

9º CONGRESSO BRASILEIRO DE PESQUISA E DESENVOLVIMENTO EM PETRÓLEO E GÁS



9º CONGRESSO
Brasileiro de P&D em
PETRÓLEO E GÁS

Maceió, AL
de 09 a 11 de novembro
2017

Realização:



ABPG
ASSOCIAÇÃO BRASILEIRA DE P&D EM
PETRÓLEO E GÁS



TÍTULO DO TRABALHO:

Prognóstico para preços do petróleo bruto utilizando a transformada wavelet discreta e máquinas de vetores de suporte

AUTORES:

Bruno Carlos Vieira dos Santos; Lucas Costa Barreto; Tiago Patrício da Silva; Samuel Vitor Saraiva; Frede de Oliveira Carvalho

INSTITUIÇÃO:

Universidade Federal de Alagoas - UFAL

Este Trabalho foi preparado para apresentação no 8º Congresso Brasileiro de Pesquisa e Desenvolvimento em Petróleo e Gás - 9º PDPETRO, realizado pela Associação Brasileira de P&D em Petróleo e Gás - ABPG, no período de 09 a 11 de novembro de 2017, em Maceió/AL. Esse Trabalho foi selecionado pelo Comitê Científico do evento para apresentação, seguindo as informações contidas no documento submetido pelo(s) autor(es). O conteúdo do Trabalho, como apresentado, não foi revisado pela ABPG. Os organizadores não irão traduzir ou corrigir os textos recebidos. O material conforme, apresentado, não necessariamente reflete as opiniões da Associação Brasileira de P&D em Petróleo e Gás. O(s) autor(es) tem conhecimento e aprovação de que este Trabalho seja publicado nos Anais do 9º PDPETRO.

PROGNÓSTICO PARA PREÇOS DO PETRÓLEO BRUTO UTILIZANDO A TRANSFORMADA WAVELET DISCRETA E MÁQUINAS DE VETORES DE SUPORTE

Abstract

The price of crude oil plays a significant role in the economy globally, as well as being a major factor affecting government plans and the commercial sectors. Estimating the price of crude oil is one of the most important issues facing economists interested in the energy sector, given that knowledge about future price fluctuations can lead to better decisions in different management spheres. As a result of this stimulus, the present work suggests the investigation of a hybrid model, based on intelligent systems techniques and signal processing, to forecast crude oil prices. The model discussed uses the discrete *wavelet* transform (TWD) as a preprocessing tool and the support vector machines (MVS) for modeling the system. The results reveal that the discrete *wavelet* transform enhances the prediction ability of the (MVS) and that the hybrid model is able to make predictions with a given confidence level for crude oil prices a week ahead.

Introdução

O petróleo é na atualidade uma das principais fontes energéticas, tal composto é oriundo da decomposição de vegetais e animais que se depositam no interior de lagos e mares, com o passar das eras sedimentos vão confinando os resíduos, que por ação das elevadas pressões e temperaturas acaba por se tornar o petróleo propriamente dito. Depois de formado o petróleo, que é uma mistura de uma infinidade de compostos em diferentes estados físicos, este se confina em rochas porosas que configuram as jazidas de petróleo (IGNÁCIO, 2017).

O valor do petróleo ao redor do mundo é ditado por uma relação de oferta e procura apresentando uma grande complexidade devido à enorme gama de variáveis que ditam tal relação. Variáveis essas que são sensíveis por algo repentino como o início de uma guerra em determinado território, um conflito ideológico/político que se agrava em um determinado país ou até mesmo a descoberta de uma nova fonte energética. Assim a tarefa de prever um cenário futuro para as variações do preço do petróleo é uma tarefa desafiadora (E SILVA et al, 2010).

Sendo assim, técnicas que busquem realizar predições, partindo de uma série de dados não lineares, são desejáveis. Para tal, aparatos computacionais, dentre os quais se destaca a modelagem, fazem uso de um conjunto complexo de dados de natureza caótica, fractais ou de múltiplas escalas, que se acumulam empiricamente na literatura financeira, para descobrir tendências e com isso poder gerar um prognóstico para eventos futuros (ZOU, 2015).

Para tal fim, algumas técnicas têm sido reportadas por diversos autores na literatura, como nos trabalhos de tendo destaque as baseadas em sistemas inteligentes com o uso de máquinas de aprendizado e estatística avançada, dentre as quais as Redes Neurais Artificiais (RNA) e as Máquinas de Vetores de Suporte (MVS) têm se mostrado bastante promissoras. Além disso, métodos de processamento de sinais, como a transformada *wavelet* discreta vêm sendo utilizadas em conjunção com essas técnicas com o intuito de melhorar a eficiência do prognóstico realizado.

Particularmente, uma das vantagens da MVS é sua fundamentação no princípio de minimização do risco estrutural, que garante a solução ótima do problema. Tal fato não se consolida com o uso das RNA, uma vez que sua operação está baseada na minimização do risco empírico e,

consequentemente, tenta encontrar um mínimo de uma função de erro e não garante, de fato, a descoberta do mínimo global da função (MADGE, 2015).

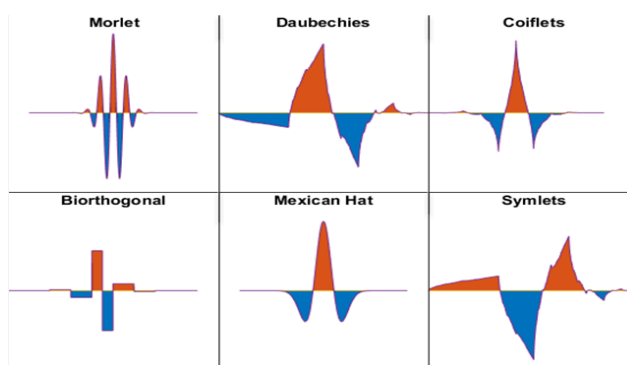
Dessa forma, este trabalho investiga a utilização da transformada *wavelet* discreta em conjunção com as máquinas de vetores de suporte para obtenção de prognósticos referentes ao preço do petróleo cru uma semana à frente. Para tal, a transformada *wavelet* será usada como método de pré-processamento dos dados da série temporal de preços com o intuito de avaliar a capacidade da técnica em melhorar a habilidade da MVS no prognóstico realizado. Os dados utilizados são registros semanais dos preços de petróleo cru do West Texas Intermediate (WTI), que servem como referência para compra e venda através de contratos, em várias localidades do mundo.

Metodologia

Transformada *wavelet*

A transformada *wavelet* é uma ferramenta de processamento de sinais capaz de reproduzir qualquer conjunto de dados localizados no domínio do tempo e mapeá-lo em um novo domínio, o do tempo e da frequência. A rigor, uma *wavelet* é uma função em forma de onda, que possui energia finita e área igual à zero sob o gráfico que a representa. Dentro desse contexto, algumas funções obedecem a esses critérios e podem ser visualizadas na Figura 1.

Figura 1: Exemplos de *wavelets*. Fonte: Krishna (2013)



A equação 1 representa a transformada *wavelet* contínua de um sinal $f(t)$ qualquer, localizado no domínio do tempo, na qual $\Psi_{a,b}(t)$ é a função *wavelet* escolhida para realizar o mapeamento, denominada de *wavelet*-mãe, e os parâmetros a e b , são para o ajuste de seu tamanho e posição no tempo, conhecidos como fatores de escala e translação. Através dos parâmetros a e b é possível a obtenção de várias versões da *wavelet*-mãe escolhida, que por sua vez serão superpostas para representar o conjunto de dados no novo domínio.

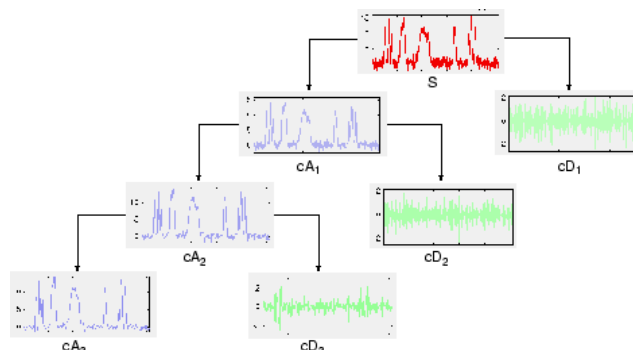
$$W_f(a,b) = \int_{-\infty}^{+\infty} f(t)\Psi_{a,b}(t)dt \quad (1)$$

Além da transformada *wavelet* contínua (TWC), a literatura traz outra alternativa: a transformada *wavelet* discreta (TWD). Na transformada *wavelet* contínua, os parâmetros a e b são variados continuamente, isto é, podem assumir quaisquer valores, ao passo que na transformada *wavelet* discreta, a escolha desses parâmetros se dá baseada em potências de dois, o que acelera o processo, sem que haja redundância nos dados ou perda de acuracidade na representação do sinal.

A TWD, em particular, realiza a decomposição do sinal original em outros sinais de frequências diferentes; estes quando combinados resultam no sinal original. As componentes de mais baixa frequência dessa decomposição são chamadas de aproximações e as de mais alta frequência de detalhes. A quantidade de componentes geradas depende do nível de decomposição escolhido. A

transformada no primeiro nível produz uma componente de detalhe (D1) e outra de aproximação (A1). Se esse processo continuar, dessa vez a componente de aproximação A1 será decomposta em um sinal de aproximação A2 e outro de detalhe D2. Dessa forma, a decomposição pode prosseguir até que se tenha apenas um único dado. Todo esse procedimento pode ser representado na Figura 2.

Figura 2: Decomposição *wavelet* de um sinal s . Fonte: http://radio.feld.cvut.cz/matlab/toolbox/wavelet/ch01_i21.html



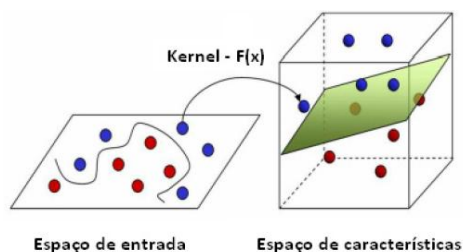
Nesse estudo, a série semanal de preços do petróleo cru foi decomposta até o terceiro nível utilizando a *wavelet* de Meyer. As componentes resultantes da transformada foram usadas como entradas para a máquina de vetores de suporte para realização da previsão do preço desejado, uma semana à frente.

Máquinas de Vetores de Suporte

As Máquinas de Vetores de Suporte (MVS) são baseadas no conceito de fronteira ou limite de decisão e, ao contrário de outras máquinas de aprendizado, que se baseiam no princípio da minimização do risco empírico, sua formulação se constitui no princípio da minimização do risco estrutural como base de funcionamento (BURGES, 1998). Uma fronteira de decisão é aquela que separa um conjunto de objetos como membros de diferentes classes. A partir disso, o que se deseja é realizar uma classificação de novos objetos de testes baseados em exemplos disponíveis ou objetos de treinamento.

De início, as MVS foram desenvolvidas para resolução de problemas de classificação binária, nos quais um conjunto de dados poderia ser repartido em duas classes através de um ponto ou uma reta. Posteriormente, uma metodologia foi desenvolvida para resolver problemas nos quais as classes do conjunto de dados não eram linearmente separáveis. O método consiste na utilização de funções de Kernel, as quais mapeiam o conjunto de dados em um espaço de maior dimensão, onde as classes do conjunto de dados podem ser linearmente separáveis e se pode construir um hiperplano para separação como mostrado na Figura 3 (REBELO, 2008). Há alguns tipos de funções de Kernel bastante difundidas, como por exemplo, a gaussiana, a linear e a polinomial.

Figura 3: Separação de classes não linearmente separáveis, com um hiperplano, através do mapeamento da função kernel em um espaço de características de maior dimensão. Fonte: Adaptado de Souza (2016).



Classificação com MVS

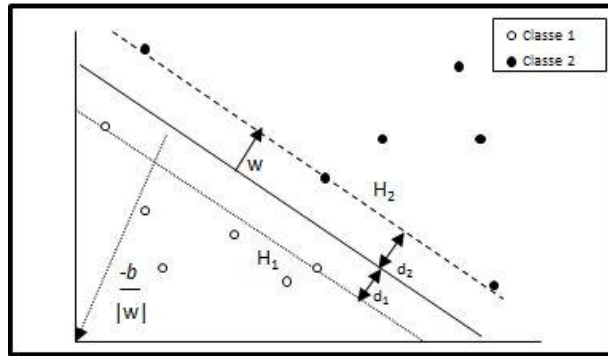
Para a classificação do conjunto de dados bidimensionais em duas classes, podemos supor L pontos de treinamento que sejam linearmente separáveis, isso significa dizer que uma linha pode ser traçada em um plano com eixos x_1 e x_2 separando essas, onde cada entrada x_i possua D atributos (isto é, com dimensão D) e possa ser classificado como $y_i = -1$ ou $y_i = +1$, assim o conjunto de dados de treinamento pode ser descrito como:

$$\{x_i, y_i\} \text{ onde } i = 1, \dots, L, y_i \in \{-1, 1\}, \quad x \in \mathbb{R}^D$$

Um hiperplano pode ser descrito por uma função $f(x) = w_0 + w_1x_1 + \dots + w_nx_n = 0$ ou matricialmente por $w^T + b = 0$ em que w_0 é uma constante descrita como b , por convenção (VAPNIK, 1995). Dada a Figura 4, w é um vetor de parâmetros normal ao hiperplano e $\frac{b}{\|w\|}$ é a distância perpendicular do hiperplano a origem.

Os vetores de suporte são os pontos de treinamentos mais próximos ao hiperplano e o objetivo das MVS é orientar o hiperplano de modo que ele fique o mais distante possível dos vetores de suporte de ambas as classes (FLETCHER, 2009).

Figura 4: Hiperplano e vetores de suporte. Fonte: Adaptado de Fletcher (2009).



Para construção de tal hiperplano, um algoritmo iterativo de treinamento é implementado visando obter a maior distância entre este hiperplano e os pontos às margens. Esse hiperplano pode ser calculado através da maximização das normas dos vetores de suporte. Assim, utilizando os dados de treinamento, a MVS separa objetos pertencentes às classes, enquanto tenta maximizar a margem entre elas.

Considerando as inequações para o caso desta classificação, onde $+1$ representa a classe 1 e -1 a classe 2, tem-se:

$$x_i \cdot w + b \geq +1, \text{ para } y_i = +1 \quad (2)$$

$$x_i \cdot w + b \leq -1, \text{ para } y_i = -1 \quad (3)$$

$$y_i(x_i \cdot w + b) - 1 \leq 0 \quad \forall_i \quad (4)$$

Se considerarmos somente os vetores de suporte da Figura 7, então os dois planos que estão sobre os pontos citados, H_1 e H_2 podem ser descritos por:

$$x_i \cdot w + b \geq 1, \text{ para } H_1 \quad (5)$$

$$x_i \cdot w + b \leq -1, \text{ para } H_2 \quad (6)$$

Ainda na figura 7 pode-se definir d_1 e d_2 como as distâncias de H_1 e H_2 , respectivamente, ao hiperplano de separação. Estas equidistâncias somadas formam o comprimento da margem da MVS. O hiperplano ótimo para separação seria então aquele em que há a maior distância entre este e os vetores de suporte e define a região de segurança da classificação de dados desconhecidos. Esse Hiperplano pode ser calculado com a maximização das normas dos vetores (NOBLE, 2006).

Para orientar o hiperplano a ser o mais distante possível dos vetores suportes é necessário maximizar a margem ($d_1 + d_2$). A geometria de vetores simples mostra que a margem corresponde a $\frac{1}{\|w\|}$, e maximizando o termo sujeito a restrição (3), tem-se:

$$\text{Minimize } \|w\|, \text{ tal que } y_i(x_i \cdot w + b) - 1 \leq 0 \quad \forall_i \quad (7)$$

Minimizar $\|w\|$ é equivalente a minimizar o termo $\frac{1}{2}\|w\|^2$, o que possibilita tornar a otimização solucionável por Programação Quadrática. Logo, tem-se:

$$\text{Minimize } \frac{1}{2}\|w\|^2, \text{ tal que } y_i(x_i \cdot w + b) - 1 \leq 0 \quad \forall_i \quad (8)$$

Dessa forma, para MVS considerando funções de Kernel para o mapeamento no espaço de características, tem-se:

$$\text{Minimize } \frac{1}{2}\|w\|^2 + C \sum_{i=1}^L \xi_i \quad (10)$$

$$\text{Sujeita a } y_i(\Phi(x_i) \cdot w + b) - 1 + \xi_i \leq 0 \quad \forall_i \quad (11)$$

Onde $K(x_i, x_j) = \Phi(x_i) \cdot \Phi(x_j)$. Isto é, a função kernel K , representa o produto interno dos dados de entrada mapeados em um espaço de característica maior, transformados por Φ .

Máquina de Vetores de Suporte de Regressão - MVSR

Além do problema de classificação de padrões, as máquinas de vetores de suporte também lidam com problemas de regressão de padrões altamente não lineares, utilizando-se do mesmo mapeamento em um espaço de maior dimensão. Quando os padrões são mapeados no novo espaço, basta realizar uma regressão linear.

Nesse trabalho, as componentes da decomposição *wavelet* e valores antigos da própria série são mapeados em um espaço de maior dimensão utilizando a função kernel polinomial. A resposta desejada é o preço do petróleo cru uma semana à frente do momento em que se obtém o prognóstico. Nesse caso, a regressão realizada pela MVSR relaciona o preço do petróleo em um passo de tempo a valores passados das componentes da decomposição, assim como também a preços de semanas anteriores. A Figura 4 ilustra o modelo da MVSR com as entradas atrasadas no tempo.

Os modelos aqui empregados foram construídos utilizando o ambiente MATLAB.

Resultados e Discussão

Primeiramente, para elaboração do modelo, foram realizados testes para a escolha da *wavelet-mãe* que seria utilizada na decomposição da série de preços de petróleo. O nível de decomposição escolhido foi igual a três, seguindo o critério de Nourani *et al* (2011). Dentre os testes realizados, MVS cujos dados de treinamento foram pré-processados pela *wavelet* de Meyer foi a que produziu o menor valor para o erro quadrático médio.

Uma vez escolhida a *wavelet*-mãe para a decomposição e o nível da transformação, três funções kernel foram comparadas levando em conta os valores do RMSE, coeficiente de determinação e Nash-Sutcliffe-Efficiency. Os resultados obtidos na sessão de validação dos três modelos utilizados estão representados nas Tabelas 2.

Tabela 2: Critérios de erro obtidos na fase de validação do modelo MVS-*wavelet*

Funções Kernel	RMSE (\$/Barril)	R ²	NSE
Gaussiana	3.1979	0.7192	0.7291
Polinomial	1.7342	0.9025	0.9203
Linear	2.1894	0.8614	0.8730

A partir disso, é possível ver que o modelo que utilizou a função de kernel polinomial para mapeamento dos dados no novo espaço foi o que retornou valores mais satisfatórios para os critérios de erro utilizados. Por último, o modelo da MVS em conjunção com a transformada *wavelet* foi comparado com outro modelo de MVS, utilizando a mesma função kernel de mapeamento e número de delays, porém sem o pré-processamento dos dados de entrada. A comparação entre os dois modelos está representada na Tabela 3.

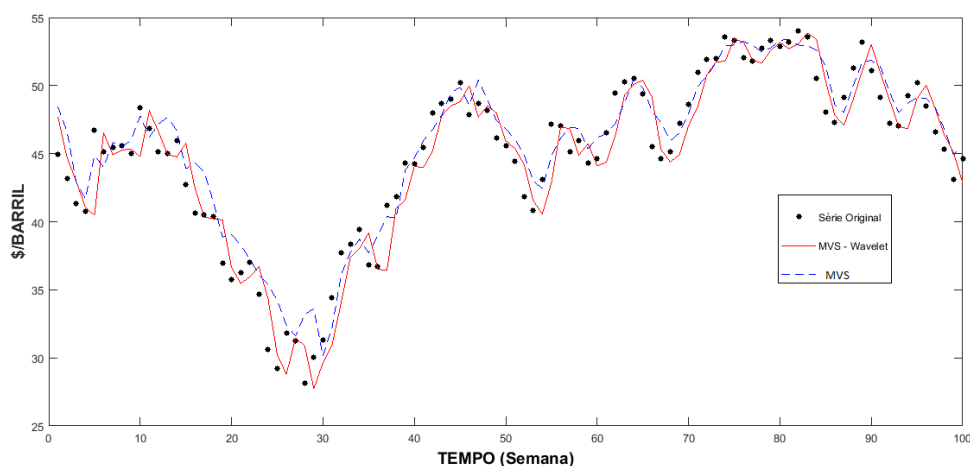
Tabela 4: Comparação entre o modelo de MVS tradicional e o modelo e a MVS em conjunção com a transformada *wavelet* discreta.

MODELO	RMSE (\$/Barril)	R ²	NSE
MVS - <i>wavelet</i>	1.7342	0.9025	0.9203
MVS	1.9559	0.9013	0.785

O número de delays, isto é, a quantidade de valores passados das componentes da transformada *wavelet* e dos preços em passos de tempos anteriores foi igual a 1 e escolhido de acordo com o menor RMSE.

Sendo assim, é fácil notar que o pré-processamento realizado para os dados de entrada da MVS potencializou a habilidade de prognóstico do modelo, uma vez que o modelo quando utilizado em conjunção com a transformada *wavelet* discreta apresentou menor valor para o RMSE e valores de R² e NSE mais próximos a 1, o que indica que o prognóstico obtido apresenta boa similaridade com os valores reais registrado em um horizonte de 100 semanas. A sobreposição das séries simuladas pelos modelos dois modelos e os valores reais do preço do petróleo bruto em um horizonte de 100 semanas está ilustrada na Figura x.

Figura 5: Sobreposição das séries simuladas e dos valores reais do preço de petróleo.



Conclusões

A partir dos testes realizados, foi possível verificar que a transformada wavelet discreta potencializou, de fato, a habilidade preditiva da rede neural frente à série temporal de preços de gás natural. O modelo desenvolvido retornou medidas de erro relativamente pequenas, o que permite afirmar que este foi capaz de realizar o prognóstico desejado, se comportando bem no horizonte de previsão testado. Assim, tal estratégia aliada a outras metodologias como, por exemplo, a análise técnica, pode compor um sistema de apoio à tomada de decisão dentro de um sistema de gestão que aumente a precisão das estimativas para os preços do gás natural do *Henry Hub*.

Referências Bibliográficas

IGNÁCIO, L.V. R.et al. O uso de inteligência artificial para a previsão do preço do petróleo. **Revista Espacios**, Vol. 38, n. 24, p. 3, 2017

E SILVA, E. G. de S.; LEGEY, L. FL; E SILVA, E. A. de S. Forecasting oil price trends using wavelets and hidden Markov models. **Energy Economics**, v. 32, n. 6, p. 1507-1519, 2010.

ZOU, Y.; YU, L.; HE, K. Wavelet entropy based analysis and forecasting of crude oil price dynamics. **Entropy**, v. 17, n. 10, p. 7167-7184, 2015.

MADGE, S. Predicting Stock Price Direction using Support Vector Machines. **Independent Work Report Spring**, 2015.

BURGES, J. C. C.. A Tutorial on Support Vector Machines for Pattern Recognition. **Data Mining and Knowledge Discovery**, V. 2, n 2, pp 121–167, 1998.

FLETCHER, T. Support Vector Machines Explained. 2009. Disponível em: <<http://www.tristanfletcher.co.uk/SVM%20Explained.pdf>>. Acesso em: 01/08/2017.

NOBLE, W. S. What is a support vector machine?. **Nature biotechnology**, v. 24, n. 12, p. 1565–7, 2006. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/17160063>>. Acesso em: 7/08/2017.

NOURANI, V.; KISI, Ö.; KOMASI, M. Two hybrid Artificial Intelligence approaches for modeling rainfall–runoff process. **Journal of Hidrology**, 2011.

ÖZDEMİR A.; KARAKAŞ S.; ÇAKMAK E.; TÜFEKÇİ D.; ARIKAN O. Time-frequency component analyser and its application to brain oscillatory activity. **Journal of Neuroscience Methods** 2005;145:107–125.