# DOCUMENTOS DE TRABAJO

# BILTOKI

D.T. 2011.05

Stochastic Surface Models for Commodity Futures:
A 2D Kalman Filter Approach.

Javier Fernández-Macho

eman ta zabal zazu

Universidad del País Vasco Euskal Herriko Unibertsitatea

Facultad de Ciencias Económicas.
Avda. Lehendakari Aguirre, 83
48015 BILBAO.

**Documento de Trabajo BILTOKI DT2011.05**
Editado por el Departamento de Economía Aplicada III (Econometría y Estadística) de la Universidad del País Vasco.

ISSN: 1134-8984

# Stochastic Surface Models for Commodity Futures: A 2D Kalman Filter Approach

Javier Fernández-Macho[a,*]

[a]*Dpt. of Econometrics and Statistics, University of the Basque Country, Lehendakari Agirre 83, E48015 BILBAO, Spain*

**Abstract**

Efficient markets have been developed for hard commodities like energy products and metals where long term maturities exist. This paper assumes a two-dimensional stochastic process where futures prices evolve both along the time axis in the real world and along a second dimension of maturities obtained in a risk-free world, thus forming a $T \times n$ grid generated by a surface model similar to those used in areas like geophysics, image restoration, neural networks, remote sensing, etc. In contrast, however, our model is neither isotropic nor invariant in both directions. The paper obtains the surface state-space form of a doubly stochastic Inhomogeneous Geometric Brownian Motion model and derives the corresponding two-dimensional Kalman filter that is recursive in both directions simultaneously. The proposed methodology uses the entire time-maturity dynamics of the full stochastic process, including links from all available maturities per period, as an alternative to standard vector Kalman filtering along the one-dimensional time line which typically results in large states unfeasible to handle computationally. The technique is illustrated using a dataset of all daily observations of NYMEX coal futures contracts during a recent year.

*JEL classification codes:* G13, G17, C13, C51.

*Keywords:* Commodity prices, Two-dimensional Kalman filter, Spatial analysis, Energy markets, Futures markets, Stochastic dynamic model.

## 1. Introduction

Commodity markets have increasingly attracted the interest of producers, consumers, financial investors and academics. Efficient markets have been developed for certain so called hard commodities like energy products and metals where prices with long term maturities exist. The evolution of this type of markets

*Corresponding author: phone.: +34946013729, fax: +34946013754.

*Email address:* javier.fernandezmacho@ehu.es (Javier Fernández-Macho)

reflects the growing demand for energy products and the growth of investment in infrastructures (note, for instance, that prices of long maturities may provide relevant information for investment decisions of producers of such commodities). The characteristics of the commodities themselves, whose prices are usually related to inflation while showing little correlation to equities and bonds, has led to an increase in their demand by financial investors and this in turn has contributed to increased liquidity and depth in these markets.

However, commodity markets show great volatility, at least in the short and medium run, because an increase in demand cannot, in most cases, be compensated by an increase in a supply which usually takes some time to adjust. The resulting imbalance is then corrected by sudden changes in prices even though these should be related to production costs for a given future demand in the long term (which, in turn, could lead to the building of increasingly more expensive installations if the demand for renewable and non renewable commodities is growing).

The development of models to valuate commodity derivatives and related investments is thus of certain importance in economics given the greater needs of hedge and risk management by producers, firms, financial investors, traders and arbitrageurs. In this respect, the stochastic models used to analyze these markets have to take into account the characteristics of the goods to be valuated, such as local behavior, the existence of a mean reversion, risk premium or volatility, and in some cases even a convenience yield or seasonality. (More information on commodities and their markets in Fabozzi et al. 2008 and Geman 2005, while London 2007 deals with energy and power derivatives.)

It is not sufficient to select a model with theoretically reasonable characteristics for the corresponding commodity. In addition, it has to be correctly calibrated and for this futures markets give us a large amount of information that usually grows not only over time but also along the maturity dimension. On one hand, along the time axis, there are futures prices that evolve in the real world, whereas for a particular given day, on the other hand, the prices evolve along a second dimension of increasing maturities as the expected values obtained in a risk-free world once the market value of risk has been deducted from the drift. In other words, futures prices of a given financial asset or commodity form a surface generated by a two-dimensional stochastic process. Both the use of an inadequate model or calibration method could lead to serious consequences, which in the case of exposure to important risks would probably incur in considerable monetary losses.

Different calibration methods have been used that make use of a limited prices vector over time, such as minimum mean square error methods (Cortazar and Schwartz, 2003), nonlinear least squares (Lucia and Schwartz, 2002) and others. But undoubtedly one of the most relevant methods has been the use of the Kalman filter. Some of the studies that used this method are summarized in what follows.

Schwartz (1997) uses the Kalman filter to estimate three stochastic models of commodity prices. A first model of one factor, where it is assumed that the spot price logarithm follows a mean reversion process. The second is a Geometric Brownian Motion (GBM) model, extended with an additional stochastic factor, namely the convenience yield. The third model extends the second by including stochastic interest rates. In all these models there exists a non-observable spot price whose behavior is reflected in the transition equation, whereas the measurement equation is constructed with futures prices. It is with this non-observable spot price that certain values, such as volatility, are obtained. Estimation is carried out using a limited series of only five prices per week using New York Mercantile Exchange (NYMEX) data as well as ten prices per week using Enron oil data. Schwartz and Smith (2000) also estimate a two factor model *via* the Kalman filter, one of which reflects the short-term impact and whose expected value tends to disappear over time, whereas the second factor reflects the uncertainty in the long-term equilibrium value. In this case the state variables are the two long and short-term price components. The data used are the same as in Schwartz (1997). Manoliu and Tompaidis (2002) use the Kalman filter to estimate two models for natural gas. The first model contains a stochastic factor that follows an Ornstein-Uhlenbeck mean-reverting process with deterministic seasonality. The second model incorporates an additional stochastic component that follows a Brownian motion process. Daily price data are used, with 15 futures contracts per day. Cortazar and Naranjo (2006) use a Kalman filter with a variable number of observations which nonetheless falls within the standard time-varying Kalman filter specification. Similarly, Diebold et al. (2006) use a dynamic version of a three-factor Nelson and Siegel (1987) yield curve model where the time-varying latent factors are related to the level, slope, and curvature of the yield function. Cortazar et al. (2008) develop a multicommodity model for energy prices where common and specific factors exist. The model is used to fit data for West Texas Intermediate and Brent oil prices and, in a second example, for West Texas Intermediate and unleaded gasoline prices. The model is estimated using the Kalman filter with the aim of predicting the futures price of one commodity, which does not have medium and long-term prices, with the help of another commodity for which those maturities exist.

All these methods would face a problem in practice when trying to deal with the full amount of information represented by the complete futures stochastic surface. As mentioned above, not only there exists a real time diffusion dynamics both for spot prices as well as for futures prices with a given maturity, but there is also a second diffusion dynamics along the maturities dimension which is usually ignored when using a standard one-dimensional vector Kalman filter that is not capable of grasping the whole stochastic surface process.

This study aims to use all the available information in the stochastic surface generated by data on historical futures prices. It sets up from a stochastic spot price behavioral model, from which we derive the equation that determines, for a given spot price and parameter values, the futures prices of different maturities on a particular day. It is then possible to explicitly obtain the stochastic differential equations that govern the behavior of a particular future both along time as well as along a second dimension of maturities. This joint evolution of spot and futures prices along both dimensions will be reflected in the transition equation of the two-dimensional state-space form of our model. We note that this contrasts with previous works that use a standard vector representation of several maturities evolving along the one-dimensional time line. This typically results in large state vectors which are difficult to handle computationally to the point that empirical analysis may be rendered unfeasible (see, *inter alia* Woods and Radewan, 1977; Wang, 1998; Zou et al., 2004, and references therein). This is because the covariance matrices involved are of an order of magnitude that increases with the square of the number of maturities. On the other hand, the size of the state vector in the present two-dimensional setup does not depend on the number of maturities so that no magnitude problem arises.

In order to analyze the statistical properties of the proposed model, a two-dimensional Kalman filter (2DKF) will be implemented, which, to the best of our knowledge, has not been attempted before in this context.[1] The proposed 2DKF methodology has important advantages in comparison to the more traditional implementation of a standard vector Kalman filter since it allows full use of the information, both in terms of the whole available observation set as well as the complete surface diffusion dynamics, thus improving the accuracy of our model parameter estimates. In this respect our model not only can deal with as many maturities per time period as available, but it also takes into account all the links between them as implied by the full surface diffusion process.

The rest of the paper is organized as follows: Section 2 presents our model and the corresponding state-space form, Section 3 develops the two-dimensional Kalman filter, Sections 4 and 5 present the maximum-

[1] Fernández-Macho (2008) has recently presented a non-Kalman-filter solution to the estimation of surface models in the spectral domain that can be of interest in the analysis of massive datasets. However, in the present case one of the transition matrices is not invariant along one of the dimensions, which is difficult to handle in the frequency domain.

likelihood estimation procedure and predictive testing, whilst in Section 6 an application for the case of coal futures is carried out. Finally Section 7 lays down the main conclusions.

## 2. Doubly stochastic local IGBM model

In order to determine the mechanism with which the prices of futures contracts on certain commodities evolve along the two-dimensional time-maturity space, we start by considering a mean-reverting stochastic process for a commodity price such as the Inhomogeneous Geometric Brownian Motion model (IGBM):[2]

$$dS_t = k(S_m - S_t)dt + \sigma_\eta S_t dz_t, \tag{1}$$

where $S_t$ is the spot price at time $t$, $S_m$ is the level that the commodity price approaches in the long run, $k$ is the speed of reversion towards the 'normal' level, $\sigma_\eta$ is the instantaneous volatility of the commodity price, and $dz_t$ denotes the increment of a Wiener process which is normally distributed with zero mean and variance $dt$. Our choice of model is justified because, as it is well known, the explicit solution to the IGBM model satisfies some reasonable conditions, namely nonnegativity and homogeneity, as well as encompassing the GBM as a particular case[3].

Let $\widehat{S}_t$ denotes the risk-neutral version of $S_t$. The stochastic process for the spot price under the equivalent martingale measure with a market value of risk $\lambda$ is given by

$$d\widehat{S}_t = \left[k(S_m - \widehat{S}_t) - \lambda \widehat{S}_t\right]dt + \sigma_\eta \widehat{S}_t dz_t,$$

so that the expected value at time $t$ of a future price $S_{t+\tau}$ is

$$F(t,\tau) = E_t(\widehat{S}_T = S_{t+\tau}) = \frac{kS_m}{k+\lambda} + \left(S_t - \frac{kS_m}{k+\lambda}\right)\mathrm{e}^{-(k+\lambda)\tau}, \tag{2}$$

where $F(t,\tau)$ is the price value of a futures contract with maturity or settlement date $T$ such that $\tau = T - t$.

Differentiating $F(t,\tau)$ in (2) with respect to $S_t$ and $t$ and applying Itô's lemma we have

$$dF = \mathrm{e}^{-(k+\lambda)\tau} dS_t + \left(S_t - \frac{kS_m}{k+\lambda}\right)(k+\lambda)\mathrm{e}^{-(k+\lambda)\tau} dt,$$

[2]This is a member of the general affine class of stochastic processes (Piazzesi, 2010), also known in the literature as Integrated GBM and related to the so called Geometric Ornstein-Uhlenbeck process (Bhattacharya, 1978). It has been reported to be well suited for energy futures contracts; (Pilipovic, 1998, section 4; see also Abadie and Chamorro 2008 and references therein).

[3]More specifically, the presence of $S_t$ in the diffusion term implies $S_t > 0$ almost surely, for all $t > 0$. On the other hand homogeneity means that if the price on one unit reverts to some mean value then the price of two units revert to twice that same mean value; (see *e.g.* Bhattacharya, 1978; Kloeden and Platen, 1992; Sick, 1995; Pilipovic, 1998; Robel, 2001; Weir, 2005).

whence, substituting from (1) and simplifying, we obtain

$$dF = \mathrm{e}^{-(k+\lambda)\tau}\big(\lambda S_t dt + \sigma_\eta S_t dz_t\big). \tag{3}$$

Furthermore, from (2) we obtain

$$\Big[F(t,\tau) - \frac{kS_m}{k+\lambda}\Big]\mathrm{e}^{(k+\lambda)\tau} = S_t - \frac{kS_m}{k+\lambda},$$

so that for two different settlement dates $t+\tau < t+\tau^\star$ such that $\tau^\star - \tau = \Delta\tau$,

$$\begin{aligned}
\Big[F(t,\tau^\star) - \frac{kS_m}{k+\lambda}\Big]\mathrm{e}^{(k+\lambda)\tau^\star} &= \Big[F(t,\tau) - \frac{kS_m}{k+\lambda}\Big]\mathrm{e}^{(k+\lambda)\tau} \\
F(t,\tau^\star) &= \frac{kS_m}{k+\lambda} + \Big[F(t,\tau) - \frac{kS_m}{k+\lambda}\Big]\mathrm{e}^{-(k+\lambda)(\tau^\star-\tau)} \\
F(t,\tau+\Delta\tau) &= F(t,\tau)\,\mathrm{e}^{-(k+\lambda)\Delta\tau} + \frac{kS_m}{k+\lambda}\big(1-\mathrm{e}^{-(k+\lambda)\Delta\tau}\big) + \sigma_\zeta dw_\tau,
\end{aligned} \tag{4}$$

where the last term, proportional to the increment of a Wiener process independent of $dz_t$, has been added in order to capture futures prices departures from the theoretical IGBM model (*i.e.* like a *local* IGBM trend). The differential equations (1) and (3) together with the difference equation (4) are the three fundamental equations that determine the mechanism with which futures contracts prices evolve along the two-dimensional time-maturity space.

### *2.1. The empirical model*

Discretizing (1), (3) and (4) we have the following system of three discrete transition equations

$$\begin{aligned}
S_{t,\tau} &= (1-k\Delta t)S_{t-1,\tau} + k\Delta t S_m + S_{t-1,\tau}\eta_t, \\
F_{t,\tau} &= F_{t-1,\tau} + (\lambda\Delta t + \eta_t)\mathrm{e}^{-(k+\lambda)\tau}S_{t-1,\tau}, \\
F_{t,\tau} &= \mathrm{e}^{-(k+\lambda)\Delta\tau}F_{t,\tau-1} + \big(1-\mathrm{e}^{-(k+\lambda)\Delta\tau}\big)\frac{kS_m}{k+\lambda} + \zeta_\tau,
\end{aligned}$$

where the subindices $t,\tau$ now represent cell positions along both dimensions of the sample grid $G \subset \mathbb{Z}^2$ with observation intervals $(\Delta t, \Delta\tau)$, and $\eta_t$, $\zeta_\tau$ are independent normally distributed innovations with zero mean and variances $\sigma_\eta^2\Delta t$, $\sigma_\zeta^2\Delta\tau$ respectively. After merging the last two equations and rearranging, the discrete system can be written as

$$\underbrace{\begin{pmatrix} S_{t,\tau} \\ F_{t,\tau} \end{pmatrix}}_{\alpha_{t,\tau}} = \underbrace{\begin{bmatrix} (1-k\Delta t) & 0 \\ \gamma_1 \lambda \Delta t \mathrm{e}^{-(k+\lambda)\tau} & \gamma_1 \end{bmatrix}}_{\Phi_\tau^{(1}} \underbrace{\begin{pmatrix} S_{t-1,\tau} \\ F_{t-1,\tau} \end{pmatrix}}_{\alpha_{t-1,\tau}} + \underbrace{\begin{bmatrix} 0 & 0 \\ 0 & \gamma_2 \mathrm{e}^{-(k+\lambda)\Delta\tau} \end{bmatrix}}_{\Phi^{(2}} \underbrace{\begin{pmatrix} S_{t,\tau-1} \\ F_{t,\tau-1} \end{pmatrix}}_{\alpha_{t,\tau-1}} \tag{5}$$

$$+ \underbrace{\begin{bmatrix} k\Delta t S_m \\ \gamma_2 \frac{kS_m}{k+\lambda}\left(1-\mathrm{e}^{-(k+\lambda)\Delta\tau}\right) \end{bmatrix}}_{c} + \underbrace{\begin{bmatrix} 1 \\ \gamma_1 \mathrm{e}^{-(k+\lambda)\tau} \end{bmatrix} S_{t-1,\tau}}_{R_{t,\tau}^{(1}} \eta_t + \underbrace{\begin{bmatrix} 0 \\ \gamma_2 \end{bmatrix}}_{R^{(2}} \zeta_\tau,$$

where $\gamma_1$, $\gamma_2$ are time-maturity allocation weights such that $\gamma_1 + \gamma_2 = 1$ (typically $\gamma_1 = \gamma_2 = ½$).

### *2.2. State-space form*

We note that (5) takes the form of a transition equation in a two-dimensional state-space (SS) model with two states $S_{t,\tau}$ and $F_{t,\tau}$:

$$\alpha_{t,\tau} = \Phi_\tau^{(1}\alpha_{t-1,\tau} + \Phi^{(2}\alpha_{t,\tau-1} + c + R_{t,\tau}^{(1}\eta_t + R^{(2}\zeta_\tau. \tag{6}$$

The corresponding measurement equation is

$$y_{t,\tau} = Z'\alpha_{t,\tau} + \varepsilon_{t,\tau}, \tag{7}$$

where $y_{t,\tau}$ is the observed price at time $t$ of a commodity's futures contract with settlement date $t+\tau$, $Z = [0 \quad 1]'$ and $\varepsilon_{t,\tau}$ is a serially uncorrelated and normally distributed zero-mean disturbance with variance $\sigma_\varepsilon^2$ that is independent of $\eta_t$ and $\zeta_\tau$. We may observe then that, in general, commodity futures may actually be allowed to be observed with errors (although not necessarily so, *e.g.* if $\sigma_\varepsilon = 0$).

## 3. The 2D Kalman Filter

The 2DKF recursions are derived in appendix Appendix A. Equations (A.2), (A.3), and (A.4) together with (A.1) and (A.5) conform the prediction stage while equations (A.9) and (A.7) conform the updating stage of the two-dimensional Kalman filter.

The 2DKF equations are to be run recursively starting from the first observation at point $(1,1)$ to the last at $(T,n)$ exhausting both dimensions in turn, that is, $\forall(t,\tau)\in G\subset\mathbb{Z}^2$:

$$\widetilde{a}_{t,\tau} = \Phi_\tau^{(1}\, a_{t-1,\tau} + \Phi^{(2}\, a_{t,\tau-1} + c, \tag{A.1$^\star$}$$

$$\widetilde{y}_{t,\tau} = Z'\,\widetilde{a}_{t,\tau}, \tag{A.2$^\star$}$$

$$v_{t,\tau} = y_{t,\tau} - \widetilde{y}_{t,\tau}, \tag{A.3$^\star$}$$

$$\widetilde{P}_{t,\tau} = \Phi_\tau^{(1}\, P_{t-1,\tau}\, \Phi_\tau^{(1\prime} + \Phi^{(2}\, P_{t,\tau-1}\, \Phi^{(2\prime} + R_{t,\tau}^{(1} R_{t,\tau}^{(1\prime} q_\eta + R^{(2} R^{(2\prime} q_\zeta, \tag{A.5$^\star$}$$

$$f_{t,\tau} = Z'\,\widetilde{P}_{t,\tau} Z + 1, \tag{A.4$^\star$}$$

$$K_{t,\tau} = \widetilde{P}_{t,\tau} Z / f_{t,\tau}, \tag{A.8$^\star$}$$

$$a_{t,\tau} = \widetilde{a}_{t,\tau} + K_{t,\tau} v_{t,\tau}, \tag{A.9$^\star$}$$

$$P_{t,\tau} = (I_m - K_{t,\tau} Z')\widetilde{P}_{t,\tau}, \tag{A.7$^\star$}$$

where $q_\eta$ and $q_\zeta$ are signal-noise ratios and $\{a_{t,\tau}, P_{t,\tau} \;:\; t=0 \vee \tau=0\}$ are known from the boundary conditions. In our case we suggest obtaining initial values from the observations themselves according to the transition equations as follows:

$$\begin{aligned} a_{t,0} &= \begin{bmatrix} b_t & b_t \end{bmatrix}', \\ a_{0,\tau} &= \begin{bmatrix} b_0 & y_{1,\tau} - b_0 \lambda \Delta t \mathrm{e}^{-(k+\lambda)\tau} \end{bmatrix}', \end{aligned}$$

where $b_t = y_{t,1}\mathrm{e}^{(k+\lambda)\Delta\tau} + \frac{kS_m}{k+\lambda}\left(1-\mathrm{e}^{(k+\lambda)\Delta\tau}\right)$ and $b_0 = (b_1 - k\Delta t S_m)/(1-k\Delta t)$.

## 4. Maximum likelihood estimation

Let $(\psi, \sigma_\varepsilon^2)$, with $\psi = (S_m, k, \lambda, q_\eta, q_\zeta)$, $q_\eta = \sigma_\eta^2 \Delta t / \sigma_\varepsilon^2$ and $q_\zeta = \sigma_\zeta^2 \Delta\tau / \sigma_\varepsilon^2$, be the vector of unknown parameters in our futures-surface model. From a prediction error decomposition similar to that commonly used in some one-dimensional models (Harvey, 1989), the joint density function of the set of observations $\mathscr{Y} = \{y_{t,\tau} \;:\; (t,\tau)\in G\}$ can be expressed in terms of the prediction errors $v_{t,\tau}$ obtained from the 2DKF recursions in (A.3$^\star$) since they are uncorrelated Gaussian deviates distributed as

$$v_{t,\tau} \sim \mathrm{NID}\big(0, \sigma_\varepsilon^2 f_{t,\tau}\big), \quad (t,\tau)\in G. \tag{8}$$

Consequently, the log-likelihood function can be written as a summation of independent terms as

$$\mathscr{L}(\mathscr{Y};\psi,\sigma_\varepsilon^2) = -\frac{Tn}{2}\log 2\pi - \frac{Tn}{2}\log\sigma_\varepsilon^2 - \frac{1}{2}\sum_{(t,\tau)\in G}\log f_{t,\tau} - \frac{1}{2\sigma_\varepsilon^2}\sum_{(t,\tau)\in G} v_{t,\tau}^2 / f_{t,\tau}.$$

After concentrating $\sigma_\varepsilon^2$ out, the profile log-likelihood function of vector $\psi$ can be written as

$$\mathscr{L}_c(\mathscr{Y};\psi) \propto -\log \widetilde{\sigma}_\varepsilon^2(\psi) - \frac{1}{Tn} \sum_{(t,\tau)\in G} \log f_{t,\tau}, \tag{9}$$

where $\widetilde{\sigma}_\varepsilon^2(\psi)$ is the maximum-likelihood estimate (mle) of $\sigma_\varepsilon^2$:

$$\widetilde{\sigma}_\varepsilon^2(\psi) = \frac{1}{Tn} \sum_{(t,\tau)\in G} \nu_{t,\tau}^2 / f_{t,\tau}.$$

Therefore, the unknown parameters in $\psi$ can be estimated by maximizing the expression in (9).

## 5. Post-sample predictive testing

In order to assess the predictive performance of the model an auxiliary reestimation may be run leaving out the last $\ell$ sample periods. The gap between observed prices and forecasts are post-sample forecast errors whose distribution is the same as for the in-sample prediction errors in (8). Therefore, the post-sample predictive test statistics

$$F_p(\tau) = \frac{(T-\ell)\sum_{j=0}^{\ell-1} \nu_{T-j,\tau}^2 / f_{T-j,\tau}}{\ell \sum_{t=1}^{T-\ell} \nu_{t,\tau}^2 / f_{t,\tau}}, \quad \tau = 1,\ldots,n, \tag{10}$$

will have an $\mathscr{F}$-Snedecor distribution with $(\ell, T-\ell)$ degrees of freedom. Furthermore, the overall test statistic

$$F_p = \frac{(T-\ell)\sum_{(t,\tau)\in Go} \nu_{t,\tau}^2 / f_{t,\tau}}{\ell \sum_{(t,\tau)\in Gi} \nu_{t,\tau}^2 / f_{t,\tau}}, \tag{11}$$

where $[Gi'|Go']'$ is the in-sample-post-sample partition of the observed grid $G$, will have an $\mathscr{F}(n\ell, n(T-\ell))$ distribution (*cf.* Harvey, 1989, p.271).

## 6. Real example: coal futures

We estimated the parameters of our IGBM model (1-3) by fitting its SS form (6-7), through use of the 2DKF filter described above, to a $250\times35$ dataset of daily observations of prices for NYMEX Central Appalachian coal futures contracts during a recent full year, namely from Dec/01/2008 to Nov/25/2009, with approximately monthly maturities, *i.e.* contracts maturing in the next month and in two, three, etc. months and so on up to the next 35 following months. Figure 1 shows the observed data surface.

Table 1 shows mle of the model parameters obtained after fitting the SS form of the IGBM model to the coal futures data. Figure 2 shows the underlying trend $\widetilde{\alpha}_{2;t,\tau}$ extracted from the coal futures dataset and

Table 1: NYMEX coal futures data: Maximum Likelihood Estimates of IGBM parameters

| Parameter | Description | Estimate | 95% Confidence Interval |
|---|---|---|---|
| $S_m$ | long-run spot price | $71.85 | [ $64.73 ; $78.98 ] |
| $k$ | mean-reversion rate | 0.1253 | [ 0.0907 ; 0.1599 ] |
| $\sigma_\eta$ | spot volatility | 1.1434 | [ 1.0116 ; 1.2753 ] |
| $\lambda$ | market value of risk | -0.0018 | [ -0.0211 ; 0.0174 ] |
| $\sigma_\zeta$ | maturity dynamics std dev | 0.0013 | [ 0.0013 ; 0.0013 ] |
| $\sigma_\varepsilon$ | measurement error std dev | 0.1133 | [ 0.1089 ; 0.1177 ] |
| $t_{1/2}$ | 'half-life' | 5.53 *yrs* | [ 4.33 ; 7.64 ] |
| $F_\infty$ | futures equilibrium price | $72.92 | [ $71.23 ; $84.29 ] |

*Estimates obtained by fitting the state-space form of the model using a two-dimensional Kalman filter. Data are daily observations of NYMEX coal futures prices for contracts at different maturities during the period Dec/1/2008 to Nov/25/2009. CI's from parametric bootstraping (100 replications) of the IGBM state-space model with the given estimated values.*

---

Figure 3 shows the corresponding estimated spot prices $\widetilde{\alpha}_{1;t}$. We note that these spot prices tend to a value of $71.85 in the long run with a significant mean reversion typical of energy markets (Schwartz and Smith, 2000); the "half-life" being approximately 5.5 years ($\log 2/k$). On the other hand, the estimated futures equilibrium price ($72.92) appears coherent with the futures market quotes shown in Figure 1. Small but significant variances of maturity innovations and measurement error are also obtained, which means that the *expected* behavior of coal futures follows an IGBM process *locally*.

Furthermore, since the futures equilibrium price is higher than the long run spot price, we have that the market value of risk is slightly negative (common in energy markets in recent years —see *e.g.* Botterud et al., 2002; Schwartz and Smith, 2000, among others) although not statistically significant. In other words, we cannot rule out a zero risk premium. As a consequence, coal futures prices tend to an equilibrium value that falls within a 95% confidence interval of its long-run spot price as shown in the extended long-run scenario of Figure 4. Finally, the estimated spot volatility (1.14) appears consistent with the historical behavior in this coal market, which has maintained a historical volatility smaller than other commodities, although it appears to have been increasing in recent years.

In order to assess the predictive performance of the model an auxiliary reestimation was run leaving out the last two weeks in the sample (*i.e.* $\ell = 10$). Figure 5 shows the one-step-ahead forecasts in the post-sample period (broken lines) as compared with the actual prices recorded (solid lines). We note that, except for the two-months maturity, the osa forecasts track the observed values closely. We note, however, that during those two last weeks there was no change in futures prices between the second and third maturities. Consequently, it is the latter that the model is able to forecast correctly. Figure 6 shows the corresponding post-sample predictive test statistics (10) for all the maturities from one month up to 35 months. We observe that, except for the anomalous two-months maturity, none of them is significant at the 5 % level. Furthermore, excluding the second maturity, the overall test statistic value (11) is 0.1853. All this indicates that the post-sample performance of the IGBM model is quite acceptable.

## 7. Conclusions

This paper attempts to contribute to a better calibration of stochastic models for commodities futures by providing a method to estimate the model parameters using both the evolution of futures prices over time as well as the term structure of daily commodity futures along a second dimension of maturities. In order to do so a two-dimensional Kalman filter has been designed which, to the best of our knowledge, has not been attempted before in this context. This two-dimensional Kalman filter makes use of a greater amount of information, both in terms of data as well as dynamics, contained in the stochastic surface generated by the term structure.

The number of maturities that can be taken into account by using a standard vector Kalman filter cannot be very high in practice. Therefore, whereas our two-dimensional model may use all quotes available each day, some previous applications are based on a one-dimensional vector setup with a limited number of maturities, thus omitting potentially relevant information from data for other maturities (for example, Schwartz, 1997, uses only five future contracts per day). This is because the covariance matrices involved in the Kalman filter of the usual one-dimensional vector setup are of an order of magnitude $O(n^2)$ so that the problem soon becomes computationally unfeasible when the number of maturities $n$ increases. In contrast, the size of the state vector in the present two-dimensional setup does not depend on $n$ so that no magnitude problem arises.

The stochastic process used as a basis from which to derive the state-space form follows an Inhomogeneous Geometric Brownian Motion (IGBM) model. This process implies nonnegativity and homogeneity and allows the existence of mean reversion, something frequently observed in commodity markets that is compatible with coal futures data. Besides, since for certain parameter values the IGBM process collapses to the GBM, our two-dimensional Kalman filter method can also be used to estimate a simpler GBM model.

In order to illustrate the method, a real-data application is carried out using NYMEX Central Appalachian Coal futures prices. This market has been chosen because of the relatively high number of maturity prices on any given day. The results obtained after calibrating the parameters are accompanied by their confidence intervals and support the existence of reversion to the mean, a non-significant market value of risk and moderate volatility for Central Appalachian Coal futures. Furthermore, post-sample predictive testing indicates that the IGBM model performs well in forecasting coal futures prices.

For simplicity's sake the state-space model has been kept as simple as possible, but some extensions can be easily implemented without much difficulty. For example, as it stands the proposed method only contemplates one single time-invariant risk factor. However, the model could be easily extended to include several risk factors that might improve the results of the estimation (*cf.* Schwartz, 1997). Or even it could contemplate the market value of risk as an additional state evolving in time. Finally, some other unobserved components (*e.g.* seasonality that is relevant in markets such as natural gas) could also be incorporated without much difficulty. In both cases these extensions of the present model need modify the state-space form but not the two-dimensional Kalman filter itself.

**Acknowledgement**

Financial support from *UPV/EHU Econometrics Research Group*, Basque Government grant GIC07/53-IT-334-07, is gratefully acknowledged. The author wishes to thank L. Abadie and J. Chamorro for their comments on an earlier draft. Any remaining errors are totally my own.

## Appendix A. Derivation of the 2DKF equations

Let $(t,\tau) \in G \subset \mathbb{Z}^2$ be a point in a $T \times n$ subset of the two-dimensional integer lattice and let $y_{t,\tau}$ be an observation in $G$ generated as from the state-space equations specified by (7) and (6). That is

$$y_{t,\tau} = Z'\alpha_{t,\tau} + \varepsilon_{t,\tau}, \tag{7$^\star$}$$

$$\alpha_{t,\tau} = \Phi_\tau^{(1}\alpha_{t-1,\tau} + \Phi^{(2}\alpha_{t,\tau-1} + c + R_{t,\tau}^{(1}\eta_t + R^{(2}\zeta_\tau, \tag{6$^\star$}$$

where $\{\varepsilon_{t,\tau}\}$, $\{\eta_t\}$ and $\{\zeta_t\}$ are normally distributed two-dimensional sequences of white noise disturbances such that $E(\varepsilon_{t,\tau}) = E(\eta_t) = E(\zeta_t) = E(\varepsilon_{t,\tau}\eta_r) = E(\varepsilon_{t,\tau}\zeta_r) = E(\eta_t\zeta_r) = 0$ and $V(\varepsilon_{t,\tau}) = \sigma_\varepsilon^2$, $V(\eta_t) = \sigma_\eta^2$, $V(\zeta_t) = \sigma_\zeta^2$, $\forall (t,\tau) \in G$, $r = 1,\ldots,T$.

*Initial state (boundary conditions).* Let $\alpha_0 = \{\alpha_{t,\tau} : t = 0 \vee \tau = 0\}$ be the set of initial states. We assume that

$$E(\alpha_0) = a_0; \quad V(\alpha_0) = P_0,$$

and they are uncorrelated with the disturbances

$$E(\varepsilon_{t,\tau}\alpha_0{}') = 0; \quad E(\eta_t \alpha_0{}') = 0, \forall (t,\tau) \in G.$$

*Appendix A.1. Prediction Equations*

At any given location $(t,\tau) \in G$, let $\mathscr{Y}_{t,\tau}$ be the set of all observations up to and including $y_{t,\tau}$, that is $\mathscr{Y}_{t,\tau} = \{y_{r,s} : r \le t \wedge s \le \tau\}$. Similarly, let $\mathscr{Y}^p_{t,\tau}$ be the set of all previous observations, that is $\mathscr{Y}^p_{t,\tau} = \mathscr{Y}_{t,\tau} - \{y_{t,\tau}\} = \mathscr{Y}_{t-1,\tau} \cup \mathscr{Y}_{t,\tau-1}$.

Furthermore, let the respective conditional mean vectors and covariance matrices be denoted as

$$E(\alpha_{t,\tau} \mid \mathscr{Y}_{t,\tau}) = a_{t,\tau};$$
$$E((\alpha_{t,\tau} - a_{t,\tau})(\alpha_{t,\tau} - a_{t,\tau})' \mid \mathscr{Y}_{t,\tau}) = \sigma^2_\varepsilon P_{t,\tau},$$

and

$$E(\alpha_{t,\tau} \mid \mathscr{Y}^p_{t,\tau}) = \widetilde{a}_{t,\tau};$$
$$E((\alpha_{t,\tau} - \widetilde{a}_{t,\tau})(\alpha_{t,\tau} - \widetilde{a}_{t,\tau})' \mid \mathscr{Y}^p_{t,\tau}) = \sigma^2_\varepsilon \widetilde{P}_{t,\tau},$$

where the variance of $\varepsilon$ takes the role of a scale factor.

Let us first assume that we have observed all data preceding point $(t,\tau) \in G$ along both dimensions so that the current information set is $\mathscr{Y}^p_{t,\tau}$. Taking conditional expectations in the transition equation (6) we have that

$$E(\alpha_{t,\tau} \mid \mathscr{Y}^p_{t,\tau}) = \Phi^{(1}_\tau E(\alpha_{t-1,\tau} \mid \mathscr{Y}^p_{t,\tau}) + \Phi^{(2} E(\alpha_{t,\tau-1} \mid \mathscr{Y}^p_{t,\tau}) + c,$$
$$\widetilde{a}_{t,\tau} = \Phi^{(1}_\tau a_{t-1,\tau} + \Phi^{(2} a_{t,\tau-1} + c, \tag{A.1}$$

that is, the predicted value for the state variable $\alpha_{t,\tau}$.[4]

Similarly, from the measurement equations (7) we obtain the prediction of the observation $y_{t,\tau}$

$$E(y_{t,\tau} \mid \mathscr{Y}^p_{t,\tau}) = Z' E(\alpha_{t,\tau} \mid \mathscr{Y}^p_{t,\tau}),$$
$$\widetilde{y}_{t,\tau} = Z' \widetilde{a}_{t,\tau}. \tag{A.2}$$

[4]Note that $\mathscr{Y}^p_{t,\tau} = \mathscr{Y}_{t-1,\tau} + \{w_{t1}, \dots, w_{t\tau-1}\}$, where the $w$'s are innovations uncorrelated with $\mathscr{Y}_{t-1,\tau}$ and, hence, with $\alpha_{t-1,\tau}$. Likewise for the second term.

*Appendix A.2. Prediction error*

From (7) and (A.2) we have

$$\begin{aligned} \nu_{t,\tau} &= y_{t,\tau} - \widetilde{y}_{t,\tau} \\ &= Z'\left(\alpha_{t,\tau} - \widetilde{a}_{t,\tau}\right) + \varepsilon_{t,\tau}. \end{aligned} \tag{A.3}$$

Taking conditional expectations we obtain the prediction error variance

$$\begin{aligned} E\big(\nu_{t,\tau}\big) &= Z'E\big(\alpha_{t,\tau} - \widetilde{a}_{t,\tau} \mid \mathscr{Y}^{p}_{t,\tau}\big) = 0, \\ V\big(\nu_{t,\tau}\big) &= Z'E\big((\alpha_{t,\tau} - \widetilde{a}_{t,\tau})(\alpha_{t,\tau} - \widetilde{a}_{t,\tau})' \mid \mathscr{Y}^{p}_{t,\tau}\big)Z + V\big(\varepsilon_{t,\tau}\big), \\ \sigma^2_{\varepsilon} f_{t,\tau} &= Z'\sigma^2_{\varepsilon}\widetilde{P}_{t,\tau}Z + \sigma^2_{\varepsilon}, \\ f_{t,\tau} &= Z'\widetilde{P}_{t,\tau}Z + 1. \end{aligned} \tag{A.4}$$

From (6) and (A.1)

$$\alpha_{t,\tau} - \widetilde{a}_{t,\tau} = \Phi^{(1}_{\tau}\big(\alpha_{t-1,\tau} - a_{t-1,\tau}\big) + \Phi^{(2}\big(\alpha_{t,\tau-1} - a_{t,\tau-1}\big) + R^{(1}_{t,\tau}\eta_t + R^{(2}\zeta_\tau,$$

so that, taking conditional expectations,

$$\begin{aligned} E\big(\alpha_{t,\tau} - \widetilde{a}_{t,\tau} \mid \mathscr{Y}^{p}_{t,\tau}\big) &= \Phi^{(1}_{\tau}\big[E\big(\alpha_{t-1,\tau} \mid \mathscr{Y}_{t-1,\tau}\big) - a_{t-1,\tau}\big] \\ &\quad + \Phi^{(2}\big[E\big(\alpha_{t,\tau-1} \mid \mathscr{Y}_{t,\tau-1}\big) - a_{t,\tau-1}\big] = 0, \\ V\big(\alpha_{t,\tau} - \widetilde{a}_{t,\tau} \mid \mathscr{Y}^{p}_{t,\tau}\big) &= \Phi^{(1}_{\tau}E\big((\alpha_{t-1,\tau} - a_{t-1,\tau})(\alpha_{t-1,\tau} - a_{t-1,\tau})' \mid \mathscr{Y}_{t-1,\tau}\big)\Phi^{(1\prime}_{\tau} \\ &\quad + \Phi^{(2}E\big((\alpha_{t,\tau-1} - a_{t,\tau-1})(\alpha_{t,\tau-1} - a_{t,\tau-1})' \mid \mathscr{Y}_{t,\tau-1}\big)\Phi^{(2\prime} \\ &\quad + \Phi^{(1}_{\tau}E\big((\alpha_{t-1,\tau} - a_{t-1,\tau})(\alpha_{t,\tau-1} - a_{t,\tau-1})' \mid \mathscr{Y}_{t-1,\tau}, \mathscr{Y}_{t,\tau-1}\big)\Phi^{(2\prime} \\ &\quad + \Phi^{(2}E\big((\alpha_{t,\tau-1} - a_{t,\tau-1})(\alpha_{t-1,\tau} - a_{t-1,\tau})' \mid \mathscr{Y}_{t,\tau-1}, \mathscr{Y}_{t-1,\tau}\big)\Phi^{(1\prime}_{\tau} \\ &\quad + R^{(1}_{t,\tau}R^{(1\prime}_{t,\tau}\sigma^2_{\eta} + R^{(2}R^{(2\prime}\sigma^2_{\zeta}, \\ \sigma^2_{\varepsilon}\widetilde{P}_{t,\tau} &= \Phi^{(1}_{\tau}\ \sigma^2_{\varepsilon}P_{t-1,\tau}\ \Phi^{(1\prime}_{\tau} + \Phi^{(2}\ \sigma^2_{\varepsilon}P_{t,\tau-1}\ \Phi^{(2\prime} \\ &\quad + \Phi^{(1}_{\tau}\ \sigma^2_{\varepsilon}C_{t-1,\tau,t,\tau-1}\ \Phi^{(2\prime} + \Phi^{(2}\ \sigma^2_{\varepsilon}C_{t,\tau-1,t-1,\tau}\ \Phi^{(1\prime}_{\tau} + R^{(1}_{t,\tau}R^{(1\prime}_{t,\tau}\sigma^2_{\eta} + R^{(2}R^{(2\prime}\sigma^2_{\zeta} \\ \widetilde{P}_{t,\tau} &= \Phi^{(1}_{\tau}\ P_{t-1,\tau}\ \Phi^{(1\prime}_{\tau} + \Phi^{(2}\ P_{t,\tau-1}\ \Phi^{(2\prime} \\ &\quad + \Phi^{(1}_{\tau}\ C_{t-1,\tau,t,\tau-1}\ \Phi^{(2\prime} + \Phi^{(2}\ C_{t,\tau-1,t-1,\tau}\ \Phi^{(1\prime}_{\tau} + R^{(1}_{t,\tau}R^{(1\prime}_{t,\tau}q_{\eta} + R^{(2}R^{(2\prime}q_{\zeta}, \end{aligned} \tag{A.5}$$

where $q_\eta = \sigma^2_\eta/\sigma^2_\varepsilon$ and $q_\zeta = \sigma^2_\zeta/\sigma^2_\varepsilon$ are signal-noise ratio parameters. In practice, simulation suggest that the covariance terms can be ignored without significant changes in the estimated results obtained.

*Appendix A.3. Updating Equations*

With information up to, but excluding, the observation at point $(t,\tau) \in G \subset \mathbb{Z}^2$ we have that the best prediction of the state vector at point $(t,\tau)$ and its covariance matrix are given by $\widetilde{a}_{t,\tau}$, $\widetilde{P}_{t,\tau}$ as expressed in (A.1) and (A.5). Once the new observation at point $(t,\tau)$ is obtained we can express all the information contained in the current set $\mathscr{Y}_{t,\tau}$ about state vector $\alpha_{t,\tau}$ by incorporating the measurement equation (7) into the following augmented model

$$\begin{array}{ccccccc} \begin{pmatrix} \widetilde{a}_{t,\tau} \\ y_{t,\tau} \end{pmatrix} & = & \begin{pmatrix} I_m \\ Z' \end{pmatrix} & \alpha_{t,\tau} & + & \begin{pmatrix} \widetilde{a}_{t,\tau} - \alpha_{t,\tau} \\ \varepsilon_{t,\tau} \end{pmatrix} \\ Y & = & X & \beta & + & u \end{array}$$

This takes the form of a general linear regression model where

$$E\left(uu'\right) = \begin{pmatrix} \sigma_\varepsilon^2 \widetilde{P}_{t,\tau} & 0 \\ 0 & \sigma_\varepsilon^2 \end{pmatrix}$$

so that, by Gauss-Markov theorem, the best linear unbiased estimator is

$$\begin{aligned} a_{t,\tau} &= \sigma_\varepsilon^2 P_{t,\tau} \left[ \begin{pmatrix} I_m & Z \end{pmatrix} \begin{pmatrix} \sigma_\varepsilon^{-2} \widetilde{P}_{t,\tau}^{-1} & 0 \\ 0 & \sigma_\varepsilon^{-2} \end{pmatrix} \begin{pmatrix} \widetilde{a}_{t,\tau} \\ y_{t,\tau} \end{pmatrix} \right] \\ &= P_{t,\tau} \left( \widetilde{P}_{t,\tau}^{-1} \widetilde{a}_{t,\tau} + Z y_{t,\tau} \right), \end{aligned} \tag{A.6}$$

where $\sigma_\varepsilon^2 P_{t,\tau}$ is the covariance matrix of $a_{t,\tau}$ such that

$$\begin{aligned} \sigma_\varepsilon^2 P_{t,\tau} &= \left[ \begin{pmatrix} I_m & Z \end{pmatrix} \begin{pmatrix} \sigma_\varepsilon^{-2} \widetilde{P}_{t,\tau}^{-1} & 0 \\ 0 & \sigma_\varepsilon^{-2} \end{pmatrix} \begin{pmatrix} I_m \\ Z' \end{pmatrix} \right]^{-1}, \\ P_{t,\tau} &= \left[ \widetilde{P}_{t,\tau}^{-1} + ZZ' \right]^{-1}. \end{aligned}$$

Applying the Matrix Inversion Lemma to the last expression we have

$$\begin{aligned} P_{t,\tau} &= \widetilde{P}_{t,\tau} - \widetilde{P}_{t,\tau} Z \left(1 + Z' \widetilde{P}_{t,\tau} Z\right)^{-1} Z' \widetilde{P}_{t,\tau} \\ &= \widetilde{P}_{t,\tau} - \widetilde{P}_{t,\tau} Z Z' \widetilde{P}_{t,\tau} / f_{t,\tau} \\ &= (I_m - K_{t,\tau} Z') \widetilde{P}_{t,\tau} \end{aligned} \tag{A.7}$$

where $f_{t,\tau}$ is the prediction error variance in (A.4) and

$$K_{t,\tau} = \widetilde{P}_{t,\tau} Z / f_{t,\tau} \tag{A.8}$$

is the Kalman gain. This $2 \times 1$ vector effectively summarizes the relevance of the new information $y_{t,\tau}$ as against the information already collected up to $\mathscr{Y}_{t,\tau}^{p}$ and determines the degree of updating necessary. For example, if the new observation does not innovate the collected information then $K_{t,\tau} = 0$ reflecting the fact that the state vector and its covariance matrix need no updating.

Finally, substituting (A.7) into (A.6) and doing some algebra

$$\begin{aligned}
a_{t,\tau} &= P_{t,\tau}\left(\widetilde{P}_{t,\tau}^{-1}\widetilde{a}_{t,\tau} + Z y_{t,\tau}\right),\\
&= (I_m - K_{t,\tau}Z')\widetilde{P}_{t,\tau}\left(\widetilde{P}_{t,\tau}^{-1}\widetilde{a}_{t,\tau} + Z y_{t,\tau}\right),\\
&= (I_m - K_{t,\tau}Z')\widetilde{a}_{t,\tau} + (I_m - K_{t,\tau}Z')\widetilde{P}_{t,\tau}Z y_{t,\tau},\\
&= \widetilde{a}_{t,\tau} - K_{t,\tau}Z'\widetilde{a}_{t,\tau} + \widetilde{P}_{t,\tau}Z f_{t,\tau}^{-1} f_{t,\tau} y_{t,\tau} - K_{t,\tau}Z'\widetilde{P}_{t,\tau}Z y_{t,\tau},\\
&= \widetilde{a}_{t,\tau} - K_{t,\tau}Z'\widetilde{a}_{t,\tau} + K_{t,\tau} f_{t,\tau} y_{t,\tau} - K_{t,\tau}Z'\widetilde{P}_{t,\tau}Z y_{t,\tau},\\
&= \widetilde{a}_{t,\tau} + K_{t,\tau}\left(f_{t,\tau} y_{t,\tau} - Z'\widetilde{P}_{t,\tau}Z y_{t,\tau} - Z'\widetilde{a}_{t,\tau}\right),\\
&= \widetilde{a}_{t,\tau} + K_{t,\tau}\left[\underbrace{(f_{t,\tau} - Z'\widetilde{P}_{t,\tau}Z)}_{1} y_{t,\tau} - Z'\widetilde{a}_{t,\tau}\right],\\
&= \widetilde{a}_{t,\tau} + K_{t,\tau}\left(y_{t,\tau} - Z'\widetilde{a}_{t,\tau}\right),\\
&= \widetilde{a}_{t,\tau} + K_{t,\tau} v_{t,\tau},
\end{aligned} \tag{A.9}$$

where $v_{t,\tau}$ is the prediction error in (A.3). Equations (A.9) and (A.7) form the updating stage of the Kalman filter.

## References


Abadie, L., Chamorro, J., 2008. Valuing flexibility: The case of an integrated gasification combined cycle power plant. Energy Economics 30, 1850–1881.

Bhattacharya, S., 1978. Project valuation with mean-reverting cash flow streams. Journal of Finance 33, 1317–1331.

Botterud, A., Bhattacharyya, A., Ilic, M., 2002. Futures and spot prices – an analysis of the Scandinavian electricity market, proceedings of North American Power Symposium 2002, Tempe, Arizona.

Cortazar, G., Milla, C., Severino, F., 2008. A multicommodity model of futures prices: Using futures prices of one commodity to estimate the stochastic process of another. Journal of Futures Markets 28 (6), 537–560.

Cortazar, G., Naranjo, L., 2006. An *N*-factor gaussian model of oil futures prices. Journal of Futures Markets 26 (3), 243–268.

Cortazar, G., Schwartz, E. S., 2003. Implementing a stochastic model for oil futures prices. Energy Economics 25 (3), 215–238.

Diebold, F. X., Rudebusch, G. D., Aruoba, S. B., 2006. The macroeconomy and the yield curve: a dynamic latent factor approach. Journal of Econometrics 131, 309–338.

Fabozzi, F., Füss, R., Kaiser, D. (Eds.), 2008. The handbook of commodity investing. Wiley, Hoboken, New Jersey.

Fernández-Macho, J., 2008. Spectral estimation of a structural thin-plate smoothing model. Computational Statistics & Data Analysis 53, 189–195.

Geman, H., 2005. Commodities and commodity derivatives. Wiley, West Sussex, England.

Harvey, A. C., 1989. Forecasting, Structural Time Series Models and the Kalman Filter. Cambridge University Press, Cambridge, UK.

Kloeden, P., Platen, E., 1992. Numerical Solution of Stochastic Differential Equations. Springer-Verlag.

London, J., 2007. Modeling Derivatives Applications. FT Press, Upper Saddle River, New Jersey.

Lucia, J., Schwartz, E. S., 2002. Electricity prices and power derivatives: Evidence from the nordic power exchange. Review of Derivatives Research 5 (1), 5–50.

Manoliu, M., Tompaidis, S., 2002. Energy futures prices: term structure models with Kalman filter estimation. Applied Mathematical Finance 9, 21–43.

Nelson, C., Siegel, A., 1987. Parsimonious modeling of yield curves. Journal of Business 60, 473–489.

Piazzesi, M., 2010. Affine term structure models. In: Ait-Sahalia, Y., Hansen, L. P. (Eds.), Handbook of Financial Econometrics. Elsevier, New York, pp. 691–766.

Pilipovic, D., 1998. Energy Risk. McGraw-Hill, New York.

Robel, G., 2001. Real options and mean-reverting prices. The 5th Annual Real Options Conference, UCLA.

Schwartz, E., Smith, J. E., 2000. Short-term variations and long-term dynamics in commodity prices. Management Science 46 (7), 893–911.

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

Sick, G., 1995. Real options. In: Jarrow, R., Maksimovic, V., Ziemba, W. (Eds.), Finance (Handbooks in Operations Research and Management Science). Vol. 9. North-Holland, pp. 631–691.

Wang, P., 1998. Applying two dimensional Kalman filtering for digital terrain modelling. In: Fritsch, D., Englich, M., Sester, M. (Eds.), Between Visions and Applications. Vol. 32/4 of 'IAPRS'. ISPRS Commission IV Symposium on GIS, Stuttgart, Germany.

Weir, J. G., 2005. Valuation of petroleum leases as real options. The 9th Annual Real Options Conference, UCLA.

Woods, J. W., Radewan, C., 1977. Kalman filtering in two dimensions. IEEE Transactions on Information Theory 23 (4), 473–482.

Zou, Y., Sheng, M., Zhong, N., Xu, S., 2004. A generalized Kalman filter for 2D discrete systems. Circuits Systems Signal Processing 23, 351–364.

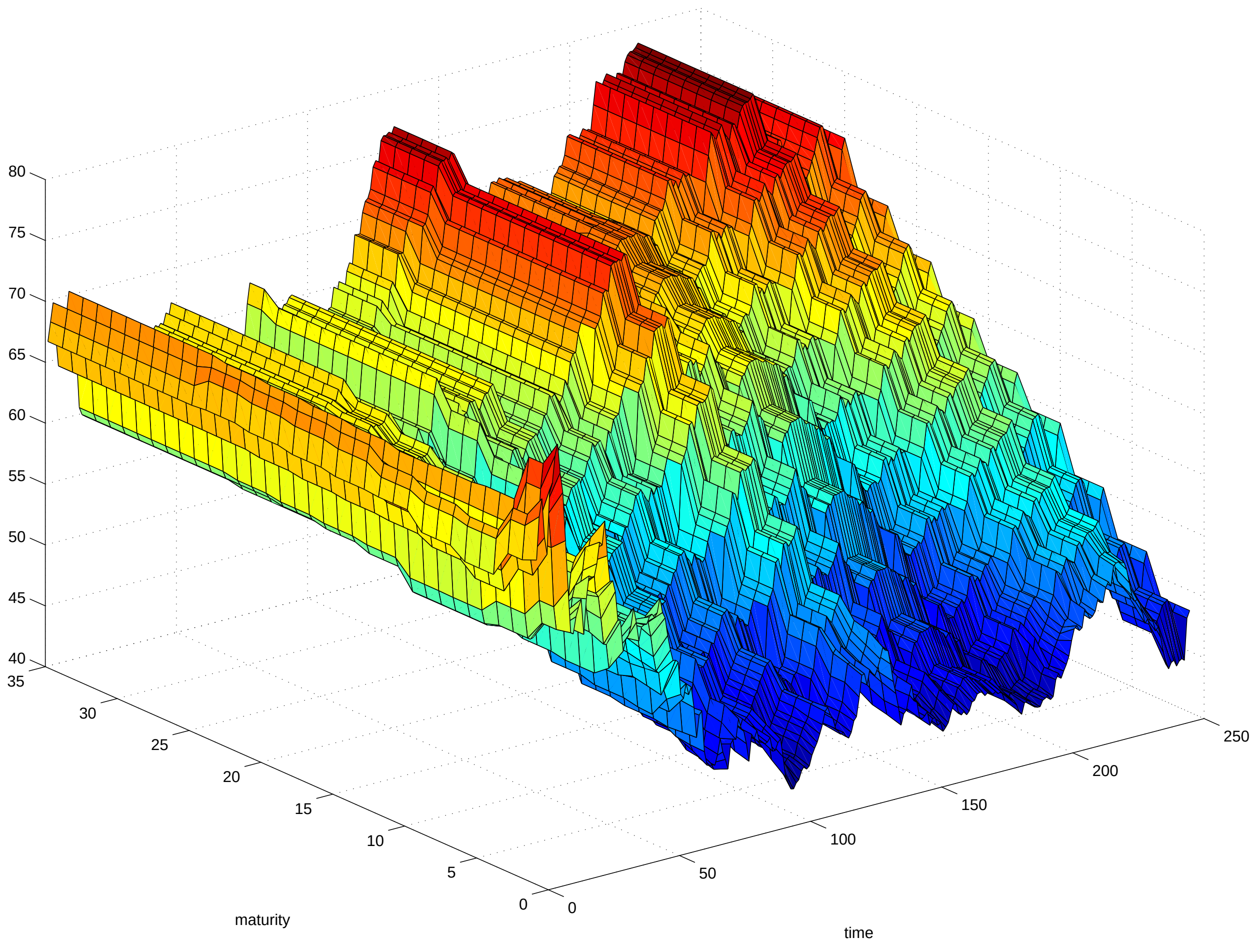



Figure 1: NYMEX coal futures observed prices: Dec/1/2008 to Nov/25/2009.

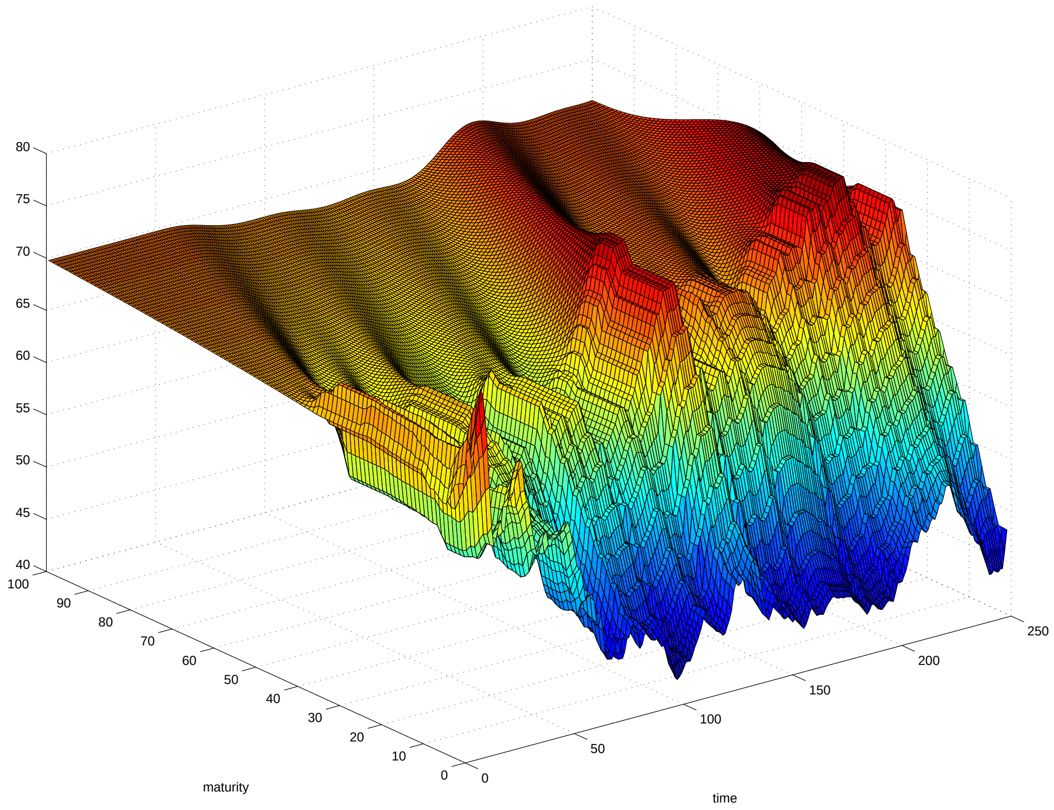



Figure 2: NYMEX coal futures underlying trend.

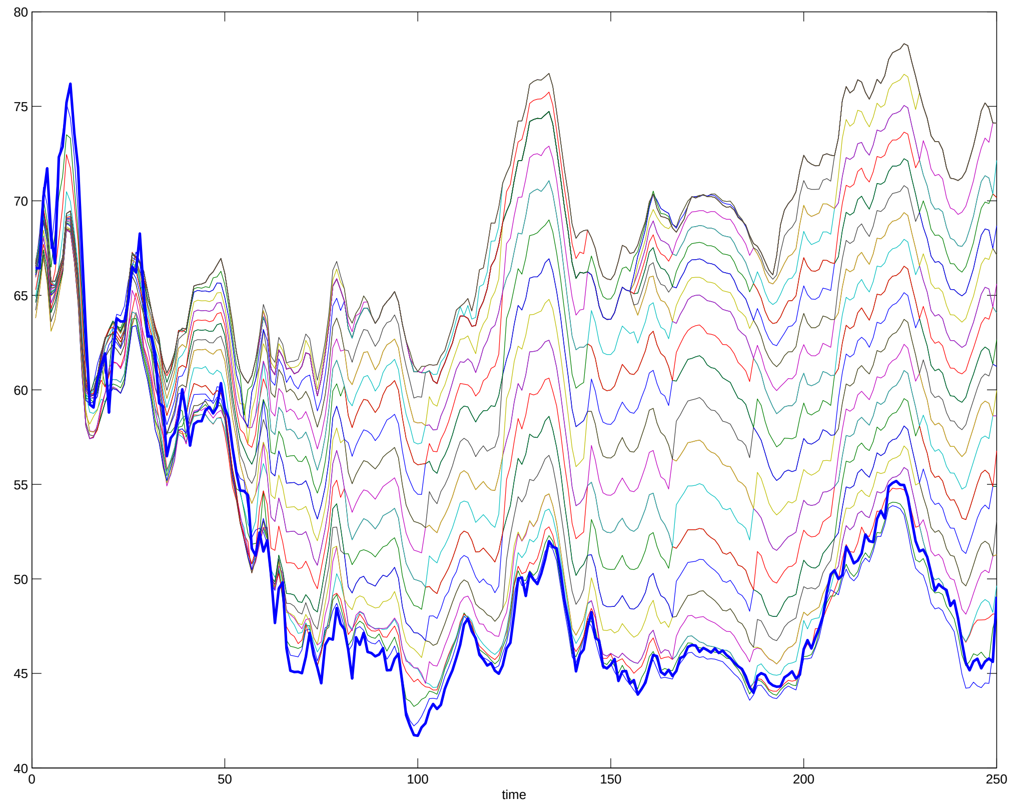



Figure 3: NYMEX coal futures and estimated spot prices (solid blue line).

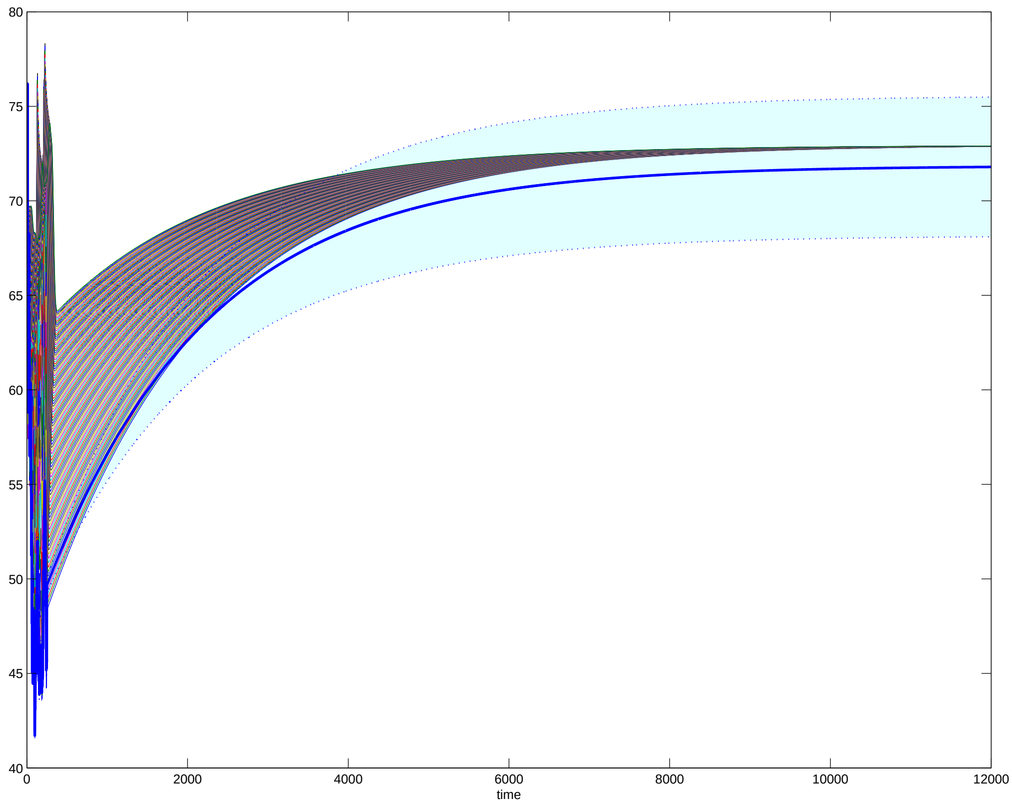



Figure 4: NYMEX coal extended futures and spot prices (solid blue line; shaded area represents its 95% CI).

Figure 5: Post-sample forecasts.

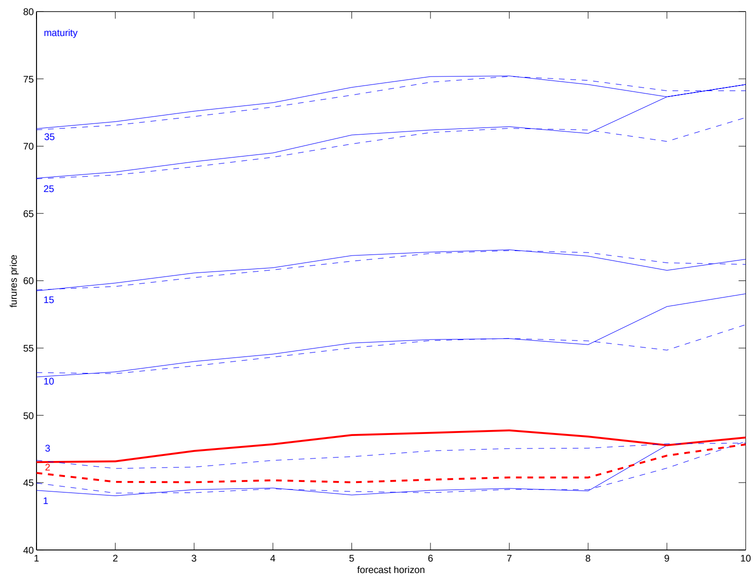
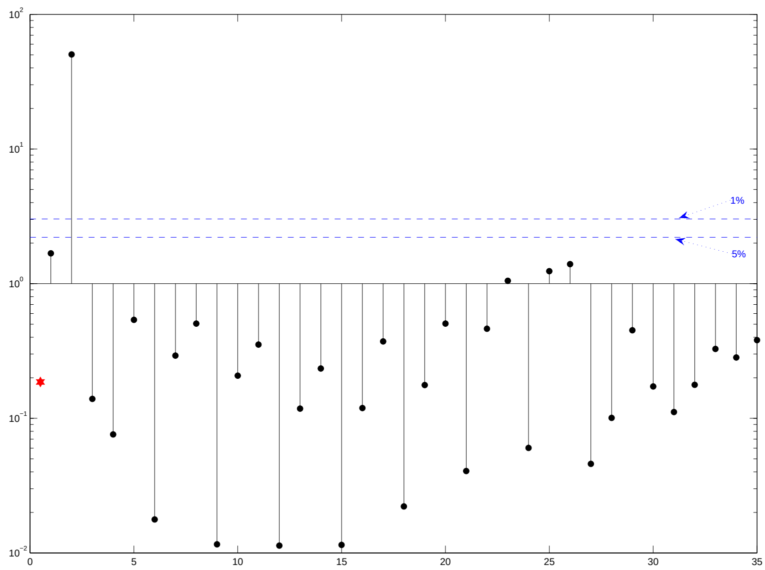



Figure 6: Post-sample predictive tests.




*Logarithmic scale. The dots indicate the test statistic value for maturities $\tau = 1 \ldots 35$. The star on the left signals the statistic value for the overall test excluding the second maturity.*