Non-Gaussian Mark to Future for Energy Forwards and Futures

Marcos Escobar, Nicolás Hernández and Luis Seco

The development of risk-management methodologies for non-Gaussian markets relies often on the assumption that underlying market factors have a Gaussian distribution. While advances have been made in modelling more general marginal distributions of risk factors, modelling non-Gaussian dependence structures is much less advanced. For commodities markets that often exhibit sudden changes from backwardation into contango (such as energy and metals), correlations or their linear transformations fail to account for realistic transformations of forward-curve convexity. This paper develops a model that overcomes this difficulty, and presents it in the context of other developments for commodity forward modelling.

The development of risk-management methodologies for non-Gaussian markets relies often on the assumption that the underlying market factors have a Gaussian distribution. While advances have been made in modelling more general marginal distributions of the risk factors, modelling non-Gaussian dependence structures is much less advanced. Normal rank correlations (NRCs) and Spearman ratios yield better results than correlations in markets with non-Gaussian marginals, but are still linear measures of dependence. For commodities markets that often exhibit sudden changes from backwardation into contango (such as energy and metals), NRC correlations and Spearman ratios fail to account for realistic transformations of forward-curve convexity. Copulas that, in principle, will reconstruct the entire dependence structure, appear to be too difficult to implement in specific situations.

This paper develops a model that overcomes this difficulty in certain markets, and presents it in the context of related developments for commodity forward (futures) modelling. It is important to point out that the focus of this paper is to provide a model to fit commodity markets for the purpose of risk management. At a critical point, we will have to introduce arbitrage opportunities into our model to improve the fit to market statistics. While this is of no serious consequence for scenario-generation purposes in the context of risk-management exercises, and might very well be consistent with liquidity constraints observed in the market under consideration (oil, in this case), our model must be used with certain limitations when used in the context of pricing contracts.

There are three popular views of dealing with commodities futures prices, namely:

- by modelling the vector of futures prices of dimension $m$ (12 in the oil case) that can be obtained at any $t$-time from the market, which is the simplest way by filtering all kind of dependencies through time in the
components of the vector, and by taking care of the multivariate structure of the residuals by NRCs or Spearman ratios.

- The theory of storage (TS) of Kaldor (1939), Brennan (1958) and Telser (1958) explains the difference between contemporaneous spot and futures prices in terms of interest forgone in storing a commodity, warehousing costs, and a convenience yield on inventory.

- An alternative view splits a futures price into an expected risk premium and a forecast of a future spot price. See, for example, Cootner (1960), Breeden (1980) and Hazuka (1984).

While the first two ways are not controversial, there is little agreement on whether futures prices contain expected premiums or can forecast spot prices. In this article, we present an alternative to the previous methods, but, first, we basically review TS and the features of the futures prices vectors.

There is one general way to obtain expression for the futures prices of commodities under TS, using no-arbitrage arguments:

\[ F_{t,T} = S_t e^{(r(t,T) + u(t,T) - \delta(t,T))(T-t)} \]

where,

- \( F_{t,T} \) denotes the forward price, at time \( t \) of a contract for delivery at time \( T \).

- \( S_t \) is the spot price of the commodity at time \( t \).

- \( r(t,T) \) is the interest rate at time \( t \) for \( (T-t) \) time to maturity, bootstrapped from the zero curve.

- \( \delta(t,T) \) is the convenience yield: the flow of services that accrues to the holder of the physical commodity, but not to the owner of a contract for future delivery. In other words, the benefit from ownership of the physical commodity that may include the ability to profit from temporary local storages, or the ability to keep a production process running at \( t \)-time for \( (T-\tau) \)-time to maturity.

- \( u(t,T) \) is the storage cost at time \( t \) for \( (T-\tau) \)-time to maturity.

In the case of a constant interest rate for any maturity time, we will get the forward price. The instantaneous interest rate and convenience yield are defined as:

\[ r_t = \lim_{T \to t} r(t, T) \]
\[ \delta_t = \lim_{T \to t} \delta(t, T). \]

respectively.

The term structure of a forward interest rate and futures convenience yields are defined as:

\[ rf(t, s) = \lim_{T \to s} r(t, s, T) \]
\[ \delta f(t, s) = \lim_{T \to s} \delta(t, s, T), \]

where \( r(t, T) \) is the interest rate yield at time \( t \) for the period from \( s \) to \( T \).

Where \( \delta(t, T) \) is the convenience yield at time \( t \) for the period from \( s \) to \( T \), then:

\[ r(t, T) = \int_t^T r_s ds \]
\[ \delta(t, T) = \int_t^T \delta_s ds. \]

As can be seen, futures prices are functions of several underlying variables. Those variables have a well-known meaning from a financial point of view and are widely accepted. That is why a natural way to model futures prices is by dealing with models of those underlying elements. In particular, when trying to forecast a
futures price, one needs only to work on the models for the underlying variables. The main difficulty with these variables is that most of them cannot be recorded or observed, leading to problems with reliability.

As for the univariate and multivariate features of commodities (oil, gas, sugar, coffee, metals) futures prices, we have the following: mean reversion and stochastic volatility, which can be modelled by an AR(1)-GARCH(1,1) model, with non-normal residuals. The vector of residuals still has the contango and backwardation feature (increasing or decreasing curve through time to maturity).

The second part of this paper is a survey of some of the most important models in the literature, highlighting some of their drawbacks (e.g., Ross 1995, Schwartz and Smith 2000, Gibson and Schwartz 1990, Schwartz 1997, Cortazar and Schwartz 1994, Schwartz and Miltersen 1998, and Urich 2000). The third part deals with the proposed model and some details about the estimation procedure. The fourth section is dedicated to the forecasting properties of the model.
using a test portfolio and measures of the risk implemented in RiskWatch and HistoRisk.

Before starting with the descriptions of some models, it is necessary to point out the following practical facts:

- Futures contract are traded on a day-by-day basis, and there are widely known conventions about how they should be traded; this is not the case of the forward contract, which is only found in over-the-counter markets.

- One of the most influential features of the futures markets is that one can only get at time t, the futures prices, for a commodity maturing at Ti – t for fixed values of Ti, that is, Ti are on the twentieth day of every month in the case of oil.

Models

The Ross model

The Ross model posits:

\[ dX_t = k \left( \mu - \frac{\sigma^2}{2k} - X_t \right) dt + \sigma dz. \]  

(1)

Using an Ornstein-Uhlenbeck (OU) process, where \( X_t \) denotes the log of the time-t current spot price \( (S_t) \), let \( F_{T_t} \) denote the time-t market price for a futures contract with time \( T \) until maturity. This yields:

\[ \ln(F_{T_t}) = \ln(S_t) + \xi_t + \xi_T - \lambda \xi \]  

(2)

z_\xi and \( \lambda \) are standard Brownian motions with \( dz_\xi dz_\lambda = \rho dt \). \( \lambda \) will be referred to as the short-term deviation in prices (temporary changes in prices that are not expected to persist), and \( \xi_T \) will represent the equilibrium price level (fundamental changes that are expected to persist).

These factors are “orthogonal” in their dynamics, which implies a small correlation between the stochastic increments of those factors.

Let \( F_{T_t} \) denote the time-t market price for a futures contract with time \( T \) until maturity.

\[ \ln(F_{T_t}) = e^{-k(T-t)} \chi_t + \xi_t + A(T-t), \]  

(4)

where

\[ A(T-t) = \begin{vmatrix} \mu \xi (T-t) - \frac{(1-e^{-k(T-t)}) \lambda \xi}{k} \\ \frac{(1-e^{-k(T-t)}) \sigma_\xi^2}{2k} + \frac{(1-e^{-k(T-t)}) \sigma_\chi^2}{2k} \\ \frac{2(1-e^{-k(T-t)}) \sigma_\xi \sigma_\chi}{2k} \end{vmatrix} \]

It assumes only two sources of randomness, which may barely hold all features of futures prices.

The Kalman filter is a recursive procedure for computing estimates of unobserved state variables based on observations that depend on these state variables. Given a prior distribution on the initial values of the state variables and a model describing the likelihood of the observations as a function of the true values, the Kalman filter generates updated posterior distributions for these state variables in accordance with Bayes’ rule. This is the procedure followed to compute the unobserved state variables in this model.
Energy forwards and futures

The Gibson-Schwartz model

In the Gibson-Schwartz model,

\[
dX_t = \left( \mu - \frac{\sigma_1^2}{2} \right) dt + \sigma_1 dz_1 \\
d\delta_t = (k(\alpha - \delta) - \lambda) dt + \sigma_2 dz_2 \\
X_t \text{ denotes the log of the time-} t \text{ current spot price (}, S_t \text{), } \delta_t \text{ denotes the instantaneous time-} t \text{ convenience yield (it includes the instantaneous cost of carry).}
\]

(5)

\[
z_1 \text{ and } z_2 \text{ are standard Brownian motions with } dz_1 dz_2 = \rho dt .
\]

\[
X_t \text{ denotes the log of the time-} t \text{ current spot price (}, S_t \text{), } \delta_t \text{ denotes the instantaneous time-} t \text{ convenience yield (it includes the instantaneous cost of carry).}
\]

\[
\ln(F_{T,t}) = \left( \frac{1 - e^{-k(T-t)}}{k} \right) \delta_t + X_t + A(T-t). \quad (6)
\]

where

\[
A(T-t) = \left\{ \begin{array}{cc}
\left( r_t - \alpha + \sigma_1 \sigma_2 \rho - \frac{\sigma_2^2}{2k^2} \right) (T-t) \\
\quad + \left( \frac{1 - e^{-2k(T-t)}}{4k^3} \right) \sigma_2^2 \\
\left( 1 - e^{-k(T-t)} \right) \frac{\alpha k + \sigma_1 \sigma_2 \rho - \frac{\sigma_2^2}{k} }{k^2}
\end{array} \right.
\]

This model does not take into consideration either stochastic interest rates or stochastic storage costs.

The Schwartz-Smith model is formally equivalent to the Gibson-Schwartz model, but the former has an important advantage, which is more transparent analytic results. Since many find the notion of convenience yield elusive, the idea of stochastic-evolving short-term deviations and equilibrium prices seems more natural and intuitive.

The parameters of the model cannot be estimated with precision, because they are based on estimated values for the convenience yield.

The Schwartz model

The Schwartz model gives:

\[
dX_t = \left( (r_t - \delta) S_t dt + S_t \sigma_1 dz_1 \right) \\
d\delta_t = k(\alpha - \delta) dt + \sigma_2 dz_2 \\
tr_t = \alpha(m - r_t) dt + \sigma_3 dz_3 . \quad (7)
\]

\[
z_1, z_2 \text{ and } z_3 \text{ are standard Brownian motions, with:}
\]

\[
dz_1 dz_2 = \rho_1 dt \\
dz_1 dz_3 = \rho_2 dt \\
dz_2 dz_3 = \rho_3 dt .
\]

Where \( X_t \) denotes the log of the time-} t \text{ current spot price (}, S_t \text{), } \delta_t \text{ denotes the instantaneous time-} t \text{ convenience yield (here, it includes the instantaneous cost of carry), and } r_t \text{ denotes the instantaneous interest rate:}

\[
\ln(F_{T,t}) = \left( \frac{1 - e^{-k(T-t)}}{k} \right) \delta_t + \left( 1 - e^{-\alpha(T-t)} \right) r_t \\
\quad + C(T-t) + X_t.
\]

Here, \( C(T-t) \) is just a complicated function of \( T-t \).

The futures prices also have a number of influential factors that may not be considered (e.g., there is no storage cost included as a source of randomness).

One of the difficulties in the empirical implementation of commodity price models is that, frequently, the factors or state variables of these models are not directly observable. For some commodities, the spot price is hard to obtain, and the futures contract closest to maturity is used as a proxy for the spot price. The problems of estimating the instantaneous convenience yield are even more complex; normally, futures prices with different maturities are used to compute it. The instantaneous interest rate is also not directly observable. Futures contracts, however, are widely traded in several exchanges and their prices are more easily observed.
The Miltersen and Schwartz model

This model generalizes and combines the two approaches by using all the information in the initial term structures of both interest rate and commodity futures prices:

\[ rf(t, s) = rf(0, s) + \mu_f(u, s)du + \sigma_f(u, s)dW_u \]  \hspace{1cm} (8)
\[ \delta f(t, s) = \delta f(0, s) + \mu_{\delta f}(u, s)du + \sigma_{\delta f}(u, s)dW_u \]  \hspace{1cm} (9)
\[ S(t) = S(0) + S\mu_f(u)du + S\sigma_f(u)dW_u. \]  \hspace{1cm} (10)

\( W_u \) is the same \( d \)-dimensional Wiener process for every differential equation. Correlations among the three processes come via the specification of the diffusion terms \( (\sigma_s) \).

The no arbitrage conditions completely determine the drift terms \( (\mu_s) \):

\[ \mu_f(t) = rf(t, t) - \delta f(t, t) \]
\[ \mu_f(t, s) = \sigma_f(t, s)\sigma_f(t, v)dv \]  \hspace{1cm} (11)
\[ \mu_{\delta f}(t, s) = \sigma_{\delta f}(t, s)\sigma_{\delta f}(t, v)dv \]
\[ + (\sigma_f(t, T) - \sigma_{\delta f}(t, T)) \]
\[ \cdot \left( \sigma_s(t) + (\sigma_f(t, s) - \sigma_{\delta f}(t, s))ds \right). \]  \hspace{1cm} (12)

The expression for the futures prices of the commodity would be:

\[ F(t, T) = S_t \exp \left( \int_0^T (rf(t, s) - \delta f(t, s))ds \right). \]  \hspace{1cm} (13)

The Cortazar and Schwartz model

This model is a particular case of the previous one: it makes the assumption that the interest rate is constant, which implies no analogue of Equations 8 and 11 and the following version of Equation 12 as a condition for no-arbitrage.

Thus:

\[ \mu_{\delta f}(t, s) = \left( -\sigma_{\delta f}(t, T) \right) \left( \sigma_s(t) - \sigma_{\delta f}(t, s)ds \right). \]  \hspace{1cm} (14)

The Thomas Urich model

This model is an extension to commodities, specifically metals, of the Garbade (1996) work on term structure for interest rates. We have:

\[ F_{t, T} = f_0(T-t) + \sum_{i=1}^{I} w_i(t)f_i(T-t) \]
\[ f_i(T-t) = \sum_{j=1}^{J} b_{ij}(T-t)^{j-1} \]  \hspace{1cm} (15)
\[ f_0(T-t) = \sum_{j=1}^{J} b_{0j}(T-t)^{j-1}. \]

In particular,

\[ F_{t, t} = b_{00} + \sum_{i=1}^{I} w_i(t)b_{i0} \],

which yields the spot price behaviour.

Here, \( w_i \) are Gaussian random walks with zero drift and unit variance per year; the random walks are statistically independent of each other. The function \( f_i \) is referred to as “modes of fluctuations.”

In general the idea behind an arbitrage-free model is that the expected rate of return on any portfolio of forwards is a function only on the risks of the portfolio, and it does not otherwise depend on the details of the composition of the
Energy forwards and futures

portfolio. That is, we assume that there exists a function $G$ such that:

$$\mu = G(\sigma_j, w_j).$$

(16)

This equation says that all portfolios with risk measures $\sigma_j$ have the same expected rate of return $\mu$.

Equation 16 is not as arbitrary as it may appear. Suppose, contrary to what is stated in Equation 16, that we had two portfolios of forwards with equal values and identical risk, but different expected rates of return. We could then form a composite (or hedge) portfolio with zero net value and zero net risk such that the net value of the portfolio would, with certainty, increase over time. This would imply the existence of “arbitrage” profits, or riskless revenues without any commitment of capital.

In the case of a linear portfolio, it can be shown that the function $G$ must be affine with respect to the risk measures or the form:

$$\mu = g_0(w) + \sum_{j=1}^{J} g_j(w) \cdot \sigma_j.$$ 

(17)

From Equations 15, 16 and 17 one derives the following consistency conditions (in order for the model to be arbitrage free):

$$I > 2$$

$$I = J, J_1$$

$$J_1 = 2*J + 1.$$

This model makes a very important assumption in order to estimate its parameters, that is, one should have available, at any time, the future prices for the same time to maturity. At times, this may not be realistic, therefore, approximations are required to fulfill the requisite conditions. Fulfilling these conditions becomes the main difficulty with this model.

A proposed model

In order to overcome Ulrich’s assumption, we make a few changes in the estimation procedure and the structure of the model, which yields the following model:

$$F_{t,T} = a_{0,t} + (T-t)a_{1,t} + a_{2,t}(T-t)^2 + \epsilon_{t,T}$$

(18)

with $\epsilon_{t,T}$ an $N(0, \sigma^2)$ distribution and $a_{0,t}$, $a_{1,t}$ and $a_{2,t}$ time-series with the following properties:

$$\log a_{0,t} = \log a_{0,t-1} + \text{res}_{t,0}$$

$$a_{1,t} = a_{1,t-1} + \text{res}_{t,1}$$

$$a_{2,t} = a_{2,t-1} + \text{res}_{t,2}.$$

Having these properties means that the model is specified by the coefficients $a_{t,t}$ on which one now needs to make statistical assumptions. In this paper, we propose dealing with modelling the coefficients as follows:

The marginal structure will be modelled using univariate features: a Box-Cox transformation, the estimation of a stochastic trend, and the calibration of non-normal residuals. We assume that the dependence structure is determined by normal rank correlations. The way to estimate the parameters of the model is by using a minimum least square at every $t$-time, using the different time-to-maturity $(T_t-t)$ as the fitting data.

This model is not arbitrage free because, following Garbade’s idea (explained by the Ulrich model above) and the equivalence between both models, there is a violation of the no-arbitrage conditions. As is shown in the section below, this is what is required to obtain a model with better forecasting properties.

Empirical results

We worked with 12 time-series of oil futures prices from January 1990 to December 2000. Each series corresponds to a specific maturity
month, that is, we have the future prices for $T_i - t$, (where $i =$January,…,December) time to maturity at $t$-time. Our goal is to provide a good fit for the one-day distribution with one-day horizons of a portfolio of forwards contracts. In other words, we try to forecast the exact values of the percentiles for tomorrow’s distribution, using past information on the vector of futures prices.

Due to the fact that the Value-at-Risk is the most important measure for risk management, which is just the 95 percentile of the distribution, we highlight the results regarding this particular percentile. We take into consideration percentiles other than 95. For example, if a portfolio depends on futures prices in a non-linear way (i.e., options on futures), then small percentiles may have an influence in the portfolio VaR.

We introduce the following notation:

$$\text{Dist}_{\alpha}(P,t-1): \text{Percentile } \alpha \text{ of the conditional distribution of the portfolio } P \text{ for time } t, \text{ given information available at time } t-1. \text{ } P_i \text{ is the real value for portfolio } P \text{ at day } i.$$ 

The number of times that the real value of portfolio $P_i$ is bigger than $\text{Dist}_{\alpha}(P,t-1)$ for a given percentile $\alpha$ (i.e., 95%) over 250 times should be a number close to $250(1 - \alpha)/100$, (12.5).

Thus, by the creation of the following Bernoulli random variable, we are able to assess how well every percentile is recovered (the test of goodness-of-forecasting). This test is equivalent to the one explained in Diebold et al. (1998, 1999), which is based on the $iid - U(0,1)$ properties of the probability integral transform series given by:

$$z_t = \text{Dist}_{\alpha}(P_i, t-1)$$

$$B_{\alpha} = 1 \text{ if } P_i > \text{Dist}_{\alpha}(P_i, t-1)$$

$$0 \text{ otherwise.}$$

Create the following graph representation: percentiles from 0.01 to 0.99 are allocated in the axis $X$, numbers:

$$\sum_{i=1}^{250} B_{\alpha}$$

are allocated in the axis $Y$.

Thus, pairs $(\alpha, Y_{\alpha})$ show the behaviour of the whole conditional distribution for the portfolio. We will use $(\alpha, Y_{\alpha})$ in Table I as a comparative tool. The closer to one the values of $Y_{\alpha}$ are, the better the model is in its forecasting abilities. As has been suggested by Diebold (1998, 1999), the previous visual assessment is more revealing, constructive and attractive than more sophisticated procedures, such as kernel density estimates.

It is important to point out that we do not estimate the parameters of our model by minimizing the goodness-of-forecasting test. We obtain the fitting of the term structure by linear regression between the vector of futures prices and the vector of time to maturities. Then we work on the new vector series, using properties of AR processes to forecast a one-day horizon.

Because of this, the results of the goodness-of-fit test are meaningful from a model validation perspective; a method that optimizes the goodness-of-fit results would rely on a separate validation methodology.

To try the model, we created a simple portfolio using RiskWatch software. The portfolio consists of three instruments: a commodity forward that starts August 23, 1998 and matures in 90 days; the second one that starts at the same time and matures in 120 days; and the last one that has the same start date and matures 250 (business) days later. The fitting methodology uses data up until August 23, 1998; subsequent market data was used only in the out-of-sample goodness-of-fit tests.

Several sets of scenarios were created using different models, arising from different considerations on the residuals, resp, to be normal/non-normal or dependent/independent. This allows
one to create a family of implementations for our proposed models, and, in this way, one can decide on the assumptions of normality or independence as a function of their forecasting abilities. In other words, we will forecast the forward prices for tomorrow (a one-day horizon) through the forecasting of the parameters of our model, using information until today, and using sets of scenarios based on different assumptions for the futures prices series, as follows:

We assume:

1. that the residuals \( r_t \) in the proposed model follow their empirical distribution, taking into account the last 100 values for roughly four months (EP100).

2. an empirical distribution of the residuals \( r_t \) in the proposed model, taking into account the last 250 values (EP250).

3. normal distribution of the residuals \( r_t \) in the proposed model, taking into account the last 250 values under the assumption of independence for the residuals (NI250).

4. normal distribution of the residuals \( r_t \) in the proposed model, taking into account the last 250 values (ND250).

5. a unit root in every univariate future prices series, multivariate normal for the residuals (using principal components), 250 previous values considered. (URN) mean reversion in every univariate future prices series, multivariate normal for the residuals (using principal components) from the AR(1), 250 previous values considered (MRN).

6. mean reversion and GARCH in every univariate future prices series, multivariate normal for the residuals (using principal components) from the AR(1)-GARCH(1,1), 250 previous values considered (MRGN).

7. mean reversion and GARCH in every univariate future prices series, empirical distribution for the multivariate residuals from the AR(1)-GARCH(1,1), 250 previous values considered (MRGEP).

<table>
<thead>
<tr>
<th>Percentiles</th>
<th>Our Model</th>
<th>Classical approaches</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>EP100</td>
<td>EP250</td>
</tr>
<tr>
<td>0.4</td>
<td>Right</td>
<td>0.925</td>
</tr>
<tr>
<td></td>
<td>Left</td>
<td>1.045</td>
</tr>
<tr>
<td>0.6</td>
<td>Right</td>
<td>1.022</td>
</tr>
<tr>
<td></td>
<td>Left</td>
<td>0.963</td>
</tr>
<tr>
<td>0.7</td>
<td>Right</td>
<td>0.989</td>
</tr>
<tr>
<td></td>
<td>Left</td>
<td>0.935</td>
</tr>
<tr>
<td>0.8</td>
<td>Right</td>
<td>1.146</td>
</tr>
<tr>
<td></td>
<td>Left</td>
<td>0.870</td>
</tr>
<tr>
<td>0.9</td>
<td>Right</td>
<td>1.283</td>
</tr>
<tr>
<td></td>
<td>Left</td>
<td>0.842</td>
</tr>
<tr>
<td>0.95</td>
<td>Right</td>
<td>1.118</td>
</tr>
<tr>
<td></td>
<td>Left</td>
<td>0.677</td>
</tr>
</tbody>
</table>

Table 1: Model comparison showing the multivariate conditional distribution.
Table 1 shows a comparison among the different models we have considered. It also shows a very interesting feature about the multivariate conditional distribution (from now on MCD, which is the one needed to forecast a one-day horizon) of the futures prices vector; the normal multivariate distribution of the vector of residuals \( (\text{res}_i, \text{ND250}) \) give a very good fit of the MCD of futures prices vector, implying a normal MCD! The mean vector and covariance matrix of the MCD depend heavily on the non-linearity of previous data (due to the fact that the new time-series vectors are not a linear combination of the futures prices vector). This model shows that the dependency through time to maturity changes according to the \( t \)-time chosen and vice versa. An attempt to capture those dependencies separately (e.g., assuming identical distributions on the prices and fitting their marginals and dependence structures separately) leads to unreliable results.

This assumption, in turn, leads us to think that the behaviour in \( t \) and \( T - t \) should be faced simultaneously. Nevertheless, there are different ways to transform a data vector simultaneously into a new independent and equally distributed one, (i.e., Vector AR, multivariate GARCH, Schwartz model). However, those models fail to get rid of this specific form of multivariate non-stationarity. Results in classical approaches (models five to eight in Table I) show how important it is to take care of the multivariate dependence relation through time and time to maturity by using the contango and backwardation features of the futures price vector.

Conclusions

We realize that the most influential characteristic in a portfolio of oil forwards (futures) is the multivariate dependency structure, specifically dealing with periods of backwardation and contango, leading to a search for a model which incorporates this feature. When used to forecast next-day risk measures (1-day VaR), the models proposed in the literature have as their main shortcomings:

- inaccuracy in the input data (i.e., errors on the estimated spot prices, instantaneous convenience yields, futures convenience yields, and so on)
- unrealistic assumptions (i.e., futures prices for the same time to maturity at every \( t \)-time)
- do not take into considerations the storage cost as a stochastic process and, in general, do not allow new sources of randomness (TS).

The “moving windows” test that validates the prediction (from a practical point of view) of any given model requires little historical data (100 – 250 previous observations) in order to obtain good predictions for a one-day-horizon on quantiles less than 0.90 – 0.95. Also, by following the ideas showed in Garbade (1996) for interest rates, and extended for metals by Urich (2000), it is possible to make the required modification in the proposed model in order to get a better fit of any portfolio of futures (tested specifically in the oil market).

Further, we have seen that a multivariate normal distribution in a few factors is able to explain the whole vector of futures prices. These factors are not obtained by principal components because correlations are not what should be conserved, and are the wrong measure of dependence.

Concerning the univariate forward (spot) oil time-series, we realize the following:

- If the goal is to get a good short-term forecast, then it is better to work with the returns (log of differencing). There is no significant presence of linear correlation (ARMA) on the returns, but there is a high correlation in the return square (GARCH(1,1)), with non-normal residuals.
- As for the medium- and long-term forecasting, there is a marked reversion to the mean, and a high correlation in the
residuals square (GARCH(1,1)).

- No deterministic or stochastic seasonality was observed in any case.

Acknowledgments

This research was funded by Algorithmics Inc., CITO (an Ontario Centre of Excellence) and Mitacs (a Canadian Centre of Excellence). The authors wish to express their gratitude to Olivier Croissant for insightful comments and advice.

References


