MODELING AND ESTIMATING COMMODITY PRICES: COPPER PRICES*

Roger J-B Wets
University of California
Davis, CA 95616
rjbwets@ucdavis.edu

Ignacio Rios
University of Chile
Santiago, Chile
irios@ing.uchile.cl

Abstract. A new methodology is laid out for the modeling of commodity prices, it departs from the 'standard' approach in that it makes a definite distinction between the analysis of the transient (short term) and stationary (long term) regimes. In particular, this allows us to come up with an explicit drift term for the transient process whereas the stationary process is primarily driftless due to inherent high volatility of commodity prices, except for an almost negligible mean reversion term, Not unexpectedly, the information used to build the transient process relies on more than just historical prices but takes into account additional information about the state of the market. This work is done in the context of copper prices but a similar approach should be applicable to wide variety of commodities although certainly not all since commodities come with very distinct characteristics. In addition, our model also takes into account inflation which leads us to a multi-dimensional nonlinear system for which we can generate explicit solutions.

Keywords: commodity prices, epi-splines, short and long term, best fit, scenario tree.

JEL Classification: C53
Date: October 17, 2012.

*This project was started while the first author was visiting the Centro de Modelamiento Matematico and Systemas Complejos de Ingeneria, Universidad de Chile. His research was supported in part by the U. S. Army Research Laboratory and the U. S. Army Research Office under grant number W911NF-10-1-0246. The research of the second author was financed by Complex Engineering Systems Institute (ICM:P-05-004-F, CONICYT: FBO16) and Fondecyt project 1120318 .

1 Introduction

The modeling of a price process associated with one or more commodities is of fundamental importance not only in the valuation of a variety of instruments and the derivatives associated with these commodities but also in the formulation of optimization and equilibrium models, aimed at finding 'optimal' extraction and/or storage strategies, that are bound to involve these prices as parameters. Although our overall approach is clearly applicable to a wide range of commodities, in this article we are going to restrict our attention to copper prices that will allow us to highlight, in a practical instance, the main features of our methodology. Copper prices are highly volatile and depend on many external factors: existing copper stocks and contracts, deposits discoveries, the local and world-wide economic environment and technological innovations, for example.

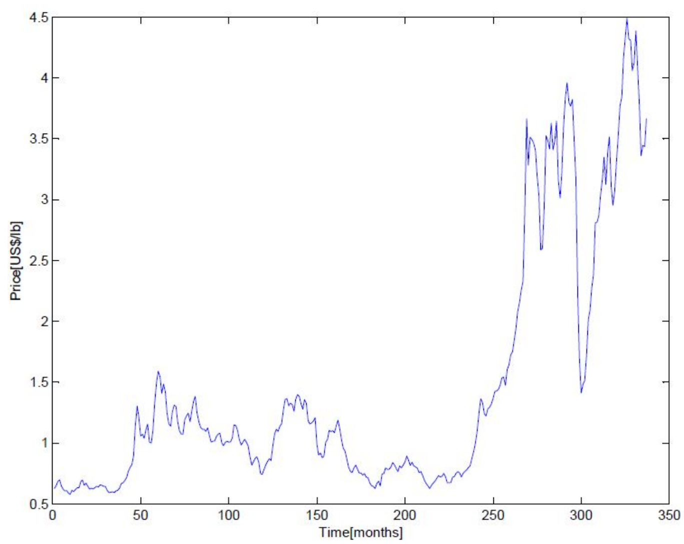

Figure 1: Historical copper prices from 1980 to 2011

This inherent high volatility renders the modeling particularly challenging. Our approach departs significantly from earlier efforts in a number of ways. To begin with, we make a distinction between the short term that can be viewed as the *transient process* and the long term that can be considered as the *stationary process*†. To find appropriate estimates for these processes we rely, as is standard, on historical prices but take also advantage of market information to build the transient component of the process. A complete description of the state-of-the market, i.e., involving existing and potential

†Splitting time in short and long term, in the case of copper prices, is in line with the results of Ulloa [11], who concludes after applying unit root tests to subsets of data of different lengths that shocks affect only in the short term, because in the long term copper prices should revert to their long term mean price presenting in the interim a high volatility.

(under exploration) reserves, accumulated stocks, deliverable and 'purely financial' contracts might turn out to be useful, but actually such detailed analysis of the market is reflected in the futures contracts quoted at various metal exchange markets: COMEX (New York), LMEX (London) and SHMETX (Shanghai). However, to exploit this information, this market information (futures) into spot prices and how we proceed is explained in the section dealing with the transient process. The main reason for making a distinction between the short and the long term comes from the fact that the high volatility suggests that no drift term can reliably be associated with the stationary process whereas recent historical prices complemented by market information should allow us to identify a drift in the formulation of the transient process‡. Our model also takes into account inflation which leads us to a multi-dimensional (nonlinear) system for which we can generate explicit solutions.

The remainder of this article is organized as follows. In §2 and 3 we present the guiding models for the long and short term processes. The long term,or stationary, process is analogous to some other commodity models found in the literature that we review briefly in that section. On the other hand, the short term, or transient, component of our model departs significantly from standard approaches and allows us to obtain better predictable behavior. In §4 we present our full model which results in a nonlinear stochastic differential system is a blending of the short and long term regimes. In §5 we describe the data used to estimate and test our models and finally, provide an empirical analysis in §6.

2 The stationary process

In our model, the long term regime will take the attributes of a stationary process which will be mostly in line with what can be found in the literature for the 'overall' process. Since this is to a large extent familiar territory, we want to get it out of way rather expediently. The only issue that needs some concern is to decide if the model should be build with or without mean reversion and there is really no consensus that has emerged from a rather elaborate analysis.

On one side, basic microeconomics theory says that when prices are high the supply will increase because higher cost producers will enter the market and that will push down prices, returning to the market equilibrium price. Conversely, if prices are relatively low some producers will not be able to enter the market and the supply will decrease, stimulating a rise in prices. The mean reversion theory, introduced by [19], is supported by many authors: [3] prove the existence of mean reversion in spot asset prices of a wide range of commodities using the term structure of future prices; [1] proves the same using the ability to hedge option contracts as a measure of mean reversion; [17] compare three models of commodity prices that takes into account mean reversion, and there is many other authors that use mean reverting processes to model commodity prices.

‡The inclusion or not of a mean reversion term in the stationary process will be taken up in the section devoted to the stationary process.

On the other side, results show that in some cases mean reversion is very slow, and in others the unit root test fails to reject the random walk hypothesis. For example, [6] apply this test to crude oil and copper prices over the past 120 years, and they reject the random walk hypothesis, which confirms that these prices are mean reverting. However, when they perform the unit root test using the data for only the past 30 or 40 years, they fail to reject the random walk hypothesis. The explanation they give to this result is that the speed of reversion is very low, so using 'recent' past data is difficult to statistically distinguish between a mean-reverting process and a random walk. Then, they conclude that one should rely more on the theoretical and economical consistency (for example, intuition concerning the operation of equilibrium mechanisms) than in statistical tests when deciding which kind of model is better.

Another example is given by [7], where they test many different models to predict medium term copper prices (from one to five years) and they conclude that the two models with better performance are the first-order autoregressive process and the random walk.

This evidence suggest that in the short term (one year) there may be no mean reversion, which is very logical because a producer can not open suddenly a new plant if prices are high or close the mine if prices are low. This is again an argument that supports our approach that disconnects short and the long term effects and will rely to a large extent on a different data base to build the two main components of our model.

For the long term we set up a stochastic differential equation that is mean reverting and, which in turn, will determine the drift of the stationary process. We rely on a variant of geometric brownian motion with mean reversion which is also in tune with our choice of inflation free 'money', cf. §5.

This model was proposed by [6], and it's also used by [14] to model oil prices[§].

So, for the stationary process the following system of stochastic differential equations provide us with the basis for the modeling process:

$$dx_i^t = \mu_i \left(\upsilon_i - x_i^t\right) dt + \left(\sum_{j=1}^{J} b_{ij} dw_j^t\right) x_i^t, \quad i = 1, \ldots, n, \tag{1}$$

$$x_i^{t_0} = x_i^0, \quad i = 1, \ldots, n \tag{2}$$

where x_i^0 is the present value of index i (is given), μ_i and b_{ij} are constants that need to be estimated, $x^t = \left(x_1^t, \ldots, x_n^t\right)$ is the state of the system at time t, $w_j, j = 1, \ldots, J$ are independent (standard) wiener processes, υ_i is and index to which x_i^t reverts in the

[§]In the Pilipovic model, prices are modeled by a system of two stochastic differential equations: the first one for the spot price, which is assumed to mean-revert toward the equilibrium price level, and the second for the equilibrium price level, which is supposed to follow a log-gaussian distribution,

$$dS_t = \alpha \left(L_t - S_t\right) dt + \sigma S_t dw_t$$
$$dL_t = \mu L_t dt + L_t \xi dz_t$$

log term and μ_i is the 'speed' at which x_i^t reverts to υ_i; our strategy will be that this mean-reversion drift is very slow and consequently, its influence is quite attenuated.

The solution of this system is: for $i = 1, \ldots, n$,

$$x_i^t = x_i^{t_0} \exp\left[\left(\mu_i + \frac{1}{2}\sum_{j=1}^{J} b_{ij}^2\right)(t - t_0) + \left(\sum_{j=1}^{J} b_{ij}\big(w_j^t - w_j^{t_0}\big)\right)(t - t_0)\right] + \mu_i\upsilon_i \int_0^t e^{r_i(t,s)} ds$$

where,

$$r_i(t,s) = -\left[\mu_i + \frac{1}{2}\sum_{j=1}^{J} b_{ij}^2\right](t - s) + \sum_{j=1}^{J} b_{ij}\left(w_j^t - w_j^s\right).$$

We are going to replace this solution by an approximate one obtained by replacing the term $\mu_i\upsilon_i \int_0^t e^{r_i(t,s)} ds$ by its expectation. In the Appendix A, we justify this approximation. We proceed in this manner since for all practical purposes the error introduced by this approximation is negligible and that the, eventual estimation of the coefficients μ_i, υ_i and b_{ij} would be very onerous, if not practically impossible.

So, we accept as 'solution' to the system of stochastic differential equations: for $i = 1, \ldots, n$,

$$x_i^t = \upsilon_i\left(1 - e^{-\mu_i t}\right) + x_i^0 \exp\left[-\left(\mu_i + \frac{1}{2}\sum_{j=1}^{J} b_{ij}^2\right)(t - t_0) + \sum_{j=1}^{J} b_{ij}\left(w_j^t - w_j^{t_0}\right)\right]$$

and considering $t_0 = 0$ we obtain,

$$x_i^t = \upsilon_i\left(1 - e^{-\mu_i t}\right) + x_i^0 \exp\left[-\left(\mu_i + \frac{1}{2}\sum_{j=1}^{J} b_{ij}^2\right) t + \sum_{j=1}^{J} b_{ij} w_j^t\right] \tag{3}$$

which is also a *log-gaussian* process. A 1-dimensional version of this process reads,

$$dx^t = \mu\big(\upsilon - x^t\big)dt + \sigma x^t dw^t, \quad x^{t_0} = x^0$$

with solution:

$$x^t = \upsilon\left(1 - e^{-\mu t}\right) + x^0 \exp\left[\big(\mu + \frac{1}{2}\sigma^2\big)t + \sigma w^t\right],$$

Finally, to calculate the mean and the covariance terms of the n-dimensional process, we rely again on the properties of gaussian processes. One obtains: for $i = 1, \ldots, n$,

$$\mathbb{E}[x_i^t] = \upsilon_i + \left(x_i^0 - \upsilon_i\right) e^{-\mu_i t} \tag{4}$$

$$\mathrm{cov}\{x_k^t, x_l^t\} = x_k^0 x_l^0 e^{-(\mu_k + \mu_l)t}\left(\exp\left[t\sum_{j=1}^{n} b_{kj} b_{lj}\right] - 1\right) \tag{5}$$

and in particular we have $\mathbb{V}[x_k^t] = (x_k^0 e^{-\mu_k t})^2 \left(e^{t\sum_{j=1}^n b_{kj}^2} - 1\right)$,
and in the 1-dimensional case,

$$\mathbb{E}[x^t] = \upsilon + (x_0 - \upsilon)\, e^{-\mu t}, \qquad \mathbb{V}[x^t] = \left(x_0^2 e^{-\mu t}\right)^2 \left(e^{\sigma^2 t} - 1\right).$$

Brief overview of the literature Although there is some overlap between the design of the stationary component of our model with some earlier work, it's difficult to make an orderly comparison since much of the novelty in our approach isn't featured, as far as we can tell, in any other proposed model. In order to emphasize, the departure of the proposed model from the relevant alternatives, we go through a brief review pointing out their salient features.

In general terms the literature oriented to modeling commodity prices can be classified in two categories: structural models and reduced form models. The first family aims to represent how partial equilibriums are reached in these markets. Then, a typical application considers models for the demand, the supply and the storage, and then an expression of the equilibrium price is derived from them. The basic equilibrium model is described by [22], and examples of this approach are presented by [2] and [15].

On the other hand, reduced form models assumes that the stochastic behaviour of commodity prices can be captured by stochastic differential equations. This approach is very popular because of its simplicity, and in absence of big changes in the market structure their predictive accuracy outperforms structural models [7]. However, most of the work has been oriented to the valuation of contingent claims, where a mean-reverting spot price model is combined with other factors to obtain a process for the valuation of different derivatives.

One of the most important examples of this approach is given by [17], who compares three models for the valuation of commodity contingent claims. In the first model, the logarithm of the spot price is considered as the unique factor and is assumed to follow a mean reverting Ornstein-Uhlenbeck process. Then, the spot price is given by:

$$dS = \kappa\,(\mu - \ln S)\, dt + \sigma S dz.$$

In his second model, [17] provides a variation of the two-factor [8] model whereas the spot price follows a mean reverting process given by:

$$dS = (\mu - \delta)\, S dt + \sigma_1 S dz_1.$$

Finally, in his third model [17] introduces a three factor model that extends the previous model by including the interest rates as a third stochastic factor. For this purpose, interest rates are modeled as a mean-reverting process, and the join process is given by:

$$\begin{aligned} dS &= (r - \delta)\, S dt + \sigma_1 S dz_1 \\ d\delta &= (\alpha - \delta)\, dt + \sigma_2 dz_2 \\ dr &= a\,(m - r)\, dt + \sigma_3 dz_3 \end{aligned}$$
$$dz_1 \cdot dz_2 = \rho_1, \quad dz_2 \cdot dz_3 = \rho_2, \quad dz_1 \cdot dz_3 = \rho_3$$

Extensions of the [17]'s two and three-factor models are presented by [9], [12], [13] and [15].

On the other hand, there are some multi-factor models that make an implicit distinction between short and long term. [14] defines a two-factor model by considering the spot price (S) and the long term equilibrium price (L). The first factor is assumed to mean-revert toward the equilibrium price level, and the second is assumed to follow a log-gaussian distribution:

$$\begin{aligned} dS_t &= \alpha \left(L_t - S_t\right) dt + \sigma S_t dz_t \\ dL_t &= \mu L_t dt + \xi L_t dw_t. \end{aligned}$$

A similar approach is followed by [18] which proposes a two factor-model which includes the short-term deviation in prices (χ_t) and the equilibrium price level (ξ_t) as factors:

$$\begin{aligned} d\chi_t &= -\kappa \chi_t dt + \sigma_\chi dz_\chi \\ d\xi_t &= \mu_\xi dt + \sigma_\xi dz_\xi \\ dz_\chi \cdot dz_\xi &= \rho_{\chi\xi} dt \end{aligned}$$

Then, from these factors the built the process for the spot price, which is given by $\ln\left(S_t\right) = \chi_t + \xi_t$.

3 The transient process

As explained earlier, the drift for the short term prices won't include a mean reversion component but the calculation of this drift term will be crucial in setting up a 'robust' process, both when working with valuations, especially shorter term valuations, and in the design of discrete versions of this process that would be appropriate as input in management models (via reliable scenarios). We again cast our model as a geometric brownian motion model (for short term copper price or commodities that exhibit similar properties), precisely because this model allows to capture the *drift* exploiting both historical and market information; it also eludes the possibility of negative prices. The innovative features of our model are mostly in the construction of this transient process. Our approach is consistent with the fundamental principle that (probabilistic) estimations should be based on all the information that can be collected rather than just 'observations'. Taking this into account, the inclusion of market information is crucial since implicitly it incorporates all the information available to which one could refer as indexes: market expectations/beliefs, stocks, production costs and other factors that affect prices.

In addition, we propose a model that incorporates in the volatility component the role played by these other indexes (variables) that may affect copper prices, such as inflation, productivity indexes, ... This leads us to a system of stochastic differential equations of the following type:

$$\begin{aligned} dx_i^t &= \left(\mu_i dt + \sum_{j=1}^{J} b_{ij} dw_j^t\right) x_i^t, \quad i = 1, \dots, n, \\ x_i^{t_0} &= x_i^0, \quad i = 1, \dots, n \end{aligned} \tag{6}$$

where x_i^0 is the present value of index i (given), μ_i and b_{ij} are constants that need to be estimated, $x^t = (x_1^t, \dots, x_n^t)$ is the state of the system at time t and $w_j, j = 1, \dots, J$ are independent (standard) wiener processes with solution: for $i = 1, \dots, n$,

$$x_i^t = x_i^{t_0} \exp\left[\left(\mu_i - \frac{1}{2}\sum_{j=1}^{J} b_{ij}^2\right)(t - t_0) + \left(\sum_{j=1}^{J} b_{ij}\left(w_j^t - w_j^{t_0}\right)\right)\right] \tag{7}$$

A 1-dimensional version of this process reads,

$$dx^t = \left(\mu dt + \sigma dw^t\right) x^t, \quad x^{t_0} = x^0 \tag{8}$$

with solution,

$$x^t = x^0 \exp\left[\left(\mu - \frac{1}{2}\sigma^2\right)(t - t_0) + \sigma\left(w^t - w^{t_0}\right)\right] = x^0 \exp\left[(\mu - \frac{1}{2}\sigma^2)t + \sigma\varepsilon\sqrt{t}\right],$$

where ε follows a standard gaussian distribution. Hence, x follows a *log-gaussian* distribution with parameters $\left((\mu - \frac{1}{2}\sigma^2)\Delta t, \sigma^2\Delta t\right)$. It follows,

$$\mathbb{E}[x^t] = x_0 e^{\mu t}, \qquad \mathbb{V}[x^t] = x_0^2 e^{2\mu t}\left(e^{\sigma^2 t} - 1\right)$$

and in the multi-dimensional case: for $i = 1, \dots, n$,

$$\mathbb{E}[x_i^t] = x_i^0 e^{\mu_i t}, \qquad \mathbb{V}[x_i^t] = \left(\mathbb{E}[x_i^t]\right)^2\left(e^{|t|\sum_{j=1}^{n} b_{ij}^2} - 1\right).$$

Of course, the system's parameters will be estimated by 'short term data' meaning relatively recent historical prices complemented by market information as explained next.

Exploiting market information. Usually the information available about a commodity, in our instance copper, is manifold: existing contracts, stocks by producers and consumers, exploration activity, location of recent discoveries, economic predictions (future demand), and so on. In order to take such wide range of information into account, one needs a dedicated research division to amalgamate this information so that it can be included in a model. We suppose that the traders in this commodity, and others that might affect it value, have actually taking all these factors into account when selling or buying futures. If we accept this as a premise, obtaining market information that can be actually in our modeling would require transforming the information we can collect about futures and *convert it into appropriate spot prices* for the next few months, say 9-12 months. We shall then rely on the recent (historical) prices in combination with these future spot prices to build the 'transient stochastic process'. This conversion relies mainly on the fact that we can deduce the spot rate curve from futures by relying on the epi-spline technology [16, 21, 20] which enables us to obtain approximates, of arbitrary accuracy, of the financial curves associated with the market values for this particular commodity via finite dimensional optimization problems.

In this framework it will suffices to consider epi-splines of second order, i.e., twice differentiable which take the following (particular) form:

$$z(t) = z_0 + v_0 t + \int_0^t \int_0^\tau x(s)\, ds\, d\tau, \quad t \in [0, T]$$

where

- $x : (0, T) \to I\!R$ is an arbitrary piecewise constant function that corresponds to the 2nd derivative of z,
- $\{v_0, z_0\}$ are constants that can be viewed as integration constraints.

Our construction is similar, at least in purport, to that in [21]: split the interval $[0, T]$ into N sub-intervals of length $\delta = T/N$ and let the function x, the second derivative of z, be constant on each one of these intervals, say,

$$x(t) = x_k, \quad t \in (t_{k-1}, t_k], \quad k = 1, \ldots, N$$

where $t_0, t_1, \ldots, t_L$ are the end points of the N sub-intervals. The curve $z \in [0, T]$ is completely determined by the choice of

$$z_0, v_0 \quad \text{and} \quad x_1, \ldots, x_N,$$

i.e., by the choice of a finite number $(N + 2)$ of parameters. Then, for $t \in (t_{k-1}, t_k]$ one has,

$$\begin{aligned} z(t) &= z_0 + \int_0^t z'(s) ds = z_0 + \sum_{j=1}^{k-1} \int_{t_{j-1}}^{t_j} z'(s)\, ds + \int_{t_{k-1}}^t z'(s\, ds) \\ &= z_0 + v_0 t + \delta \sum_{j=1}^{k-1} \left(t - t_j + \frac{\delta}{2} \right) x_j + \frac{1}{2} (t - t_{k-1})^2 x_k; \end{aligned}$$

such a function belongs to $\mathcal{C}^{1,pl}$, i.e., it's continuously differentiable with piece-wise linear derivative. In our particular case, we want to generate a *spot curve* for the commodity prices by minimizing the deviations from the available data, i.e.,

$$\text{find} \quad z \in \mathcal{C}^{1,pl}([0, T], N) \quad \text{such that} \quad \|\vec{s} - z(\vec{t})\|_p \quad \text{is minimized} \tag{9}$$

where by two vector: $\vec{s} = (s_1, \ldots, s_L)$ which corresponds to the value of the assets considered and $\vec{t} = (t_1, \ldots, t_L)$ which are the dates when the assets generate the 'cash-flow' of the 'deliveries'.

To model commodity prices, here copper, we actually derive the corresponding discount factor curve df, which will be used to generate the spot prices. It's a function with the following properties:

- it should be nonnegative, decreasing and with $df(0) = 1$;
- the net present value of the 'cash-flows' must be as close as possible to zero,
- all the associated, forward-rates and spot, curves should be smooth.

Our problem can thus be reformulated as

$$\text{find} \quad df \in \mathcal{C}^{1,pl}\left([0,T], N\right) \quad \text{so that} \quad \|v\|_p \quad \text{is minimized}$$

with I our collection of instruments and $\vec{v} = \left(v_1, \ldots, v_{|I|}\right)$ is the corresponding vector of net present values, i.e.,

$$v_i = \sum_{t=1}^{L_i} df(t_{il}) s_{il}, \quad \forall i \in I.$$

For $t \in \big(\delta(k-1), \delta k\big], \delta = T/N$, $df \in \mathcal{C}^{1,pl}\left([0,T], N\right)$ can be expressed as:

$$df(t) = 1 + v_0 t + \delta \sum_{j=1}^{k-1} \left(t - t_j + \frac{\delta}{2}\right) x_j + \frac{1}{2}\left(t - t_{k-1}\right)^2 x_k \quad \text{where} \quad \tau = t - \delta(k-1) \quad (10)$$

where v_0 and $x_1, \ldots, x_N$ are the parameters that need to be estimated. It's noteworthy that df is linear in those parameters! As criterion, we rely on minimizing maximum error, i.e., $p = \infty$, the problem then takes the following form, with df as defined above,

$$\begin{aligned} \min \ & \max_{i\in I, l=1,\ldots,L_i} |df(t_{il}| & df'(t) \geq 0, \qquad \forall t \in [0,T] \\ & df(T) \geq 0, & \\ & v_0 \leq 0, x_k \in I\!R, k = 1, \cdots, N; & \end{aligned}$$

note that since $df(0) = 1$, the two first constraints will imply that $df \geq 0$ on $[0,T]$. Finding a discount factor curve is fundamentally an infinite dimensional optimization problem but the use of the epi-spline representation reduces it to a finite dimensional one that can exploit the well-tested standard optimization routines. How this is actually carried out is explained in [20, §6.2].

Spot rates, drift and initial conditions. Given the discount curve df, the *spot prices curve* is immediately available since

$$sp(t) = df(t)^{-1/t} - 1 \qquad \text{for} \quad t \in [0,T].$$

We have thus at our disposal recent historical prices, say for the last 9-12 months, today's price and the market prices for the next 9-12 months. It's this information that will be used, as explained in the next section to build the drift of the transient process. Again, we rely on an epi-spline fit to these prices to determine the 'drift' of the transient process. Combining this prices-information (short term past and relatively short term market spot-prices) to obtain the drift of the process, i.e., we 'fit' as well as possible our drift curve to these spot prices. The drift fit is obtain by solving the following optimization problem: find z_0, v_0 and $x_j, j = 1, \ldots, N$ such that

$$\|(sp(t) - z(t)), t = -9, \ldots t = 9\|_{\square} \quad \text{where}$$

$$z(t) = z_0 + v_0 t + \delta \sum_{j=1}^{k-1} (t - t_j + 0.5\delta)\, x_j + \frac{\tau^2}{2} x_k, \quad \forall t \in [0, T]\,.$$

where the interval $[-9, 9]$, the time span we want to take into account, has been subdivided in a $18/N$ mesh and $sp(t)$ for $t = -9, \ldots, 0$ are the observed historical prices and $sp(t)$ are the calculated market spot-prices. The drift of the transient process will then be taken to be the optimal solution of this optimization program $z^*(t)$ for $t \geq 0$. A noteworthy consequence of this approach is that he initial condition of our process will be $z^*(0)$ and not today observed price $sp(0)$. One justification for proceeding in this fashion is that one needs to view today (observed) spot price as the 'actual' spot price perturbed by some random factors; our empirical calculation confirm that choosing $z^*(0)$ as the initial point for the transient process yields better results.

4 Blending transient and stationary processes

There remain only to pass from the short term (transient) process to the long term (stationary) process to end up with a 'global' process. How to do this is still very much an open question that was only dealt with experimentally and via data analysis, and consequently, only in the context of copper prices. Extensive experimentation suggest that the transient process 'reign' is relatively short, the market reverts rather rapidly to its natural state. So, let's denote by X^T the transient process and by X^S the stationary process, and assume that the general process X is a blending of the transient and the stationary processes, i.e.,

$$X_t = \lambda_t X_t^T + (1 - \lambda_t)\, X_t^S, \tag{11}$$

with $X_{t_0} = X_0$ the initial state vector derived for the transient process and λ_t is a (decreasing) function of time. It's natural to think that in the short term $\lambda_t = 1$, i.e., the process is purely transient, and for the long term $\lambda_t = 0$, which means that there is no influence of the transient behavior. In the case of copper prices, we set

$$\lambda_t = \begin{cases} 1 & t \in [0, T_1], \\ \gamma^{T_1 - t} & t \in [T_1, T_2], \\ 0 & t \in [T_2, \infty) \end{cases}$$

with: T_1, T_2 and γ are parameters to be estimated. The estimation of these parameters is a serious challenge because no such study is available at this time for copper prices or any other commodity. However, we relied on our data analysis; one could also rely on experts' advice.

Once the parameters T_1, T_2 and γ have been defined we can build the blended process; cf. the implementation in the follow up sections. To do this, we have to recall that X^S and X^T can be approximated (locally) by multivariate gaussian distributions, so the convex combination of these processes will also be (locally) gaussian. In particular, if $X^T \sim \mathcal{N}(\mu_T, \Sigma_T)$ and $X^S \sim \mathcal{N}(\mu_S, \Sigma_S)$, hence

$$X_t = \lambda_t X_t^T + (1 - \lambda_t)\, X_t^S \sim \mathcal{N}(\lambda_t \mu_T + (1 - \lambda_t)\, \mu_S, \lambda_t^2 \Sigma_T + (1 - \lambda_t)^2\, \Sigma_S).$$

5 Data

The data used to estimate and test the models described above consist of monthly average observations of the LME spot copper price, from 01/1980 to 11/2012. Having a good amount of data is particularly important to estimate the stationary process, because one of our goals is to test how much historical data is better to consider in order to obtain better predictions for the long term copper price.

In particular, for our experiments with all the information available and the stationary process we proceeded to deflect prices by the US CPI, in order to avoid inflation effects. This wasn't done for the transient process experiments because market data comes in nominal terms, and also because in the short term prices do not change considerably due to inflation.

In addition to this, in order to include market information in the estimation of the transient process we used the first twelve LME copper future contracts. Then, for each month from 01/2005 to 10/2012 we used the average future price for each contract, and we combine this information with the short term historical spot prices to get an estimation of the parameters involved in each short term model.

Finally, as a first approach we consider just two factors to estimate the short and long term multivariate process: the spot copper price and the exchange rate between US dollars and UF¶. For this purpose, we consider monthly average data for the exchange rate, from 01/1984 to 10/2012. The used of the exchange rate in our copper price model is based on the work of Chen et al.[4], where is shown that "the Chilean exchange rate has strong predictive power for future copper prices". Furthermore, the inclusion of

¶UF is a Chilean monetary unit adjusted for inflation.

this factor allow us to incorporate a measure of the Chilean inflation, which might be also important in modeling copper prices. As [17] proposes in his three factor model, a third factor that could be included to our model is the interest rate. However, to keep simplicity we just consider two factors in this first approach.

Unit root tests As we discussed above, there is not a consensus if whether commodity prices exhibit mean reversion, and copper prices are not an exception. So, we proceeded to apply the best known unit root tests - Augmented Dickey-Fuller, Kwiatkowski–Phillips–Schmidt–Shin and Variance Ratio test - to the data we use in our experiments[‖].

	ADF test	KPSS test	VRatio test
LME spot copper prices 1980 - 2012	0.337 (0.775)	5.405 (0.010)	2.693 (0.007)
LME deflected spot copper prices 1980 - 2012	-0.756 (0.374)	4.949 (0.010)	2.737 (0.006)
LME spot copper prices 2005 - 2012	0.284 (0.752)	0.471 (0.010)	2.445 (0.010)

Table 1: Unit root tests applied to the data used in our experiments.

The results in Table 1 allow to conclude that the ADF test fails to reject the existence of a unit root in every set of data considered. In the same line, the null hypothesis of the KPSS test is rejected in all cases, so it confirms the results obtained with the ADF test. However, the results of the Variance Ratio test it follows that we can reject their null hypothesis at a confidence level of 95% in every case, which allow to conclude that the time series considered do not follow a random walk. This result contradicts those obtained with the other tests, and confirms the difficulty of determining the existence of unit roots in times series by using those tests.

6 Results

In this section we present the results obtained for the models presented above. In particular, we estimate them using the techniques described in Appendix B for the transient process and the generalized method of moments for our stationary process, and then we proceed to evaluate them in terms of their in-sample accuracy and their power of forecast out-of-sample.

[‖]We also implemented the Phillips-Perron test but the results obtained were the same as for the ADF test.

In our first experiment we proceed to estimate the one-dimensional model for the transient process considering just the spot copper price, and we evaluate the different ways to estimate the parameters and the effect of including market information.

Next we present the results related to the use of epi-splines, to incorporate a dynamic drift term, and then we show the results of our first approach to the multivariate transient process.

Finally, the results of the stationary process are shown, and an example of how to blend both processes is presented.

In order to compare the results obtained with the different approaches we examined the estimation results considering several criteria:

- In and out-of sample mean absolute error (MAD)
- In and out-of sample mean absolute percentage error (MAPE)
- In and out-of sample mean squared error (MSE)
- In and out-of sample weighted average error (WAE), weighted by $1/t_i$, where t_i is the period considered.
- In and out-of-sample maximum absolute error (MAE)
- In and out-of-sample root mean square error (RMSE)

6.1 Transient process

Incorporating market information One of the main questions around the transient process is the effect of incorporating market information in the estimation of the drift term. For this reason, for each period t we proceeded to estimate the parameters of our model considering just the recent historical information, $x_{t-1}, \ldots, x_{t-12}$, and also considering market prices, $f_t^1, \ldots, f_t^{12}$.

Then we calculated the expected value of our process for each case with the expression

$$\mathbb{E}\left[x^t\right] = \theta e^{\mu t}$$

for our novel approach, and

$$\mathbb{E}\left[x^t\right] = p_0 e^{\mu t}$$

for the classical method, and then we calculated the errors of each approach.

The results are shown in Table 2. The first two columns show the errors related to both estimation methods considering just the historical information, the next two columns considers both historical and market information, and the last column considers the estimation with the stationary model.

		Historical		Hist + Market	
		Novel	Classic	Novel	Classic
In-sample	MAD	0.618	0.515	0.412	0.539
	MAPE	0.172	0.195	0.145	0.204
	MSE	0.733	0.463	0.263	0.525
	WAE	0.825	0.452	0.437	0.454
	MAE	2.299	2.085	1.765	2.257
	RMSE	0.856	0.680	0.512	0.724
Out-of-sample	MAD	0.692	0.935	0.700	0.763
	MAPE	0.181	0.317	0.249	0.267
	MSE	0.907	1.457	0.814	1.006
	WAE	0.478	0.821	0.426	0.542
	MAE	2.490	3.024	2.472	2.730
	RMSE	0.952	1.207	0.902	1.003

Table 2: Comparison of the different approaches in estimating the short term drift.

From Table 2 we can see that considering market information improves considerably the performance out-of-sample of our model, reducing significantly almost the errors measured. In addition to this, the accuracy obtained by our novel approach is remarkably better, which proves that estimating a representative initial point is very important to get good predictions.

Another way to compare the predictive power of the different approaches is to built a confidence interval around the expected value estimated in each case. For doing this, we first estimate the variance at each point t in the future, which is given by

$$\mathbb{V}\left[x^t\right] = \left(\mathbb{E}\left[x^t\right]\right)^2 \left(e^{\sigma^2 t} - 1\right).$$

Then, for each time t we compute the upper and the lower value by the expressions

$$\overline{x}^t = \mathbb{E}\left[x^t\right] + \mathbb{V}\left[x^t\right]^{\frac{1}{2}}, \qquad \underline{x}^t = \mathbb{E}\left[x^t\right] - \mathbb{V}\left[x^t\right]^{\frac{1}{2}}$$

and with these values we have a confidence interval for each t given by $[\underline{x}^t, \overline{x}^t]$. Figure 6.1 shows examples of confidence intervals obtained for the model without and with incorporating market information respectively.

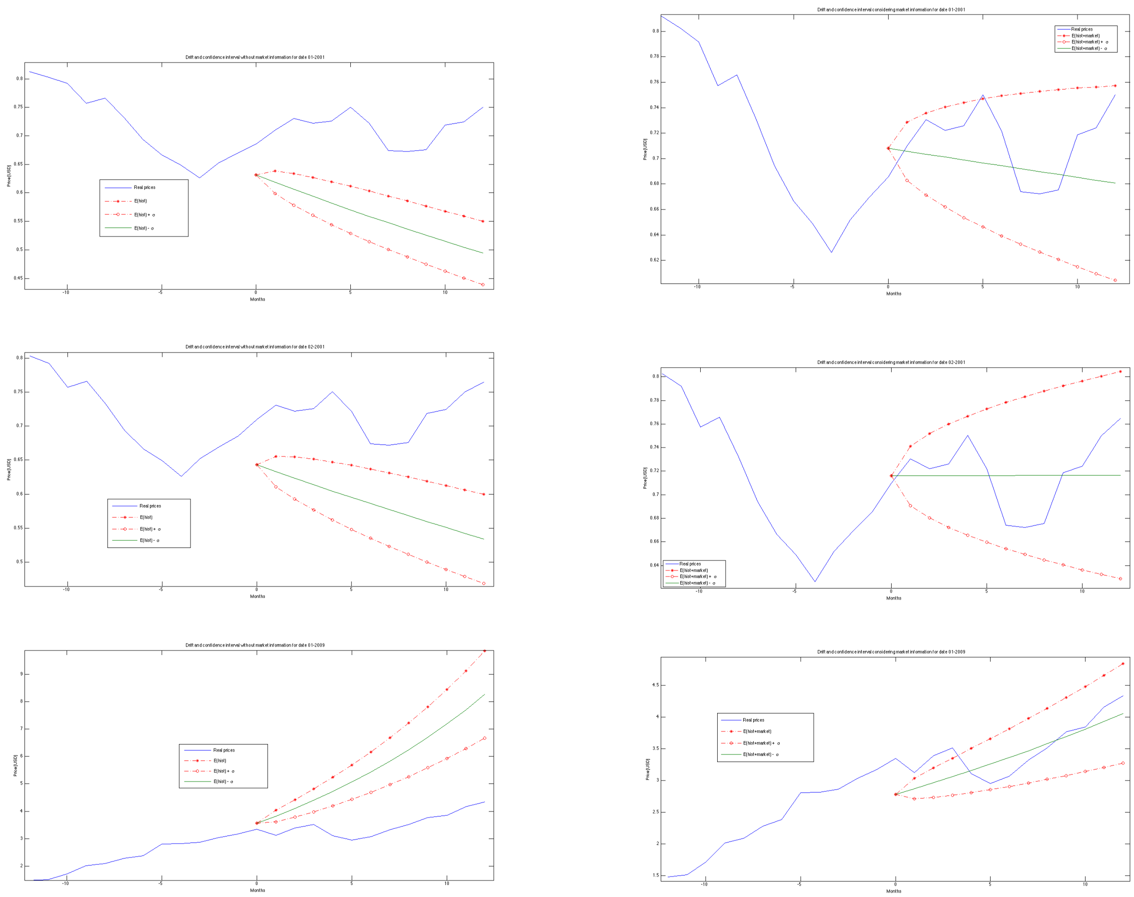

Figure 2: Example of confidence interval obtained for our model. The left plot is obtained estimating the parameters only with historical data, and the right plot is the one obtained considering historical and market information.

6.2 Multivariate transient process

In section 3 we have shown that the solution of the system of stochastic differential equations (6) is given by:

$$x_i^t = x_i^0 \exp\left[\left(\mu_i - \frac{1}{2}\sum_{j=1}^{J} b_{ij}^2\right) t + \left(\sum_{j=1}^{J} b_{ij} w_j^t\right)\right]$$

As w_j^t are standard and independent Wiener processes, we know that the term

$$z_i^t = \left(\mu_i - \frac{1}{2}\sum_{j=1}^{J} b_{ij}^2\right) t + \left(\sum_{j=1}^{J} b_{ij} w_j^t\right)$$

is normally distributed with mean $\left(\mu_i - \frac{1}{2}\sum_{j=1}^{J} b_{ij}^2\right) t$ and variance $\left(\sum_{j=1}^{J} b_{ij}^2\right) t$. Then, the variable $x_i^t = x_i^0 e^{z_i^t}$ is log-normally distributed, and the joint process $X^t = [x_1^t, \dots, x_J^t]$ is a multivariate log-gaussian process with mean and covariance matrix,

$$m^t = \begin{bmatrix} x_1^0 e^{\mu_1 t} \\ \vdots \\ x_J^0 e^{\mu_J t} \end{bmatrix} \tag{12}$$

$$\Sigma^t = \begin{bmatrix} \left(x_1^0 e^{\mu_1 t}\right)^2 \left(e^{|t|\sum_{j=1}^{J} b_{1j}^2} - 1\right) & \cdots & x_1^0 x_J^0 e^{(\mu_1+\mu_J)t} \left(e^{|t|\sum_{j=1}^{J} b_{1j}b_{Jj}} - 1\right) \\ \vdots & \ddots & \vdots \\ x_1^0 x_J^0 e^{\mu_1 t} \left(e^{|t|\sum_{j=1}^{J} b_{1j}b_{Jj}} - 1\right) & \cdots & \left(x_J^0 e^{\mu_J t}\right)^2 \left(e^{|t|\sum_{j=1}^{J} b_{Jj}^2} - 1\right) \end{bmatrix}, \tag{13}$$

As we discussed above, in our first approach we just considered two equations ($n = 2$) to model the evolution of copper prices: one for the copper price in US dollars, and other for the exchange rate between US dollars and UF. In addition, we considered the copper price and the inflation (exchange rate US dollar - UF) as factors, i.e., $J = 2$, which means that the system is closed and no other factors affect the evolution on the indexes considered.

As we have seen before, $\mu_i, b_{ij}, i, j \in \{p, r\}$ are parameters that need to be estimated. For doing this, we use the recent historical information (for the last year)

$$x_i^{-12}, x_i^{-11}, \dots, x_i^0, \quad i \in \{p, r\},$$

and market information (futures) for copper prices,

$$x_p^1, \dots, x_p^{12}.$$

Having this data, we estimated the parameters using the methods described in Appendix B and we obtained the cumulative probability and the probability density functions of the multivariate process for the next period (month). Figure 3 shows an example of the curves obtained considering the data of 10/2011.

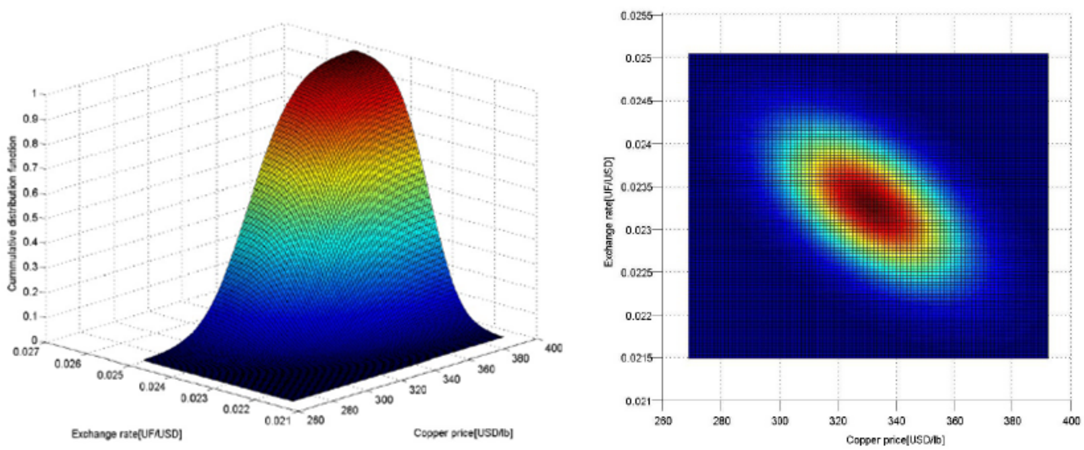

Figure 3: CDF and PDF of the transient process estimated for 10/2011

6.3 Stationary process

In order to compare the performance of this approach in the short term against the models presented above we first estimate our stationary model considering recent historical data and then we estimate the expected value of our process for the next twelve months.

Then, we proceed to estimate our stationary model for the long term considering all the historical information available. Nevertheless, the are some parameters - the number of periods to predict and the amount of data to be considered - that need to be defined before doing the estimations. For this purpose, we first estimate the one dimensional model for the spot copper price considering a variable amount of historical data, in order to determine how much data is better to calibrate the model. Finally, we show the results obtained for our first approach in the multivariate stationary process.

Comparing the transient and the stationary processes To compare the transient and the stationary processes we proceed to estimate the parameters associated to the stationary process following the methodology used for the epi-splines, and then we calculated the expected value of the process for the next t months using the expression

$$\mathbb{E}\left[x^t\right] = \upsilon\left(1 - e^{-\mu t}\right) + x^0 e^{-\mu t}.$$

In Table 3 we show the errors obtained with this approach (out of sample) using a variable amount of historical data.

	Out-of-sample				
	$k = 12$	$k = 15$	$k = 18$	$k = 21$	$k = 24$
MAD	1.115	0.881	0.825	0.984	0.798
MAPE	1.127	0.560	0.506	0.854	0.630
MSE	14.181	1.662	1.375	5.259	1.027
WAE	0.610	0.543	0.463	0.5996	0.648
MAE	65.166	8.764	8.913	31.585	4.193
RMSE	3.765	1.289	1.172	2.293	1.013

Table 3: Errors out-of-sample of the stationary process obtained considering different amount of historical data for the short term.

Comparing these results with the ones in Table 2 we can check that our transient approach outperform the stationary model in the short term. From Table 3 we can also see that as we include more historical data in the parameter estimation we obtain fewer errors, so the amount of historical data is an important issue to check in this model.

Amount of data to calibrate the stationary process An important issue of the stationary process is how much historical information we have to consider to estimate the parameters. Table 4 shows the MAPE obtained by estimating the parameters with a varying amount of historical data, k, and considering a variable amount of years to predict.

		Amount of data				
		$k = 12$	$k = 24$	$k = 48$	$k = 60$	$k = 120$
Periods	$N = 12$	0.165	0.164	0.175	0.182	0.200
	$N = 24$	0.248	0.246	0.266	0.278	0.296
	$N = 48$	0.345	0.367	0.403	0.410	0.403
	$N = 60$	0.402	0.426	0.462	0.466	0.460
	$N = 120$	0.533	0.510	0.493	0.480	0.459

Table 4: MAPE out-of-sample of the stationary process obtained considering different amount of historical data and a variable number of periods to predict.

As we expected, when we increase the number of periods to be predicted the error also increase. However, the amount of historical data to be considered is not clear. We can see that when we want to forecast fewer amount of periods, considering a short amount of data allow us to obtain better results. In contrast, when we want to predict a larger number of periods we need more historical data in order to obtain good results. Then, a possible rule could be to consider an amount of data similar to the number of periods we want to predict.

6.4 Multivariate stationary process

From section 2 we know that the solution of our multivariate stationary process is given by: for each $j \in \{1, \ldots, J\}$

$$x_i^t = \upsilon_i \left(1 - e^{-\mu_i t}\right) + x_i^0 \exp\left[-\left(\mu_i + \frac{1}{2}\sum_{j=1}^{J} b_{ij}^2\right) t + \sum_{j=1}^{J} b_{ij} w_j^t\right]$$

As $w_j^t,\ j \in \{1, \ldots, J\}$ are standard and independent Wiener processes, we know that the term

$$z_i^t = -\left(\mu_i + \frac{1}{2}\sum_{j=1}^{J} b_{ij}^2\right) t + \sum_{j=1}^{J} b_{ij} w_j^t,$$

is normally distributed with mean $-\left(\mu_i + \frac{1}{2}\sum_{j=1}^{J} b_{ij}^2\right) t$ and variance $\sum_{j=1}^{J} b_{ij}^2 t$. Then, the variable $y_i^t = x_i^0 e^{z_i^t}$ is log-normally distributed and the joint process $Y^t = [y_1^t, \ldots, y_J^t]$ is a multivariate log-gaussian distributed with parameters ρ^t, Σ_Y^t

$$\rho^t = \begin{bmatrix} -\left(\mu_1 + \frac{1}{2}\sum_{j=1}^{J} b_{ij}^2\right) t \\ \vdots \\ -\left(\mu_J + \frac{1}{2}\sum_{j=1}^{J} b_{Jj}^2\right) t \end{bmatrix}$$

$$\Sigma_Y^t = \begin{bmatrix} \sum_{j=1}^{J} b_{1j}^2 & \cdots & \sum_{j=1}^{J} b_{1j} b_{Jj} \\ \vdots & \ddots & \vdots \\ \sum_{j=1}^{J} b_{1j} b_{Jj} & \cdots & \sum_{j=1}^{J} b_{Jj}^2 \end{bmatrix},$$

and the original process x^t is shifted multivariate log-normally distributed with mean and variance

$$m^t = \begin{bmatrix} \upsilon_1 \left(1 - e^{-\mu_1 t}\right) + x_1^0 e^{-\mu_1} \\ \vdots \\ \upsilon_J \left(1 - e^{-\mu_J t}\right) + x_J^0 e^{-\mu_J} \end{bmatrix}$$

$$\Sigma^t = \begin{bmatrix} \left(x_1^0 e^{-\mu_1 t}\right)^2 \left(e^{|t| \sum_{j=1}^{J} b_{1j}^2} - 1\right) & \cdots & x_1^0 x_J^0 e^{-(\mu_1 + \mu_J)t} \left(e^{|t| \sum_{j=1}^{J} b_{1j} b_{Jj}} - 1\right) \\ \vdots & \ddots & \vdots \\ x_1^0 x_J^0 e^{-\mu_1 t} \left(e^{|t| \sum_{j=1}^{J} b_{1j} b_{Jj}} - 1\right) & \cdots & \left(x_J^0 e^{-\mu_J t}\right)^2 \left(e^{|t| \sum_{j=1}^{J} b_{Jj}^2} - 1\right) \end{bmatrix},$$

It can be proved that when

$$\frac{s_{ij}}{m_i m_j} = \frac{x_i^0 x_j^0 e^{-(\mu_i+\mu_j)t}\left(e^{|t|\sum_{j=1}^{J} b_{1j}b_{Jj}}-1\right)}{\upsilon_i\left(1-e^{-\mu_i t}\right)+x_i^0 e^{-\mu_i}\upsilon_j\left(1-e^{-\mu_j t}\right)+x_j^0 e^{-\mu_j}}, \quad \forall i,j\in\{1,\dots,J\}$$

is sufficiently small, the process $x^t = \{x_1^t, \dots, x_J^t\}$ can be closely approximated to a multivariate gaussian distribution with mean m^t and covariance matrix Σ^t.

As for the transient process, in our first approach we considered a closed system with 2 factors: spot copper price in US dollars and the exchange rate between US dollars and UF. Then, using the historical data we can obtain the cumulative density and the probability density functions. In Figure 4 we show the results obtained considering the data from 01/1984 to 10/2011.

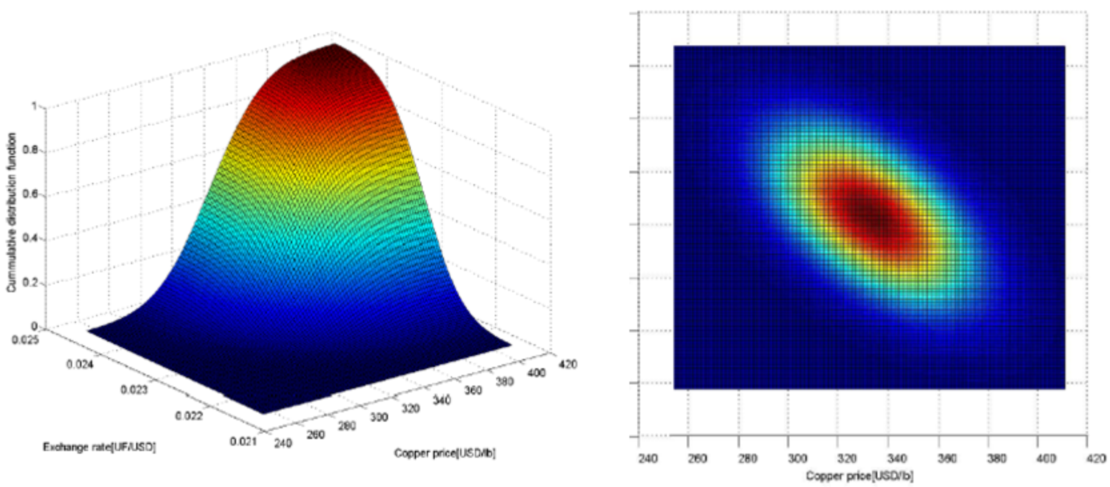

Figure 4: CDF and PDF of the stationary process estimated for 10/2011

Finally, as it's explained in section 4 the transient and the stationary processes can be blend to estimate copper prices in the mid term. In our case, as an example we proceeded to blend the process estimated in sections 6.2 and 6.4, considering $\gamma = 2$, $T_1 = 1$ and $T_2 = 4$. Figure 5 shows the results obtained.

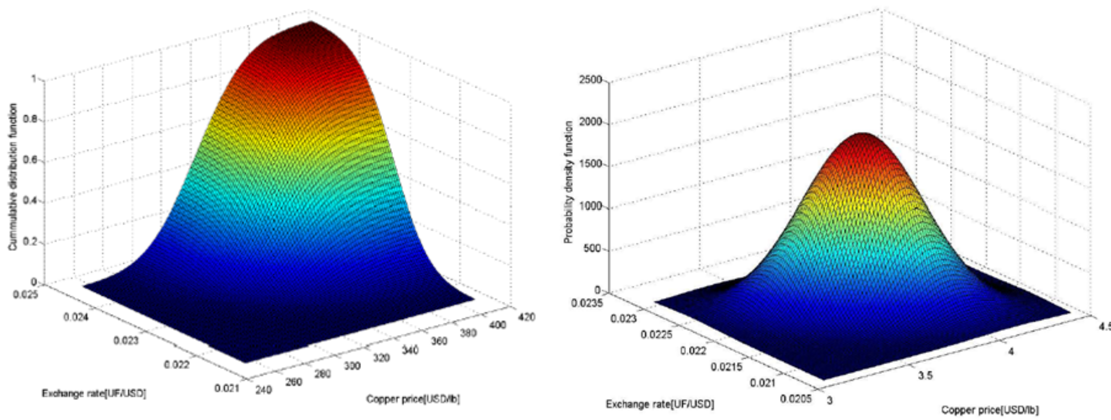

Figure 5: CDF and PDF of the blended process estimated for 10/2011

Appendix

A Approximation of the solution of the stationary process

The solution of the stationary process is given by,

$$x_i^t = x_i^0 \exp\left[-\left(\mu_i + \frac{1}{2}\sum_{j=1}^{J} b_{ij}^2\right)(t - t_0) + \sum_{j=1}^{J} b_{ij}\left(w_j^t - w_j^{t_0}\right)\right] + \mu_i \upsilon_i \int_0^t e^{r_i(t,s)ds}$$

where $r_i(t,s) = -\left[\mu_i + \frac{1}{2}\sum_{j=1}^{J} b_{ij}^2\right](t-s) + \sum_{j=1}^{J} b_ij\left(w_j^t - w_j^s\right)$. We are going to approximate this solution replacing the term $\mu_i \upsilon_i \int_0^t e^{r_i(t,s)ds}$ by its expectation. Then,

$$\begin{aligned}
\mathbb{E}\left(\mu_i \upsilon_i \int_0^t e^{r_i(t,s)ds}\right) &= \mu_i \upsilon_i \int_0^t e^{-\left(\mu_i + \frac{1}{2}\sum_{j=1}^{J} b_{ij}^2\right)(t-s)} \mathbb{E}\left(\exp\left[\sum_{j=1}^{J} b_{ij}\left(w_j^t - w_j^s\right)\right]\right) ds \\
&= \mu_i \upsilon_i \int_0^t e^{-\left(\mu_i + \frac{1}{2}\sum_{j=1}^{J} b_{ij}^2\right)(t-s)} \mathbb{E}\left(\prod_{j=1}^{J} \exp\left[b_{ij}\left(w_j^t - w_j^s\right)\right]\right) ds \\
&= \mu_i \upsilon_i \int_0^t e^{-\left(\mu_i + \frac{1}{2}\sum_{j=1}^{J} b_{ij}^2\right)(t-s)} \prod_{j=1}^{J} \mathbb{E}\left(\exp\left[b_{ij}\left(w_j^t - w_j^s\right)\right]\right) ds
\end{aligned}$$

But noting that $\left(w_j^t - w_j^s\right)$ is a gaussian process with mean 0 and variance $(t-s)$ we know that,

$$
\begin{aligned}
&= \mu_i \upsilon_i \int_0^t e^{-\left(\mu_i + \frac{1}{2}\sum_{j=1}^J b_{ij}^2\right)(t-s)} \prod_{j=1}^J \exp\left[\frac{1}{2} b_{ij}^2 (t-s)\right] ds \\
&= \mu_i \upsilon_i \int_0^t e^{-\left(\mu_i + \frac{1}{2}\sum_{j=1}^J b_{ij}^2\right)(t-s)} \exp\left[\frac{1}{2} \sum_{j=1}^J b_{ij}^2 (t-s)\right] ds \\
&= \mu_i \upsilon_i \int_0^t e^{-(\mu_i)(t-s)} ds \\
&= \mu_i \upsilon_i e^{-\mu_i t} \int_0^t e^{\mu_i s} ds \\
&= \upsilon_i e^{-\mu_i t} \left(e^{\mu_i t} - 1\right) \\
&= \upsilon_i \left(1 - e^{-\mu_i t}\right)
\end{aligned}
$$

Finally, we can approximate the solution of the stationary process to,

$$
x_i^t = \upsilon_i \left(1 - e^{-\mu_i t}\right) + x_i^0 \exp\left[-\left(\mu_i + \frac{1}{2}\sum_{j=1}^J b_{ij}^2\right)(t - t_0) + \sum_{j=1}^J b_{ij}\left(w_j^t - w_j^{t_0}\right)\right]
$$

B Parameter estimation of the transient process

The SDE (6) can be re-written as,

$$
dS_i^t = \left(\mu_i - \frac{1}{2}\sum_{j=1}^J b_{ij}^2\right) dt + \sum_{j=1}^J b_{ij} dw_j^t
$$

where $S_i^t = \ln x_i^t$. Then, we know that the dS_i^t follows a gaussian distribution with the following properties (see Dixit [5] and Hull [10]):

$$
\begin{aligned}
\mathbb{E}\left[dS_i^t\right] &= \left(\mu_i - \frac{1}{2}\sum_{j=1}^J b_{ij}^2\right) dt \\
\mathbb{V}\left[dS_i^t\right] &= \sum_{j=1}^J b_{ij}^2 dt \\
cov\left[dS_i^t, dS_k^t\right] &= \sum_{j=1}^J b_{ij} b_{kj} dt
\end{aligned}
$$

Considering the discrete case we have,

$$\mathbb{E}\left[S_i^{t+\Delta t} - S_i^t\right] = \left(\mu_i - \frac{1}{2}\sum_{j=1}^{J} b_{ij}^2\right)\Delta t$$

$$\mathbb{V}\left[S_i^{t+\Delta t} - S_i^t\right] = \sum_{j=1}^{J} b_{ij}^2 \Delta t$$

$$cov\left[S_i^{t+\Delta t} - S_i^t, S_k^{t+\Delta t} - S_k^t\right] = \sum_{j=1}^{J} b_{ij} b_{kj} \Delta t$$

Then, the easiest method to estimate the parameters of this model is using the fact that $S_i^t = \ln x_i^t$ and historical prices in such a way that,

$$\mu_i = \mathbb{E}\left[\frac{1}{\Delta t}\ln\left(\frac{x_i^{t+\Delta t}}{x_i^t}\right)\right] + \frac{1}{2}\sum_{j=1}^{J} b_{ij}^2$$

$$\sum_{j=1}^{J} b_{ij}^2 = \mathbb{V}\left[\frac{1}{\sqrt{\Delta t}}\ln\left(\frac{x_i^{t+\Delta t}}{x_i^t}\right)\right]$$

$$\sum_{j=1}^{J} b_{ij} b_{kj} = cov\left[\frac{1}{\sqrt{\Delta t}}\ln\left(\frac{x_i^{t+\Delta t}}{x_i^t}\right), \frac{1}{\sqrt{\Delta t}}\ln\left(\frac{x_k^{t+\Delta t}}{x_k^t}\right)\right]$$

Another way to estimate these parameters is recalling that, for $i \in \{p, r\}$

$$\mathbb{E}[x_i^t] = x_i^0 e^{\mu_i t},$$

where μ_i is the drift and x_i^0 the initial value of index i.

Then, we estimate $\mu_i,\ i \in \{p, r\}$ and the initial state denoted by $\theta_i,\ i \in \{p, r\}$. Estimating the initial state is very important because in most applications is used the actual spot price as initial condition, forgetting that this also has noise as it is a random variable.

Finally, assuming that the errors in the observations (x_i^t) come from white noise around the drift term $\mu_i t$, one has

$$x_i^t = \theta_i e^{\mu_i t + \varepsilon_i^t}, \quad t \in T$$

The main idea of this approach is to minimize the error associated to the estimation. For doing so, we are going to minimize $\sum_{t \in T} |\varepsilon_i^t|^2$, i.e.,

$$\left(\hat{\theta}_i, \hat{\mu}_i\right) \in \text{argmin}\,(\theta_i, \mu_i) \sum_t \left|\mu_i t - \ln\left(\frac{x_i^t}{\theta_i}\right)\right|^2$$

Differentiating with respect to θ_i and μ_i we get,

$$\frac{d\upsilon_i}{d\theta_i} = 2\sum_t \left(\ln\left(\frac{x_i^t}{\theta_i}\right) - \mu_i t \right) \frac{1}{\theta_i}$$
$$\frac{d\upsilon_i}{d\mu_i} = 2\sum_t \left(\ln\left(\frac{x_i^t}{\theta_i}\right) - \mu_i t \right) t$$

Setting these derivatives equal to 0 we obtain,

$$\frac{d\upsilon_i}{d\theta_i} = 0 \Rightarrow \mu_i = \frac{\sum_{t\in T} \ln\left(\frac{x_i^t}{\theta_i}\right) t}{\sum_{t\in T} t^2}, \qquad \frac{d\upsilon_i}{d\theta_i} = 0 \Rightarrow \mu_i = \frac{\sum_{t\in T} \ln\left(\frac{x_i^t}{\theta_i}\right)}{\sum_{t\in T} t}.$$

Solving the system and denoting $a = \sum_{t\in T} t$ and $b = \sum_{t\in T} t^2$ we obtain, for $i \in \{p, r\}$

$$\hat{\theta}_i = \exp\left(\left(a^2 - b\eta\right)^{-1} \sum_t (at - b)\ln(x_i^t) \right)$$
$$\hat{\mu}_i = b^{-1} \left(\sum_{t\in T} t \ln\left(\frac{x_i^t}{\hat{\theta}_i}\right) \right)$$

where η is the number of observation, i.e., if we consider just the historical information of the last 12 months $\eta = 13$.

Covariance matrix To estimate the covariance matrix with this method we know that,

$$cov\{x_i^t, x_j^t\} = x_i^0 x_j^0 e^{(\mu_i+\mu_j)t} \left[\exp\left(t \sum_{k=1}^{J} b_{ik} b_{jk} \right) - 1 \right]$$

Assuming that observations are corrupted by a white noise ε_{kl}^t that affects $|t|\sum_{j\in\{p,r\}} b_{kj} b_{lj}$ and recalling that $\hat{x}_k^t = \hat{\theta}_k e^{\hat{\mu}_k |t|}$, i.e., for $t = -12, \ldots, 0$,

$$\left(x_k^t - \hat{x}_k^t\right)\left(x_l^t - \hat{x}_l^t\right) = \hat{\theta}_k \hat{\theta}_l e^{(\mu_k+\mu_l)t} \left[\exp\left(|t| \sum_{j\in\{p,r\}} b_{kj} b_{lj} + \varepsilon_{kl}^t \right) \right]$$

Then, seeking estimates that minimize $\sum_t |\varepsilon_{kl}^t|^2$, one obtains the estimate $\hat{\beta}_{kl}$ for $\sum_{j\in\{p,r\}} b_{kj} b_{lj}$:

$$\hat{\beta}_{kl} = \frac{\sum_t |t| \ln\left[1 + \frac{\left(x_k^t - \hat{x}_k^t\right)\left(x_l^t - \hat{x}_l^t\right)}{\hat{x}_k^t \hat{x}_l^t} \right]}{\sum_t t^2}$$

Thus, the estimate for $cov\left(x_p^t, x_r^t\right)$ is,

$$\hat{\sigma}_{pr}^t = \hat{\theta}_p \hat{\theta}_r e^{(\mu_p+\mu_r)|t|} \left(e^{\hat{\beta}_{pr}|t|} - 1\right)$$

and the variance, for $k \in \{p, r\}$,

$$\hat{\sigma}_{kk}^t = \hat{\theta}_k^2 \left(e^{\hat{\beta}_{kk}|t|} - 1\right)$$

References

[1] Henrik Andersson. Are commodity prices mean reverting? *Applied Financial Economics*, 17(10):769–783, 2007.

[2] M.T Barlow. A diffusion model for electricity prices. *Mathematical Finance*, 12(4):287–298, October 2002.

[3] Hendrick Bessembinder, Jay F. Coughenour, Paul J. Seguin, and Margaret Smoller. Mean reversion in equilibrium asset prices: Evidence from the futures term structure. *The Journal of Finance*, 50(1):pp. 361–375, 1995.

[4] Yu-Chin Chen, Kenneth Rogoff, and Barbara Rossi. Can exchange rates forecast commodity prices? *The Quarterly Journal of Economics*, 125(3):1145–1194, August 2010.

[5] Avinash K. Dixit. *The art of smooth pasting*. London School of Economics and Political Science, 1992.

[6] Avinash K. Dixit and Robert S. Pindyck. *Investment under Uncertainty*. Princeton University Press, 1994.

[7] Eduardo Engel and Rodrigo Valdes. Prediciendo el precio del cobre: Mas alla del camino aleatorio? March 2001.

[8] Rajna Gibson and Eduardo S. Schwartz. Stochastic convenience yield and the pricing of oil contingent claims. *The Journal of Finance*, 45(3):959–976, July 1990.

[9] Jimmy E. Hilliard and Jorge A. Reis. Valuation of commodity futures and options under stochastic convenience yields, interest rates and jump diffusions in the spot. *The Journal of Financial and Quantitative Analysis*, 33(1):61–86, 1998.

[10] John C. Hull. *Options, Futures, and Other Derivatives*. Prentice Hall, 7 edition, May 2008.

[11] Patricio Meller. *Dilemas y debates en torno al cobre*, chapter 7, Tendencias y volatilidad del precio del cobre. Dolmen, 2002.

[12] Kristian R. Miltersen and Eduardo S. Schwartz. Pricing of options on commodity futures with stochastic term structures of convenience yields and interest rates. *The Journal of Financial and Quantitative Analysis*, 33(1):33–59, March 1998.

[13] Martin J. Nielsen and Eduardo S. Schwartz. Theory of storage and the pricing of commodity claims. *Review of Derivates Research*, 7:5–24, 2004.

[14] Dragana Pilipovic. *Energy Risk: Valuing and Managing Energy Derivatives.* McGraw-Hill, 2007.

[15] Diana Ribeiro. *Models for the price of a storable commodity.* PhD thesis, University of Warwick, 2004.

[16] J. Royset and R. Wets. Epi-splines and exponential epi-splines: Pliable approximation tools. *SIAM Review (submitted)*, 2013.

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

[18] Eduardo S. Schwartz and James E. Smith. Short term variations and long term dynamics in commodity prices. *Management Science*, 46(7):893–911, July 2000.

[19] George E. Uhlenbeck and Leonard Ornstein. On the theory of brownian motion. *Physical Review*, 36:823–841, September 1930.

[20] Roger J-B. Wets and Stehpen W. Bianchi. *Term and volatility structure, Handbook of Assets and Liability Management.* Elsevier, 2006.

[21] Roger J-B Wets, Stehpen W. Bianchi, and Liming Yang. Serious zero curves. 2002.

[22] J.C. Williams and B. D. Wright. *Storage and Commodity Markets.* Cambridge University Press, 1991.