

Seleção de modelos e previsão de acréscimo de reservas de óleo e gás

Gustavo Melo

Industrial Engineering Department, Pontifical Catholic University of Rio de Janeiro – PUC-Rio,
CP38097, 22453-900 Rio de Janeiro – Brazil
gustavo.melo@labnexo.com

Jorge Montechiari

Industrial Engineering Department, Pontifical Catholic University of Rio de Janeiro – PUC-Rio,
CP38097, 22453-900 Rio de Janeiro – Brazil
jorge.montechiari@labnexo.com

Silvio Hamacher

Industrial Engineering Department, Pontifical Catholic University of Rio de Janeiro – PUC-Rio,
CP38097, 22453-900 Rio de Janeiro – Brazil
hamacher@puc-rio.br

Fernando Cyrino

Industrial Engineering Department, Pontifical Catholic University of Rio de Janeiro – PUC-Rio,
CP38097, 22453-900 Rio de Janeiro – Brazil
cyrino@puc-rio.br

RESUMO

Este artigo apresenta um método de análise estatística para avaliar a aplicabilidade da metodologia de Knoring et al. (1999) para previsão de acréscimo de reservas de óleo e gás em um conjunto de campos brasileiros. O autor propõe nove modelos para representar séries históricas de acréscimo de reservas, baseando-se na combinação de duas funções matemáticas: Seleção de Controle, que representa o ganho de conhecimento obtido com a evolução da exploração e Seleção de Condição, decrescente, à medida que o volume de novas descobertas tende a diminuir com o aumento da exploração. Neste artigo propõe-se um método de escolha do modelo mais adequado, fundamentado em critérios e análises estatísticas. Utilizando-se um período de cinco anos para testar a capacidade preditiva do modelo, o MAPE de previsão foi de 2,65%, com destaque para a previsão cinco anos à frente, com erro de 0,22%, confirmando a precisão do modelo escolhido pelo método utilizado.

PALAVRAS-CHAVE. Reservas de Óleo e Gás, Previsão, Séries Temporais.

ABSTRACT

This paper presents a statistical method to evaluation the application of the Knoring's et al (1999) methodology for forecast crude oil and natural gas reserves in the Brazilian context. The author proposes nine models to represent the Time Series of reserves growing based on a combination of two mathematical functions. The first, called the Control Selection Function represents the gain of knowledge obtained an area exploration. The second on, Condition Selection Function, is a decreasing function which indicates that the volume of new discoveries tends to fall as the exploratory process advances. As there are nine possible models, the task of choose one of them is the propose of this paper, using an Information Criterion and statistical tests. Using a five years period for the out-of-sample forecast, the obtained MAPE was 2.65%, whereupon the last year with only 0.22%, confirming the accuracy of the model and the choice method used.

KEY WORDS: Crude Oil and Natural Gas Reserves, Forecast, Time Series.

1. Introdução

Realizar uma boa estimativa da quantidade de hidrocarbonetos existentes nos reservatórios é questão primordial para as empresas exploradoras de petróleo. Soma-se a isso o fato de que o valor das empresas petrolíferas frente ao cenário econômico mundial é medido, principalmente, pela quantidade de reservas¹ de óleo e gás que têm sob suas posses, sendo os principais atrativos para a captação de recursos, além de servir de base para a elaboração do planejamento exploratório e orçamentário dessas empresas.

Neste sentido, diversos autores desenvolveram diferentes abordagens na tentativa de explicar o comportamento de uma curva de crescimento de reservas, ou de produção de petróleo, para uma determinada região. O trabalho de Hubbert (1956) passou a ser globalmente conhecido após o autor prever com sucesso o pico de produção de petróleo de 48 estados norte-americanos, que ocorreria 15 anos após suas previsões. A metodologia utilizada pelo autor é relativamente simples. Sua maior preocupação era, a partir de um valor pré-definido do volume total de óleo existente, determinar o pico de produção e o tempo que restava para a exaustão das reservas recuperáveis. A aplicação dessa metodologia era facilitada pela hipótese, assumida por Hubbert (1956), da simetria da curva em relação a essa produção máxima. Posteriormente, tal hipótese foi criticada por diversos outros autores. Segundo Deffeyes (2008), por exemplo, além da curva Logística utilizada por Hubbert (1956), as curvas Gaussiana e Lorentziana também podem ser ajustadas aos históricos de produção e acréscimo de reservas. Para Rosa & Gomes (2004), a maior dificuldade na aplicação do método é a definição do volume total de óleo, também denominado “últimas reservas recuperáveis”, existentes num país ou numa região.

De forma distinta de Hubbert (1956) e Deffeyes (2008), Knoring et al. (1999) afirmam não existir um modelo universal de curva para representar o crescimento de reservas e que as curvas que melhor se ajustam a dados históricos podem possuir um comportamento assimétrico, sugerindo nove modelos para representar séries observadas.

A questão passa a ser, então, a escolha do modelo, dentre os propostos, que melhor representa o comportamento de séries históricas de evolução de reservas. Nesse contexto, o objetivo deste artigo é propor um método fundamentado em critérios e análises estatísticas e aplicá-lo a um conjunto de campos produtores de petróleo do cenário brasileiro.

Na próxima seção será descrita a metodologia de ajuste e avaliação dos modelos considerados.

2. Metodologia de Knoring

Segundo Knoring et al. (1999), o processo de acréscimo de reservas (ΔR) pode ser matematicamente representado pelo produto de duas funções, sendo estas: função seleção de controle ($\varphi(R)$) e função seleção de condição ($f(R)$), como mostrado na Eq. (1).

$$\Delta R = \varphi(R) \times f(R) \quad \text{Eq. (1)}$$

A função seleção de controle representa o conhecimento obtido com a evolução da exploração de uma região e pode ser representada por uma única função crescente e linear, pois, de acordo com os autores, à medida que uma região é explorada, maior é o conhecimento e o

¹ “Reserva” é a parte conhecida, técnica e economicamente recuperável de um determinado recurso natural finito (SZKLO, MACHADO & SCHAEFFER, 2007).

controle sobre a mesma. Dessa forma, o controle em um instante $(t + 1)$ sempre será maior que o controle em um instante (t) .

A função seleção de condição $(f(R))$ é uma função decrescente que representa o volume associado a novas descobertas, e indica que esta quantidade tende a diminuir à medida que é feita a exploração de uma determinada região. Assim, a condição em um instante $(t + 1)$ será sempre menor que a condição no instante (t) . Para representar a função seleção de condição, Knoring et al. (1999) sugerem 9 modelos de funções, representadas pelas Eq. (2) a Eq. (10).

$$\text{Modelo 1: } \frac{\Delta R}{R_{t+1}} = a - bR_t \quad \text{Eq. (2)}$$

$$\text{Modelo 2: } \frac{\Delta R}{R_{t+1}} = 1 - aR_t^b \quad \text{Eq. (3)}$$

$$\text{Modelo 3: } \frac{\Delta R}{R_{t+1}} = e^{a-bR_t} \quad \text{Eq. (4)}$$

$$\text{Modelo 4: } \frac{\Delta R}{R_{t+1}} = 1 - \left(1 - c + aR_t^{\frac{1}{b}}\right)^b \quad \text{Eq. (5)}$$

$$\text{Modelo 5: } \frac{\Delta R}{R_{t+1}} = c - aR_t^b \quad \text{Eq. (6)}$$

$$\text{Modelo 6: } \frac{\Delta R}{R_{t+1}} = b \ln \frac{A}{R_t} \quad \text{Eq. (7)}$$

$$\text{Modelo 7: } \frac{\Delta R}{R_{t+1}} = a \left(\frac{A - R_t}{A}\right)^b \quad \text{Eq. (8)}$$

$$\text{Modelo 8: } \frac{\Delta R}{R_{t+1}} = e^{a-bR_t} - c \quad \text{Eq. (9)}$$

$$\text{Modelo 9: } \frac{\Delta R}{R_{t+1}} = e^{a-bR_t} - \frac{c}{R_t} \quad \text{Eq. (10)}$$

O comportamento geral das funções seleção de controle, seleção de condição e de acréscimo de reservas pode ser observado na Figura 1.

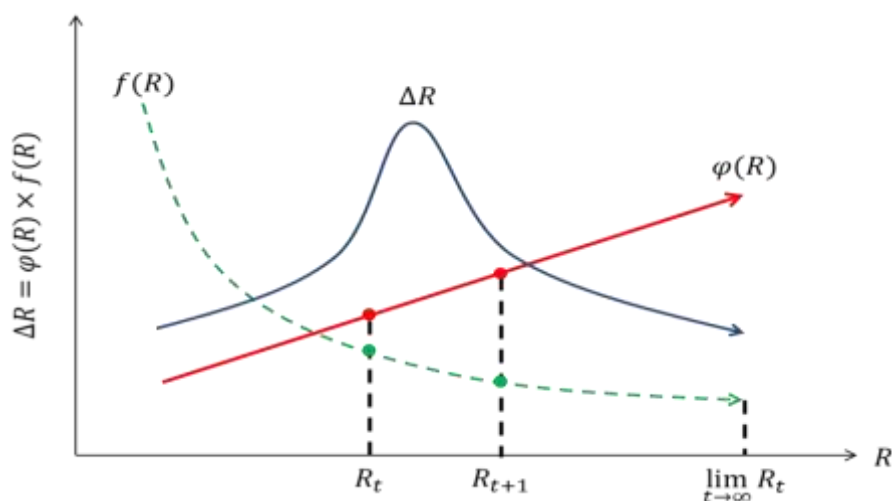


Figura 1 – Método de Knoring
Fonte: Adaptado de Knoring et al., 1999, p.20

Os nove modelos de função seleção de condição combinados ao modelo linear da função seleção de controle dão origem a nove modelos para representar a evolução das reservas com o tempo, representados pelas Eq. (11) a Eq. (19).

$$\text{Knoring 1: } VRec_{t+1} = \frac{VRec_t}{1-a+bVRec_t} \quad \text{Eq. (11)}$$

$$\text{Knoring 2: } VRec_{t+1} = aVRec_t^b \quad \text{Eq. (12)}$$

$$\text{Knoring 3: } VRec_{t+1} = \frac{VRec_t}{1-e^{a-bVRec_t}} \quad \text{Eq. (13)}$$

$$\text{Knoring 4: } VRec_{t+1} = \frac{VRec_t}{\left(1-c+aVRec_t^{\frac{1}{b}}\right)^b} \quad \text{Eq. (14)}$$

$$\text{Knoring 5: } VRec_{t+1} = \frac{VRec_t}{1-c+aVRec_t^b} \quad \text{Eq. (15)}$$

$$\text{Knoring 6: } VRec_{t+1} = \frac{VRec_t}{1-b \ln \frac{A}{VRec_t}} \quad \text{Eq. (16)}$$

$$\text{Knoring 7: } VRec_{t+1} = \frac{VRec_t}{1-c(A-VRec_t)^b} \quad \text{Eq. (17)}$$

$$\text{Knoring 8: } VRec_{t+1} = \frac{VRec_t}{1+c-e^{a-bVRec_t}} \quad \text{Eq. (18)}$$

$$\text{Knoring 9: } VRec_{t+1} = \frac{VRec_t}{1-\left(e^{a-bVRec_t} - \frac{c}{VRec_t}\right)} \quad \text{Eq. (19)}$$

A variável utilizada nas equações acima ($VRec$) refere-se ao conceito de volume recuperável, que, para um ano qualquer, quantifica as reservas que ainda podem ser recuperadas mais a produção acumulada até este ano, ou seja, todo o recurso descoberto comercialmente viável até então. A primeira diferença da série de volume recuperável resulta na série de acréscimo de reservas de um ano t para o ano $t + 1$, representando a quantidade de reservas provadas descobertas ou agregadas neste último ano. Essa afirmação é provada a seguir, utilizando-se o conceito de volume recuperável.

$$VRec_t = ProdAcum_t + R_t \quad \text{Eq. (20)}$$

Sendo:

- $ProdAcum_t$ – produção acumulada até o final do ano t ;
- R_t – reservas provadas ao final do ano t .

Consequentemente,

$$VRec_{t+1} = ProdAcum_{t+1} + R_{t+1} \quad \text{Eq. (21)}$$

Onde:

- $ProdAcum_{t+1} = ProdAcum_t + Prod_{t+1}$, sendo $Prod_{t+1}$ a produção no ano $t + 1$;
- $R_{t+1} = R_t + RDesc_{t+1} - Prod_{t+1}$, sendo $RDesc_{t+1}$ todo o volume de reservas descoberto no ano $t + 1$.

Dessa forma, pode-se concluir que:

$$VRec_{t+1} - VRec_t = RDesc_{t+1} \quad \text{Eq. (22)}$$

Logo, o cálculo da quantidade de reservas provadas descoberta ou agregada ao volume recuperável de uma região qualquer em um determinado ano $t + 1$ pode ser obtido aplicando a Eq. (22). Também é relevante dizer que séries de volume recuperável formam curvas assintóticas, visto que tendem a um valor máximo, chamado Últimas Reservas Recuperáveis (*URR*), que será alcançado quando não houver mais acréscimo de reservas à série em questão.

3. Análise Estatística

Utilizando-se a metodologia de Knoring et al. (1999), a questão passa a ser, após o ajuste dos modelos a uma série de volume recuperável, como proceder na escolha do melhor modelo para representá-la. Cabe aqui dizer que para ajustar esses modelos não lineares a séries históricas é necessário suporte computacional adequado.

Primeiramente, antes de realizar o ajuste, é preciso dividir a série em dois períodos, para que seja possível realizar a escolha baseada nos critérios propostos a seguir. O primeiro é denominado período *in sample*, utilizado para a estimação dos parâmetros dos modelos. O segundo é chamado período *out of sample*, utilizado para testar a capacidade preditiva dos mesmos, já que não é utilizado na estimação dos parâmetros. A partir destes dois períodos, estatísticas serão obtidas e utilizadas na aplicação do método de escolha proposto a seguir.

Após o ajuste, será averiguada a significância dos coeficientes de todos os modelos, através de um teste de significância dos parâmetros. Aqueles cujos coeficientes forem considerados estatisticamente insignificantes serão eliminados do processo de escolha.

Dando prosseguimento à análise, é preciso comparar, de alguma forma, os modelos restantes entre si. Dentre as diversas metodologias utilizadas para este fim propõe-se uma análise comparativa a partir de Critérios de Informação (*CI*), que selecionam, dado um conjunto de modelos possíveis para uma série real, aquele que minimiza, simultaneamente, a discrepância (entre os dados e os modelos) e a complexidade dos mesmos (Akaike, 1974). A complexidade mínima é justificada pelo Princípio da Parcimônia: sendo iguais outras características desejáveis, modelos mais simples, isto é, com menor número de parâmetros, devem ser preferidos.

Formalmente, os *CI's* são expressos pela relação geral da Eq. (23).

$$CI(k) = \text{discrepância} + \text{complexidade} \quad \text{Eq. (23)}$$

Akaike (1974) sugeriu o seguinte critério, conhecido como *AIC*:

$$AIC = \left(-\frac{2 \log L(\hat{\theta})}{T^*} \right) + \left(\frac{2k}{T^*} \right) \quad \text{Eq. (24)}$$

Em que:

- k - número de parâmetros independentes estimados;
- T^* - número de observações utilizadas na estimação do modelo;
- $L(\hat{\theta})$ - verossimilhança avaliada no máximo;

Posteriormente, Mills e Prasad (1992) provaram, através de simulação de Monte Carlo, que o critério representado pela Eq.(25), conhecido como *AIC* corrigido (*AICc*), proposto inicialmente por Hurvich e Tsai (1989), é mais adequado para séries com poucas observações.

$$AICc = (-2 \log L(\hat{\theta})) + (2k \frac{T^*}{T^* - k - 1}) \quad \text{Eq. (25)}$$

Segundo os autores, para um tamanho de amostra igual a 50, o $AICc$ indicou o verdadeiro modelo mais frequentemente do que os outros critérios, como o de Akaike (1974), o de Schwarz (1978) e o de Hannan-Quinn (1979). Esse é o caso da série em análise neste estudo, que possui 39 observações, como poderá ser visto na próxima seção. Contudo, para um tamanho de amostra de 100 ou mais, outros critérios, como o de Schwarz (1978), por exemplo, apresentaram melhor desempenho (Mills e Prasad, 1992).

Também será verificado se os resíduos gerados pelo modelo escolhido são correlacionados entre si. Para este diagnóstico será aplicado o teste de Ljung-Box (1978). Vale ressaltar que, embora não sejam estabelecidos pressupostos iniciais acerca da distribuição da série de resíduos, os mesmos estão sendo analisados com vista ao maior rigor estatístico da modelagem.

Caso exista autocorrelação nos resíduos, o modelo será considerado inadequado para descrever a série histórica em questão. Dessa forma, o próximo modelo a minimizar o critério $AICc$ será considerado o melhor para representá-la.

Para medir o poder de previsão do modelo escolhido no período *out of sample*, será utilizado o Erro Percentual Absoluto Médio ($MAPE$ – *Mean Absolute Percentage Error*), representado pela Eq. (26).

$$MAPE = \frac{1}{n} \sum_{t=1}^n \left| \frac{A_t - F_t}{A_t} \right| \quad \text{Eq. (26)}$$

Onde:

- A_t – valor real no período t ;
- F_t – valor previsto para o período t ;
- n – número de observações.

Como mede a discrepância média entre os valores previstos e os valores observados, quanto menor o valor do $MAPE$, melhor será a previsão. A seguir, a Figura 2 representa resumidamente os passos para a aplicação do método proposto para modelagem e previsão de séries de volume recuperável.

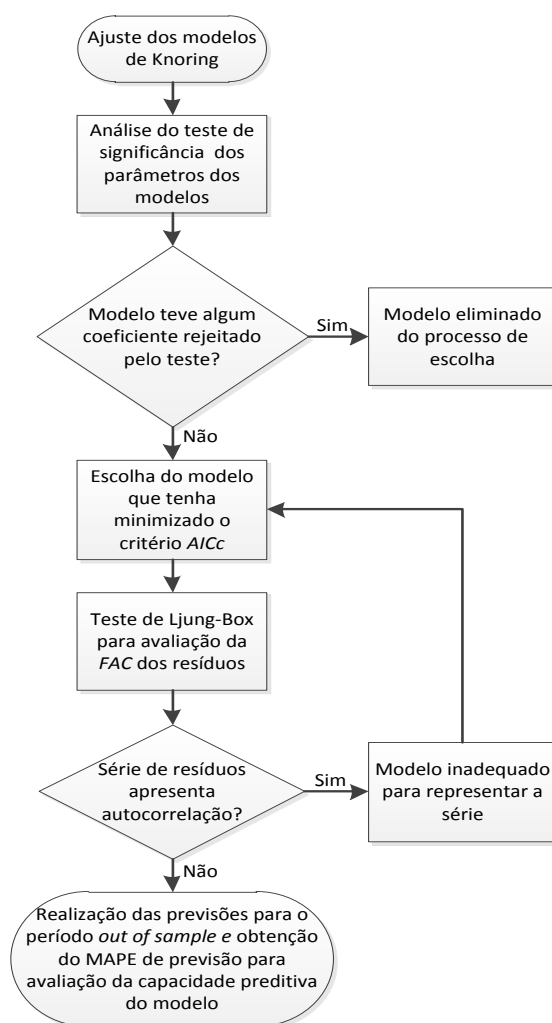


Figura 2 – Método proposto

4. Estudo de caso

A metodologia apresentada acima será aplicada a uma Unidade Operacional (UO²) pertencente ao cenário brasileiro de produção de petróleo sob gerência da Petrobras, a UO-Bacia de Campos. Esta UO é constituída por campos maduros² e possui o histórico de produção mais significativo dentre todas³ as UO's. Os dados obtidos são anuais e compreendem-se no período entre 1974 e 2012, totalizando 39 observações.

A série histórica em questão é de volume recuperável, e uma vez que esses dados são altamente estratégicos para a companhia, não serão divulgados neste estudo.

Determinou-se como período *in sample* os anos entre 1974 e 2007. Já os anos de 2009 a 2012, foram utilizados para testar a capacidade preditiva dos modelos, constituindo, portanto, o período *out of sample*.

²As regiões maduras são assim classificadas por apresentar grande histórico de volumes e um número de novas descobertas em tendência de declínio.

³ Não foi considerado o histórico de produção dos campos do Pré-Sal, cuja exploração iniciou-se em 2007, constituindo uma série histórica com um pequeno número de observações, o que impossibilita o ajuste adequado aos modelos propostos.

Utilizou-se o software *EViews* para a realização dos ajustes, das previsões dos modelos e para a obtenção de todas as estatísticas utilizadas na análise que se segue.

Prosseguindo com a análise, na Tabela 1 encontra-se a representação das Funções de Autocorrelação (*FAC*) e de Autocorrelação Parcial (*FACP*) da série histórica de volume recuperável da UO-Bacia de Campos.

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
		1	0.934	0.934	36.688	0.000
		2	0.860	-0.096	68.621	0.000
		3	0.780	-0.077	95.663	0.000
		4	0.701	-0.042	118.10	0.000
		5	0.619	-0.065	136.10	0.000
		6	0.540	-0.025	150.21	0.000
		7	0.461	-0.046	160.85	0.000
		8	0.384	-0.050	168.45	0.000
		9	0.309	-0.032	173.55	0.000
		10	0.237	-0.042	176.66	0.000
		11	0.172	-0.007	178.35	0.000
		12	0.109	-0.046	179.05	0.000
		13	0.051	-0.022	179.21	0.000
		14	-0.008	-0.064	179.21	0.000
		15	-0.042	0.126	179.33	0.000
		16	-0.084	-0.122	179.82	0.000

Tabela 1 – *FAC* e *FACP* da série histórica da UO-Bacia de Campos

Pode-se perceber o decaimento lento dos *lags* da *FAC*, o que caracteriza a não estacionariedade e presença de autocorrelação nos dados. Esta informação (autocorrelação) pode ser formalmente comprovada pelo teste de Ljung-Box (1978). Este teste tem como objetivo investigar se as primeiras “*m*” autocorrelações de uma série temporal são conjuntamente estatisticamente significantes. Ou seja, se algum subconjunto das autocorrelações é não nulo. As hipóteses nula e alternativa para esse teste são:

$$-H_0: \rho(1) = \rho(2) = \dots = \rho(m) = 0$$

$$-H_a: \text{pelo menos um dos } \rho's \neq 0$$

Considerando um nível de significância α de 5%, rejeita-se a hipótese nula, pois todos os *p-values* (coluna “*Prob*” da Tabela 1) obtidos são menores que α . Dessa forma, conclui-se que a série em questão não representa um ruído branco, podendo ser modelada a fim de se realizar previsões para a mesma.

A *FACP* é a estimativa da autocorrelação entre as observações $VRec_t$ e $VRec_{t-k}$, controlando-se pelas observações intermediárias. Mede, assim, a “autocorrelação líquida” entre $VRec_t$ e $VRec_{t-k}$. Pode-se observar na coluna “*Partial Correlation*” da Tabela 1 que apenas o *lag* 1 apresenta autocorrelação parcial significativa, indicando que é bastante razoável que esta série seja modelada por modelos autorregressivos de ordem 1, embora outras suposições acerca da estacionariedade da série devam ser feitas no caso de outros modelos. Todos os modelos de Knoring et al. (1999), que podem ser vistos da Eq. (11) à Eq. (19), são não lineares autorregressivos de ordem 1, logo, devido a esta última característica, podem ser adequados para representar a série em análise.

Após os ajustes dos modelos à série, verificou-se a significância de seus parâmetros, através do teste cujas hipóteses estão representadas a seguir, sendo $C(k)$ um coeficiente qualquer que se deseja testar:

$$-H_0: C(k) = 0$$

$$-H_a: C(k) \neq 0$$

Apenas os modelos Knoring 2 (Eq.(12)), Knoring 3 (Eq.(13)) e Knoring 6 (Eq.(16)) foram aprovados pelo teste acima, isto é, a um nível de significância de 5%, tiveram a hipótese nula rejeitada para todos os seus coeficientes. Os outros modelos, portanto, foram eliminados do processo de escolha.

Dentre os que restaram, aquele que obteve o menor valor para o critério $AICc$ foi Knoring 3. A partir da Tabela 2, os valores dos coeficientes e o teste de significância para os mesmos podem ser analisados.

	Coefficient	Std. Error	t-Statistic	Prob.
C(1)	-1.062501	0.354169	-2.999986	0.0053
C(2)	0.002172	0.000560	3.878764	0.0005

Tabela 2 – Parâmetros do modelo Knoring 3

De acordo com os p -values apresentados na Tabela 2, ao nível de significância de 5%, os dois coeficientes são considerados estatisticamente significantes.

É necessário analisar, também, a FAC dos resíduos gerados por Knoring 3, que pode ser vista na Tabela 3.

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
		1	0.022	0.022	0.0179	0.894
		2	-0.108	-0.109	0.4531	0.797
		3	-0.120	-0.117	1.0105	0.799
		4	-0.006	-0.014	1.0119	0.908
		5	-0.228	-0.260	3.1607	0.675
		6	-0.076	-0.097	3.4086	0.756
		7	0.262	0.221	6.4496	0.488
		8	0.024	-0.070	6.4773	0.594
		9	-0.053	-0.034	6.6124	0.677
		10	-0.233	-0.259	9.3330	0.501
		11	0.185	0.172	11.121	0.433
		12	-0.076	-0.044	11.437	0.492
		13	-0.061	-0.077	11.653	0.556
		14	-0.066	-0.148	11.917	0.613
		15	-0.003	-0.161	11.917	0.685
		16	-0.051	-0.016	12.091	0.738

Tabela 3 – FAC e $FACP$ dos resíduos gerados pelo modelo Knoring 3

Os p -values de todos os $lags$ nos leva a concluir, ao nível de significância α de 5%, que o resíduo é decorrelatado, pois falhou-se em rejeitar a hipótese nula do teste de Ljung-Box (1978), como desejado. Dessa forma, esse modelo pode ser considerado adequado para representar e realizar previsões para a série de volume recuperável da UO-Bacia de Campos.

A Figura 3 a seguir mostra a série real, a série ajustada no período *in sample* por Knoring 3 e as previsões para o período *out of sample*, de 2008 a 2012.

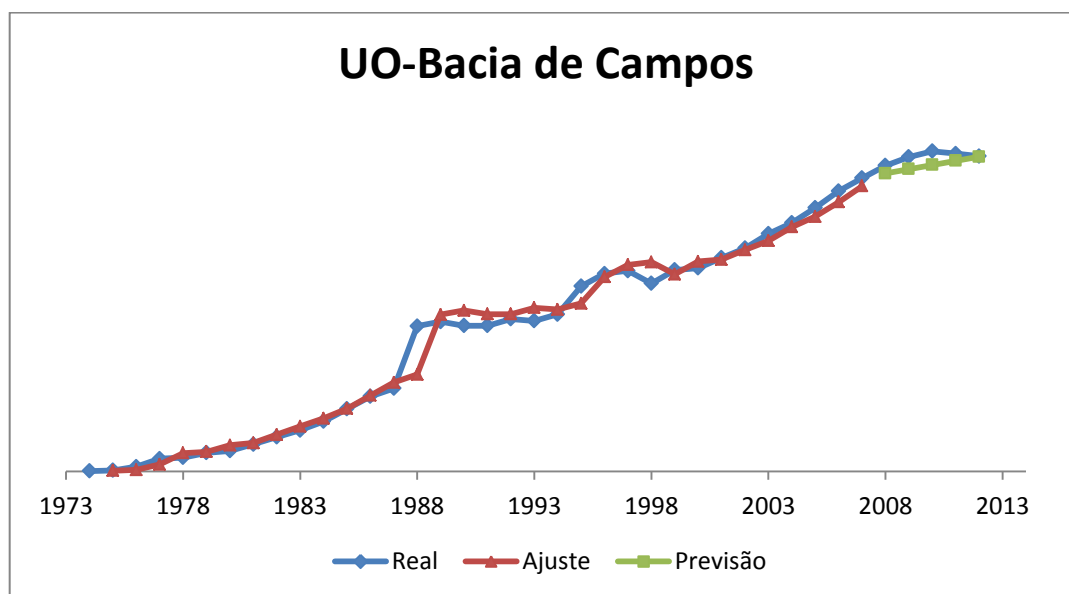


Figura 3 - Série real e série ajustada por Knoring 3 para o período *in sample*.

Analizando primeiramente o período *in sample*, obteve-se um *MAPE* do ajuste de 10,9%. Este valor, relativamente alto, é justificado por duas observações atípicas que ocorrem na série real, prejudicando o ajuste do modelo. Na Figura 3, pode-se observar que ocorre um crescimento repentino e acentuado da série em 1988, seguido de uma queda em 1990. Além disso, a partir de 1995, a série muda bruscamente sua tendência, com a ocorrência de outro “salto”, sendo que em 1998 ocorre mais uma queda, seguida de uma tendência de crescimento até o último ano do ajuste, isto é, 2007. Séries de volume recuperável nunca deveriam apresentar decaimento, visto que representa a produção acumulada acrescida das reservas provadas, discussão já abordada anteriormente. A única maneira de uma série desse tipo decair é se um determinado volume de reservas provadas não pertencer mais a esta classe. Isso ocorre por erro de estimação ou de declaração de reservas, o que prejudica os ajustes dos modelos. Ainda assim, calculando o *MAPE* do ajuste de 1999 até 2007, o valor obtido foi de 2,41%, demonstrando que o modelo capturou bem a tendência da série nestes últimos anos.

Em relação ao período utilizado para testar a capacidade preditiva do modelo, mais uma inconsistência é visível na série real, visto que em 2011 e em 2012 ocorrem quedas do volume recuperável.

Dando prosseguimento à análise, o *MAPE* obtido para as previsões foi de 2,65%, valor relativamente baixo, se comparado ao obtido no ajuste. É relevante dizer também que o *MAPE* da previsão obtido por Knoring 3 é o menor dentre os obtidos pelos outros dois modelos aprovados pelo teste de significância dos parâmetros. Para avaliá-lo melhor, a Tabela 4 demonstra a discrepância entre os valores previstos e os observados para os cinco anos do período *out of sample*, já que o *MAPE* é a média desses valores.

Ano	Discrepância
2008	2,59%
2009	3,90%
2010	4,27%
2011	2,27%
2012	0,22%

Tabela 4 – Discrepância entre os valores previstos e observados para o período *out of sample*

Cabe aqui dizer que o modelo subestimou, em suas previsões, todos os valores observados nesses cinco anos. Porém, pode-se constatar que para 2012, ano mais distante do período *in sample*, o erro de previsão foi o menos significativo, atribuindo credibilidade ao modelo, que capturou a verdadeira tendência da série, portanto.

Dessa forma, para realizar previsões para a série em questão, basta substituir os valores dos coeficientes da Tabela 6 na equação de Knoring 3 (Eq. (13)), obtendo a Eq. (27).

$$\widehat{VRec}_{t+1} = \frac{VRec_t}{1 - e^{-1,062501 - 0,002172VRec_t}} \quad \text{Eq. (27)}$$

Finalmente, o horizonte de previsão não deve ser maior do que 5 anos à frente do último ano do período *out of sample*, a fim de que sejam obtidas previsões confiáveis. Além disso, a cada nova observação, deve-se repetir o processo de ajuste e escolha do modelo, pois, dessa forma, mais informações da série serão capturadas e incorporadas aos modelos.

5. Conclusão

Realizar projeções de produção ou do acréscimo de reservas de óleo e gás de uma região é um assunto que vem sendo discutido desde meados do século passado, devido à importância do petróleo na conjuntura do mundo atual.

Neste contexto, o objetivo deste trabalho foi realizar a aplicação da metodologia de Knoring et al. (1999) a um conjunto de campos brasileiros, propondo um método de escolha do melhor modelo para representar a série em estudo, dentre os nove propostos pelos autores. Em resumo, essa metodologia é baseada em uma análise estatística detalhada dos resultados obtidos a partir do ajuste dos modelos, como teste de significância dos parâmetros, critérios de informação e análise da função de autocorrelação dos resíduos, além da avaliação dos *MAPE*'s do ajuste e da previsão.

Verificou-se que o valor relativamente alto obtido para o *MAPE* do ajuste pelo modelo escolhido é justificado pela complexidade da série e pelos “*outliers*” que podem ser observados na mesma, e que, apesar disso, o modelo capturou bem a série nos últimos nove anos. Ademais, o *MAPE* da previsão, de 2,65%, confirma que este modelo é adequado para efetuar previsões confiáveis para a série analisada, com destaque para o erro percentual de previsão obtido para 2012, de 0,22%, apenas.

Em estudos futuros, sugere-se pesquisar e testar modelos que considerem o esforço exploratório, isto é, o número de poços perfurados ao longo do tempo. Sugere-se, também, testar modelos de regressão dinâmica, incorporando, assim, à modelagem, outras variáveis exógenas explicativas. Além disso, como neste trabalho o foco foi a previsão de curto prazo, é necessário desenvolver uma metodologia de projeção de acréscimo de reservas de óleo e gás para o longo prazo, considerando um horizonte de previsão de 20 anos ou mais.

6. Referências

- AKAIKE, H. (1974). *A new look at the statistical model identification*. IEEE Transactions on Automatic Control., Boston, v.19, n.6, p.716-723.
- DEFFEYES, K. S. (2008). *Hubbert's Peak: The Impending World Oil Shortage*. New Jersey: Princeton University Press, p. 232.
- HANNAN, E.J.; QUIN, G.G. (1979). *The determination of the order of an autoregression*. J.R. Statistic. Soc. B, 41, p. 190-195.
- HUBBERT, M. K. (1956) *Nuclear Energy and the Fossil Fuels*. Shell Development Company, Houston, Texas.
- HURVICH, C.M. & TSAI, C.L. (1989). *Regression and time series model selection in small samples*. Biometrika, 76, 297–307.
- KNORING, L. D.; CHILINGAR, G. V.; GORFUNKEL, M. V. (1999). *Strategies for Optimizing Petroleum Exploration*. Houston: Gulf Publishing Company, p.321.
- LJUNG, G. M.; BOX, G. E. P. (1978). *On a measure of lack of fit in time series models*. Biometrika, 65, 297-303.
- MILLS, J.A. & PRASAD, K. (1992). *A comparison of model selection criteria*. Econometric Reviews, 11, 205–233.
- ROSA, S. E. S.; GOMES, G. L. (2004). *O Pico de Hubbert e o Futuro da Produção Mundial de Petróleo*. Revista do BNDES, v.11, n.22, p.21-49.
- SCHWARZ, G. (1978). *Estimating the dimensional of a model*. Annals of Statistics, Hayward, v.6, n.2, p.461-464.
- SZKLO A., MACHADO G. e SCHAEFFER R. (2007). *Future oil production in Brazil-Estimates based on a Hubbert model*. Energy Policy, 35, 2360-2367.