HISTORICAL AND RISK-NEUTRAL ESTIMATION IN COMMODITY MARKET WITH A TWO FACTOR STOCHASTIC VOLATILITY MODEL.

GAETANO FILECCIA AND CARLO SGARRA

Abstract. In this paper we analyze spot prices and futures quotation data to get inference under both the historical and the risk neutral measure in commodity crude oil market (data are referred to WTI index which tracks the crude oil barrel price on NYMEX market). While big part of research and techniques in finance deals with the risk neutral modeling or with the model choice under the historical measure, the purpose of this work is to study the estimation problem under both measures at the same time, through a suitable parametric choice of the Radon-Nikodym derivative. To conduct this estimation we resort to a recent technique in Bayesian inference: the Particle Markov Chain Monte Carlo proposed by Andrieu, Doucet and Holenstein [7], in which particle filters algorithms are used to estimate the marginal likelihood for Markov Chain Monte Carlo inference. We adopt a stochastic volatility two factor model to describe the dynamics of the spot price, for which the futures prices and future options prices are available in closed form. Three versions of the original model, with and without jumps in prices and with seasonality term, are taken into account and results are compared.

1. Introduction

Stochastic volatility models are a well-known choice in commodity markets, Geman [18] and Hikspoors [16] 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 [17], where the commodity spot values are modeled by a mean reverting process and the convenience yield is incorporated in the discount factor. While Cortazar and Schwartz [21] 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 [20] 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 pointed out [19], 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 modeling choice.
Further developments were implemented, as in Ribeiro and Hodges [23], to

1991 *Mathematics Subject Classification.* (commodity markets - financial econometrics).
Key words and phrases. PMCMC, bayesian, commodities, stochastic volatility, jumps.

describe the dynamics for the oil sector of commodity market. Moreover in this paper [23], it is 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 process for the convenience yield would automatically exclude arbitrage opportunities in the market. In the model proposed in this article, we stand on this remark by Ribeiro and Hodges, modeling 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 (modeled 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 data set used for inference in the model without jumps. A third version of the model substitutes the jump term with a seasonality term (modeled by the usual periodic function, as suggested also in [16]), allowing for a comparison between the effect of jumps and of the seasonality term in the goodness of fit of analyzed data set. Data set used for the estimation come from the WTI index spot values, quoted in NYMEX market, and the futures written on it, more details about data sets are given in a dedicated section. The time window considered spans from 01/02/2007 to 31/12/2010.

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, Polson and Rossi [25] a fast-growing literature is exploring the application of Bayesian estimation techniques, in particular 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 [1] and by Eraker, Johannes and Polson [2] in which some popular stochastic volatility models are analyzed 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, Martin and Wright [3] and, more recently, by Yu, Li and Wells [4], who extended the results by Eraker [1] including different jump models in the analysis and compare them; all these paper refer to equity markets, analyzing 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 particle filters are the papers by Andrieu and Johansen [9], and Arulampalam et al. [10], while for application to stochastic volatility models Javaheri [11], Johannes, Polson an Stroud [6], Aihara [28]. Since the success of PF techniques, different authors have worked on efficient estimation with Particle filters, and conjunction of MCMC and PF algorithms; this culminated in a paper by Andrieu, Doucet and Holestein [7], where have been introduced the particle Markov Chain Monte Carlo algorithms that implement for inference on parameter set of a model an MCMC where the marginal likelihood is estimated by a nested Auxiliary particle filter. We have adopted an algorithm belonging to this family to carry out inference on the model chosen.

2. The Models Proposed

We studied three possible different variants the basic model including a volatility process and a convenience yield process into spot price dynamics. Both the spot variance and the convenience yield processes follow a CIR dynamics. Under the historical measure $\mathbb{P}$ the dynamics is the following:

$$\begin{cases} \dfrac{dS_t}{S_t} & = (\mu + c - \delta_t)dt + \sqrt{V_t}dW_{S_t}^{(\mathbb{P})} \\ d\delta_t & = \alpha(\bar{\delta} - \delta_t)dt + \sigma\sqrt{\delta_t}dW_{\delta_t}^{(\mathbb{P})} \\ dV_t & = \beta(\bar{V} - V_t)dt + \xi\sqrt{V_t}dW_{V_t}^{(\mathbb{P})} \\ dW_{S_t}^{(\mathbb{P})} & dW_{V_t}^{(\mathbb{P})} = \rho dt \end{cases} \tag{2.1}$$

Besides the basic model we considered the same model allowing for jumps in the spot dynamics (modeled with poisson distributed jump time arrivals and normal jump size), and the basic model with a seasonality term in spot dynamics. The only dynamics changing from one model to the other is the spot equation. In the model with jumps the dynamics equation becomes:

$$\begin{cases} \dfrac{dS_t}{S_t} & = (\mu + c - \delta_t)dt + \sqrt{V_t}dW_{S_t}^{(\mathbb{P})} + dJ_{S_t}^{(\mathbb{P})} \\ d\delta_t & = \alpha(\bar{\delta} - \delta_t)dt + \sigma\sqrt{\delta_t}dW_{\delta_t}^{(\mathbb{P})} \\ dV_t & = \beta(\bar{V} - V_t)dt + \xi\sqrt{V_t}dW_{V_t}^{(\mathbb{P})} \\ dW_{S_t}^{(\mathbb{P})} & dW_{V_t}^{(\mathbb{P})} = \rho dt \end{cases} \tag{2.2}$$

In the case with seasonality we add the usual sinusoidal term (like discussed also in [16], hence the equation become:

$$\begin{cases} \dfrac{dS_t}{S_t} & = g(t_{\text{year}}) + (\mu + c - \delta_t)dt + \sqrt{V_t}dW_{S_t}^{(\mathbb{P})} \\ g(t_y) & = \exp\{\zeta_1(sen(2\pi t_{\text{year}} + \omega_1)) + \zeta_2(cos(2\pi t_{\text{year}} + \omega_2))\} \end{cases} \tag{2.3}$$

with t_{year} the number of days from the first of January of the same year divided by 365. The dynamics of the convenience yield process and the variance process does not change from previous models.

Since seasonality term is just a deterministic function of time, changing from a probability measure to another one does not affect it. Hence, following, we discuss the model that allow spot prices to jump (2.2) (since also the risk-neutral dynamics for the basic model can be derived easily: just neglecting the jump terms), and we refer to this model In order to describe the dynamics under the risk-neutral measure, 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 [12] and by Pan [13], 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:

$$\begin{cases} dW_{\delta_t}^{(\mathbb{Q})} &= dW_{\delta_t}^{(\mathbb{P})} - \frac{\eta_\delta}{\sigma}\sqrt{\delta_t}dt \\ dW_{V_t}^{(\mathbb{Q})} &= dW_{V_t}^{(\mathbb{P})} - \frac{1}{\sqrt{1-\rho^2}}\left(\rho\eta_{S_t} + \frac{\eta_V}{\xi}\right)\sqrt{V_t}dt \\ dW_{S_t}^{(\mathbb{Q})} &= dW_{S_t}^{(\mathbb{P})} + \eta_{S_t}\sqrt{V_t}dt \end{cases} \tag{2.4}$$

Hence in (2.2) $\mu = r + \eta_S\sqrt{V_t} + \mu_J^*$ where μ_J^* is the compensator of the jump process.
According to the specified choice, the jump structure remains the same under the two measures; only the drift term in the spot dynamics will be affected by the measure change, in order to assure that the discounted price process is a martingale under the risk-neutral measure $\mathbb{Q}$.
Hence, under the risk-neutral measure $\mathbb{Q}$:

$$\begin{cases} \frac{dS_t}{S_t} &= (r + c - \delta_t - \mu^*)dt + \sqrt{V_t}dW_{S_t}^{(\mathbb{Q})} + dJ_{S_t}^{(\mathbb{Q})} \\ d\delta_t &= \left(\alpha(\bar{\delta} - \delta_t) + \eta_\delta\delta_t\right)dt + \sigma\sqrt{\delta_t}dW_{\delta_t}^{(\mathbb{Q})} \\ dV_t &= \left(\beta(\bar{V} - V_t) + \eta_V V_t\right)dt + \xi\sqrt{V_t}dW_{V_t}^{(\mathbb{Q})} \\ dW_{S_t}^{(\mathbb{Q})} & dW_{V_t}^{(\mathbb{Q})} = \rho dt \end{cases} \tag{2.5}$$

[1]

We make the usual choice of describing the price dynamics through the log-price process: $x_t = \log[S_t]$. It is worth remarking that, since the cost of storage is separated off the dynamics of the convenience yield, δ_t this is modeled 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 [23].
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 use a Euler discretization method. The discretized model (under the risk neutral measure) can be written as follows:

(2.6)

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

[1] for the basic model we just ignore μ^* and $dJ_{S_t}^{(\mathbb{Q})}$, for the model with seasonality we add the seasonality factor ($g(t_{\text{year}})$) to the spot dynamics like in (2.3)

Where each $\epsilon_t^{(S)}$, $\epsilon_t^{(V)}$ and $\epsilon_t^{(\delta)}$ are normally distributed random variables with zero mean and variance Δt. In the discretized jump addend, the $N_t^{(J)}$ are independent Poisson distributed r.v. with parameter λ_J, while $\epsilon_{i,t}^{(J)}$ are independent normal r.v. with mean $(\mu_J - \eta_J)$ and variance σ_J^2. All the random variables just listed are independent on each other.
Under the historical measure the same discretization holds, provided that the risk premium $\{\eta_\delta,\ \eta_V,\ \eta_J\}$ and the compensator μ^* are set to zero, while the drift coefficient $r_f + c$ has to be substituted by μ.
The futures price is given by the following expression:

$$F(0,\tau) = \exp\{A_0(\tau) + x_t + A_2(\tau)\delta_t\} \tag{2.7}$$

where $\tau = T - t$ is the futures time to maturity. Details about computations and the specifics of the functions $A_0(\tau)$ and $A_2(\tau)$ are provided in the appendix. In case we consider model allowing for seasonality, the future slightly change, and it becomes:

$$F(0,\tau) = g(T_{\text{year}}) \exp\{A_0(\tau) + x_t + A_2(\tau)\delta_t\} \tag{2.8}$$

where T_{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 modeled the price of the futures making the hypothesis that the market price, $F^{\text{M}}(0,\tau)$, are represented by the theoretical price got by (2.7) plus an error distributed as a white noise $\epsilon_{\text{fut}} \sim \mathcal{N}(0,\sigma_\epsilon^2)$[2]:

$$F^{\text{M}}(0,\tau) = F(0,\tau) + \epsilon_{\text{fut}} \tag{2.9}$$

3. Futures Prices

When the underlying follows, under the the dynamics $\mathbb{Q}$ (2.5) the Kolmogorov backward equation for the generic contract value, $f(t, x_t, \delta_t, V_t, J_t)$, at time t, when the underlying follows, under the $\mathcal{Q}$ measure, the dynamics (2.5) is:

$$\begin{aligned}
&\frac{\partial f}{\partial t} + \partial f \partial x (r_f + c - \lambda \mu_J^* - \delta_t - \frac{1}{2} V_t) + \frac{1}{2}\frac{\partial^2 f}{\partial x^2} V_t + \\
&+ \frac{\partial f}{\partial \delta}\left[\alpha(\bar{\delta} - \delta_t) - \eta_\delta \delta_t\right] + \frac{1}{2}\frac{\partial^2 f}{\partial \delta^2}\sigma^2 \delta_t + \\
&+ \frac{\partial f}{\partial V}\left[\beta(\bar{V} - V_t) - \eta_V V_t\right] + \frac{1}{2}\frac{\partial^2 f}{\partial V^2}\xi^2 V_t + \\
&+ \frac{\partial^2 f}{\partial V \partial x}\rho \xi V + \\
&+ \lambda\left[f(t, x_t + ln(1+J), \delta_t, V_t) - f(t, x_t, \delta_t, V_t)\right] = 0
\end{aligned} \tag{3.1}$$

The PDE (3.1) has to be solved with the usual terminal condition $f(t = T) = H(x_T, \delta_T, V_T)$ with H payoff at time T maturity, and since we are dealing with

[2] σ_ϵ^2 is fixed at value: half of the mean highest-lowest (future price) range observed daily

futures the final payoff is:

$$H(x_T, \delta_T, V_T) = \exp\{x_T\} \tag{3.2}$$

We make the hypothesis of a solution form:

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

if we try this guess solution into (3.1), we get the follow ODE system:

(3.3)
$$\begin{cases} -\dfrac{\partial A_0(\tau)}{\partial \tau} + A_1(\tau)(r_f + c) + A_2(\tau)\alpha\bar{\delta} + A_3\beta\bar{V} = 0 \\ -\dfrac{\partial A_1(\tau)}{\partial \tau} = 0 \\ -\dfrac{\partial A_2(\tau)}{\partial \tau} - A_1(\tau) - A_2(\tau) + \frac{1}{2}A_2^2(\tau)\sigma^2 = 0 \\ -\dfrac{\partial A_3(\tau)}{\partial \tau} - \frac{1}{2}A_1(\tau) + \frac{1}{2}A_1^2(\tau) - A_3(\tau)(\beta + \eta_V) + \frac{1}{2}A_3^2(\tau)\xi^2 + A_1(\tau)A_3(\tau)\rho\xi = 0 \end{cases}$$

where we changed variable from t to $\tau = T - t$ [3] .
Since the payoff is (3.2) we obtain the initial condition for the ODE system:

$$\begin{cases} A_0(0) &= 0 \\ A_1(0) &= 1 \\ A_2(0) &= 0 \\ A_3(0) &= 0 \end{cases} \tag{3.4}$$

Solving previous system we get :

$$\begin{cases} A_0(\tau) &= (r_f + c)\tau - \dfrac{2\alpha\bar{\delta}}{\sigma^2}\left[B\tau + \log\left(\dfrac{D - B\exp\{C\tau\}}{D - B}\right)\right] \\ A_1(\tau) &= 1 \\ A_2(\tau) &= -\frac{2}{\sigma^2}\dfrac{\exp\{C\tau\} - 1}{\frac{\exp\{C\tau\}}{D} - \frac{1}{B}} \\ A_3(\tau) &= 0 \end{cases} \tag{3.5}$$

with:

$$\begin{cases} C &= \sqrt{(\alpha + \eta_\delta)^2 + 2\sigma^2} \\ D &= \dfrac{(\eta_\delta - \alpha) + C}{2} \\ B &= \dfrac{(\eta_\delta - \alpha) - C}{2} \end{cases} \tag{3.6}$$

[3]Note also that in the system above there is no dependence on the jump process parameters, since the compensator μ_J is by definition: $\mu_J = [f(t, x_t + ln(1 + J), \delta_t, V_t) - f(t, x_t, \delta_t, V_t)]$.
Hence there is no difference in futures pricing formula between model including jump process and model excluding it.

4. Inference Algorithm

The reference 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 Θ is the set of parameters, that is our main inference target: $\Theta = \{\epsilon, \mu, c, \alpha, \bar{\delta}, \sigma, \eta_\delta, \beta, \bar{V}, \xi, \eta_V, \rho, \lambda, \mu_J, \sigma_J, \eta_J\}$[4]. 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 make inference in a so large state space, Particle Markov Chain Monte Carlo methods (from now on PMMC) represents an efficient technique, since it allows to simulate the latent processes in a single simulation block. 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 make use of a particle filter algorithm[5] to estimate the marginal likelihood $L(y|\Theta)$, this will be used in defining an acceptance ratio probability that will ensure that after a certain amount of time, whatever point in state space we started from, we will sample from the right distribution, getting unbiased estimate for Θ.
The MCMC method used is a variant of the well known Metropolis-Hastings algorithm

- At step (s=0)
  after setting a starting point $\Theta = \theta(0)$ arbitrarily, then a Sampling Important Resampling algorithm is implemented. SIR 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, θ^*, is sampled from a proposal (symmetric) distribution $q(\cdot|\theta(s-1))$ and the new sampling value is accepted with a probability:

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

  where $\tilde{p}(\cdot)$ is the prior distribution, while $p(z_{1:T}|\theta^*)$ is the likelihood for the observation chain given the parameter set θ^*.
  Again the marginal likelihood and $x_{1:T}^*$ comes form 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:

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

[4]in the case of the model with seasonality we replace the parameters referring to jumps($\{\lambda, \mu_J, \sigma_J, \eta_J\}$) with the parameters referring to seasonality: $\{\zeta_1,\ \zeta_2,\ \omega_1,\ \omega_2\}$

[5]Details about the generic Particle Filter algorithm and probability distributions used in the simulation we conducted are in appendix A.

We can also discard the information about hidden state process and use the sample to get inference about θ, 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 PMMH (particle Metropolis Hastings algorithm), 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 [**?**].

5. The Data

The data set used for the analysis are relative to WTI Cushing Crude Oil spot and futures quotations on NYMEX market from 1/02/2007 to 31/12/2010. Spot data are collected from the US Energy information administration website 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 parameter set of the different models analyzed is in figure Fig.1

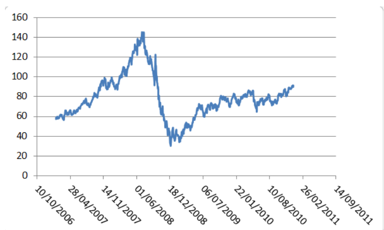

Figure 1. Spot FOB data for WTI crude oil

Besides 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 analyzed a spot datum and 12 futures data, and in the range there are 988 dates. Hence the data set consists in 988 spot values and 11856 future contracts.
In addition to this data set we reserved a panel of data to evaluate out of the sample performances. The data set 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 01/01/2011 to 28/09/2012. The number of working days in the range is 187.
Future contracts are usually characterized 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 Fig.2 are indicate 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).

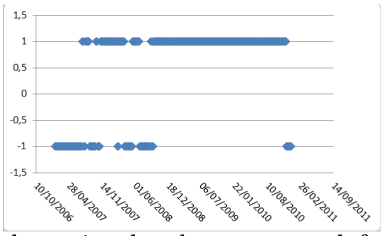

Figure 2. the dates in the data set used for the parameter estimation are divided in normal future market days (+1 value) and inverted future market days (−1 value)

Both in the analyzed period and in the data set used to evaluate out of the sample performances show both the future market behaviour. The two behaviour are shown in the figures Fig.3 and Fig.4 and illustrate, respectively, the future market structure at the date 11/05/2011 and at 25/08/2011.

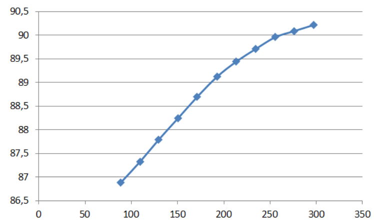

Figure 3. Normal future market: WTI future quotations 11 May 2011, different maturities. On x-axis are the days to delivery

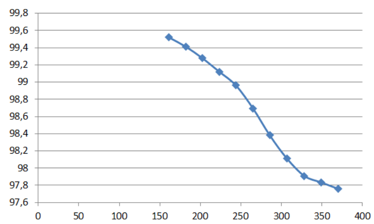

Figure 4. Inverted future market: WTI future quotations 25 August 2011, different maturities. On x-axis are the days to delivery

6. Numerical Results

The analysis has been conducted with model (2.6), with the same model excluding jumps (that is imposing $\lambda_J = 0$) and with the same configuration of the latter model, including the seasonality factor $g(t_{\text{year}})$, as in (2.3).

Discarded the first part of each chain (10000 iterations for both the models), we elaborated the simulation to get inference on parameter set. The results are summarized in table 1.

TABLE 1. Parameter inference: posterior means (and posterior standard deviation) of the model parameter set Θ for three analyzed models. All values are scaled by a 100 factor

Model Parameters	Basic Model	Model with Seasonality	Model with Jumps
μ	0.0843 (0.0016)	0.0757 (0.0717)	()
c	0.2215 (0.0023)	0.1985 (0.0943)	()
α	0.0258 (0.0196)	0.1047 (0.0651)	()
$\bar{\delta}$	0.1973 $(7.48E-4)$	0.1915 (0.0058)	()
σ	0.0304 (0.0033)	0.0593 (0.0226)	()
η_δ	0.0282 $(2.45E-4)$	0.0271 (0.0035)	()
β	1.5165 (0.2376)	1.3730 (0.8001)	()
$\bar{V}$	0.0438 $(9.02E-4)$	0.0359 (0.0092)	()
ξ	0.3394 (0.0048)	0.0359 (0.1159)	()
ρ	−50.63 (2.1394)	−31.88 (16.60)	()
ζ_1	− (−)	0.8679 (0.2534)	()
ζ_2	− (−)	−0.9197 (0.1607)	()
ω_1	− (−)	4.3136 (2.2499)	()
ω_2	− (−)	4.4705 (4.2490)	()
λ_J	− (−)	− (−)	()
μ_J	− (−)	− (−)	()
σ_J	− (−)	− (−)	()

Together with inference about the parameters we got also inference on the path of the latent processes (convenience yield and variance process). In the following figures (Fig.5, Fig.6 and Fig.**??**) are shown the inference on the latent processes for the different models, the blue line represent the mean of all path, that we can interpret as our estimate for the process, in red the lines representing the estimate plus (and minus) the standard deviation of the position of all the sampled path: we are interpreting the standard deviation as the error associate with our estimation.

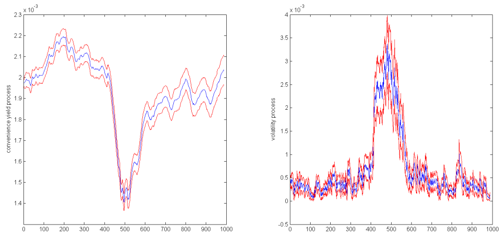

FIGURE 5. inference on dynamics for the convenience yield and the volatility process got under the model without seasonality or jumps

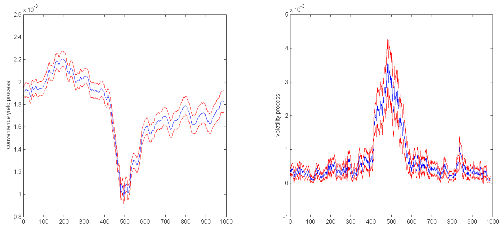

FIGURE 6. inference on dynamics for the convenience yield and the volatility process got under the model with seasonality

In the table 1 are summarized the results for posterior distributions $p(\Theta|Z_{0:T})$ for the three models.
Since cost of storage (c) estimate is greater than long run convenience yield ($\bar{\delta}$) one, then $c - \delta_t$ assumes often positive values, which reflects in a normal market effect in futures structure, as it is present for most of the dates in the considered period. From Fig.5 and Fig.6 we can observe that the estimate for δ_t is greater than the estimate for c in the same perdiods when in the markets are observed an inverted future market, as it is possible to see from Fig.2. The sharp increase in volatility and decrease in convenience yield estimate corresponds to the fall in spot valuation in the period $01/08/2008 - 31/12/2008$

7. In the sample and out of the sample performances

To compare performances by the two model we analyzed both in the sample and out-of the sample results. for in the sample results we, as Yu Li Wells [4], analyzed the ϵ residuals to verify if the assumption of normality is satisfied , hence the model well describe the dynamics of data we are studying. Out-of the sample we ran a particle filter algorithm using the parameter set we got from inference and check the RMSE and MAE for futures and option on futures, to value which perform better perform for risk-management purposes.

To analyze the goodness of fit and to compare the different models, we studied the residuals $\epsilon_t^{(S)}, \epsilon_t^{(V)},\ \epsilon_t^{(\delta)}$ from (2.6):

$$
\begin{cases}
\epsilon_t^{(S)} & = \dfrac{x_{t+1} - x_t - (\mu + c - \delta_t)\Delta t}{\sqrt{V_t \Delta t}} \\
\epsilon_t^{(V)} & = \dfrac{V_{t+1} - V_t - \beta(\bar{V} - V_t)\Delta t}{\sqrt{V_t \Delta t}} \\
\epsilon_t^{(\delta)} & = \dfrac{\delta_{t+1} - \delta_t - \alpha(\bar{\delta} - \delta_t)\Delta t}{\sigma\sqrt{\delta_t \Delta t}}
\end{cases}
\tag{7.1}
$$

The model hypothesis is they are distributed according to a standard normal, we use the Kolmogorov-Smirnov test to check this hypothesis, and when it is not satisfied we compute the skewness and kurtosis of the distribution to compare the different models. For the risk neutral dynamics we evaluate the square root of the mean of quadratic errors (RMSE) and the absolute mean error (AME)for both the data sets: the data set used for parameter estimation (in the sample set ITS) and the data set outside the first one (out of the sample OTS). Results are shown in Table 7.

Table 2. 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

Residual		Basic Model	Model with Seasonality	Model with Jumps
	KS test	0.346	0.502	
ϵ_S	skewness	-0.053	-0.045	
	kurtosis	2.674	2.732	
	KS test	$7.6E-71$	$2.4E-20$	
ϵ_δ	skewness	-0.907	-1.043	
	kurtosis	7.240	9.529	
	KS test	$2.5E-12$	$4.1E-14$	
ϵ_V	skewness	0.216	0.266	
	kurtosis	11.42	8.424	

Futures error		Basic Model	Model with Seasonality	Model with Jumps
ITS	RMSE	3.21	2.86	
	AME	2.33	2.03	
OTS	RMSE			
	AME			

8. Conclusions

References

[1] Eraker, B.: *Do stock pices and volatility jumps? Reconciling evidence from spot and option prices*, The Journal of Finance (June 2004) Vol.LIX No. 3.

[2] Eraker, B., Johannes, M. S. and Polson, N.: *The impact of Jumps in Volatility and Returns*, The Journal of Finance (June 2003) Vol.LVIII No. 3.

[3] Forbes, C.S., Martin, G.M. and Wright, J.: *Bayesian Estimation of a Stochastic Volatility Model Using Option and Spot Prices*, Working paper (February 2002), Monash University, 2002.

[4] Yu, C., Li, H. and Wells, M. T.: *MCMC Estimation of Lévy Jump Models using Stock and Option prices* Mathematical Finance, Vol. 21, No. 3 (July 2011), 383422

[5] Pitt, M.K., and Shephard, N.: *Filtering via Simulation: Auxiliary Particle Filters*, Journal of the American Statistical Association, Vol. 94, No. 446 (June 1999), pp. 590-599

[6] Johannes, M.S., Polson, N. and Stroud, J.R.: *Optimal Filtering of Jump Diffusions: Extracting Latent States from Asset Prices*, Review of Financial Studies (2009), 22, 2559-2599.

[7] Andrieu, C., Doucet, A. and Holenstein, R.: *Particle Markov Chain Monte Carlo methods*, Journal of Royal Statistical Society B (2010) 72, Part 3, pp.269-342.

[8] Pitt, M.K., Silva, R., Giordani, P. and Kohn, R.: *Auxiliary Particle filtering within adaptive Metropolis-Hastings Sampling*, working paper (May 2010), available at http://arxiv.org/abs/1006.1914

[9] Andrieu, C. and Johansen, A.M.: *A Tutorial on Particle Filtering and Smoothing: Fifteen years later*, Handbook of Nonlinear Filtering, (eds. D. Crisan et B. Rozovsky), Oxford University Press (2008).

[10] Arulampalam, M.S., Maskell, S., Gordon, N. and Clapp T.: *A Tutorial on Particle Filters for Online Nonlinear/Non-Gaussian Bayesian Tracking*, IEEE Transaction on signal processing (February 2002), VOL. 50, NO. 2.

[11] Javaheri, A.: *Inside volatility arbitrage* Wiley Finance (2005)

[12] Heston, S.: *A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options* Review of financial studies (1993), Vol. 6 (2), pp 327-343.

[13] Pan, J.: *The Jump-Risk Premia Implicit in Options: Evidence From an Integrated Time Series Study* Journal of Financial Economics, Vol.63, pp 350.

[14] Bakshi, G. and Madan, D.: *Spanning and derivative-security valuation* Journal of Financial Economics (2000), Vol.55, pp 205238.

[15] Carr, P. and Madan, D.: *Option valuation using the fast Fourier transform* Journal of Computational Finance (1999), Vol.55, pp 205238.

[16] Hikspoors, S.: *Multi-Factor Energy Price Models and Exotic Derivatives Pricing*, University of Toronto Ph.D. Thesis (advisor: Prof. S.Jaimungal) May 2008, available at S.Jaimungal's webpage (http://www.utstat.utoronto.ca/sjaimung/papers/Hikspoors_Samuel_200811_PhD_thesis.pdf).

[17] Schwartz, E.S.: *The stochastic behavior of commodity prices: Implications for valuation and hedging* Journal of Finance (1997) 52, Part 3, pp. 923-973.

[18] Geman, H.: *commodities and commodity derivatives* Wiley Finance (2005)

[19] Geman, H.: *scarsity and price volatility in oil markets* EDF Trading Tchnical Report (2000)

[20] Eydeland, A. and Geman, H.: *Pricing Power Derivatives* RISK (October 1998), Vol. 11, pp 71-73.

[21] Cortazar, G. and Schwartz, E.S.: *Implementing a Stochastic Model for Oil Futures Prices* Energy Economics (May 2003) Volume 25, Issue 3, pp.215-238.

[22] Yan, X.S. : *Valuation of Commodity Derivatives in a New Multi-Factor Model*, Review of Derivatives Research (2002) Volume 5, pp.251-271.

[23] Ribeiro, D. and Hodges, S.D: *A Two-Factor Model for Commodity Prices and Futures Valuation*, EFMA 2004 Basel Meetings Paper (January 2004). Available at SSRN: http://ssrn.com/abstract=498802ordoi:10.2139/ssrn.498802

[24] Pindyck, R.S.: *Volatility in natural gas and oil markets* The Journal of Energy and Development Volume 30, No. 1.

[25] Jacquier, E., Polson, N.G. and Rossi, P.E.: *Bayesian Analysis of Stochastic Volatility Models* Journal of Business & Economic Statistics (1994) Vol.12, Issue 4, pp.371-89

[26] Jacquier, E., Polson, N.G. and Rossi, P.E.: *Bayesian Analysis of Stochastic Volatility Models with fat-tails and correlated errors* Journal of Econometrics (2004) 122 , pp.185212
[27] Chernov, M. and Ghysels, E.: *A study towards a unified approach to the joint estimation of objective and risk neutral measures for the purpose of options valuation* Journal of Financial Economics (2000) 56 , pp.407-458
[28] Aihara, S.I., Bagchi, A. and Saha, S.: *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 Seoul, Korea, July 6-11, 2008
[29] Smith, B.J.: *boa: An R Package for MCMC Output Convergence Assessment and Posterior Inference* Journal of Statistical Software (November 2007), Volume 21, Issue 11. Available at : `http://www.jstatsoft.org/v21/i11/paper`
[30] Geweke, J.: *Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments* Bayesian Statistics 4, J. M. Bernardo, J. O. Berger, A. P. Dawid, and A. F. M. Smith (eds.), pp. 169-193. Oxford University Press.

Appendix A. Particle Filter

The particle filter the algorithm implemented is the Sampling Important Resampling (SIR) algorithm. It was introduced to overcome degeneracy typical of standard bootstrap Particle Filter, which is a technique to sample from a distribution $p(X_{1:T}|Z_{1:T})$ by sequential application of Importance Sampling Monte Carlo sampler. Each simulated process is identified as "particle".
With respect to the basic bootstrap method in SIR algorithm a resample step is added. Roughly speaking, with re-sampling we will focus our sample on the most probable particles, avoiding that the most of particles lose their "importance weight" moving on an improbable forward direction.
Describing our discretized model with the following probabilities:

$$\begin{cases} x_k \sim & p(x_k|x_{k-1}) \\ z_k \sim & p(z_k|x_k) \end{cases} \tag{A.1}$$

The SIR algorithm is (for the generic time step $t = k$[6]):

(1) for $j = 1, \dots M$ sample $x_k^{(j)}$ from $q(x_k|x_{0:k-1}^{(j)}, z_{0:k})$
(2) $x_{0:k}^{(j)} = (x_{0:k-1}^{(j)}, x_k^{(j)})$
(3) compute the importance weight for all the particles (for $j = 1, \dots M$)

$$\omega_k^{(j)} = \omega_{k-1}^{(j)} \frac{p(z_k|x_k^{(j)})p(x_k^{(j)}|x_{k-1}^{(j)})}{q(x_k|x_{0:k-1}^{(j)}, z_{0:k})}$$

(4) normalize the importance weights:

$$w_k^{(i)} = \frac{\omega_k^{(i)}}{\sum_{j=1}^M \omega_k^{(j)}}$$

(5) for $j = 1, \dots M$ sample $x_{0:k}^{(j)}$ from the empirical distribution $\displaystyle\sum_{i=1}^M w_k^{(i)} \delta_{x_{0:k}^{(i)}}$

[6]The time step t goes from 1 to T, hence the algorithm has to be repeated for all the time step

(6) since all the particles are, now, equally probable, set $w_k^{(j)} = \omega_k^{(j)} = \frac{1}{M}$ for all the j
(7) go to the next time step $k+1$
(8) at time T the likelihood

$$p(z_{1:T}|\theta) = p(z_1|\theta)\prod_{t=2}^{T} p(z_t|z_{1:t-1}, \theta)$$

can be estimated by the weights ω:

$$\hat{p}(z_t|z_{1:t-1}, \theta) = \sum_{j=1}^{M} \omega_t^{(j)}$$

the probability function come from the Euler-discretized version of the model in (2.6). As in [28], we integrated out the jump process, instead of simulating it to spare computation time and improve algorithm efficiency, hence the simulated latent process is $X_{0:T} = \{V_{0:T}, \delta_{0:T}\}$.
The probability distribution used in simulation are:

$$q(x_k|x_{0:k-1}^{(j)}, z_{0:k}) = q_1(\delta_k|\delta_{k-1})q_2(Vk|Vk-1, x_k) \tag{A.2}$$

with

$$\begin{cases} q_1(\delta_k|\delta_{k-1}) \sim & \mathcal{N}\left(\delta_{k-1} + [\alpha(\bar{\delta} - \delta_{k-1}) + \eta_\delta \delta_{k-1}]\Delta t; \sigma^2 \delta_{k-1}\right) \\ q_2(V_k|V_{k-1}) \sim & \mathcal{N}\left(V_{k-1} + [\beta(\bar{V} - V_{k-1}) + \eta_V V_{k-1}]\Delta t\right) \\ p(z_k|V_k, \delta_k) \sim & \sum_{j=0}^{+\infty} \frac{e^{-\lambda\Delta t}(\lambda \Delta t)^j}{j!} \mathcal{N}\left(z_{k-1} + (r_f + c - \delta_k - \mu^*)\Delta t + \frac{\rho}{\xi} M_V + j\mu_J; (1-\rho^2)V_k + j\sigma_J^2\right) \\ M_V = & V_k - V_{k-1} - [\beta(\bar{V} - V_{k-1}) - \eta_V V_{k-1}]\Delta t \end{cases} \tag{A.3}$$