Math Finan Econ (2015) 9:247–270
DOI 10.1007/s11579-014-0140-2

Modeling and estimating commodity prices: copper prices

Roger J.-B. Wets · Ignacio Rios

Received: 9 June 2013 / Accepted: 1 December 2014 / Published online: 23 January 2015

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 short term and long term regimes. In particular, this allows us to come up with an explicit drift term for the short-term process whereas the long-term process is primarily driftless due to inherent high volatility of commodity prices excluding an almost negligible mean reversion term. Not unexpectedly, the information used to build the short-term 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 consider a multi-dimensional system for which one can generate explicit solutions.

Keywords Commodity prices · Epi-splines · Short and long term · Best fit · Scenario tree

JEL Classification C53

1 Conceptual framework

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,

---

Electronic supplementary material The online version of this article (doi:10.1007/s11579-014-0140-2) contains supplementary material, which is available to authorized users.

R. J.-B. Wets (✉)
University of California, Davis, CA 95616, USA
e-mail: rjbwets@ucdavis.edu

I. Rios
University of Chile, Santiago, Chile
e-mail: irios@ing.uchile.cl

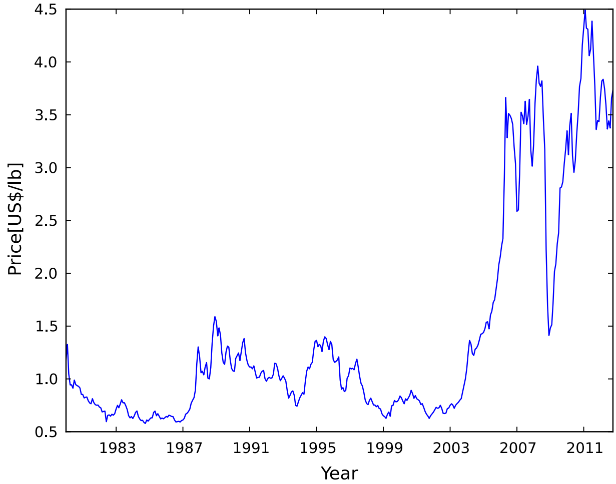


Fig. 1 Historical copper prices in US$ from 1980 to 2011

aimed at finding 'optimal' extraction and/or storage strategies, that are bound to involve these prices as parameters. Although our overall approach should, at least conceptually, be applicable to a wide range of commodities, in this article we 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 (Fig. 1), for example.

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 short- and long-term processes. Splitting time in short and long terms, in the case of copper prices is consistent with the conclusions reached by Schwartz [18,19] and in line with the more data driven study of Ulloa in [22, Capitulo 7: Tendencia y volatilidad del precio del cobre] who concludes after applying unit root tests to subsets of data of different lengths that shocks only affect the short-term, because in the long term copper prices essentially revert to their long term mean price, with in the interim, a high volatility. 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 short-term component of the process. A complete description of the state-of-the market, i.e., involving existing and potential (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), LME (London) and SHME (Shanghai). However, to exploit this information, this market information (futures) must be converted into "projected" spot prices and how we proceed is explained in the section dealing with the short-term 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 long-term process whereas recent historical prices complemented with market information should allow us to identify a drift in the formulation of the short-term process*[1]. Our model also takes into account inflation

[1] The inclusion or not of a mean reversion term in the long-term process will be taken up in the section devoted to the long-term process.

which leads us to a multi-dimensional (slightly nonlinear) system for which we can generate explicit solutions.

Both the need come up with reliable valuation of copper-related instruments and obtaining optimal management decisions (mines, processors, etc.) give rise to the need to design a copper-price process. There are however significant differences in the main concerns. Operational decision makers are mostly concerned with having an exceptionally well-designed short-term process in combination with a reasonably well-thought out long-term process. The overall concerns of firm-strategy planners or traders are in some ways more comprehensive and want the capability of both short- and (mostly) long-term valuations of instruments or projects that they might be considering. For the operations oriented decision maker, a 'static' model will do: what is today's best description of the copper-price process that will be helpful as input in a stochastic optimization model to reaching (present) optimal extraction, storage, . . . , decisions. In the second one, a more 'dynamic' version will be more appropriate because it cannot content itself with the perceived view of the future market from today's viewpoint; it has to be understood that the short-term view of this process might need to be modified significantly, over even relatively short time spans, in response to an evolving environment.

Although for both short- and long-term models, we allow for an arbitrary finite number of factors, when dealing specifically with copper, we only consider two factors:

(1) In order to factor in the inflationary/deflationary component, we rely on copper spot prices quoted in UF(CH)—UF is a Chilean monetary unit adjusted for inflation.
(2) Since market information is quoted in US$, the present reserve currency, the second factor is the exchange rate between the UF and the US$.
(3) One might consider including other factors such as interest rates. In fact, to do so, our methodology makes this quite easy, since good interest rates models come with the same characteristics, namely, a 'split' between short- and long term models but with slightly different characteristics that would unnecessarily render the validation of the model more complex without, in view of our results, improving 'accuracy'.

A word about our 'practical' results. As far as the long-term is concerned, our results are comparable to the best results of Schwartz [18,19] and significantly better than those obtained via alternative models. For the short-term, there is essentially no competition; more about this when we analyze the results.

The remainder of this article is organized as follows. In § 2 and § 3, we introduce the guiding models for the short- and long-term processes. The long-term process, except for being multivariate, is analogous to other proposals found in the literature that we review briefly. On the other hand, the short-term component of our model departs significantly from standard approaches and allows us to obtain significantly better predictable behavior. We discuss and even propose a 'blending' of the short- and long-term processes, but don't provide definite guidelines how this should be worked out practically since they would, unavoidably, be data-availability driven.

2 The short-term process

As already mentioned earlier, our principal contributions lay (i) in recognizing that the short- and long-term processes have to be dealt with separately and (ii) in the less than standard estimation of the parameters of the short-term process. The underlying reason is that for the short-term process one is able to come up with drift terms that are crucial to setting up a 'robust' process, both when working with (short-term) valuations and in the design of discretized versions (scenario trees) of this process that would be appropriate as input in operational models. We cast our short-term model as a geometric brownian motion, precisely because this model allows us to capture the *drift components* exploiting both historical and market information; it also eludes the possibility of negative prices. The innovative features of our approach are mostly in the construction of this short-term 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 it implicitly incorporates all the information available: market expectations/beliefs, stocks, production costs and other factors that influence prices. In addition, we propose a model that could incorporate in the volatility component the role played by other indexes (variables) that may affect future commodity prices, such as inflation, productivity indexes, ... This leads us to a (stochastic) system described by

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

with initial conditions at time $t = 0$

$$x_i^0 \quad \text{for } i = 1, \ldots, n,$$

where each x_i corresponds to an ‘index’ of interest: commodity (copper) prices, inflation rate(s), foreign exchange rates, interest rates, etc. The drift of the process is captured by the coefficient μ_i whereas its diffusion coefficients b_{ij} model the volatility-interaction between these indexes as well as, potentially, some additional ones; w_j, $j = 1, \ldots, J$ are independent (standard) wiener processes. A 1-dimensional version of this process would read

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

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

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

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

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

It’s immediate that our x-process is the solution of a system of stochastic differential equations of the following type:

$$dx_i^t = \left(\mu_i dt + \sum_{j=1}^{J} b_{ij} dw_j^t\right) x_i^t, \quad x_i^0 = x_i^0, \quad i = 1, \ldots, n$$

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.

2.1 Exploiting market information

Usually the information available about a commodity, in our instance copper, is manifold: existing contracts, producers and consumers’ stocks, exploration activity, location of recent discoveries, economic predictions (future demand), and so on. In order to take such a wide range of information into account, one needs a dedicated research division to amalgamate it so that it can be included in a model. We suppose that the traders in this commodity, and other agents that might affect its prospective value, have actually taking all these factors into account when selling or buying futures[2] If we accept this as a premise, obtaining market information that can actually be exploited in our modeling requires transforming the information one can collect about futures' prices and *converting it into "projected" spot prices* for the next few months, say 9–12 months. We shall then rely on the recent (historical) prices in combination with these projected spot prices to build the 'short-term stochastic process'. This conversion relies mainly on the fact that we can deduce the projected spot rate curve from the futures prices by relying on the epi-spline technology [17,23,24].

In this instance, 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 [23]: 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 $0, t_1, \ldots, t_L$ are the end points of the N sub-intervals. The curve $z \in [0, T]$ is completely determined by the choice of

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

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 (projected) *spot curve* for the commodity prices by minimizing the deviations from the available data, i.e.,

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

where $\vec{s} = (s_1, \ldots, s_L)$ corresponds to the present net value of the assets being considered and $\vec{t} = (t_1, \ldots, t_L)$ are the dates when these assets generate 'cash-flow' via deliveries.

[2] How futures' prices are determined is not of immediate concern including the role played by any of the factors mentioned above; one could consult [5,8] for an analysis of how they might depend on stock levels and spot prices associated with contracts involving actual deliveries.

To model commodity prices, here copper, we actually derive the corresponding discount factor curve df which will be used to generate the ‘projected’ 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

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

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

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

usually in the case of futures only a single delivery takes place but our formulation also allows for instruments with layered deliveries. For $t \in \left(\delta(k-1), \delta k\right]$, $\delta = T/N$, $df \in \mathcal{C}^{1,pl}\left([0,T], N\right)$ can be expressed as:

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

and $v_0, x_1, \ldots, x_N$ are the parameters that need to be determined. 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,

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

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, cf. [24, Sect. 6.2].

2.2 ‘Projected’ spot rates, drift and initial conditions

Given the discount curve df, the (projected) *spot prices curve* is immediately available since

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

We have at our disposal recent historical prices, say for the last $9 \leq \tau \leq 12$ months, today’s price and the market prices for the next τ months. It’s this information that will be used, as explained in the next section to build the drift of the short-term process. Again, we rely on an epi-spline fit to these prices to determine the ’drift’ of the short-term process. Combining this prices-information (short term past and relatively short term market projected 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:

$$\begin{aligned}
&\text{find}\, u_0,\ v_0 \,\text{and}\, x_j,\ j = 1, \ldots, N \text{ such that} \\
&\|(sp(t) - u(t)), \quad t = -\tau, \ldots, \tau\|_p \text{ with} \\
&u(t) = u_0 + v_0 t + \delta \sum_{j=1}^{k-1} \left(t - t_j + 0.5\delta\right) x_j + \frac{\tau^2}{2} x_k, \quad \forall t \in [0, T]
\end{aligned}$$

where the interval $[-\tau, \tau]$, the time span we want to take into account, has been subdivided in a $2\tau/N$ mesh and $sp(t)$ for $t = -\tau, \ldots, 0$ are the observed historical prices and $sp(t)$ are the (calculated) market projected spot-prices for $t = 1, \ldots, \tau$. The drift of the short-term process will then be taken to be the optimal solution of this optimization program $u^*(t)$ for $t \geq 0$.

One noteworthy consequence of this approach is that the initial condition of our process will be $u^*(0)$ and not today observed price $sp(0)$. One justification for proceeding in this fashion is that one should view today's (observed) spot price as the 'actual' spot price perturbed by some noise; our empirical calculation confirm that choosing $u^*(0)$ as the initial point for the short-term process yields substantially better results.

We proceed similarly to obtain the drift component of the UF(CH) versus US\$ as well as any other index, e.g., interest rates, that might have been included in the model.

2.3 Volatility components

Since $\mathbb{V}[x_i^t] = \left(\mathbb{E}[x_i^t]\right)^2 \left(e^{|t| \sum_{j=1}^{n} b_{ij}^2} - 1\right)$, once the drift terms have been determined, the estimation of the volatility component (covariance) will rely on a (standard) least squares estimate, cf. Appendix 2.

3 The long-term process

Our long-term process is pretty much in line with what can be found in the literature for the 'overall' process. Since this is to a large extent familiar territory, we won't provide a too detailed analysis. 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 rather elaborate analyses.

Basic microeconomics theory tells us that high prices should lead to increased supply since because higher cost producers will enter the market and that, in turn, will push down prices, returning to the market to its '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 [21], is supported by many authors: [3] confirms 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; [18] compare three models of commodity prices that takes into account mean reversion, and there are many other authors that use mean reverting processes to model commodity prices.

On the other side, results show that in some cases mean reversion is, at best, very slow, and in others the unit root test fails to reject the random walk hypothesis. For example,

[7] 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 for this result is that the speed of reversion is very slow, so using 'recent' past data is difficult to statistically distinguish between a mean-reverting process and a random walk. 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 [9], 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 (±a couple of years) 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. Again this argument 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 stochastic process that is mean reverting determines the drift of the long-term process. We rely on a variant of geometric brownian motion with mean reversion which is also in tune with our choice of inflation free 'currency'.

Our model for the long term process: for $i = 1, \ldots, n$,

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

and with $t_0 = 0$,

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

i.e., also a (shifted) *log-gaussian* process; the 1-dimensional version reads,

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

The long-term system can be viewed as an approximate solution of the system of stochastic differential equations,

$$dy_i^t = \mu_i \left(\upsilon_i - y_i^t\right) dt + \left(\sum_{j=1}^{J} b_{ij} dw_j^t\right) y_i^t, \quad i = 1, \ldots, n, \tag{4}$$

$$y_i^{t_0} = y_i^0, \quad i = 1, \ldots, n \tag{5}$$

where y_i^0 is the present value of index i (is given), μ_i and b_{ij} are constants that need to be estimated, $y^t = \left(y_1^t, \ldots, y_n^t\right)$ 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 y_i^t reverts in the log term and μ_i is the 'speed' at which y_i^t reverts to υ_i; in view of our earlier discussion, the strategy is to include mean-reversion drifts at a very slow rate and, consequently, its influence is quite attenuated.

This model was proposed by [7], and also used in [15] to model oil prices[3]. The solution of this system is: for $i = 1, \ldots, n$,

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

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 substituted the term $\mu_i \upsilon_i \int_0^t e^{r_i(t,s)} ds$ by its expectation; calculated in Appendix 1. 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.

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

$$\mathbb{E}[y_i^t] = \upsilon_i + \left(y_i^0 - \upsilon_i\right) e^{-\mu_i t} \tag{6}$$

$$\text{cov}\{y_k^t, y_l^t\} = y_k^0 y_l^0 e^{-(\mu_k + \mu_l)t}\left(\exp\left[t \sum_{j=1}^{n} b_{kj} b_{lj}\right] - 1\right) \tag{7}$$

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

$$\mathbb{E}[y^t] = \upsilon + (y_0 - \upsilon)\, e^{-\mu t}, \quad \mathbb{V}[y^t] = \left(y_0^2 e^{-\mu t}\right)^2 \left(e^{\sigma^2 t} - 1\right).$$

4 Blending short- and long-term processes

There remains only to pass from the short- to the long-term process to end up with a 'global' process. How to do this is still very much an open question that we only deal with experimentally complemented by data analysis, and consequently, only in the context of copper prices and exchange rates. Moreover, there seems to be a consensus in the literature, see the next section, that the long-term model reflects pretty well the behavior starting somewhere between the third and fourth year and up to that point is rather unsatisfactory. This was corroborated by our own numerical experimentation which suggests that the 'reign' of the short-term process is relatively short, the market reverts rather rapidly to its natural state.

[3] 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,

$$dS_t = \alpha (L_t - S_t)\, dt + \sigma S_t dw_t$$
$$dL_t = \mu L_t dt + L_t yi dz_t$$

On the basis of these considerations, we posit that the overall process X is a blending of the short-term process x and the long-term process y:

$$X^t = \lambda_t x^t + (1 - \lambda_t)\, y^t, \tag{8}$$

with $X_{t_0} = X_0$ the initial state vector derived for the short-term 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., $X = x$, and eventually, for $t \geq 4$, $\lambda_t = 0$, which means that there is no longer any influence of the short-term behavior. In the case of copper prices (and exchange rates), 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, in principle, 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, for that matter. However, we relied on our own data analysis to set $t_1 = 1$, $t_2 = 4$ and $\gamma = 2$; one could also rely on experts' advice. Once the parameters t_1, t_2 and γ have been selected a blended process can be built; cf. the follow up sections for a specific example.

5 Brief overview of the literature

Although there is some overlap between the design of the long-term 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 an expression of the equilibrium price is derived from them. The basic equilibrium model is described in [25] and examples of this approach are featured in [2] and [16].

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 relative simplicity, and in absence of big changes in the market structure their predictive accuracy outperforms structural models [9]. 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 Schwartz [18] 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, [18] provides a variation of the two-factor [10] model with the spot price following a mean reverting process given by:

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

where δ is the convenience yield which is also assumed to be stochastic. In fact, its behaviour is modeled by the SDE:

$$d\delta = (\alpha - \delta)\, dt + \sigma_2 dz_2$$

and the correlation between both processes is incorporated with the condition $dz_1 \cdot dz_2 = \rho_1$. Finally, in his third model [18] introduces a three factor model that extends the previous one 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:

$$\begin{aligned} dS &= (r - \delta)\, Sdt + \sigma_1 Sdz_1 \\ d\delta &= (\alpha - \delta)\, dt + \sigma_2 dz_2 \\ dr &= a\,(m - r)\, dt + \sigma_3 dz_3 \\ dz_1 \cdot dz_2 = \rho_1, &\quad dz_2 \cdot dz_3 = \rho_2, \quad dz_1 \cdot dz_3 = \rho_3 \end{aligned}$$

Extensions of [18]'s two and three-factor models are presented by [11,13,14] and [16].

On the other hand, there are some multi-factor models that make an implicit distinction between short and long term. [15] 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:

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

A similar approach is followed by [20] 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:

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

From these factors one builds the process for the spot price, which is given by $\ln(S_t) = \chi_t + \xi_t$.

6 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 long-term process, because one of our goals is to test the performance of our method varying the forecast's horizon.

In particular, for our experiments with all the information available and the long-term process we proceeded to deflect prices by the US CPI, in order to avoid inflation effects. This wasn't done for the short-term 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 short-term process we used the first twelve LME copper future contracts. Then, for each month from 01/2000 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, here we consider just two factors to estimate the short- and long-term multivariate processes: the spot copper price and the exchange rate between US$ dollars and UF(CH).

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

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

For this purpose, we take into account monthly average data for the exchange rate, from 01/1984 to 10/2012. Those used for the exchange rate in our copper price model is based on the work of Chen et al. [4], where it is shown that "the Chilean exchange rate has strong predictive power for future copper prices." Furthermore, the inclusion of this factor allows us to incorporate a measure of the Chilean inflation, which might also be important in modeling copper prices. As proposed by Schwartz [18] in his three factor model, interest rates are included as a third factor which could also be easily be included to our model. However, for the sake of simplicity, we restrict ourselves to the two principal factors: copper prices and US$-UF(CH) exchange rates. The actual data used can be found on the journal web-page under Electronic Supplementary Material.

Unit root tests As discussed already earlier, there is no consensus if commodity prices exhibit mean reversion and copper prices are no exception. So, we proceeded to apply the best known unit root tests—Augmented Dickey–Fuller (ADF), Kwiatkowski-Phillips-Schmidt-Shin and Variance Ratio test—to the data we use in our experiments[4].

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 by the ADF test. However, from 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 stochastic processes considered do not follow a random walk. This result contradicts those obtained by other tests, and confirms the difficulty of determining the existence of unit roots in times series by using those tests.

7 Validation and experiments

7.1 Short-term drift estimation

Having the historical information and also market information converted in projected spot prices,

$$sp(t), \quad t = -12(9), \ldots, 0, \ldots, 12(9),$$

we can compute the drift term of the process. To accomplish this we rely on three methods:

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

1. The classical approach in the literature, which is
$$\begin{aligned} \hat{\mu} &= \mathbb{E}\left[\ln\left(\frac{sp(t+1)}{sp(t)}\right)\right] + \frac{1}{2}\hat{\sigma}^2 \\ \hat{\sigma} &= \mathbb{V}\left[\ln\left(\frac{sp(t+1)}{sp(t)}\right)\right] \end{aligned} \tag{9}$$
   Under this approach the spot price at time 0 is assumed to be the initial state of the system, i.e., $\theta = sp(0)$. This approach is referred to as "Classic" approach in the results.
2. A second approach is to estimate the initial state and the drift term by fitting a curve to the data. In particular, if we assume that spot prices can be approximated according to the expression,
$$sp(t) = \theta e^{\mu t + \varepsilon_t},$$
   where θ is the initial state, μ the drift term and ε_t an error, we can estimate the parameters by minimizing the sum of the errors, i.e.,
$$\left(\hat{\theta}, \hat{\mu}\right) \in \underset{(\theta,\mu)}{\text{argmin}} \sum_{t=-12}^{12} \left|\mu t - \ln\left(\frac{sp(t)}{\theta}\right)\right|. \tag{10}$$
   We refer to this method as "Novel" approach.
3. Another approach to estimate the drift term is to fit another curve (we relied on an epi-spline fit) aiming to minimize the deviations from the data, i.e., find a curve z such that
$$\|z(t) - sp(t),\ t = -12, \ldots, 12\|_p \tag{11}$$
   is minimized under certain norm p. Then, the initial state can be obtained from the value of this curve at time 0, i.e. $\hat{\theta} = z(0)$, and the drift term can be assumed to be time dependent and computed by,
$$\hat{\mu}_t = \frac{1}{t} \ln\left[\frac{z(t)}{z(0)}\right].$$
   In our case, we use epi-splines to fit this curve. Then, we refer as this method of estimating θ and μ as the "Epi" approach.

To measure the impact of including market information into the drift estimation and also the performance of our method against benchmark approaches in the literature we proceed as follows:

1. For each time t we proceed to estimate the parameters ($\hat{\mu}$ and $\hat{\theta}$) using only historical information, $\{t-12, \ldots, t-1\}$, and also combining this with market information converted into future spot prices, $\{t+1, \ldots, t+12\}$,
2. Compute a forecast for the next 12 months using the expression,
$$f_{(t+i)} = \hat{\theta} e^{\hat{\mu} t}$$
3. Compute the in-sample error between the forecast in the observed spot prices,
$$\epsilon_i = S_{t-i} - f_{t-i}, \quad i = 12, \ldots, 0$$
   and the out-of-sample error,
$$\epsilon_i = S_{t+i} - f_{t+i}, \quad i = 1, \ldots, 12$$

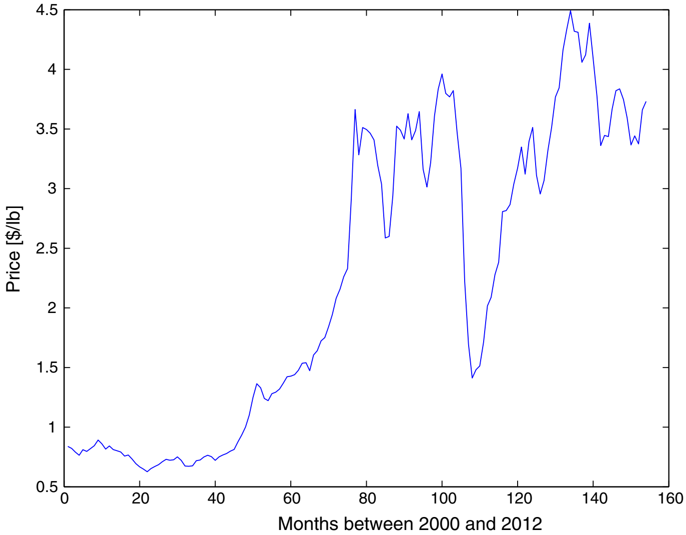


Fig. 2 Montly average LME spot prices from 2000 to 2012

To measure the errors we used the following metrics:

- 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)

Data To perform the experiments concerning drift estimation in the short term we use monthly average spot prices obtained form the LME between 2000 and 2012, in combination to future contracts for each month up to one year. However, looking at the data depicted in Fig. 2 we can see that there is an structural change in 2004 who produced a positive—permanent shock in prices. Then, to avoid problems in the estimation we proceed to run our experiments in two separated time intervals:

- From January 2000 to July 2004,
- From October 2005 to October 2011.

7.2 Analysis of the short term drift estimation

As indicated earlier, we implement three methods to estimate the drift and the initial state of our short-term model. In addition to this, we evaluate the impact of including market information in the estimation of these parameters. Figure 3 records the errors in and out of sample obtained for the period between 2000 and 2004, while Fig. 4 records those between 2005 and 2011. It's important to remark that we don't perform any estimates for 2012 in order to use that data to measure the errors of the forecasts made in 2011. In these plots, for each measure of errors we have three columns, one for each drift estimation method. In

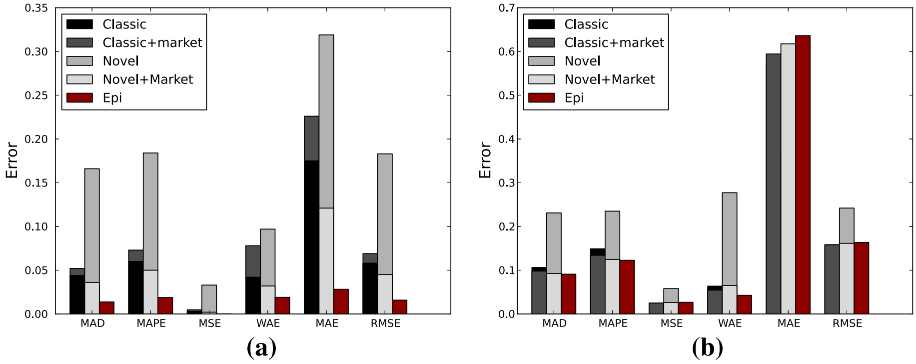


Fig. 3 In and out-of-sample errors obtain between 01/2000 and 07/2004. **a** In-sample. **b** Out-of-sample

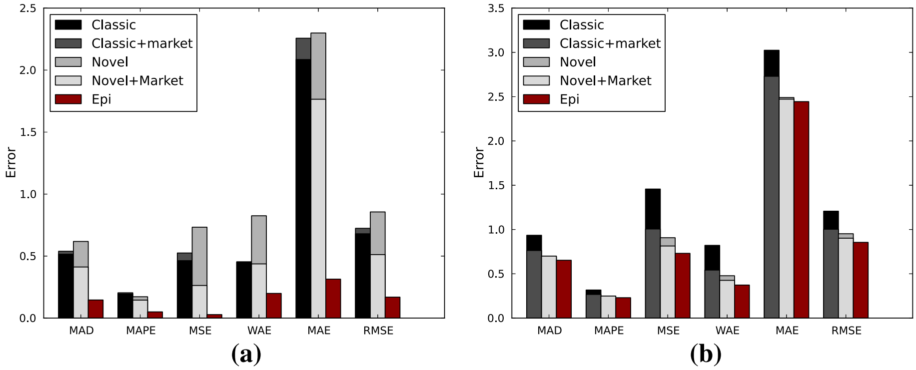


Fig. 4 In and out-of-sample errors obtain between 10/2005 and 10/2012. **a** In-sample. **b** Out-of-sample

addition to this, the first two columns of each error measure contain two colors: one referring to the estimation only considering historical data, and the other one reflecting the results of combining historical and market information.

From Figs. 3b and 4b we observe that the use of epi-splines outperforms the other methods, decreasing by around 6 % the different measures of error considered. In addition to this, we can see that the inclusion of market information reduce the errors out of sample regardless of the method employed to estimate the drift and the initial state of the process. However it's not clear what estimation method benefited the most by including market information: in the period 2000–2004 the "novel" approach presents savings of around 40 %, whereas in the next period, the "classic" method is the one with the highest savings.

Finally, we observe that estimating the initial state of the system invariably leads to better results than just considering the spot price at time 0 as the initial state of the system.

Another way to visualize the impact of including market information is by plotting the forecasts and confidence intervals obtained with and without market information. In Fig. 5 two examples of this are provided, considering the data from 01-2001 and 10-2010. These plots consider the forecasts for the next 12 months, and the confidence intervals were obtained with one standard deviation.

From these plots is clear that the inclusion of market information in the drift estimation contributes to improve our results.

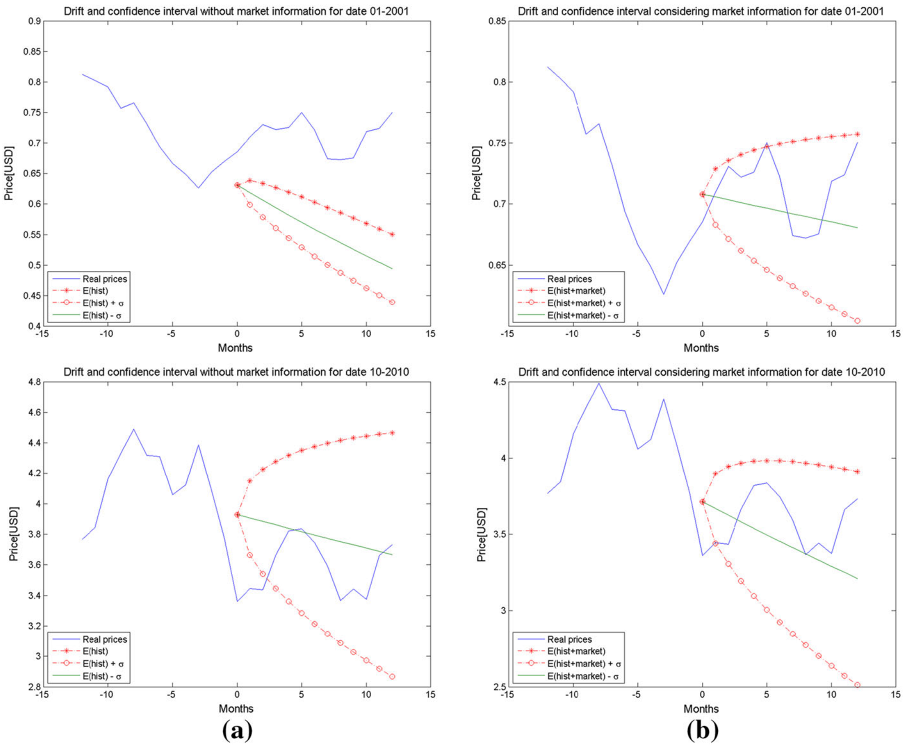


Fig. 5 Forecasts and confidence intervals obtained with and without market information. **a** Only historical information. **b** Historical and market information

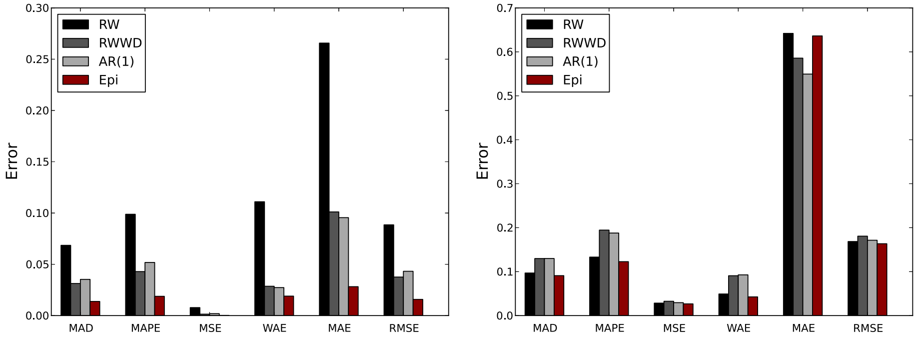
7.3 Benchmark comparison

In [9] a comparison between different methods to estimate short and middle term copper prices (from 1 to 5 years) is performed. Among their conclusions the authors state that in the short-term (less than one year) the best model to represent the stochasticity of copper prices is the random walk, whereas in longer periods the AR(1) process outperforms more sophisticated methods as the one, two and three factor models in [18], ARIMA's and non-linear ESTAR models, among others.

Short-term Therefore, in Figs. 6 and 7 we compare our results to those obtained via the random walk (RW), the random walk with drift (RWWD) and the first order autoregressive process (AR(1)) in the periods 2000–2004 and 2005–2012 respectively.

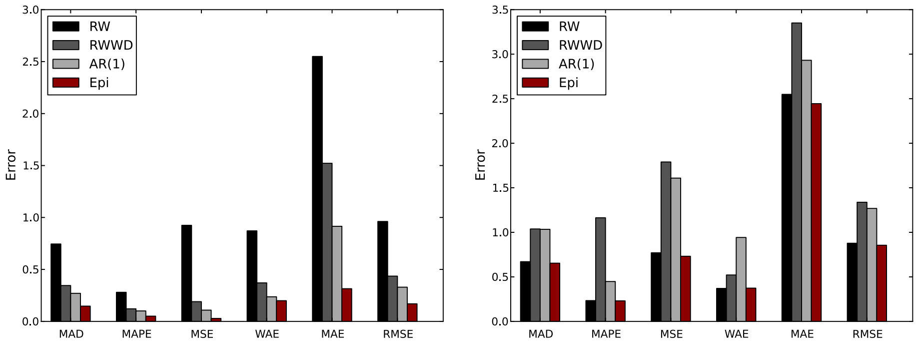
The results depicted in Figs. 6b and 7b are consistent with the conclusions in [9]: the random walk outperforms the random walk with drift and the AR(1) process. However, if we look at the results we can see that the use of epi-splines leads to a reduction between 2.5 and 4 % depending on the measure of error considered.

Long-term estiamtes Fig. 8 records the results out-of-sample obtained from comparing our long-term model against the AR(1) model, which is the best in the literature according to [9]. In particular, for each month between 1980 and 2011 we estimated the expected value of each process for the next N periods, with $N = 12, 24, 48, 60$ and 120 months.



(a) **(b)**

Fig. 6 In and out-of-sample errors of benchmark methods between 01/2000 and 07/2004. **a** In-sample. **b** Out-of-sample



(a) **(b)**

Fig. 7 In and out-of-sample errors of benchmark methods between 10/2005 and 10/2012. **a** In-sample. **b** Out-of-sample

From this plot is clear that the AR(1) process is outperformed by our long-term model, which leads to savings around 40 % in the MAPE.

7.4 Multivariate model

7.4.1 Short-term

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)

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

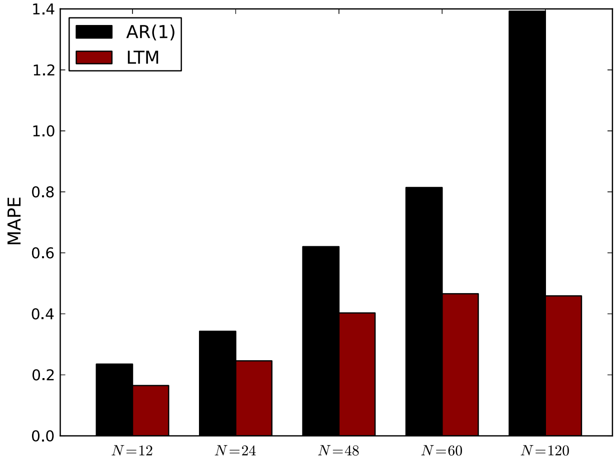


Fig. 8 Long-term comparison

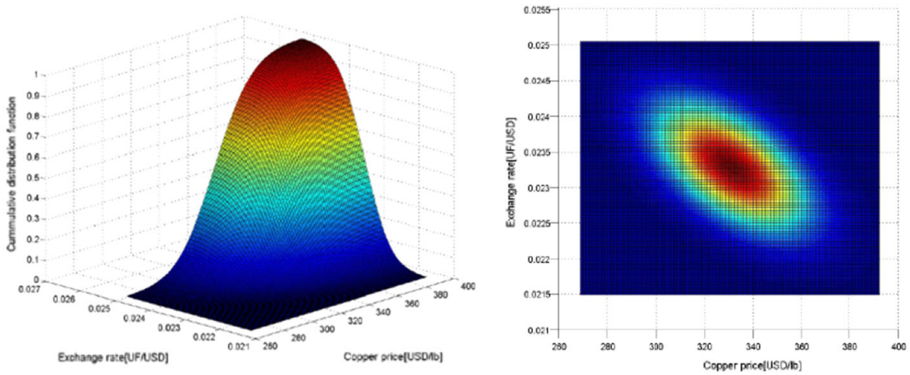


Fig. 9 CDF and PDF of the transient process estimated for 10/2011

and market information (futures) for copper prices,

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

Having this data, we estimated the parameters using the methods described in Appendix 2 and we obtained the cumulative probability and the probability density functions of the multivariate process for the next period (month). Figure 9 shows an example of the curves obtained considering the data of 10/2011.

7.4.2 Long-term

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 Fig. 10 we show the results obtained considering the data from 01/1984 to 10/2011.

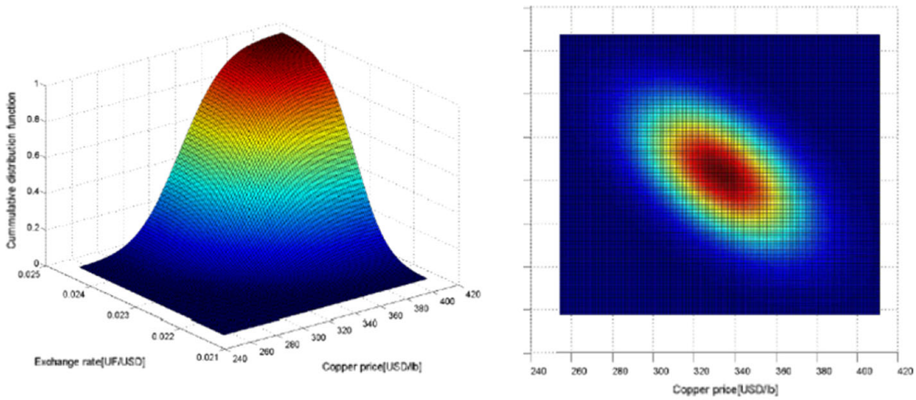


Fig. 10 CDF and PDF of the stationary process estimated for 10/2011

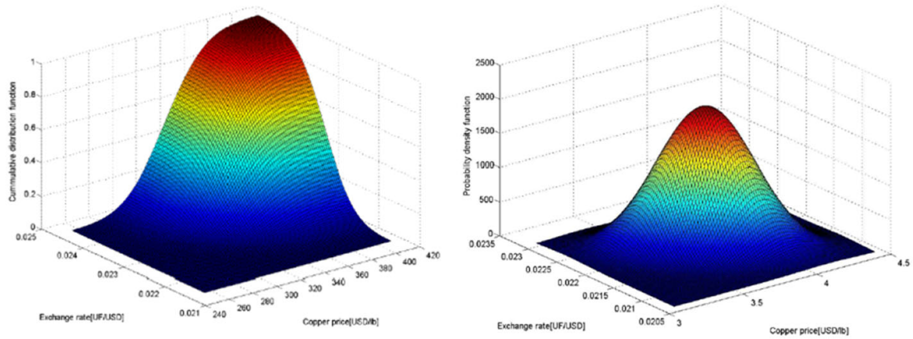


Fig. 11 CDF and PDF of the blended process estimated for 10/2011

Finally, as it's explained in Sect. 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 this section and in the previous one, considering $\gamma = 2$, $T_1 = 1$ and $T_2 = 4$. Figure 11 shows the results obtained.

8 Conclusions

We have described a stochastic process aimed at estimating copper prices. It models the blending of two sub-models: short- and long-term models. As indexes, our focus was on copper prices, the inflation adjusted prices UF(CH) and the corresponding exchange rate with the US$; note, however, that the guiding models would allow for the inclusion of many additional indexes if they were consider relevant.

We depart from previous work in several ways. The explicit differentiation of short and long term regimes, the inclusion of market information in the estimation of the short term drift and the capability to extend the model to incorporate any number of factors (indexes) are examples os this.

In addition , we have seen that the use of the rather novel epi-splines technology, to distill market information to generate 'projected' spot prices improves remarkably the drift estimation in the short-term, outperforming any, so far any suggested, models.

Because of our access to rather parsimonious data, we were unable to provide a satisfactory methodological answer to the blending of the short- and long-term processes, we relied on our experimental results as well as the observations recorded in earlier studies; however access to additional information, which is usually proprietary, should result in further improved estimates for the coefficients of the blending function (λ).

Acknowledgments 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.

9 Appendix

9.1 Appendix 1: Approximate solution of the long-term process

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

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

$$
\begin{aligned}
&= \mu_i \upsilon_i e^{-\mu_i t} \int_0^t e^{\mu_i s} ds \\
&= \upsilon_i e^{-\mu_i t} \left(e^{\mu_i t} - 1 \right) \\
&= \upsilon_i \left(1 - e^{-\mu_i t} \right)
\end{aligned}
$$

Finally, we can approximate the solution of the long-term process to,

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

9.2 Appendix 2: Parameter estimation of the short-term process

The SDE that governs the short term process can be 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 [6] and Hull [12]):

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

Considering the discrete case we have,

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

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,

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

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

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,

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

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 $\sum_{j\in\{p,r\}} b_{kj} b_{lj}$:

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

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. Andersson, H.: Are commodity prices mean reverting? Appl. Financ. Econ. **17**(10), 769–783 (2007)
2. Barlow, M.T.: A diffusion model for electricity prices. Math. Financ. **12**(4), 287–298 (2002)
3. Bessembinder, H., Coughenour, J.F., Seguin, P.J., Smoller, M.: Mean reversion in equilibrium asset prices: evidence from the futures term structure. J. Financ. **50**(1), 361–375 (1995)
4. Chen, Y.-C., Rogoff, K., Rossi, B.: Can exchange rates forecast commodity prices? Q. J. Econ. **125**(3), 1145–1194 (2010)
5. Cretì, A., Villeneuve, B.: Commodity storage with durable shocks: a simple markovian model. U. Paris-Dauphine, Technical report, CEREMADE (2013)
6. Dixit, A.K.: The Art of Smooth Pasting. London School of Economics and Political Science, London (1992)
7. Dixit, A.K., Robert, S.: Investment Under Uncertainty. Princeton University Press, Pindyck (1994)
8. Ekeland, I., Lautier, D., Villeneuve, B.: A simple equilibrium model for a commodity market with spo trades and future contracts. U. Paris-Dauphine, Technical report, CEREMADE (2013)
9. Engel, E., Valdes, R.: Prediciendo el precio del cobre: Mas alla del camino aleatorio? Universidad de Chile (2001)
10. Gibson, R., Schwartz, E.S.: Stochastic convenience yield and the pricing of oil contingent claims. J. Financ. **45**(3), 959–976 (1990)
11. Hilliard, J.E., Reis, J.A.: Valuation of commodity futures and options under stochastic convenience yields, interest rates and jump diffusions in the spot. J. Financ. Quant. Anal. **33**(1), 61–86 (1998)
12. Hull, J.: Options, Futures, and Other Derivatives, 7th edn. Prentice Hall, Boston (2008)
13. Miltersen, K.R., Schwartz, E.S.: Pricing of options on commodity futures with stochastic term structures of convenience yields and interest rates. J. Financ. Quant. Anal. **33**(1), 33–59 (1998)
14. Nielsen, M.J., Schwartz, E.S.: Theory of storage and the pricing of commodity claims. Rev. Deriv. Res. **7**, 5–24 (2004)
15. Pilipovic, D.: Energy Risk: Valuing and Managing Energy Derivatives. McGraw-Hill, New York (2007)
16. Ribeiro, D.: Models for the price of a storable commodity. Ph.D. thesis, University of Warwick (2004)
17. Royset, J., Wets, R.J.-B.: Epi-splines and exponential epi-splines: Pliable approximation tools. Math. Programm. (submitted) (2014)
18. Schwartz, E.S.: The stochastic behavior of commodity prices: Implications for valuation and hedging. J. Financ. **52**(3), 923–973 (1997)
19. Schwartz, E.S.: Valuing long-term commodity assets. Financ. Manag. **27**(1), 57–66 (1998)

20. Schwartz, E.S., Smith, J.E.: Short term variations and long term dynamics in commodity prices. Manag. Sci. **46**(7), 893–911 (2000)
21. Uhlenbeck, George E., Ornstein, Leonard: On the theory of brownian motion. Phys. Rev. **36**, 823–841 (1930)
22. Ulloa, A.: Tendencia y volatilidad del precio del cobre. In: Meller, P. (ed.) Dilemas y debates en torno al cobre, Chapter 7. Dolmen, Santiago (2002)
23. Wets, R.J.-B., Bianchi, S.W., Yang, L.: Serious Zero Curves. EpiRisk Research LLC, El Cerrito (2002)
24. Wets, R.J.-B., Bianchi, S.W.: Term and volatility structure, handbook of assets and liability management. In: Zenios, S.A., Ziemba, W. (eds.) Handbook of Aset and Liability Management, pp. 26–68. Elsevier, Amsterdam (2006)
25. Williams, J.C., Wright, B.D.: Storage and Commodity Markets. Cambridge University Press, Cambridge (1991)