

# Historical and risk-neutral estimation in a two factors stochastic volatility model for oil markets

Gaetano Fileccia and Carlo Sgarra\*

Dipartimento di Matematica,  
Politecnico di Milano, Piazza Leonardo Da Vinci,  
32-20133 Milano, Italy

Email: gaetano.fileccia@polimi.it

Email: carlo.sgarra@polimi.it

\*Corresponding author

## 1 Introduction

Stochastic volatility models are a well-known choice in commodity markets. Geman (2005) and Hikspoors (2008) discussed some of the most popular models of this kind in commodity finance. The seminal paper, to which they refer is the one by Schwartz (1997), where the commodity spot values are modelled by a mean reverting process and the convenience yield is incorporated in the discount factor. While Cortazar and Schwartz (2003) proposed an extension of the original Schwartz model adding one factor to describe the long-term interest rate by an Ornstein-Uhlenbeck process, Eydeland and Geman (1998) proposed a different extension adding a mean reverting OU-process to describe the instantaneous variance of the diffusion term in the spot dynamics.

As Geman (2000) pointed out, the use of mean reversion in the dynamics of spot values is controversial, in particular for crude oil market (the object of our analysis) it is not always needed. In the same work it was suggested a simple test to discriminate if a mean reverting dynamics is preferable to a simple diffusion process, the results shows that for time series starting from 2005 there is no strong evidence for a mean reverting modelling choice.

Further developments were implemented, as in Ribeiro and Hodges (2004), to describe the dynamics for the oil sector of commodity market. Moreover, in their paper, they also pointed out that the possibility for convenience yield to assume negative values (to explain the contango effect in commodity market) from a theoretical viewpoint is due to the cost of storage implied in prices dynamics; if this term was separated off the convenience yield, then a Cox-Ingersoll-Ross (CIR) process for the convenience yield would automatically exclude arbitrage opportunities in the market. In the model proposed in this paper, we stand on this remark by Ribeiro and Hodges, modelling the convenience yield (once the cost of storage is separated off) with a CIR dynamics but without imposing any linear dependence of the spot dynamics diffusion term on the convenience yield; on the contrary, we prefer to include an extra volatility factor (with a CIR dynamics, like in the well known Heston model. The possibility of jumps with finite activity (modelled by a compound Poisson process with normally distributed jump size) is allowed in a second version of the model, which has been tested with the same dataset used for inference in the model without jumps. A third version of the model substitutes the jump term with a seasonality term (modelled by the usual periodic function, as suggested also by Hikspoors (2008)), allowing for a comparison between the effect of jumps and of the seasonality term in the goodness of fit of analysed dataset. Dataset used for the estimation come from the West Texas intermediate (WTI) index spot values, quoted in the New York Market Exchange (NYMEX), and the futures written on it, more details about datasets are provided in a dedicated section. The time window considered spans from the beginning of February 2007 to the end of September 2010. In this period the oil market exhibited an inverted market state (also known as contango), and we expect our models can capture this phenomenon.

Inference under both the historical and the risk-neutral measures are quite common in literature for models of this kind, and they could be easily extended to the present one. Some

authors resorted to Bayesian analysis to get inference in stochastic volatility framework: since the seminal paper by Jacquier et al. (1994) a fast-growing literature is exploring the application of Bayesian estimation techniques, in particular Markov Chain Monte Carlo (MCMC) methods, to get inference for models belonging to this family. A branch of this literature is devoting to study joint estimation under historical and risk neutral measure, using both stock and derivatives prices data. The link between the two measures is provided by a suitable parametric choice of the Radon-Nikodym derivative.

The main references in this section are the two papers by Eraker (2004) and by Eraker et al. (2003) in which some popular stochastic volatility models are analysed using a Gibbs sampling algorithm. Further references in which the Gibbs sampling method is used to get inference for stochastic volatility models, eventually including jumps, are the papers by Forbes et al. (2002) and, more recently, by Yu et al. (2011), who extended the results by Eraker (2004) including different jump models in the analysis and compare them; all these paper refer to equity markets, analysing S&P500 or DAX data. Another Bayesian technique which is becoming popular to get inference in a stochastic volatility framework, and when latent factor are present in general, is the particle filter (PF) method. This is a Bayesian filter algorithm based on sequential importance sampling for Bayesian networks; differently from Kalman filter, it can be used also for non-gaussian and non-linear dynamics. A complete survey on the theoretical background and implementation details on PF are the papers by Andrieu and Johansen (2008), and Arulampalam et al. (2002), while for application to stochastic volatility models we mention the papers by Javaheri (2005), Johannes et al. (2009) and Aihara et al. (2008). Different authors have developed efficient estimation methods based on PF, by combining the PF approach with MCMC algorithms; in particular, Andrieu et al. (2010), in order to get inference on the parameters of a stochastic volatility model with jumps, introduced a particle MCMC algorithm in which the marginal likelihood is estimated by a nested Auxiliary PF. We are going to apply a similar algorithm in order to carry out our inference on the model class under investigation.

## 2 The models proposed

We study three different models for oil price dynamics. We start by studying a basic model (Model B), then we analyse two possible extensions. The basic model (Model B) is a two-factor model, including a volatility process and a convenience yield process into the spot price dynamics. Both the spot variance and the convenience yield processes follow a square-root CIR type dynamics: the coefficients  $\alpha$  and  $\beta$  appearing in the stochastic differential equation describing this dynamics are the so-called mean-reversion speeds, for the convenience yield and the volatility respectively, while the coefficients  $\sigma$  and  $\xi$  are the constant of proportionality characterising the diffusion terms of the square root processes, respectively for the convenience yield and for the volatility again. Under the historical measure  $\mathbb{P}$  the dynamics described by the model B is then the following:

$$\begin{cases} \frac{dS_t}{S_t} = (\mu + c - \delta_t)dt + \sqrt{V_t}dW_{S_t}^{(\mathbb{P})} \\ d\delta_t = \alpha(\bar{\delta} - \delta_t)dt + \sigma\sqrt{\delta_t}dW_{\delta_t}^{(\mathbb{P})} \\ dV_t = \beta(\bar{V} - V_t)dt + \xi\sqrt{V_t}dW_{V_t}^{(\mathbb{P})} \\ dW_{S_t}^{(\mathbb{P})} dW_{V_t}^{(\mathbb{P})} = \rho dt \end{cases} \quad (1)$$

This basic model we just introduced (and never introduced before, to our knowledge) shares some relevant features with the model proposed by Yan (2002), but it assumes the interest rates to be constant and the convenience yield dynamics to be driven by a further square-root process. As Schwartz (1997) pointed out, after a systematic comparison of performances between models with deterministic and stochastic interest rates in capturing the futures observed structure, models with stochastic interest rates do not show any significant improvement, if we exclude the longest maturity futures (about ten-years maturity). Since the biggest maturity taken into account in the present paper is five years, we prefer to avoid stochastic modelling for interest rates in order to make the model more parsimonious. We decided then to keep the rates at the FED levels. Moreover, since the cost of storage is separated out from the dynamics of the convenience yield,  $\delta_t$  can be modelled by a CIR process that prevents it from assuming negative values, automatically excluding arbitrage opportunities in the market, as pointed out in Ribeiro and Hodges (2004). In the paper by Yan (2002) the parameter estimation problem was not examined, so the main goal of this paper is to develop an estimation procedure for a whole class of models strictly related to that proposed by Yan. The stochastic volatility feature combined with the stochastic dynamics for the convenience yield and the eventual inclusion of jumps make the model considered flexible enough in order to describe the oil market dynamics, by keeping at the same time a high level of analytical tractability: we can obtain in fact for futures and option prices close-form solutions. The procedure, we are going to develop in this paper, allows moreover to estimate accurately and efficiently the model parameters both under the historical and the risk-neutral dynamics specification. Similar results, for a different model class, have been presented in the paper by Peters et al. (2013). In equation (1) we identify the spot value at time  $t$  with  $S_t$ , while  $\delta_t$  identifies the net convenience yield process (after separating the positive revenues from the cost of storage and insurance,  $c^1$ .) and with  $V_t$  the variance process.  $W_{S_t}$ ,  $W_{\delta_t}$ ,  $W_{V_t}$  are three Brownian motions, with  $W_{\delta_t}$  independent from the other two process, while  $W_{S_t}$  and  $W_{V_t}$  are correlated as described in the fourth equation of the previous system (1).

In order to understand separately the role plaid by the different factors and their significance in explaining the spot price dynamics, we consider moreover two variants of model B: the first allowing for jumps (Model J) in the spot dynamics (modelled with exponentially distributed jump arrivals time and lognormally distributed jump size), and a model including a seasonality term in spot dynamics (Model S). This last model includes in the spot dynamics the usual sinusoidal term (as discussed by Hikspoors (2008)).

Modelling prices dynamics by processes including both stochastic volatility and jumps is getting more and more popular for several different kind of traded assets, since only the combined effects of both features can provide reasonable volatility smiles for the on the whole maturity spectrum (See, for example, Cont and Tankov (2004)). The seasonality terms, appearing in the oil price dynamics, is on the other hand, a distinctive feature of energy commodities: the rising of energy demand in both the cold and the hot season, and the fall of demand during the rest of the year, create the characteristic periodic behaviour of the deterministic component in the oil price dynamics.

All the models taken into account (model B, S and J) share similar dynamics for the convenience yield process and the volatility process, while the dynamics for the spot process is different.

All the dynamics are resumed by the following general system of stochastic differential equations:

$$\begin{cases} dx_t = (\mu + c - \delta_t - \frac{1}{2}V_t)dt + \sqrt{V_t}dW_{x_t}^{(\mathbb{P})} + dJ_{x_t}^{(\mathbb{P})} \\ d\delta_t = \alpha(\bar{\delta} - \delta_t)dt + \sigma\sqrt{\delta_t}dW_{\delta_t}^{(\mathbb{P})} \\ dV_t = \beta(\bar{V} - V_t)dt + \xi\sqrt{V_t}dW_{V_t}^{(\mathbb{P})} \\ S_t = \exp\{g(t_{\text{year}}) + x_t\} \\ g(t_{\text{year}}) = \zeta_1(\sin(2\pi t_{\text{year}} + \omega_1)) + \zeta_2(\cos(2\pi t_{\text{year}} + \omega_2)) \\ dW_{x_t}^{(\mathbb{P})} dW_{V_t}^{(\mathbb{P})} = \rho dt \end{cases} \quad (2)$$

with  $t_{\text{year}}$  the time elapsed since the beginning of the year (we adopt the word ‘year’ here and in the rest of the paper to mean calendar year).  $J_{x_t}^{(\mathbb{P})}$  is a jump process, where the jump arrivals time are distributed according to an exponential random variable with  $\lambda_J$  parameter, and the jump size are normally distributed with mean  $\mu_J$  and variance  $\sigma_J^2$ . We make the usual choice of describing the price dynamics through the log-price process:  $x_t = \log[S_t]$  eventually by multiplying this term by the seasonality function  $g$  whenever this will be assumed to be different from zero.

From equation (2) we can obtain the three alternative models we are going to study by fixing some of the parameters:

- *model B*: we fix  $\lambda_J$ , the parameter of Poisson-distributed jump time arrivals, to zero and  $\zeta_1 = \zeta_2 = 0$
- *model J*: we fix  $\zeta_1 = \zeta_2 = 0$
- *model S*: we fix  $\lambda_J = 0$ .

The model B, which is our basic model, is obtained by the general dynamics described in equation (2) by removing both the seasonality terms and the jump component in the spot price dynamics, while in model J the seasonality terms are removed and the jump component is taken into account, and in model S the reverse situation is considered (the jump component is removed and the seasonality terms are taken into account).

In order to describe the dynamics under the risk-neutral measure (Shreve, 2003; Bjork, 2009), we need to define the Radon-Nikodym derivative of this measure with respect to the historical one. The parametric form we choose is the same proposed by Heston (1993) and by Pan (2002), which preserves the model structure under the measure change. This can be specified by the following relations between the Wiener processes under the two measures, provided by the Girsanov theorem:

$$\begin{cases} dW_{\delta_t}^{(\mathbb{Q})} = dW_{\delta_t}^{(\mathbb{P})} + \frac{\eta_{\delta}}{\sigma}\sqrt{\delta_t}dt \\ dW_{V_t}^{(\mathbb{Q})} = dW_{V_t}^{(\mathbb{P})} - \frac{1}{\sqrt{1-\rho^2}}\left(\rho\eta_{S_t} - \frac{\eta_V}{\xi}\right)\sqrt{V_t}dt \\ dW_{x_t}^{(\mathbb{Q})} = dW_{x_t}^{(\mathbb{P})} + \eta_{S_t}\sqrt{V_t}dt \end{cases} \quad (3)$$

where we have denoted by  $\eta_S$  the risk premium associated with the Brownian part of the spot dynamics. Hence, the dynamics described under the measure  $\mathbb{P}$  by the system (2), under the risk-neutral measure  $\mathbb{Q}$  becomes:

$$\begin{cases} dx_t = (r + c - \delta_t - \frac{1}{2}V_t - \mu^*)dt + \sqrt{V_t}dW_{x_t}^{(\mathbb{Q})} + dJ_{x_t}^{(\mathbb{Q})} \\ d\delta_t = (\alpha(\bar{\delta} - \delta_t) - \eta_\delta\delta_t)dt + \sigma\sqrt{\delta_t}dW_{\delta_t}^{(\mathbb{Q})} \\ dV_t = (\beta(\bar{V} - V_t) - \eta_V V_t)dt + \xi\sqrt{V_t}dW_{V_t}^{(\mathbb{Q})} \\ S_t = \exp\{g(t_{\text{year}}) + x_t\} \\ dW_{S_t}^{(\mathbb{Q})}dW_{V_t}^{(\mathbb{Q})} = \rho dt \\ g(t_{\text{year}}) = \zeta_1(\sin(2\pi t_{\text{year}} + \omega_1)) + \zeta_2(\cos(2\pi t_{\text{year}} + \omega_2)), \end{cases} \quad (4)$$

where  $\mu^*$  and  $dJ^{(\mathbb{Q})}$  denote respectively the predictable compensator and the compensated jump measure associated with the jump process.<sup>2</sup> Moreover, since  $\delta_t - c$  can assume both positive and negative values, normal and inverted futures market structures are both allowed. To cope with the inference in this setting we adopt a Euler discretisation method. A comprehensive and detailed reference providing efficient methods for computing approximate solutions of stochastic differential equations of the kind under examination is the book by Glasserman (2004). By approximating the deterministic differentials by forward finite increments, the Brownian motions by normally distributed random variables with zero mean and variance  $\Delta t$ , and by simulating the jump dynamics by a compound Poisson process  $N_t^{(J)}$ , the discretised dynamics (under the risk neutral measure) can then be written as follows:

$$\begin{cases} x_{t+1} = x_t + (r + c - \delta_t - \frac{1}{2}V_t - \mu^*)\Delta t + \sqrt{1 - \rho^2}\sqrt{V_t}\epsilon_t^{(S)} \\ \quad + \rho\sqrt{V_t}\epsilon_t^{(V)} + \sum_{i=1}^{N_t^{(J)}} \epsilon_{i,t}^{(J)} \\ \delta_{t+1} = \delta_t + (\alpha(\bar{\delta} - \delta_t) - \eta_\delta\delta_t)\Delta t + \sigma\sqrt{\delta_t}\epsilon_t^{(\delta)} \\ V_{t+1} = V_t + (\beta(\bar{V} - V_t) - \eta_V V_t)\Delta t + \xi\sqrt{V_t}\epsilon_t^{(V)} \end{cases} \quad (5)$$

where each  $\epsilon_t^{(S)}$ ,  $\epsilon_t^{(V)}$  and  $\epsilon_t^{(\delta)}$  are normally distributed random variables with zero mean and variance  $\Delta t$ . In the discretised jump term, the  $N_t^{(J)}$  are independent Poisson distributed random variables with intensity parameter  $\lambda_J$ , and  $\epsilon_{i,t}^{(J)}$  are independent normal random variables with mean  $(\mu_J - \eta_J)$  and variance  $\sigma_J^2$ . All the random variables just listed are independent on each other.

Under the historical measure the same discretisation holds, with the risk premia  $\{\eta_\delta, \eta_V, \eta_J\}$  and the compensator  $\mu^*$  set to zero, while the drift coefficient  $r_f$  is substituted by  $\mu$ .

The relation between the compensator  $\mu^*$  and the parameters characterising the Gaussian distribution of the jump size,  $(\mu_J, \sigma_J^2)$ , is provided by the following formula:

$$\mu^* = \exp\left\{\mu_J + \frac{1}{2}\sigma_J^2\right\} - 1 \quad (6)$$

With these model specifications, the futures prices are given by the following expression:

$$F(0, \tau) = \exp\{A_0(\tau) + x_t + A_2(\tau)\delta_t\} \quad (7)$$

where  $\tau = T - t$  is the futures time to maturity. Details about computations and the specifics for the functions  $A_0(\tau)$  and  $A_2(\tau)$  are provided in the next section. Formula (7) holds for two of the three models studied: model B and model J. Futures prices for the model with seasonality (model S) include a seasonal term: under model S the futures price formula becomes:

$$F(0, \tau) = \exp\{g(T_{\text{year}}) + A_0(\tau) + x_t + A_2(\tau)\delta_t\}. \quad (8)$$

where  $T_{\text{year}}$  is the time in years the maturity day differ from the first of January of the same year. Since the prices of a futures are affected by different kind of noises (possible incomplete specification of the model, market inefficiency, random noise, etc.) we modelled (as in Eraker et al. (2003)) the logarithm of the price of the futures making the hypothesis that the observed market price,  $\ln F^M(0, \tau)$ , are represented by the theoretical price got by equation (7) plus an error distributed as a white noise  $\epsilon_{\text{fut}} \sim \mathcal{N}(0, \sigma_\epsilon^2)$ :<sup>3</sup>

$$\ln F^M(0, \tau) = \ln F(0, \tau) + \epsilon_{\text{fut}}. \quad (9)$$

### 3 Futures prices

When the underlying asset follows the dynamics (4), the value of a generic futures contract,  $f(t, x_t, \delta_t, V_t, J_t)$ , can be computed by the following backward partial differential equation (PDE):

$$\begin{aligned} & \frac{\partial f}{\partial t} + \frac{\partial f}{\partial x} \left( r_f + c - \mu^* - \delta_t - \frac{1}{2}V_t \right) + \frac{\partial f}{\partial \delta} [\alpha(\bar{\delta} - \delta_t) - \eta_\delta \delta_t] \\ & + \frac{\partial f}{\partial V} [\beta(\bar{V} - V_t) - \eta_V V_t] + \frac{1}{2} \frac{\partial^2 f}{\partial x^2} V_t + \frac{1}{2} \frac{\partial^2 f}{\partial \delta^2} \sigma^2 \delta_t + \frac{1}{2} \frac{\partial^2 f}{\partial V^2} \xi^2 V_t \\ & + \frac{\partial^2 f}{\partial V \partial x} \rho \xi V + \lambda \mathbb{E}^{(\mathbb{Q})} [f(t, x_t + \ln(1 + J_t), \delta_t, V_t) - f(t, x_t, \delta_t, V_t)] = 0. \end{aligned}$$

This PDE can be easily obtained as follows: by applying the Ito Lemma to the pricing function  $f(t, x, v, \delta)$  we get:

$$\begin{aligned} & \frac{\partial f}{\partial t} + \frac{\partial f}{\partial x} (r_f + c - \mu^* - \delta_t - \frac{1}{2}V_t) + \frac{\partial f}{\partial \delta} [\alpha(\bar{\delta} - \delta_t) - \eta_\delta \delta_t] \\ & + \frac{\partial f}{\partial V} [\beta(\bar{V} - V_t) - \eta_V V_t] + \frac{1}{2} \frac{\partial^2 f}{\partial x^2} V_t + \frac{1}{2} \frac{\partial^2 f}{\partial \delta^2} \sigma^2 \delta_t + \frac{1}{2} \frac{\partial^2 f}{\partial V^2} \xi^2 V_t \\ & + \frac{\partial^2 f}{\partial V \partial x} \rho \xi V + \lambda \mathbb{E}^{(\mathbb{Q})} [f(t, x_t + \ln(1 + J_t), \delta_t, V_t) - f(t, x_t, \delta_t, V_t)] \\ & + \frac{\partial f}{\partial x} \sqrt{V_t} dW_{x_t}^{(\mathbb{Q})} + \frac{\partial f}{\partial V} \xi \sqrt{V_t} dW_{V_t}^{(\mathbb{Q})} + \frac{\partial f}{\partial \delta} \sigma \sqrt{\delta_t} dW_{\delta_t}^{(\mathbb{Q})} + dJ_{x_t}^{(\mathbb{Q})}. \end{aligned}$$

The fundamental theorem of asset pricing (See Bjork, 2009) implies that, in order to avoid arbitrage opportunities, every contingent claim written on  $S$  must be a martingale with respect to  $Q$ . Since the terms involving the Wiener processes  $dW^{(\mathbb{Q})}$  and the compensated jump measure  $dJ_{x_t}^{(\mathbb{Q})}$  are  $Q$ -martingales by definition, the only term which prevents the pricing function  $f$  to be a martingale is the drift term, so this term must vanish, and this

requirement implies the equation. The previous PDE, in order to admit a solution, requires the specification of a terminal condition  $f(t = T) = H(x_T, \delta_T, V_T)$ , with  $H$  payoff at maturity  $T$ , and since we are dealing with futures, the final payoff is provided by the following expression:

$$H(x_T, \delta_T, V_T) = \exp\{x_T\}. \quad (10)$$

The approach we are going to follow in the present section is the same as that presented in the paper by Yan (2002), Section 1.4, where the computations are exposed in full detail; this approach can be successfully applied to every model belonging to the large class of the so-called 'Affine Models'. For a general survey on these models and their properties a standard reference is the paper by Duffie et al. (2003). In Appendix B is included the computation of the Fourier transform of the options on futures prices. According to the general theory of Affine Processes the solution of equation (10) must be of the following form:

$$f_t = \exp\{A_0(t) + A_1(t)x_t + A_2(t)\delta_t + A_3(t)V_t\}. \quad (11)$$

If we substitute this tentative solution into our PDE we get a polynomial in the independent variables and, if we impose the vanishing of all the terms with the same power in each independent variable, we get the following system of ordinary differential equations:

$$\begin{cases} -\frac{A'_0(\tau)}{\tau} + A_1(\tau)(r_f + c) + A_2(\tau)\alpha\bar{\delta} + A_3\beta\bar{V} = 0 \\ -\frac{A'_1(\tau)}{\tau} = 0 \\ -\frac{A'_2(\tau)}{\tau} - A_1(\tau) - A_2(\tau)(\alpha + \eta_\delta) + \frac{1}{2}A_2^2(\tau)\sigma^2 = 0 \\ -\frac{A'_3(\tau)}{\tau} - \frac{1}{2}A_1(\tau) + \frac{1}{2}A_1^2(\tau) - A'_3(\tau)(\beta + \eta_V) + \frac{1}{2}A_3^2(\tau)\xi^2 \\ + A_1(\tau)A_3(\tau)\rho\xi = 0, \end{cases} \quad (12)$$

where we changed variable from  $t$  to  $\tau = T - t$ .<sup>4</sup>

Since the payoff is equation (10) we obtain the initial condition for the ODE system:

$$\begin{cases} A_0(0) = 0 \\ A_1(0) = 1 \\ A_2(0) = 0 \\ A_3(0) = 0 \end{cases} \quad (13)$$

The previous system of linear differential equations can be solved via standard methods (see, for example, Boyce and Di Prima, 2012), which provide the following solution:

$$\begin{cases} A_0(\tau) = (r_f + c)\tau - \frac{2\alpha\bar{\delta}}{\sigma^2} \left[ B\tau + \log \left( \frac{D - B \exp\{C\tau\}}{D - B} \right) \right] \\ A_1(\tau) = 1 \\ A_2(\tau) = -\frac{2}{\sigma^2} \frac{\exp\{C\tau\} - 1}{\frac{\exp\{C\tau\}}{D} - \frac{1}{B}} \\ A_3(\tau) = 0 \end{cases} \quad (14)$$

with:

$$\begin{cases} C = \sqrt{(\alpha + \eta_\delta)^2 + 2\sigma^2} \\ D = \frac{(\eta_\delta + \alpha) + C}{2} \\ B = \frac{(\eta_\delta + \alpha) - C}{2} \end{cases} \quad (15)$$

#### 4 The inference algorithm

The object of our inference is the vector  $\{\Theta, V_{0:T}, \delta_{0:T}, J_{0:T}\}$  where  $V_{0:T}$ ,  $\delta_{0:T}$ ,  $J_{0:T}$  are the three latent processes (the variance process, the convenience yield and the jump process) and  $\Theta$  is the set of parameters, that is our main inference target:  $\Theta = \{\epsilon, \mu, c, \alpha, \bar{\delta}, \sigma, \eta_\delta, \beta, \bar{V}, \xi, \rho, \lambda, \mu_J, \sigma_J, \eta_J\}$ .<sup>5</sup> For simplicity we shall indicate with  $X_{0:T}$  the vector of the three latent processes  $\{V_{0:T}, \delta_{0:T}, J_{0:T}\}$ , and by  $Z_{1:T}$  the set of observed market data (coming from both asset price and futures price quotations).

To get inference in a so large state space, particle Markov Chain Monte Carlo methods (PMMC henceforth) represents an efficient technique, since it allows to simulate the latent processes in a single simulation block. A presentation of PMMC technique applied to econometric inference in commodity financial markets has recently appeared in Peters et al. (2013). The PMMC allows to sample from  $p(\Theta, X_{0:T}|Z_{1:T})$ , that is the joint probability distribution of the parameter set and the latent processes given the observed data. To sample with a Monte Carlo technique from such a probability distribution we used a PF algorithm<sup>6</sup> to estimate the marginal likelihood  $L(y|\Theta)$ , this which is used in defining an acceptance ratio probability that will ensure that after a sufficient amount of time, whatever point in state space we started from, we will sample from the distribution  $p(\Theta|Z_{1:T})$ , getting unbiased estimate for  $\Theta$ .

The MCMC method used is a variant of the well known Metropolis-Hastings algorithm (a thorough discussion about theory and implementation of MCMC methods can be found in Gilks (1996))

- At step ( $s = 0$ )  
after setting a starting point  $\Theta = \theta(0)$  arbitrarily, then a Sampling Important Resampling algorithm is implemented. SIR<sup>7</sup> algorithm allows to simulate the latent processes  $x_{1:T}^{(0)}$  from the distribution  $p(X_{1:T}|Z_{1:T}, \theta(0))$  and an estimate of the marginal likelihood  $p(Z_{1:T}|\theta(0))$
- At step ( $s$  from 1 to MC, the length of the Markov chain sequence we want to simulate)  
a new value set,  $\theta^*$ , is sampled from a proposal (symmetric) distribution  $q(\cdot|\theta(s-1))$  and the new sampling value is accepted with a probability:

$$\min \left( 1, \frac{p(z_{1:T}|\theta^*)\tilde{p}(\theta^*)}{p(z_{1:T}|\theta(s-1))\tilde{p}(\theta(s-1))} \right) \quad (16)$$

where  $\tilde{p}(\theta)$  is the prior distribution associated to  $\theta$ , while  $p(z_{1:T}|\theta^*)$  is the likelihood for the observation chain given the parameter set  $\theta^*$ .

Again the marginal likelihood and  $x_{1:T}^*$  comes from a PF algorithm run.

If accepted:  $\theta(s) = \theta^*$ , otherwise:  $\theta(s) = \theta(s-1)$ .



According to this algorithm we are sampling from a Markov chain whose limiting distribution is:

$$p(\Theta, X_{1:T} | Z_{1:T}) = p(\theta | z_{1:T}) p_{\theta}(x_{1:T} | z_{1:T}) \quad (17)$$

We can also discard the information about hidden state process and use the sample to get inference about  $\theta$ , if we need just the last one.

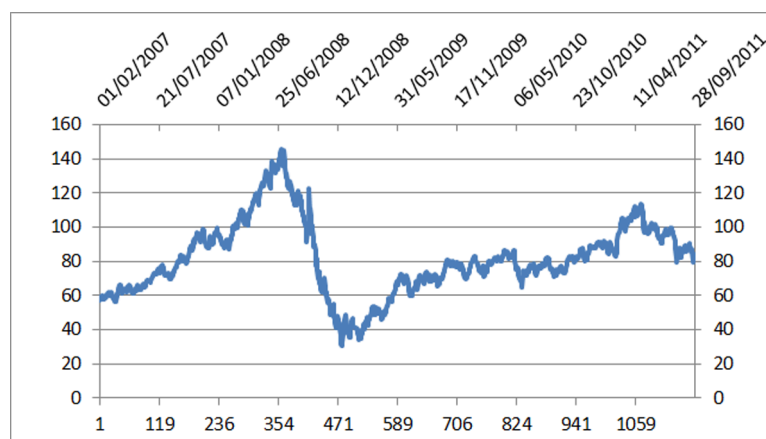
To reduce the number of rejected proposal, and increase mixing of the chain different techniques are known, as Metropolis within Gibbs variant, that is the one we implemented in our sampling algorithm. The output coming from the Particle Metropolis Hastings algorithm (PMMH), as any MCMC output need to be resized, removing the burn in, that is the part of the chain needed by algorithm to get to the stationary distribution. To check that convergence to stationary distribution was reached (for all the parameters) we adopted the Geweke test (Geweke, 1992).

## 5 The data

The data used for the analysis are relative to WTI Cushing Crude Oil spot and futures quotations on NYMEX market from 1 February, 2007 to 28 September, 2011. We used the the first 85% if the dataset to get the inference on our model and reserved the last 15% of dates for out of the sample performances study. The range of dates for the estimation goes from 1 February, 2007 to 31 December, 2010, and the spot data are presented in Figure 1.

Spot data are collected from the US Energy information administration webpage where a large collection of energy related time series (among which the WTI FOB spot prices) is provided. Daily data are taken into account for any available working day in the interval. A plot of the spot data used in estimating the parameters of the different analysed models is shown in Figure 1.

**Figure 1** Spot FOB data for WTI crude oil. The  $x$ -axis label (the lower one) refers to the number of working days from the starting date 1 February, 2007 (see online version for colours)



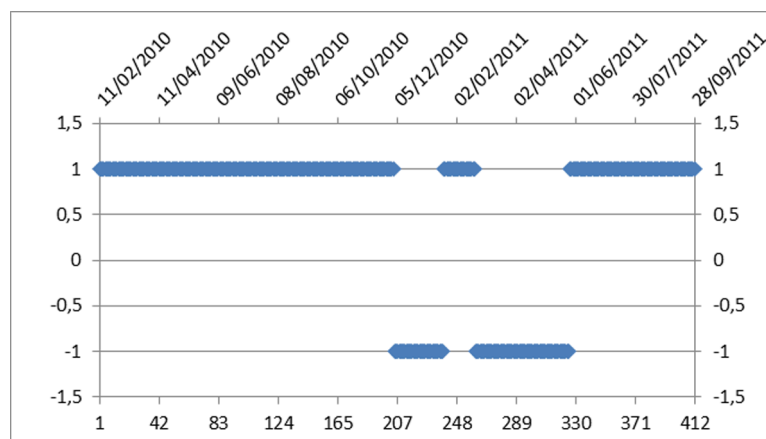
Beside the spot data, for any working day we recorded a panel of 12 future contract values. Their maturity day is fixed on the first working day of each month of 2012. So for any trading

day we analysed a spot datum and 12 futures data, and in the range there are 988 dates. Hence the dataset consists in 988 spot values and 11,856 future contracts.

In addition to this dataset we reserved a panel of data to evaluate out of the sample performances. The dataset include again a FOB spot datum and a panel of 12 future data (with different maturity one for each month of 2012, the maturities are set on the first trading day), the range of dates goes from 1 January, 2011 to 28 September, 2012. The number of working days in the range is 187.

Future contracts are usually characterised by their behaviour when time to maturity goes to zero. If at a certain date the futures quotation increases when the time to maturity become longer, it usually said the future market is normal, otherwise is said inverted. In Figure 2 are indicated the dates for which we can observe a normal future market (equivalent to the value +1) and the dates with an inverted future market (equivalent to the value -1).

**Figure 2** The dates in the dataset used for the parameter estimation are divided in normal future market days (+1 value) and inverted future market days (-1 value). The x-axis label (the lower one) refers to the number of working days from the starting date 1 February, 2007 (see online version for colours)

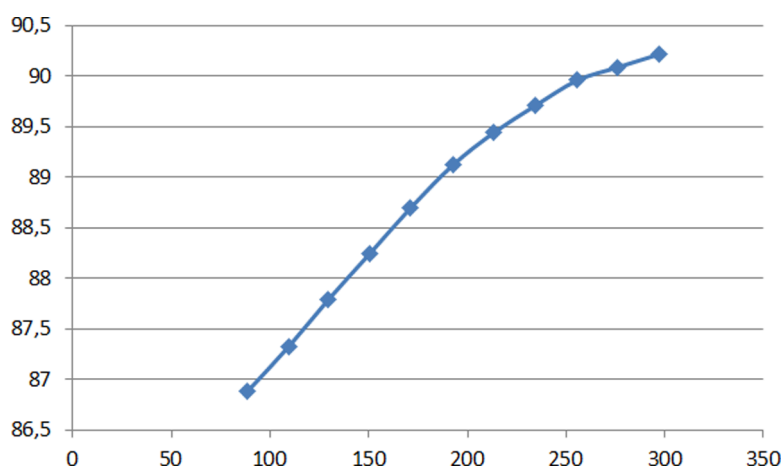


Both the future market behaviours have been found in the different analysed periods. The two behaviours are illustrated by Figures 3 and 4 which show, respectively, an example of normal market structure (the future values on 11 May, 2011 with higher market values for futures with longer time to maturity) and an example of inverted market structure (the future values on 25 September, 2011 with higher market values for futures with shorter time to maturity).

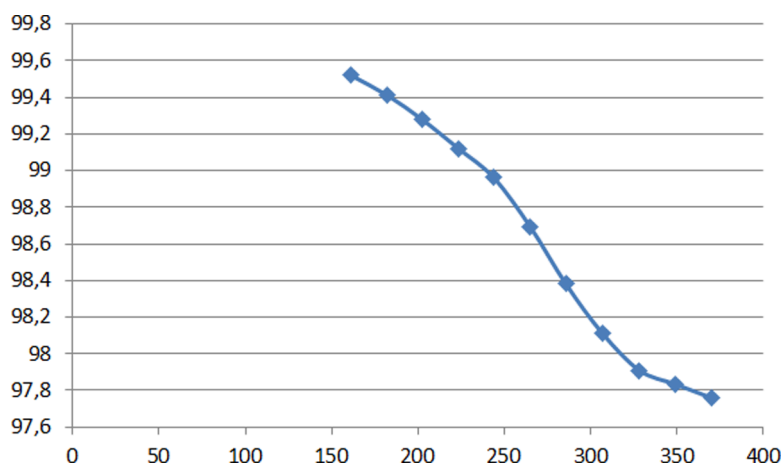
From the available range of dates we extracted a second dataset from 11 February, 2010 to 28 September, 2011; again we used about the first 85% of the dates to get inference and the remaining part to evaluate out of the sample performances evaluation. The dataset used for the inference goes from 11 February, 2010 to 30 July, 2011. The spot quotations are reported in Figure 5. With respect to the first dataset, the period from June 2008 to December 2008, when a disastrous fall in FOB quotations was observed in the market (the barrel price went from about 150\$ to about 30\$, before rising up again to a more stable value of about 100\$), has been excluded. By restricting the dataset to one year data and excluding the exceptional period just discussed, from our analysis, we hope to find a description of the market more reliable than that provided by the first dataset. We shall evaluate this hypothesis by analysing

both the residuals and the out-of-the sample performances in Section 7. In Figure 6 are reported, as done for the previous dataset, with dates characterised by a normal future market structure and with dates characterised by an inverse future structure. We excluded the model with seasonality from the analysed models, since the dates range spans just one year and a half, hence any inference on yearly periodic functions would be affected by the too short range considered. The number of working dates in this range is 350 and, as before, we considered for each date one spot FOB quotation and 12 futures for each date, hence our inference dataset consists in 350 spot data and 4200 futures data. The out of the sample performances are evaluated on the dataset consisting in 81 spot data and 972 futures data, these data correspond to the range of dates from 1 July, 2011 to 28 September, 2011.

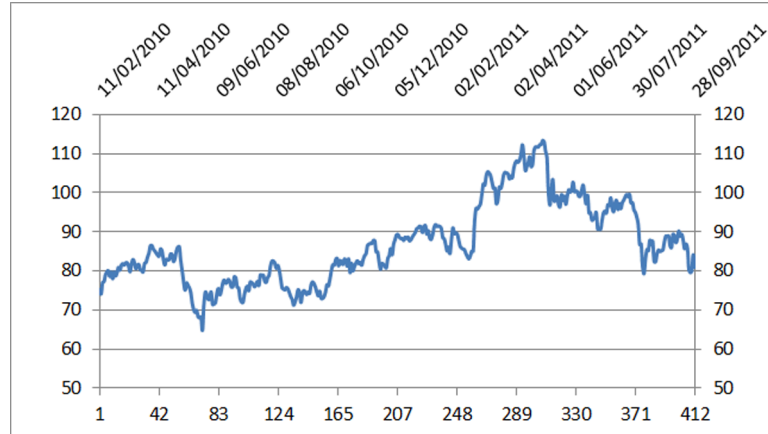
**Figure 3** *Normal future market: WTI future quotations 11 May, 2011, different maturities.*  
On  $x$ -axis are the days to delivery (see online version for colours)



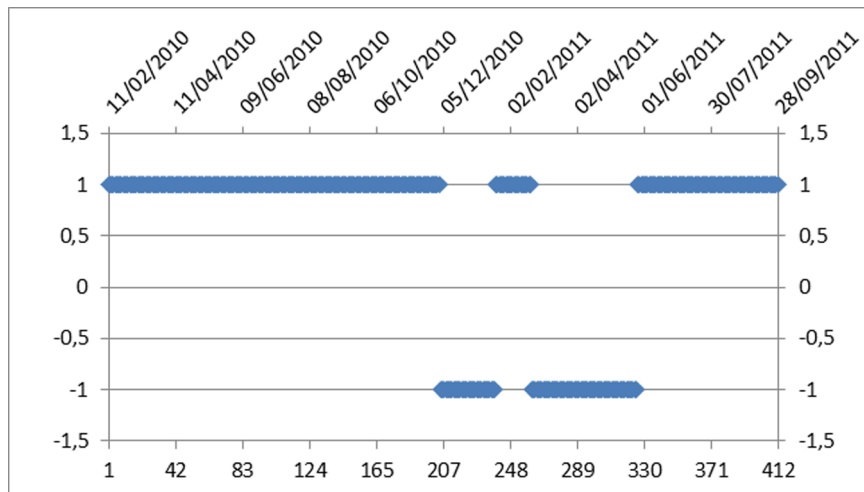
**Figure 4** *Inverted future market: WTI future quotations 25 August, 2011, different maturities.*  
On  $x$ -axis are the days to delivery (see online version for colours)



**Figure 5** Spot FOB data for WTI crude oil for the reduced dataset. The  $x$ -axis label (the lower one) refers to the number of working days from the starting date 11 February, 2010 (see online version for colours)



**Figure 6** The dates in the dataset used for the parameter estimation are divided in normal future market days (+1 value) and inverted future market days (-1 value). The  $x$ -axis label (the lower one) refers to the number of working days from the starting date 11 February, 2010 (see online version for colours)



## 6 Numerical results

To conduct our inference for each of our models discussed in Section 2, we simulated Markov chains of 30,000 steps (from these we discard about 10,000 sampled points as burn-in of the Markov chain).<sup>8</sup> We used the mean of the posterior distribution  $p(\Theta|Z_{0:T})$  as point estimate for each parameter; these results are reported in the Table 1. Together with means we reported also the standard deviation of the posterior distributions.

Since for all the three models the cost of storage ( $c$ ) estimate is greater than long run convenience yield ( $\delta$ ) one, then  $c - \delta_t$  assumes often positive values, which leads to the

prevalence of the normal market effect in futures structure, as it is present for most of the dates in the considered period. From Figures 7 and 8 we can observe that the estimate for  $\delta_t$  is greater than the estimate for  $c$  in the same periods when in the markets are observed an inverted future market, as it is possible to see comparing Figures 2 and 6 with Figure 2.

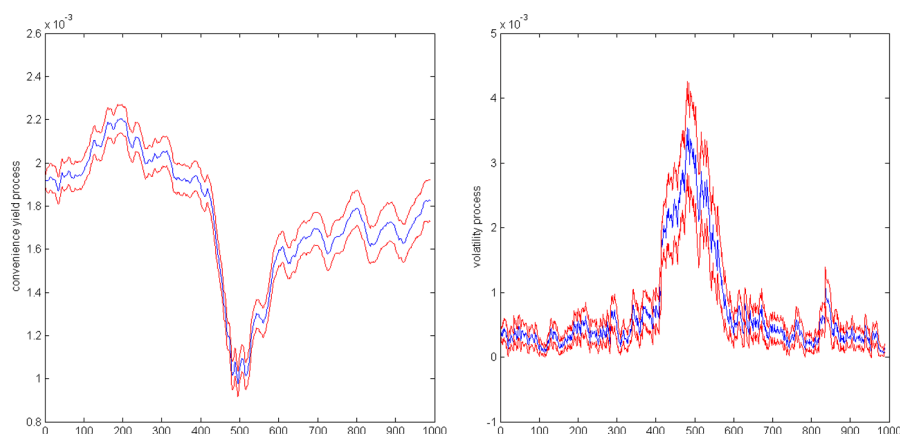
**Table 1** *Parameter inference*: posterior means (and posterior standard deviation) of the model parameter set  $\Theta$  for three analysed models (estimated with the dataset 2007–2011). All values are expressed on a daily basis and are scaled by a 100 factor

<i>Model parameters</i>	<i>Basic model</i>	<i>Model with seasonality</i>	<i>Model with jumps</i>
$\eta_S$	0.0892 (0.0069)	0.0859 (0.0717)	0.1041 (0.0046)
$c$	0.2043 (0.0061)	0.1972 (0.0943)	0.1895 (0.00257)
$\alpha$	0.0942 (0.0415)	0.1242 (0.0651)	0.4633 (0.0652)
$\bar{\delta}$	0.1977 ( $1.10E - 3$ )	0.1965 (0.0058)	0.1868 ( $5.25E - 4$ )
$\sigma$	0.0511 (0.0061)	0.0694 (0.0226)	0.27 (0.0021)
$\eta_\delta$	0.0293 ( $6.95E - 4$ )	0.0256 (0.0035)	0.0357 ( $3.96E - 4$ )
$\beta$	0.5294 (0.2811)	1.4486 (0.8001)	0.7909 (0.5878)
$\bar{V}$	0.0429 ( $7.22E - 4$ )	0.0392 (0.0092)	0.0235 ( $5.61E - 4$ )
$\xi$	0.2616 (0.0666)	0.3405 (0.1159)	0.155 (0.015)
$\rho$	-64.44 (2.0302)	-33.06 (16.60)	-55.98 (0.64)
$\zeta_1$	- (-)	0.6712 (0.2534)	- (-)
$\zeta_2$	- (-)	-0.7275 (0.1607)	- (-)
$\omega_1$	- (-)	2.6048 (2.2499)	- (-)
$\omega_2$	- (-)	3.4644 (4.2490)	- (-)
$\lambda_J$	- (-)	- (-)	5.689 (0.2001)
$\mu_J$	- (-)	- (-)	-0.7241 ( $8.91E - 3$ )
$\sigma_J$	- (-)	- (-)	7.373 (0.1594)

The inference results show for all the models a low level in long-run volatility ( $\bar{V}$ ), for the basic model the estimated value for  $\bar{V}$  is 0.0429 (corresponding to an annualised long-run volatility close to 3.5%), a value that well represent the historical standard deviation we recognise in the return time series analysed, if we exclude the period from June 2008 to

December 2008, in which spot quotation faced a dramatic fall resulting in a sudden increase in volatility level. As expected the model with seasonality show similar values to the basic model for the volatility process parameters, since the seasonality function do not affect the volatility dynamics, just the log-spot one. The main difference among this two models is in the correlation parameter, in model with seasonality we found a lower value for correlation. The posterior for the parameters  $\zeta_1$ ,  $\zeta_2$ ,  $\omega_1$ ,  $\omega_2$  show mean values close to zero, and higher variances with respect to the others parameter posteriors. Beside the basic model and the model including the seasonality term, as discussed in Section 2, we analyse the model including jumps. Jumps could provide a better description for exceptional returns, while ordinary returns can be better explained by the short term volatility dynamics. By comparing the results of the inference carried out within the basic model and within the model with jumps, we can observe similar point estimates for all the parameters involved in the spot and in the convenience yield dynamics, while the differences appears in the parameters for the volatility process, for which we can observe a specifically lower value for the long-run volatility in the case with jumps with respect to the basic model; obviously, this is due to the presence of jumps, which partially explain the volatility of the spot dynamics. The jumps intensity assumes a value significantly different from zero, and the point estimate suggest an average of 5.6% days with jumps.

**Figure 7** Inference on dynamics for the convenience yield and the volatility process got under the model without seasonality or jumps (dataset 2007–2011). On the  $x$ -axis the working day starting from 1 February, 2007 (see online version for colours)



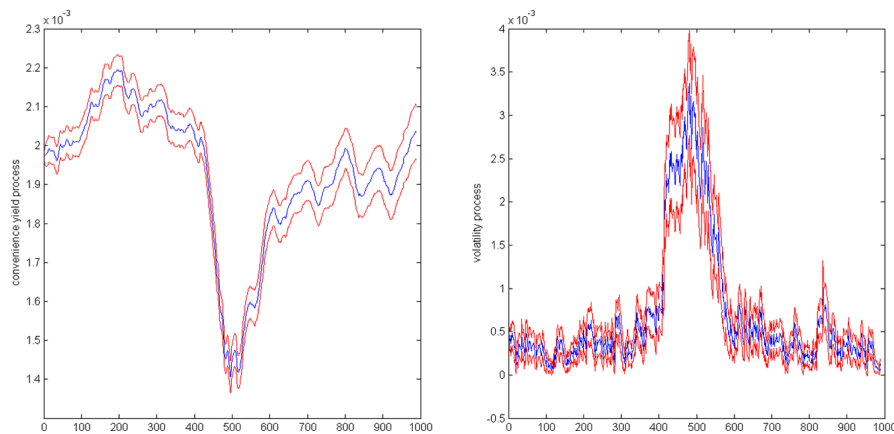
Together with inference about the parameters we got inference on the path of the latent processes (convenience yield and variance process), too. In Figures 7–9) are shown the inference on the latent processes for the different models considered.

The central line represent the mean of all path, that we can interpret as our estimate for the process; the two external lines, instead, identify an interval given by the mean plus (and minus) the standard deviation of the position for all the sampled path. In all the dynamics we can notice how during the period from June 2008 to December 2008, the filtered dynamics behaves like in presence of outliers.

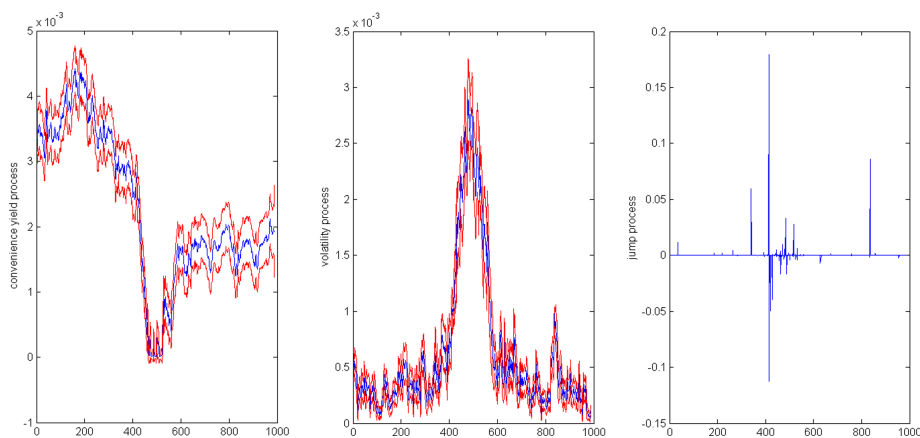
It is interesting to remark the occurrences when the filtered process for the convenience yield is above the ‘threshold’ represented by the estimated cost of storage plus the risk free rate. In this cases, it is more valuable to have the commodity immediately or within a short

maturity than later, since a shortage in storages or an excess of demand (summarised in a positive convenience yield) makes more profitable to own the commodity; so the future prices are generally higher for shorter maturities than longer (inverted future market). If we compare the filtered dynamics we reported above, with the Figure 2 (where are indicated the normal future market day, and the inverted ones), the remark just pointed out can provide quite a good interpretation.

**Figure 8** Inference on dynamics for the convenience yield and the volatility process got under the model with seasonality (dataset 2007–2011). On the  $x$ -axis the working day starting from 1 February, 2007 (see online version for colours)

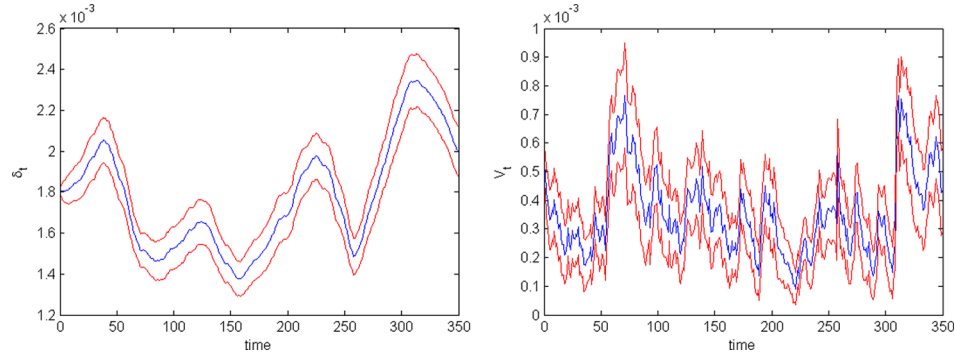


**Figure 9** Inference on dynamics for the convenience yield, the volatility process and the jump process got under the model with jumps (dataset 2007–2011). On the  $x$ -axis the working day starting from 1 February, 2007 (see online version for colours)

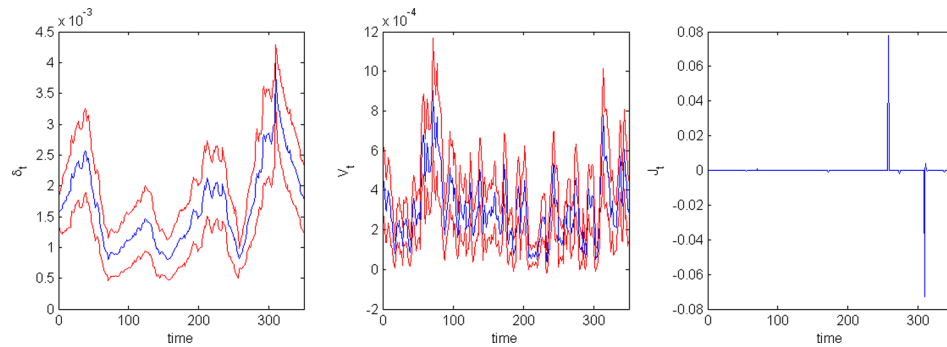


As discussed in the Section 5 we conduct the inference for the basic model and for the model with jumps also for a smaller set of dates, excluding the period with the sudden fall in WTI quotations, see Figures 10 and 11.

**Figure 10** Inference on dynamics for the convenience yield and the volatility process got under the model without seasonality or jumps (dataset 2010–2011). On the  $x$ -axis the working day starting from 11 February, 2010 (see online version for colours)



**Figure 11** Inference on dynamics for the convenience yield and the volatility process got under the model with jumps (dataset 2010–2011). On the  $x$ -axis the working day starting from 11 February, 2010 (see online version for colours)



By comparing inferences reported in Table 2 with the one reported in Table 1 we can notice the slightly lower estimate for the intensity of the jump process, since the new dataset spot dynamic is less volatile than the previous one (see Figures 1 and 5). Comparing the two tables, we can also notice that both the elasticity constants of the mean reversion processes (the one in the convenience yield process,  $\alpha$ , and the one in the volatility process,  $\beta$ ) have changed: higher value for the ‘mean reverting speed’ are preferred since we excluded the period with high volatility from analysed dataset.

## 7 In the sample and out of the sample performances

To compare performances by the three model we analysed both in the sample and out-of the sample results. For in the sample results we, as Yu et al. (2011), analysed the  $\epsilon$  residuals to verify if the assumption of normality is satisfied, if it is then the model well describes the dynamics of data we are studying. In out of the sample analysis we ran a PF algorithm using the parameter set we got from inference and check the square root of the quadratic errors mean (RMSE) and the mean absolute error (MAE) for futures and options on futures.



**Table 2** *Parameter inference*: posterior means (and posterior standard deviation) of the model parameter set  $\Theta$  for the two analysed models (estimated with the reduced dataset 2010–2011). All values are expressed on a daily basis and are scaled by a 100 factor

<i>Model parameters</i>	<i>Basic model</i>	<i>Model with jumps</i>
$\eta_S$	0.1209 (0.0069)	0.2012 (0.0194)
$c$	0.1983 (0.0061)	0.1715 (0.00587)
$\alpha$	1.7618 (0.0415)	3.521 (1.38)
$\bar{\delta}$	0.1950 (1.10E - 3)	0.1606 (0.0066)
$\sigma$	0.0616 (0.0061)	0.4581 (0.0411)
$\eta_\delta$	0.0285 (6.95E - 4)	0.0318 (0.0013)
$\beta$	3.3356 (0.2811)	2.559 (1.08)
$\bar{V}$	0.0461 (7.22E - 4)	0.0283 (0.00194)
$\xi$	0.2646 (0.0666)	0.5267 (0.034)
$\rho$	-66.04 (2.0302)	-54.99 (0.981)
$\lambda_J$	- (-)	4.039 (0.447)
$\mu_J$	- (-)	-1.084 (0.0062)
$\sigma_J$	- (-)	9.363 (0.5387)

To analyse the goodness of fit and to compare the different models, we studied the residuals  $\epsilon_t^{(S)}$ ,  $\epsilon_t^{(V)}$ ,  $\epsilon_t^{(\delta)}$  from equation (5):

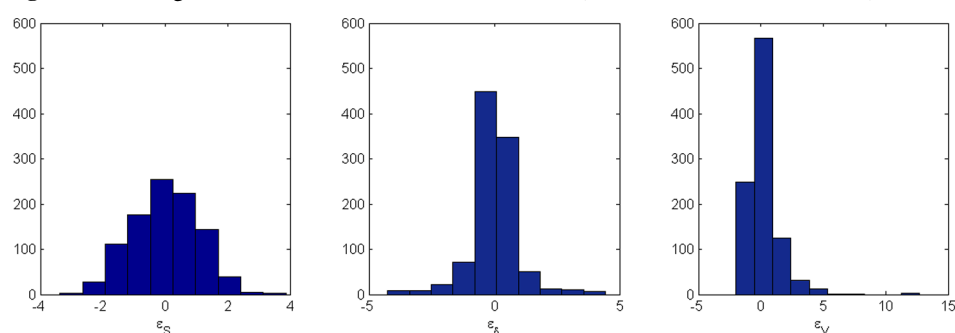
$$\begin{cases} \epsilon_t^{(S)} = \frac{x_{t+1} - x_t - (\mu + c - \delta_t)\Delta t - J_{t+1}}{\sqrt{V_t}\Delta t} \\ \epsilon_t^{(\delta)} = \frac{\delta_{t+1} - \delta_t - \alpha(\bar{\delta} - \delta_t)\Delta t}{\sigma\sqrt{\delta_t}\Delta t} \\ \epsilon_t^{(V)} = \frac{V_{t+1} - V_t - \beta(\bar{V} - V_t)\Delta t}{\xi\sqrt{V_t}\Delta t} \end{cases} \quad (18)$$

Since the model hypothesis,  $\epsilon^{(S)}$ ,  $\epsilon^{(\delta)}$  and  $\epsilon^{(V)}$  are distributed each one according to a standard normal, we use the Kolmogorov-Smirnov test to check this hypothesis<sup>9</sup> and the skewness and kurtosis of the distributions to compare the different models.

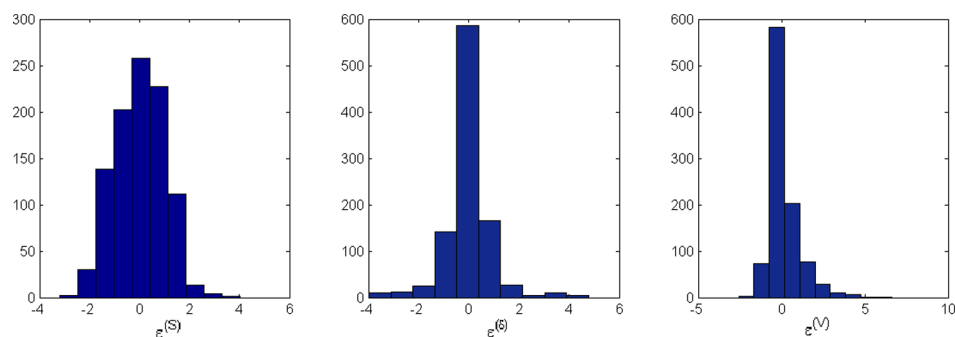
For the risk neutral dynamics, specified by equation (4), we evaluate again both the RMSE and the AME for both the datasets: the dataset used for parameter estimation (in the sample set ITS) and the dataset outside the first one (out of the sample OTS).<sup>10</sup> Results are shown in Table 4.

For the basic model and, as it is shown by the  $p$ -value for the Kolmogorov-Smirnov test (see Table 3), we have to reject the normality hypothesis for at least two of the residuals: the convenience yield and the volatility residuals. The presence of the period with the fall of WTI spot has a great impact in the sampled path, in particular we notice in the histograms (Figures 12–14) of the residuals  $\epsilon_\delta$ ,  $\epsilon_V$  a skewed distribution, and with some values far from the mode values that brought a much more large kurtosis than for a normal distribution. The model with jumps seems to mitigate the effect of leptokurtic distribution, explaining part of the outliers variance value with an increased jump activity. The model with seasonality, even if it provides higher  $p$ -value for the KS test for the spot residual, it does not show any improvement in the performances of the model as far as  $\epsilon_\delta$ ,  $\epsilon_V$  are concerned.

**Figure 12** Histogram for the residuals for the basic model (see online version for colours)



**Figure 13** Histogram for the residuals for the model with seasonality term (see online version for colours)

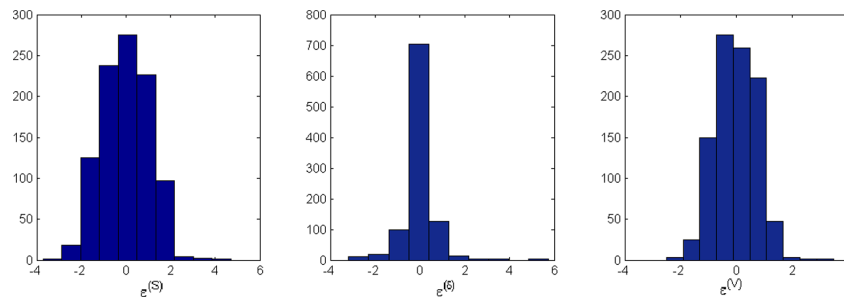


The errors associated with the futures (see Table 4), both in the sample and out of the sample, suggest that we do not get a real improvement by including the seasonality terms; instead, model with jumps shows better performances in out of sample errors analysis, due to the greater flexibility given by using three latent process instead of two in our PF approach.<sup>11</sup>

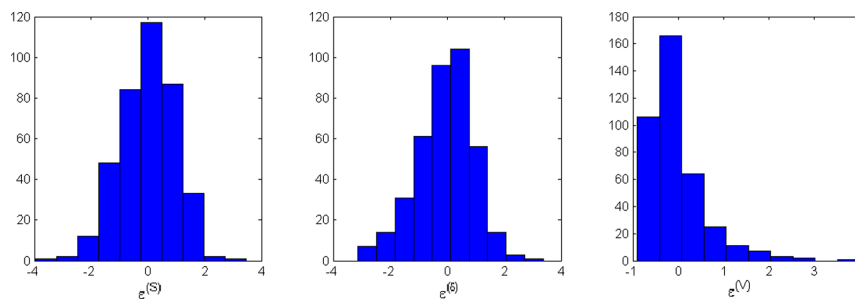
By using the reduced sample set we can confirm the positive impact in performances when jumps process has been included in the model dynamics, see Figures 15 and 16. Performances are in Table 5 and in Table 6. Using this dataset the basic model again does not fit perfectly the data (see Table 3), while the model with jumps in spot well describes

the in the sample performances, revealing itself a good model, succeeding in capturing both the dynamics for WTI quotations (the  $p$ -values for all residuals are sufficient high to avoid rejection of the normality hypothesis, see Table 5)<sup>12</sup> and the futures in the analysed time interval: as we can see from Table 6 the error metrics RMSE and AME for the jump model are far better than those for the basic model, and this happens both for in the sample and out of the sample analysis). This good performance of jump model reflects the ability of the jump process to capture exceptional local events in price dynamics, leading to a better description of the time series, in the cases analysed.

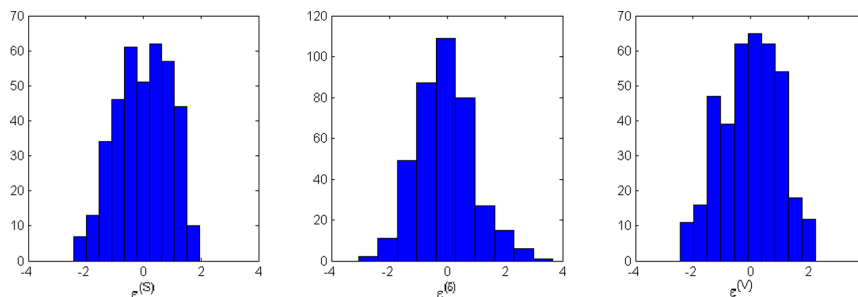
**Figure 14** Histogram for the residuals for the model with jumps (see online version for colours)



**Figure 15** Histogram for the residuals for the basic model using the second dataset (see online version for colours)



**Figure 16** Histogram for the residuals for the model with jumps using the second dataset (see online version for colours)



**Table 3** Analysis of the residuals under the historical measure for the three model, for each residual is reported the p-value of the Kolmogorov-Smirnov test, the skewness and the kurtosis of the distribution

<i>Residual</i>		<i>Basic model</i>	<i>Model with seasonality</i>	<i>Model with jumps</i>
$\epsilon_S$	KS test	0.151	0.47	0.1399
	Skewness	-0.045	-0.048	0.061
	Kurtosis	2.998	2.86	3.152
$\epsilon_\delta$	KS test	$1.6E - 15$	$2.4E - 20$	$1.3E - 24$
	skewness	-0.141	0.33	0.389
	Kurtosis	9.26	9.77	10.29
$\epsilon_V$	KS test	$3.1E - 15$	$3.1E - 23$	$1.8E - 6$
	Skewness	2.95	1.98	0.039
	Kurtosis	12.31	8.924	2.697

**Table 4** In the sample and out of the sample errors for futures data

<i>Futures error</i>		<i>Basic model</i>	<i>Model with seasonality</i>	<i>Model with jumps</i>
ITS	RMSE	1.036	0.956	1.08
	AME	0.525	0.457	0.499
OTS	RMSE	1.033	0.878	0.35
	AME	0.583	0.488	0.21

**Table 5** Analysis of the residuals under the historical measure for model B and model J, for each residual is reported the p-value of the Kolmogorov-Smirnov test, the skewness and the kurtosis of the distribution. We refer to the second dataset

<i>Residual</i>		<i>Basic model</i>	<i>Model with jumps</i>
$\epsilon_S$	KS test	0.436	0.241
	Skewness	-0.319	-0.175
	Kurtosis	3.517	2.181
$\epsilon_\delta$	KS test	0.220	0.1085
	Skewness	-0.311	0.360
	Kurtosis	3.340	3.480
$\epsilon_V$	KS test	$1.7E - 14$	0.303
	Skewness	2.131	-0.174
	Kurtosis	9.701	2.327

**Table 6** second dataset futures performances

<i>Futures error</i>		<i>Basic model</i>	<i>Model with jumps</i>
ITS	RMSE	1.933	0.997
	AME	1.417	0.455
OTS	RMSE	0.735	0.41
	AME	0.574	0.33

## 8 Conclusions

In this paper, we analysed three models, we denoted by (B), (J) and (S), and their performances in capturing the dynamics of West Texas Index spot characteristics and the

future market structure. The aforementioned models belong to the large class of stochastic volatility models, incorporating jumps, stochastic convenience yield and seasonality terms in the price dynamics. Two datasets have been used to get inference, the first one including a period characterised by a sudden fall in spot quote dynamics, while the second one excluding it. We found that model with jumps succeeds in properly describing the second set of observed data. As far as the first set is concerned, none of the three models considered performs very well, although the model with jumps fits the data slightly better than the other two models. Hence we analysed the performances of the basic model (without jumps) and the model with jumps by using a smaller dataset and we found that, while the basic model still exhibits some difficulties in fitting properly the observed data, the model including jumps in the spot dynamics fits very well. We reported also the performances in capturing future market dynamics for a set of data outside the range used to get the inference for the model parameters. Again the model with jumps exhibits the best performances. It is worth remarking that the analysis of our data seems to suggest that the sample correlation between convenience yield process and spot logarithm process is not vanishing. A model including correlation between these two processes will be the subject of future investigations.

## References

- Aihara, S.I., Bagchi, A. and Saha, S. (2008) 'Estimating volatility and model parameters of stochastic volatility models with jumps using particle filter', *Proceedings of the 17th World Congress The International Federation of Automatic Control*, 6–11 July, Seoul, Korea.
- Andrieu, C. and Johansen, A.M. (2008) 'A tutorial on particle filtering and smoothing: fifteen years later', in Crisan, D. and Rozovsky, B. (Eds.): *Handbook of Nonlinear Filtering*, Oxford University Press, Oxford.
- Andrieu, C., Doucet, A. and Holenstein, R. (2010) 'Particle Markov Chain Monte Carlo methods', *Journal of Royal Statistical Society B*, Vol. 72, Part 3, pp.269–342.
- Arulampalam, M.S., Maskell, S., Gordon, N. and Clapp T. (2002) 'A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking', *IEEE Transaction on Signal Processing*, Vol. 50, No. 2, pp.174–188.
- Bakshi, G. and Madan, D. (2000) 'Spanning and derivative-security valuation', *Journal of Financial Economics*, Vol. 55, pp.205–238.
- Bjork, T. (2009) *Arbitrage Theory in Continuous Time*, Oxford University Press, Oxford.
- Boyce, W.E. and Di Prima, R.C. (2012) *Elementary Differential Equations and Boundary Value Problems*, 2nd ed., Wiley, New York.
- Cont, R. and Tankov, P. (2004) *Financial Modelling with Jump Processes*, Chapman Hall/CRC, Boca Raton.
- Cortazar, G. and Schwartz, E.S. (2003) 'Implementing a stochastic model for oil futures prices', *Energy Economics*, Vol. 25, No. 3, pp.215–238.
- Duffie, D., Filipovic, D. and Schachermayer, W. (2003) 'Affine processes and applications in finance', *Ann. Appl. Probab.*, Vol. 13, No. 3, pp.817–1230.
- Eraker, B., Johannes, M.S. and Polson, N. (2003) 'The impact of jumps in volatility and returns', *The Journal of Finance*, Vol. 58, No. 3, pp.1269–1300.
- Eraker, B. (2004) 'Do stock prices and volatility jumps? Reconciling evidence from spot and option prices', *The Journal of Finance*, Vol. 59, No. 3, pp.1367–1403.
- Eydeland, A. and Geman, H. (1998) 'Pricing power derivatives', *RISK*, Vol. 11, pp.71–73.
- Forbes, C.S., Martin, G.M. and Wright, J. (2002) *Bayesian Estimation of a Stochastic Volatility Model Using Option and Spot Prices*, Working Paper, Monash University.

- Geman, H. (2005) *Commodities and Commodity Derivatives*, Wiley Finance, New York.
- Geman, H. (2000) *Scarcity and Price Volatility in Oil Markets*, EDF Trading Technical Report.
- Geweke, J. (1992) 'Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments', in Bernardo, J.M., Berger, J.O., Dawid, A.P. and Smith, A.F.M. (Eds.): *Bayesian Statistics 4*, Oxford University Press, Oxford, pp.169–193.
- Gilks, W.R., Richardson, S. and Spiegelhalter, D. (1996) *Markov Chain Monte Carlo in Practice*, Chapman & Hall/CRC Interdisciplinary Statistics, Boca Raton.
- Glasserman, P. (2004) *Monte Carlo Methods in Financial Engineering*, Springer-Verlag, New York.
- Heston, S. (1993) 'A closed-form solution for options with stochastic volatility with applications to bond and currency options', *Review of Financial Studies*, Vol. 6, No. 2, pp.327–343.
- Hikspoors, S. (2008) *Multi-Factor Energy Price Models and Exotic Derivatives Pricing*, University of Toronto PhD Thesis, Advisor: Prof. S. Jaimungal, available at S.Jaimungal's webpage [http://www.utstat.utoronto.ca/sjaimung/papers/Hikspoors\\_Samuel\\_200811\\_PhD\\_thesis.pdf](http://www.utstat.utoronto.ca/sjaimung/papers/Hikspoors_Samuel_200811_PhD_thesis.pdf)
- Jacquier, E., Polson, N.G. and Rossi, P.E. (1994) 'Bayesian analysis of stochastic volatility models', *Journal of Business and Economic Statistics*, Vol. 12 No. 4, pp.371–389.
- Jacquier, E., Polson, N.G. and Rossi, P.E. (2004) 'Bayesian analysis of stochastic volatility models with fat-tails and correlated errors', *Journal of Econometrics*, Vol. 122, pp.185–212.
- Javaheri, A. (2005) *Inside Volatility Arbitrage*, Wiley Finance, New York.
- Johannes, M.S., Polson, N. and Stroud, J.R. (2009) 'Optimal filtering of jump diffusions: extracting Latent states from asset prices', *Review of Financial Studies*, Vol. 22, pp.2559–2599.
- Pan, J. (2002) 'The jump-risk premia implicit in options: evidence from an integrated time series study', *Journal of Financial Economics*, Vol. 63, pp.3–50.
- Peters, G.W., Briers, M., Shevchenko, P. and Doucet, A. (2013) 'Calibration and filtering for multi factor commodity models with seasonality: incorporating panel data from futures contracts', *Methodology and Computing in Applied Probability*, Vol. 15, No. 4, pp.841–874.
- Pitt, M.K. and Shephard, N. (1999) 'Filtering via simulation: auxiliary particle filters', *Journal of the American Statistical Association*, Vol. 94, No. 446, pp.590–599.
- Pitt, M.K., Silva, R., Giordani, P. and Kohn, R. (2012) 'On some properties of Markov chain Monte Carlo simulation methods based on the particle filter', *Journal of Econometrics*, Vol. 171, No. 2, pp.134–151.
- Ribeiro, D. and Hodges, S.D. (2004) *A Two-Factor Model for Commodity Prices and Futures Valuation*, EFMA 2004 Basel Meetings Paper, <http://ssrn.com/abstract=498802> or doi:10.2139/ssrn.498802.
- Schwartz, E.S. (1997) 'The stochastic behavior of commodity prices: implications for valuation and hedging', *Journal of Finance*, Vol. 52, pp.923–973.
- Shreve, S. (2003) *Stochastic Calculus for Finance II: Continuous Time Models*, Springer Finance, New York.
- Smith, B.J. (2007) 'boa: an R package for MCMC output convergence assessment and posterior inference', *Journal of Statistical Software*, Vol. 21, No. 11, pp.1–37.
- Yan, X.S. (2002) 'Valuation of commodity derivatives in a new multi-factor model', *Review of Derivatives Research*, Vol. 5, pp.251–271.
- Yu, C., Li, H. and Wells, M.T. (2011) 'MCMC estimation of Lévy jump models using stock and option prices', *Mathematical Finance*, Vol. 21, No. 3, pp.383–422.

## Notes

- <sup>1</sup>The usual definition of the convenience yield, as discussed also in Schwartz (1997), is obtained adding the cost of storage to the net convenience yield:  $\tilde{\delta}_t = \delta_t + c$ .
- <sup>2</sup>For the basic model we just fix to zero  $\mu^*$  and  $dJ_{x_t}^{(Q)}$ , for the model with seasonality we add the seasonality factor ( $g(t_{\text{year}})$ ) to the spot dynamics like in equation (2).
- <sup>3</sup> $\sigma_\epsilon^2 = 0,08$ , for the dataset under investigation, is consistent with the value observed for the mean of the daily spread of the highest-lowest prices for futures in the considered range.
- <sup>4</sup>Note that in the Futures price obtained there is no dependence on the jump process parameters in strict analogy with Yan (2002).
- <sup>5</sup>In the case of the model with seasonality we replace the parameters referring to jumps ( $\{\lambda, \mu_J, \sigma_J\}$ ) with the parameters referring to seasonality:  $\{\zeta_1, \zeta_2, \omega_1, \omega_2\}$ .
- <sup>6</sup>Details about the generic PF algorithm and probability distributions used in the simulation we conducted are in Appendix A.
- <sup>7</sup>The sequential importance resampling (SIR) algorithm is one of the algorithm of the PF family. A full and deep description is in Andrieu and Johansen (2008).
- <sup>8</sup>The lack of convergence has been tested by Geweke test, and at 1% significance level we cannot reject the hypothesis of convergence of the Markov Chain. For any further details about the test, the estimation error and the algorithm used to compute them we refer to the manual written by Smith (2007).
- <sup>9</sup>We can reject the normality hypothesis for  $p$ -values lower than a given threshold. Usual  $p$ -value threshold are 1%, 5%, 10%.
- <sup>10</sup>It is worth to remark that in the analysis out of the sample just the parameter set is fixed from previous inference. The inference on the path is got from the sampling importance resampling (SIR) algorithm. Hence it is not surprising the performances out of the sample could be not worse than in the sample performances.
- <sup>11</sup>This improvement has a cost in required time to run the single PF algorithm: spanning observed data coming from 998 dates and using 2000 ‘particles’, for the basic model we need 6.45 s, for the model including the seasonality term the time is equal (the number of latent process simulated is the same), the model including jumps requires 9.03 s. All the algorithms written in Matlab have been running on a i3(3,07 GHz) pc.
- <sup>12</sup>Beside  $p$ -values for the Kolmogorov-Smirnov test, all above 10%, we remark that also the kurtosis and skewness values are consistent with the normality hypothesis.

## Appendix A: Particle filter

We implemented the ASIR algorithm proposed by Pitt and Shephard (1999) to generalise the standard sampling important resampling algorithms, attempting to mitigate the sample impoverishment effect, sampling the ancestor of the particle at each time step, adapting the proposed ‘particle’ sample to the new incoming information. The starting point to define the algorithm is the discretised dynamics of the model (5) adding the future formula (8). Let us consider the time  $t$  a sample for the latent process  $\{\delta_{t-1}, V_{t-1}\}$ , that is:  $\{\delta_{t-1}^{(i)}, V_{t-1}^{(i)}\}$ , with associated weights  $\pi_{t-1}^{(i)}$ , where  $i = 1, \dots, N_s$  and  $N_s$  is the number of particles. Before propagating the particles, that is before obtaining a sample for the latent processes at time  $t$ , we select the ancestor for this new sample. In practice, we define an expected value for the

latent process sample at time  $t$ :  $\{\bar{\delta}_t^{(i)}, \bar{V}_t^{(i)}\}$ , and we use this estimate to attach a weight function  $\omega_{t|t-1}^{(i)}$  to the particles and select the best fitting for the new incoming information:

$$\begin{cases} \bar{\delta}_t^{(i)} = \alpha(\bar{\delta} - \delta_{t-1}^{(i)})\Delta t \\ \bar{V}_t^{(i)} = \beta(\bar{V} - V_{t-1}^{(i)})\Delta t \\ \omega_{t|t-1}^{(i)} = \sum_k \phi\left(y_t | m_{x_t}^{(i)} + k\mu_j, \Sigma_{x_t}^{(i)} + k\sigma_j^2\right) \frac{\lambda_j^k e^{-\lambda_j}}{k!} \pi_t^{(i)} \\ m_{x_t}^{(i)} = (r + c - \bar{\delta}_t^{(i)} - \frac{1}{2}\bar{V}_t^{(i)})\Delta t \\ \Sigma_{x_t}^{(i)} = \bar{V}_t^{(i)} \Delta t \end{cases} \quad (19)$$

where  $m_{x_t}^{(i)}$  is the mean of the observed process log spot excluded jumps  $x$  at time  $t$ , and  $\Sigma_{x_t}^{(i)}$  his variance, while  $\mu_j$  is the mean of the jump amplitude and  $\sigma_j^2$  his variance.

Hence we get a  $N_s$  long sample  $\{\delta_{t-1}^{(s)}, V_{t-1}^{(s)}\}$  from the distribution  $\{\delta_{t-1}^{(i)}, V_{t-1}^{(i)}\}$  associated to probability weights  $\omega_{t|t-1}^{(i)}/\sum_i \omega_{t|t-1}^{(i)}$ . This are our selected ancestors, we will propagate using equation (5). First we simulate the jumps. The number of jumps  $N_t^j$  happening in the time passing from  $t-1$  to  $t$  is distributed according to a Poisson with parameter  $\lambda_j$ :

$$N_t^{(s)} = k \text{ with probability } q(N_t^{(s)}) = \phi\left(x_t | m_{x_t}^{(s)} + k\mu_j, \Sigma_{x_t}^{(s)} + k\sigma_j^2\right) \frac{\lambda_j^k e^{-\lambda_j}}{k!}$$

For numerical purposes we need to restrict to a finite number of possible jumps, since the time interval consists in one day, we can fix the maximum jumps to 1 per day.

Hence, the jump amplitude is simulated from a normal, as suggested in Johannes et al. (2009) we incorporate the new return information in our sampling:

$$J_t^{(s)} \sim q(J_t | x_t) = \mathcal{N}\left(N_t^{(s)}\mu_j + N_t^j \frac{\sigma_j^2}{\bar{V}_t^{(s)}\Delta t} \left(x_t - (r + c + \left(\eta_S - \frac{1}{2}\right)\bar{V}_t - \bar{\delta}_t^{(s)})\right), N_t^{(s)}\sigma_j^2 \left(1 - N_t^{(s)} \frac{\sigma_j^2}{\bar{V}_t^{(s)}\Delta t}\right)\right)$$

Once we have a sample for the jumps, we sample the value of the ‘particles’ for the variance and the convenience yield:

$$\begin{cases} \delta_t^{(s)} \sim p(\delta_t^{(s)} | \delta_{t-1}^{(s)}) = \mathcal{N}\left(\bar{\delta}_t^{(s)}, \sigma^2 \delta_{t-1}^{(s)} \Delta t\right) \\ V_t^{(s)} \sim q(V_t^{(s)} | V_{t-1}^{(s)}, \delta_{t-1}^{(s)}, x_t, J_t) \\ = \mathcal{N}\left(\bar{V}_t^{(s)} + (x_t - (r + c - \bar{\delta}_t^{(s)} - \frac{1}{2}\bar{V}_t^{(s)})\Delta t - J_t(s)_t) \right. \\ \left. \times \xi \sqrt{1 - \rho^2}, \rho^2 \xi^2 V_{t-1}^{(s)} \Delta t\right) \end{cases} \quad (20)$$



Hence, by Monte Carlo importance sampling theory (see Glasserman, 2004), we compute the new weights  $\pi_t^{(s)}$ :

$$\left\{ \begin{array}{l} \pi_t^{(s)} = \frac{\lambda_j^{N_t^{(s)}} e^{-\lambda_j}}{N_t^{(s)}!} \\ \frac{\omega_{t|t-1}^{(s)} q(N_t^{(s)}|x_t)}{\phi(J_t^{(s)}|\mu_j, \sigma_j^2) \phi(V_t^{(s)}, \bar{V}_t^{(s)}, \xi^2 V_{t-1}^{(s)} \Delta t)} \\ \times \frac{q(J_t^{(s)}|x_t) q(V_t^{(s)}|V_{t-1}^{(s)}, \delta_{t-1}^{(s)}, x_t, J_t)}{q(x_t|(r+c-\delta_t^{(s)}-\frac{1}{2}V_{t-1}^{(s)})\Delta t+J_t^{(s)})} \\ + \rho \sqrt{V_{t-1}^{(s)} \Delta t} \epsilon_t^{(V,s)}, (1-\rho^2)V_{t-1}^{(s)} \Delta t \end{array} \right. \quad (21)$$

where  $\epsilon_t^{(V,s)}$  is the random value got for  $\epsilon_t^{(V)}$  and corresponding to the  $s$ -th particle. An unbiased (Pitt et al., 2012) estimate for the marginal likelihood of the observation given the parameter set is:

$$\left\{ \begin{array}{l} \hat{p}(x_{1:T}, F_{1:T}) = \hat{p}(x_1, F_1) \prod_{t=2}^T \hat{p}(x_t, F_t | y_{t-1}, F_{t-1}) \\ \hat{p}(x_t, F_t | y_{t-1}, F_{t-1}) = \left( \sum_{s=1}^{N_s} \frac{\pi_t^{(s)}}{N_s} \right) \left( \sum_{s=1}^{N_s} \omega_{t|t-1}^{(s)} \right) \end{array} \right. \quad (22)$$

where  $F_{1:T}$  indicates the vector of the latent processes  $(\delta_{1:T}, V_{1:T}, J_{1:T})$ .

At this stage we can add an other re-sample step from the particles with weights  $\pi_t^{(s)}$ , conditional on a measure of the impoverishment of the sample, if we resample we have to set all the weights equal to  $1/N_s$ .

To obtain the algorithm for the model without jumps it is sufficient to set all  $N_t^{(s)} = 0$  for all  $t$  in the algorithm presented in this section.

## Appendix B: Futures options prices

To evaluate, under the  $\mathcal{Q}$  measure, a general contingent claim  $H$ , whose dynamics follows equation (4), with the seasonality function fixed at zero  $g(t_{\text{year}}) = 0$ , we need to solve the PDE:

$$\begin{aligned} & \frac{\partial H}{\partial t} + \frac{\partial H}{\partial x} \left( r_f + c - \lambda \mu^* - \delta_t - \frac{1}{2} V_t \right) \\ & + \frac{\partial H}{\partial V} [\beta(\bar{V} - V_t) - \eta_V V_t] + \frac{\partial H}{\partial \delta} [\alpha(\bar{\delta} - \delta_t) - \eta_\delta \delta_t] \\ & + \frac{1}{2} \frac{\partial^2 H}{\partial x^2} V_t + \frac{1}{2} \frac{\partial^2 H}{\partial V^2} \xi^2 V_t + \frac{1}{2} \frac{\partial^2 H}{\partial \delta^2} \sigma^2 \delta_t + \frac{\partial^2 H}{\partial V \partial x} \rho \xi V \\ & - r_f H + \lambda [H(t, x_t + \ln(1+J), \delta_t, V_t) - H(t, x_t, \delta_t, V_t)] = 0, \end{aligned}$$

equipped with a terminal condition specified by the contingent claim payoff. In particular, in this section we are interested in the valuation of a European call option  $C(t, \tau)$  written on a futures contract maturing at time  $\tau' \geq \tau$ , where  $\tau$  is the maturity of the option and

$K$  the strike price. Following Heston (1993), Bakshi and Madan (2000) and Yan (2002), we can decompose the value  $C(t, \tau)$  in the following way:

$$C(t, \tau, \tau') = G(t, \tau)\Pi_1(t, \tau) - KB(t, \tau)\Pi_2(t, \tau), \quad (23)$$

where  $B(t, \tau)$ ,  $G(t, \tau)$ ,  $\Pi_1(t, \tau)$  and  $\Pi_2(t, \tau)$  are defined as follows:

$$B(t, \tau) = E_Q \left[ \exp \left( - \int_t^{t+\tau} r(u) du \right) \right] \quad (24)$$

$$G(t, \tau) = E_Q \left[ \exp \left( - \int_t^{t+\tau} r(u) du \right) H(t, \tau') \right] \quad (25)$$

$$\Pi_1(t, \tau) = \frac{1}{G(t, \tau)} E_Q \left[ \exp \left( - \int_t^{t+\tau} r(u) du \right) H(t, \tau') | H(t, \tau') > K \right] \quad (26)$$

$$\Pi_2(t, \tau) = \frac{1}{B(t, \tau)} E_Q \left[ \exp \left( - \int_t^{t+\tau} r(u) du \right) | H(t, \tau') > K \right]. \quad (27)$$

Here  $B(t, \tau)$  is the price at time  $t$  of a pure (unit) discount bond maturing at time  $\tau$  and  $G(t, \tau)$  is the price at time  $t$  of a forward contract delivering at time  $t + \tau$  the contingent claim  $H(t + \tau, \tau' - \tau)$ . By defining the following discounted function of the logarithm of the futures price:

$$f(t, \tau, w) = E_Q \left[ \exp \left( - \int_t^{t+\tau} r(u) du \right) \exp(iw \ln H(t, \tau')) \right], \quad (28)$$

Bakshi and Madan (2000) proved moreover that:

$$B(t, \tau) = f(t, \tau; 0) \quad (29)$$

$$G(t, \tau) = f \left( t, \tau; \frac{1}{i} \right) \quad (30)$$

$$f_1(t, \tau; w) = \frac{1}{iG(t, \tau)} f \left( t, \tau; w + \frac{1}{i} \right) \quad (31)$$

$$f_2(t, \tau; w) = \frac{1}{B(t, \tau)} f(t, \tau; w) \quad (32)$$

where  $f_1(t, \tau; w)$  and  $f_2(t, \tau; w)$  are respectively the characteristic functions of the two probability densities  $\Pi_1(t, \tau)$  and  $\Pi_2(t, \tau)$ . Hence each term can be recovered once computed  $f(t, \tau; w)$ , that is the characteristic function of the logarithm of the future price probability density. Since  $f(t, \tau; w)$  is the price of a contingent claim paying  $\exp\{iwH(t + \tau, \tau' - \tau)\}$  at maturity  $t + \tau$ , its value can be found solving again the same PDE (equation (23)) with the terminal condition:

$$f(t + \tau; 0) = \exp\{i\phi H(t + \tau, \tau' - \tau)\} \quad (33)$$

If we substitute the trial solution:

$$f(t, \tau; w) = \exp\{\theta_0(\tau) + \theta_\delta(\tau)\delta_t + \theta_V(\tau)V_t + iw[x + \beta_0(t) + \beta_\delta(t)\delta_t]\} \quad (34)$$

in our PDE, we get the following system of ordinary differential equations:

$$\begin{cases} -\frac{\partial\phi_0(\tau)}{\partial\tau} + iw(r_f + c - \mu^*) + \phi_\delta\alpha\bar{\delta} + \theta_V\beta\bar{V} = 0 \\ -\frac{\partial\phi_\delta(\tau)}{\partial\tau} + iw - \phi_\delta(\alpha - \eta_\delta) + \frac{1}{2}\phi_\delta^2(\tau)\sigma^2 = 0 \\ -\frac{\partial\theta_V(\tau)}{\partial\tau} - \frac{1}{2}iw - \frac{1}{2}w^2 - \theta_V(\beta - \eta_V) + \frac{1}{2}\xi^2\theta_V^2 + iw\theta_V\rho\xi = 0 \end{cases} \quad (35)$$

where

$$\mu^* = \exp\left\{\mu_J + \frac{1}{2}\sigma_J^2\right\} - 1 \quad (36)$$

$$\phi_\delta(\tau) = \theta_\delta(\tau) + i\phi\beta_\delta(\tau' - \tau) \quad (37)$$

$$\phi_0(\tau) = \theta_0(\tau) + i\phi\beta_0(\tau' - \tau) \quad (38)$$

equipped with the terminal conditions:

$$\begin{cases} \phi_0(0) = i\phi\beta_0(\tau') \\ \phi_\delta(0) = i\phi\beta_\delta(\tau') \\ \theta_V(0) = 0 \end{cases} \quad (39)$$

Solving the ODE system we get:

$$\theta_V = -\frac{2}{\xi^2} \left( \frac{\exp\{C_V\tau\} - 1}{\frac{\exp\{C_V\tau\}}{A_V} - \frac{1}{B_V}} \right) \quad (40)$$

with

$$\begin{cases} C_V = \sqrt{(\eta_V + iw\rho\xi - \beta)^2 + \xi^2(w^2 + iw)} \\ A_V = \frac{1}{2}[(\beta - \eta_V - iw\rho\xi) + C_V] \\ B_V = \frac{1}{2}[(\beta - \eta_V - iw\rho\xi) - C_V] \end{cases} \quad (41)$$

and

$$\phi_\delta = -\frac{2}{\sigma^2} \frac{A_\delta \exp\{C_\delta\tau\} + B_\delta\bar{C}}{\exp\{C_\delta\tau\} + \bar{C}}, \quad (42)$$

where

$$\left\{ \begin{array}{l} C_\delta = \sqrt{(\alpha - \eta_\delta)^2 + 2iw\sigma^2} \\ A_\delta = \frac{1}{2}[(\eta_\delta - \alpha) + C_\delta] \\ B_\delta = \frac{1}{2}[(\eta_\delta - \alpha) - C_\delta] \\ \tilde{C}_\delta = \sqrt{(\alpha - \eta_\delta)^2 + 2\sigma^2} \\ \tilde{A}_\delta = \frac{1}{2}[(\eta_\delta - \alpha) + \tilde{C}_\delta] \\ \tilde{B}_\delta = \frac{1}{2}[(\eta_\delta - \alpha) - \tilde{C}_\delta] \\ \bar{C} = \frac{i\phi \frac{\exp\{\tilde{C}_\delta \tau'\} - 1}{\exp\{\tilde{C}_\delta \tau'\} - \frac{1}{\tilde{B}_\delta}} - A_\delta}{B_\delta - iw \frac{\exp\{\tilde{C}_\delta \tilde{\tau}\} - 1}{\exp\{\tilde{C}_\delta \tilde{\tau}\} - \frac{1}{\tilde{B}_\delta}}} \end{array} \right. \quad (43)$$

Finally we solve for  $\phi_0$ :

$$\begin{aligned} \phi_0 = & iw(r_f + c - \mu^*)\tau - \frac{2\beta\bar{V}}{\xi^2} \left( (B_V + C_V)\tau + \log \frac{A_V \exp\{-C_V \tau\} - B_V}{C_V} \right) \\ & - \frac{2\alpha\bar{\delta}}{\sigma^2} \left( (B_d + C_d)\tau + \log \frac{1 + \exp\{-C_d \tau\} \tilde{C}}{1 + \tilde{C}} \right) \\ & + iw \left( (r_f + c)\tau' - \frac{2\alpha\bar{\delta}}{\sigma^2} \left( (\tilde{B}_\delta + \tilde{C}_\delta)\tau' \right. \right. \\ & \left. \left. + \log \frac{\tilde{A}_\delta \exp\{-\tilde{C}_\delta \tau'\} - \tilde{B}_\delta}{\tilde{A}_\delta - \tilde{B}_\delta} \right) \right) \\ & + \lambda_J \left( \exp \left\{ iw\mu_J - \frac{1}{2}w^2\sigma_J^2 \right\} - 1 \right). \end{aligned} \quad (44)$$