formulas are readily available. In contrast, more exotic derivatives like American options or Asian options necessitate pricing through a Monte Carlo approach for accurate evaluation. This can be effectively achieved by utilizing forward prices path simulations which can be obtained by Equation (10).

Finally, we examine the spot prices generated by the HJM framework. To streamline our discussion, we concentrate solely on two markets: the German and Italian power spot markets. In Figure 15, we present a single realization of the stochastic process $S^k = \{S^k(t); t \in [t_0, T]\}$ for the two distinct commodities, in accordance with the dynamics outlined in Equation (14). Notably, the spot prices demonstrate a tendency to move in tandem, as they are influenced by the same stochastic factors. In this scenario, the correlation level in daily spot log-returns is notably high, approximately $\rho = 0.98$. This correlation is further illustrated in Figure 16, where we present a contour plot of a bivariate Gaussian distribution fitted to daily log-returns. Such a correlation is derived from futures prices, which tends to be higher than the historical correlation observed in the spot market, which is not considered within this model framework.

It is important to note that many European power spot markets are interconnected or “coupled”, meaning that market coupling optimizes the allocation of cross-border capacities between countries. One potential effect of this coupling on price dynamics is that the electricity price in two different countries can be the same at the same hour. However, the proposed model does not account for this behavior. To incorporate such effects, a fundamental component must be introduced, as suggested by Carmona and Coulon [22] and Kiesel and Kusterman [41].

Once Monte Carlo simulations are available both for power and futures products derivatives, the pricing of complex energy financial claims, such as virtual power plants, swing options or gas storage might be easily performed following the algorithms presented by

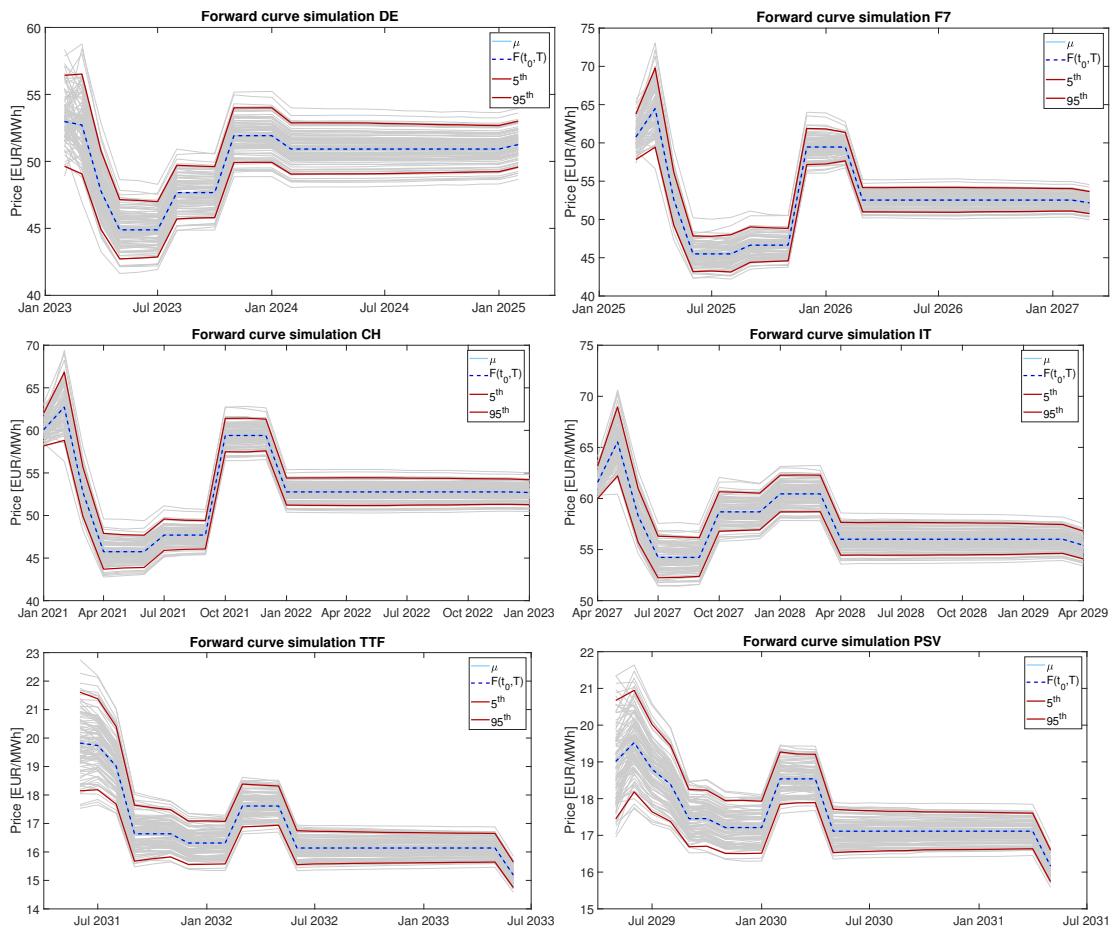


Figure 14: Simulated two days forward curves: $n_{sim} = 400$.

Tseng and Barz [66], Dörr [27], Boogert and De Jong [15] respectively, as will be sketched in the next section.

6.2 Virtual power plant pricing

In this section, we use Monte Carlo simulation to price a virtual power plant (VPP), building upon the approach introduced by Tseng and Barz [66]. A VPP represents a financial contract designed to emulate the operational dynamics of a power plant, converting a fuel source, typically natural gas, into electricity. In this example, we consider hourly markets for both electricity and gas, focusing specifically on the Italian power and gas markets.

The heat rate of a plant, denoted as H , signifies the conversion efficiency between electricity and fuel. Assuming H is predetermined, and representing the electricity and fuel prices as $S(t)^E$ and $S(t)^F$ respectively, the payoff for each generated megawatt-hour (MWh) of electricity can be expressed as follows:

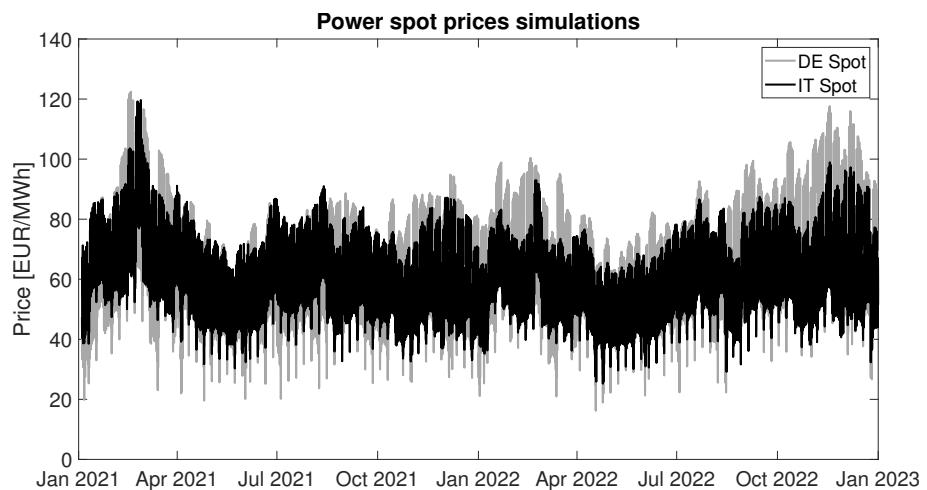
$$\Phi(t) = \max(S(t)^E - H \cdot S(t)^F, 0).$$

We can decide to turn on the plant if $S(t)^E > H \cdot S(t)^F$. For this reason Hsu [39] proposes a VPP value over a time period $[0, T]$:

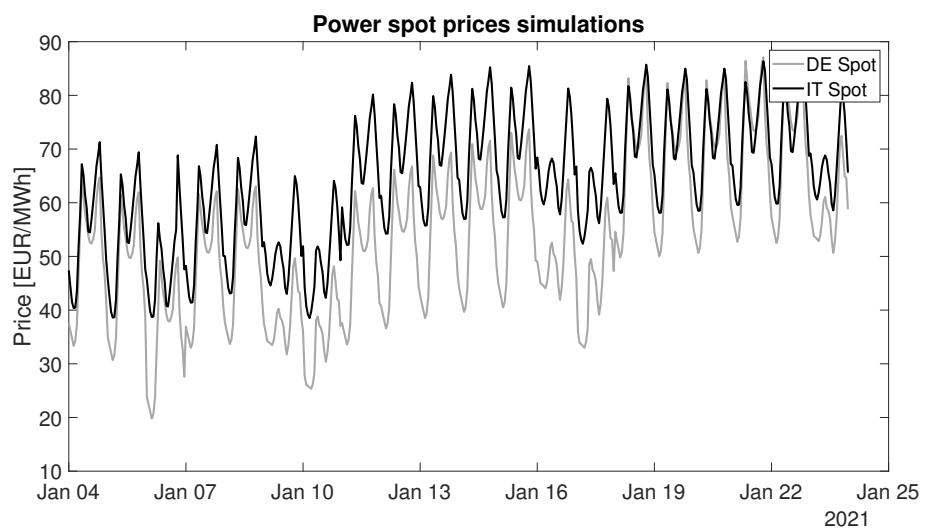
$$\sum_{t=1}^T \mathbb{E}^Q [e^{-rt} \max(S(t)^E - H \cdot S(t)^F, 0)],$$

so that the value of the VPP can be estimated by a series of European spark-spread call options. In reality, a power plant operates under various physical constraints, including non-zero startup/shutdown costs, which are typically incorporated into virtual power plant contracts. Failing to account for these physical constraints may result in an overestimation of the VPP contract, making it essential to include them from a modeling perspective. In this example, we consider only a subset of these constraints. For further details, we recommend referring to Tseng and Barz [66]. In particular we consider:

- t^{on} : the minimum number of hours the plant must remain on after it was turned on.
- t^{off} : the minimum number of hours the plant must remain off after it was turned off.
- $q(t)$: the amount of electricity generated at time t .
- q^{min} : minimum rated capacity of the unit.
- q^{max} : maximum rated capacity of the unit.
- S_u : startup cost associated with turning on the unit.
- S_d : shutdown cost associated with turning off the unit.
- H : it is considered constant and not a function of $q(t)$.



(a) Spot prices simulations.



(b) Zoom spot prices simulations.

Figure 15: German and Italian power spot prices simulations.

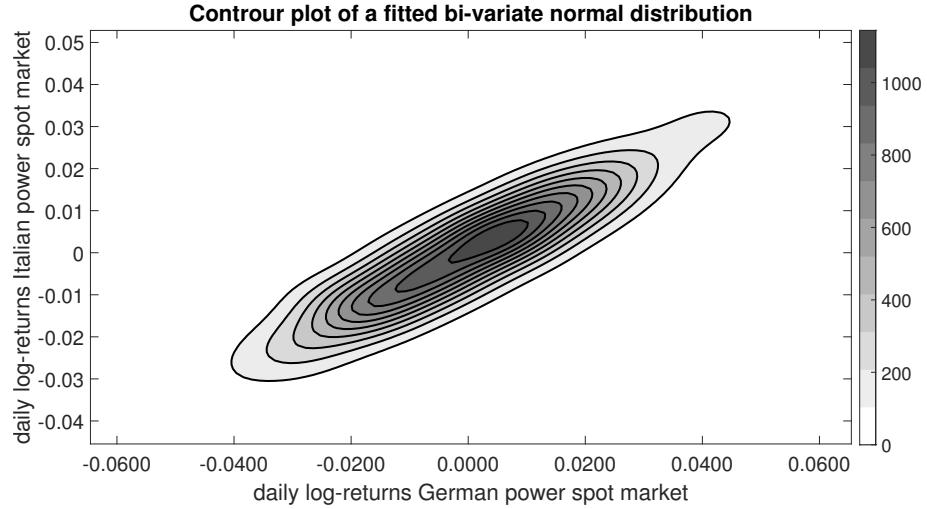


Figure 16: Contour plot of a bi-variate normal distribution fitted on daily log-returns of German and Italian power spot prices simulation. Correlation in log-returns $\rho = 0.98$.

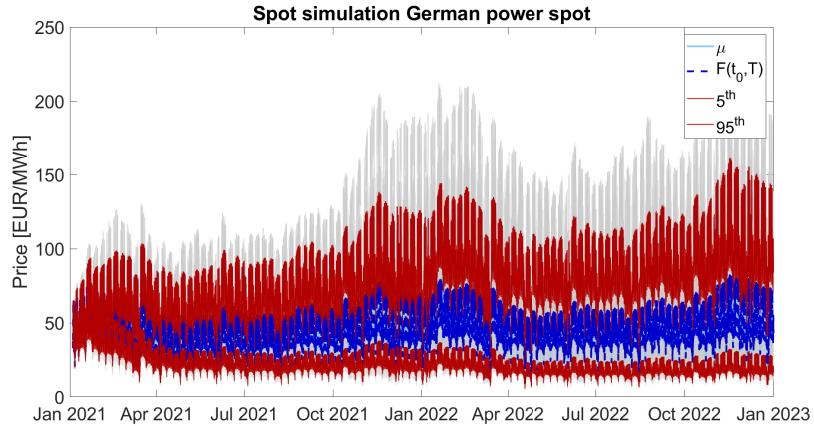


Figure 17: Sample paths for the German power spot price.

The problem to be solved takes the following form:

$$\sup_{q(t)} \mathbb{E} \left[\sum_{t=1}^T q(t) e^{-rt} \max(S(t)^E - H \cdot S(t)^F, 0) \right],$$

subject to the constraints $q^{min} \leq q(t) \leq q^{max}$ for $t \in [0, T]$. The power plant must operate for at least t^{on} hours when it is switched on and for at least t^{off} hours when it is switched off, taking into account startup and shutdown costs.

To incorporate unit constraints, the problem can be formulated as a multistage stochastic problem, which can be solved using a backward stochastic dynamic programming (BSDP) approach, such as the one proposed by Longstaff and Schwartz [46]. In this case, no closed pricing formulas are available. We compute the price of a VPP contract with the parameters listed in Table 4. Additionally, we determine the value of the contract

for various values of t^{on} and t^{off} . We anticipate that as the contract loses flexibility, its value decreases.

Furthermore, to ensure the coherence of the pricing, we compute an upper bound for the contract value using a strip of European call options on the spark-spread. We also employ a “Naive Monte Carlo” evaluation, where we calculate the optimal behavior of the virtual power plant path by path: the VPP value is obtained by averaging all the optimal payoffs. This approach is expected to yield higher values than those obtained by the BSDP method.

The results, presented in Table 18 and Figure 18, confirm these expectations. The computational time increases as t^{on}/t^{off} increases, as the state variable space, required to implement the algorithm, expands with t^{on}/t^{off} , leading to a slowdown in the solution algorithm.

Parameters	Value
τ^s	October 1, 2023
τ^e	October 31, 2023
t^{on}	4, 16, 24, 54, 96, 124, 160
t^{off}	4, 16, 224, 54, 96, 124, 160
q^{min}	180 (MW)
q^{max}	360 (MW)
S_u	2000 (EUR)
S_d	7000 (EUR)
H	40%

Table 4: VPP contract parameters for a delivery period $[\tau^s, \tau^e]$ from October 1, 2023 to October 31, 2023.

$t^{on} (t^{off})$	VPP value (EUR)	Computational Time (s)
1 (1)	$88.02 \cdot 10^6$	0.21
4 (4)	$86.95 \cdot 10^6$	14.93
16 (16)	$85.57 \cdot 10^6$	22.15
24 (24)	$84.11 \cdot 10^6$	33.67
54 (54)	$81.03 \cdot 10^6$	51.71
96 (96)	$75.79 \cdot 10^6$	96.12
124 (124)	$72.45 \cdot 10^6$	132.17
160 (160)	$68.16 \cdot 10^6$	186.42

Table 5: VPP evaluation varying t^{on} and t^{off} .

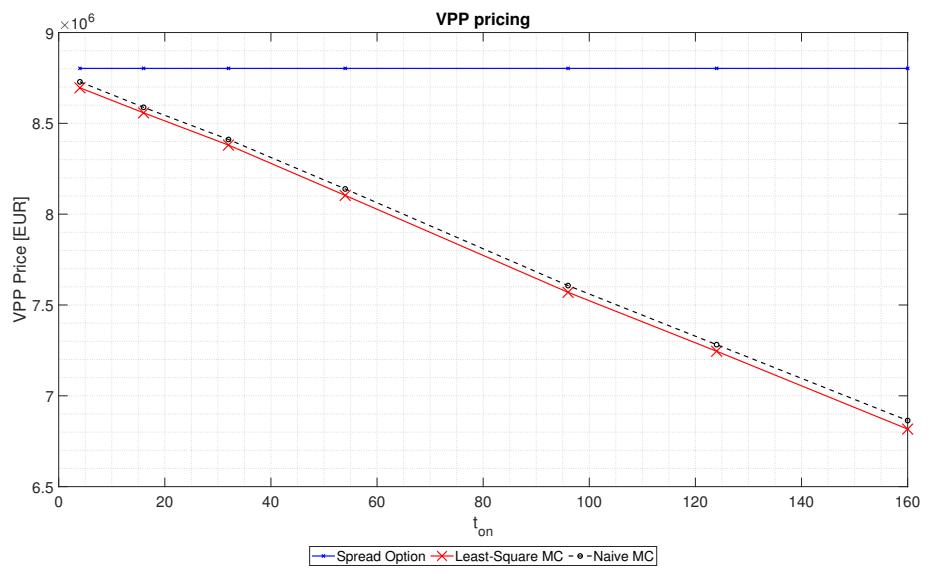


Figure 18: Example of the value of a VPP varying the parameters t^{on} and t^{off} .

6.3 Swing option pricing

In this section, we evaluate a swing option using the algorithm proposed by Dörr [27]. We consider the daily spot price of Dutch natural gas (TTF), denoted as $S(t)$, for a specified delivery period, for example, from October 1, 2023, to October 31, 2023.

We assume that the owner of the swing option can exercise a maximum of u_{max} daily upswings, obtaining a payoff of $(S(t) - K)^+$, and a maximum of d_{max} daily downswings, also resulting in a payoff of $(S(t) - K)^+$, where K represents a predetermined strike price. Typically, these payoffs are multiplied by a specified quantity Q , which we assume to be unity in this example.

We further assume that the strike price K remains the same for both upswings and downswings, although it is possible to use different strike prices. Essentially, this type of contract comprises the summation of multiple exercise American put and call options.

We recall that the value of an American put option P^{Am} with strike K and maturity T can be computed as:

$$P^{Am}(K, T) = \sup_{t \leq \tau \leq T} \mathbb{E}^{\mathbb{Q}} \left[e^{-r(T-t)} (K - S(\tau)) \right],$$

as shown in Bjork [12, Chapter 21].

If we assume that the option can be exercised daily, a straightforward lower bound for this contract is as follows:

$$lb_{swing} = C^{Am}(K, T) + P^{Am}(K, T),$$

where $C^{Am}(K, T)$ and $P^{Am}(K, T)$ are the prices of an American Call and American Put option with strike price K and maturity T equal to the end of the delivery period. On the other hand a natural upper bound for the contract is given by:

$$ub_{swing} = \sum_{t=1}^T C^{Eu}(K, t) + P^{Eu}(K, t),$$

where $C^{Eu}(K, t)$ and $P^{Eu}(K, t)$ are the prices of an European Call and European Put option with strike price K and maturity t , for all t from now to the end of the delivery period. The value of the swing contract can be determined by extending the algorithm introduced by Longstaff and Schwartz [46], as proposed by Dörr [27]. We utilize the parameters listed in Table 6, and the results of the pricing, varying the maximum number of up-downswings (u_{max} and d_{max}), are presented in Table 7.

Figure 19 illustrates how the value of the swing option increases with higher values of u_{max} and d_{max} . Notably, the option's value falls within the range defined by the two proposed bounds. Specifically, when the number of up-downswings equals the number of days in the delivery period, the upper bound is attained.

Parameters	Value
τ^s	October 1, 2023
τ^e	October 31, 2023
u_{max}	1, ..., 31
d_{max}	1, ..., 31
K	39.83 (EUR/MWh)

Table 6: Swing contract parameters for a delivery period $[\tau^s, \tau^e]$ from October 1, 2023, to October 31, 2023.

u_{max} (d_{max})	Swing value (EUR)	Computational Time (s)
1 (1)	19.43	0.06
5 (5)	89.54	0.26
10 (10)	169.06	0.90
15 (15)	239.46	2.11
20 (20)	298.67	3.88
25 (25)	351.58	6.21
30 (30)	394.74	9.13

Table 7: Swing evaluation varying u_{max} and d_{max} .

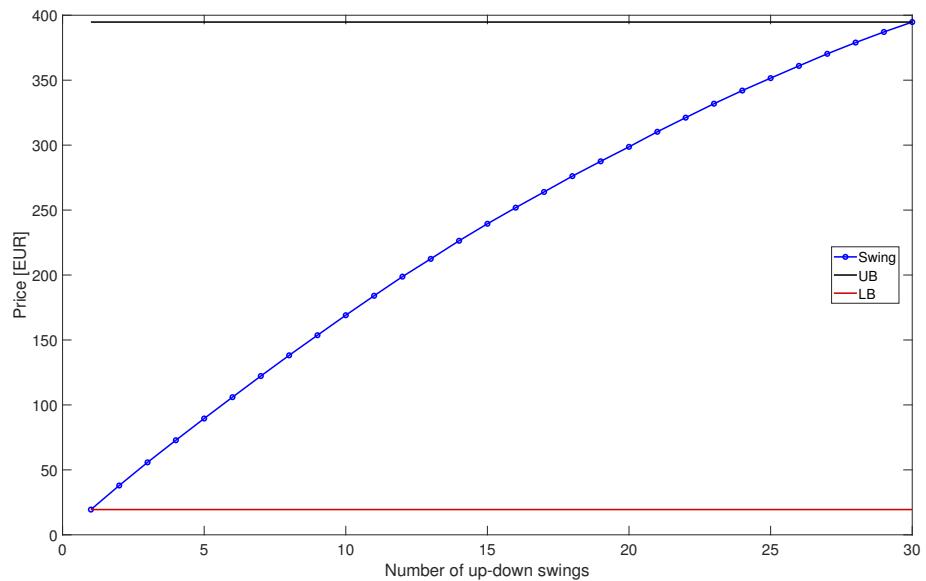


Figure 19: Example of the value of a swing contract varying the number of u_{max} and d_{max} .

6.4 Gas storage evaluation

More structured energy contracts, such as those examined by Boogert and De Jong [15] or Kluge [42], can be priced using similar Monte Carlo techniques. In this section, we adopt the approach proposed by Boogert and De Jong [15] to price a gas storage contract.

Viewing the storage contract from the perspective of the contract holder, we assume it is initiated at time $t = 0$ and settled at time $t = T + 1$. Each day, the holder can opt to inject gas, take no action, or withdraw gas, with a volume denoted as $\Delta v(t)$, subject to certain volumetric constraints. In this example, the payoff at time t is assumed to be:

$$h(S(t), \Delta v(t)) = \begin{cases} -S(t)\Delta v(t) & \text{inject} \\ 0 & \text{do nothing} \\ S(t)\Delta v(t) & \text{withdrawn} \end{cases}$$

Additionally, we assume that the volume of the storage at time t satisfies $v_{min} \leq v(t) \leq v_{max}$, and that injection or withdrawal is subject to $i^{min} \leq \Delta v(t) \leq i^{max}$. These constraints can be easily adjusted over time. Furthermore, a constraint on the terminal value of the volume $v(T + 1)$ can be imposed by introducing a “penalty function” $q(S(T + 1), v(T + 1))$.

The value of the storage contract $U(0, S(0), v(0))$ is the expected value of the accumulated future payoff $h(S(t), \Delta v(t))$ under the most optimal strategy (or policy) $\pi = \{\pi(1, S(1), v(1)), \dots, \pi(T, S(T), v(T))\}$, namely:

$$U(0, S(0), v(0)) = \sup_{\pi} \mathbb{E} \left[\sum_{t=0}^T e^{-rt} h(S(t), \Delta v(t)) + e^{-r(T+1)} q(S(T + 1), v(T + 1)) \right]. \quad (15)$$

This problem can be addressed using standard backward dynamic programming (BSDP) techniques in a stochastic setting (refer to Busoniu et al. [21]). Computational efficiency can be enhanced by employing a Least-Squares approach, as described in Boogert and De Jong [15].

In our example, we examine a storage contract based on TTF natural gas with a delivery period spanning the entirety of 2024. The parameters are listed in Table 8.

Parameters	Value
τ^s	January 1, 2024
τ^e	December 31, 2024
v_{min}	0 (MWh)
v_{max}	250,000 (MWh)
v_0	100,000 (MWh)
v_{T+1}	100,000 (MWh)
i_{min}	-8,000 (MWh)
i_{max}	8,000 (MWh)

Table 8: Gas storage contract parameters for a delivery period $[\tau^s, \tau^e]$ from January 1, 2024, to December 31, 2024.

In Table 9, we present the gas storage price evaluated using 2500 simulations. The *deterministic* approach involves maximizing the payoff path by path and subsequently

averaging the results. This approach addresses the following problem:

$$\mathbb{E} \left[\sup_{\pi} \sum_{t=0}^T e^{-rt} h(S(t), \Delta v(t)) + e^{-r(T+1)} q(S(T+1), v(T+1)) \right]$$

and clearly tends to overvalue the contract as it assumes perfect foresight of the future. The accurate value of the contract is obtained by solving Equation (15) using the BSDP. We observe that the contract's value is smaller than in the deterministic case. Additionally, following the methodology outlined in Langville and Stewart [44], we conduct an out-of-sample test of the optimal policy π . It is observed that the contract's value, priced on the out-of-sample paths, closely approximates the value obtained by the solution of Equation (15). This suggests that the pricing algorithm is effective, as it yields out-of-sample values that closely match the in-sample values for the option.

In Figure 20, deterministic and stochastic optimal trajectories of the storage level are depicted, related to the realized price process $S(t)$.

Approach	Swing value (EUR)	Computational Time (s)
Deterministic	19,826,114	25,920.02
SDP	7,457,729	1,020.77
SDP (out sample)	7,462,901	23.04

Table 9: Gas storage evaluation using different pricing methods.

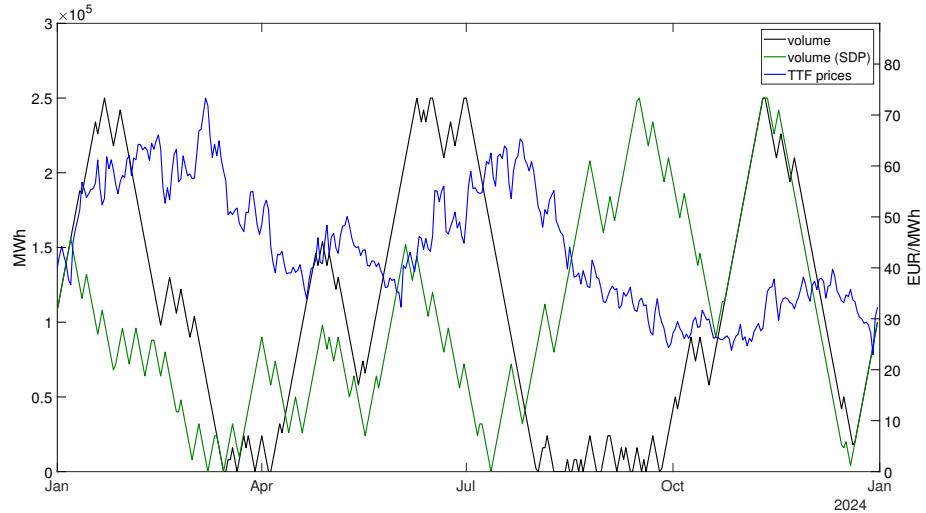


Figure 20: Gas storage volume for a simulation of TTF prices computed using the deterministic approach and the BSDP. Notice how the deterministic approach (in black) maximizes the value extracted from the contract, as storage is filled during low-price periods and emptied during high-price ones.

7 Conclusions

In this article, we extensively explored the implementation and potential applications of the Heath-Jarrow-Morton framework in energy markets, with a specific focus on the European power and natural gas markets. We introduced risk-neutral dynamics for fixed-delivery forward prices $F_{M_j}(t, T_{M_j})$ and derived the dynamics of spot $S(t)$ and futures prices $F(t, \tau^s, \tau^e)$. By demonstrating the interdependence of power and natural gas futures markets, we employed the PCA algorithm to select a few stochastic factors, namely independent Brownian motions, which explain a significant portion of the market variance.

Additionally, we applied the model to real market data from European power and natural gas markets, discussing the daily log-returns correlation structures and validating the model's narrow fit to the market.

Moreover, we analyzed the futures and spot simulations generated by the model and we used Monte Carlo simulation to compute risk metrics and to evaluate virtual power plants, swing options and storage contracts.

The Heath-Jarrow-Morton model offers ease of implementation and rapid calibration, owing to the assumption of log-returns normality and enables fast simulations. However, it does have limitations that must be considered when using the outputs for risk-metric computation or derivative pricing. Log-returns are normally distributed, which may not fully capture market characteristics such as jumps, volatility smiles, and clustering. To address these limitations, extensions involving stochastic volatility or jump processes can be explored, albeit with increased computational complexity.

Furthermore, calibration based on historical futures prices may not fully replicate vanilla options quoted in the market, particularly in illiquid power option markets. The model also inherits correlation and volatility in simulated spot log-returns directly from the forward prices due to the lack of consideration of spot quotations during calibration.

Despite these drawbacks, the proposed framework remains one of the simplest methods for simulating dependent futures and spot prices in a multi-commodity setting, widely adopted by practitioners and serving as a benchmark for testing innovative approaches. Future research avenues may involve incorporating jumps in prices' dynamics or exploring stochastic volatility approaches in the multi-dimensional framework, while ensuring both mathematical and numerical tractability. This could pave the way for advancements in both academia and industry.

Compliance with Ethical Standard

The authors have no relevant financial or non-financial interests to disclose.

References

- [1] D. Applebaum. *Lévy Processes and Stochastic Calculus*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2 edition, 2009. doi: 10.1017/CBO9780511809781.
- [2] F. Asche, P. Osmundsen, and M. Sandsmark. The UK Market for Natural Gas, Oil and Electricity: Are the Prices Decoupled? *The Energy Journal*, 27(2):27–40, 2006. ISSN 01956574, 19449089. URL <http://www.jstor.org/stable/23297017>.
- [3] L.J. Bachmeier and J. M. Griffin. Testing for Market Integration Crude Oil, Coal, and Natural Gas. *The Energy Journal*, 27(2):55–71, 2006. ISSN 01956574, 19449089. URL <http://www.jstor.org/stable/23297019>.
- [4] L. Ballotta and E. Bonfiglioli. Multivariate Asset Models Using Lévy Processes and Applications. *The European Journal of Finance*, 13(22):1320–1350, 2013.
- [5] O.E. Barndorff-Nielsen. Processes of Normal Inverse Gaussian Type. *Finance and Stochastics*, 2(1):41–68, 1998.
- [6] D. S. Bates. Jumps and Stochastic Volatility: Exchange Rate Processes Implicit in Deutsche Mark Options. *Review of Financial Studies*, 9(1):69–107, 1996.
- [7] C. Bencivenga, G. Sargent, and R. L. D’Ecclesia. Energy Markets: Crucial Relationship Between Prices. In Marco Corazza and Claudio Pizzi, editors, *Mathematical and Statistical Methods for Actuarial Sciences and Finance*, pages 23–32, Milano, 2010. Springer Milan. ISBN 978-88-470-1481-7.
- [8] F. E. Benth and S. Koekebakker. Stochastic Modeling of Financial Electricity Contracts. *Energy Economics*, 30(3):1116–1157, May 2008. URL <https://ideas.repec.org/a/eee/eneeco/v30y2008i3p1116-1157.html>.
- [9] F.E. Benth and J. Saltyte-Benth. The Volatility of Temperature and Pricing of Weather Derivatives. *Quantitative Finance*, 7(5):553–561, 2007. doi: 10.1080/14697680601155334.
- [10] F.E. Benth, J.S. Benth, and S. Koekebakker. *Stochastic Modeling of Electricity and Related Markets*. Number 6811 in World Scientific Books. World Scientific Publishing Co. Pte. Ltd., July 2008. ISBN ARRAY(0x50724998). URL <https://ideas.repec.org/b/wsi/wsbook/6811.html>.
- [11] F.E. Benth, M. Piccirilli, and T. Vargiolu. Mean-reverting Additive Energy Forward Curves in a Heath–Jarrow–Morton Framework. *Mathematics and Financial Economics*, 13:543–577, 2019.
- [12] T. Bjork. *Arbitrage Theory in Continuous Time*. Number 9780199574742 in OUP Catalogue. Oxford University Press, 2009. ISBN ARRAY(0x4dc87ac8).
- [13] F. Black. The Pricing of Commodity Contracts. *Journal of Financial Economics*, 3(1):167–179, 1976. ISSN 0304-405X. doi: <https://doi.org/10.1016/>

0304-405X(76)90024-6. URL <https://www.sciencedirect.com/science/article/pii/0304405X76900246>.

- [14] F. Black and M. Scholes. The Pricing of Options and Corporate Liabilities. *Journal of Political Economy*, 81(3):637–654, 1973.
- [15] A. Boogert and C. De Jong. Gas Storage Valuation Using a Monte Carlo Method. (0704), January 2007. URL <https://ideas.repec.org/p/bbk/bbkfep/0704.html>.
- [16] B. Bosco, L. Parisio, M. Pellegatti, and F. Baldi. Long-run Relations in European Electricity Prices. *Journal of Applied Econometrics*, 25(5):805–832, 2010. ISSN 08837252, 10991255. URL <http://www.jstor.org/stable/40865175>.
- [17] D. Brigo and F. Mercurio. *Interest-Rate Models: Theory and Practice*. Springer Finance, 2001. ISBN 9783540417729.
- [18] D. Brigo, A. Dalessandro, Neugebauer M., and F. Triki. A Stochastic Process Toolkit for Risk Management. Technical report, 2007.
- [19] E. Broszkiewicz-Suwał and A. Weron. Calibration of the Multifactor HJM Model for Energy Market. HSC Research Reports HSC/05/03, Hugo Steinhaus Center, Wrocław University of Technology, 2005. URL <https://EconPapers.repec.org/RePEc:wuu:wpaper:hsc0503>.
- [20] B. Buchmann, K. Lu, and D. Madan. Calibration for Weak Variance-Alpha-Gamma Processes. *Methodology and Computing in Applied Probability*, 21(4), 2019. doi: 10.1007/s11009-018-9655-y.
- [21] L. Busoniu, R. Babuska, B. Schutter, and D. Ernst. *Reinforcement Learning and Dynamic Programming Using Function Approximators*. CRC Press, Inc., USA, 1st edition, 2010. ISBN 1439821089.
- [22] R. Carmona and M. Coulon. *A Survey of Commodity Markets and Structural Models for Electricity Prices*, pages 41–83. Springer New York, New York, NY, 2014. ISBN 978-1-4614-7248-3. doi: 10.1007/978-1-4614-7248-3_2. URL https://doi.org/10.1007/978-1-4614-7248-3_2.
- [23] A. Cartea and M. Figueroa. Pricing in Electricity Markets: a Mean Reverting Jump Diffusion Model with Seasonality. *Applied Mathematical Finance, No. 4, December 2005*, 12(4):313–335, 2005.
- [24] L. Clewlow and C. Strickland. Valuing Energy Options in a One Factor Model Fitted to Forward Prices. Research Paper Series 10, Quantitative Finance Research Centre, University of Technology, Sydney, April 1999. URL <https://ideas.repec.org/p/uts/rpaper/10.html>.
- [25] L. Clewlow and C. Strickland. A Multi-Factor Model for Energy Derivatives. (28), December 1999. URL <https://ideas.repec.org/p/uts/rpaper/28.html>.
- [26] R. Cont and P. Tankov. *Financial Modelling with Jump Processes*. Chapman and Hall, 2003.

- [27] U. Dörr. Valuation of Swing Options and Examination of Exercise Strategies by Monte Carlo Techniques. 2003. URL <https://api.semanticscholar.org/CorpusID:204747442>.
- [28] E. Edoli, D. Tasinato, and T. Vargiolo. Calibration of a Multifactor Model for the Forward Markets of Several Commodities. *Optimization*, 62(11):1553–1574, 2013. doi: 10.1080/02331934.2013.854786. URL <https://doi.org/10.1080/02331934.2013.854786>.
- [29] G. W. Emery and Q. Liu. An Analysis of the Relationship Between Electricity and Natural-Gas Futures Prices. *Journal of Futures Markets*, 22(2):95–122, 2002. doi: <https://doi.org/10.1002/fut.2209>. URL <https://onlinelibrary.wiley.com/doi/abs/10.1002/fut.2209>.
- [30] O. Feron and P. Gruet. Estimation of the Number of Factors in a Multi-Factorial Heath-Jarrow-Morton model in Electricity Markets. Working Papers hal-02880824, HAL, June 2020. URL <https://ideas.repec.org/p/hal/wpaper/hal-02880824.html>.
- [31] D. Frestad, F. E. Benth, and S. Koekebakker. Modeling Term Structure Dynamics in the Nordic Electricity Swap Market. *The Energy Journal*, 0(Number 2):53–86, 2010. URL <https://ideas.repec.org/a/aen/journl/2010v31-02-a03.html>.
- [32] M. Gardini, P. Sabino, and E. Sasso. Correlating Lévy Processes with Self-Decomposability: Applications to Energy Markets. *Decision in Economics and Finance*, 2021.
- [33] H. Geman. *Commodities and Commodity Derivatives. Modeling and Pricing for Agicultural, Metals and Energy*. John Wiley & Sons, 2005. URL <https://shs.hal.science/halshs-00144182>.
- [34] P. Glasserman. *Monte Carlo Methods in Financial Engineering*. Springer-Verlag New York, 2004.
- [35] R. Green. Risk Management and the Normal Inverse Gaussian distribution - Evidence from the Nordic Electricity Market. Technical report, 2006. Working paper, Lund University Sweden.
- [36] D. Heath, R. Jarrow, and A. Morton. Bond Pricing and the Term Structure of Interest Rates: A New Methodology for Contingent Claims Valuation. *Econometrica*, 60(1): 77–105, 1992.
- [37] S. L. Heston. A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options. *The Review of Financial Studies*, 6(2): 327–343, 1993.
- [38] W.J. Hinderks, R. Korn, and A. Wagner. A Structural Heath–Jarrow–Morton Framework for Consistent Intraday Spot and Futures Electricity Prices. *Quantitative Finance*, 20(3):347–357, 2020. doi: 10.1080/14697688.2019.1687927. URL <https://doi.org/10.1080/14697688.2019.1687927>.

- [39] M. Hsu. Spark Spread Options Are Hot! *The Electricity Journal*, 11(2):28–39, 1998. ISSN 1040-6190. doi: [https://doi.org/10.1016/S1040-6190\(98\)00004-9](https://doi.org/10.1016/S1040-6190(98)00004-9). URL <https://www.sciencedirect.com/science/article/pii/S1040619098000049>.
- [40] R.A. Johnson and D.W. Wichern. *Applied Multivariate Statistical Analysis*. Prentice Hall, Upper Saddle River, NJ, 5. ed edition, 2002. ISBN 0130925535. URL http://gso.gbv.de/DB=2.1/CMD?ACT=SRCHA&SRT=YOP&IKT=1016&TRM=ppn+330798693&sourceid=fbw_bibsonomy.
- [41] R. Kiesel and M. Kusterman. Structural Models for Coupled Electricity Markets. *Journal of Commodity Markets*, 3(1):16–38, 2016. ISSN 2405-8513. doi: <https://doi.org/10.1016/j.jcomm.2016.07.007>. URL <https://www.sciencedirect.com/science/article/pii/S2405851315300933>.
- [42] T. Kluge. Pricing Swing Options and other Electricity Derivatives. Technical report, 2006. PhD Thesis, Available at <http://perso-math.univ-mlv.fr/users/bally.vlad/publications.html>.
- [43] S. Koekebakker and O. Fridthjof. Forward Curve Dynamics in the Nordic Electricity Market. *Managerial Finance*, 31:73–94, 2005.
- [44] A.N. Langville and W.J. Stewart. The Kronecker Products and Stochastic Automata Networks. *Journal of Computational and Applied Mathematics*, 167(2):429–447, 2004.
- [45] L. Latini, M. Piccirilli, and T. Vargiuoli. Mean-reverting No-arbitrage Additive Models for Forward Curves in Energy Markets. *Energy Economics*, 79:157–170, 2019. ISSN 0140-9883. doi: <https://doi.org/10.1016/j.eneco.2018.03.001>. URL <https://www.sciencedirect.com/science/article/pii/S014098831830080X>.
- [46] F. A. Longstaff and E.S. Schwartz. Valuing American Options by Simulation: a Simple Least-Squares Approach. *Review of Financial Studies*, 14(1):113–147, 2001.
- [47] J. Lucia and E. Schwartz. Electricity Prices and Power Derivatives: Evidence from the Nordic Power Exchange. *Review of Derivatives Research*, 5(1):5–50, Jan 2002. ISSN 1573-7144. doi: 10.1023/A:1013846631785. URL <https://doi.org/10.1023/A:1013846631785>.
- [48] E. Luciano and P. Semeraro. Multivariate Time Changes for Lévy Asset Models: Characterization and Calibration. *Journal of Computational and Applied Mathematics*, 233(1):1937–1953, 2010.
- [49] D. Madan, P. Carr, M. Stanley, and E. Chang. The Variance Gamma Process and Option Pricing. *Review of Finance*, 2, 11 1999. doi: 10.1023/A:1009703431535.
- [50] R.C. Merton. Options Pricing when Underlying Shocks are Discontinuous. *Journal of Financial Economics*, 3:125–144, 1976.
- [51] J. W. Mjelde and D.A. Bessler. Market Integration Among Electricity Markets and Their Major Fuel Source Markets. *Energy Economics*, 31(3):482–491, 2009. ISSN 0140-9883. doi: <https://doi.org/10.1016/j.eneco.2009.02.002>. URL <https://www.sciencedirect.com/science/article/pii/S0140988309000310>.

- [52] C. W. Oosterlee and L. A. Grzelak. *Mathematical Modeling and Computation in Finance: with Exercises and Python and MATLAB Computer Codes*. World Scientific, 2020. ISBN 978-1-78634-794-7.
- [53] T. Panagiotidis and E. Rutledge. Oil and Gas Markets in the UK: Evidence from a Cointegrating Approach. *Energy economics*, 29(2):329–347, 2007. ISSN 0140-9883.
- [54] V. Panov and E. Samarin. Multivariate Asset-Pricing Model Based on Subordinated Stable Processes. *Applied Stochastic Models in Business and Industry*, 35(4):1060–1076, 2019.
- [55] H. Park, J. W. Mjelde, and D. A. Bessler. Price Dynamics Among U.S. Electricity Spot Markets. *Energy Economics*, 28(1):81–101, 2006. ISSN 0140-9883. doi: <https://doi.org/10.1016/j.eneco.2005.09.009>. URL <https://www.sciencedirect.com/science/article/pii/S0140988305000903>.
- [56] N. Cufaro Petroni and P. Sabino. Cointegrating Jumps: an Application to Energy Facilities. accepted presentation at the 10th ENERGY AND FINANCE CONFERENCE, London, 9-11 September 2015.
- [57] M. Piccirilli, M. D. Schmeck, and T. Vargiolu. Capturing the Power Options Smile by an Additive Two-Factor Model for Overlapping Futures Prices. *Energy Economics*, 95:105006, 2021. ISSN 0140-9883. doi: <https://doi.org/10.1016/j.eneco.2020.105006>. URL <https://www.sciencedirect.com/science/article/pii/S0140988320303467>.
- [58] J. Saifert and M. Uhrig-Homburg. Modelling Jumps in Electricity Prices: Theory and Empirical Evidence. 2006.
- [59] P. A. Samuelson. *Proof that Properly Anticipated Prices Fluctuate Randomly*, chapter Chapter 2, pages 25–38. doi: 10.1142/9789814566926_0002. URL https://www.worldscientific.com/doi/abs/10.1142/9789814566926_0002.
- [60] K. Sato. *Lévy Processes and Infinitely Divisible Distributions*. Cambridge U.P., Cambridge, 1999.
- [61] W. Schoutens. *Lévy Processes in Finance: Pricing Financial Derivatives*. John Wiley and Sons Inc, 2003.
- [62] E. Schwartz. The Stochastic Behavior of Commodity Prices: Implications for Valuation and Hedging. *The Journal of Finance*, 52(3):923–973, 1997. doi: <https://doi.org/10.1111/j.1540-6261.1997.tb02721.x>. URL <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1540-6261.1997.tb02721.x>.
- [63] P. Schwartz and J.E. Smith. Short-term Variations and Long-term Dynamics in Commodity Prices. *Management Science*, 46(7):893–911, 2000.
- [64] P.D. Sclavounos and P. E. Ellefsen. Multi-Factor Model of Correlated Commodity Forward Curves for Crude Oil and Shipping Markets. 2009.

- [65] R. Seydel. *Tools for Computational Finance*. Universitext (1979). Springer, 2004. ISBN 9783540406044.
- [66] C. Tseng and G. Barz. Short-Term Generation Asset Valuation: A Real Options Approach. *Operations Research*, 50(2):297–310, 2000.