PONTIFÍCIA UNIVERSIDADE CATÓLICA
DO RIO DE JANEIRO

Rafael de Sequeira Baptista Ferraz

Estimativa de Preços de Contratos Futuros Sobre Petróleo Utilizando o Método do Filtro de Kalman

Dissertação de Mestrado

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia de Produção da PUC - Rio como requisito parcial para obtenção do título de Mestre em Engenharia de Produção.

Orientador: Prof. Tara Keshar Nanda Baidya

Coorientador: Prof. Fernando Antonio Lucena Aiube

Rio de Janeiro, 30 de Junho de 2009

PONTIFÍCIA UNIVERSIDADE CATÓLICA
DO RIO DE JANEIRO

Rafael de Sequeira Baptista Ferraz

Estimativa de Preços de Contratos Futuros sobre Petróleo Utilizando o Método do Filtro de Kalman

Dissertação apresentada como requisito parcial para obtenção do título de Mestre pelo Programa de Pós-Graduação em Engenharia de Produção da PUC-Rio. Aprovada pela Comissão Examinadora abaixo assinada.

Prof. Tara Keshar Nanda Baidya
Orientador
Departamento de Engenharia Industrial - PUC-Rio

Prof. Fernando Antônio Lucena Aiube
Coorientador
Departamento de Engenharia Industrial - PUC-Rio

Prof. Paulo Henrique Soto Costa
Departamento de Engenharia Industrial - PUC-Rio

Prof. Florival Rodrigues de Carvalho
Departamento de Engenharia Química – UFPE

Sr. Krongnon Wailamer de Souza Regueira
Consultor Autônomo

Prof. José Eugenio Leal
Coordenador Setorial do Centro Técnico Científico - PUC-Rio

Rio de Janeiro, 30 de junho de 2009

Rafael de Sequeira Baptista Ferraz

Graduou-se em Engenharia de Produção pela Pontifícia Universidade do Rio de Janeiro em 2006.

Ficha Catalográfica

Ferraz, Rafael de Sequeira Baptista

Estimativa de preços de contratos futuros sobre petróleo utilizando o método do filtro de Kalman / Rafael de Sequeira Baptista Ferraz ; orientador: Tara Keshar Nanda Baidya ; co-orientador: Fernando Antonio Lucena Aiube. – 2009.

173 f. : il. (color.) ; 30 cm

Dissertação (Mestrado em Engenharia Industrial)– Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2009.

Inclui bibliografia

1. Engenharia industrial – Teses. 2. Mercado futuro. 3. Processo estocástico. 4. Filtro de Kalman. 5. Séries temporais. 6. Previsão. I. Baidya, Tara Keshar Nanda. II. Aiube, Fernando Antonio Lucena. III. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Engenharia Industrial. IV. Título.

CDD:658.5

À minha esposa Marília, à minha mãe Maria da Glória,
ao meu pai Newton, aos meus irmãos André e Julia,
à minha madrasta Rosana e à minha sogra Cláudia.

Agradecimentos

Ao meu orientador Tara Baidya, professor do quadro principal do Departamento de Engenharia Industrial (DEI) da PUC-Rio, pela ajuda e cobrança exercidos no período de elaboração desta dissertação.

Ao meu coorientador Fernando Aiube, professor do DEI e funcionário da Petrobras, pelos ensinamentos e ajuda na implementação dos modelos.

Ao professor do DEI, Paulo Henrique Soto Costa, pelos auxílios e ensinamentos durante a graduação e o mestrado.

Ao professor Fabio Rodrigo Siqueira Batista, do DEI e do Centro de Pesquisas de Energia Elétrica (CEPEL), por aceitar participar da banca avaliadora desta dissertação.

Ao professor da Universidade Federal de Pernambuco e Superintendente da Agência Nacional de Petróleo e Gás Natural e Biocombustíveis (ANP), Florival Rodrigues Carvalho, por aceitar fazer parte da banca avaliadora.

Ao professor doutor Krongnon Wailamer de Souza Regueira e funcionário da Superintendência de Planejamento e Pesquisa da ANP pelo seu interesse em fazer parte da banca avaliadora desta dissertação.

À PUC-Rio e ao CNPq pela bolsa de mestrado concedida.

Resumo

Ferraz, Rafael de Sequeira Baptista. Baidya, Tara Keshar Nanda. **Estimativa de Preços de Contratos Futuros Sobre Petróleo Utilizando o Método do Filtro de Kalman.** Rio de Janeiro, 2009. 173p. Dissertação de Mestrado – Departamento de Engenharia Industrial, Pontifícia Universidade Católica do Rio de Janeiro.

O Mercado Futuro adquire cada vez mais importância no cenário das Finanças Corporativas mundiais. O interesse principal das empresas neste segmento das finanças é a necessidade de proteção contra a volatilidade dos mercados financeiros. Neste sentido, uma das *commodities* mais negociadas é o petróleo. A dificuldade em precificar os contratos futuros do barril faz com que muitos modelos sejam criados para demonstrar a evolução dos preços ao longo do tempo. A utilização de processos estocásticos para representar possíveis trajetórias das séries temporais vem alcançando cada vez mais notoriedade, pois incorpora a aleatoriedade nas análises. O presente trabalho pretende testar a eficácia do modelo proposto na previsão do preço dos contratos futuros um passo à frente, ou seja, em estimar o preço para certa data na data imediatamente anterior. Neste sentido, o objetivo do estudo é coerente com o principal objetivo da análise de séries temporais que é construir modelos capazes de realizar previsões. Além disso, será estimado o preço à vista, variável a qual não é observável no mercado, e posteriormente serão confrontados os valores obtidos com uma *proxy*. O preço do mercado *spot* possui utilidade para os *traders* que necessitam obter o valor de um ativo que não é transacionado desta forma em bolsa. As estimativas dos parâmetros dos processos estocásticos serão feitas através de uma ferramenta estatística que passou a ser muito utilizada em modelos financeiros, o Filtro de Kalman. O procedimento consistirá em adotar um modelo de processo estocástico consagrado para uma série de preços de contratos futuros de uma *commodity*, estimando seus parâmetros e variáveis de estado com o Filtro, utilizando-o para previsão dos preços dos contratos futuros e para estimar o preço à vista, e posteriormente confrontando as estimativas e os valores reais coletados do mercado. Desta forma, se avaliará a capacidade do modelo em se adequar a novas mudanças estruturais na série. As ferramentas serão sempre explicitadas de

maneira acessível, demonstrando cada passo tomado e sempre que possível fazendo paralelo com outros conhecimentos mais básicos.

Palavras-chave

Mercado Futuro; Processo Estocástico; Filtro de Kalman; Séries Temporais; Previsão

Abstract

Ferraz, Rafael de Sequeira Baptista. Baidya, Tara Keshar Nanda. **Estimation of Petroleum Future Contracts Using The Kalman Filter Method.** Rio de Janeiro, 2009. 173p. MSc. Dissertation – Departamento de Engenharia Industrial, Pontifícia Universidade Católica do Rio de Janeiro.

The Future Market is becoming increasingly important in the global scenario of Corporate Finance. The main interest in this segment of finance is the need of being protected against the volatility of financial markets. Accordingly, one of the most traded *commodity* is oil. Because of difficulty in determine the value of future contracts on oil barrel, many models were created to demonstrate the evolution of their prices over time. The use of stochastic processes to represent possible trajectories of the time series is reaching more and more notoriety because it incorporates the randomness in the analysis. This study seeks to test the effectiveness of the proposed model in predicting the price of future contracts one step ahead, i.e. to estimate the price for a certain date on the preceding date. Consequently, the objective of the study is consistent with the primary objective of time series analysis that is to build models capable of forecasting. Furthermore, it will be estimated the spot price, variable that is not observable in the market, then the values obtained will be faced with a "proxy". The spot price is useful for traders who need to obtain the value of an asset that is not transacted in this way at the exchange. Estimates of the parameters of stochastic processes will be made through a statistical tool that has become widely used in financial models, the Kalman filter. The procedure is to adopt well known model of stochastic process for a series of prices of commodity futures contracts, estimating its parameters and state variables with the filter, using it to forecast the prices of future contracts and to estimate the spot price, and later comparing the estimates and the real values collected from the market. Thus, it will be evaluated ability of the model to fit the new structural changes in the time series. The tools will always be explained in an accessible way, demonstrating each step and where possible making parallel with other more basic knowledge.

Keywords

Future Market; Stochastic Process; Kalman Filter; Time Series; Forecast

Sumário

Lista de Figuras

Lista de Tabelas

Lista de Símbolos e Abreviações

a.a.	Ao ano
bbl(s)	Volume em Barril ou Barris
E&P	Exploração e Produção de Petróleo
f.d.p.	Função Densidade de Probabilidade
FK	Filtro de Kalman
MME	Medida de Martingal Equivalente
MAB	Movimento Aritmético Browniano
MAE	Erro Absoluto Médio
MAPE	Erro Absoluto Médio Percentual
MB	Movimento Browniano
MGB	Movimento Geométrico Browniano
MRM	Movimento de Reversão à Média
MQO	Mínimos Quadrados Ordinários
NYMEX	New York Mercantile Exchange
RMSE	Raiz do Erro Quadrático Médio
S&S	Schwartz e Smith
US$	Dólares Norte Americanos
WTI	West Texas Intermediate
*	Variável sob Precesso Neutro ao Risco
$\boldsymbol{x}_{\boldsymbol{t}^+}$	Vetor Calculado Após Observação da data t
$\boldsymbol{x}_{\boldsymbol{t}^-}$	Vetor previsto para a data t
$\boldsymbol{P}_{\boldsymbol{t}^+}$	Matriz de Covariâncias dos erros de previsão para o vetor de estado após obter-se a informação de t
$\boldsymbol{P}_{\boldsymbol{t}^-}$	Matriz Calculada Antes da Observação da data t
$\boldsymbol{K}_{\boldsymbol{t}}$	Vetor de Ganhos de Kalman calculado após observação de t
$\boldsymbol{z}_{\boldsymbol{t}}$	Vetor de variáveis observadas em t
$\boldsymbol{x}_{\boldsymbol{t}}$	Vetor de variáveis de estado estimado para t
$\boldsymbol{H}_{\boldsymbol{t}}$	Matriz de hiperparâmetros
$\boldsymbol{u}_{\boldsymbol{t}}$	Distúrbios aleatórios em t
$\boldsymbol{\phi}$	Matriz que apresenta a evolução das variáveis de estado no tempo
$\boldsymbol{G}$	Matriz da dinâmica dos erros das equações de estado

$N(\mu,\sigma^2)$	Distribuição Normal com Média μ e variância σ^2
X_{t2}	Valor Observado da Variável Independente Dois na Data t
$f(y,\theta)$	f.d.p. onde y representa os valores da série e θ os parâmetros
$L(\hat{\theta},y)$	Função de Verossimilhança

Lista de Equações

$$F = Se^{((r+u-\delta)(T-t))} \qquad (1)$$

$$B(S, \delta, \tau) \qquad (2)$$

$$\frac{dS}{S} = \mu dt + \sigma_1 dz_1 \qquad (3)$$

$$d\delta = k(\alpha - \delta)dt + \sigma_2 dz_2 \qquad (4)$$

$$\frac{1}{2}B_{SS}S^2{\sigma_1}^2 + \frac{1}{2}B_{\delta\delta}\sigma_2^2 + B_{S\delta}S\rho\sigma_1\sigma_2 + B_S S(r-\delta) +$$
$$B_\delta[k(\alpha-\delta) - \lambda\sigma_2] - B_\tau - rB = 0 \qquad (5)$$

$$ln(S_t) = \chi_t + \xi_t \qquad (6)$$

$$d\chi_t = -k\chi_t dt + \sigma_\chi dz_\chi \qquad (7)$$

$$d\xi_t = \mu_\xi dt + \sigma_\xi dz_\xi \qquad (8)$$

$$dz_\chi dz_\xi = \rho_{\chi\xi} dt \qquad (9)$$

$$E[(\chi_t, \xi_t)] = \left[e^{-kt}\chi_0, \xi_0 + \mu_\xi t\right] \qquad (10)$$

$$Cov[(\chi_t, \xi_t)] = \begin{bmatrix} (1-e^{-2kt})\frac{\sigma_\chi^2}{2k} & (1-e^{-kt})\frac{\rho_{\chi\xi}\sigma_\chi\sigma_\xi}{k} \\ (1-e^{-kt})\frac{\rho_{\chi\xi}\sigma_\chi\sigma_\xi}{k} & \sigma_\xi^2 t \end{bmatrix} \qquad (11)$$

$$E[ln(S_t)] = e^{-kt}\chi_0 + \xi_0 + \mu_\xi t \qquad (12)$$

$$Var[ln(S_t)] = (1-e^{-2kt})\frac{\sigma_\chi^2}{2k} + \sigma_\xi^2 t + 2(1-e^{-kt})\frac{\rho_{\chi\xi}\sigma_\chi\sigma_\xi}{k} \qquad (13)$$

$$l\,n(E[S_t]) = e^{-kt}\chi_0 + \xi_0 + \mu_\xi t + \frac{1}{2}(1-e^{-2kt})\frac{\sigma_\chi^2}{2k} + \sigma_\xi^2 t +$$
$$2(1-e^{-kt})\frac{\rho_{\chi\xi}\sigma_\chi\sigma_\xi}{k} \qquad (14)$$

$$d\chi_t = \left(-k\chi_t - \lambda_\chi\right)dt + \sigma_\chi dz_\chi^* \qquad (15)$$

$$d\xi_t = \left(\mu_\xi - \lambda_\xi\right)dt + \sigma_\xi dz_\xi^* \qquad (16)$$

$$dz_\chi^* dz_\xi^* = \rho_{\chi\xi} dt \qquad (17)$$

$$E^*[(\chi_t, \xi_t)] = \left[e^{-kt}\chi_0 - (1-e^{-kt})\frac{\lambda_\chi}{k}, \xi_0 + \mu_\xi^* t\right] \qquad (18)$$

$$Cov^*[(\chi_t, \xi_t)] = Cov[(\chi_t, \xi_t)] \qquad (19)$$

$$E^*[ln(S_t)] = e^{-kt}\chi_0 - (1-e^{-kt})\frac{\lambda_\chi}{k} + \xi_0 + \mu_\xi^* t \qquad (20)$$

$$Var^*[ln(S_t)] = Var[ln(S_t)] \qquad (21)$$

$$ln\left(F_{T,0}\right) = ln(E^*[S_T]) \qquad (22)$$

$$ln\left(F_{T,0}\right) = e^{-kT}\chi_0 + \xi_0 + A(T) \tag{23}$$

$$A(T) = \mu_\xi^* T - (1 - e^{-kT})\frac{\lambda_\chi}{k} + \frac{1}{2}((1 - e^{-2kT})\frac{\sigma_\chi^2}{2k} +$$
$$\sigma_\xi^2 T + 2(1 - e^{-kT})\frac{\rho_{\chi\xi}\sigma_\chi\sigma_\xi}{k})$$

$$\boldsymbol{x_t} = \boldsymbol{\phi x_{t-1}} + \boldsymbol{G v_t} \tag{24}$$

$$\begin{bmatrix}\chi_t \\ \xi_t\end{bmatrix} = \begin{bmatrix}0 \\ \mu_\xi^* \Delta t\end{bmatrix} + \begin{bmatrix}e^{-k\Delta t} & 0 \\ 0 & 1\end{bmatrix}\begin{bmatrix}\chi_{t-1} \\ \xi_{t-1}\end{bmatrix} + \begin{bmatrix}v_{1t} \\ v_{2t}\end{bmatrix} \tag{25}$$

$$Cov[(\chi_{\Delta t}, \xi_{\Delta t})] = \begin{bmatrix}(1 - e^{-2kt})\frac{\sigma_\chi^2}{2k} & (1 - e^{-kt})\frac{\rho_{\chi\xi}\sigma_\chi\sigma_\xi}{k} \\ (1 - e^{-kt})\frac{\rho_{\chi\xi}\sigma_\chi\sigma_\xi}{k} & \sigma_\xi^2 t\end{bmatrix} \tag{26}$$

$$\boldsymbol{z_t} = \boldsymbol{H_t x_t} + \boldsymbol{u_t} \tag{27}$$

$$\begin{bmatrix}ln\,F_{T_1} \\ \vdots \\ ln\,F_{T_n}\end{bmatrix} = \begin{bmatrix}A(T_1) \\ \vdots \\ A(T_n)\end{bmatrix} + \begin{bmatrix}e^{-kT_1} & 1 \\ \vdots & \vdots \\ e^{-kT_n} & 1\end{bmatrix}\begin{bmatrix}\chi_t \\ \xi_t\end{bmatrix} + \begin{bmatrix}u_{1t} \\ \vdots \\ u_{nt}\end{bmatrix} \tag{28}$$

$$\boldsymbol{K_t} = \boldsymbol{P_{t^-} H_t'}[\boldsymbol{H_t P_{t^-} H_t'} + \boldsymbol{R_t}]^{-\boldsymbol{1}} \tag{29}$$

$$\boldsymbol{P_{t^+}} = [\boldsymbol{I} - \boldsymbol{K_t H_t}]\boldsymbol{P_{t^-}} \tag{30}$$

$$\boldsymbol{x_{t^+}} = \boldsymbol{x_{t^-}} + \boldsymbol{K_t}[\boldsymbol{z_t} - \boldsymbol{H_t x_{t^-}}] \tag{31}$$

$$\boldsymbol{P_{(t+1)^-}} = \boldsymbol{\phi P_{t^+} \phi'} + \boldsymbol{GQG'} \tag{32}$$

$$\boldsymbol{x_{(t+1)^-}} = \boldsymbol{\phi x_{t^+}} \tag{33}$$

$$dx = \alpha dt + \sigma dz \therefore dz = \varepsilon\sqrt{dt} \therefore \varepsilon \sim N(0,1) \tag{34}$$

$$E[dx] = E[\alpha dt + \sigma dz] = \alpha dt + \sigma E[dz] = \alpha dt \tag{35}$$

$$Var[dx] = \sigma^2 dt \tag{36}$$

$$dx \sim N(\alpha dt, \sigma^2 dt) \tag{37}$$

$$dx = a(x,t)dt + b(x,t)dz \tag{38}$$

$$dx = \alpha x dt + \sigma x dz \tag{39}$$

$$E[x_t] = x_0 \mathrm{e}^{\alpha T} \tag{40}$$

$$Var[x_t] = {x_0}^2 \mathrm{e}^{2\alpha T + \sigma^2 T} - {x_0}^2 \mathrm{e}^{2\alpha T} = {x_0}^2 \mathrm{e}^{2\alpha T}\left(\mathrm{e}^{\sigma^2 T} - 1\right) \tag{41}$$

$$dx = \eta(\bar{x} - x)dt + \sigma dz \tag{42}$$

$$Var[x_t] = \frac{\sigma^2}{2\eta}(1 - e^{-2\eta t}) \tag{43}$$

$$\boldsymbol{b_t} = (\boldsymbol{X_t}'\boldsymbol{X_t})^{-1}\boldsymbol{X_t}'\boldsymbol{Y_t} \tag{44}$$

$$Y_t = \beta_1 + \beta_2 X_{t2} + \beta_3 X_{t3} + \cdots + \beta_k X_{tk} + \varepsilon_t \tag{45}$$

$$\boldsymbol{Y_t} = \boldsymbol{X_t b_t} + \boldsymbol{e_t} \tag{46}$$

$$\boldsymbol{b_{t+1}} = \boldsymbol{b_t} + \boldsymbol{K_{t+1}}(\boldsymbol{Y_{t+1}} - \boldsymbol{x_{t+1}'}\boldsymbol{b_t}) \tag{47}$$

$$\boldsymbol{K_{t+1}} = \frac{(\boldsymbol{X_t'X_t})^{-1}\boldsymbol{x_{t+1}}}{\boldsymbol{I}+\boldsymbol{x_{t+1}'}(\boldsymbol{X_t'X_t})^{-1}\boldsymbol{x_{t+1}}} \quad (48)$$

$$\boldsymbol{P_t} = (\boldsymbol{X_t'X_t})^{-1} \quad (49)$$

$$\boldsymbol{P_{t+1}} = \boldsymbol{P_t} - \boldsymbol{P_t}\frac{\boldsymbol{x_{t+1}x_{t+1}'}}{\boldsymbol{I}+\boldsymbol{x_{t+1}'P_t x_{t+1}}}\boldsymbol{P_t} \quad (50)$$

$$f(y,\theta) = \prod_{t=1}^{T} f(y_t,\theta) \quad (51)$$

$$\hat{\mu} = \frac{\sum_{t=1}^{T} y_t}{T} \quad (52)$$

$$\widehat{\sigma^2} = \frac{\sum_{t=1}^{T}(y_t-\mu)^2}{T} \quad (53)$$

$$E(\boldsymbol{z}_t|t-1) = \boldsymbol{H_t x_{t^-}} \quad (54)$$

$$\boldsymbol{P_{t^-}} = E[(\boldsymbol{x_t} - \boldsymbol{x_{t^-}})(\boldsymbol{x_t} - \boldsymbol{x_{t^-}})'] \quad (55)$$

$$\boldsymbol{R_t} = E[\boldsymbol{u_t u_t}'] \quad (56)$$

$$\boldsymbol{F}_t = \boldsymbol{H_t P_{t^-} H_t'} + \boldsymbol{R_t} \quad (57)$$

$$L(\boldsymbol{y};\ \boldsymbol{\theta}) = -\frac{T}{2} ln[2\pi] - \frac{1}{2}\sum_{t=1}^{T}|\boldsymbol{F}_t| - \frac{1}{2}\sum_{t=1}^{T} \boldsymbol{v}_t' \boldsymbol{F}_t^{-1}\boldsymbol{v}_t \quad (58)$$

$$\Delta t = \frac{1}{52} \quad (59)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(F_{oi}-F_{ei})^2}{n}} \quad (60)$$

$$MAE = \frac{\sum_{i=1}^{n}|F_{oi}-F_{ei}|}{n} \quad (61)$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{F_{oi}-F_{ei}}{F_{oi}}\right| \quad (62)$$

$$JB = \frac{n}{6}\left(S^2 + \frac{(K-3)^2}{4}\right) \quad (63)$$