



Revista ADM.MADE

Revista do Mestrado em Administração e
Desenvolvimento Empresarial - Universidade
Estácio de Sá

Revista ADM.MADE, Rio de Janeiro, ano 12, v.16, n.2, p.51-67, maio/agosto, 2012

Revista do Mestrado em Administração e Desenvolvimento Empresarial da Universidade

Estácio de Sá – Rio de Janeiro (MADE/UNESA). ISSN: 2237-5139

Conteúdo publicado de acesso livre e irrestrito, sob licença Creative Commons 3.0.

Editora responsável: Isabel de Sá Affonso da Costa

Organizador do número temático: Marco Aurélio Carino Bouzada

Proposta de Modelo de Previsão de Preço de Gasolina ao Consumidor Final no Estado do Rio de Janeiro

Márcio Pitzer¹

André Lindolfo Ziebell Ricardo²

Marco Aurélio Carino Bouzada³

Uma versão preliminar deste artigo foi apresentada e publicada nos Anais do XV SIMPOI - Simpósio de Administração da Produção, Logística e Operações Internacionais, agosto de 2012, São Paulo - SP.

Artigo recebido em 12/11/2012 e aprovado em 14/12/2012. Artigo submetido a avaliação *double blind*.

¹ Mestre em Administração pelo Programa de Mestrado em Administração e Desenvolvimento Empresarial da Universidade Estácio de Sá (MADE/UNESA). Gerente de Serviços Especiais na Oi. Endereço: Av. Pres. Vargas, 642, 22^o. andar – Centro – CEP: 20071-001 - Rio de Janeiro - RJ. Email: pitzer.marcio@gmail.com

² Pós-graduado em Gestão de Serviços pelo Instituto Coppead da Universidade Federal do Rio de Janeiro (COPPEAD/UFRJ). Especialista em Logística da Oi. Endereço: Rua Gonzaga Bastos, 209/Bl. E/ Apt. 601 - Vila Isabel - CEP: 20541-902 - Rio de Janeiro - RJ. Email: andre-lindolfo@bol.com.br

³ Doutor em Administração pelo COPPEAD/UFRJ. Professor do MADE/UNESA. Endereço: Av. Pres. Vargas, 642, 22^o. andar – Centro – CEP: 20071-001 - Rio de Janeiro - RJ. Email: marco.bouzada@estacio.br.

Proposta de Modelo de Previsão de Preço da Gasolina ao Consumidor Final no Estado do Rio de Janeiro

O objetivo deste estudo é propor um modelo utilizando a regressão múltipla para prever o preço do litro da gasolina no estado do Rio de Janeiro, com base em variáveis observáveis e de fácil acesso. O trabalho é focado no estudo de caso da empresa CAM Ltda. e busca uma ferramenta que possa auxiliá-la na gestão de sua frota de 1.050 veículos. Apesar de um valor não muito elevado para o nível de explicação (65%) – que impediu uma previsão mais precisa do valor do litro da gasolina – o modelo se mostrou útil à empresa, quando comparado ao utilizado anteriormente.

Palavras-chave: preço da gasolina; previsão; regressão múltipla.

Keywords: gasoline prices; forecast model; multiple regression.

Proposal Of A Model To Forecast The Gasoline Price For The Rio De Janeiro Final Consumer

The objective of this study is to propose a model using multiple regression to forecast the gasoline liter price in Rio de Janeiro, based on observable and accessible variables. The work is focused on the CAM Ltda. company's case study and it searches for a tool to help CAM in its 1,050 vehicles fleet management. Despite the explanation level not being so high (65%) - which prevented a better forecast for the gasoline liter value – the model proved useful to the company, when compared to the previously used one.

1. Introdução

Conforme estudo realizado pela Confederação Nacional dos Transportes (CNT), no ano de 2000 o transporte de cargas no Brasil representou mais de 4% do PIB nacional e movimentou mais de R\$ 42 bilhões. Dentro deste cenário, o modal rodoviário é líder absoluto (CONFEDERAÇÃO NACIONAL DO TRANSPORTE, 2011).

O petróleo e seus derivados constituem a principal matriz energética do país. O item combustível é um custo variável e de grande importância na composição do custo do frete. No entanto, além dele, podemos destacar outros principais custos, tais como:

- Depreciação;
- Pessoal direto (motoristas e ajudantes);
- Seguro do veículo;
- IPVA/ seguro obrigatório;
- Custos administrativos;
- Pneus;
- Lubrificantes;
- Manutenção;
- Pedágio.

Em longo prazo, o etanol promete ser uma alternativa à dependência do petróleo, porém seu desenvolvimento ainda está centrado a suprir a necessidade para veículos de passeio; ainda não há, em larga escala, a sua utilização para veículos pesados. Com isso, o petróleo ainda é uma das principais *commodities* comercializada pelo Brasil, sendo a principal matéria-prima para a produção da gasolina.

A gasolina é o principal combustível da frota nacional de veículos. Além disso, existe pressão do consumo sobre a produção, observada por Burnquist e Bacchi (2010). Estes fatores ratificam a importância de um modelo que possa prever, com acurácia, o preço final da gasolina.

A construção de tal modelo se constitui um desafio, tendo em vista a influência de variáveis difíceis de serem determinadas e modeladas. Entre estas variáveis, podemos citar, como exemplo, os fatores regionais, culturais e econômicos que afetam os principais países produtores de petróleo que formam a Organização dos Países Exportadores de Petróleo (OPEP); e a adição de outros combustíveis na composição da gasolina como, por exemplo, o álcool, que está sujeito a variações de preço em função da safra.

Há ainda os fatores políticos internos que influenciam diretamente os ajustes no preço da gasolina. No primeiro semestre de 2011, o governo brasileiro acenou com uma política permanente para garantir o abastecimento de etanol no mercado e minimizar os impactos da entressafra da cana-de-açúcar nos preços dos combustíveis. Para tanto, deu à Agência Nacional do Petróleo, Gás Natural e Biocombustíveis (ANP) o poder de regular os estoques de etanol no país e aumentou a flexibilidade na determinação do percentual de álcool anidro na gasolina.

O objetivo deste estudo é propor um modelo através do qual seja possível prever o valor do preço do litro da gasolina no estado do Rio de Janeiro, com base em variáveis observáveis e de fácil acesso, contribuindo para a gestão do orçamento e do fluxo de caixa da empresa CAM Ltda., objeto do estudo de caso desta pesquisa.

2. Revisão de Literatura

Esta parte do artigo está dividida em duas seções: uma revisão sobre trabalhos envolvendo a previsão do preço e da demanda de combustível, e uma revisão sobre trabalhos fazendo uso da regressão múltipla – a metodologia utilizada neste artigo – como ferramenta de previsão.

A partir das variáveis identificadas e dos *insights* observados nos diferentes trabalhos relatados nesta seção foi iniciada a construção do modelo aplicado no caso estudado.

2.1. Previsão do preço e da demanda de combustível

Analisando as principais revistas e trabalhos sobre o tema, foi verificado que a maioria procurava analisar o efeito do preço do petróleo ou dos combustíveis sobre alguma variável econômica. Entretanto, as variáveis utilizadas na análise podem ser um indicador de como executar o caminho inverso, ou seja, analisar variações em itens econômicos para prever o impacto destes no preço da gasolina no futuro.

Araújo (2006) procurou verificar os efeitos da volatilidade de preço do petróleo na economia brasileira. O estudo verificou que a política de preços dos derivados de petróleo foi conservadora no sentido de proteger o consumidor da alta volatilidade dos preços internacionais e da taxa de câmbio. O maior impacto, segundo o estudo, foi o aumento do Índice de Preço ao Consumidor Amplo (IPCA) pelo efeito do aumento dos derivados de petróleo, e a reação do Banco Central com aumento da taxa de juros. Neste estudo foi possível identificar a importância das variáveis cotação do dólar (ou câmbio) e inflação medida pelo IPCA.

Santos (2007) observou uma forte influência da política cambial nos preços do combustível no Brasil, o que reforça a importância da variável cotação do dólar em um modelo de previsão do preço do combustível.

Outra proposta observada nos trabalhos foi a tentativa de previsão, não do preço, mas da demanda pela gasolina ou por outros derivados de petróleo.

Burnquist e Bacchi (2010) procuraram analisar a demanda por gasolina no Brasil a partir de um modelo econômico básico que relaciona consumo, preço e renda, utilizando técnicas de cointegração. Como representante da componente consumo, foi utilizada a série de dados históricos de venda de gasolina e ou derivados de petróleo pelas refinarias. Já a componente renda foi modelada com as informações sobre o Produto Interno Bruto (PIB).

Como o interesse do trabalho aqui apresentado é a previsão do preço do litro da gasolina, a variável demanda, que foi prevista no trabalho de Burnquist e Bacchi (2010), passa a ser um importante item na análise apresentada neste artigo.

Por fim, foi verificada, no trabalho de Samohyl e Dantas (1998), a tentativa de previsão da demanda de gasolina no Brasil, onde as variáveis utilizadas foram o consumo de petróleo, o PIB e o preço do barril em dólares. Esta tentativa reforça a análise anterior referente ao PIB e à cotação do dólar, e acrescenta as questões de produção e de consumo de barris de petróleo.

Outro componente importante é a intervenção direta nos preços do governo, item que pode influenciar o modelo utilizado. Silva (2003) observou que os preços da gasolina, do diesel, e do Gás Liquefeito de Petróleo (GLP) não estiveram alinhados aos preços internacionais no período analisado. Como conclusão, o trabalho afirma que a análise do comportamento dos preços dos principais derivados no Brasil - gasolina, diesel e GLP - revelou a interferência de fatores extraeconômicos na definição dos preços da Petrobras.

2.2. Previsão utilizando a regressão múltipla

Bouzada e Saliby (2009) atacaram o problema da previsão de demanda de chamadas para um determinado produto no *call center* de uma grande empresa brasileira. A abordagem utilizada foi a da regressão múltipla com variáveis *dummy*. O modelo desenvolvido utiliza informações disponíveis capazes de influenciar a demanda, tais como o dia da semana, a ocorrência ou não de feriado e a proximidade da data com eventos críticos, como a chegada da conta à residência do cliente e seu vencimento.

Carvalho et al. (2004) parametrizaram e testaram um modelo de regressão múltipla aplicado sobre os componentes principais mais significativos obtidos de séries de produtividades da cultura do café em três cidades de Minas Gerais. O objetivo dos autores foi prever a produtividade dos cafeeiros nestas cidades.

Paixa e Mesquita (2007) lidaram com o problema da gestão do numerário em caixas eletrônicos, dificultado pela grande incerteza dos saques nos caixas, que apresentam variações referentes ao dia da semana, a feriados, a pagamentos de acordo com o dia útil do mês, ao tipo de equipamento e à sua localização dentro da agência. Os autores apresentaram modelo de regressão múltipla com variáveis *dummy* para prever a demanda diária por dinheiro nos caixas.

Bouzada (2008) procurou prever o tempo médio de atendimento (TMA) em um *call center*. A regressão múltipla com variáveis *dummy* foi utilizada para tal fim, gerando modelo que utiliza informações disponíveis capazes de influenciar o TMA, como a data de chegada da conta telefônica à residência do cliente e o seu vencimento.

3. Metodologia

O trabalho apresenta estudo de caso acerca do qual é desenvolvido um modelo que utiliza uma ferramenta quantitativa chamada regressão múltipla. Modelos que fazem uso desta ferramenta envolvem análise do relacionamento entre múltiplas variáveis explicativas e uma variável dependente. Boa parte das pesquisas que examinam o efeito exercido por duas ou mais variáveis independentes sobre uma variável dependente utiliza a análise de regressão múltipla, que é definida por Tabachnick e Fidell (1996) como um conjunto de técnicas estatísticas que possibilita a avaliação do relacionamento de uma variável dependente com diversas variáveis independentes.

Apesar de necessário à obtenção de dados empíricos confiáveis, o conhecimento dos riscos, das limitações e dos desafios associados ao uso de regressão múltipla não é suficiente para garantir qualidade às análises e aos resultados de pesquisa. A confiabilidade dos resultados empíricos não é produto apenas do uso adequado de técnicas estatísticas de análise de dados. Seguramente, as análises estatísticas não constituem as melhores nem as únicas respostas aos problemas de confiabilidade enfrentados pelo pesquisador em seus trabalhos. Para Kromrey e Foster-Johnson (2008), talvez a melhor resposta não esteja em estatísticas melhores, mas em um melhor pensamento ou em uma melhor elaboração sobre a natureza do problema investigado.

Para desenvolver o modelo aqui apresentado, foram definidas as variáveis independentes potencialmente explicativas do preço do litro da gasolina no estado do Rio de Janeiro. Para tais variáveis, que estão indicadas e definidas a seguir, foram pesquisados os seus valores históricos mensais para o período de dezembro de 2004 a outubro de 2010:

- Dólar (média mensal) – a cotação do dólar utilizando a média do mês, conforme divulgado pelo Banco Central do Brasil (BANCO CENTRAL DO BRASIL, 2010);

- PIB acumulado dos últimos 12 meses – o Produto Interno Bruto (PIB) acumulado dos últimos 12 meses, utilizando como referência o mês da variável (BANCO CENTRAL DO BRASIL, 2010);
- PIB mensal – o PIB do mês (BANCO CENTRAL DO BRASIL, 2010);
- Barril petróleo – o valor em dólar do barril de petróleo cru divulgado pela U.S. Energy Information Administration (U.S. ENERGY INFORMATION ADMINISTRATION, 2010);
- Venda de gasolina (refinarias) – a quantidade de gasolina comercializada pelas refinarias no mês, conforme divulgado pela Agência Nacional do Petróleo, Gás Natural e Biocombustíveis (AGÊNCIA NACIONAL DO PETRÓLEO, 2010);
- Venda de derivados de petróleo (refinarias) – a quantidade de derivados de petróleo comercializada pelas refinarias no mês (AGÊNCIA NACIONAL DO PETRÓLEO, 2010);
- Produção de gasolina mensal – a quantidade de gasolina produzida no mês (AGÊNCIA NACIONAL DO PETRÓLEO, 2010);
- Produção derivados de petróleo – a quantidade de derivados de petróleo produzida no mês (AGÊNCIA NACIONAL DO PETRÓLEO, 2010);
- IPCA do mês – o Índice Nacional de Preços ao Consumidor Amplo do mês, divulgado pelo Instituto Brasileiro de Geografia e Estatística (INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA, 2010).

Além das variáveis acima, tem-se o valor que se deseja prever, que é o preço do litro da gasolina no estado do Rio de Janeiro, também informado pela ANP (AGÊNCIA NACIONAL DO PETRÓLEO, GÁS NATURAL E BIOCMBUSTÍVEIS, 2010).

A análise foi realizada no Microsoft Excel 2007, com o auxílio do suplemento denominado Análise de Dados, que viabiliza a aplicação da técnica de regressão. A ferramenta Excel foi escolhida por estar disponível em todos os computadores da empresa, permitindo uma futura eventual continuidade das análises.

Depois de realizada a análise de regressão, foi conduzido teste para verificar a ausência da autocorrelação serial, um dos pressupostos da regressão linear.

Finalmente, foi feita análise comparativa contrapondo, para período de seis anos, os valores realmente gastos com gasolina em cada um dos meses deste período, os valores que foram orçados através da metodologia atual (conforme mais bem explicado na seção 4.1 a seguir) e os que teriam sido previstos utilizando o modelo de regressão apresentado neste trabalho.

3.1. Limitações

A linearidade das variáveis, outro pressuposto da regressão linear, não foi testada, já que violações no tocante a esta premissa não são tão graves quanto a presença de autocorrelação serial, que foi efetivamente testada.

Para comparação e validação do modelo de regressão, foram utilizados os mesmos períodos de tempo incluídos na própria construção do modelo, expediente geralmente não muito recomendável. Tal procedimento foi conduzido por causa do tamanho reduzido do histórico disponível utilizado na amostra (70 meses), que inviabilizaria sua partição em amostra de ajuste e amostra de validação, o que seria mais adequado. Ressalta-se, ainda, que, para que a comparação com a previsão obtida pela metodologia atual fosse mais representativa e consistente, foi necessário utilizar todo o período histórico para a validação do modelo.

4. Descrição do Caso

4.1. Situação atual

A empresa CAM Ltda. possui frota de 1.050 veículos que realizam entregas e serviços administrativos no estado do Rio de Janeiro. Estes serviços e suas entregas não são a atividade principal da empresa, mas são importantes no relacionamento com o cliente final. Todos os veículos são abastecidos com gasolina como combustível.

A despesa com combustível utilizado pela frota é significativa no orçamento total da empresa, montando 9% do total de despesas.

A empresa tem, como objetivo, estimar os gastos que serão incorridos nos próximos meses com combustível, para poder administrar melhor seu fluxo de caixa. A metodologia que a CAM Ltda. utiliza para prever o preço futuro do litro da gasolina consiste em atualizar, pelo IPCA, o preço do litro da gasolina no mês de dezembro do ano anterior. A empresa verificou que as estimativas realizadas para os seus gastos com combustível estavam distantes da realidade, conforme mostrado na Tabela 1.

Tabela 1 - Valores previstos e reais de gasto com combustível

Ano	Orçamento	Valor Gasto	Diferença	Erro
2005	R\$ 11.293.325,07	R\$ 11.080.800,00	R\$ 212.525,07	1,88%
2006	R\$ 12.618.432,35	R\$ 12.117.600,00	R\$ 500.832,35	3,97%
2007	R\$ 12.307.794,70	R\$ 11.972.400,00	R\$ 335.394,70	2,73%
2008	R\$ 12.497.534,95	R\$ 12.076.400,00	R\$ 421.134,95	3,37%
2009	R\$ 12.672.570,45	R\$ 12.208.800,00	R\$ 463.770,45	3,66%
2010	R\$ 10.796.314,82	R\$ 10.491.200,00	R\$ 305.114,82	2,83%

Fonte: Elaboração própria

Cabe ressaltar que a empresa considerou a quantidade de combustível gasto uma informação estratégica e sugeriu que fossem utilizados 400 mil litros por mês como parâmetro nas análises.

4.2. O método proposto

Através da técnica de regressão múltipla aplicada às variáveis descritas na seção 3 deste trabalho, foi verificado se seria possível definir modelo cuja previsão de gastos com combustível fosse mais precisa que a da metodologia atualmente utilizada pela empresa CAM.

- Primeira Regressão

Os resultados mais importantes estão apresentados nas Tabelas 2, 3 e 4.

Tabela 2 - Estatísticas da Regressão 1

Estatística de regressão	
R múltiplo	0,85479
R – Quadrado	0,73066
R – Quadrado ajustado	0,69026
Erro padrão	0,05802
Observações	70,00000

Fonte: Elaboração própria

É possível verificar que o R^2 indica que o modelo está explicando 73,07% da variação dos dados. Neste caso, pode-se inferir a real influência das variáveis explicativas na quantidade de ligações e aceitar o modelo com elevada confiança, já que o nível de significância da estatística F teve resultado igual a 0,0000.

Neste modelo, observa-se que as variáveis PIB acumulado dos últimos 12 meses, PIB mensal, produção de gasolina mensal, produção de derivados de gasolina e IPCA do mês, estão com valor p maior que 5%, indicando que não são bons previsores para o preço do litro da gasolina. Em vista disso, foram retiradas do modelo as duas piores variáveis (valor p mais alto) – produção de gasolina mensal e produção de derivados de gasolina – para verificar o impacto. A seção seguinte apresenta os resultados.

Tabela 3 - Análise de Variância (ANOVA) da Regressão 1

	gl	SQ	MQ	F	F de significação
Regressão	9,00000	0,54792	0,06088	18,08547	0,00000
Resíduo	60,00000	0,20197	0,00337		
Total	69,00000	0,74989			

Fonte: Elaboração própria

Tabela 4: Coeficientes da Regressão 1

	Coeficientes	Erro padrão	Stat t	Valor-P	95% inferiores	95% superiores
Interseção	2,99683	0,30451	9,84142	0,00000	2,38772	3,60595
Dólar (média mensal)	(0,29413)	0,07025	(4,18715)	0,00009	(0,43365)	(0,15362)
Barril petróleo	(0,00194)	0,00073	(2,65987)	0,01001	(0,00340)	(0,00048)
Venda de gasolina (refinarias)	0,00000	0,00000	2,69946	0,00901	0,00000	0,00000
Venda de derivados de petróleo (refinarias)	(0,00000)	0,00000	(3,17503)	0,00237	(0,00000)	(0,00000)
PIB acumulado (últimos 12 meses)	0,00000	0,00000	1,46096	0,14924	(0,00000)	0,00000
PIB mensal	0,00000	0,00000	0,722260	0,47273	(0,00000)	0,00000
Produção de gasolina	0,00000	0,00000	0,51073	0,61141	(0,00000)	0,00000
Produção derivados de petróleo	0,00000	0,00000	0,39876	0,69148	(0,00000)	0,00000
IPCA do mês	(0,05659)	0,03657	(1,54764)	0,12697	(0,12974)	0,01655

Fonte: Elaboração própria

- Segunda Regressão

Os resultados mais importantes estão apresentados nas Tabelas 5, 6 e 7.

Verifica-se que o R^2 teve leve queda, explicando 72,57% das variações. Adicionalmente, ainda permaneceram variáveis com valor p maior que 5%. Foram eliminadas, então, as variáveis PIB acumulado dos últimos 12 meses e PIB mensal. A seção seguinte apresenta os resultados.

Tabela 5 - Estatísticas da Regressão 2

Estatística de regressão	
R múltiplo	0,85188
R – Quadrado	0,72571
R – Quadrado ajustado	0,69474
Erro padrão	0,05760
Observações	70,00000

Fonte: Elaboração própria

Tabela 6 - ANOVA da Regressão 2

	gl	SQ	MQ	F	F de significação
Regressão	7,00000	0,54420	0,07774	23,43351	0,00000
Resíduo	62,00000	0,20569	0,00332		
Total	69,00000	0,74989			

Fonte: Elaboração própria

Tabela 7 - Coeficientes da Regressão 2

	Coeficientes	Erro padrão	Stat t	Valor-P	95% inferiores	95% superiores
Interseção	3,20328	0,23042	13,90221	0,00000	2,74269	3,66387
Dólar (média mensal)	(0,30954)	0,06753	(4,58349)	0,00002	(0,44453)	(0,17454)
Barril petróleo	(0,00192)	0,00072	(2,66944)	0,00969	(0,00336)	(0,00048)
Venda de gasolina (refinarias)	0,00000	0,00000	3,06858	0,00319	0,00000	0,00000
Venda de derivados de petróleo (refinarias)	(0,00000)	0,00000	(3,41395)	0,00113	(0,00000)	(0,00000)
PIB acumulado (últimos 12 meses)	0,00000	0,00000	1,36844	0,17611	(0,00000)	0,00000
PIB mensal	0,00000	0,00000	0,76975	0,44437	(0,00000)	0,00000
IPCA do mês	(0,06043)	0,03611	(1,67328)	0,09931	(0,13262)	0,01176

Fonte: Elaboração própria.

- Terceira Regressão

Os resultados mais importantes estão apresentados nas Tabelas 8, 9 e 10.

Verifica-se que o R^2 teve queda significativa e que o valor p das variáveis venda de derivados de petróleo (refinarias) e IPCA do mês também aumentaram consideravelmente. Com isso, optou-se por fazer uma última análise retirando estas duas variáveis. A seção seguinte apresenta os resultados.

Tabela 8 - Estatísticas da Regressão 3

Estatística de regressão	
R múltiplo	0,81876
R – Quadrado	0,67037
R – Quadrado ajustado	0,64462
Erro padrão	0,06215
Observações	70,00000

Fonte: Elaboração própria

Tabela 9 - ANOVA da Regressão 3

	GI	SQ	MQ	F	F de significação
Regressão	5,00000	0,50271	0,10054	26,03178	0,00000
Resíduo	64,00000	0,24718	0,00386		
Total	69,00000	0,74989			

Fonte: Elaboração própria.

Tabela 10: Coeficientes da Regressão 3

	Coefi- cientes	Erro padrão	Stat t	Valor-P	95% inferiores	95% superiores
Interseção	3,44706	0,23442	14,70453	0,00000	2,97875	3,91537
Dólar (média mensal)	(0,42508)	0,05829	(7,29218)	0,00000	(0,54153)	(0,30863)
Barril petróleo	(0,00297)	0,00069	(4,32177)	0,00006	(0,00435)	(0,00160)
Venda de gasolina (refinarias)	0,00000	0,00000	2,52308	0,01413	0,00000	0,00000
Venda de derivados de petróleo (refinarias)	(0,00000)	0,00000	(1,52010)	0,13341	(0,00000)	0,00000
IPCA do mês	(0,02031)	0,03684	(0,55123)	0,58340	(0,09390)	0,05329

Fonte: Elaboração própria.

- Quarta Regressão

Os resultados mais importantes estão apresentados nas Tabelas 11, 12 e 13.

Tabela 11 - Estatísticas da Regressão 4

Estatística de regressão	
R múltiplo	0,81136
R – Quadrado	0,65830
R – Quadrado ajustado	0,64277
Erro padrão	0,06231
Observações	70,00000

Fonte: Elaboração própria.

Tabela 12 - ANOVA da Regressão 4

	gl	SQ	MQ	F	F de significação
Regressão	3,00000	0,49366	0,16455	42,38457	0,00000
Resíduo	66,00000	0,25624	0,00388		
Total	69,00000	0,74989			

Fonte: Elaboração própria.

Tabela 13 - Coeficientes da Regressão 4

	Coeficientes	Erro padrão	Stat t	Valor-P	95% inferiores	95% superiores
Interseção	3,3025885705	0,20776	15,89596	0,00000	2,88778	3,71740
Dólar (média mensal)	(0,3872461372)	0,04988	(7,76392)	0,00000	(0,48683)	(0,28766)
Barril petróleo	(0,0029092773)	0,00066	(4,42369)	0,00004	(0,00422)	(0,00160)
Venda de gasolina (refinarias)	0,0000000957	0,00000	2,04528	0,04482	0,00000	0,00000

Fonte: Elaboração própria.

Pode-se verificar que o nível de significância da estatística F manteve-se muito baixo (praticamente zero); entretanto o modelo está explicando apenas 65,83% da variação dos dados, o que é aceitável, mas não chega a poder ser considerado bom, preditivamente.

Apesar disso, e pelo fato de todas as variáveis explicativas apresentarem um valor p abaixo do nível de significância tradicionalmente usado nos testes de hipóteses (5%), optou-se por manter o modelo e aplicá-lo aos dados, para verificar se, da forma que está, pode auxiliar a empresa no seu objetivo de melhoria na previsão do fluxo de caixa, conforme pode ser observado no final desta seção.

O novo modelo foi, então, utilizado para calcular as previsões do preço, em reais, do litro da gasolina (G) para os anos de 2005 a 2010, através da fórmula:

$$G=3,3025886-0,3872461xDólar-0,0029093xBarril\ petr\ oleo+0,0000001xVenda\ de\ gasolina$$

Na fórmula obtida, pode-se observar que o dólar mensal e o barril de petróleo têm influência negativa no preço da gasolina enquanto o esperado para estas variáveis seria uma relação direta e positiva. Uma possível explicação para este fato é a interferência governamental no preço final dos combustíveis; entretanto, novos estudos devem ser realizados para investigar esta influência.

A diferença entre os valores reais e os gerados pelo modelo também é conhecida como resíduo. Para ser usado um modelo de regressão múltipla, é necessário supor que a correlação entre os resíduos, ao longo do espectro das variáveis independentes, é zero; ou seja, que os resíduos são independentes entre si, não apresentando, conseqüentemente, autocorrelação serial (CORRAR; PAULO; DIAS FILHO, 2007).

Segundo estes autores, uma forma de verificar a validade dessa suposição é por meio do teste de Durbin-Watson. No modelo que está sendo analisado, foi encontrado o valor 1,698 para a Estatística DW.

De acordo com Corrar, Paulo e Dias Filho (2007), essa estatística obtida deve ser comparada com valores críticos que variam em função do nível de significância do teste, do tamanho da amostra e do número de variáveis independentes.

Com 5% de significância, 70 dados na amostra e 3 variáveis independentes, o valor crítico inferior (dL) obtido foi 1,525 e o valor crítico superior (dU) encontrado foi 1,703 (STANFORD, 2011).

Como o valor obtido para a Estatística DW (1,698) se situa entre os dois valores críticos encontrados, o teste para verificar a existência de autocorrelação dos resíduos é não conclusivo: não é possível garantir que os resíduos são independentes, mas também é incorreto afirmar que existe uma relação de dependência entre eles (CORRAR; PAULO; DIAS FILHO, 2007).

A comparação entre os resultados obtidos pela metodologia atual de previsão e a proposta pelo modelo apresentado neste trabalho pode ser visualizada na Tabela 14, onde pode ser observado o Erro Percentual Absoluto (APE) incorrido por cada metodologia, em cada ano da análise.

Tabela 14 - Valor real e orçamento previsto com os dois modelos

Ano	Valor Gasto (R\$)	Modelo antigo			Modelo Novo		
		Orçamento (R\$)	Diferença (R\$)	Erro (APE)	Orçamento (R\$)	Diferença (R\$)	Erro (APE)
2005	11.080.800	11.293.325,07	212.525,07	1,88%	11.381.376,39	300.576,39	2,64%
2006	12.117.600	12.618.432,35	500.832,35	3,97%	11.770.103,42	(347.496,58)	2,95%
2007	11.972.400	12.307.794,70	335.394,70	2,73%	12.100.610,44	128.210,44	1,06%
2008	12.076.400	12.497.534,95	421.134,95	3,37%	12.057.502,99	(18.897,01)	0,16%
2009	12.208.800	12.672.570,45	463.770,45	3,66%	12.153.528,92	(55.271,08)	0,45%
2010	10.491.200	10.796.314,82	305.114,82	2,83%	10.484.077,85	(7.122,15)	0,07%

Fonte: Elaboração própria.

Apesar da pouca explicação do novo modelo (66%), pode-se observar que este permitiu uma melhor estimativa de gastos em cinco dos seis anos analisados, potencializando um auxílio na previsão do fluxo de caixa da empresa CAM. Nos últimos seis anos, o modelo atual acusou um Erro Percentual Absoluto Médio (MAPE) anual de 3,07%, enquanto o modelo aqui proposto teria incorrido em MAPE anual de 1,22% para o mesmo período.

Pode parecer estranho um modelo explicar “apenas” 66% da variação dos dados e, mesmo assim, cometer erro médio de previsão tão baixo quanto 1,22%. Cabe aqui ressaltar que este erro médio é anual, enquanto a regressão foi desenvolvida em cima de uma base de dados mensal. Desta maneira, o modelo consegue explicar 66% da variação mensal, deixando de explicar os outros 34%. Mas o erro médio anual torna-se muito menor porque, ao longo do ano, erros de previsão mensais para mais acabam “compensando” alguns erros de previsão mensais para menos.

5. Considerações Finais

O objetivo deste artigo foi desenvolver, testar e apresentar um modelo causal de previsão do preço do litro da gasolina no Estado do Rio de Janeiro. Diversas potenciais causas para a variação neste preço foram pesquisadas na literatura e incluídas inicialmente no modelo como variáveis candidatas a explicadores do preço do litro da gasolina.

Foi possível verificar que as variáveis PIB acumulado dos últimos 12 meses, PIB mensal, venda de derivados de petróleo (refinarias), produção de gasolina mensal, produção derivados de petróleo e IPCA do mês não se comportaram como bons explicadores lineares para a variação no preço da gasolina.

Já a cotação do dólar, o valor do barril de petróleo e a quantidade de gasolina vendida pelas refinarias mostraram-se capazes de auxiliar a explicação da variação do preço do combustível. Entretanto, o percentual explicado desta variação ainda está aquém do ideal para um modelo de previsão mais robusto.

Para justificar tal afirmação, parece pertinente comparar a qualidade do modelo obtido – nível de explicação igual a 65,83%, MAPE anual obtido igual a 1,22%, contra 3,07% da abordagem atual – com a de outros apresentados na revisão de literatura.

O modelo desenvolvido por Bouzada e Saliby (2009) apresentou nível de explicação igual a 85% e MAPE de 10,98% contra 13,54% obtido pela ferramenta anteriormente em uso, encerrando ganhos de acurácia da ordem de 3 pontos percentuais para o período estudado.

Carvalho et al. (2004) encontraram um nível de explicação igual a 88%, 75% e 66% para cada uma das três cidades cujos cafeeiros se quis prever as produtividades.

O modelo de Paixa e Mesquita (2007) para prever a demanda por dinheiro nos caixas eletrônicos apresentou um nível de explicação igual a 89%, com MAPE de 25,15%.

Bouzada (2008) encontrou, no seu modelo, um nível de explicação igual a 79% e um MAPE de 3,61%. A ferramenta anteriormente em uso incorria em um MAPE de 5,58%, fazendo com que o novo modelo proporcionasse ganhos de acurácia da ordem de 2 pontos percentuais.

Desta forma, é possível perceber que o nível de explicação resultante do modelo apresentado neste artigo é inferior ao de todos os modelos pesquisados. O MAPE obtido, no entanto, foi inferior ao dos outros modelos mencionados (fazendo a ressalva que se trata de um erro anual de previsão). A pequena magnitude desta medida de erro e, principalmente, os potenciais ganhos de acurácia a serem obtidos pela utilização do modelo aqui proposto em detrimento da abordagem atual, reforçam a relevância dos resultados obtidos e apresentados nesta pesquisa.

Apesar dos valores insuficientes para previsão mais precisa do valor do litro da gasolina, o modelo se mostrou útil à empresa CAM, sendo mais preciso do que o atualmente utilizado - acarretando um erro de previsão anual praticamente três vezes menor - e auxiliando na melhor gestão do fluxo de caixa. Desta forma, o objetivo de definir um modelo capaz de prever o preço do litro da gasolina no estado do Rio de Janeiro foi atingido. É necessário, porém, destacar o valor não muito alto obtido para o nível de explicação e deixar claro que os bons resultados do modelo se aplicam ao estudo de caso em questão, não podendo ser generalizados para qualquer situação.

Não obstante, novos estudos devem ser realizados para melhoria do modelo de previsão, auxiliando na explicação mais precisa do valor do litro da gasolina. Tais modelos podem considerar, por exemplo, as seguintes variáveis métricas adicionais: valor do álcool; PIB regional e valor do frete.

Pode ser considerada também a inclusão de variáveis nominais ou *dummy* – como em Bouzada e Saliby (2009), Paixa e Mesquita (2007) e Bouzada (2008), dentre outros – que busquem identificar situações específicas, tais como ano de eleição no Brasil e conflitos nas principais áreas produtoras, entre outras.

Por fim, é recomendável ampliar a série histórica utilizada, o que concederia maior representatividade às observações e maior confiabilidade aos resultados, além de permitir

que a amostra fosse partida em amostra de ajuste de modelo e amostra de validação, conforme foi explicado na seção 3.1.

Referências

- AGÊNCIA NACIONAL DO PETRÓLEO, GÁS NATURAL E BIOCOMBUSTÍVEIS - ANP. **Boletins ANP**. Disponível em <<http://www.anp.gov.br/?id=711>> Acesso em 30/12/2010.
- ARAÚJO, T. J. **Quais são os efeitos da volatilidade de preço do petróleo na economia brasileira?** Uma análise de 2002 a 2006. 2006. Dissertação (Mestrado Profissionalizante em Economia) – Faculdade de Economia e Finanças IBMEC, Rio de Janeiro, 2006.
- BANCO CENTRAL DO BRASIL. **Sistema gerenciador de séries temporais**. Disponível em <<https://www3.bcb.gov.br/sgspub/localizarseries/localizarSeries.do?method=prepararTelaLocalizarSeries>> Acesso em 30/12/2010.
- BOUZADA, M. O uso da Regressão Múltipla (com variáveis dummy) para prever o tempo médio de atendimento em um call center. In: Encontro Nacional da ANPAD - ENANPAD, 32., 2008, Rio de Janeiro. **Anais...** Rio de Janeiro: ANPAD, 2008.
- BOUZADA, M.; SALIBY, E. Prevendo a demanda de ligações em um call center por meio de um modelo de Regressão Múltipla. **Gestão & Produção**, v. 16, n.3, p.382-397, jul-set/2009.
- BURNQUIST, H.; BACCHI, M. **A demanda por gasolina no Brasil: uma análise utilizando técnicas de co-integração**. São Paulo: Centro de Estudos Avançados em Economia Aplicada (CEPEA), 2010. Disponível em <<http://www.cepea.esalq.usp.br/pdf/DemandaGasolina.pdf>>. Acesso em 30/12/2010
- CARVALHO, L.; SEDYIAMA, G.; CECON, P.; ALVES, H. Modelo de regressão para a previsão de produtividade de cafeeiros no Estado de Minas Gerais. **Revista Brasileira de Engenharia Agrícola e Ambiental**, v. 8, n.2-3, p.1-15, maio-dezembro, 2004.
- CONFEDERAÇÃO NACIONAL DO TRANSPORTE - CNT. **Transporte de cargas no Brasil: ameaças e oportunidades para o desenvolvimento do país**. Pesquisa CNT/COPPEAD. 2011. Disponível em <<http://www.cnt.org.br/portal/webcnt/default.aspx>>. Acesso em 05/01/2011.
- CORRAR, L.; PAULO, E.; DIAS FILHO, J. **Análise multivariada para os cursos de Administração, Ciências Contábeis e Economia**. São Paulo: Atlas, 2007.
- INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. Disponível em <www.ibge.gov.br> Acesso em 20/12/2010.
- KROMREY, J.; FOSTER-JOHNSON, L. Statistically differentiating between interaction and nonlinearity in multiple regression analysis: a Monte Carlo investigation of a recommended strategy. **Educational and Psychological Measurement**, v. 58, p. 392-413, fev/2008.
- PAIXA, R.; MESQUITA, M. Previsão de demanda e reposição de numerário em uma rede de caixas-eletrônicos. In: Encontro Nacional de Engenharia de Produção- ENEGEP, 27., 2007, Foz do Iguaçu (PR). **Anais...** Foz do Iguaçu (PR): ABEPRO, 2007.
- SAMOHYL, R.; DANTAS, A. Previsão com cointegração e modelo a correção de erro: aplicação empírica no consumo de gasolina do Brasil. In: Encontro Nacional de Engenharia de Produção - ENEGEP, 18., 1998, Niterói (RJ). **Anais...** Niterói (RJ): ABEPRO, 1998.
- SANTOS, C. **A influência da variação cambial nos preços dos combustíveis no Brasil no período de 1990 a 2003**. 2007. Dissertação (Mestrado em Engenharia da Energia), Universidade Federal de Itajubá, Itajubá (MG), 2007.

SILVA, C.M.S. **Estratégia de Preços da Petrobras no mercado de combustíveis brasileiro pós-liberalização e instrumentos de amortecimento de variações internacionais.** Dissertação (Mestrado em Economia), Instituto de Economia, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2003.

STANFORD. **Critical values for the Durbin-Watson test.** Disponível em <http://www.stanford.edu/~clint/bench/dwcrit.htm>. Acesso em: 16/10/2011

TABACHNICK, B.; FIDELL, L. **Using multivariate statistics.** New York: Harper Collins, 1996.

U.S. ENERGY INFORMATION ADMINISTRATION. Disponível em <http://tonto.eia.gov/dnav/pet/hist/LeafHandler.ashx?n=PET&s=RWTC&f=D> Acesso em 30/12/2010.