

COMBINAÇÃO DE PREVISORES COM APLICAÇÕES NO MERCADO DE DERIVADOS DO PETRÓLEO

Rafael de O. Valle dos Santos
PETROBRAS

Av. Nilo Peçanha 151, Centro, Rio de Janeiro – RJ. CEP 20020-100
rvsantos@petrobras.com.br

Marley M. B. R. Vellasco
PUC-Rio

Rua Marquês de São Vicente 225, Gávea, Rio de Janeiro – RJ. CEP
marley@ele.puc-rio.br

RESUMO

Diversos resultados empíricos na área de séries temporais indicam que combinar previsores (experts) é, em média, melhor que tentar selecionar um único modelo de previsão. Na medida em que se decide por um esquema de combinação linear, há vários métodos disponíveis para determinar o quanto cada previsor deve contribuir para a resposta consensual, ou em outras palavras, quais devem ser os pesos dos previsores envolvidos. Este trabalho explora o uso de diversos métodos ponderação para combinação linear de previsores, utilizando dados do mercado brasileiro de derivados do petróleo.

PALAVRAS CHAVE. Séries temporais. Combinação de previsores. Combinação linear. Mercado de derivados.

Área principal: P&G – PO na Área de Petróleo & Gás

ABSTRACT

Several empirical results on the time series field indicate that combining forecasts is, on average, better than selecting a single forecasting model. Once the linear combination framework is chosen, there are many ways to define the amount of contribution of each combining model to the consensual response; in other words, there are many possible weighting methods. The present work explore the usage of different weight generation schemes for the linear combination of forecasts, using data from the brazilian downstream oil market.

KEYWORDS. Time series. Forecast combination. Linear combination. Downstream oil market.

Main area: P&G – PO na Área de Petróleo & Gás

1. Introdução

Análise e previsão de séries temporais é notadamente uma área com muitas aplicações práticas. Muitas destas aplicações estão ligadas ao gerenciamento de cadeias de suprimento ou à modelagem de variáveis econômicas; contudo, há na literatura uma lista extensa de possibilidades ([Harvey, 1991](#); [Pole et al. 1994](#); [Souza e Camargo, 2004](#)):

1. Número de casamentos na Grécia;
 2. Volume do rio Nilo;
 3. Mortalidade em Londres;
 4. Bolsas roubadas em Chicago;
 5. Volume de chuvas em Fortaleza;
 6. Número de tratores na Espanha;
 7. Vazão mensal em reservatório hidrelétrico de Furnas;
- etc.

Em particular, para as empresas integradas de Petróleo & Gás, o uso de técnicas de séries temporais é notadamente útil nas atividades de planejamento relacionadas ao marketing e comercialização de derivados (*downstream*). Não obstante, pode-se encontrar aplicações destas técnicas em outros departamentos: financeiro, materiais e E&P (*upstream*).

No processo de tomada de decisão, normalmente múltiplas previsões para uma mesma variável estão disponíveis para o gestor. Neste momento, como observado por Timmerman ([2006](#)), uma questão surge naturalmente: qual a melhor forma de se aproveitar as informações de cada previsor?

Como será visto ao longo deste trabalho, a multiplicidade de previsões, relativamente comum no âmbito das grandes empresas, não deve ser considerada indesejada. Ao contrário, ela pode levar à construção de sistemas com previsões combinadas. Se bem projetada, a *combinação de previsores*¹ resulta em previsões consensuais superiores (ou de menor risco) do que as individuais; esta conclusão é abordada em diversos trabalhos ([Chase, 2009](#); [Sanchez, 2008](#); [Timmermann, 2006](#); [Hibon & Evgeniou, 2005](#); [Zou & Yang, 2004](#)).

Com o objetivo de estudar a aplicação das combinações de previsores ao mercado nacional de derivados do petróleo, este trabalho utiliza dados reais do mercado brasileiro, relativos às vendas de *óleo diesel* e *gás liquefeito de petróleo* (GLP) na região sudeste do país.

2. Combinação Linear de Previsores

A combinação linear de N previsores, calculada no instante t , h passos a frente, assume a seguinte forma:

$$\hat{y}_{t+h|t} = f(\hat{y}_{t+h|t}, \hat{w}_{t+h|t}) = f_{linear}(\hat{w}_{t+h|t}) \quad (1)$$

Onde $\hat{y}_{t+h|t}$ é o vetor de previsões disponíveis e $\hat{w}_{t+h|t}$ é o vetor de pesos estimados. Expandindo:

$$\hat{y}_{t+h|t} = [\hat{y}_{t+h|t,1} \quad \hat{y}_{t+h|t,2} \quad \dots \quad \hat{y}_{t+h|t,N}]' \quad (2)$$

$$\hat{w}_{t+h|t} = [\hat{w}_{t+h|t,1} \quad \hat{w}_{t+h|t,2} \quad \dots \quad \hat{w}_{t+h|t,N}]' \quad (3)$$

¹ Os termos “combinação de previsores”, “combinação de previsões” e “comitês de previsores” (*ensembles*) são usados intercaladamente na literatura.

sendo $\hat{y}_{t+h|t,k}$ a previsão fornecida em t , h passos a frente, pelo k -ésimo predictor e $\hat{w}_{t+h|t,k}$ o peso a associado a esta previsão.

Combinações lineares são geralmente desejáveis por sua simplicidade matemática e facilidade de interpretação, principalmente pelo fato de haver uma relação *biunívoca* entre pesos e predictors. Por outro lado, esquemas não-lineares como os relatados por Donaldson e Kamstra ([1996](#)) e Mubwandarikwa ([2007](#)), apesar de apresentarem bons resultados em diversas situações, têm contra si o fato de serem complexos ou de difícil interpretação.

O formato mais comum para combinação linear de predictors é o seguinte:

$$\hat{y}_{t+h|t} = \sum_{k=1}^N \hat{w}_{t+h|t,k} \hat{y}_{t+h|t,k} \quad (4)$$

sujeito a

$$\sum_{k=1}^N w_{t+j|t,k} = 1 \quad e \quad w_{t+j|t,k} \geq 0 \quad (5)$$

Além da simplicidade na interpretação dos pesos, o esquema de combinação nas equações (4) e (5) é altamente desejável por dois outros motivos: (i) garantia de *não-tendenciosidade* na previsão combinada (se os predictors envolvidos forem não-tendenciosos) e (ii) garantia de *convexidade* na combinação.

3. Métodos de Ponderação para Combinações Lineares

Para combinação linear de predictors, foco teórico deste trabalho, existem na literatura alguns métodos consagrados para geração de pesos, todos eles baseados em desempenhos individuais. Os métodos selecionados para este estudo estão descritos a seguir.

3.1. Média Simples

Tirar a média das previsões é o método mais simples de combinação linear de predictors; ainda assim, é difícil de ser batido ([Timmermann, 2006](#)). Fazer a média simples no horizonte $h > 0$ é, em outras palavras, utilizar o seguinte valor de peso para cada um dos N predictors disponíveis:

$$\hat{w}_{t+h|t,k} = 1/N \quad (k = 1, 2, \dots, N) \quad (6)$$

3.2. Mínimos Quadrados Irrestritos

A estimação de pesos por mínimos quadrados irrestritos é um exemplo clássico de otimização aplicada à estatística. A idéia é que se obtenha os pesos minimizando o erro quadrático ao longo das observações disponíveis. No caso linear essa minimização é expressa como se segue, considerando-se janela de tempo de tamanho v e horizonte de previsão $h > 0$:

$$\hat{\mathbf{w}}_{t+h|t} = \begin{bmatrix} \hat{w}_{t+h|t,1} \\ \hat{w}_{t+h|t,2} \\ \mathbf{M} \\ \hat{w}_{t+h|t,N} \end{bmatrix} = \min_{\mathbf{w}} \sum_{i=0}^{v-1} \left(y_{t-i} - \sum_{k=1}^N w_{t-i|t-i-h,k} \cdot \hat{y}_{t-i|t-i-h,k} \right)^2 \quad (7)$$

A expressão acima tem solução exata ([Johnson e Wichern, 2007](#)):

$$\hat{\mathbf{w}}_{t+h|t} = (\hat{\mathbf{V}}'_{t|t-h} \hat{\mathbf{V}}_{t|t-h})^{-1} \hat{\mathbf{V}}'_{t|t-h} \mathbf{v}_{t|t-h} \quad (8)$$

sendo

$$\hat{\mathbf{V}}_{t|t-h} = \hat{\mathbf{V}}_{t|t-h}(\mathbf{v}) = \begin{bmatrix} \hat{y}_{t-1+|t-h-1+1,1} & \hat{y}_{t-1+|t-h-1+1,2} & \Lambda & \hat{y}_{t-1+|t-h-1+1,N} \\ \hat{y}_{t-2+|t-h-2+1,1} & \hat{y}_{t-2+|t-h-2+1,2} & \Lambda & \hat{y}_{t-2+|t-h-2+1,N} \\ \mathbf{M} & \mathbf{M} & \Lambda & \mathbf{M} \\ \hat{y}_{t-v+1|t-h-v+1,1} & \hat{y}_{t-v+1|t-h-v+1,2} & \Lambda & \hat{y}_{t-v+1|t-h-v+1,N} \end{bmatrix} \quad (9)$$

$$\mathbf{v}_{t|t-h} = \mathbf{v}_{t|t-h}(\mathbf{v}) = \begin{bmatrix} y_{t-1+1} \\ y_{t-2+1} \\ \mathbf{M} \\ y_{t-v+1} \end{bmatrix} \quad (10)$$

As equações acima podem ser ajustadas para considerar janela expansiva, i.e., uma janela de tempo que aumenta junto com o histórico. Neste caso, a equação (7) ficaria assim:

$$\hat{\mathbf{w}}_{t+h|t} = \begin{bmatrix} \hat{w}_{t+h|t,1} \\ \hat{w}_{t+h|t,2} \\ \mathbf{M} \\ \hat{w}_{t+h|t,N} \end{bmatrix} = \min_w \sum_{i=h+1}^t \left(y_i - \sum_{k=1}^N w_{i|t-h,k} \cdot \hat{y}_{i|t-h,k} \right)^2 \quad (11)$$

3.3. Mínimos Quadrados Restritos

O uso de mínimos quadrados irrestritos garante, dentro da amostra, a melhor solução matemática para a estimação de pesos em combinações lineares (ver seção anterior). Mas os vetores de pesos gerados por aquele método podem trazer problemas práticos fora da amostra, uma vez que não atendem a duas restrições desejáveis (Timmermann, 2006): (a) *somar 1* (b) *ter elementos maiores ou iguais a zero*. A restrição de *somar 1* é necessária para que a previsão por combinação linear possa ser garantidamente *não-tendenciosa* (se os previsores sendo combinados forem não-tendenciosos); já a restrição de *não ter componentes negativos*, associada à condição de não-tendenciosidade, torna a combinação *convexa*, fazendo com que magnitude da previsão combinada esteja dentro da região de valores definida pelas previsões individuais.

Para transformar o problema de mínimos quadrados irrestritos em um problema de mínimos quadrados *restritos*, reescreve-se a equação (7) da seguinte forma:

$$\hat{\mathbf{w}}_{t+h|t} = \begin{bmatrix} \hat{w}_{t+h|t,1} \\ \hat{w}_{t+h|t,2} \\ \mathbf{M} \\ \hat{w}_{t+h|t,N} \end{bmatrix} = \min_w \sum_{i=0}^{v-1} \left(y_{t-i} - \sum_{k=1}^N w_{t-i|t-h,k} \cdot \hat{y}_{t-i|t-h,k} \right)^2 \quad (12)$$

sujeito a

$$\sum_{k=1}^N w_{t|t-h,k} = 1 \quad e \quad w_{t|t-h,k} \geq 0 \quad (13)$$

Para este tipo de otimização não há solução exata; os algoritmos utilizados são iterativos ([Gill et al. 1984](#); [Mathworks, 2009a](#)).

3.4. Shrinkage

Shrinkage ([Diebold e Pauly, 1990](#)) não é um método de ponderação isolado, mas uma metodologia de integração de métodos. A idéia básica é que o vetor de pesos final seja uma média entre os pesos calculados por algum método de ponderação e os pesos fixos da média simples; em outras palavras, os pesos estimados são *dilatados* na direção da combinação por média simples:

$$\hat{w}_{t+h|t,k} = \alpha \hat{w}_{t+h|t,k}^{\text{método}} + (1 - \alpha)(1/N) \quad (14)$$

O hiperparâmetro α varia no intervalo $[0,1]$ e controla a quantidade de dilatação na direção dos pesos iguais; no caso mais trivial, α assume o valor de $1/2$.

4. Experimentos

Os métodos pesquisados neste trabalho foram aplicados em dados reais do mercado brasileiro de derivados do petróleo, publicados pela *Agência Nacional de Petróleo, Gás Natural e Biocombustíveis* ([ANP, n.d.](#)). Duas séries de vendas mensais foram escolhidas, ambas relativas a região sudeste do país: *óleo diesel* e *gás liquefeito de petróleo* (GLP). As figuras 1 e 2 a seguir ilustram estas séries:

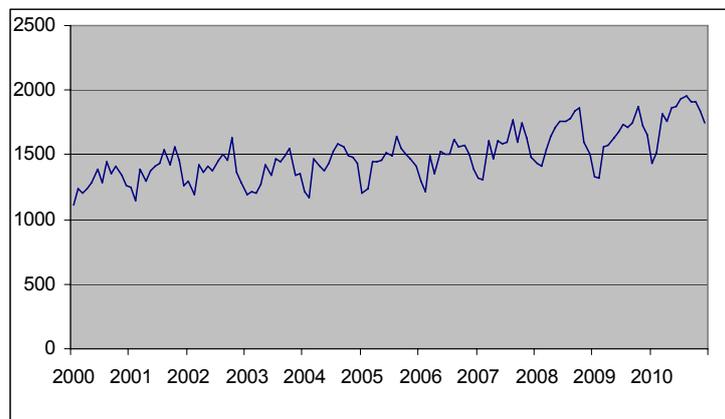


Figura 1 – Vendas de óleo diesel de jan/2000 a dez/2010 na região sudeste (Mm³)

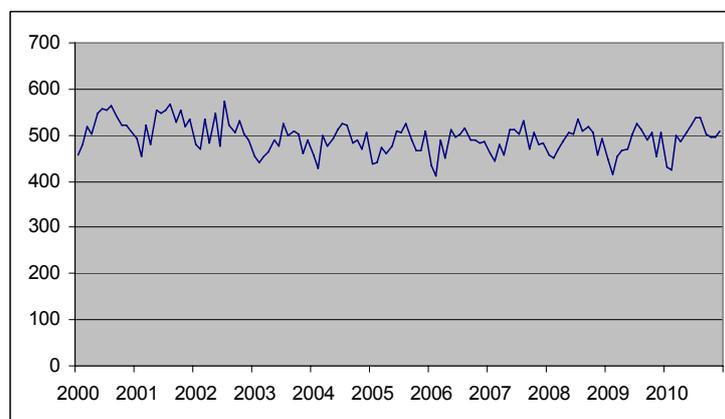


Figura 2 – Vendas de GLP de jan/2000 a dez/2010 na região sudeste (Mm³)

4.1. Metodologia

Para cada uma das séries estudadas, foram ajustados 4 modelos de previsão:

1. Holt-Winters;
2. Regressão harmônica;
3. Decomposição clássica;
4. ARIMA Box & Jenkins.

A escolha destes modelos, que possuem vasta literatura ([Souza e Camargo, 2004](#); [Makridakis et al. 1998](#); [Harvey, 1991](#); [Kachigan, 1986](#)) seguiu dois critérios: terem naturezas diferentes entre si (buscando complementaridade) e serem capazes de representar tendência e sazonalidade (características recorrentes em séries de vendas).

Para cada um dos previsores individuais foram geradas previsões de 1 até 12 passos a frente. Estes conjuntos de previsões, juntamente com as realizações das séries, foram utilizados para testar os métodos de ponderação e combinação linear estudados neste trabalho.

Para comparação de resultados foi escolhida o *symmetric mean absolute percentage error* (SMAPE) como métrica de desempenho:

$$SMAPE = \frac{1}{H} \sum_{t=h+1}^H \frac{|y_t - \hat{y}_{t|t-h}|}{\left(\frac{|y_t| + |\hat{y}_{t|t-h}|}{2} \right)} \quad (15)$$

Sendo:

H o tamanho da amostra de teste;

h o horizonte de previsão considerado;

y_t a realização da série no instante t ;

$\hat{y}_{t|t-h}$ a previsão em t , realizada em $t-h$.

O SMAPE é uma métrica amplamente utilizada em trabalhos relacionados a séries temporais, principalmente por seu emprego na competição *M3* para modelos de previsão ([Makridakis e Hibon, 2000](#)).

Para garantir a validade das conclusões tomadas, foram utilizados testes de hipótese para comparação formal entre os principais métodos de combinação e cada um dos previsores individuais. Os testes utilizados – *teste t*, *teste de sinais* e *teste de Wilcoxon* ([Kachigan, 1986](#); [Gibbons, 1992](#); Flores, [1986](#) e [1989](#)) – verificam se a mediana das *diferenças de desempenho* entre os dois métodos é significativa. Mede-se uma diferença de desempenho para cada um dos horizontes de previsão considerados (de 1 até 12 passos a frente); a unidade de medida da diferença é a mesma de cada um dos desempenhos sendo comparados (SMAPE).

Finalmente, destaca-se que todos os modelos, algoritmos e testes de hipótese utilizados neste trabalho foram implementados com os seguintes pacotes de software: MATLAB ([Mathworks, n.d.](#)), Forecast Pro XE ([BFS, n.d.](#)) e MS-Excel.

4.2. Resultados – DIESEL

Todos os modelos individuais de previsão disponíveis foram ajustados e usados para gerar previsões de vendas de óleo diesel até 12 passos a frente, de maneira não recursiva (i.e., sem re-estimação de parâmetros a cada passo). A Tabela 1. mostra os resultados médios acumulados para horizontes de até 12 meses; a Tabela 2. traz a matriz de correlações entre os erros instantâneos dentro da amostra.

Tabela 1 – Resultados médios para previsões individuais

Horizonte	SMAPE(%)			
	HW	REG	DEC	BJ
1	1.61	2.56	1.01	4.46
2	3.47	3.36	4.34	3.29
3	5.11	6.01	6.85	4.80
4	5.46	6.68	7.62	4.96
5	5.92	7.40	8.27	5.43
6	6.05	7.74	8.47	5.54
7	6.32	8.22	8.81	5.85
8	6.18	8.20	8.59	5.82
9	6.02	8.24	8.42	5.70
10	5.47	7.82	7.86	5.23
11	5.48	7.80	7.83	5.20
12	5.56	7.97	7.86	5.18

Horizontes de até 12 meses. Os previsores individuais são: *Holt-Winters* (HW), *regressão harmônica* (REG), *decomposição clássica* (DEC) e *ARIMA Box & Jenkins* (BJ). Os erros acumulados nos primeiros 6 meses foram usados para selecionar os modelos a serem combinados.

Tabela 2 – Matriz de correlações entre previsores

	HW	REG	DEC	BJ
HW	1.00			
REG	0.73	1.00		
DEC	0.72	0.96	1.00	
BJ	0.93	0.68	0.65	1.00

Correlações entre os erros instantâneos dentro da amostra. Os modelos REG e DEC são os mais correlacionados (0,96).

Com base nos menores erros médios apresentados nos primeiros 6 meses de previsão, definiu-se a seguinte seqüência para combinação dos previsores:

1. BJ + HW; (2 modelos)
2. BJ + HW + REG; (3 modelos)
3. BJ + HW + DEC; (3 modelos)
4. Todos. (4 modelos)

Para cada uma das 4 combinações propostas acima, testou-se todos os métodos de ponderação estudados na seção 3. (média simples + 2 métodos de ponderação + 2 variantes *shrinkage*), admitindo-se 3 tamanhos de janela: 3, 6 e *expansiva*². A Tabela 3 mostra, para cada tipo de combinação, os melhores resultados obtidos.

² O total de experimentos realizados por série foi de $4 \times 5 \times 3 = 60$.

Tabela 3 – Melhor resultado de combinação

Horizonte	SMAPE(%)			
	BJ+HW	BJ+HW+REG	BJ+HW+DEC	Todos
1	7.49	4.73	3.08	1.73
2	5.23	2.59	4.16	1.83
3	3.67	3.34	3.23	3.92
4	2.98	6.14	4.42	5.38
5	2.64	5.77	5.90	6.53
6	2.45	5.12	5.62	6.00
7	2.11	4.93	5.78	6.14
8	1.97	4.41	5.78	6.15
9	1.97	4.17	5.71	6.10
10	2.44	3.89	5.24	5.62
11	2.26	4.31	5.37	5.73
12	2.26	4.20	5.20	5.55

Os melhores métodos de combinação foram: Mínimos quadrados irrestritos (MQI) com janela 3 (BJ+HW); Shrinkage de MQI (SMQI) com janela 3 (BJ+HW+REG) e SMQI com janela 6 (BJ+HW+REG; todos).

As Tabelas 4, 5, e 6 exibem os testes de hipótese (seção 4.1) para o melhor resultado de combinação obtido: **BJ+HW, com MQI janela 3**. Em cada teste, a hipótese nula é de que a mediana das diferenças de desempenho é zero; ela é rejeitada se o *valor-p* for menor ou igual a 5%. Mostra-se ainda o intervalo de confiança (95%) para a mediana das diferenças (individual – combinação) e indica-se se a amostra testada é normal, de acordo com a metodologia Jarque-Bera. ([Cromwell et al. 1994](#)).

Tabela 4 – Teste t (melhor combinação x individuais)

Previsor	Diferença significativa?	Valor-p	Limite inferior	Limite superior	Amostra normal?
BJ	SIM	0.010	0.58	3.42	NÃO
HW	SIM	0.034	0.19	4.01	NÃO

Tabela 5 – Teste de sinais (melhor combinação x individuais)

Previsor	Diferença significativa?	Valor-p	Limite inferior	Limite superior	Amostra normal?
BJ	SIM	0.039	1.12	3.73	NÃO
HW	SIM	0.039	1.44	4.05	NÃO

Tabela 6 – Teste de Wilcoxon (melhor combinação x individuais)

Previsor	Diferença significativa?	Valor-p	Limite inferior	Limite superior	Amostra normal?
BJ	SIM	0.021	0.36	3.33	NÃO
HW	NÃO	0.052	-0.16	3.72	NÃO

Por fim, o gráfico a seguir exibe, de maneira comparativa, a evolução dos principais SMAPEs fora da amostra.

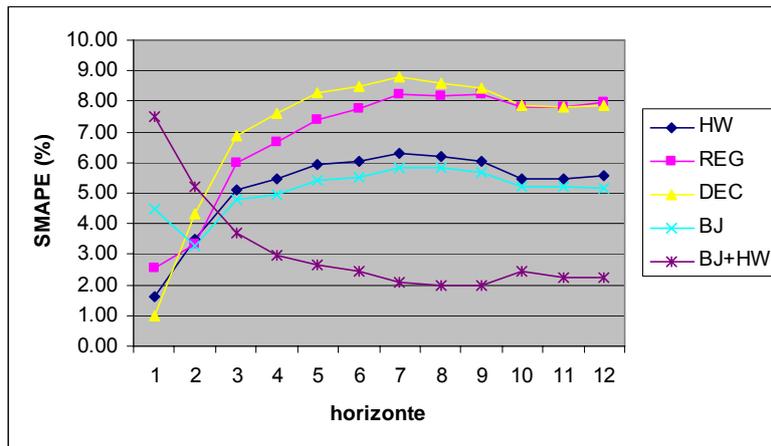


Figura 3 – Evolução dos SMAPEs fora da amostra (melhor combinação x individuais)

4.3. Resultados – GLP

Para análise do gás liquefeito de petróleo (GLP), utilizou-se a mesma lógica de experimentos executada para o diesel (seção 4.2). Assim, a Tabela 7 mostra os resultados médios acumulados para horizontes de até 12 meses, enquanto a Tabela 8 traz a matriz de correlações entre os erros instantâneos dentro da amostra.

Tabela 7 – Resultados médios para previsões individuais

Horizonte	SMAPE(%)			
	HW	REG	DEC	BJ
1	3.17	0.75	1.40	2.10
2	1.87	0.63	0.91	2.04
3	3.34	2.98	3.00	3.26
4	3.92	4.41	4.18	3.64
5	3.72	4.42	4.13	3.30
6	4.05	4.93	4.73	3.52
7	4.13	5.08	5.00	3.51
8	4.31	5.23	5.31	3.61
9	4.24	5.26	5.34	3.49
10	3.88	5.05	5.11	3.21
11	4.14	5.37	5.40	3.44
12	4.18	5.50	5.57	3.39

Horizontes de até 12 meses. Os previsores individuais são: Holt-Winters (HW), regressão harmônica (REG), decomposição clássica (DEC) e ARIMA Box & Jenkins (BJ). Os erros acumulados nos primeiros 6 meses foram usados para selecionar os modelos a serem combinados

Tabela 8 – Matriz de correlações entre previsores

	HW	REG	DEC	BJ
HW	1.00			
REG	0.67	1.00		
DEC	0.68	0.99	1.00	
BJ	0.85	0.64	0.65	1.00

Correlações entre os erros instantâneos dentro da amostra. Os modelos RGE e DEC são os mais correlacionados (0,99).

A seqüência para combinação dos previsores foi a mesma utilizada para o diesel:

1. BJ + HW; (2 modelos)
2. BJ + HW + REG; (3 modelos)
3. BJ + HW + DEC; (3 modelos)
4. Todos. (4 modelos)

A Tabela 9 mostra, para cada tipo de combinação, os melhores resultados obtidos.

Tabela 9 – Melhor resultado de combinação

Horizonte	SMAPE(%)			
	BJ+HW	BJ+HW+REG	BJ+HW+DEC	Todos
1	2.05	0.53	0.55	2.10
2	2.61	1.95	3.62	2.04
3	3.42	2.69	3.53	3.26
4	3.53	3.83	5.18	3.64
5	2.87	3.18	4.57	3.30
6	2.59	2.95	3.93	3.52
7	2.35	2.93	3.44	3.51
8	2.26	2.57	3.02	3.61
9	2.02	2.53	2.69	3.49
10	2.39	2.41	2.56	3.21
11	2.56	2.70	2.49	3.44
12	2.71	2.86	2.69	3.54

Os melhores métodos de combinação foram: Shrinkage de MQI (SMQI) com janela 3 (BJ+HW); Mínimos quadrados irrestritos (MQI) com janela 6 (BJ+HW+REG; BJ+HW+DEC) e MQR com janela 3 (Todos).

As Tabelas 10, 11 e 12 exibem os testes de hipótese para o melhor resultado de combinação obtido: **BJ+HW, com SMQI janela 3.**

Tabela 10 – Teste t (melhor combinação x individuais)

Previsor	Diferença significativa?	Valor-p	Limite inferior	Limite superior	Amostra normal?
BJ	SIM	0.007	0.20	1.00	SIM
HW	SIM	0.001	0.57	1.69	SIM

Tabela 11 – Teste de sinais (melhor combinação x individuais)

Previsor	Diferença significativa?	Valor-p	Limite inferior	Limite superior	Amostra normal?
BJ	SIM	0.039	0.05	1.16	SIM
HW	SIM	0.039	0.39	1.78	SIM

Tabela 12 – Teste de Wilcoxon (melhor combinação x individuais)

Previsor	Diferença significativa?	Valor-p	Limite inferior	Limite superior	Amostra normal?
BJ	SIM	0.012	0.14	1.05	SIM
HW	SIM	0.003	0.52	1.75	SIM

O gráfico a seguir exhibe, de maneira comparativa, a evolução dos principais SMAPEs fora da amostra.

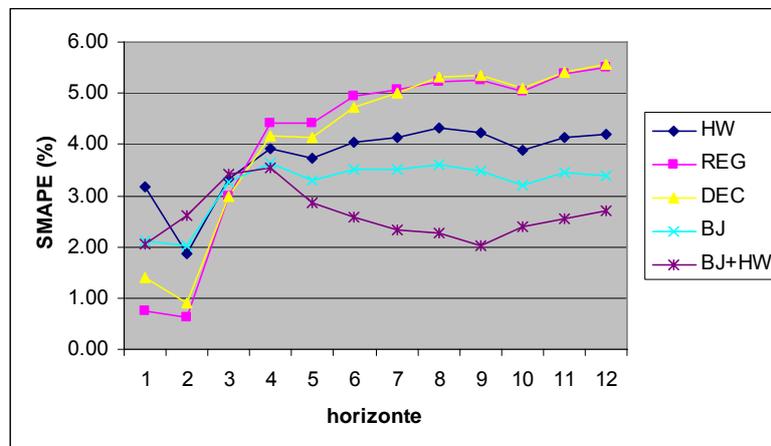


Figura 4 – Evolução dos SMAPEs fora da amostra (melhor combinação x individuais)

5. Conclusão

A combinação linear de previsores, aplicada de maneira criteriosa, acrescentou ganho significativo na análise das séries consideradas neste estudo: vendas de diesel e GLP na região sudeste do Brasil. O horizonte máximo de previsão foi de 12 meses.

Observou-se, para as duas séries, que os melhores resultados foram obtidos pela combinação dos dois previsores com os melhores desempenhos médios, considerada uma amostra de validação de 6 meses. O método de ponderação de *mínimos quadrados irrestritos* (MQI) e sua versão com *shrinkage* (SMQI) foram os de melhor desempenho.

Comprovou-se também uma conclusão teórica interessante: não há ganho em combinar *cegamente* previsores. Por exemplo, ao considerar a combinação com todos os previsores disponíveis, os desempenhos médios caem em relação às combinações menores; pode-se considerar como justificativa teórica para este fato a alta correlação entre os modelos de regressão harmônica (REG) e decomposição clássica (DEC) – 0.96 e 0.99 respectivamente para diesel e GLP – aliada ao fato dos desempenhos individuais desses dois métodos não serem bons (se comparados com os outros).

Por fim, considera-se este trabalho como uma contribuição para o estudo das combinações de previsores aplicadas ao marketing e comercialização de derivados do petróleo. Espera-se que extensões deste artigo, já em andamento, tragam outros ganhos.

Referências

- ANP (n.d.) *Dados estatísticos mensais* [WWW] Agência Nacional do Petróleo, Gás Natural e Biocombustíveis. Disponível em: <http://www.anp.gov.br/>. [acessado em 01/04/2011].
- BFS (n.d.) *Forecast Pro XE* [WWW] Business Forecast Systems, Inc. Disponível em: <http://www.forecastpro.com/products/overview/XE.htm>. [acessado em 16/11/2010].
- CHASE, C. W. (2009) *Demand-driven forecasting: a structured approach to forecasting*. John Wiley & Sons, Inc.
- CROMWELL, J. B., LABYS, W. C. e TERRAZA, M. (1994) *Univariate Tests for Time Series Models*. Sage Publications, Inc.
- DIEBOLD, F. X. e PAULY, P. (1990) The use of prior information in forecast combination. *International Journal of Forecasting*, 6, pp. 503-508.
- DONALDSON, R. G. e KAMSTRA, M. (1996) Forecast Combinig with Neural Networks. *Journal of Forecasting*, 15, pp. 49-61.

- FLORES, B. E.** (1986) Use of the sign test to supplement the percentage better statistic. *International Journal of Forecasting*, 2, pp. 477–489.
- FLORES, B. E.** (1989) The utilization of the Wilcoxon test to compare forecasting methods: A note. *International Journal of Forecasting*, 5, pp. 529-535.
- GIBBONS, J. D.** (1992) *Nonparametric Statistics: An Introduction*. Sage Publications, Inc.
- GILL, P. E. et al.** (1984) Procedures for optimization problems with a mixture of bounds and general linear constraints. *ACM Transactions on Mathematical Software*, 10, pp. 282-298.
- HARVEY, A. C.** (1991) *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge University Press.
- HIBON, M. e EVGENIOU, T.** (2005) To combine or not to combine: selecting among forecasts and their combinations. *International Journal of Forecasting*, 21, pp. 5–24.
- JOHNSON, R. A. e WICHERN, D. W.** (2007) *Applied Multivariate Statistical Analysis (6th Edition)*. Prentice Hall.
- KACHIGAN, S. K.** (1986) *Statistical Analysis: An Interdisciplinary Introduction to Univariate & Multivariate Methods*. Radius Press.
- MAKRIDAKIS, S. G. e HIBON, M.** (2000) The M3-competition: Results, conclusions and implications. *International Journal of Forecasting*, 16, pp. 451–476 (special issue).
- MAKRIDAKIS, S. G., WHEELWRIGHT, S. C. e HYNDMAN, R. J.** (1998) *Forecasting: methods and applications*. 3rd ed. John Wiley & Sons.
- MATHWORKS** (2009a) *MATLAB Optimization Toolbox 4 – User’s Guide* [WWW] Mathworks, Inc. Disponível em: http://www.mathworks.com/access/helpdesk/help/pdf_doc/optim/optim_tb.pdf [acessado em 03/03/10].
- MATHWORKS** (n.d.) *MATLAB - The Language Of Technical* [WWW] Mathworks, Inc. Disponível em: <http://www.mathworks.com/products/matlab/> [acessado em 16/11/2010].
- MUBWANDARIKWA, E.** (2007) *Modality conditions and prior weights in the geometric combination of bayesian forecasting models*. Thesis (PhD), The Open University Walton Hall, UK.
- POLE, A., WEST, M. e HARRISON, J.** (1994) *Applied Bayesian forecasting and time series analysis*. Chapman & Hall.
- SÁNCHEZ, I.** (2008) Adaptive combination of forecasts with application to wind energy. *International Journal of Forecasting*, 24, pp. 679-693.
- SOUZA, R. C. e CAMARGO, M. E.** (2004) *Análise e Previsão de Séries Temporais: os modelos ARIMA* (2^a edição). Reinaldo Castro Souza & Maria Emilia Camargo.
- TIMMERMANN, A.** (2006) Forecast combination. In: ELLIOTT, G., GRANGER, C.W.J, TIMMERMANN, A. (eds.) *Handbook of economic forecasting, volume 1*. North-Holland, pp. 135–196.
- ZOU, H. e YANG, Y.** (2004) Combining time series models for forecasting. *International Journal of Forecasting* 20, pp. 69– 84.