



## Interfaces with Other Disciplines

Long-term swings and seasonality in energy markets<sup>☆</sup>Manuel Moreno<sup>a</sup>, Alfonso Novales<sup>b</sup>, Federico Platania<sup>c,\*</sup><sup>a</sup> Department of Economic Analysis and Finance, University of Castilla-La Mancha, Cobertizo San Pedro Mártir s/n, Toledo 45071, Spain<sup>b</sup> Department of Quantitative Economics and Instituto Complutense de Análisis Económico (ICAE), University Complutense of Madrid, Campus de Somosaguas, Madrid 28223, Spain<sup>c</sup> Léonard de Vinci Pôle Universitaire, Research Center, 92 916 Paris La Défense, France

## ARTICLE INFO

## Article history:

Received 30 March 2018

Accepted 30 May 2019

Available online 7 June 2019

## Keywords:

Finance

Energy markets

Seasonality

Long-term swings

Kalman filter

## ABSTRACT

This paper introduces a two-factor continuous-time model for commodity pricing under the assumption that prices revert to a stochastic mean level, which shows smooth, periodic fluctuations over long periods of time. We represent the mean reversion price by a Fourier series with a stochastic component. We also consider a seasonal component in the price level, an essential characteristic of many commodity prices, which we represent again by a Fourier series. We obtain analytical pricing expressions for futures contracts. Using futures price data on Natural Gas, we provide evidence on the presence of long-term fluctuations and show how to estimate the long-term component simultaneously with a seasonal component using the Kalman filter. We analyse the in-sample and out-of-sample empirical performance of our pricing model with and without a seasonal component and compare it with Schwartz and Smith (2000) model. Our findings show the in-sample and out-of-sample superiority of our model with seasonal fluctuations, thereby providing a simple and powerful tool for portfolio management, risk management, and derivative pricing.

© 2019 Elsevier B.V. All rights reserved.

## 1. Introduction

Characterising the stochastic behaviour of commodity prices is an issue of special relevance for practitioners in financial markets, since some commodity markets are very liquid and trade high volumes every day. Markets for futures, options, and options on futures with some commodity as the underlying asset are also very active. Commodities are not standard financial assets, so it should not be surprising that they might need specific valuation models. In particular, the characteristics of many commodity products and markets imply a relatively complex pricing nature that may combine short-term seasonal behaviour with fluctuations around a long-term trend. Short-term seasonal fluctuations over an annual period generally reflect changes in demand and supply across the different seasons in a year (Gould et al., 2008; Taylor, 2010). However, changes in production technology or shifts in taste may give rise to long-term trends that cause market

prices to fluctuate. Indeed, in line with our argument, Mu and Ye (2015) analysed the crude oil market and found evidence of a long-term trend combined with cyclical movements. The goal of this paper is to propose a model for commodity prices that can help characterise and estimate such components when they are present, without imposing any a priori constraint on their periodicity.

A pioneer contribution in commodity pricing is Schwartz (1997), who proposes a mean-reverting behaviour of commodity prices arguing that, in an equilibrium setting, when prices are relatively high (low) supply will increase (decrease) putting a downward (upward) pressure on prices. With this argument in mind, the author compares three mean-reverting models for the stochastic behaviour of a commodity price: a simple one-factor model for the logarithm of spot prices, a model proposed by Gibson and Schwartz (1990) that adds a second factor accounting for the convenience yield of the commodity, and a further extension that, additionally, incorporates the stochastic behaviour of interest rates as in Vasicek (1977). An interesting variation is the two-factor model in Schwartz and Smith (2000), which specifies log-spot prices as the sum of two state variables representing the equilibrium price level (which is assumed to follow a geometric Brownian motion process) and a short-term deviation from equilibrium prices that reverts towards zero.

<sup>☆</sup> The authors are very grateful to the editor Emanuele Borgonovo, and the anonymous referees for their truly helpful comments and suggestions over the previous versions of the paper.

\* Corresponding author.

E-mail addresses: [manuel.moreno@uclm.es](mailto:manuel.moreno@uclm.es) (M. Moreno), [anovales@ccee.uclm.es](mailto:anovales@ccee.uclm.es) (A. Novales), [federico.platania@devinci.fr](mailto:federico.platania@devinci.fr) (F. Platania).

The majority of prior research has focused on investigating the impact of a stochastic convenience yield or a jump diffusion process in commodity prices. For example, [Casassus, Collin-Dufresne, and Routledge \(2005\)](#) recognized the relevance of expressing the convenience yield as a function of the spot price level and interest rates. [Liu and Tang \(2011\)](#) introduced a stochastic volatility in the convenience yield process and [Chiu, Wong, and Zhao \(2015\)](#) also modelled its stochastic behaviour. [Fouquau and Six \(2015\)](#) prove theoretically that the two methods commonly used to obtain the convenience yield (filtered from derivative prices or using the interest-adjusted basis an observable proxy) do not provide the same results. [Almansour \(2016\)](#) extended the [Gibson and Schwartz \(1990\)](#) model to allow for regime switching and to account for shifts in the futures term structure. See [Lai and Mellios \(2016\)](#) and references therein for different analysis of the convenience yield in commodity pricing. In addition to the convenience yield, there exists a considerable body of literature that investigates the relevance of abrupt changes in commodity prices and have analysed the significance of jumps in the processes of these prices. See, for example, [Deng \(2001\)](#); [Hambly, S, and Kluge \(2009\)](#); [Hilliard and Hilliard \(2015\)](#); [Hilliard and Reis \(1998\)](#); [Kyriakou, Nomikos, Papastolou, and Poliasis \(2016\)](#); [Mayer, Schmid, and Weber \(2011\)](#); [Schmitz, Wang, and Kimn \(2014\)](#); [Sévi \(2015\)](#); [Yan \(2002\)](#), and [Gómez-Valle, Habbilashkary, and Martínez-Rodríguez \(2017, 2017\)](#); [Gómez-Valle, Habbilashkary, and Martínez Rodríguez \(2018\)](#).

A different strand of literature investigates the impact of seasonal patterns in commodity prices. In this regard, [Lucía and Schwartz \(2002\)](#) provide a simple and clever contribution addressing the seasonal behaviour embedded in many commodity prices. A seasonal component characterised by a deterministic trigonometric function with an imposed annual period is added in an ad-hoc fashion to alternative specifications for the process of log-spot prices, using data from the Scandinavian electricity market for model validation. [Cartea and Figueroa \(2005\)](#) extend this one-factor pricing model specifying a zero level mean-reverting jump-diffusion process for the underlying log-spot price, replacing the trigonometric function in [Lucía and Schwartz \(2002\)](#) by a fifth-order Fourier series. [Geman and Nguyen \(2005\)](#) use soybean prices to analyse the relationship between price volatility and inventories and apply their results to the dynamics of the term structure of soybean forward prices. [Borovkova and Geman \(2007\)](#) proposed a two-factor model that includes the average forward price and the stochastic forward premium. [Cartea and González-Pedraz \(2012\)](#) add a deterministic long-term trend to the spot price process. Finally, [Arismendi, Back, Propkopcuk, Paschke, and Rudolf \(2016\)](#) introduced a model with seasonal volatility to price commodity options.

Deterministic specifications may be well suited to capture annual seasonality. However, to model swings at the mean level of log-prices over long periods of time, a more promising approach might incorporate them in the behaviour of a mean reversion level for prices. In this regard, using a square-root model representation, [Moreno and Platania \(2015\)](#) incorporate harmonic oscillators to describe the fluctuations in interest rates and their volatility over time, while using an Ornstein-Uhlenbeck representation, [Moreno, Novales, and Platania \(2018\)](#) assume that the mean reversion level follows a time-dependent process driven by a Fourier series. We follow a similar approach by proposing a two-factor extension of the Ornstein-Uhlenbeck specification in which the mean reversion level is represented as a stochastic process with a drift component and a diffusion process that accounts for unexpected noisy shocks. The drift component produces smooth long-term fluctuations that are meant to represent changes in the technology producing the commodity or changes in consumer preferences towards the commodity, while the diffusion process generates fluctuations around the long-run trend.

Many commodities also display seasonal fluctuations around the equilibrium price level that we represent using a Fourier series that enters additively to the process for log-prices. We identify the different components, corresponding to either annual seasonality or longer term components, through the relevant peaks in the power spectrum of futures prices. Under this framework, we compute closed-form expressions for the prices of futures contracts.<sup>1</sup>

Energy markets present a perfect framework to analyse the suitability of models with a seasonal component as well as with long-term fluctuations in the mean reversion level. By nature, many sources of energy can be difficult to store or transport. For instance, the low density of Natural gas makes its storage and transportation highly impractical, so that seasonality has a strong impact on prices, specially, during periods of high demand or production shortages. In addition, there is a number of seasonal variables driving the commodity price (such as business activity, weather conditions, market regulations), and there is extensive research on pricing and hedging energy derivatives and commodities. Some interesting contributions in this area can be found in [Burger, Graeber, and Schindlmayr \(2007\)](#); [Carmona and Coulon \(2013\)](#); [Cartea and Villaplana \(2008\)](#); [Cartea and Williams \(2008\)](#); [Escribano, Ignacio Peña, and Villaplana \(2011\)](#); [Forsythe \(2007\)](#); [Geman \(2005\)](#); [Kamat and Oren \(2002\)](#); [Weron \(2007\)](#); [Wong and Lo \(2009\)](#), [Back, Propkopcuk, and Rudolf \(2013\)](#); [Chkili, Ham-moudeh, and Nguyen \(2014\)](#); [Islyayev and Date \(2015\)](#); [Janczura \(2014\)](#); [Li and Mendoza-Arriaga \(2013\)](#); [Li and Linetsky \(2013\)](#); [Li, Mendoza-Arriaga, Mo, and Mitchell \(2016\)](#); [Pellegrino and Sabino \(2014\)](#), and [Ewald, Zhang, and Zong \(2018\)](#), among many others.

With this motivation, we perform an empirical study using our model to price futures contracts on Natural Gas. We provide empirical evidence on the existence of long-term swings together with annual-frequency seasonal components in Natural Gas prices. We apply the Kalman filter on the state-space representation of our model to examine its in-sample fitting ability. Including and excluding the seasonal component in the price level, we can evaluate the relevance of the seasonal component to fit the market prices of energy futures. We also carry out an out-of-sample exercise by using the Kalman filter to obtain one-step ahead forecasts and compare the in-sample and out-of-sample performance of the two versions of our model against the two-factor model in [Schwartz and Smith \(2000\)](#).<sup>2</sup> Both exercises confirm the superiority of the Fourier model to fit market data as well as to forecast futures prices. Finally, we analyse the economic factors driving the evolution of the long-term swing process. Our findings reveal that business cycle—personal income—and labour market indicators—employment and number of hours—turn out to be significant linear predictors of the mean reversion level in natural gas futures prices.

This paper is organised as follows. [Section 2](#) derives the posited model. [Section 3](#) presents the empirical analysis, where [Sections 3.3](#) and [3.4](#) present the in-sample and out-of-sample exercises, respectively, and [Section 3.5](#) presents the economic determination of the long-term swing process. Finally, [Section 4](#) summarises the main findings and provides some concluding remarks.

## 2. Model with long-term swings and seasonal fluctuations

In this section, we introduce our valuation model for commodity prices, obtaining a closed-form expression for pricing futures

<sup>1</sup> We make no distinction between futures and forward agreements.

<sup>2</sup> As described in the academic literature, most multi-factor models are hardly comparable to our model as they propose a stochastic convenience yield or a jump diffusion process. [Schwartz and Smith \(2000\)](#) is the most popular two factor model and it has been chosen as a natural benchmark in many papers, see for instance, [Aube, Baidya, and Tito \(2008\)](#); [Lucía and Schwartz \(2002\)](#); [Manoliu and Tompaidis \(2002\)](#); [Mirantes, Población, and Serna \(2012, 2015\)](#); [Sørensen \(2002\)](#), and [Sbuelz \(2015\)](#), among others.

contracts. We want our model to incorporate potentially relevant determinants of energy derivative prices in the form of seasonal fluctuations as well as long-term swings in commodity prices.

We represent seasonal fluctuations by a deterministic term entering additively to the stochastic process for log-prices. However, long-term swings may be the effect of structural changes in production technology, preferences, or consumption patterns; we model them as changes in the mean reversion level. The idea is that the mean reversion level of log-prices experiences smooth swings. The observed time series for prices is then interpreted as reflecting short-term fluctuations around this long-term trend for mean prices due to seasonality and purely transitory shocks.

2.1. The general setup

We assume that the stochastic process for the commodity spot price at time  $t$ ,  $S_t$ , is given by the sum of two components

$$\ln(S_t) = f(t) + Y_t \tag{1}$$

where the  $f(t)$  component is a deterministic function accounting for any seasonal behaviour in the commodity price, which is modelled by a Fourier series

$$f(t) = \sum_{n=0}^{\infty} \text{Re}[A_n e^{in\omega_z t}] \tag{2}$$

and the  $Y_t$  component follows a mean-reverting process. In this regard, the mean reversion level can be interpreted as a long-run or equilibrium price that depends on the structural and environmental conditions in the energy market. The interesting assumption in our model is that the mean reversion level in  $Y_t$  follows a time-dependent process capturing long-term variations in the trend of the log-spot price,

$$dY_t = \kappa(z(t) + \eta_t - Y_t)dt + \sigma_Y dW_t^a \tag{3}$$

$$d\eta_t = \sigma_\eta dW_t^b \tag{4}$$

$$z(t) = \sum_{n=0}^{\infty} \text{Re}[B_n e^{in\omega_z t}] = B_0 + \hat{z}(t) \tag{5}$$

$$\hat{z}(t) = \sum_{n=1}^{\infty} \text{Re}[B_n e^{in\omega_z t}] \tag{6}$$

where  $\kappa, \sigma_Y, \omega_f, \omega_z, \sigma_\eta$  are positive constants.  $W_t^a$  and  $W_t^b$  are standard Wiener processes with  $dW_t^a dW_t^b = \rho dt$ .  $B_0$  is the real part of  $B_{x,0} + iB_{y,0}$ . Note that, for all  $n$ ,  $A_n, B_n$  are complex numbers, so they include a phase factor. Specifically, consider  $A_n = A_{x,n} + iA_{y,n}$  and  $B_n = B_{x,n} + iB_{y,n}$  where  $A_{x,n}, A_{y,n}, B_{x,n}, B_{y,n} \in \mathbb{R}$ . Hence,  $A_{x,n}, A_{y,n}, B_{x,n}$ , and  $B_{y,n}$  denote the amplitude and phase in the Fourier representations for  $f(t)$  and  $z(t)$ , respectively.

Assuming a constant market price of risk in both stochastic processes,  $\Lambda(Y_t, t) = \lambda_Y$  and  $\Lambda(\eta_t, t) = \lambda_\eta$ ,  $Y_t$  and  $\eta_t$  can be represented by the following risk-neutral two-factor process:<sup>3</sup>

$$dY_t = \mu_t dt + \sigma_Y d\widetilde{W}_t^a \tag{7}$$

$$d\eta_t = -\lambda_\eta \sigma_\eta dt + \sigma_\eta d\widetilde{W}_t^b \tag{8}$$

$$\mu_t = \kappa(\alpha + \hat{z}(t) + \eta_t - Y_t) \tag{9}$$

$$\alpha = B_0 - \frac{\lambda_Y \sigma_Y}{\kappa} \tag{10}$$

where  $\widetilde{W}_t^a = W_t^a + \lambda_Y t$  and  $\widetilde{W}_t^b = W_t^b + \lambda_\eta t$  are standard Wiener processes under the risk-neutral measure  $\tilde{P}$ . In addition, note that under the risk-neutral measure the process for  $\eta_t$  has a deterministic drift component determined by the product of the market price of risk of the stochastic mean reverting process and the volatility level in the long-run price.

Conditional on the information available at time  $t$ , the logarithm of the commodity spot price at time  $T$  ( $T > t$ ) follows a normal distribution. Under the risk-neutral probability measure  $\tilde{P}$ , the conditional mean and variance of that distribution are,<sup>4</sup>

$$\begin{aligned} \tilde{E}[\ln(S_T) | \mathcal{F}_t] &= f(T) + e^{-\kappa(T-t)} (\ln(S_t) - f(t)) + (1 - e^{-\kappa(T-t)}) \\ &\quad \times \left( \alpha + \left( \frac{1}{\kappa} + t \right) \lambda_\eta \sigma_\eta + \eta_t \right) - \lambda_\eta \sigma_\eta (T - t) e^{-\kappa(T-t)} \\ &\quad + \sum_{n=1}^{\infty} \text{Re} \left[ \frac{\kappa B_n}{\kappa + in\omega_z} (e^{in\omega_z T} - e^{-\kappa(T-t) + in\omega_z t}) \right] \end{aligned} \tag{11}$$

$$\begin{aligned} \tilde{V}[\ln(S_T) | \mathcal{F}_t] &= \Sigma_{tT}^2 = \sigma_\eta^2 \left[ (T-t) - \frac{2}{\kappa} (1 - e^{-\kappa(T-t)}) + \frac{1}{2\kappa} (1 - e^{-2\kappa(T-t)}) \right] \\ &\quad + \frac{\sigma_Y^2}{2\kappa} (1 - e^{-2\kappa(T-t)}) + \frac{2\sigma_Y \sigma_\eta \rho}{\kappa} \left[ (1 - e^{-\kappa(T-t)}) \right. \\ &\quad \left. - \frac{1}{2} (1 - e^{-2\kappa(T-t)}) \right] \end{aligned} \tag{12}$$

where we have applied the isometry property for stochastic integrals to compute the variance. Obtaining the forward price of a commodity maturing at time  $T$  is a straightforward application of the properties of the log-normal distribution under the risk-neutral measure. Hence, the following proposition holds.

**Proposition 2.1.** *The forward price at time  $t$  of a commodity maturing at time  $T$  is given by*

$$\begin{aligned} F(S_t, t, T) &= \tilde{E}[S_T | \mathcal{F}_t] = \exp \left\{ \tilde{E}[\ln(S_T) | \mathcal{F}_t] + \frac{1}{2} \tilde{V}[\ln(S_T) | \mathcal{F}_t] \right\} \\ &= \exp \left\{ f(T) + e^{-\kappa(T-t)} (\ln(S_t) - f(t)) + (1 - e^{-\kappa(T-t)}) \right. \\ &\quad \times \left( \alpha + \left( \frac{1}{\kappa} + t \right) \lambda_\eta \sigma_\eta + \eta_t \right) - \lambda_\eta \sigma_\eta (T - t) e^{-\kappa(T-t)} \\ &\quad \left. + \sum_{n=1}^{\infty} \text{Re} \left[ \frac{\kappa B_n}{\kappa + in\omega_z} (e^{in\omega_z T} - e^{-\kappa(T-t) + in\omega_z t}) \right] + \frac{1}{2} \Sigma_{tT}^2 \right\} \end{aligned} \tag{13}$$

where  $\Sigma_{tT}^2$  is given as in Eq. (12). Alternatively,

$$\begin{aligned} \ln(F(S_t, t, T)) &= f(T) + e^{-\kappa(T-t)} (\ln(S_t) - f(t)) + (1 - e^{-\kappa(T-t)}) \\ &\quad \times \left( \alpha + \left( \frac{1}{\kappa} + t \right) \lambda_\eta \sigma_\eta + \eta_t \right) - \lambda_\eta \sigma_\eta (T - t) e^{-\kappa(T-t)} \\ &\quad + \sum_{n=1}^{\infty} \text{Re} \left[ \frac{\kappa B_n}{\kappa + in\omega_z} (e^{in\omega_z T} - e^{-\kappa(T-t) + in\omega_z t}) \right] + \frac{1}{2} \Sigma_{tT}^2 \end{aligned} \tag{14}$$

Then, the log-futures price turns out to be made up of four components:

- (i) A correction on the log-spot price,  $e^{-\kappa(T-t)} \ln(S_t)$ .
- (ii) The effect from a periodic component that we interpret as incorporating seasonal fluctuations,  $f(T) - e^{-\kappa(T-t)} f(t)$ .

<sup>3</sup> Assuming a constant market price of risk is a common practice in this type of modelling exercise. See, for instance, Schwartz (1997); Schwartz and Smith (2000), or Lucía and Schwartz (2002).

<sup>4</sup> For details on the solution of linear mean-reverting stochastic differential equations see, for instance, Steele (2003).

- (iii) The effect from the level of noise variance of the two stochastic processes,  $(1 - e^{-\kappa(T-t)})(\frac{1}{\kappa} + t)\lambda_\eta\sigma_\eta - \lambda_\eta\sigma_\eta(T - te^{-\kappa(T-t)}) + \frac{1}{2}\Sigma_{\eta T}^2$ , and
- (iv) The influence of the time-varying mean reversion level:

$$(1 - e^{-\kappa(T-t)})(\alpha + \eta_t) + \sum_{n=1}^{\infty} \text{Re} \left[ \frac{\kappa B_n}{\kappa + i n \omega_z} (e^{i n \omega_z T} - e^{-\kappa(T-t) + i n \omega_z t}) \right] \tag{15}$$

The last three components tend toward zero as we approach the maturity of the futures contract, and, at  $t = T$ , we will have  $F(S_T, T, T) = S_T$ , as expected.<sup>5</sup>

It is also convenient to see the theoretical futures price as the aggregate of two functions:  $\ln(F(S_t, t, T)) = M(S_t, t, T; \theta_M) + N(t, T; \theta_N)$ , with:

$$M(S_t, t, T; \theta_M) = e^{-\kappa(T-t)} \ln(S_t) + (1 - e^{-\kappa(T-t)}) \times \left( \alpha + \left( \frac{1}{\kappa} + t \right) \lambda_\eta \sigma_\eta + \eta_t \right) - \lambda_\eta \sigma_\eta (T - te^{-\kappa(T-t)}) + \frac{1}{2} \Sigma_{\eta T}^2 \tag{16}$$

that depends on the spot price as well as on time to maturity, and

$$N(t, T; \theta_N) = \sum_{n=0}^{\infty} \text{Re} [A_n (e^{i \omega_f T} - e^{-\kappa(T-t) + i \omega_f t})] + \sum_{n=1}^{\infty} \text{Re} \left[ \frac{\kappa B_n}{\kappa + i n \omega_z} (e^{i n \omega_z T} - e^{-\kappa(T-t) + i n \omega_z t}) \right] \tag{17}$$

with the structural parameter vector  $\theta_M = (\alpha, \sigma_Y, \sigma_\eta, \lambda_\eta, \kappa)$ , and the seasonal parameter vector  $\theta_N = (\kappa, \{A_n\}_{n=0}^{\infty}, \omega_f, \{B_n\}_{n=1}^{\infty}, \omega_z)$ .

The first term in  $M(S_t, t, T; \theta_M)$  depends on spot prices and the following terms are corrections by time to maturity and by the level of risk. Since the price series are constructed using rolling maturities, the power functions of  $T - t$  remain within a narrow range. Thus,  $M(S_t, t, T; \theta_M)$  reproduces the behaviour of spot prices and is a relevant driver of the shape of the forward/futures curve.

However,  $N(t, T; \theta_N)$  describes seasonal fluctuations in futures price and the time evolution of the mean-reversion futures price level. Its analytical representation comes from the Fourier series structure of  $f(t)$  and  $z(t)$  and thus, the function  $N(t, T; \theta_N)$  is stationary.

2.2. Particular cases

Now, we propose two different particular specifications from the previous setup:

- *Fourier model* emerges from the general setup above when we incorporate a Fourier series with a single term for the mean reversion level, through the second term in  $N(t, T; \theta_N)$ , and do not consider the  $f(t)$  function, so that  $\ln(S(t)) = Y_t$ , and the  $N(t, T; \theta_N)$  function becomes:

$$N(t, T; \theta_N) = \text{Re} \left[ \frac{\kappa B}{\kappa + i \omega_z} (e^{i \omega_z T} - e^{-\kappa(T-t) + i \omega_z t}) \right] \tag{18}$$

- *Fourier model with one seasonal component* adds to the previous specification a Fourier series representation for  $f(t)$ , again with a single frequency, to capture either seasonal component, so that the  $N(t, T; \theta_N)$  function becomes:

$$N(t, T; \theta_N) = \text{Re} [A (e^{i \omega_f T} - e^{-\kappa(T-t) + i \omega_f t})] + \text{Re} \left[ \frac{\kappa B}{\kappa + i \omega_z} (e^{i \omega_z T} - e^{-\kappa(T-t) + i \omega_z t}) \right] \tag{19}$$

Our empirical exercise consists of estimating each model using daily observations of futures contracts written on Natural Gas. By considering nested model specifications, we can easily analyse the contribution of the seasonal component and the long-term swings in the mean reversion price to fitting futures pricing data. While the main determinant of the level of fitted futures prices is the presence of the spot price process in  $M(S_t, t, T; \theta_M)$ , we see that all the other terms also matter to achieve a good data fit. However, the contribution of the noise variance, incorporated in the last term in  $M(S_t, t, T; \theta_M)$ , is generally a minor contribution, as shown below.

We start by examining the power spectrum of prices for each futures contract to offer initial evidence on the presence of periodic terms of different natures. We let the data suggest the relevant frequencies without restricting the range of frequencies to be estimated from observed futures pricing data; we take the estimated short frequencies (long periods) as corresponding to the long-term swing process.

3. Empirical analysis

The aim of this section is to provide a comprehensive description of the prices in the Natural Gas commodity markets. The underlying hypothesis is that the price of Natural Gas has an inherent structure, including seasonal fluctuation and long-term swings in the mean reversion level, shared by every available futures contract.

Furthermore, we analyse the in-sample and out-of-sample empirical performance of the Fourier model and the Fourier model with a seasonal component against Schwartz and Smith (2000) model.<sup>6</sup>

3.1. The data

The data set used for the empirical study consists of daily observations of prices for futures contracts Ng-5, Ng-8, Ng-12, Ng-18, and Ng-24 written on Natural Gas traded on the New York Mercantile Exchange (NYMEX). The figures in the mnemonics of each contract signal the number of months to maturity. For instance, Ng-5 represents the fifth contract closest to maturity, and so on. All the time series have been extracted from Datastream. We take Henry Hub as the pricing point for natural gas futures contracts. Table 1 presents some data statistics for Henry Hub futures prices.

Moreover, Table 2 presents the augmented Dickey–Fuller (ADF) test statistic for the logarithm of prices for Natural Gas futures as well as for the first difference of each series. Clearly, the unit root hypothesis cannot be rejected for the price level at any time to maturity, meaning that futures prices are non-stationary. However, the presence of a unit root is rejected when considering the first difference of each price series.

The possible existence of periodic patterns suggests the convenience of analysing the spectral density of futures prices. Fig. 1 presents the Natural Gas pricing surface, while Fig. 2 displays the spectral density for each of the five Natural Gas futures contracts considered. The maximum at the zero frequency reflects the existence of a unit root in the price series. Interestingly, the five time series of futures prices reach additional peaks in their spectral density at other low frequencies, indicating the presence of additional medium- to long-term components driving the behaviour of these prices.

We do not impose any restriction on the range of numerical frequencies to be estimated. Consistent with our interpretation of the theoretical model, we associate the lowest estimated frequency to the Fourier term representing the mean reversion level.

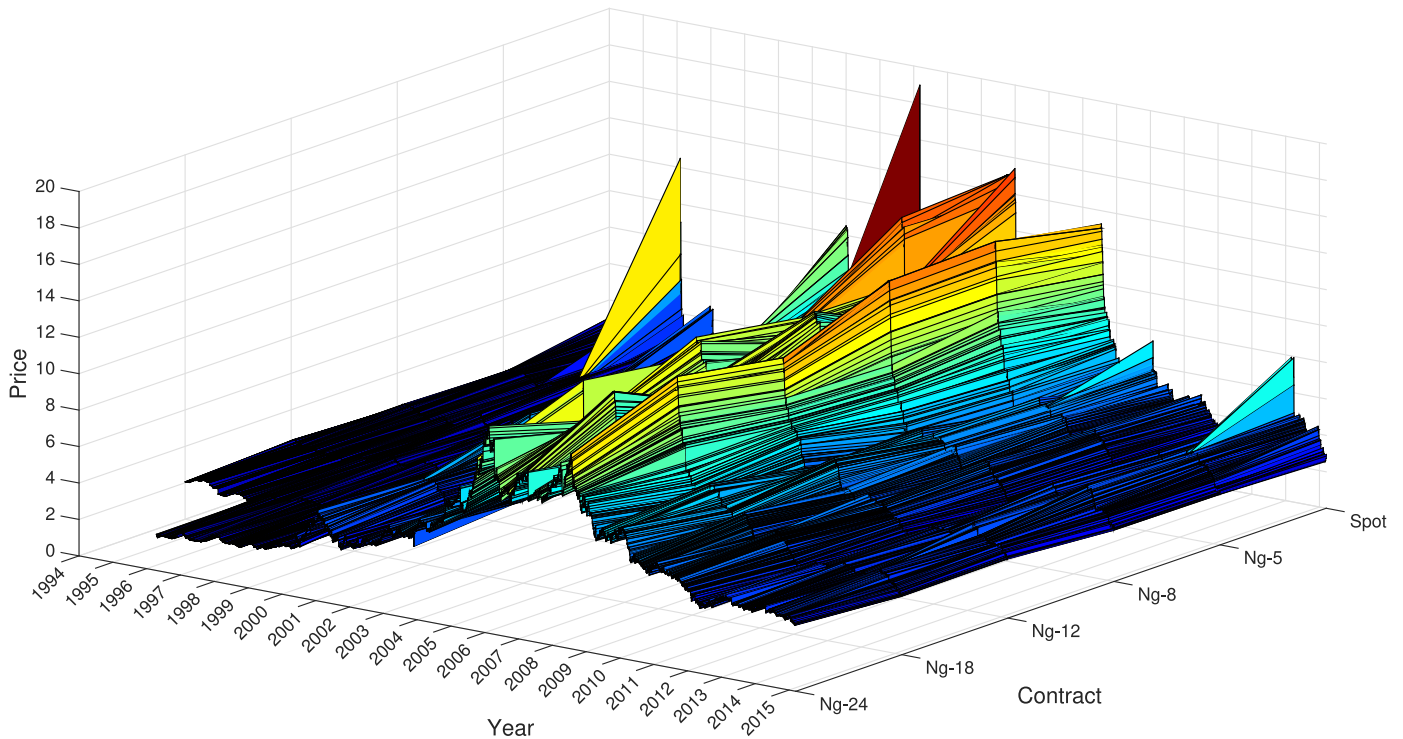
<sup>5</sup> Notice that  $\Sigma_{\eta T}^2$  converges to zero as  $t$  goes to  $T$ .

<sup>6</sup> Appendix A presents Schwartz and Smith (2000) model and its main features.



**Table 1**  
Key statistics for the prices of each futures contract written on Natural Gas.

	Period	Mean	Median	Std	Max	Min	Skewness	Kurtosis
Ng-5	03/01/94–23/02/15	4.6011	4.1095	2.4906	14.6650	1.5000	1.1944	4.2725
Ng-8	03/01/94–23/02/15	4.6704	4.1150	2.5072	14.1660	1.5720	1.0617	3.6719
Ng-12	03/01/94–23/02/15	4.6833	4.0760	2.4437	12.1830	1.7060	0.8875	2.9747
Ng-18	03/01/94–23/02/15	4.7014	4.1415	2.4436	12.9460	1.6960	0.8943	3.0496
Ng-24	12/04/96–23/02/15	4.9709	4.3670	2.2502	11.3740	1.8150	0.6548	2.5064



**Fig. 1.** Natural Gas market price surface between 03/01/94 and 23/02/15.

**Table 2**  
Augmented Dickey–Fuller test for the logarithm of futures prices and for the first difference of each price series. For the time series in first differences, we impose  $\delta = 0$  in the test model.

	ADF (Level)		ADF (First Difference)	
	t-stat	p-value	t-stat	p-value
Ng-5	-1.3058	0.8854	-23.2280	< 0.001
Ng-8	-1.0526	0.9348	-22.3136	< 0.001
Ng-12	-0.6174	0.9774	-23.6746	< 0.001
Ng-18	-0.4741	0.9847	-22.8120	< 0.001
Ng-24	-0.7164	0.9709	-23.0158	< 0.001

Test model:  $y_t = c + \delta t + \alpha y_{t-1} + \sum_{i=1}^{10} b_i \Delta y_{t-i} + e_t$ .

To estimate the Fourier model without a seasonal component, we consider a single frequency in the estimation. For the Fourier model with a seasonal component, we consider two frequencies: a short one (long period) corresponding to the Fourier component in the level of mean reversion, and a higher frequency (shorter period) that we assign to the  $f(t)$  function.

### 3.2. State-space representation

Any theoretical model explaining the formation of futures prices assumes a specific relationship between spot and futures prices, as we have described in the previous sections for alternative models of energy futures prices. However, as we observe in Figs. 1 and 3, spot Natural Gas market data experience frequent

and pronounced spikes that are not reflected in data for futures prices. In addition, we observe dates where there is a remarkable discrepancy between spot and futures prices, even for those futures closest to maturity – as shown in Fig. 3. In this regard, significant discrepancies between the spot and futures prices mostly fall in days when spot prices present spikes or jumps. Over the years, however, a variety of events have affected the price formation in the energy market.<sup>7</sup> These gaps between spot and futures prices suggest that while the fundamental and theoretical evolution of energy prices is shared by prices in the spot and futures markets, spot prices incorporate some idiosyncratic jump element that is not reflected in futures prices. Hence, a direct linear relationship between the two observed price series would face that inconsistency, hindering the empirical characterisation of such a relationship. To avoid this empirical difficulty, we assume that spot prices are not directly observed and embed spot and futures prices in a state-space representation that we estimate using the Kalman Filter, see Hamilton (1994).

Given the analytical expressions for futures prices obtained in Section 2 and Appendix A, we represent each model as,

- Schwartz and Smith (2000) model:
  - Observation equation:

$$x_t = A_t + H_t[\chi_t, \xi_t]^T + \varepsilon_t \tag{20}$$

<sup>7</sup> The online appendix provides a description of unusual or remarkable events that could have affected the energy market equilibrium.

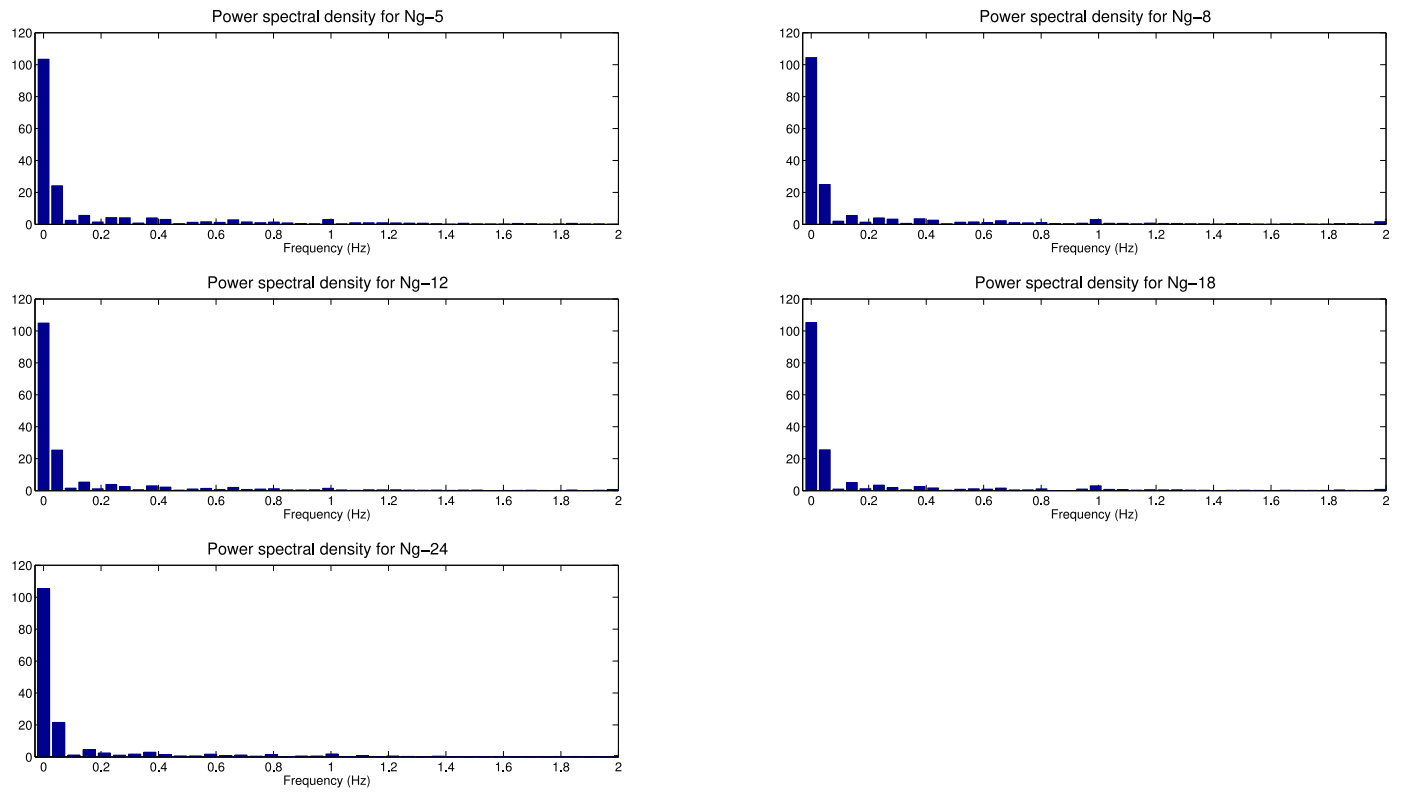


Fig. 2. Estimated spectral power density for Natural Gas futures contracts for the five maturities considered.

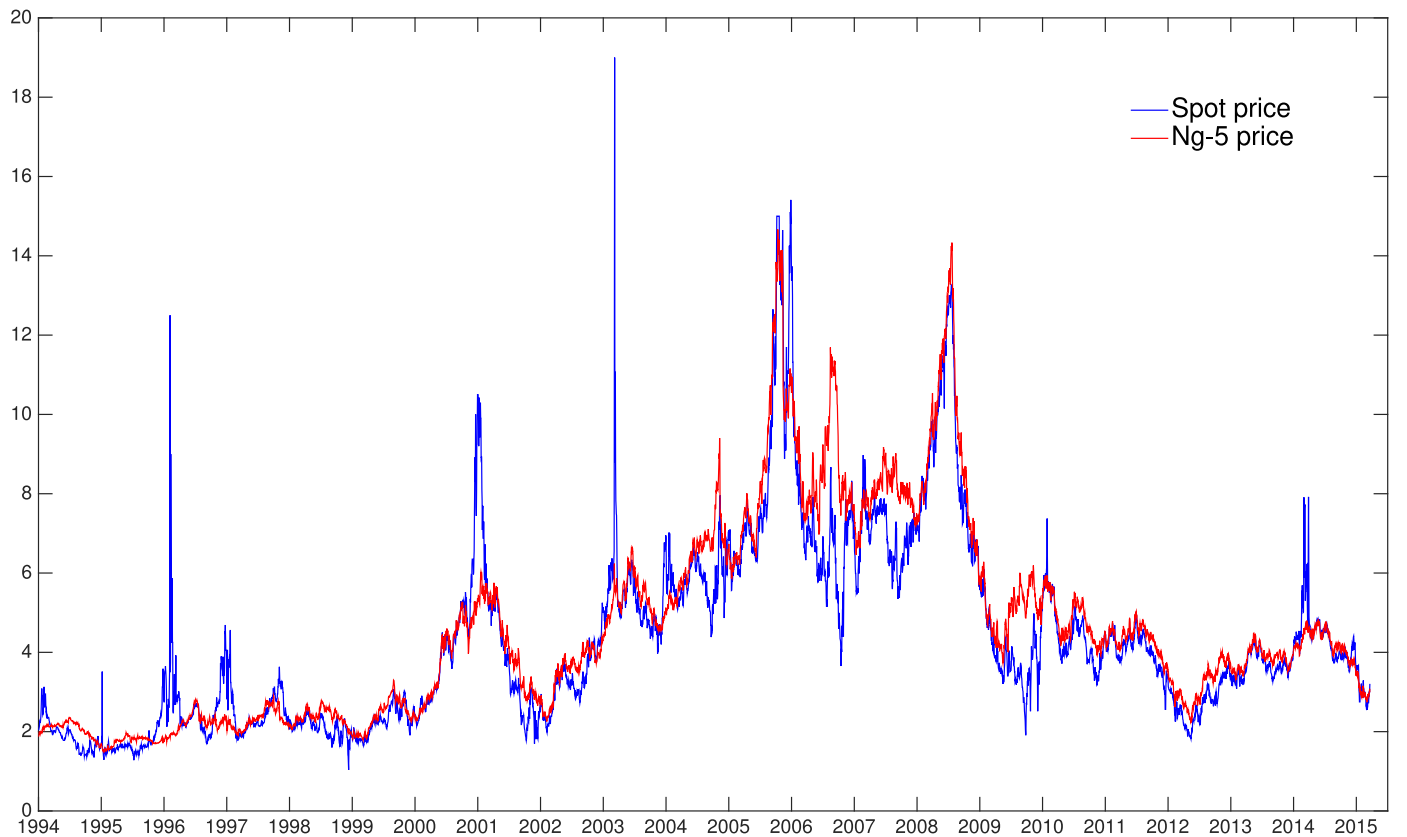


Fig. 3. Spot Natural Gas market price time-series versus the price of the futures contract maturing in five months (01/94 to 23/02/15).

– State equation:

$$[\chi_{t+1}, \xi_{t+1}]^T = G_t + F_t[\chi_t, \xi_t]^T + \epsilon_{t+1} \tag{21}$$

where

$$\chi_t = \ln(F(S_t, t, T)) \tag{22}$$

$$A_t = A(T - t) \tag{23}$$

$$H_t = [e^{-\kappa(T-t)}, 1] \tag{24}$$

$$G_t = [0, \mu_\xi dt] \tag{25}$$

$$F_t = \begin{bmatrix} e^{-\kappa(T-t)} & 0 \\ 0 & 1 \end{bmatrix} \tag{26}$$

$$E[\epsilon_t] = 0 \tag{27}$$

$$\text{Var}[\epsilon_t] = \begin{bmatrix} (1 - e^{-2\kappa dt}) \frac{\sigma_\chi^2}{2\kappa} & (1 - e^{-\kappa dt}) \frac{\rho_{\chi\xi} \sigma_\chi \sigma_\xi}{\kappa} \\ (1 - e^{-\kappa dt}) \frac{\rho_{\chi\xi} \sigma_\chi \sigma_\xi}{\kappa} & \sigma_\xi^2 dt \end{bmatrix} = Q \tag{28}$$

• Fourier model and Fourier model w/seasonal component:

– Observation equation:

$$\chi_t = A_t + H_t[Y_t, \eta_t]^T + \epsilon_t \tag{29}$$

– State equation:

$$[Y_{t+1}, \eta_{t+1}]^T = G_t + F_t[Y_t, \eta_t]^T + \epsilon_{t+1} \tag{30}$$

where

$$\text{Fourier} = \begin{cases} Y_t = \ln(S_t) \\ A_t = (1 - e^{-\kappa(T-t)}) \left( \alpha + \left( \frac{1}{\kappa} + t \right) \lambda_\eta \sigma_\eta \right) \\ \quad - \lambda_\eta \sigma_\eta (T - te^{-\kappa(T-t)}) + \text{Re} \left[ \frac{\kappa B}{\kappa + i\omega_z} \right. \\ \quad \left. \times (e^{i\omega_z T} - e^{-\kappa(T-t) + i\omega_z t}) \right] + \frac{1}{2} \sum_{t=T}^2 \end{cases} \tag{31}$$

$$\text{Fourier w/ seasonal component} = \begin{cases} Y_t = \ln(S_t) - \text{Re} [Ae^{i\omega_f t}] \\ A_t = \text{Re} [Ae^{i\omega_f T}] + (1 - e^{-\kappa(T-t)}) \\ \quad \times \left( \alpha + \left( \frac{1}{\kappa} + t \right) \lambda_\eta \sigma_\eta \right) \\ \quad - \lambda_\eta \sigma_\eta (T - te^{-\kappa(T-t)}) + \text{Re} \left[ \frac{\kappa B}{\kappa + i\omega_z} \right. \\ \quad \left. \times (e^{i\omega_z T} - e^{-\kappa(T-t) + i\omega_z t}) \right] + \frac{1}{2} \sum_{t=T}^2 \end{cases} \tag{32}$$

$$\chi_t = \ln(F(S_t, t, T)) \tag{33}$$

$$H_t = [e^{-\kappa(T-t)}, 1 - e^{-\kappa(T-t)}] \tag{34}$$

$$G_t = [\kappa z(t) dt, 0] \tag{35}$$

$$F_t = \begin{bmatrix} 1 - \kappa dt & \kappa dt \\ 0 & 1 \end{bmatrix} \tag{36}$$

$$E[\epsilon_t] = 0 \tag{37}$$

$$\text{Var}[\epsilon_t] = \begin{bmatrix} \sigma_Y^2 dt & \sigma_Y \sigma_\eta \rho dt \\ \sigma_Y \sigma_\eta \rho dt & \sigma_\eta^2 dt \end{bmatrix} = Q \tag{38}$$

### 3.3. Estimation with Natural Gas data

#### 3.3.1. Numerical results

Table 3 presents estimated parameters and their standard deviations for each model over the whole sample, along with measures to compare the relative quality of each model: the maximised log-likelihood from the Kalman filter estimation, the Akaike information criterion (AIC), an estimate of the standard deviation of pricing errors,  $(\frac{1}{T} \sum_t \hat{u}_t^2)^{1/2}$ , and the sum of squared residuals,  $\sum_t \hat{u}_t^2$ . The results are conclusive: compared with the benchmark model, the different specifications of Fourier terms in the stochastic process for log-spot prices dramatically improve the in-sample fitting of futures prices. Specifically, we observe that the Fourier model almost halves the aggregated squared residuals of Schwartz and Smith, with a 47% reduction. When compared to the Fourier model w/seasonal component, the reduction is even greater, an 81% reduction over the benchmark model. We also use the AIC as an estimator of the relative quality of the Fourier model over the benchmark model. Even though the AIC penalises those models with a higher number of parameters, we observe that both specifications of the Fourier model provide lower AIC values than Schwartz and Smith. These results suggest the superior performance of both specifications of the Fourier model over the benchmark model, with a clear dominance of the Fourier with a seasonal component model.

Although the Schwartz and Smith model and the Fourier model—w/ and w/o seasonal component—belong to the family of two-factor models, the theoretical background and modelling behind them is essentially different. The objective of the Fourier model is to account for long-term swings in the mean reversion level as well as for seasonal patterns in energy markets, while the Schwartz and Smith model aims to describe the evolution of the equilibrium price level and short-term deviations—see Appendix A and Section 2. Even though comparing the in-sample and out-of-sample performance provides a meaningful and valid empirical appraisal of each model, the theoretical background prevents us from comparing the estimated parameters between Schwartz and Smith and the Fourier models. As we observe in Table 3, the estimated parameters in the Fourier and Fourier w/seasonal component are quite stable, in particular, we observe the presence of a consistent long-term swing of 19–19.5 years in both models. In both cases, the long-term swing dominates the mean reversion level and the level  $\alpha$  is statistically non-significant. In addition, the seasonal Fourier model also presents a seasonal component with a one-year period. In both models we observe a higher variance level in futures prices than in the long-run mean reversion level, that is  $\sigma_Y > \sigma_\eta$ , as should be expected from our theoretical model. In effect, long-run mean reversion prices can be regarded as equilibrium prices that depend on structural and environmental conditions, such as production technology or consumption preferences, and, therefore, are subject to a lower volatility level and shocks than current prices. In addition, interestingly, is the correlation between futures prices and the long-run mean reversion price, which we estimate is 30.54% and 41.13% for the Fourier and the seasonal Fourier model, respectively. Indeed, the long-term mean reversion price level depends on structural and environmental conditions, but does not depend on short-term or seasonal fluctuations. The simplest version of the Fourier model does not incorporate these short-term and seasonal fluctuations in prices. Hence, the idiosyncratic component in futures prices will be somewhat affected by these fluctuations—which are present in prices but not in the long-run level—resulting in an underestimation of the correlation parameter. In contrast to the Fourier model, the Fourier model w/seasonal component successfully identifies and incorporates the effect of seasonal fluctuations; hence, after

**Table 3**

$\sum_{t=1}^T \hat{u}_t^2$  is the sum of squared pricing errors,  $(\frac{1}{T} \sum_{t=1}^T \hat{u}_t^2)^{1/2}$  is an estimate of the residual standard deviation, and AIC is the Akaike information criterion.

Parameters	Schwartz and Smith	Fourier	Fourier w/Seasonal
$\sigma_x$	0.4469 (0.0007)		
$\lambda_x$	0.0651 (0.0132)		
$\mu_\xi$	-0.0651 (0.0021)		
$\sigma_\xi$	0.2255 (0.0004)		
$\mu_\xi^*$	-0.0463 (0.0003)		
$\rho_{\xi x}$	-0.4364 (0.0027)		
$\kappa$	0.5929 (0.0020)	1.7898 (0.0013)	1.2853 (0.0005)
$\alpha$		0.4432 (1.0632)	0.4793 (0.5454)
$\sigma_y$		0.6295 (0.0004)	0.5085 (0.0002)
$\sigma_\eta$		0.1924 (0.0001)	0.1771 (0.0001)
$\lambda_\eta$		0.1465 (0.0003)	0.1896 (0.0002)
$\rho$		0.3054 (0.0009)	0.4113 (0.0005)
$B_x$		-0.0453 (0.0002)	-0.1130 (0.0001)
$B_y$		-0.0873 (0.0001)	-0.0845 (0.0001)
$\omega_z$		0.3223 (0.0058)	0.3336 (0.0010)
Period $\approx$		19.5 years	19 years
$A_x$			0.0599 (0.0001)
$A_y$			0.0166 (0.0001)
$\omega_f$			6.2471 (0.0001)
Period $\approx$			1 year
Log-likelihood $\times 10^{-4}$	6.9419	7.1974	8.2286
AIC $\times 10^{-5}$	-1.3881	-1.4392	-1.6454
$\sum_{t=1}^T \hat{u}_t^2$	172.224	91.284	32.0269
$(\frac{1}{T} \sum_{t=1}^T \hat{u}_t^2)^{1/2}$	0.0790	0.0575	0.0341

accounting for these effects, the correlation between prices and the long-run level increases. Finally, the pair of parameters  $(B_x, B_y)$  and  $(A_x, A_y)$  define the amplitude and phase in the long-term swing and seasonal component, respectively. We analyse the impact of these components in the next section.

3.3.2. Model analysis

The better overall fit to the data, especially when incorporating a seasonal component, means that the Fourier specification does a better job in jointly reproducing the long-term swings in the mean together with the seasonal components that might be present in Natural Gas futures price data. In addition, we want to advance the understanding of the role of the different components of the pricing model to explain futures prices. To that end, using data from the Ng-18 futures contract, Fig. 4 decomposes fitted futures prices from the seasonal Fourier model to show the contribution of each component. The upper-left graph displays observed and filtered spot prices, while the lower-left graph shows estimated and observed futures prices (all in logs). Spot prices present a higher short-term volatility than futures prices due to the presence of important spikes in spot prices that are not reproduced in futures prices. Given the close association between spot and futures prices in our pricing formula, the filtered spot price process closely reproduces the time evolution of the observed spot prices, but cannot reproduce their spikes. If we had used observed spot prices rather than filtered spot prices to estimate the model, these spikes would have produced significant fitting errors in futures prices. The lower-right graph displays the evolution of each term in the Fourier model with the seasonal component: (i) the spot-price effect,  $e^{-\kappa(T-t)} \ln(S_t)$ , (ii) the effect from the periodic component,  $Re[Ae^{i\omega_f t}] - e^{-\kappa(T-t)} Re[Ae^{i\omega_f t}]$ , (iii) the volatility effect in both stochastic terms,  $(1 - e^{-\kappa(T-t)}) (\frac{1}{\kappa} + t) \lambda_\eta \sigma_\eta - \lambda_\eta \sigma_\eta (T - te^{-\kappa(T-t)}) + \frac{1}{2} \sum_{t=1}^T \hat{u}_t^2$ , and (iv) the effect of the time-varying mean reversion level,  $(1 - e^{-\kappa(T-t)}) (\alpha + \eta_t) + Re[\frac{\kappa B}{\kappa + i\omega_z} (e^{i\omega_z T} - e^{-\kappa(T-t) + i\omega_z t})]$ . The volatility effect can be regarded as a level effect of minor contribution compared with the other components. In addition, the time-varying mean-reversion level reproduces to a good extent the level of futures prices while the periodic effect is a sinusoid that provides the right time pattern to reproduce the evolution of futures prices.

Now, we consider the behaviour of the  $N(t, T; \theta_N)$  function (see Eq. (17)) which is a filtered signal for the behaviour of futures prices after taking into account the effect of spot prices, the time-varying mean reversion level, and the volatility effect. The upper graph of Fig. 5 shows this function together with the adjustment from the Fourier model. Clearly, the Fourier term for the long-run mean reversion level captures the time evolution of a general trend although, as expected, it does not reproduce the short-term seasonal fluctuations. Indeed, it is striking that after having discounted the influence of spot prices, the  $N(t, T; \theta_N)$  function may display such a clear seasonal effect with an apparent annual period. The Fourier model with a seasonal component is able to capture this seasonal pattern by adding a Fourier term with a single frequency for either a seasonal or a cyclical component in  $f(t)$  (middle graph). Finally, the lower graph presents the power spectral density, where we easily identify a short frequency describing a long-term fluctuation and a second spectral peak with an annual frequency.

3.4. Out-of-sample analysis

We have just seen the improvement achieved by our Fourier-based model for in-sample fitting relative to a more standard alternative. In this section we examine whether our model also helps in anticipating changes in Natural Gas futures markets. To this end, we compare the one-day forecasting performance of the Fourier model and the Fourier model with a seasonal component with that of the Schwartz and Smith model over three full years – 2012, 2013, and 2014. To obtain forecast futures prices over 2012, we start by estimating each model (Eqs. (20) and (29)) using data from 03/01/1994 (12/04/1996 for Ng-24) up to 31/12/2011. After that, we initiate the Kalman Filter algorithm and obtain one-day ahead forecasts for futures prices over 2012 using the state-space representation presented in Section 3.2,

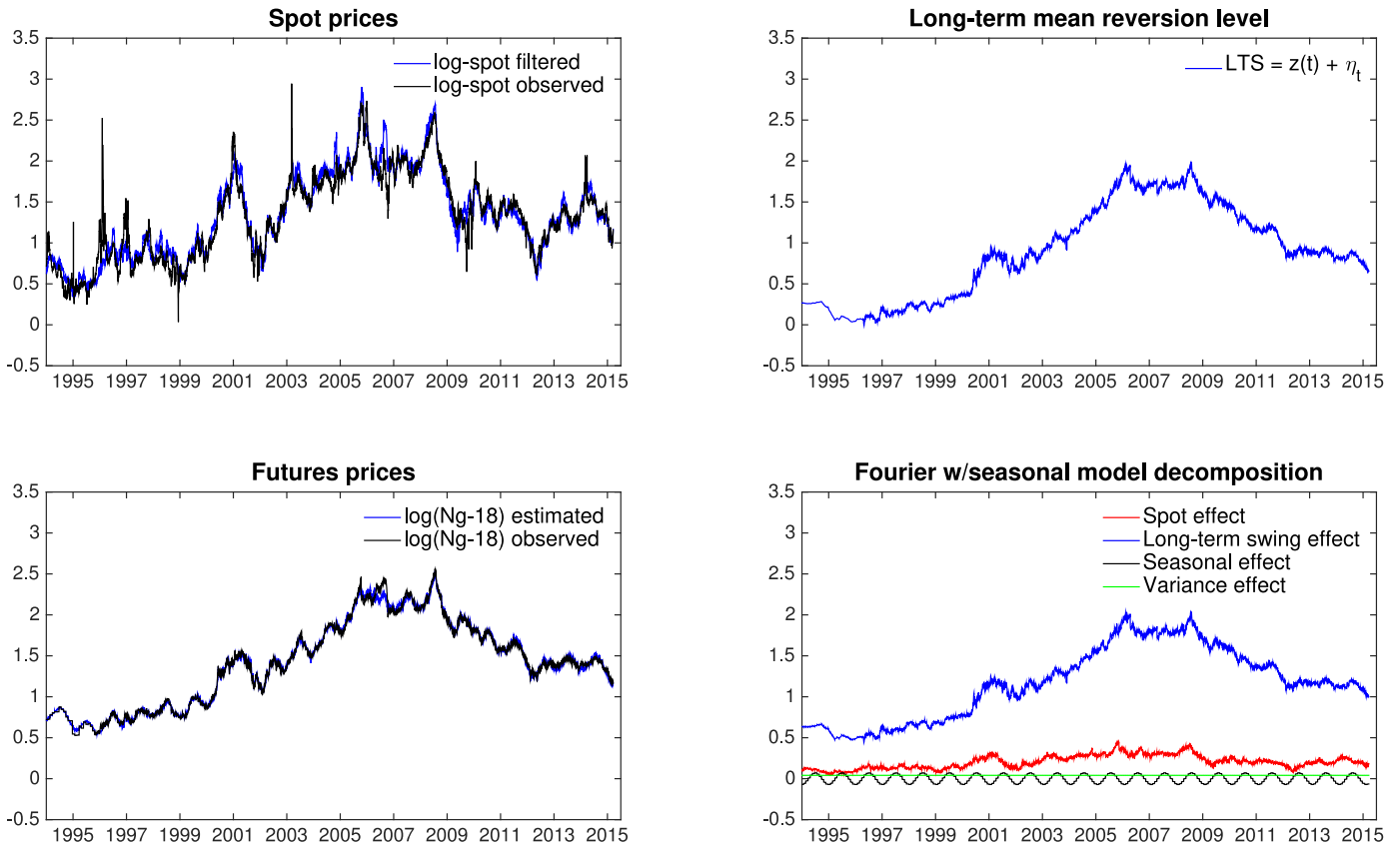
$$\hat{x}_{t+1|t} = E(x_{t+1} | \mathcal{F}_t) = A_{t+1} + H_{t+1} [\hat{y}_{t+1|t}, \hat{\eta}_{t+1|t}]^T \tag{39}$$

with associated mean square error

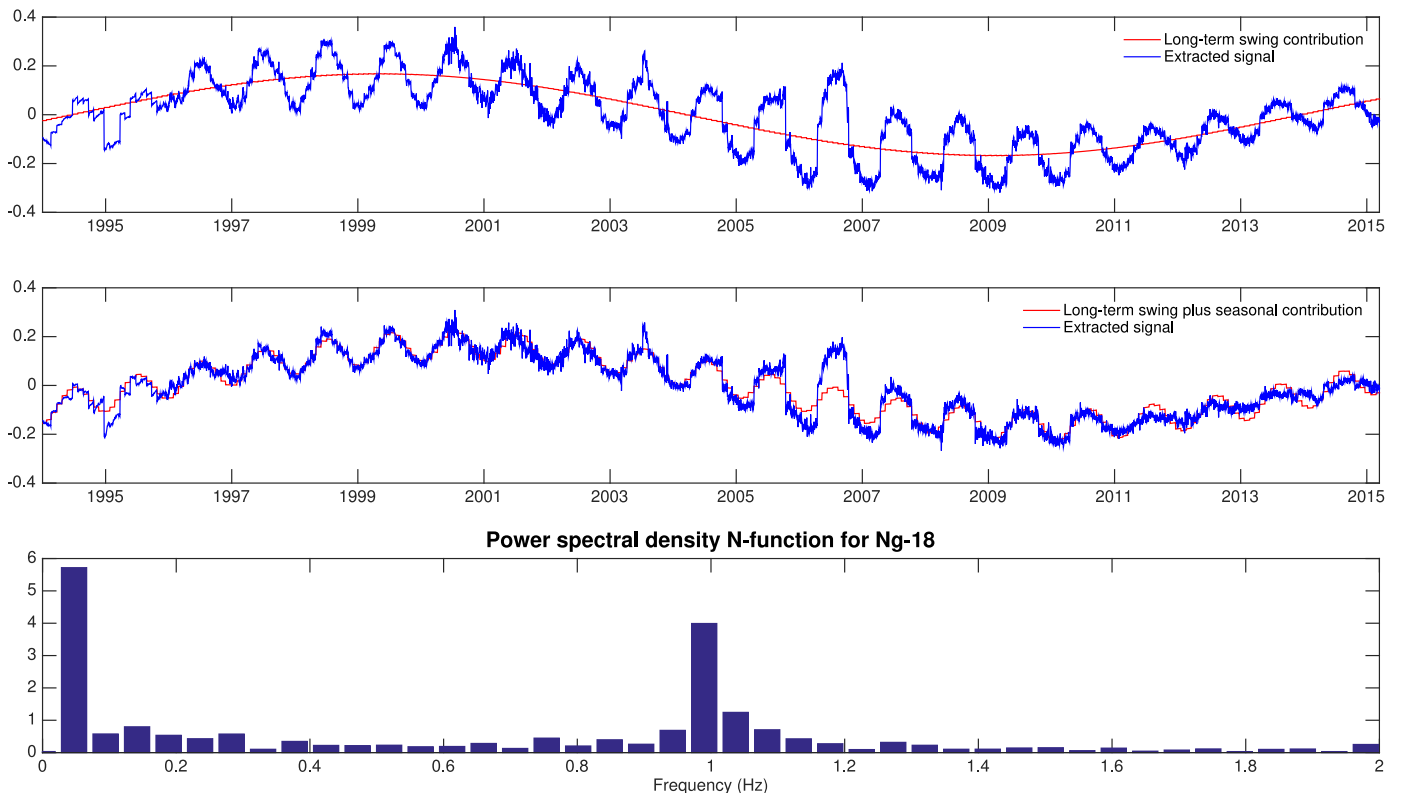
$$E[(x_{t+1} - \hat{x}_{t+1|t})(x_{t+1} - \hat{x}_{t+1|t})^T] \tag{40}$$

and we run the process for the entire year. For a detailed explanation on the Kalman Filter, we recommend Hamilton (1994).





**Fig. 4.** The upper-left row shows observed and filtered spot prices, both in logs. The upper-right row shows the process followed by the  $LTS = z(t) + \eta_t$  process provided in Eq. (3). The lower left graph shows observed and fitted futures prices. The lower right graph displays: (i) the effect of spot prices, (ii) the seasonal component, (iii) the volatility effect, (iv) the effect of the long-term mean reversion price level.



**Fig. 5.** The upper and middle graphs display the  $N(t, T; \theta)$  function together with the Fourier component estimated in the Fourier and Fourier w/seasonal model, respectively. The lower graph shows the  $N(t, T; \theta)$  function power spectral density for the Ng-18 futures contract.

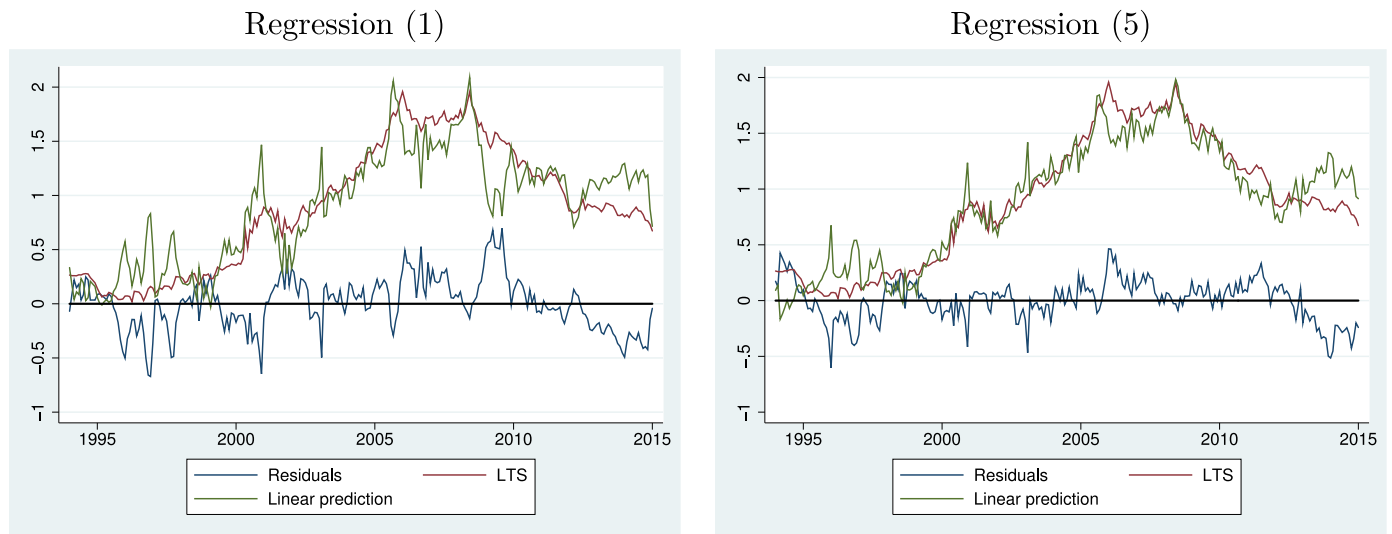


Fig. 6. Long-term swing process (red line) along with the fitted values (green line) and the residuals (blue line) obtained from regression models (1) and (5), respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

Table 4

One-day ahead forecasting errors for Schwartz and Smith, Fourier, and Fourier with seasonal models over 2012–2014. The results for different maturities ( $T_5, T_8, T_{12}, T_{18}$  and  $T_{24}$ ) are presented. For each model, the table displays the sum of squared values of daily forecasting errors over each year as well as over the whole forecasting sample (last column). The last row in each panel aggregates squared residuals across maturities for a given year. The lowest forecasting errors are presented in bold, and the lowest aggregated squared residuals are shown in red. Note that:  $\sum \hat{u}^2 = \sum (x_{t+1} - \hat{x}_{t+1|t})^2$ .

	Contract	2012	2013	2014	2012–2014
Schwartz and Smith	$T_5$	0.1417	0.5131	2.4152	3.0700
	$T_8$	1.0417	0.7346	1.7965	3.5728
	$T_{12}$	0.6583	<b>0.0357</b>	<b>0.0526</b>	0.7466
	$T_{18}$	<b>0.0518</b>	0.6096	1.0739	1.7353
	$T_{24}$	0.9503	0.0230	<b>0.0180</b>	0.9913
	$\sum \hat{u}^2$	2.8438	1.9160	5.3562	10.1160
Fourier	$T_5$	0.1474	<b>0.0649</b>	0.0849	0.2972
	$T_8$	1.1616	<b>0.4744</b>	0.8593	2.4953
	$T_{12}$	<b>0.1368</b>	0.1035	0.4647	0.7050
	$T_{18}$	0.7578	0.5704	0.6843	2.0125
	$T_{24}$	<b>0.0428</b>	0.0223	0.0200	0.0851
	$\sum \hat{u}^2$	2.2464	<b>1.2355</b>	2.1132	5.5951
Fourier w/Seasonal	$T_5$	<b>0.1402</b>	0.0722	<b>0.0801</b>	<b>0.2925</b>
	$T_8$	<b>0.9253</b>	0.6180	<b>0.2853</b>	<b>1.8286</b>
	$T_{12}$	0.3335	0.2409	0.1273	<b>0.7017</b>
	$T_{18}$	0.3616	<b>0.4313</b>	<b>0.3157</b>	<b>1.1086</b>
	$T_{24}$	0.0436	<b>0.0223</b>	0.0181	<b>0.0840</b>
	$\sum \hat{u}^2$	<b>1.8042</b>	1.3847	<b>0.8265</b>	<b>4.0154</b>

To obtain forecasts for 2013, we estimate each model using data up to 31/12/2012. The same recalibration procedure is applied to obtain one-day ahead forecasts for futures prices over 2014. Table 4 summarises the results. We observe that the Schwartz and Smith model and the simpler version of the Fourier model achieve the minimum sum of squares of forecast errors in 4 of the 15 (maturity, year)-pairs. The Fourier model with a seasonal component attains the minimum sum of squared forecast errors in the remaining seven cases. Thus, the Fourier model with a seasonal component specification obtains the best forecasting performance in most (maturity, year)-pairs. When we aggregate over time in the last column, the Fourier model with a seasonal component is the best performing model for all maturities. Furthermore, the reduction in forecasting errors is large, especially relative to the Schwartz and Smith model. If we aggregate across maturities in the last row of each panel, the Fourier model with a seasonal

Table 5

Realised average dollar mispricing by each model (in USD units).

	2012	2013	2014
Schwartz and Smith	0.1260	0.1114	0.1846
Fourier	0.1022	0.0900	0.1203
Fourier with seasonal	0.0985	0.0962	0.0767

component has the best performance for 2012 and 2014, with the simpler version of the Fourier model being best for 2013, although by a small margin. It is also interesting to analyse the mispricing in terms of dollar value. Table 5 presents the average dollar mispricing committed by each model. We observe that both versions of the Fourier model consistently outperform the benchmark, with the Fourier model w/seasonal component providing more stable and reliable predictions of futures prices. These results are very encouraging and suggest the convenience of using our model specification to forecast future Natural Gas prices against more standard alternative specifications.

### 3.5. The economic determinants of long-term swings (LTS)

If our model is successful, the estimate of the underlying price level that the market reverts to should reflect the evolution of the energy market, as well as the possible reaction to business cycle events. To analyse this, in this section, we estimate a set of regression models explaining the time series of the  $LTS = z(t) + \eta_t$  process provided in Eq. (3) and empirically obtained in Section 3.3. The independent variables considered here are:

1. NATLGAS: Henry Hub spot prices as traded on the New York Mercantile Exchange (NYMEX)<sup>8</sup>
2. CRUDEOIL: Crude Oil spot prices as traded on the New York Mercantile Exchange (NYMEX)<sup>8</sup>
3. STORAGE: US working Natural gas Gn underground storage, including US totals and regional breakdowns<sup>9</sup>
4. INCOME: US Personal income excluding transfer receipts chained 2012 Dollars SAAR.<sup>9</sup>
5. TOT\_EMPL: US employment total in labour force, seasonally adjusted.<sup>9</sup>

<sup>8</sup> Obtained from Datastream.

<sup>9</sup> Obtained from Bloomberg.

**Table 6**

Least-squares estimates (standard deviation in parentheses) of five regression models explaining the relationship between the long-term swing process ( $LTS = z(t) + \eta_t$ ) in Natural Gas futures prices and economic conditions. For each model, we present the adjusted  $R$ -squared statistic, the root-mean-square error (Root MSE), and the  $F$ -test for global significance.

Long-term swing process					
VARIABLES	(1) Energy market	(2) Income	(3) Job market	(4) Income + Job market	(5) Complete
L_NATLGAS	0.709 (0.039)				0.375 (0.049)
L_CRUDEOIL	0.255 (0.033)				0.134 (0.034)
L_STORAGE	0.194 (0.039)				0.145 (0.032)
L_INCOME		2.891 (0.157)		−6.607 (0.555)	−4.678 (0.521)
L_TOT_EMPL			6.198 (0.589)	23.177 (1.483)	14.407 (1.608)
L_HOURS			−17.695 (1.566)	−27.847 (1.587)	−20.246 (1.433)
Constant	−2.483 (0.271)	−25.687 (1.434)	−10.072 (11.814)	−114.470 (13.622)	−57.258 (14.648)
Observations	253	253	253	253	253
adj. $R$ -sq	0.813	0.562	0.726	0.840	0.892
Root MSE	0.243	0.372	0.294	0.225	0.185
Prob > $F$	0.000	0.000	0.000	0.000	0.000

6. HOURS: US average weekly hours non-farm total private production and non-supervisory, seasonally adjusted.<sup>9</sup>

In Section 3.3, we obtained daily estimates of LTS. We now consider the end-of-month values of LTS between January 1994 and January 2015, with 253 monthly observations. Table 6 shows the adjusted  $R$ -squared statistic for each regression model. All the variables are in logarithms and all the estimated coefficients are significant at the 1% significance level. The first regression shows that mean reversion prices are strongly correlated with the spot prices of Natural Gas and Crude Oil. This is to be expected, since LTS is estimated from Natural Gas futures prices. In addition, given the energy market interdependence we might expect a strong correlation between the Natural Gas and Crude Oil markets. The second regression shows a significant correlation also with personal income, one of the main business cycle indicators. We do not include employment in that regression because of its high correlation with personal income (correlation=0.98). The third regression uses as explanatory variables some labour market indicators: employment and hours. They achieve an even higher explanatory power than income and, it is, in fact, quite remarkable that long-run mean reversion prices are so strongly correlated with business cycle indicators. The coefficient in hours is negative, given its well-known negative correlation with employment. The fourth regression shows that income has additional explanatory power to that contained in the labour market indicators. The collinearity in that regression justifies the negative sign of the income coefficient. The final regression, again subject to collinearity, achieves an explanatory power better than the regression on energy prices alone.<sup>10</sup> Finally, Fig. 6 presents a visual comparison of the results obtained using the energy market regression model (1) and the full regression model (5). Even though both models reproduce the evolution of long-term mean reversion prices quite well, it is clear that income and job market variables help to reduce the spikes present in the energy spot market, providing a much smoother and accurate representation of the long-term mean reversion process.

<sup>10</sup> Collinearity precludes us from giving a strong interpretation of individual coefficient estimates, but it does not bias the overall measures of fit of the regression model.

#### 4. Conclusions

Changes in the production technology or shifts in preferences may cause long-run swings in commodity prices. We view such changes as influencing the mean reversion level of commodity prices and we incorporate that feature into a continuous-time Ornstein–Uhlenbeck model for the logarithm of commodity spot prices. In our model, the mean reversion price level experiences such long-term swings, represented by a Fourier series, as well as purely random shocks, represented by a diffusion process. Thus, we have a two-factor model for commodity prices. In addition, we consider the possibility of seasonal fluctuations in the observed level of spot prices, which we represent by a second Fourier series.

Our model is analytically tractable and allows closed-form expressions for the price of futures contracts. It nests the one-factor models by Schwartz (1997) and Lucía and Schwartz (2002), preserving their analytical tractability and allowing a more flexible representation of the time evolution of commodity spot prices. We have examined in detail the empirical performance of two versions of our model, with and without a seasonal component, against the two-factor model in Schwartz and Smith (2000) using prices of futures contracts on Natural Gas at different maturities. To identify the relevant frequencies driving futures prices and their associated periods, we first conducted a spectral analysis of each futures contract. The models are estimated in state-space representation using the Kalman Filter.

Even in its simplest representation, without the seasonal component in the observed spot prices, our model outperforms Schwartz and Smith (2000), providing a better and more reliable fitting to observed futures prices. Thus, the extension incorporated in our model seems to be quite relevant for price modelling. Allowing for a seasonal component leads to still significant additional improvements in fitting the data. Furthermore, we have also performed an out-of-sample analysis where, again, our model provides significant improvement over the benchmark, achieving better predictions in most cases.

The main improvement in the fitting ability of our model comes from allowing long-term swings in the mean reversion level. Furthermore, without imposing any restriction or prior belief on the frequencies to be estimated from the Fourier series for the different components, our model is quite successful in capturing a

seasonal component whenever it is present in the data. The model should also be appropriate for modelling other types of commodities with specific seasonal components, like agricultural products.

### Acknowledgments

The authors gratefully acknowledge financial support by grants P12-SEJ-1733, PPII-2014-030-P, and ECO2017-85927-P (Moreno), grants ECO2015-67305-P, PrometeoII/2013/015, and Programa de Ayudas a la Investigación from Banco de España (Novales), and grants Spanish Ministry of Education-FPU-MEC, Fonds de la Recherche Scientifique FNRS (T.0014.14), and PPII-2014-030-P (Platania).

### Appendix A

This appendix introduces the benchmark model that we use in our empirical analysis. The model was presented by [Schwartz and Smith \(2000\)](#).

*Benchmark model: Schwartz and Smith (2000)*

This model assumes that the commodity spot price is given by the decomposition of two stochastic factors as

$$\ln(S_t) = \chi_t + \xi_t \quad (41)$$

where  $\xi_t$  represents the equilibrium price level process and  $\chi_t$  the short-term deviation from the equilibrium price level. Under this framework, the authors assume that short-term deviations follow a zero mean reverting process given as

$$d\chi_t = -\kappa \chi_t dt + \sigma_\chi dW_t^a \quad (42)$$

while the equilibrium price level is assumed to follow a Brownian motion process given as

$$d\xi_t = \mu_\xi dt + \sigma_\xi dW_t^b \quad (43)$$

where  $dW_t^a$  and  $dW_t^b$  are standard Wiener processes with  $dW_t^a dW_t^b = \rho_{\chi\xi} dt$ .

Assuming a constant market price of risk for the equilibrium price level and short-term deviation process, the risk-neutral processes are given as

$$d\chi_t = (-\kappa \chi_t - \lambda_\chi) dt + \sigma_\chi d\tilde{W}_t^a \quad (44)$$

$$d\xi_t = (\mu_\xi - \lambda_\xi) dt + \sigma_\xi d\tilde{W}_t^b \quad (45)$$

Under the risk-neutral measure the log-spot price process is normally distributed, hence the forward price at time  $t$  of a contract maturing at time  $T$  is given as

$$F(S_t, t, T) = \tilde{E}[S_T | \mathcal{F}_t] = \exp \left\{ \tilde{E}[\ln(S_T) | \mathcal{F}_t] + \frac{1}{2} \tilde{V}[\ln(S_T) | \mathcal{F}_t] \right\} \quad (46)$$

where

$$\tilde{E}[\ln(S_T) | \mathcal{F}_t] = e^{-\kappa(T-t)} \chi_t + \xi_t + \mu_\xi^*(T-t) - (1 - e^{-\kappa(T-t)}) \frac{\lambda_\chi}{\kappa} \quad (47)$$

$$\tilde{V}[\ln(S_T) | \mathcal{F}_t] = (1 - e^{-2\kappa(T-t)}) \frac{\sigma_\chi^2}{2\kappa} + \sigma_\xi^2(T-t) + 2(1 - e^{-\kappa(T-t)}) \frac{\rho_{\chi\xi} \sigma_\chi \sigma_\xi}{\kappa} \quad (48)$$

hence, the logarithm of futures prices can be represented as

$$\ln(F(S_t, t, T)) = e^{-\kappa(T-t)} \chi_t + \xi_t + A(T, t) \quad (49)$$

and

$$A(T, t) = \mu_\xi^*(T-t) - (1 - e^{-\kappa(T-t)}) \frac{\lambda_\chi}{\kappa} + \frac{1}{2} \left( (1 - e^{-2\kappa(T-t)}) \frac{\sigma_\chi^2}{2\kappa} + \sigma_\xi^2(T-t) + 2(1 - e^{-\kappa(T-t)}) \frac{\rho_{\chi\xi} \sigma_\chi \sigma_\xi}{\kappa} \right) \quad (50)$$

$$\mu_\xi^* = \mu_\xi - \lambda_\xi \quad (51)$$

### Supplementary material

Supplementary material associated with this article can be found, in the online version, at doi: [10.1016/j.ejor.2019.05.042](https://doi.org/10.1016/j.ejor.2019.05.042).

### References

- Aiube, F., Baidya, T., & Tito, E. (2008). Analysis of commodity prices with the particle filter. *Energy Economics*, 30(2), 597–605.
- Almansour, A. (2016). Convenience yield in commodity price modeling: A regime switching approach. *Energy Economics*, 53, 238–247.
- Arismendi, J. C., Back, J., Propkopczuk, M., Paschke, R., & Rudolf, M. (2016). Seasonal stochastic volatility: Implication for the pricing of commodity options. *Journal of Banking and Finance*, 66, 53–65.
- Back, J., Propkopczuk, M., & Rudolf, M. (2013). Seasonality and the valuation of commodity options. *Journal of Banking and Finance*, 37(2), 273–290.
- Borovkova, S., & Geman, H. (2007). Seasonal and stochastic effects in commodity forward curves. *Review of Derivatives Research*, 9(2), 167–186.
- Burger, M., Graeber, B., & Schindlmayr, G. (2007). *Managing energy risk: An integrated view on power and other energy markets*. Wiley. Finance
- Carmona, R., & Coulon, M. (2013). A survey of commodity markets and structural models for electricity prices. In F. E. Benth, V. Kholodnyi, & P. Laurence (Eds.), *Proceedings from the special thematic year at the wolfgang pauli institute, vienna*.
- Cartea, A., & Figueroa, M. G. (2005). Pricing in electricity markets: a mean reverting jump diffusion model with seasonality. *Applied Mathematical Finance*, 12(4), 313–335.
- Cartea, A., & González-Pedraz, C. (2012). How much should we pay for interconnecting electricity markets? A real options approach. *Energy Economics*, 34(1), 14–30.
- Cartea, A., & Villaplana, P. (2008). Spot price modeling and the valuation of electricity forward contracts: The role of demand and capacity. *Journal of Banking and Finance*, 32(12), 2502–2519.
- Cartea, A., & Williams, T. (2008). UK gas markets: The market price of risk and applications to multiple interruptible supply contracts. *Energy Economics*, 30(3), 829–846.
- Casassus, J., Collin-Dufresne, P., & Routledge, B. R. (2005). Equilibrium commodity prices with irreversible investment and non-linear technology. Working paper 11864. National Bureau of Economic Research.
- Chiu, M. C., Wong, H. Y., & Zhao, J. (2015). Commodity derivatives pricing with cointegration and stochastic covariances. *European Journal of Operational Research*, 246(2), 476–486.
- Chkili, W., Hammoudeh, S., & Nguyen, D. K. (2014). Volatility forecasting and risk management for commodity markets in the presence of asymmetry and long memory. *Energy Economics*, 41, 1–18.
- Deng, S. (2001). Stochastic models of energy commodity prices and their applications: Mean-reversion with jumps and spikes. *Journal of Regulatory Economics*, 19(3), 239–270.
- Escribano, A., Ignacio Peña, J., & Villaplana, P. (2011). Modelling electricity prices: International evidence. *Oxford Bulletin of Economics and Statistics*, 73(5), 622–650.
- Ewald, C. O., Zhang, A., & Zong, Z. (2018). On the calibration of the Schwartz two-factor model to WTI crude oil options and the extended Kalman filter. *Annals of Operations Research*, available at. doi:10.1007/s10479-018-2770-x.
- Forsythe, P. (2007). A semi-lagrangian approach for natural gas valuation and optimal operation. *SIAM Journal on Scientific Computing*, 30(1), 339–368.
- Fouquau, J., & Six, P. (2015). A comparison of the convenience yield and interest-adjusted basis. *Finance Research Letters*, (14), 142–149.
- Geman, H. (2005). *Commodities and commodity derivatives: Modeling and pricing of agricultural, metals and energy*. Wiley. Finance
- Geman, H., & Nguyen, V. N. (2005). Soybean inventory and forward curve dynamics. *Management Science*, 51(7), 1076–1091.
- Gibson, R., & Schwartz, E. S. (1990). Stochastic convenience yield and the pricing of oil contingent claims. *Journal of Finance*, 45(3), 959–976.
- Gómez-Valle, L., Habibilashkary, Z., & Martínez-Rodríguez, J. (2017). A new technique to estimate the risk-neutral processes in jump/diffusion commodity futures models. *Journal of Computational and Applied Mathematics*, (309), 435–441.
- Gómez-Valle, L., Habibilashkary, Z., & Martínez Rodríguez, J. M. (2018). A multiplicative seasonal component in commodity derivative pricing. *Journal of Computational and Applied Mathematics*, 330, 835–847.
- Gould, P. G., Koehler, A. B., Ord, J. K., Snyder, R. D., Hyndman, R. J., & Vahid-Araghi, F. (2008). Forecasting time-series with multiple seasonal patterns. *European Journal of Operational Research*, 191(1), 207–222.
- Hamby, B., S. H., & Kluge, T. (2009). Modelling spikes and pricing swing options in electricity markets. *Quantitative Finance*, 9(8), 937–949.

- Hamilton, J. D. (1994). *Time Series Analysis*. Princeton, NJ: Princeton University Press.
- Hilliard, J. E., & Hilliard, J. (2015). Estimating early exercise premiums on gold and copper options using a multifactor model and density matched lattice. *Financial Review*, 50(1), 27–56.
- Hilliard, J. E., & Reis, J. (1998). Valuation of commodity futures and options under stochastic convenience yields, interest rates, and jump diffusion in the spot. *Journal of Financial and Quantitative Analysis*, 33(1), 61–86.
- Islyayev, S., & Date, P. (2015). Electricity futures price models: Calibration and forecasting. *European Journal of Operational Research*, 247(1), 144–154.
- Janczura, J. (2014). Pricing electricity derivatives within a Markov regime-switching model: A risk premium approach. *Mathematical Methods of Operations Research*, 79(1), 1–30.
- Kamat, R., & Oren, S. S. (2002). Exotic options for interruptible electricity supply contracts. *Operations Research*, 50(5), 835–850.
- Kyriakou, I., Nomikos, N. K., Papastolou, N. C., & Poliasis, P. K. (2016). Affine-structure models and the pricing of energy commodity derivatives. *European Financial Management*, 22(5), 853–881.
- Lai, A. N., & Mellios, C. (2016). Valuation of commodity derivatives with an unobservable convenience yield. *Computers and Operations Research*, 66, 402–414.
- Li, L., & Mendoza-Arriaga, R. (2013). Ornstein–Uhlenbeck processes time changed with additive subordinators and their applications in commodity derivative models. *Operations Research Letters*, 41(5), 521–525.
- Li, L., & Linetsky, V. (2013). Optimal stopping and early exercise: an eigenfunction expansion approach. *Operations Research*, 61(3), 625–646.
- Li, L., Mendoza-Arriaga, R., Mo, Z., & Mitchell, D. (2016). Modelling electricity prices: A time change approach. *Quantitative Finance*, 16(7), 1089–1109.
- Liu, P., & Tang, K. (2011). The stochastic behavior of commodity prices with heteroskedasticity in the convenience yield. *Journal of Empirical Finance*, 18(2), 211–224.
- Lucia, J., & Schwartz, E. S. (2002). Electricity prices and power derivatives: Evidence from the nordic power exchange. *Review of Derivatives Research*, 5(1), 5–50.
- Manoliu, M., & Tompaidis, S. (2002). Energy futures prices: Term structure models with Kalman filter estimation. *Applied Mathematical Finance*, 9(1), 21–43.
- Mayer, K., Schmid, T., & Weber, F. (2011). Modeling electricity spot prices - combining mean-reversion, Spikes and Stochastic Volatility, CEFS Working Paper series, No. 2011-02.
- Mirantes, A. G., Población, J., & Serna, G. (2012). The stochastic seasonal behaviour of natural gas prices. *European Financial Management*, 18(3), 410–443.
- Mirantes, A. G., Población, J., & Serna, G. (2015). Commodity derivative valuation under a factor model with time-varying market prices of risk. *Review of Derivatives Research*, 18, 75–93.
- Moreno, M., Novales, A., & Platania, F. (2018). A term structure model under cyclical fluctuations in interest rates. *Economic Modelling*, 72, 140–150.
- Moreno, M., & Platania, F. (2015). A cyclical square-root model for the term structure of interest rates. *European Journal of Operational Research*, 241(1), 109–121.
- Mu, X., & Ye, H. (2015). Small trends and big cycles in crude oil prices. *Energy Journal*, 36(1), 49–72.
- Pellegrino, T., & Sabino, P. (2014). Pricing and hedging multiasset spread options using a three-dimensional fourier cosine series expansion method. *Journal of Energy Markets*, 7(2), 71–92.
- Sbuelz, A. (2015). The Schwartz and Smith (2000) model with state-dependent risk premia. *Mathematical Finance Letters*, 7, 1–7.
- Schmitz, A., Wang, Z., & Kimn, J. H. (2014). A jump diffusion model for agricultural commodities with Bayesian analysis. *Journal of Futures Markets*, 34(3), 235–260.
- Schwartz, E. S. (1997). The stochastic behaviour of commodity prices: Implications for valuation and hedging. *Journal of Finance*, 52(3), 923–973.
- Schwartz, E. S., & Smith, J. (2000). Short-term variations and long-term dynamics in commodity prices. *Management Science*, 46(7), 893–911.
- Sévi, B. (2015). Explaining the convenience yield in the WTI crude oil market using realized volatility and jumps. *Economic Modelling*, 44, 243–251.
- Sørensen, C. (2002). Modeling seasonality in agricultural commodity futures. *Journal of Futures Markets*, 22(5), 393–426.
- Steele, J. M. (2003). *Stochastic calculus and financial applications*. Springer-Verlag.
- Taylor, J. W. (2010). Triple seasonal methods for short-term electricity demand forecasting. *European Journal of Operational Research*, 204(1), 139–152.
- Vasicek, O. (1977). An equilibrium characterization of the term structure. *Journal of Financial Economics*, 5(2), 177–188.
- Weron, R. (2007). *Modeling and forecasting electricity loads and prices: A statistical approach*. Wiley. Finance
- Wong, H. Y., & Lo, Y. W. (2009). Option pricing with mean reversion and stochastic volatility. *European Journal of Operational Research*, 197(1), 179–187.
- Yan, X. (2002). Valuation of commodity derivatives in a new multi-factor model. *Review of Derivatives Research*, 5(3), 251–271.